STABILITY RESULTS FOR STOCHASTIC IMPULSIVE SYSTEMS*

JOÃO PEDRO HESPANHA†

Abstract. This paper considers stochastic impulsive systems, which evolve according to ordinary differential equations most of the time, but occasionally exhibit stochastic discontinuities in their trajectories. The main contribution of this paper consists of a set of results to establish global existence of solution and stability for SISs. We consider stability notions that emphasize the existence of bounds on the solution to SISs that hold uniformly over time, expressed in terms of class $\mathcal{K}$ and class $\mathcal{KL}$ functions. The results are stated in terms of the existence of Lyapunov functions that satisfy appropriate algebraic conditions that can be checked without computing solutions to the SISs.

Key words. Stochastic Impulsive Systems; Stochastic Hybrid Systems; Markov processes; Stability Theory; Lyapunov Stability.

1. Introduction. *Impulsive Systems* are dynamical systems whose state evolves according to an ordinary differential equation most of the time, but occasionally exhibits discontinuities, known as impulses or jumps. In *Stochastic Impulsive Systems (SISs)*, randomness arises from *when* jumps occur and to *where* the state jumps. The "when" question is determined by a state-dependent intensity functions that determines how likely a jump will occur, much like the intensity in continuous-time Markov chains, except that in SISs this intensity is a continuously varying quantity. Once a jump takes place, the question of "where" the state jumps is determined by a distribution of the jump destinations that typically depends on the value of the state just before the jump.

SISs are closely related (and heavily inspired) by the Piecewise-Deterministic Markov Process (PDMPs) introduced by Davis [5] and, in fact, SISs can be viewed as a special case of PDMPs. While the work on the stability of Markov processes [8, 17–19], inspired many of the results and techniques used here, we never formally prove that the state of SISs are Markov processes. In fact, the stability results presented here are formulated for "output processes" that generally do not exhibit Markov's property. SIS's are also closely related to Stochastic Hybrid Systems [1, 2, 11, 13]. In fact, the SHSs considered in [9, 11] are special cases of the SISs considered here and therefore one can use the tests for existence of solution and stability developed here to prove these properties for those SHSs. Switching difusions [7, 23] differ from SIs in that the emphasis there is the change in dynamics at a set of event times and not the impulsive effects. Also, in switching diffusions, the solution between discrete events is obtained from a stochastic differential equation, instead of an ordinary differential equation; therefore such systems exhibit randomness even without jumps.

The main contribution of this paper consist of a set of theorems based on which one can prove global existence of solution and appropriately defined notions of stability for SISs. The stability notions considered are formulated in terms of the properties of "output" processes that may include only a portion of the state, and typically are not Markovian. This is especially important for SHSs in which some components of the state take values in discrete sets (e.g., boolean variables) that may not converge to any specific value. We focus our work on notions of stability that emphasize the existence of *uniform* bounds on the transient behavior of output processes. In particular, on the existence of bounds on the norms of such processes in terms of class $\mathcal{K}$ and class $\mathcal{KL}$, that hold uniformly over time. Because of the stochastic nature of SISs, we consider both *sample-path notions of stability*, which place constraints on the confidence that sample paths do not "misbehave" (i.e., do not violate norm bounds); as well as *ensemble notions of stability* that place constraints on the expected values of the stochastic process.

All the existence and stability results are stated in terms of the existence of Lyapunov functions that satisfy appropriate *algebraic conditions that can be checked without computing solutions to the SISs.* To simplify the process of finding such Lyapunov functions, we have devoted significant effort to enlarge the set over which one can search for Lyapunov functions. Essentially, any non-negative locally Lipschitz continuous function can be used as candidate Lyapunov function. This permits Lyapunov functions that are not

*This work was supported by the National Science Foundation under Grants ECCS-0835847, ECCS-0725485, and CNS-0720842. The research reported in this paper is original and has not been published or submitted for publication elsewhere (journal or conference).

†University of California, Santa Barbara, CA 93106, hespanha@ece.ucsb.edu

differentiable everywhere (e.g., an Euclidean norm) and also functions that are not positive definite. This level of generality is important for systems for which some components of the state may not converge, which is common, e.g., in SHSs.

An important feature of these results is that global existence of solution and, in particular, establishing that there is zero probability of infinite jumps in finite time is a *conclusion* of our theorems and not an assumption. We also do not require assumptions related to the sum of discontinuities of the Lyapunov function along solutions to the SIS, which would be hard to verify for general Lyapunov functions, without somehow computing solutions to the SISs.

While our proofs are much inspired by the notion of extended generator and by Dynkin's formula for PDMPs [5], we do not need to assume that a candidate Lyapunov function belongs to the domain of the extended generator. Instead, Dynkin's formula is also a *conclusion* of our global existence of solution theorem and not an assumption. To achieve this, we base our global existence result on an auxiliar "stopped process" that is always globally defined and for which a Dynkin's-like formula holds for a very large class of Lyapunov functions. This stopped process evolves just like the state of the SIS inside a compact set $\mathcal{Q}$, but becomes constant the first time that the interior of $\mathcal{Q}$ is abandoned. Under appropriate assumptions, the stopped process converges to the state of the SIS as we "enlarge" the compact set. The idea of using stopped processes to prove global existence and stability dates back to the seminal work of Kushner [18] and was also used more recently in [20] to establish global existence of solution to general Markov processes. Our derivation of a Dynkin's-like formula for the stopped processes is heavily inspired by the work of Davis. However, we do provide a self-contained proof that only uses basic results in probability theory and dispenses a background knowledge of Markov processes and stochastic integration. Whether or not this is an advantage of our proof is probably a mater of taste.

An interesting feature of our stability results is that we obtain fairly explicit formulas for the class $\mathcal{K}$ and class $\mathcal{KL}$ that appear in the stability definitions. This provides insight on how the stability bounds vary over time (for the asymptotic convergence results) and also on how these bounds vary with the degree of confidence that appear in the sample-path notions of stability.

The remainder of this paper is organized as follows: SISs and their solutions are introduced formally in Sections 2 and 3, respectively; and Section 4 discusses the different notions of stability considered. Stopped processes are defined in Section 5 and a Dynkin's-like formula for these stochastic processes is provided in Section 6. The main results regarding global existence of solution and stability appear in Sections 7 and 8, respectively.

Notation. All random variables in this write-up are measurable on the same probability space $(\Omega, \mathcal{F}, \mathrm{P})$. We denote random variables and stochastic processes in boldface and, for short of notation, we sometimes omit the dependence on the outcome $\omega \in \Omega$ [as in $\mathbf{t}_k$ or $\mathbf{N}(t)$, instead of $\mathbf{t}_k(\omega)$ or $\mathbf{N}(t;\omega)$]. In expressions involving random variable that require some quantification on ω, we use the superscript $\overset{\text{wpo}}{\cdot}$ to denote universal quantification with respect to some subset of Ω with probability one [as in $\mathbf{X} \overset{\text{wpo}}{=} \mathbf{Z}$ to express that $\exists \bar{\Omega} \in \mathcal{F}$ such that $\mathrm{P}(\bar{\Omega}) = 1$ and $\mathbf{X}(\omega) = \mathbf{Y}(\omega)$, $\forall \omega \in \bar{\Omega}$.] Given a time dependent function $x : [0,T) \to \mathbb{R}^n$, $T \in (0, \infty]$ that has limits from the left at every point in $(0,T)$, we denote by $(\cdot)^-$ the left-limit operator that maps x into the function $x^- : (0,T) \to \mathbb{R}^n$ defined by $x^-(t) \coloneqq \lim_{\tau \downarrow t} x(\tau)$, $\forall t \in (0,T)$.

2. Stochastic Impulsive Systems.

In this paper, we use the following equations to represent a *Stochastic Impulsive System (SIS)*

$$\begin{cases} \dot{\mathbf{x}} = f(\mathbf{x}) & \text{w.p. } 1 - \lambda(\mathbf{x})dt \\ \mathbf{x} = \mathbf{x}_k & \text{w.p. } \nu_{\mathbf{x}^-}(d\mathbf{x}_k)\,\lambda(\mathbf{x})dt. \end{cases} \tag{2.1}$$

The *solution to this SIS* is a stochastic process $\mathbf{x}(t)$, $t \geqslant 0$ to be defined shortly that takes values in the *state space* $\mathcal{X} \subset \mathbb{R}^n$. Typically, $\mathcal{X}$ is the whole $\mathbb{R}^n$, but it is sometimes convenient to define the SIS on open subsets of $\mathbb{R}^n$, thus the introduction of $\mathcal{X}$. Three entities appear in (2.1) and are needed to characterize a SIS:

1. a *vector field* $f : \mathcal{X} \to \mathbb{R}^n$;

2. a *transition intensity* $\lambda : \mathcal{X} \to [0, \infty)$; and
3. a *family of jump measures* $\{\nu_x : \mathcal{F}_\mathcal{X} \to [0,1],\ x \in \mathcal{X}\}$, each defined on the σ-algebra $\mathcal{F}_\mathcal{X}$ of Lebesgue measurable sets in $\mathcal{X}$.

The expression in (2.1) should only be understood as the representation of a SIS, as the precise meaning of a *solution* to a SIS will be defined precisely below. Informally, the solution to (2.1) is a stochastic process $\mathbf{x}(t)$ that evolves continuously according to the ODE $\dot{\mathbf{x}} = f(\mathbf{x})$ most of the time, but occasionally exhibits discontinuities due to instantaneous jumps. These discontinuities occur at *jump times* $\{\mathbf{t}_k\}$ and correspond to resets of the state to *jump points* $\{\mathbf{x}_k\}$. The transition intensity λ determines the jump times in the sense that the probability that a jump will occur in a small interval $(t, t + dt]$ of length dt is equal to $\lambda\big(\mathbf{x}(t)\big)dt$. Once a jump takes place at a time $\mathbf{t}_k$, the jump measure $\nu_{\mathbf{x}^-(\mathbf{t}_k)}(\cdot)$ defines the distribution of the jump point $\mathbf{x}(t_k) = \mathbf{x}_k$.

It is useful to consider the following special cases of SISs that are common in applications.

SIS with a deterministic jump map. In SISs, the measures ν_x, $x \in \mathcal{X}$ that define the jump points may be "degenerate" in the sense that all probability mass may be concentrated at a single point. For example, we may have

$$\nu_x(z) = \delta\big(z - \rho(x)\big), \qquad \forall x, z \in \mathcal{X},$$

where δ denotes the Dirac distribution and $\rho : \mathcal{X} \to \mathcal{X}$ a given *(determinisc) jump map*. In this case, a jump at time $\mathbf{t}_k$ causes a reset of the state to $\mathbf{x}(\mathbf{t}_k) = \rho\big(\mathbf{x}^-(\mathbf{t}_k)\big)$ with probability one, and one can "abbreviate" the SIS representation to

$$\begin{cases} \dot{\mathbf{x}} = f(\mathbf{x}) & \text{w.p. } 1 - \lambda(\mathbf{x}^-)dt \\ \mathbf{x} = \rho(\mathbf{x}^-) & \text{w.p. } \lambda(\mathbf{x}^-)dt. \end{cases}$$

SIS with a (stochastic) discrete jump map. A common generalization of SISs with deterministic jump maps arises when the jump measures ν_x, $\forall x \in \mathcal{X}$ are finite discrete distributions defined by

$$\nu_x(z) = \sum_{i=1}^{K} p_i \delta\big(z - \rho_i(x)\big), \quad \forall i \in \{1, 2, \dots, K\},\ x, z \in \mathcal{X}, \tag{2.2}$$

where each $p_i(x) \in [0, 1]$ is the probability that the jump measure "selects" the jump point to be $\rho_i(x) \in \mathcal{X}$, with $\sum_i p_i(x) = 1$, $\forall x \in \mathcal{X}$. In this case, we "abbreviate" the SIS representation to

$$\begin{cases} \dot{\mathbf{x}} = f(\mathbf{x}) & \text{w.p. } 1 - \lambda(\mathbf{x}^-)dt \\ \mathbf{x} = \rho_1(\mathbf{x}^-) & \text{w.p. } p_1(\mathbf{x}^-)\lambda(\mathbf{x}^-)dt \\ \mathbf{x} = \rho_2(\mathbf{x}^-) & \text{w.p. } p_2(\mathbf{x}^-)\lambda(\mathbf{x}^-)dt \\ \quad\vdots & \\ \mathbf{x} = \rho_K(\mathbf{x}^-) & \text{w.p. } p_K(\mathbf{x}^-)\lambda(\mathbf{x}^-)dt. \end{cases}$$

Stochastic hybrid systems. It sometimes happens that one can split the components of the state $\mathbf{x}$ as $(\mathbf{q}, \mathbf{z})$, where $\mathbf{q}$ remain constant along solutions to the vector field and only changes at jump times. In such SISs, $\mathbf{q}$ often takes place in a discrete set $\mathcal{Q}$ and is called the *discrete state*, whereas $\mathbf{z}$ is called the *continuous state*. These systems are also called Stochastic Hybrid Systems (SHS) [10] and can be "abbreviated" to

$$\begin{cases} \dot{\mathbf{z}} = f(\mathbf{q}, \mathbf{z}) & \text{w.p. } 1 - \lambda(\mathbf{q}^-, \mathbf{z}^-)dt \\ (\mathbf{q}, \mathbf{z}) = (\mathbf{q}_k, \mathbf{z}_k) & \text{w.p. } \nu_{\mathbf{q}^-, \mathbf{z}^-}(d\mathbf{q}_k\, d\mathbf{z}_k)\lambda(\mathbf{q}^-, \mathbf{x}^-)dt, \end{cases} \tag{2.3}$$

where the (missing) equation for $\dot{\mathbf{q}}$ indicates that $\mathbf{q}$ remains constant between jumps. SHSs can also only have deterministic jump maps or stochastic discrete jump maps, as indicated above.

Remark 1 (Time-variant case). *All the results in this paper remain unchanged if the vector field in (2.1) is time-varying since we do not make use of time-invariance. We omit the dependence of the vector field f on t only for convenience of notation.* □

3. Solution to a SIS. The following assumption on the vector field and the transition intensity of the SIS is needed to define a solution to the SIS and is implicit in all results presented on this paper.

ASSUMPTION 1. *The maps defining the SIS satisfy the following properties:*

A1 *The vector field $f : \mathcal{X} \to \mathbb{R}^n$ is locally Lipschitz on $\mathcal{X}$.*

Consequently, for each $x_0 \in \mathcal{X}$ and $t_0 \geqslant 0$, the (deterministic) initial value problem

$$\dot{x} = f(x), \quad x(t_0) = x_0 \tag{3.1}$$

has a unique solution on some nonempty interval starting at t_0. In what follows, we shall denote by $[t_0, T_{t_0,x_0})$ with $T_{t_0,x_0} \in (t_0, \infty]$ the maximal interval of existence of solution to (3.1) and by $t \mapsto \varphi(t; t_0, x_0) \in \mathcal{X}$, $t \in [t_0, T_{t_0,x_0})$ the corresponding solution. This assumption also allow us to conclude that $t \mapsto \varphi(t; t_0, x_0)$ is an absolutely continuous function and that we can only have $T_{t_0,x_0} < \infty$ when (3.1) has finite escape time from $\mathcal{X}$, in which case $\varphi(t; t_0, x_0)$ must exit every compact subset of $\mathcal{X}$ before T_{t_0,x_0}.

A2 *The transition intensity $\lambda : \mathcal{X} \to [0, \infty)$ is locally bounded*[1] *and $t \mapsto \lambda\big(\varphi(t; t_0, x_0)\big)$ is integrable on any finite interval $[t_0, T]$, $\forall T < T_{t_0,x_0}$.*

This assumption is satisfied, e.g., if $\lambda : \mathcal{X} \to [0, \infty)$ is continuous on $\mathcal{X}$ (but not necessarily uniformly bounded). □

With Assumption 1 in place, the state $\mathbf{x}(t)$, $t \geqslant 0$ of the SIS is defined to be a stochastic process constructed as follows. Consider the following processes defined on a common probability space $(\Omega, \mathcal{F}, \mathrm{P})$:

- A random variable $\mathbf{x}_0$ that will be associated with the initial condition $\mathbf{x}(0)$.
- A standard Poisson process $\mathbf{N}(t)$, $t \geqslant 0$ with inter-event times $\{\boldsymbol{\Delta}_k : k \geqslant 0\}$ identically distributed standard exponential random variables with cummulative distribution

$$\mathrm{P}(\boldsymbol{\Delta}_k \leqslant x) = F(x) \coloneqq 1 - e^{-x}, \quad \forall x \in \mathbb{R}.$$

- For each $x \in \mathcal{X}$, a process $\mathbf{Z}(x) \coloneqq \{\mathbf{Z}_k(x) : k = 1, 2, \dots\}$ consisting of a sequence of i.i.d. random variables, each with measure ν_x.
  The sequences $\mathbf{Z}(x)$ corresponding to the different values of x should be independent of each other and also independent of $\mathbf{N}$.

The SIS state $\mathbf{x}(t)$, $t \geqslant 0$ is obtained from the processes defined above through the following (deterministic) construction:

1. Initialize $\mathbf{t}_0 = 0$ and $k = 0$.
2. For each $k \geqslant 1$, let $\mathbf{t}_k$ to be the first time in the interval $\big(\mathbf{t}_{k-1}, T_{\mathbf{t}_{k-1},\mathbf{x}_{k-1}}\big)$ for which

$$\int_{\mathbf{t}_{k-1}}^{\mathbf{t}_k} \lambda\big(\varphi(t; \mathbf{t}_{k-1}, \mathbf{x}_{k-1})\big) dt = \boldsymbol{\Delta}_k \tag{3.2}$$

   or $\mathbf{t}_k = T_{\mathbf{t}_{k-1},\mathbf{x}_{k-1}}$ in case

$$\int_{\mathbf{t}_{k-1}}^{T} \lambda\big(\varphi(t; \mathbf{t}_{k-1}, \mathbf{x}_{k-1})\big) dt < \boldsymbol{\Delta}_k, \quad \forall T \in \big(\mathbf{t}_{k-1}, T_{\mathbf{t}_{k-1},\mathbf{x}_{k-1}}\big).$$

3. For this value of $\mathbf{t}_k$, set

$$\mathbf{x}(t) = \varphi(t; \mathbf{t}_{k-1}, \mathbf{x}_{k-1}), \quad \forall t \in [\mathbf{t}_{k-1}, \mathbf{t}_k),$$

4. In case $\mathbf{t}_k < T_{\mathbf{t}_{k-1},\mathbf{x}_{k-1}}$, set

$$\mathbf{x}_k = \mathbf{Z}_k\big(\varphi_{\mathbf{t}_k,\mathbf{t}_{k-1}}(\mathbf{x}_{k-1})\big),$$

   increment k and go back to 2.

[1] We recall that a function $f : \mathcal{X} \to \mathcal{Y}$, $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{Y} \subset \mathbb{R}^m$ is *locally bounded on* $\mathcal{X}$ if for every compact subset $\mathcal{C} \subset \mathcal{X}$, there exists a constant $\kappa_{\mathcal{C}}$ such that $\|f(x)\| \leqslant \kappa_{\mathcal{C}}$, $\forall x \in \mathcal{C}$.

The random variables $\mathbf{t}_k \in [0,\infty)$, $k \geqslant 0$ are called *jump times*, the random variables $\mathbf{x}_k \in \mathcal{X}$, $k \geqslant 0$ are called *jump points*, and the random variable

$$\mathbf{T}_{\max} \coloneqq \sup_{k\geqslant 0} \mathbf{t}_k \leqslant \infty. \tag{3.3}$$

is called the *maximal time* for the process[2]. When $\mathbf{T}_{\max} \overset{\text{wpo}}{=} +\infty$, we say that the process is *globally defined*. Since each $\mathbf{\Delta}_k \overset{\text{wpo}}{>} 0$, the sequence of jump times $\{\mathbf{t}_k\}$ is strictly increasing with probability one and therefore the process $\mathbf{x}(t)$ is *cadlag* on the interval $[0,\mathbf{T}_{\max})$ with probability one, i.e., it is right-continuous on $[0,\mathbf{T}_{\max})$ and has a left-limit at every time $t \in (0,\mathbf{T}_{\max})$. In what follows, when we analyze specific realizations of the random process $\mathbf{x}$, we generally consider outcomes $\omega \in \Omega$ for which $\mathbf{x}$ is cadlag. The set of such outcomes will be denoted by Ω_{cadlag} and, as noted above, $\mathrm{P}(\Omega_{\text{cadlag}}) = 1$.

Based on the sequence of jump times, we define the *jumps counter* $\bar{\mathbf{N}}(t)$ to be the stochastic process that counts the number of jumps up to and including time t, i.e.,

$$\bar{\mathbf{N}}(t) \coloneqq \max\big\{k : \mathbf{t}_k \leqslant t\big\}, \tag{3.4}$$

which is also cadlag with probability one.

Remark 2 (Generation of Sample Paths). *To generate sample paths of* $\mathbf{x}$*, on some interval* $[0,T)$*,* $T<\infty$ *one does not need to start by generating sample paths of the whole processes* $\mathbf{N}$ *and* $\mathbf{Z}(x)$*,* $x \in \mathcal{X}$ *and, only after that, generate* $\mathbf{x}$*; as the above construction seems to require. Instead, one can use the iterative algorithm described above and only generate the random variables* $\mathbf{\Delta}_k$ *and* $\mathbf{x}_k$ *as they are needed:*

1. *At the start of the step 2, generate a sample of the standard exponential random variable* $\mathbf{\Delta}_k$*.*
2. *At step 4, generate a sample for* $\mathbf{x}_k$ *according to the measure* $\nu_{\varphi_{\mathbf{t}_k,\mathbf{t}_{k-1}}(\mathbf{x}_{k-1})}$*.*

All samples should be independent of each other. □

4. Uniform Stability for Not Necessarily Markovian Stochastic Processes. We are interested in stability notions that are applicable not only to the state $\mathbf{x}$ of the SIS (2.1), but also to "output" processes defined by

$$\mathbf{y}(t) \coloneqq h\big(\mathbf{x}(t)\big) \in \mathbb{R}^m, \quad \forall t \in [0,\mathbf{T}_{\max})$$

for some function $h : \mathcal{X} \to \mathbb{R}^m$. Investigating the stability of such output processes is useful, e.g., when one is not concern about boundedness and/or convergence of some components of the whole state $\mathbf{x}$ of the SIS, and the output map h is used to "extract" only those components of the state that are relevant for the stability analysis. For example in the SHSs (2.3), $\mathbf{y}$ often ignores the components of $\mathbf{q}$ that take discrete values (e.g., boolean variables).

Our definitions of stability emphasizes the existence of specific (uniform) bounds on how much $\|\mathbf{y}(t)\|$, $t \geqslant \tau$ can grow and how much it should decay with respect $\|\mathbf{y}(\tau)\|$, as a function of the interval $t-\tau$. These bounds are formulated in terms of class $\mathcal{K}$ and class $\mathcal{KL}$ functions [3], as is commonly done in studying the stability of deterministic systems and was also done in, e.g., [3, 17][4] for stochastic processes. We shall see that our stability results provide explicit formulas for the class $\mathcal{K}$ and class $\mathcal{KL}$ bound estimates and therefore provide insight into the "transient" properties of the process $\mathbf{y}$.

To capture a notion of uniformity in stability, even when $\mathbf{y}$ is not Markovian, we formulate our stability definitions in terms of probabilities conditioned to the natural filtration $\{\mathcal{F}_\tau : \tau \geqslant 0\}$ of the process $\mathbf{y}$. We recall that each σ-algebra $\mathcal{F}_\tau$, $\tau \geqslant 0$ contains all (measurable) events of $\mathbf{y}$ up to and including time τ. In

[2]The above construction may result in a solution with a finite number of jumps, even when $\mathbf{T}_{\max} = +\infty$. In this case, the sets of jump times and jump points are finite.

[3]A function $\alpha : [0,\infty) \to [0,\infty)$ is of class $\mathcal{K}$ if it is continuous, strictly increasing, and $\alpha(0) = 0$. A function $\beta : [0,\infty)\times[0,\infty) \to [0,\infty)$ is of class $\mathcal{KL}$ if $s \mapsto \beta(s,t)$ is of class $\mathcal{K}$ for each fixed $t \geqslant 0$, and $t \mapsto \beta(s,t)$ decreases to zero for each fixed $s \geqslant 0$.

[4]The definitions of stability considered here differ significantly from the ones in [17] since in that reference the quantification with respect to time appears *outside* the probability measure, whereas here [e.g., in (4.1) or (4.2)] it appears *inside*, which is consistent with the commonly used notions of stability in probability [3, 8, 18, 19, 23].

case $\mathbf{y}$ is Markov process, the conditioning to $\mathcal{F}_\tau$ amounts to simply conditioning to $\mathbf{y}(\tau)$, which is common in the definitions of stochastic stability for Markov processes formulated in terms of the transition function of the process, e.g., in [8, 18, 19].

We start by considering notions of stability that limit the probability that specific sample paths "misbehave." In particular, the stochastic process $\mathbf{y} : [0, \mathbf{T}_{\max}) \to \mathcal{X}$ with natural filtration $\{\mathcal{F}_t : t \geqslant 0\}$ is said to be

D1 *uniformly stable in probability* if it is globally defined with probability one, i.e., $\mathbf{T}_{\max} \overset{\text{wpo}}{=} \infty$, and for every $\epsilon > 0$, there exists a class $\mathcal{K}$ function α_ϵ such that

$$\mathrm{P}\left(\exists t \in [\tau, \infty) : \|\mathbf{y}(t)\| > \alpha_\epsilon\big(\|\mathbf{y}(\tau)\|\big) \mid \mathcal{F}_\tau\right) \overset{\text{wpo}}{\leqslant} \epsilon, \quad \forall \tau \geqslant 0; \tag{4.1}$$

D2 *uniformly asymptotically stable in probability* if for every $\epsilon > 0$ there exists a class $\mathcal{KL}$ function β_ϵ such that

$$\mathrm{P}\left(\exists t \in [\tau, \infty) : \|\mathbf{y}(t)\| > \beta_\epsilon\big(\|\mathbf{y}(\tau)\|, t-\tau\big) \mid \mathcal{F}_\tau\right) \leqslant \epsilon, \quad \forall \tau \geqslant 0 \tag{4.2}$$

and consequently[5] $\mathbf{y}(t)$ converges to zero with probability one, i.e.,

$$\mathrm{P}\left(\lim_{t\to\infty} \mathbf{y}(t) = 0\right) = 1, \tag{4.3}$$

and also converges to zero in probability, i.e.,

$$\lim_{t\to\infty} \mathrm{P}\left(\|\mathbf{y}(t)\| \geqslant \epsilon\right) = 0;$$

D3 *uniformly exponentially stable in probability (with decay rate $\mu > 0$)* when it is uniformly exponentially stable and (4.2) holds for functions β_ϵ of the form

$$\beta_\epsilon(s,t) = c_\epsilon e^{-\mu t}, \qquad \forall t, s \geqslant 0,$$

for an appropriately selected constant $c_\epsilon > 0$. Note that, while c_ϵ may depende of ϵ, the decay rate μ is not allowed to do so for uniform exponential stability.

The inequalities (4.1) and (4.2) on the conditional probabilities imply that similar inequalities hold for the unconditional probability, due to the Smoothing Property of conditional expectation [16, p. 45].

We also consider notions of stability that limit the expected value of positive functions of the process $\mathbf{y}$. In particular, given a non-negative function $W : \mathbb{R}^m \to [0, \infty)$, the stochastic process $\mathbf{y} : [0, \mathbf{T}_{\max}) \to \mathcal{X}$ is said to be

D4 *uniformly mean-W stable* if it is globally defined with probability one, i.e., $\mathbf{T}_{\max} \overset{\text{wpo}}{=} \infty$, and there exists a class $\mathcal{K}$ function α such that

$$\mathrm{E}\Big[W\big(\mathbf{y}(T)\big) \mid \mathcal{F}_\tau\Big] \overset{\text{wpo}}{\leqslant} \alpha\big(\|\mathbf{y}(\tau)\|\big), \qquad \forall T \geqslant \tau \geqslant 0;$$

D5 *uniformly stochastically mean-W stable* if there exists a class $\mathcal{K}$ function α such that

$$\int_\tau^\infty \mathrm{E}\Big[W\big(\mathbf{y}(T)\big) \mid \mathcal{F}_\tau\Big] dT \overset{\text{wpo}}{\leqslant} \alpha\big(\|\mathbf{y}(\tau)\|\big), \qquad \forall \tau \geqslant 0;$$

[5] To verify that this is so, note that, for any given $\epsilon > 0$, $\|\mathbf{y}(t)\| \leqslant \beta_\epsilon\big(\|\mathbf{y}(\tau)\|, t-\tau\big)$, $\forall t \in [\tau, \infty)$ implies that $\lim_{t\to\infty} \mathbf{y}(t) = 0$ and therefore

$$\mathrm{P}\left(\lim_{t\to\infty} \mathbf{y}(t) = 0\right) \geqslant \mathrm{P}\left(\forall t \in [\tau, \infty) : \|\mathbf{y}(t)\| \leqslant \beta_\epsilon(\|\mathbf{y}(\tau)\|, t-\tau)\right).$$

Uniform asymptotically stability in probability implies that the right-hand side is greater than or equal to $1-\epsilon$, $\forall \epsilon > 0$ and therefore the left-hand side cannot be strictly smaller than 1.

D6 *uniformly asymptotically mean-W stable* if there exists a class $\mathcal{KL}$ function β such that

$$\mathrm{E}\Big[W\big(\mathbf{y}(T)\big) \bigm| \mathcal{F}_\tau\Big] \overset{\text{wpo}}{\leqslant} \beta\big(\|\mathbf{y}(\tau)\|, T-\tau\big), \qquad \forall T \geqslant \tau \geqslant 0; \tag{4.4}$$

D7 *uniformly exponentially mean-W stable* if it is uniformly asymptotically mean-W stable and (4.4) holds for a function β of the form

$$\beta(s,t) = ce^{-\lambda t}, \qquad \forall t, s \geqslant 0,$$

for appropriately selected constants $c, \lambda > 0$.

For scalar non-negative processes, the function W may be the identity, in which case we simply say that the processes are "mean stable," without reference to W.

Inspired by the terminology used for Markov processes, when the inequalities (4.1)–(4.4) hold, not just for every deterministic time $\tau \geqslant 0$, but also for every stopping time $\boldsymbol{\tau} \overset{\text{wpo}}{\in} [0,\infty)$ for the natural filtration $\{\mathcal{F}_\tau, \tau \geqslant 0\}$ of $\mathbf{y}$, we add the modifier *strongly*, as in *strongly uniformly stable in probability* or *strongly uniformly mean-W stable*. These stronger notions of stability allow us to use the inequalities (4.1)–(4.4), e.g., when $\boldsymbol{\tau}$ is the first time that $\mathbf{y}$ enters a given set. We recall that a random variable $\mathbf{T} \geqslant 0$ is a *stopping time* for the natural filtration of the process $\mathbf{y}(t)$, $t \geqslant 0$ if one can decide whether or not $\mathbf{T} \leqslant T$ on the basis of the knowledge of $\mathbf{y}(t)$, $\forall t \in [0,T]$.

5. Stopped Process. The maximal time $\mathbf{T}_{\max}$ defined in (3.3) may be finite because of finite escape time of the solution to (3.1) or because the jump times have a finite accumulation point. It is therefore convenient to consider a version of the process $\mathbf{x}$ that has been frozen before the solution misbehaves. Formally, given a compact subset $\mathcal{Q}$ of $\mathcal{X}$, we define the *$\mathcal{Q}$-stopping time* to be the first time that $\mathbf{x}(t)$, $t \in [0, \mathbf{T}_{\max})$ leaves the interior of $\mathcal{Q}$. When $\mathbf{x}(t)$ remains in the interior of $\mathcal{Q}$ for all $t \in [0,\mathbf{T}_{\max})$, the $\mathcal{Q}$-stopping time is defined to be equal to $\mathbf{T}_{\max}$. Specifically,

$$\begin{aligned}\mathbf{T}_{\mathcal{Q}} &\coloneqq \inf\Big(\big\{t \in [0,\mathbf{T}_{\max}) : \mathbf{x}(t) \notin \mathrm{Int}(\mathcal{Q})\big\} \cup \{\mathbf{T}_{\max}\}\Big)\\ &= \sup\Big(\big\{T \in [0,\mathbf{T}_{\max}) : \mathbf{x}(t) \in \mathrm{Int}(\mathcal{Q}),\ \forall t \in [0,T)\big\} \cup \{0\}\Big),\end{aligned}$$

where $\mathrm{Int}(\mathcal{Q})$ denotes the interior of $\mathcal{Q}$. As this terminology implies, $\mathbf{T}_{\mathcal{Q}}$ is a stopping times for the natural filtration of the process $\mathbf{x}(t)$, $t \geqslant 0$.

Based on $\mathbf{T}_{\mathcal{Q}}$, we define the *$\mathcal{Q}$-stopped process* and the *$\mathcal{Q}$-stopped jump counter* to be the stochastic processes

$$\mathbf{x}_{\mathcal{Q}}(t) \coloneqq \begin{cases}\mathbf{x}(t) & t \in [0,\mathbf{T}_{\mathcal{Q}})\\ \mathbf{x}(\mathbf{T}_{\mathcal{Q}}) & t \in [\mathbf{T}_{\mathcal{Q}},\infty),\end{cases} \qquad \bar{\mathbf{N}}_{\mathcal{Q}}(t) \coloneqq \begin{cases}\bar{\mathbf{N}}(t) & t \in [0,\mathbf{T}_{\mathcal{Q}})\\ \bar{\mathbf{N}}(\mathbf{T}_{\mathcal{Q}}) & t \in [\mathbf{T}_{\mathcal{Q}},\infty),\end{cases} \tag{5.1}$$

respectively. For these stopped process to be well defined for all times, we must exclude the possibility that $\mathbf{T}_{\max} = \mathbf{T}_{\mathcal{Q}}$ with $\mathbf{T}_{\mathcal{Q}} < \infty$, in which case $\mathbf{x}(\mathbf{T}_{\mathcal{Q}})$ and $\bar{\mathbf{N}}(\mathbf{T}_{\mathcal{Q}})$ in the lower branches in (5.1) would not be well defined. This possibility is excluded by the following lemma, proved in the appendix.

Lemma 5.1 (Stopped Process). *For every compact set $\mathcal{Q} \subset \mathcal{X}$, the following statements hold:*

1. $\mathrm{P}(\mathbf{T}_{\mathcal{Q}} < \infty, \mathbf{T}_{\mathcal{Q}} = \mathbf{T}_{\max}) = 0$.
2. *The expected number of jumps of $\bar{\mathbf{N}}_{\mathcal{Q}}$ is finite on every finite interval $[\tau, T]$, $T \geqslant \tau \geqslant 0$. In particular,*

$$\mathrm{E}\big[\bar{\mathbf{N}}_{\mathcal{Q}}(T) - \bar{\mathbf{N}}_{\mathcal{Q}}(\tau)\big] \leqslant \lambda_{\max}(T-\tau), \quad \forall T \geqslant \tau \geqslant 0 \qquad \lambda_{\max} \coloneqq \sup_{x \in \mathcal{Q}} \lambda(x) < \infty,$$

and consequently $\bar{\mathbf{N}}_{\mathcal{Q}}(T) \overset{\text{wpo}}{<} \infty$, $\forall T \in [0,\infty)$.

3. *The $\mathcal{Q}$-stopped process is cadlag on the whole interval $[0,\infty)$ with probability one.*

These statements can be concisely stated as: $\mathrm{P}(\Omega_{\mathcal{Q}}) = 1$*, for*

$$\Omega_{\mathcal{Q}} \coloneqq \Big\{\omega \in \Omega : t \mapsto x_{\mathcal{Q}}(t;\omega) \textit{ is cadlag on } [0,\infty) \textit{ and } \bar{\mathbf{N}}_{\mathcal{Q}}(T;\omega) < \infty,\ \forall T \geqslant 0\Big\}. \qquad \square$$

6. Dynkin's Formula for the Stopped Process. The Lyapunov stability results in this paper rely on studying the evolution of scalar-valued "Lyapunov functions" $V : \mathcal{X} \to \mathbb{R}$ along solutions to the SIS (2.1). To maximize the flexibility provided by this analysis method, it is crucial to be able to draw such Lyapunov functions from a large universe of functions. The set of functions from which we can draw Lyapunov functions is denoted by $\mathcal{D}$ and consists of the set of functions $V : \mathcal{X} \to \mathbb{R}$ that satisfy the following properties:

P1 The function V is locally bounded on $\mathcal{X}$.

P2 For every fixed $t_0 \geqslant 0$, $x_0 \in \mathcal{X}$ the map $t \mapsto V\big(\varphi(t; t_0, x_0)\big)$ is absolutely continuous on $[t_0, T_{t_0,x_0})$.

P3 For every $x \in \mathcal{X}$ the integral $\int_{\mathcal{X}} V(z)\nu_x(dz)$ exists and the function

$$x \mapsto \int_{\mathcal{X}} V(z)\nu_x(dz) - V(x)$$

is locally bounded on $\mathcal{X}$.

P4 For every fixed $t_0 \geqslant 0$, $x_0 \in \mathcal{X}$ the map $t \mapsto \int_{\mathcal{X}} V(z)\nu_{\varphi(t;t_0,x_0)}(dz) - V\big(\varphi(t; t_0, x_0)\big)$ is integrable on any finite interval $[t_0, T]$, $\forall T < T_{t_0,x_0}$.

As we shall see shortly in Lemma 6.1, every locally Lipschitz function V satisfies Properties P1 and P2. Properties P3 and P4 can also be deduced from mild continuity conditions, for large classes of jump measures:

1. For a SIS with (stochastic $K > 1$ or deterministic $K = 1$) discrete jump maps for which the jump measure is given by (2.2), we have that

$$\Big| \int_{\mathcal{X}} V(z)\nu_x(dz) - V(x)\Big| = \Big| \sum_{i=1}^{K} V\big(\rho_i(x)\big)p_i(x) - V(x)\Big| \leqslant \sum_{i=1}^{K} \Big|V\big(\rho_i(x)\big)\Big| + \big|V(x)\big|.$$

   Therefore, when V is locally Lipschitz (see P2), the property P3 holds if the functions $x \mapsto \rho_i(x)$, $\forall i \in \{1, 2, \ldots, K\}$ are locally bounded on $\mathcal{X}$ (e.g., if they are continuous on $\mathcal{X}$). Continuity of the functions $x \mapsto \rho_i(x)$ would then also suffice for P4.
2. Suppose that every $x \in \mathcal{X}$, ν_x has a probability density function f_x and therefore

$$\int_{\mathcal{X}} V(z)\nu_x(dz) - V(x) = \int_{\mathcal{X}} V(z)f_x(z)dz - V(x), \quad \forall x \in \mathcal{X}.$$

   In this case, the property P3 holds if the function $x \mapsto \int_{\mathcal{X}} V(z)f_x(z)dz - V(x)$ is locally bounded on $\mathcal{X}$ (e.g., if it is continuous on $\mathcal{X}$). Continuity of this function would also suffice for P4.

   Combinations of continuous and discrete distributions are also possible with appropriate continuity assumptions.

Property P2, allows us to conclude that

$$V\big(\varphi(t; t_0, x_0)\big) = V(x_0) + \int_{t_0}^{t} L_f V\big(\varphi(s; t_0, x_0)\big)ds, \qquad \forall t_0 \geqslant 0, x_0 \in \mathcal{X}, t \in [t_0, T_{t_0,x_0}), \tag{6.1}$$

for an appropriate defined function $L_f V : \mathcal{X} \to \mathbb{R}$. When V is differentiable, we simply have

$$L_f V(x) = \frac{\partial V(x)}{\partial x} \cdot f(x), \quad \forall x \in \mathcal{X},$$

but (6.1) can hold even when V is not differentiable. The following Lemma is useful to upper-bound $L_f V$ for functions that are not differentiable [12, 22]. As we shall see later, appropriate upper bounds on $L_f V(x)$ suffice to establish the stability of SISs.

Lemma 6.1. *Suppose that V is locally Lipschitz and that there exists a continuous function $\alpha : \mathcal{X} \to \mathbb{R}$ such that*

$$\frac{\partial V(x)}{\partial x} \cdot f(x) \leqslant \alpha(x), \quad \forall x \in \mathcal{X}\backslash\mathcal{N}$$

for some zero Lebesgue-measure subset $\mathcal{N}$ of $\mathcal{X}$ that contains the points at which V is not differentiable. Then Properties P1, P2 and equation (6.1) hold with

$$L_f V(x) \leqslant \alpha(x), \quad \forall x \in \mathcal{X}. \qquad \square$$

In view of Properties P1–P4 and the discussion above, for each function $V \in \mathcal{D}$ we can construct a function $LV : \mathcal{X} \to \mathbb{R}$ defined by

$$(LV)(x) \coloneqq L_f V(x) + \lambda(x)\Big(\int_{\mathcal{X}} V(z)\nu_x(dz) - V(x)\Big), \quad \forall x \in \mathcal{X}.$$

This construction defines an operator L, which we call the *extended generator of the SIS*, that maps $\mathcal{D}$ to the set of functions from $\mathcal{X}$ to $\mathbb{R}$. This terminology is inspired by the Markov Process literature, but it is important to emphasize that in that literature the extended generator is typically defined as having for domain the set of functions for which Dynkin's equation [akin to (7.5a) below] holds for the Markov process, whereas here we take the domain of L to be the whole set $\mathcal{D}$, *even if Dynkin's equation does not hold for some elements of* $\mathcal{D}$.

The following theorem is the main result of this section and provides a Dynkin's-like formula (6.2) for the stopped process. The proof of this result (in appendix) is heavily inspired by the results in [5] but avoids some of the standing assumptions in this reference (including Assumption 24.8, item 3 of Assumption 24.8, and item 3 in Theorem 26.14). While these assumptions would be easy to verify for a SIS with uniformly bounded transition intensity λ and for uniformly bounded functions $V \in \mathcal{D}$, Lyapunov functions always need to be unbounded and many applications also require unbounded intensities. We can avoid these assumptions essentially because the processes considered here are less general than the ones in [5] and, most importantly, because we derive a Dynkin's-like formula for the *stopped process* and not for the original process.

Theorem 6.2 (Dynkin's Formula for the Stopped Process). *For every compact set $\mathcal{Q} \subset \mathcal{X}$, every function $V \in \mathcal{D}$, and every pair of times $0 \leqslant \tau \leqslant T < \infty$ such that $\tau \overset{\text{wpo}}{<} \mathbf{T}_{\max}$, we have that*

$$\mathrm{E}\left[V\big(\mathbf{x}_{\mathcal{Q}}(T)\big) \mid \mathcal{F}_\tau\right] \overset{\text{wpo}}{=} V\big(\mathbf{x}_{\mathcal{Q}}(\tau)\big) + \mathrm{E}\left[\int_\tau^T (L_{\mathcal{Q}}V)\big(\mathbf{x}_{\mathcal{Q}}(t)\big)dt \mid \mathcal{F}_\tau\right] \tag{6.2a}$$

$$\overset{\text{wpo}}{=} V\big(\mathbf{x}_{\mathcal{Q}}(\tau)\big) + \int_\tau^T \mathrm{E}\left[(L_{\mathcal{Q}}V)\big(\mathbf{x}_{\mathcal{Q}}(t)\big) \mid \mathcal{F}_\tau\right]dt, \tag{6.2b}$$

where $\{\mathcal{F}_t : t \geqslant 0\}$ denotes the natural filtration of the joint process $\big(\mathbf{x}_{\mathcal{Q}}(t), \bar{\mathbf{N}}_{\mathcal{Q}}(t)\big)$, $t \geqslant 0$;

$$(L_{\mathcal{Q}}V)(x) \coloneqq \begin{cases} (LV)(x) & x \in \mathrm{Int}(\mathcal{Q}) \\ 0 & x \in \mathcal{X} \backslash \mathrm{Int}(\mathcal{Q}); \end{cases} \tag{6.3}$$

and $\mathrm{Int}(\mathcal{Q})$ denotes the interior of $\mathcal{Q}$. The two equations in (6.2) also hold when $\{\mathcal{F}_t : t \geqslant 0\}$ is the natural filtration of the $\mathcal{Q}$-stopped process $\mathbf{x}_{\mathcal{Q}}(t)$, $t \geqslant 0$. $\square$

A lot more could be deduced regarding the process $x_{\mathcal{Q}}$. We shall state without proof that $x_{\mathcal{Q}}$ is a strong Markov process whose extended generator is given by (6.3) with a domain that contains $\mathcal{D}$. However, all that we need for the remainder of this paper are the specific results stated in Theorem 6.2.

7. Global Existence. We say that $V : \mathcal{X} \to \mathbb{R}$ in $\mathcal{D}$ is an *exponentially-bounded on-the-average Lyapunov function* for the SIS (2.1) if it satisfies the following three properties:

P5 V is *non-negative* in the sense that

$$V(x) \geqslant 0, \quad \forall x \in \mathcal{X}. \tag{7.1}$$

We emphasize that we do not require $V(0) = 0$ and therefore the right-hand side of (7.1) can be strictly positive for every $x \in \mathcal{X}$.

P6 V is *radially unbounded in* $\mathcal{X}$ in the sense that, $\forall m \geqslant 0$, the set

$$\mathcal{Q}_m \coloneqq \{x \in \mathcal{X} : V(x) \leqslant m\} \tag{7.2}$$

is compact.

P7 LV is *geometrically-bounded* in the sense that there exists a finite constant $c \in \mathbb{R}$ such that

$$(LV)(x) \coloneqq L_f V(x) + \lambda(x)\Big(\int_{\mathcal{X}} V(z)\nu_x(dz) - V(x)\Big) \leqslant cV(x), \quad \forall x \in \mathcal{X}. \tag{7.3}$$

The constant c in (7.3) is called the *expansion rate* of V.

The terminology "exponentially-bounded on-the-average Lyapunov function" is motivated by the fact that, as we will see in Theorem 7.1 below, the *expected value* of $V\big(\mathbf{x}(t)\big)$ can grow at most exponentially along solutions of the SIS (2.1). The use of a condition like (7.3) to prove global existence of solution for Markov processes can be found in [20].

The result that follows allow us to establish global existence of solution for the process $\mathbf{x}$. This result also provides upper bounds on the expected value of the Lyapunov function [cf. (7.5c)] and on the probability that the Lyapunov function grows beyond a given constant m [cf. (7.6)]. These upper bounds hold, not only for arbitrary deterministic times, but also for (random) stopping times. The Lyapunov stability results in Section 8 make extensive use of this theorem and do require bounds that hold for stopping times.

Theorem 7.1 (Global Existence). *Let $V \in \mathcal{D}$ be an exponentially-bounded on-the-average Lyapunov function for the SIS (2.1) with expansion rate c, and assume that $\mathrm{E}\big[V\big(\mathbf{x}(0)\big)\big] < \infty$. The following properties hold for the natural filtration $\{\mathcal{F}_t : t \geqslant 0\}$ of the process $\mathbf{x}(t)$, $t \geqslant 0$.*

1. *The state $\mathbf{x}$ is globally defined in $[0, \infty)$ with probability one, i.e.,*

$$\mathrm{P}(\mathbf{T}_{\max} = +\infty) = 1. \tag{7.4}$$

2. *For every $T \geqslant \tau \geqslant 0$,*

$$\mathrm{E}\Big[V\big(\mathbf{x}(T)\big) \mid \mathcal{F}_\tau\Big] \overset{\text{wpo}}{=} V\big(\mathbf{x}(\tau)\big) + \int_\tau^T \mathrm{E}\Big[(LV)\big(\mathbf{x}(t)\big) \mid \mathcal{F}_\tau\Big]dt \tag{7.5a}$$

$$\overset{\text{wpo}}{=} V\big(\mathbf{x}(\tau)\big) + \mathrm{E}\Big[\int_\tau^T (LV)\big(\mathbf{x}(t)\big)dt \mid \mathcal{F}_\tau\Big] \tag{7.5b}$$

$$\overset{\text{wpo}}{\leqslant} e^{c(t-\tau)}V\big(\mathbf{x}(\tau)\big). \tag{7.5c}$$

The inequality (7.5c), also holds for the more general case where τ is a stopping time for the filtration $\{\mathcal{F}_t : t \geqslant 0\}$.

3. *For any every $m > 0$ and every stopping time $\boldsymbol{\tau} \overset{\text{wpo}}{\in} [0, \infty)$ for $\{\mathcal{F}_t : t \geqslant 0\}$, we have that*

$$\mathrm{P}\left(\exists t \in [\boldsymbol{\tau}, \infty) : V\big(\mathbf{x}(t)\big) \geqslant m \mid \mathcal{F}_{\boldsymbol{\tau}}\right) \overset{\text{wpo}}{\leqslant} \frac{e^{c(t-\boldsymbol{\tau})}V\big(\mathbf{x}(\boldsymbol{\tau})\big)}{m}, \qquad \forall m > 0. \tag{7.6}$$

(The inequality (7.6), also holds for the special case where $\boldsymbol{\tau}$ is a deterministic time). □

The following equality is a direct consequence of (7.5a) in Theorem 7.1:

$$\mathrm{E}\Big[V\big(\mathbf{x}(T_2)\big) \mid \mathcal{F}_\tau\Big] - \mathrm{E}\Big[V\big(\mathbf{x}(T_1)\big) \mid \mathcal{F}_\tau\Big] = \int_{T_1}^{T_2} \mathrm{E}\Big[(LV)\big(\mathbf{x}(t)\big)dt \mid \mathcal{F}_\tau\Big]dt, \qquad \forall T_1, T_2 \geqslant \tau,$$

and therefore we can conclude from the Fundamental Theorem of Calculus for Lebesgue Integrals [6, p. 102] that $t \mapsto \mathrm{E}\big[V\big(\mathbf{x}(t)\big) \mid \mathcal{F}_\tau\big]$ is absolutely continuous and differentiable almost everywhere with probability one. This observation is formalized in the corollary that follows.

COROLLARY 7.2. *Let $V \in \mathcal{D}$ be an exponentially-bounded on-the-average Lyapunov function for the SIS (2.1) for which $\mathrm{E}\big[V\big(\mathbf{x}(0)\big)\big] < \infty$. For every $\tau \geqslant 0$, the function $t \mapsto \mathrm{E}\big[V\big(\mathbf{x}(t)\big) \mid \mathcal{F}_\tau\big]$ is absolutely continuous on $[\tau, \infty)$ and differentiable almost everywhere, with probability one. Moreover,*

$$\frac{d}{dt}\,\mathrm{E}\Big[V\big(\mathbf{x}(t)\big) \mid \mathcal{F}_\tau\Big] \overset{\text{wpo}}{=} \mathrm{E}\Big[(LV)\big(\mathbf{x}(t)\big) \mid \mathcal{F}_\tau\Big], \qquad \forall t \overset{\text{ae}}{\geqslant} \tau. \tag{7.7}$$

□

The $\overset{\text{wpo}}{=}$ in (7.7) refers to the fact that the random process $\mathrm{E}\big[V\big(\mathbf{x}(t)\big) \mid \mathcal{F}_\tau\big]$ is absolutely continuous on a subset of Ω with probability one. On the other hand, the $\overset{\text{ae}}{\geqslant}$ refers to the fact that, for those sample paths of $\mathbf{v}(t)$ that are absolutely continuous, there may be a set of times with zero Lebesgue measure on which the derivative $\frac{d}{dt}\,\mathrm{E}\big[V\big(\mathbf{x}(t)\big) \mid \mathcal{F}_\tau\big]$ does not exist.

Proof of Theorem 7.1 – Global existence. The proof of Theorem 7.1 uses two key results from the theory of martingales: Doob's stopping Theorem and Doob's martingale inequality. The following result, which is used to prove Theorem 7.1, combines Doob's stopping Theorem [14, Theorem 1.39, p. 10] with Markov's inequality [4, Theorem 1, p. 86] to obtain a version of Doob's martingale inequality that holds for stopping times[6].

LEMMA 7.3. *Let $\boldsymbol{\tau} \overset{\text{wpo}}{\in} [0, \infty)$ be a finite stopping time and $\mathbf{M}(t)$, $t \geqslant 0$ a nonnegative supermartingale for a filtration $\{\mathcal{F}_t : t \geqslant 0\}$. Then*

$$\mathrm{P}\left(\exists t \in [\boldsymbol{\tau}, \infty) : \mathbf{M}(t) \geqslant m \mid \mathcal{F}_{\boldsymbol{\tau}}\right) \overset{\text{wpo}}{\leqslant} \frac{\mathbf{M}(\boldsymbol{\tau})}{m}, \qquad \forall m > 0.$$

This lemma will be applied to the supermartingale defined in the following lemma.

LEMMA 7.4. *Under the assumptions of Theorem 7.1, for a given $m > 0$, define*

$$\mathbf{M}_{\mathcal{Q}_m}(t) \coloneqq e^{-ct} V\big(\mathbf{x}_{\mathcal{Q}_m}(t)\big) \geqslant 0, \quad \forall t \geqslant 0,$$

where $\mathcal{Q}_m$ denotes the compact set defined by (7.2). Then $\mathbf{M}_{\mathcal{Q}_m}$ is a nonnegative supermartingale of the natural filtration $\{\bar{\mathcal{F}}_t : t \geqslant 0\}$ of the $\mathcal{Q}_m$-stopped process $\mathbf{x}_{\mathcal{Q}_m}$, i.e.,

$$\mathrm{E}[\mathbf{M}_{\mathcal{Q}_m}(t)] < \infty \qquad\qquad \mathrm{E}[\mathbf{M}_{\mathcal{Q}_m}(t) \mid \bar{\mathcal{F}}_\tau] \overset{\text{wpo}}{\leqslant} \mathbf{M}_{\mathcal{Q}_m}(\tau), \qquad\qquad \forall t \geqslant \tau \geqslant 0. \tag{7.8}$$

Proof of Lemma 7.4. Because of (6.2b) in Theorem 6.2, we know that the process $t \mapsto \mathrm{E}\big[V\big(\mathbf{x}_{\mathcal{Q}_m}(t)\big) \mid \bar{\mathcal{F}}_\tau\big]$ is absolutely continuous and the Fundamental Theorem of Calculus for Lebesgue Integrals [6, p. 102] allows us to conclude that this process is differentiable almost everywhere with probability one, with

$$\frac{d}{dt}\,\mathrm{E}\Big[V\big(\mathbf{x}_{\mathcal{Q}_m}(t)\big) \mid \mathcal{F}_\tau\Big] \overset{\text{wpo}}{=} \mathrm{E}\Big[(LV)\big(\mathbf{x}_{\mathcal{Q}_m}(t)\big) \mid \mathcal{F}_\tau\Big], \qquad \forall t \overset{\text{ae}}{\geqslant} \tau \geqslant 0. \tag{7.9}$$

For $t < \mathbf{T}_{\mathcal{Q}_m}$, we have that $\mathbf{x}_{\mathcal{Q}_m}(t) = \mathbf{x}(t) \in \mathcal{Q}_m$ and therefore

$$\mathrm{E}\big[\mathbf{M}_{\mathcal{Q}_m}(t) \mid t < \mathbf{T}_{\mathcal{Q}_m}\big] = e^{-ct}\,\mathrm{E}\big[V\big(\mathbf{x}(t)\big) \mid t < \mathbf{T}_{\mathcal{Q}_m}\big] \leqslant e^{-ct} \sup_{x \in \mathcal{Q}_m} V(x), \quad \forall t \geqslant 0$$

and for $t \geqslant \mathbf{T}_{\mathcal{Q}_m}$, we have that $\mathbf{x}_{\mathcal{Q}_m}(t) = \mathbf{x}(\mathbf{T}_{\mathcal{Q}_m}) \in \mathcal{Q}_m$ and therefore

$$\begin{aligned}\mathrm{E}\big[\mathbf{M}_{\mathcal{Q}_m}(t) \mid t \geqslant \mathbf{T}_{\mathcal{Q}_m}\big] &= e^{-ct}\,\mathrm{E}\big[V\big(\mathbf{x}(\mathbf{T}_{\mathcal{Q}_m})\big) \mid t \geqslant \mathbf{T}_{\mathcal{Q}_m}\big] \\ &\leqslant e^{ct} \begin{cases} \mathrm{E}[V(\mathbf{x}(0))] & \mathbf{T}_{\mathcal{Q}_m} = 0 \\ \sup_{x \in \mathcal{Q}_m} V(x) & \mathbf{T}_{\mathcal{Q}_m} > 0 \text{ is not a jump time} \\ \sup_{x \in \mathcal{Q}_m} \int_{\mathcal{X}} V(z)\nu_x(dz) & \mathbf{T}_{\mathcal{Q}_m} > 0 \text{ is a jump time.} \end{cases}\end{aligned}$$

[6] We believe that this result is not novel, but since we were unable to find it stated in the literature, we include a short proof in the appendix.

Combining these two cases, we thus conclude that, because $V \in \mathcal{D}$,

$$\mathrm{E}[\mathbf{M}_{\mathcal{Q}_m}(t)] \leqslant e^{-ct} \max\left\{ \mathrm{E}[V(\mathbf{x}(0))],\ \sup_{x\in\mathcal{Q}_m} V(x),\ \sup_{x\in\mathcal{Q}_m} \int_{\mathcal{X}} V(z)\nu_x(dz) \right\} < \infty,$$

$\forall t \geqslant 0$, which proves the left-hand side inequality in (7.8). Moreover, because of (7.9), (6.3), and (7.3), we also have

$$\begin{aligned}\frac{d}{dt}\mathrm{E}[\mathbf{M}_{\mathcal{Q}_m}(t) \mid \bar{\mathcal{F}}_\tau] &\overset{\text{wpo}}{=} e^{-ct}\Big(\frac{d}{dt}\mathrm{E}\Big[V\big(\mathbf{x}_{\mathcal{Q}_m}(t)\big) \mid \mathcal{F}_\tau\Big] - c\,\mathrm{E}\Big[V\big(\mathbf{x}_{\mathcal{Q}_m}(t)\big) \mid \mathcal{F}_\tau\Big]\Big) \\ &\overset{\text{wpo}}{=} e^{-ct}\Big(\mathrm{E}\Big[(LV)\big(\mathbf{x}_{\mathcal{Q}_m}(t)\big) \mid \mathcal{F}_\tau\Big] - c\,\mathrm{E}\Big[V\big(\mathbf{x}_{\mathcal{Q}_m}(t)\big) \mid \mathcal{F}_\tau\Big]\Big) \overset{\text{wpo}}{\leqslant} 0\end{aligned}$$

$\forall t \overset{\text{ae}}{\geqslant} \tau \geqslant 0$. The right-hand side inequality in (7.8) follows from this and the absolutely continuous of $t \mapsto \mathrm{E}[\mathbf{M}_{\mathcal{Q}_m}(t) \mid \bar{\mathcal{F}}_\tau]$. ∎

Proof of Theorem 7.1. Applying Lemma 7.3 to $\mathbf{M}_{\mathcal{Q}_m}(t)$ (cf. Lemma 7.4) we conclude that

$$\mathrm{P}\left(\exists t \in [0,\infty) : V\big(\mathbf{x}_{\mathcal{Q}_m}(t)\big) \geqslant me^{ct} \mid \bar{\mathcal{F}}_\tau\right) \overset{\text{wpo}}{\leqslant} \frac{e^{-c\tau}V\big(\mathbf{x}_{\mathcal{Q}_m}(\tau)\big)}{m}, \quad \forall m > 0, \tau \geqslant 0. \tag{7.10}$$

To prove (7.4), pick some $T \in (0,\infty)$ and consider a generic outcome $\omega \in \Omega_{\mathrm{Q}_m}$ for which $\mathbf{T}_{\max}(\omega) \leqslant T$, with Ω_{Q_m} defined as in Lemma 5.1. The corresponding $\mathbf{x}(t;\omega)$ cannot remain in the interior of $\mathcal{Q}_m$ on $[0,T]$ since otherwise $\mathbf{x}(t;\omega)$ would be equal to $\mathbf{x}_{\tau,\mathcal{Q}_m}$ on $[0,T]$ and we would have $\mathbf{T}_{\max}(\omega) > T$. Therefore

$$\begin{aligned}\mathbf{T}_{\max}(\omega) \leqslant T \quad &\Rightarrow \quad \exists t \in [0,T] : V(\mathbf{x}(t;\omega)) \geqslant m \\ &\Rightarrow \quad \exists t \in [0,T] : V(\mathbf{x}_{\mathcal{Q}_m}(t;\omega)) \geqslant m \\ &\Rightarrow \quad \exists t \in [0,T] : V(\mathbf{x}_{\mathcal{Q}_m}(t;\omega)) \geqslant me^{-|c|T}e^{ct},\end{aligned}$$

where the last inequality follows from the fact that $m \geqslant me^{-|c|T}e^{ct}$, $\forall t \in [0,T]$, $c \in \mathbb{R}$. This means that

$$\mathrm{P}(\mathbf{T}_{\max} \leqslant T \mid \bar{\mathcal{F}}_\tau) \leqslant \mathrm{P}\left(\exists t \in [0,T] : V\big(\mathbf{x}_{\mathcal{Q}}(t)\big) \geqslant me^{-|c|T}e^{ct} \mid \bar{\mathcal{F}}_\tau\right) \overset{\text{wpo}}{\leqslant} \frac{e^{-c\tau}\mathbf{M}_{\mathcal{Q}_m}(\tau)}{me^{-|c|T}},$$

where the last inequality follows from (7.10). Since $\mathbf{M}_{\mathcal{Q}_m}(\tau) \overset{\text{wpo}}{\geqslant} 0$ and $\mathrm{E}[\mathbf{M}_{\mathcal{Q}_m}(\tau)] < \infty$, we must have that $\mathbf{M}_{\mathcal{Q}_m}(\tau) \overset{\text{wpo}}{<} \infty$. Therefore, since the inequality above holds for every $m > 0$, we conclude that we must have $\mathrm{P}(\mathbf{T}_{\max} \leqslant T \mid \bar{\mathcal{F}}_\tau) \overset{\text{wpo}}{=} 0$, $\forall T \in (0,\infty)$, which proves (7.4)[7].

To prove (7.5), we start from (6.2a) in Theorem 6.2, which we re-write as

$$\mathrm{E}\left[V\big(\mathbf{x}_{\mathcal{Q}_m}(T)\big) \mid \bar{\mathcal{F}}_\tau\right] \overset{\text{wpo}}{=} V\big(\mathbf{x}_{\mathcal{Q}_m}(\tau)\big) + \int_\tau^T \mathrm{E}\Big[(L_{\mathcal{Q}_m}V)\big(\mathbf{x}_{\mathcal{Q}_m}(t)\big) \mid \bar{\mathcal{F}}_\tau\Big]dt, \quad \forall T \geqslant \tau \geqslant 0. \tag{7.11}$$

We first show that this equation still holds if we replace the natural filtration $\bar{\mathcal{F}}_t$ of $\mathbf{x}_{\mathcal{Q}_m}$ by the natural filtration $\mathcal{F}_t$ of $\mathbf{x}$. After that, by making $m \to \infty$, we will be able to replace the $\mathcal{Q}_m$-stopped process $\mathbf{x}_{\mathcal{Q}_m}$ by $\mathbf{x}$ in all the expectations in (7.11).

According to the definition of conditional expectation,

$$\boldsymbol{\eta}_m \coloneqq \mathrm{E}\left[V\big(\mathbf{x}_{\mathcal{Q}_m}(T)\big) \mid \bar{\mathcal{F}}_\tau\right],$$

[7] Here, we are using the Smoothing Property of the conditional expectation [16, p. 45] to conclude that, given an event $A \in \mathcal{F}$ with indicator function $\mathbf{I}_A$ and a σ-algebra $\bar{\mathcal{F}} \subset \mathcal{F}$, we have that $\mathrm{E}\left[\mathrm{P}(A|\mathcal{F})\right] = \mathrm{E}\left[\mathrm{E}[\mathbf{I}_A|\mathcal{F}]\right] = \mathrm{E}[\mathbf{I}_A] = \mathrm{P}(A)$. Therefore, $\mathrm{P}(A|\mathcal{F}) \overset{\text{wpo}}{\leqslant} c$ implies that $\mathrm{P}(A) \leqslant c$.

is the (almost surely unique) random variable that is $\bar{\mathcal{F}}_\tau$-measurable and for which

$$\mathrm{E}[\boldsymbol{\eta}_m \mathbf{I}_A] = \mathrm{E}\big[V\big(\mathbf{x}_{\mathcal{Q}_m}(T)\big)\mathbf{I}_A\big], \qquad \forall A \in \bar{\mathcal{F}}_\tau \tag{7.12}$$

where $\mathbf{I}_A : \Omega \to \{0,1\}$ denotes the indicator function of the set A [16, p. 44]. Since $\bar{\mathcal{F}}_\tau \subset \mathcal{F}_\tau$, $\boldsymbol{\eta}_m$ is also $\mathcal{F}_\tau$-measurable. Therefore, to prove that

$$\boldsymbol{\eta}_m \coloneqq \mathrm{E}\big[V\big(\mathbf{x}_{\mathcal{Q}_m}(T)\big) \mid \bar{\mathcal{F}}_\tau\big] \overset{\text{wpo}}{=} \mathrm{E}\big[V\big(\mathbf{x}_{\mathcal{Q}_m}(T)\big) \mid \mathcal{F}_\tau\big], \qquad \forall T \geqslant \tau$$

it suffices to show that (7.12) actually holds for all $A \in \mathcal{F}_\tau$. To this effect, we pick some $A \in \mathcal{F}_\tau$ and compute

$$\begin{aligned}\mathrm{E}[\boldsymbol{\eta}_m \mathbf{I}_A] &= \mathrm{E}\big[\boldsymbol{\eta}_m \mathbf{I}_{A\cap\{\mathbf{T}_{\mathcal{Q}_m}\leqslant\tau\}} + \boldsymbol{\eta}_m \mathbf{I}_{A\cap\{\mathbf{T}_{\mathcal{Q}_m}>\tau\}}\big]\\ &= \mathrm{E}\big[\boldsymbol{\eta}_m \mathbf{I}_{A\cap\{\mathbf{T}_{\mathcal{Q}_m}\leqslant\tau\}}\big] + \mathrm{E}\big[\boldsymbol{\eta}_m \mathbf{I}_{A\cap\{\mathbf{T}_{\mathcal{Q}_m}>\tau\}}\big].\end{aligned} \tag{7.13}$$

When $\mathbf{T}_{\mathcal{Q}_m} > \tau$, we have that $\mathbf{x}_{\mathcal{Q}_m}(s) = \mathbf{x}(s)$, $\forall s \in [0,\tau]$ and therefore $A \cap \{\mathbf{T}_{\mathcal{Q}_m} > \tau\} \in \bar{\mathcal{F}}_\tau$. Because of (7.12), we thus conclude that

$$\mathrm{E}\big[\boldsymbol{\eta}_m \mathbf{I}_{A\cap\{\mathbf{T}_{\mathcal{Q}_m}>\tau\}}\big] = \mathrm{E}\big[V\big(\mathbf{x}_{\mathcal{Q}_m}(T)\big)\mathbf{I}_{A\cap\{\mathbf{T}_{\mathcal{Q}_m}>\tau\}}\big]. \tag{7.14}$$

On the other hand, when $\mathbf{T}_{\mathcal{Q}_m} \leqslant \tau$, we have that $\mathbf{x}_{\mathcal{Q}_m}(T) = \mathbf{x}_{\mathcal{Q}_m}(\tau)$ and therefore

$$\begin{aligned}\mathrm{E}\big[\boldsymbol{\eta}_m \mathbf{I}_{A\cap\{\mathbf{T}_{\mathcal{Q}_m}\leqslant\tau\}}\big] &= \mathrm{E}\big[V\big(\mathbf{x}_{\mathcal{Q}_m}(\tau)\big)\mathbf{I}_{A\cap\{\mathbf{T}_{\mathcal{Q}_m}\leqslant\tau\}}\big]\\ &= \mathrm{E}\big[V\big(\mathbf{x}_{\mathcal{Q}_m}(T)\big)\mathbf{I}_{A\cap\{\mathbf{T}_{\mathcal{Q}_m}\leqslant\tau\}}\big].\end{aligned} \tag{7.15}$$

Replacing (7.14) and (7.15) in (7.13), we conclude that

$$\begin{aligned}\mathrm{E}[\boldsymbol{\eta}_m \mathbf{I}_A] &= \mathrm{E}\big[V\big(\mathbf{x}_{\mathcal{Q}_m}(T)\big)\mathbf{I}_{A\cap\{\mathbf{T}_{\mathcal{Q}_m}>\tau\}}\big] + \mathrm{E}\big[V\big(\mathbf{x}_{\mathcal{Q}_m}(T)\big)\mathbf{I}_{A\cap\{\mathbf{T}_{\mathcal{Q}_m}\leqslant\tau\}}\big]\\ &= \mathrm{E}\big[V\big(\mathbf{x}_{\mathcal{Q}_m}(T)\big)\mathbf{I}_A\big]\end{aligned}$$

and therefore (7.12) does indeed hold for all $A \in \mathcal{F}_\tau$. The proof that

$$\mathrm{E}\Big[\int_\tau^T (L_{\mathcal{Q}_m}V)\big(\mathbf{x}_{\mathcal{Q}_m}(t)\big)dt \mid \bar{\mathcal{F}}_\tau\Big] \overset{\text{wpo}}{=} \mathrm{E}\Big[\int_\tau^T (L_{\mathcal{Q}_m}V)\big(\mathbf{x}_{\mathcal{Q}_m}(t)\big)dt \mid \mathcal{F}_\tau\Big],$$

follows the same steps and therefore we do not repeat it here. We thus conclude that

$$\mathrm{E}\big[V\big(\mathbf{x}_{\mathcal{Q}_m}(T)\big) \mid \mathcal{F}_\tau\big] \overset{\text{wpo}}{=} V\big(\mathbf{x}_{\mathcal{Q}_m}(\tau)\big) + \mathrm{E}\Big[\int_\tau^T (L_{\mathcal{Q}_m}V)\big(\mathbf{x}_{\mathcal{Q}_m}(t)\big)dt \mid \mathcal{F}_\tau\Big], \tag{7.16}$$

$\forall m > 0,\ T \geqslant \tau \geqslant 0$. Our goal is now to replace the $\mathbf{x}_{\mathcal{Q}}$ by $\mathbf{x}$ in the equation above, by making $m \to \infty$. Using the notation $a \wedge b \coloneqq \min\{a,b\}$, we conclude from the definitions of the stopped process and of $L_{\mathcal{Q}_m}V$ that

$$\begin{aligned}\mathbf{x}_{\mathcal{Q}_m}(\tau) &= \mathbf{x}\big(\tau \wedge \mathbf{T}_{\mathcal{Q}_m}\big),\\ \mathbf{x}_{\mathcal{Q}_m}(T) &= \mathbf{x}\big(T \wedge \mathbf{T}_{\mathcal{Q}_m}\big),\\ (L_{\mathcal{Q}_m}V)\big(\mathbf{x}_{\mathcal{Q}_m}(t)\big) &= \begin{cases}(LV)\big(\mathbf{x}(t)\big) & t < \mathbf{T}_{\mathcal{Q}_m}\\ 0 & t \geqslant \mathbf{T}_{\mathcal{Q}_m}\end{cases}\end{aligned}$$

and therefore

$$\begin{aligned}\mathrm{E}\Big[V\big(\mathbf{x}(T \wedge \mathbf{T}_{\mathcal{Q}_m})\big) \mid \mathcal{F}_\tau\Big] &\overset{\text{wpo}}{=} V\big(\mathbf{x}(\tau \wedge \mathbf{T}_{\mathcal{Q}_m})\big) + \mathrm{E}\Big[\int_\tau^{T\wedge\mathbf{T}_{\mathcal{Q}_m}} (LV)\big(\mathbf{x}(t)\big)dt \mid \mathcal{F}_\tau\Big]\\ &\overset{\text{wpo}}{=} V\big(\mathbf{x}(\tau \wedge \mathbf{T}_{\mathcal{Q}_m})\big) + c\,\mathrm{E}\Big[\int_\tau^{T\wedge\mathbf{T}_{\mathcal{Q}_m}} V\big(\mathbf{x}(t)\big)dt \mid \mathcal{F}_\tau\Big]\\ &\quad - \mathrm{E}\Big[\int_\tau^{T\wedge\mathbf{T}_{\mathcal{Q}_m}} \Big(cV\big(\mathbf{x}(t)\big) - (LV)\big(\mathbf{x}(t)\big)\Big)dt \mid \mathcal{F}_\tau\Big] \quad \forall m > 0.\end{aligned} \tag{7.17}$$

Since (i) the maps

$$\begin{aligned}
&m \mapsto V\big(\mathbf{x}(T \wedge \mathbf{T}_{\mathcal{Q}_m})\big),\\
&m \mapsto \int_\tau^{T \wedge \mathbf{T}_{\mathcal{Q}_m}} V\big(\mathbf{x}(t)\big)dt,\\
&m \mapsto \int_\tau^{T \wedge \mathbf{T}_{\mathcal{Q}_m}} \Big(cV\big(\mathbf{x}(t)\big) - (LV)\big(\mathbf{x}(t)\big)\Big)dt
\end{aligned}$$

are non-negative and monotone non-decreasing with probability one and (ii)

$$\begin{aligned}
&\mathrm{P}\left(\lim_{m\to\infty} V\big(\mathbf{x}(T \wedge \mathbf{T}_{\mathcal{Q}_m})\big) = V\big(\mathbf{x}(T)\big) \mid \mathcal{F}_\tau\right)\\
&\qquad\qquad \geqslant \mathrm{P}(\exists m : \mathbf{T}_{\mathcal{Q}_m}(\omega) \geqslant T \mid \mathcal{F}_\tau) \geqslant \mathrm{P}(\mathbf{T}_{\max} > T \mid \mathcal{F}_\tau) = 1,\\
&\mathrm{P}\left(\lim_{m\to\infty} \int_\tau^{T \wedge \mathbf{T}_{\mathcal{Q}_m}} V\big(\mathbf{x}(t)\big)dt = \int_\tau^T V\big(\mathbf{x}(t)\big)dt \mid \mathcal{F}_\tau\right)\\
&\qquad\qquad \geqslant \mathrm{P}(\exists m : \mathbf{T}_{\mathcal{Q}_m}(\omega) \geqslant T \mid \mathcal{F}_\tau) \geqslant \mathrm{P}(\mathbf{T}_{\max} > T \mid \mathcal{F}_\tau) = 1,\\
&\mathrm{P}\left(\lim_{m\to\infty} \int_\tau^{T \wedge \mathbf{T}_{\mathcal{Q}_m}} \Big(cV\big(\mathbf{x}(t)\big) - (LV)\big(\mathbf{x}(t)\big)\Big)dt = \int_\tau^T \Big(cV\big(\mathbf{x}(t)\big) - (LV)\big(\mathbf{x}(t)\big)\Big)dt \mid \mathcal{F}_\tau\right)\\
&\qquad\qquad \geqslant \mathrm{P}(\exists m : \mathbf{T}_{\mathcal{Q}_m}(\omega) \geqslant T \mid \mathcal{F}_\tau) \geqslant \mathrm{P}(\mathbf{T}_{\max} > T \mid \mathcal{F}_\tau) = 1,
\end{aligned}$$

we conclude from the Monotone Convergence Theorem that

$$\begin{aligned}
\lim_{m\to\infty} \mathrm{E}\Big[V\big(\mathbf{x}(T \wedge \mathbf{T}_{\mathcal{Q}_m})\big) \mid \mathcal{F}_\tau\Big] &= \mathrm{E}\Big[\lim_{m\to\infty} V\big(\mathbf{x}(T \wedge \mathbf{T}_{\mathcal{Q}_m})\big) \mid \mathcal{F}_\tau\Big]\\
&= \mathrm{E}\Big[V\big(\mathbf{x}(T)\big) \mid \mathcal{F}_\tau\Big]\\
\lim_{m\to\infty} \mathrm{E}\Big[\int_\tau^{T \wedge \mathbf{T}_{\mathcal{Q}_m}} V\big(\mathbf{x}(t)\big)dt \mid \mathcal{F}_\tau\Big] &= \mathrm{E}\Big[\lim_{m\to\infty} \int_\tau^{T \wedge \mathbf{T}_{\mathcal{Q}_m}} V\big(\mathbf{x}(t)\big)dt \mid \mathcal{F}_\tau\Big]\\
&= \mathrm{E}\Big[\int_\tau^T V\big(\mathbf{x}(t)\big)dt \mid \mathcal{F}_\tau\Big]\\
\lim_{m\to\infty} \mathrm{E}\Big[\int_\tau^{T \wedge \mathbf{T}_{\mathcal{Q}_m}} \Big(cV\big(\mathbf{x}(t)\big) - (LV)\big(\mathbf{x}(t)\big)\Big)dt \mid \mathcal{F}_\tau\Big] &= \cdots\\
&= \mathrm{E}\Big[\int_\tau^T \Big(cV\big(\mathbf{x}(t)\big) - (LV)\big(\mathbf{x}(t)\big)\Big)dt \mid \mathcal{F}_\tau\Big].
\end{aligned}$$

Moreover,

$$\begin{aligned}
\mathrm{P}\left(\lim_{m\to\infty} V\big(\mathbf{x}(\tau \wedge \mathbf{T}_{\mathcal{Q}_m})\big) = V\big(\mathbf{x}(\tau)\big)\right) &\geqslant \mathrm{P}\left(\exists m : \mathbf{T}_{\mathcal{Q}_m}(\omega) \geqslant T \mid \mathcal{F}_\tau\right)\\
&\geqslant \mathrm{P}\left(\mathbf{T}_{\max} > T \mid \mathcal{F}_\tau\right) = 1.
\end{aligned}$$

By considering the limit $m \to \infty$ in (7.17), and using the equalities above, we conclude that

$$\begin{aligned}
\mathrm{E}\big[V\big(\mathbf{x}(T)\big) \mid \mathcal{F}_\tau\big] &\overset{\text{wpo}}{=} V\big(\mathbf{x}(\tau)\big) + \mathrm{E}\Big[\int_\tau^T (LV)\big(\mathbf{x}(t)\big)dt \mid \mathcal{F}_\tau\Big]\\
&\overset{\text{wpo}}{=} V\big(\mathbf{x}(\tau)\big) + \int_\tau^T \mathrm{E}\big[(LV)\big(\mathbf{x}(t)\big) \mid \mathcal{F}_\tau\big]dt,
\end{aligned} \tag{7.18}$$

where the last equality was obtained using Fubini's Theorem to interchange the integration with respect to time and the expected value [16, p. 53]. Because of (7.8), we also have that

$$\mathrm{E}[V\big(\mathbf{x}(t \wedge \mathbf{T}_{\mathcal{Q}_m})\big) \mid \bar{\mathcal{F}}_\tau\big] \overset{\text{wpo}}{\leqslant} e^{c(t-\tau)} V\big(\mathbf{x}(\tau \wedge \mathbf{T}_{\mathcal{Q}_m})\big), \quad \forall t \geqslant \tau \geqslant 0,$$

which, also taking the limit $m \to \infty$, leads to (7.5c).

If we now define

$$\mathbf{M}_V(t) \coloneqq e^{-ct} V\big(\mathbf{x}(t)\big), \qquad \forall t \geqslant 0,$$

we conclude from (7.5c) that this process satisfies

$$\mathrm{E}[\mathbf{M}_V(T)] < \mathrm{E}\big[V\big(\mathbf{x}(0)\big)\big] < \infty \quad \text{and} \quad \mathrm{E}[\mathbf{M}_V(T) \mid \mathcal{F}_\tau] \overset{\text{wpo}}{\leqslant} \mathbf{M}_V(\tau), \quad \forall T \geqslant \tau \geqslant 0$$

and is therefore a nonnegative supermartingale of $\{\mathcal{F}_t : t \geqslant 0\}$. We then conclude from Lemma 7.3 that for any every $\bar{m} > 0$ and every stopping time $\boldsymbol{\tau} \overset{\text{wpo}}{\in} [0, \infty)$, we have that

$$\mathrm{P}\left(\exists t \in [\boldsymbol{\tau}, \infty) : V\big(\mathbf{x}(t)\big) \geqslant e^{ct}\bar{m} \mid \mathcal{F}_{\boldsymbol{\tau}}\right) \overset{\text{wpo}}{\leqslant} \frac{e^{-c\boldsymbol{\tau}} V\big(\mathbf{x}(\boldsymbol{\tau})\big)}{\bar{m}}.$$

The inequality (7.6) then follows by setting $\bar{m} \coloneqq e^{-ct} m$. Note also that, because of Doob's stopping Theorem [14, Theorem 1.39, p. 10], the inequality (7.5c), also holds for the more general case where τ is a stopping time for $\{\mathcal{F}_t : t \geqslant 0\}$. ∎

8. Lyapunov Stability Theorems. This section contains the main results of this paper. These results rely heavily on Theorem 7.1: the marginal stability theorem is a straightforward corollary of Theorem 7.1 and, while the theorem regarding asymptotic stability is more involved, it still uses Theorem 7.1 in two key steps.

8.1. Marginal Stability. We say that $V : \mathcal{X} \to \mathbb{R}$ in $\mathcal{D}$ is a *non-increasing on-the-average Lyapunov function* for the SIS (2.1) if it is exponentially-bounded on-the-average with an expansion rate $c = 0$ in (7.3). In summary, it must satisfy P5, P6, and the following property, which is essentially P7 with $c = 0$:
P8 LV is *non-positive* in the sense that

$$(LV)(x) \coloneqq L_f V(x) + \lambda(x)\Big(\int_{\mathcal{X}} V(z)\nu_x(dz) - V(x)\Big) \leqslant -W(x) \leqslant 0, \quad \forall x \in \mathcal{X}, \tag{8.1}$$

for some non-negative function $W : \mathcal{X} \to \mathbb{R}$.
Also here, we *do not require* $V(0) = 0$ and therefore V may not be positive definite. This is important to investigate the boundedness/stability of "subcomponents" of $\mathbf{x}$ (cf. discussion in Section 4).

The terminology "non-increasing on-the-average" is motivated by the fact that the *expected value* of $V\big(\mathbf{x}(t)\big)$ decreases along solutions of the SIS (2.1) (c.f. Corollary 8.1 below). This non-increase on the average is a direct consequence of (8.1) and would hold *a-fortiori* if we demanded that

$$L_f V(x) \leqslant 0 \qquad \text{and} \qquad \int_{z \in \mathcal{X} : V(z) \leqslant V(x)} \nu_x(dz) = 1, \qquad \forall x \in \mathcal{X},$$

where the left-hand side inequality essentially requires $V\big(\mathbf{x}(t)\big)$ to decrease along the (deterministic) flows of $\dot{\mathbf{x}} = f(\mathbf{x})$ and the right-hand side inequality requires $V\big(\mathbf{x}(t)\big)$ not to increase with probability one at each jump time. However, P8 is weaker than this since for (8.1) to hold the decrease only needs occur in expected value, which does not exclude the possibility that specific sample paths may exhibit strict increase on the value of $V\big(\mathbf{x}(t)\big)$ along flows of $\dot{\mathbf{x}} = f(\mathbf{x})$ and/or strict increase on the value of $V\big(\mathbf{x}(\mathbf{t}_k)\big)$ at jump times.

The result that follows is essentially a corollary of Theorem 7.1, specialized for the case of zero expansion rate. It provides results on the *uniform* stability of the process $\mathbf{y} \coloneqq V(\mathbf{x})$ and *non-uniform* bounds on the process $\mathbf{z} \coloneqq W(\mathbf{x})$, where V and W appear in the left-hand-side and right-hand-side of (8.1), respectively.

Corollary 8.1 (Marginal Stability). *Let $V \in \mathcal{D}$ be a non-increasing on-the-average Lyapunov function for the SIS (2.1) for which $\mathrm{E}\big[V\big(\mathbf{x}(0)\big)\big] < \infty$ and let W be as in (8.1). All conclusions of Theorem 7.1 hold with $c = 0$. In particular, for every stopping time $\boldsymbol{\tau} \overset{\text{wpo}}{\in} [0, \infty)$ for the natural filtration $\{\mathcal{F}_t : t \geqslant 0\}$ of the process $\mathbf{x}(t)$, $t \geqslant 0$, the following statements hold:*

1. *The process* $\mathbf{y} \coloneqq V(\mathbf{x})$ *is strongly uniformly stable in probability (strong version of definition D1). Specifically,*

$$\mathrm{P}\left(\exists t \in [\boldsymbol{\tau}, \infty) : V\big(\mathbf{x}(t)\big) > \alpha_\epsilon\big(V\big(\mathbf{x}(\boldsymbol{\tau})\big)\big) \mid \mathcal{F}_{\boldsymbol{\tau}}\right) \overset{\text{wpo}}{\leqslant} \epsilon, \quad \forall \epsilon > 0, \tag{8.2}$$

for the function $\alpha_\epsilon \in \mathcal{K}$ *defined by* $\alpha_\epsilon(s) \coloneqq \frac{s}{\epsilon}$, $\forall s \geqslant 0$.

2. *The process* $y \coloneqq V(\mathbf{x})$ *is strongly uniformly mean stable (strong version of definition D4). Specifically,*

$$\mathrm{E}\Big[V\big(\mathbf{x}(T)\big) \mid \mathcal{F}_{\boldsymbol{\tau}}\Big] \overset{\text{wpo}}{\leqslant} V\big(\mathbf{x}(\boldsymbol{\tau})\big), \qquad \forall T \geqslant \boldsymbol{\tau}. \tag{8.3}$$

(The inequalities (8.2)–(8.3) also holds for the special case where $\boldsymbol{\tau}$ *is a non-negative deterministic time.) Moreover, when the function* W *in (8.1) belongs to* $\mathcal{D}$ *and there exists constants* $a, b > 0$ *such that*

$$(LW)(x) \coloneqq L_f W(x) + \lambda(x)\Big(\int_{\mathcal{X}} W(z)\nu_x(dz) - W(x)\Big) \geqslant -aV(x) - b, \quad \forall x \in \mathcal{X}, \tag{8.4}$$

we further conclude that

3. *The expected value of the process* $\mathbf{z} \coloneqq W(\mathbf{x})$ *is integrable on any interval* $[\tau, \infty)$, $\tau \geqslant 0$, *and*

$$\int_{\tau}^{\infty} \mathrm{E}\Big[W\big(\mathbf{x}(t)\big) \mid \mathcal{F}_{\tau}\Big] \overset{\text{wpo}}{\leqslant} V\big(\mathbf{x}(\tau)\big), \quad \forall \tau \geqslant 0. \tag{8.5}$$

4. *The expected value of the process* $\mathbf{z} \coloneqq W(\mathbf{x})$ *is upper bounded by*

$$\mathrm{E}\Big[W\big(\mathbf{x}(t)\big) \mid \mathcal{F}_{\tau}\Big] \leqslant \Big(\frac{1}{2\epsilon} + a\,\epsilon\Big) V\big(\mathbf{x}(\tau)\big) + b\,\epsilon, \quad \forall t \geqslant \tau \geqslant 0,\ \epsilon > 0. \tag{8.6}$$

5. *The expected value of the process* $\mathbf{z} \coloneqq W(\mathbf{x})$ *converges to zero, i.e.,*

$$\lim_{t\to\infty} \mathrm{E}\big[W\big(\mathbf{x}(t)\big) \mid \mathcal{F}_{\tau}\big] \overset{\text{wpo}}{=} 0, \quad \forall \tau \geqslant 0. \tag{8.7}$$

6. *The process* $\mathbf{z} \coloneqq W(\mathbf{x})$ *converges to zero in probability, i.e.,*

$$\lim_{t\to\infty} \mathrm{P}\left(W\big(\mathbf{x}(t)\big) \geqslant \epsilon \mid \mathcal{F}_{\tau}\right) = 0, \quad \forall \epsilon > 0, \tau \geqslant 0. \tag{8.8}$$

By the Smoothing Property of the conditional expectation [16, p. 45], all the bounds in this theorem also hold for unconditional probabilities/expectations and for conditional probabilities/expectations with respect to the natural filtration of $V(\mathbf{x})$, *since this filtration in contained in the natural filtrations of* $\mathbf{x}$. □

While (8.5)–(8.7) provide upper bounds on the expected value of the process $\mathbf{z} \coloneqq W(\mathbf{x})$, we cannot conclude that this process satisfies any of the (uniform) stability definition considered in Section 4 because these bounds do not appear as a function of $\mathbf{z}(\tau)$. Because of this, the expected value of $\mathbf{z}(t)$ may exhibit a large transient, even when $\mathbf{z}(\tau)$ starts quite small.

Proof of Corollary 8.1. To prove (8.2) we start by considering an event $\omega \in \Omega$ for which $V\big(\mathbf{x}(\tau;\omega)\big) = 0$. In view of (7.6) with $c = 0$, for such event we have that

$$\mathrm{P}\left(\exists t \in [\boldsymbol{\tau}, \infty) : V\big(\mathbf{x}(t)\big) \geqslant m \mid \mathcal{F}_{\boldsymbol{\tau}}\right) = 0, \qquad \forall m > 0,$$

and therefore[8]

$$\begin{aligned}\mathrm{P}\left(\exists t \in [\boldsymbol{\tau}, \infty) : V\big(\mathbf{x}(t)\big) > 0 \mid \mathcal{F}_{\boldsymbol{\tau}}\right) &= \mathrm{P}\left(\exists m > 0, t \in [\boldsymbol{\tau}, \infty) : V\big(\mathbf{x}(t)\big) \geqslant m \mid \mathcal{F}_{\boldsymbol{\tau}}\right)\\ &\leqslant \lim_{m\downarrow 0^+} \mathrm{P}\left(\exists t \in [\boldsymbol{\tau}, \infty) : V\big(\mathbf{x}(t)\big) \geqslant m \mid \mathcal{F}_{\boldsymbol{\tau}}\right) = 0,\end{aligned}$$

[8]We recall that, given a sequence of measurable events $E_n \in \mathcal{F}$, with $E_n \subset E_{n+1}$, $\forall n \geqslant 0$, $\mathrm{P}(\bigcup_{n\geqslant 0} E_n) = \lim_{n\to\infty} \mathrm{P}(E_n)$.

from which (8.2) follows for events ω with $V\big(\mathbf{x}(\tau;\omega)\big)=0$. For events, $\omega\in\Omega$ for which $V\big(\mathbf{x}(\tau;\omega)\big)>0$, one can go from (7.6) with $c=0$ to (8.2) by using the following change of variables

$$\epsilon=\frac{V\big(\mathbf{x}(\boldsymbol{\tau})\big)}{m}\quad\Leftrightarrow\quad m=\frac{V\big(\mathbf{x}(\boldsymbol{\tau})\big)}{\epsilon}>0.$$

Statement 2 in Corollary 8.1 follows directly from the statement 2 in Theorem 7.1 with $c=0$ and requires no further proof.

To prove (8.5), we use (7.5a) and (8.1) to conclude that

$$\mathrm{E}\Big[V\big(\mathbf{x}(T)\big)\mid\mathcal{F}_\tau\Big]\overset{\text{wpo}}{\leqslant}V\big(\mathbf{x}(\tau)\big)-\int_\tau^T\mathrm{E}\Big[W\big(\mathbf{x}(t)\big)\mid\mathcal{F}_\tau\Big]dt\quad\forall T\geqslant\tau\geqslant 0,\tag{8.9}$$

and therefore

$$\int_\tau^T\mathrm{E}\Big[W\big(\mathbf{x}(t)\big)\mid\mathcal{F}_\tau\Big]dt\overset{\text{wpo}}{\leqslant}V\big(\mathbf{x}(\tau)\big)\quad\forall T\geqslant\tau\geqslant 0.$$

Since the random variable in the left-hand side of the above inequality is monotone increasing in T and bounded, by the Monotone Convergence Theorem the limit as $T\to\infty$ exists and satisfies (8.5).

To proceed with the proofs of (8.6)–(8.8), we pick some $m>0$ and use the compact set $\mathcal{Q}_m$ in (7.2) to construct the $\mathcal{Q}_m$-stopping time $\mathbf{T}_{\mathcal{Q}_m}$ and the $\mathcal{Q}_m$-stopped process $\mathbf{x}_{\mathcal{Q}_m}$. From (6.2b) applied to the function $W\in\mathcal{D}$ and (8.4), we conclude that

$$\begin{aligned}\mathrm{E}\big[W\big(\mathbf{x}_{\mathcal{Q}}(T)\big)\mid\mathcal{F}_\tau\big]&\overset{\text{wpo}}{=}W\big(\mathbf{x}_{\mathcal{Q}}(\tau)\big)+\int_\tau^T\mathrm{E}\big[(L_{\mathcal{Q}}W)\big(\mathbf{x}_{\mathcal{Q}}(t)\big)\mid\mathcal{F}_\tau\big]dt\\&\overset{\text{wpo}}{\geqslant}W\big(\mathbf{x}_{\mathcal{Q}}(\tau)\big)+\int_\tau^T\mathrm{E}\big[aV\big(\mathbf{x}_{\mathcal{Q}}(t)\big)+b\mid\mathcal{F}_\tau\big]dt.\end{aligned}$$

As we have done to go from (7.16) to (7.18), we can use the Monotone Convergence Theorem to take the limit as $m\to\infty$ and conclude that

$$\mathrm{E}\big[W\big(\mathbf{x}(T)\big)\mid\mathcal{F}_\tau\big]\overset{\text{wpo}}{\geqslant}W\big(\mathbf{x}(\tau)\big)+\int_\tau^T\mathrm{E}\big[aV\big(\mathbf{x}(t)\big)+b\mid\mathcal{F}_\tau\big]dt.\tag{8.10}$$

Defining

$$u(t)\coloneqq\mathrm{E}\Big[V\big(\mathbf{x}(t)\big)\mid\mathcal{F}_\tau\Big],\qquad v(t)\coloneqq\mathrm{E}\Big[W\big(\mathbf{x}(t)\big)\mid\mathcal{F}_\tau\Big],\qquad\forall t\geqslant\tau\geqslant 0,$$

we conclude from (8.9) and (8.10) that the assumptions of Lemma A.2 in the appendix hold, from which (8.6) and (8.7) follow.

Finally, from Markov's inequality [4, Theorem 1, p. 86] we further conclude that

$$\mathrm{P}\left(W\big(\mathbf{x}(t)\big)\geqslant\epsilon\mid\mathcal{F}_\tau\right)\leqslant\frac{\mathrm{E}\big[W\big(\mathbf{x}(t)\big)\mid\mathcal{F}_\tau\big]}{\epsilon},\quad\forall\epsilon>0,t\geqslant\tau.$$

Taking the limit as $t\to\infty$ and using (8.7), we conclude that (8.8) holds. ■

8.2. Asymptotic Stability. We say that $V:\mathcal{X}\to\mathbb{R}$ in $\mathcal{D}$ is a *strictly-decreasing on-the-average Lyapunov function* for the SIS (2.1) if it is a non-decreasing on-the-average Lyapunov function and, in addition, the following property holds:
P9 LV is *negative* in the sense that

$$(LV)(x)\coloneqq L_fV(x)+\lambda(x)\Big(\int_{\mathcal{X}}V(z)\nu_x(dz)-V(x)\Big)\leqslant-\alpha\big(V(x)\big)\leqslant 0,\qquad\forall x\in\mathcal{X},\tag{8.11}$$

for some class $\mathcal{K}$ function α.

Also here, the terminology "strictly decreasing on-the-average" is motivated by the fact that the *expected value* of $V\big(\mathbf{x}(t)\big)$ strictly decreases along solutions of the SIS (2.1). Again, this not exclude the possibility that specific sample paths may exhibit strict increase on the value of $V\big(\mathbf{x}(t)\big)$ along flows of $\dot{\mathbf{x}} = f(\mathbf{x})$ and/or at jump times from $\mathbf{x}^-(\mathbf{t}_k)$ to $\mathbf{x}(\mathbf{t}_k) = \mathbf{z}_k$.

The result that follows allows us to establish uniform asymptotic stability of the process $\mathbf{y} = V(\mathbf{x})$.

Theorem 8.2 (Asymptotic Stability). *Let $V \in \mathcal{D}$ be a strictly-decreasing on-the-average Lyapunov function for the SIS (2.1) for which $\mathrm{E}\big[V\big(\mathbf{x}(0)\big)\big] < \infty$. All conclusions of Theorem 7.1 (with $c = 0$) and Corollary 8.1 hold and, in addition:*

4. *The process $\mathbf{y} \coloneqq V(\mathbf{x})$ is uniformly asymptotically stable in probability (Definition D2). Specifically, for every $\epsilon > 0$, there exists a class $\mathcal{KL}$ function β_ϵ such that*

$$\mathrm{P}\left(\exists t \in [\tau, \infty) : V\big(\mathbf{x}(t)\big) > \beta_\epsilon\big(V(\mathbf{x}(\tau)), t - \tau\big) \mid \mathcal{F}_\tau\right) \overset{\text{wpo}}{\leqslant} \epsilon, \quad \forall \tau \geqslant 0. \tag{8.12}$$

5. *The process $\mathbf{y} \coloneqq V(\mathbf{x})$ is uniformly stochastically mean-α stable (Definition D5). Specifically,*

$$\int_\tau^\infty \mathrm{E}\left[\alpha\Big(V\big(\mathbf{x}(t)\big)\Big) \mid \mathcal{F}_\tau\right] dt \overset{\text{wpo}}{\leqslant} V\big(\mathbf{x}(\tau)\big), \quad \forall \tau \geqslant 0. \tag{8.13}$$

By the Smoothing Property of the conditional expectation [16, p. 45], all the bounds in this theorem also hold for unconditional probabilities/expectations and for conditional probabilities/expectations with respect to the natural filtration of $V(\mathbf{x})$, since this filtration in contained in the natural filtrations of $\mathbf{x}$. □

From the proof of Theorem 8.2 (to be presented shortly), we can conclude that the class $\mathcal{KL}$ function β_ϵ in (8.12) can be of the form

$$\beta_\epsilon(s, t) = \frac{1}{\epsilon}\beta(s, \epsilon t), \quad \forall s, t \geqslant 0,$$

for an appropriately defined class $\mathcal{KL}$ function β that *does not depend on* ϵ. This bound indicates that, in order to decrease the probability that $V\big(\mathbf{x}(t)\big)$ exceeds $\beta_\epsilon\big(V(\mathbf{x}(\tau)), t - \tau\big)$, we may need to (i) increase β_ϵ by dividing the ϵ-independent β by ϵ and (ii) slow down the decrease of $\beta_\epsilon(s, t)$ by scaling the time-dependence through a multiplication by ϵ of the time-argument. The qualification "may" in the above sentence, emphasizes the fact that the function β_ϵ used in the proof of Theorem 8.2 does not necessarily provide tight bounds so it may be possible to decrease the probability that $V\big(\mathbf{x}(t)\big)$ exceeds $\beta_\epsilon\big(V(\mathbf{x}(\tau)), t-\tau\big)$ using less conservative means. This is the case when the class $\mathcal{K}$ function α in (8.11) is of the form $\alpha(s) = \mu s$, $\forall s \geqslant 0$ for some $\mu > 0$, which leads to exponential stability and a class $\mathcal{KL}$ function β_ϵ for which the exponential convergence to zero of $t \mapsto \beta_\epsilon(s, t)$ can be made independent of ϵ, as stated in the following result, which is a direct consequence of Theorem 7.1 with an expansion rate $c = -\mu < 0$.

Corollary 8.3 (Exponential Stability). *Let $V \in \mathcal{D}$ be a strictly-decreasing on-the-average Lyapunov function for the SIS (2.1) for which $\mathrm{E}\big[V\big(\mathbf{x}(0)\big)\big] < \infty$. When the class $\mathcal{K}$ function α in (8.11) is of the form $\alpha(s) = \mu s$, $\forall s \geqslant 0$ for some $\mu > 0$, we have that*

4'. *The process $\mathbf{y} \coloneqq V(\mathbf{x})$ is uniformly exponentially stable in probability with decay rate $\mu > 0$ (Definition D3). Specifically, (8.12) holds with*

$$\beta_\epsilon(s, t) \coloneqq \frac{e^{-\mu t} s}{\epsilon}, \qquad \forall t, s \geqslant 0.$$

5'. *The process $\mathbf{y} \coloneqq V(\mathbf{x})$ is uniformly exponentially mean-α stable (Definition D7). Specifically,*

$$\mathrm{E}\big[V\big(\mathbf{x}(t)\big) \mid \mathcal{F}_\tau\big] \leqslant e^{-\mu(t-\tau)} V\big(\mathbf{x}(\tau)\big), \quad \forall t \geqslant \tau \geqslant 0.$$

□

Proof of Theorem 8.2 – Asymptotic Stability.

Proof of Theorem 8.2. To construct the class $\mathcal{KL}$ function β_ϵ that appears in (8.12), for each $s \geqslant 0$ we pick (i) a monotone increasing sequence $0 =: T_0(s) < T_1(s) < T_2(s) < \cdots$ that converges to $+\infty$ as $n \to \infty$; and (ii) a monotone decreasing sequence

$$\epsilon_0(s) > \epsilon_1(s) > \epsilon_2(s) > \cdots > 0, \ \forall s > 0, \qquad \epsilon_n(0) = 0, \ \forall n \geqslant 0$$

that converges to zero as $n \to \infty$. We then define β_ϵ to be any class $\mathcal{KL}$ function such that[9]

$$\beta_\epsilon(s,t) \geqslant \bar{\beta}_\epsilon(s,t) \coloneqq \epsilon_n(s) \quad \forall t \in \big[T_n(s), T_{n+1}(s)\big), \ n \geqslant 0. \tag{8.14}$$

For a given $\epsilon > 0$, our goal is to prove that it is indeed possible to pick the $\epsilon_n, T_n(s)$, $\forall n \geqslant 0, s \geqslant 0$ so that (8.12) holds. Defining $\mathbf{v}_\tau \coloneqq V(\mathbf{x}(\tau))$, for any event $\omega \in \Omega$ for which $\mathbf{v}_\tau(\omega) = 0$ the equation (8.12) follows directly from (8.2) no matter how we choose the class $\mathcal{KL}$ function β_ϵ. Therefore, to construct β_ϵ so that (8.12) holds, we only need to consider events $\omega \in \Omega$ for which $\mathbf{v}_\tau(\omega) > 0$.

To construct an appropriate β_ϵ, first note that, because of (8.14), we have that

$$\begin{aligned}
&\mathrm{P}\Big(\exists t \in [\tau, \infty) : V\big(\mathbf{x}(t)\big) > \beta_\epsilon\big(\mathbf{v}_\tau, t-\tau\big) \mid \mathcal{F}_\tau\Big) \\
&\qquad \overset{\text{wpo}}{\leqslant} \mathrm{P}\Big(\exists n \geqslant 0, t \in [\tau + T_n(\mathbf{v}_\tau), \infty) : V\big(\mathbf{x}(t)\big) > q\epsilon_n(\mathbf{v}_\tau) \mid \mathcal{F}_\tau\Big) \\
&\qquad\qquad \overset{\text{wpo}}{\leqslant} \sum_{n \geqslant 0} \mathrm{P}\Big(\exists t \in [\tau + T_n(\mathbf{v}_\tau), \infty) : V\big(\mathbf{x}(t)\big) > q\epsilon_n(\mathbf{v}_\tau) \mid \mathcal{F}_\tau\Big),
\end{aligned} \tag{8.15}$$

where the second inequality follows from the countable subadditivity property of the probability measure[10]. We therefore conclude that any upper bound on the sum in the right-hand side of (8.15) is also an upper bound on the probability on the left-hand side of (8.15). To construct an upper bound for the term $n = 0$ in (8.15), we can use directly (7.6) with $c = 0$ to conclude that[11]

$$\mathrm{P}\left(\exists \in [\tau, \infty) : V\big(\mathbf{x}(t)\big) > \epsilon_0(\mathbf{v}_\tau) \mid \mathcal{F}_\tau\right) \overset{\text{wpo}}{\leqslant} \frac{\mathbf{v}_\tau}{\epsilon_0(\mathbf{v}_\tau)} \tag{8.16}$$

To construct an upper bound on the terms with $n \geqslant 1$ in the right-hand side of (8.15), for each $s \geqslant 0$, we pick a third sequence $\{\delta_n(s) > 0 : n \geqslant 1\}$ and define the sequence of random variables $\{\mathbf{T}_n : n \geqslant 0\}$ with $\mathbf{T}_0 \coloneqq \tau$ and, $\forall n \geqslant 1$,

$$\begin{aligned}
\mathbf{T}_n &\coloneqq \inf\Big(\big\{t \geqslant \tau : V\big(\mathbf{x}(t)\big) \leqslant \delta_n(\mathbf{v}_\tau)\big\} \cup \{+\infty\}\Big) \\
&= \sup\Big(\big\{T \geqslant \tau : V\big(\mathbf{x}(t)\big) > \delta_n(\mathbf{v}_\tau), \ \forall t \in [\tau, T)\big\} \cup \{\tau\}\Big) \overset{\text{wpo}}{\geqslant} \tau,
\end{aligned}$$

which could be viewed as a $\mathcal{Q}$-stopping time for the (not necessarily compact) set $\mathcal{Q} \coloneqq \{x \in \mathcal{X} : V(x) \geqslant \delta_n(\mathbf{v}_\tau)\}$. We then have that the probabilities of the complements of the events appearing in the right-hand side of (8.15) satisfy, for every $n \geqslant 1$, [12]

$$\begin{aligned}
&\mathrm{P}\Big(\forall t \in [\tau + T_n(\mathbf{v}_\tau), \infty) : V\big(\mathbf{x}(t)\big) \leqslant \epsilon_n(\mathbf{v}_\tau) \mid \mathcal{F}_\tau\Big) \\
&\qquad \overset{\text{wpo}}{\geqslant} \mathrm{P}\Big(\mathbf{T}_n \leqslant \tau + T_n(\mathbf{v}_\tau) \text{ and } \forall t \in [\mathbf{T}_n, \infty) : V\big(\mathbf{x}(t)\big) \leqslant \epsilon_n(\mathbf{v}_\tau) \mid \mathcal{F}_\tau\Big),
\end{aligned}$$

[9] Since $s \mapsto \bar{\beta}_\epsilon(s,t)$ is typically not continuous for each fixed $t \geqslant 0$, we cannot just take $\beta_\epsilon = \bar{\beta}_\epsilon$. However, for each $t \geqslant 0$, we just need to take $s \mapsto \beta_\epsilon(s,t)$ to be any continuous upper bound on $s \mapsto \bar{\beta}_\epsilon(s,t)$ that is still of class $\mathcal{K}$.

[10] We recall that, given a sequence of measurable events $E_n \in \mathcal{F}$, $\forall n \geqslant 0$, the subadditivity property of P states that $\mathrm{P}(\bigcup_{n \geqslant 0} E_n) \leqslant \sum_{n \geqslant 0} \mathrm{P}(E_n)$.

[11] Recall that we only need to consider events for which $\mathbf{v}_\tau > 0$.

[12] Here, we are simply using the fact that given any two events $A, B \in \mathcal{F}$, $P(A) \geqslant P(B \cap A)$.

from which we obtain the following inequality for the probabilities of the complementary events

$$\begin{aligned}
&\mathrm{P}\Big(\exists t \in [\tau + T_n(\mathbf{v}_\tau), \infty) : V\big(\mathbf{x}(t)\big) > \epsilon_n(\mathbf{v}_\tau) \mid \mathcal{F}_\tau\Big) \\
&\qquad \overset{\text{wpo}}{\leqslant} \mathrm{P}\Big(\mathbf{T}_n > \tau + T_n(\mathbf{v}_\tau) \text{ or } \exists t \in [\mathbf{T}_n, \infty) : V\big(\mathbf{x}(t)\big) > \epsilon_n(\mathbf{v}_\tau) \mid \mathcal{F}_\tau\Big) \\
&\qquad \overset{\text{wpo}}{\leqslant} \mathrm{P}\Big(\mathbf{T}_n > \tau + T_n(\mathbf{v}_\tau) \mid \mathcal{F}_\tau\Big) + \mathrm{P}\Big(\exists t \in [\mathbf{T}_n, \infty) : V\big(\mathbf{x}(t)\big) > \epsilon_n(\mathbf{v}_\tau) \mid \mathcal{F}_\tau\Big).
\end{aligned} \tag{8.17}$$

The proposition that follows (proved at the end of this section) provides an upper-bound on the first term in (8.17):

PROPOSITION 8.4. *For every $\tau \geqslant 0$,*

$$\mathrm{P}(\mathbf{T}_n > \tau + T_n(\mathbf{v}_\tau) \mid \mathcal{F}_\tau) \overset{\text{wpo}}{\leqslant} \frac{\mathbf{v}_\tau}{\alpha(\delta_n(\mathbf{v}_\tau))T_n(\mathbf{v}_\tau)}. \tag{8.18}$$

□

Since the right-hand side of (8.18) goes to zero as $T_n(\mathbf{v}_\tau) \to \infty$, this also shows that $\mathbf{T}_n \overset{\text{wpo}}{<} \infty$ and therefore we can use (7.6) with $\boldsymbol{\tau} \coloneqq \mathbf{T}_n$ and $c = 0$ to conclude that

$$\mathrm{P}\big(\exists t \in [\mathbf{T}_n, \infty) : V\big(\mathbf{x}(t)\big) > \epsilon_n(\mathbf{v}_\tau) \mid \mathcal{F}_{\mathbf{T}_n}\big) \overset{\text{wpo}}{\leqslant} \frac{V\big(\mathbf{x}(\mathbf{T}_n)\big)}{\epsilon_n(\mathbf{v}_\tau)} \overset{\text{wpo}}{\leqslant} \frac{\delta_n(\mathbf{v}_\tau)}{\epsilon_n(\mathbf{v}_\tau)}.$$

Since by definition $\mathbf{T}_n \geqslant \tau$, $\forall n \geqslant 1$ we have that $\mathcal{F}_{\mathbf{T}_n} \supset \mathcal{F}_\tau$ [14, Property 1.17, p. 4] and we conclude from the Smoothing Property[13] of the conditional expectation [16, p. 45] that

$$\mathrm{P}\big(\exists t \in [\mathbf{T}_n, \infty) : V\big(\mathbf{x}(t)\big) > \epsilon_n(\mathbf{v}_\tau) \mid \mathcal{F}_\tau\big) \leqslant \frac{\delta_n(\mathbf{v}_\tau)}{\epsilon_n(\mathbf{v}_\tau)}.$$

Using this bound and (8.18) in (8.17), leads to

$$\mathrm{P}\Big(\exists t \in [\tau + T_n(\mathbf{v}_\tau), \infty) : V\big(\mathbf{x}(t)\big) > \epsilon_n(\mathbf{v}_\tau) \mid \mathcal{F}_\tau\Big) \overset{\text{wpo}}{\leqslant} \frac{\mathbf{v}_\tau}{\alpha(\delta_n(\mathbf{v}_\tau))T_n(\mathbf{v}_\tau)} + \frac{\delta_n(\mathbf{v}_\tau)}{\epsilon_n(\mathbf{v}_\tau)}.$$

We are finally ready to use (8.16) (valid for $n = 0$) and the upper bound above (valid for $n \geqslant 1$) in (8.15):

$$\mathrm{P}\Big(\exists t \in [\tau, \infty) : V\big(\mathbf{x}(t)\big) > \beta_\epsilon\big(\mathbf{v}_\tau, t - \tau\big) \mid \mathcal{F}_\tau\Big) \overset{\text{wpo}}{\leqslant} \frac{\mathbf{v}_\tau}{\epsilon_0(\mathbf{v}_\tau)} + \sum_{n=1}^{\infty}\Big(\frac{\mathbf{v}_\tau}{\alpha(\delta_n(\mathbf{v}_\tau))T_n(\mathbf{v}_\tau)} + \frac{\delta_n(\mathbf{v}_\tau)}{\epsilon_n(\mathbf{v}_\tau)}\Big). \tag{8.19}$$

It is now time to pick the s-dependent sequences $\{\epsilon_n(s) : n \geqslant 0\}$, $\{T_n(s) : n \geqslant 0\}$, and $\{\epsilon_n(s) : n \geqslant 0\}$ so that (4.3) follows from (8.19). To this effect, we note that since α is of class $\mathcal{K}$, there exists a constant c and a class $\mathcal{K}_\infty$ function[14] γ so that[15]

$$\alpha\big(\gamma(s_1)\gamma(s_2)\big) \geqslant \min\{s_1 s_2, c\}, \quad \forall s_1, s_2 \geqslant 0. \tag{8.20}$$

---

[13] Cf. footnote 7.

[14] A class $\mathcal{K}$ function α is of class $\mathcal{K}_\infty$ if $\lim_{s\to\infty} \alpha(s) = +\infty$.

[15] To verify this, first note that since α is of class $\mathcal{K}$, there always exists a positive constant $\bar{c}$ and a class $\mathcal{K}_\infty$ function $\bar{\alpha}$ such that

$$\bar{\alpha}(s) = \alpha(s), \quad \forall s \leqslant \bar{c} \qquad \text{and} \qquad \bar{\alpha}(s) \geqslant \alpha(s), \quad \forall s > \bar{c}.$$

Since $\bar{\alpha}$ is of class $\mathcal{K}_\infty$, it has an inverse function $\bar{\alpha}^{-1}$ also of class $\mathcal{K}_\infty$ and we have that

$$\alpha\big(\bar{\alpha}^{-1}(s)\big) \geqslant \begin{cases} \bar{\alpha}(\bar{\alpha}^{-1}(s)) = s & \bar{\alpha}^{-1}(s) \leqslant \bar{c} \\ \alpha(\bar{c}) & \bar{\alpha}^{-1}(s) > \bar{c} \end{cases} \quad \Leftrightarrow \quad \alpha\big(\bar{\alpha}^{-1}(s)\big) \geqslant \begin{cases} s & s \leqslant c \\ c & s > c \end{cases} = \min\{s, c\}, \quad c \coloneqq \bar{\alpha}(\bar{c}).$$

Moreover, from [21, Corollary 10], we also know that because $\bar{\alpha}^{-1}$ is of class $\mathcal{K}_\infty$, there exists another class $\mathcal{K}_\infty$ function γ such that $\bar{\alpha}^{-1}(s_1 s_2) \leqslant \gamma(s_1)\gamma(s_2)$, $\forall s_1, s_2 \geqslant 0$. Therefore $\alpha\big(\gamma(s_1)\gamma(s_2)\big) \geqslant \alpha\big(\bar{\alpha}^{-1}(s_1 s_2)\big) \geqslant \min\{s_1 s_2, c\}$, $\forall s_1, s_2 \geqslant 0$.

Suppose now that we pick two sequences[16] $\{a_n : n \geqslant 1\}$ and $\{b_n : n \geqslant 1\}$ such that

$$\begin{cases} a_n \geqslant a_{n+1}, & \forall n \geqslant 1, \\ \sum_{n=1}^{\infty} a_n = 1, \end{cases} \qquad \begin{cases} \frac{\gamma(b_n)}{a_n} \geqslant \frac{\gamma(b_{n+1})}{a_{n+1}}, & \forall n \geqslant 1, \\ \frac{\gamma(b_n)}{a_n} \to 0 \end{cases}$$

and define

$$\begin{aligned} \delta_n(s) &= \gamma(s)\gamma(b_n), \quad \forall n \geqslant 1, s > 0 \\ T_n(s) &= \frac{3}{\epsilon a_n} \max\left\{\frac{1}{b_n}, \frac{s}{c}\right\}, \quad \forall n \geqslant 1, s \geqslant 0 \\ \epsilon_n(s) &= \begin{cases} \frac{3s}{\epsilon} & n = 0, \forall s \geqslant 0 \\ \frac{3\gamma(s)}{\epsilon} \frac{\gamma(b_n)}{a_n} & \forall n \geqslant 1, s \geqslant 0. \end{cases} \end{aligned}$$

Since a_n decreases monotonically to zero, for each $s \geqslant 0$, the sequence $\{T_n(s) : n \geqslant 1\}$ increases monotonically to $+\infty$ and since $\frac{\gamma(b_n)}{a_n}$ decreases monotonically to zero, for each $s \geqslant 0$, $\epsilon_n(s)$ also decreases monotonically to zero. If we now replace these sequences in (8.19), we obtain

$$\begin{aligned} &\mathrm{P}\Big(\exists t \in [\tau, \infty) : V\big(\mathbf{x}(t)\big) > \beta_\epsilon\big(\mathbf{v}_\tau, t-\tau\big) \mid \mathcal{F}_\tau\Big) \\ &\qquad \overset{\text{wpo}}{\leqslant} \frac{\epsilon}{3} + \sum_{n=1}^{\infty}\left(\frac{\mathbf{v}_\tau}{\alpha(\gamma(\mathbf{v}_\tau)\gamma(b_n))\frac{3}{\epsilon a_n}\max\left\{\frac{1}{b_n}, \frac{\mathbf{v}_\tau}{c}\right\}} + \frac{\epsilon a_n}{3}\right) \\ &\qquad \overset{\text{wpo}}{\leqslant} \frac{\epsilon}{3} + \sum_{n=1}^{\infty}\left(\frac{\epsilon a_n}{3} + \frac{\epsilon a_n}{3}\right) = \epsilon \end{aligned}$$

where, in the last inequality, we used (8.20) to conclude that

$$\begin{aligned} \frac{\mathbf{v}_\tau}{\alpha(\gamma(\mathbf{v}_\tau)\gamma(b_n))\frac{3}{\epsilon a_n}\max\left\{\frac{1}{b_n}, \frac{\mathbf{v}_\tau}{c}\right\}} &\leqslant \frac{\epsilon a_n}{3}\frac{\mathbf{v}_\tau}{\min\{\mathbf{v}_\tau b_n, c\}\max\left\{\frac{1}{b_n}, \frac{\mathbf{v}_\tau}{c}\right\}} \\ &= \frac{\epsilon a_n}{3}\frac{1}{\min\{b_n, \frac{c}{\mathbf{v}_\tau}\}\max\left\{\frac{1}{b_n}, \frac{\mathbf{v}_\tau}{c}\right\}} = \frac{\epsilon a_n}{3}. \end{aligned}$$

To prove (8.13), we simply use (7.5a) and (8.11) to conclude that

$$\mathrm{E}\Big[V\big(\mathbf{x}(T)\big) \mid \mathcal{F}_\tau\Big] \overset{\text{wpo}}{\leqslant} V\big(\mathbf{x}(\tau)\big) - \int_\tau^T \mathrm{E}\Big[\alpha\Big(V\big(\mathbf{x}(t)\big)\Big) \mid \mathcal{F}_\tau\Big]dt, \qquad \forall T \geqslant \tau \geqslant 0$$

The inequality (8.13) then follows from this and the fact that $\mathrm{E}\Big[V\big(\mathbf{x}(T)\big) \mid \mathcal{F}_\tau\Big] \geqslant 0$. ■

Proof of Proposition 8.4. Defining

$$\mathbf{I}_{V(\mathbf{x}(t)) \geqslant \delta_n(\mathbf{v}_\tau)} \coloneqq \begin{cases} 1 & V(\mathbf{x}(t)) \geqslant \delta_n(\mathbf{v}_\tau) \\ 0 & V(\mathbf{x}(t)) < \delta_n(\mathbf{v}_\tau), \end{cases}$$

we have that

$$\begin{aligned} \mathbf{T}_n > \tau + T_n(\mathbf{v}_\tau) \quad &\Rightarrow \quad V(\mathbf{x}(t)) \geqslant \delta_n(\mathbf{v}_\tau),\ \forall t \in [\tau, \tau + T_n(\mathbf{v}_\tau)) \\ &\Rightarrow \quad \int_\tau^{\tau+T_n(\mathbf{v}_\tau)} \mathbf{I}_{V(\mathbf{x}(t)) \geqslant \delta_n(\mathbf{v}_\tau)} dt \geqslant T_n(\mathbf{v}_\tau). \end{aligned}$$

[16]For example $a_n \coloneqq 1/2^n$, $b_n \coloneqq \gamma^{-1}(1/2^{2n})$, $\forall n \geqslant 1$, for which $\frac{\gamma(b_n)}{a_n} = 1/2^n$.

From this and Markov's inequality [4, Theorem 1, p. 86], we conclude that

$$\begin{aligned}\mathrm{P}(\mathbf{T}_n > \tau + T_n(\mathbf{v}_\tau) \mid \mathcal{F}_\tau) &\overset{\text{wpo}}{\leqslant} \mathrm{P}\Big(\int_\tau^{\tau+T_n(\mathbf{v}_\tau)} \mathbf{I}_{V(\mathbf{x}(t))\geqslant\delta_n(\mathbf{v}_\tau)} dt \geqslant T_n(\mathbf{v}_\tau) \;\Big|\; \mathcal{F}_\tau\Big) \\ &\overset{\text{wpo}}{\leqslant} \frac{\mathrm{E}\Big[\int_\tau^{\tau+T_n(\mathbf{v}_\tau)} \mathbf{I}_{V(\mathbf{x}(t))\geqslant\delta_n(\mathbf{v}_\tau)} dt \;\Big|\; \mathcal{F}_\tau\Big]}{T_n(\mathbf{v}_\tau)}. \end{aligned} \tag{8.21}$$

On the other hand,

$$\begin{aligned}\mathbf{I}_{V(\mathbf{x}(t))\geqslant\delta_n(\mathbf{v}_\tau)} = 1 \quad &\Rightarrow \quad V(\mathbf{x}(t)) \geqslant \delta_n(\mathbf{v}_\tau) \\ &\Rightarrow \quad \alpha\big(V(\mathbf{x}(t))\big) \geqslant \alpha(\delta_n(\mathbf{v}_\tau)) \geqslant \alpha(\delta_n(\mathbf{v}_\tau))\mathbf{I}_{V(\mathbf{x}(t))\geqslant\delta_n(\mathbf{v}_\tau)} \\ &\Rightarrow \quad (LV)\big(\mathbf{x}(t)\big) \leqslant -\alpha\big(V(\mathbf{x}(t))\big) \leqslant -\alpha(\delta_n(\mathbf{v}_\tau))\mathbf{I}_{V(\mathbf{x}(t))\geqslant\delta_n(\mathbf{v}_\tau)}.\end{aligned}$$

Because of (8.1), the last inequality is also true when $\mathbf{I}_{V(\mathbf{x}(t))\geqslant\delta_n(\mathbf{v}_\tau)} = 0$, and therefore we can use (7.5b) to conclude that

$$\begin{aligned}\mathrm{E}\big[V\big(\mathbf{x}(\tau+T_n(\mathbf{v}_\tau))\big) - \mathbf{v}_\tau \;\big|\; \mathcal{F}_\tau\big] &\overset{\text{wpo}}{=} \mathrm{E}\Big[\int_\tau^{\tau+T_n(\mathbf{v}_\tau)} (LV)\big(\mathbf{x}(t)\big) dt \;\Big|\; \mathcal{F}_\tau\Big] \\ &\overset{\text{wpo}}{\leqslant} -\alpha(\delta_n(\mathbf{v}_\tau))\,\mathrm{E}\Big[\int_\tau^{\tau+T_n(\mathbf{v}_\tau)} \mathbf{I}_{V(\mathbf{x}(t))\geqslant\delta_n(\mathbf{v}_\tau)} dt \;\Big|\; \mathcal{F}_\tau\Big],\end{aligned}$$

which implies that

$$\begin{aligned}\mathrm{E}\Big[\int_\tau^{\tau+T_n(\mathbf{v}_\tau)} \mathbf{I}_{V(\mathbf{x}(t))\geqslant\delta_n(\mathbf{v}_\tau)} dt \;\Big|\; \mathcal{F}_\tau\Big] &\overset{\text{wpo}}{\leqslant} \frac{\mathrm{E}\big[\mathbf{v}_\tau - V\big(\mathbf{x}(\tau+T_n(\mathbf{v}_\tau))\big) \;\big|\; \mathcal{F}_\tau\big]}{\alpha(\delta_n(\mathbf{v}_\tau))} \\ &\overset{\text{wpo}}{\leqslant} \frac{\mathbf{v}_\tau}{\alpha(\delta_n(\mathbf{v}_\tau))}.\end{aligned}$$

Equation (8.18) follows from this and (8.21). ■

9. Conclusions and Future Work. The main contribution of this paper is a collection of Lyapunov-based results to prove global existence of solution and stability of stochastic hybrid systems. The results are stated in terms of the existence of Lyapunov functions that satisfy appropriate algebraic conditions that can be checked without computing solutions to the SISs. A key open issue related to this work is whether not the sufficient existence/stability conditions provided here are also necessary.

A. Appendix. To prove Lemma 5.1, we start by showing that the jumps counter $\bar{\mathbf{N}}(t)$ defined in (3.4), can be related to the standard Poisson process $\mathbf{N}(t)$, $t \geqslant 0$ used to construct the solution to the SIS, through the following intensity-dependent time scaling

$$\bar{\mathbf{N}}(t) = \mathbf{N}\Big(\int_0^t \lambda(\mathbf{x}(s))ds\Big), \qquad \forall t \in [0, \mathbf{T}_{\max}). \tag{A.1}$$

Proof of (A.1). Denoting by

$$\mathbf{T}_k \coloneqq \sum_{i=1}^k \mathbf{\Delta}_i$$

the event times of the standard Poisson process $\mathbf{N}(t)$, $t \geqslant 0$, we have that

$$\mathbf{N}\Big(\int_0^t \lambda(\mathbf{x}(s))ds\Big) = \max\Big\{k : \mathbf{T}_k \leqslant \int_0^t \lambda\big(\mathbf{x}(s)\big)ds\Big\} = \max\Big\{k : \sum_{i=1}^k \mathbf{\Delta}_i \leqslant \int_0^t \lambda\big(\mathbf{x}(s)\big)ds\Big\}.$$

Our goal is to show that this expression equals (3.4). To this effect, take an arbitrary jump time $\mathbf{t}_k$ that appears in (3.4). Since the transition intensity is non-negative, we have that

$$\mathbf{t}_k \leqslant t \quad \Rightarrow \quad \int_0^{\mathbf{t}_k} \lambda\big(\mathbf{x}(s)\big)ds \leqslant \int_0^t \lambda\big(\mathbf{x}(s)\big)ds \quad \Leftrightarrow \quad \sum_{i=1}^k \mathbf{\Delta}_i \leqslant \int_0^t \lambda\big(\mathbf{x}(s)\big)ds,$$

where we also used the fact that

$$\int_0^{\mathbf{t}_k} \lambda\big(\mathbf{x}(s)\big)ds = \sum_{i=1}^k \int_{\mathbf{t}_{i-1}}^{\mathbf{t}_i} \lambda\big(\mathbf{x}(s)\big)ds = \sum_{i=1}^k \mathbf{\Delta}_i. \tag{A.2}$$

Since we have established that

$$\big\{k : \mathbf{t}_k \leqslant t\big\} \subset \Big\{\sum_{i=1}^k \mathbf{\Delta}_i \leqslant \int_0^t \lambda\big(\mathbf{x}(s)\big)ds\Big\},$$

we conclude that

$$\bar{\mathbf{N}}(t) := \max\big\{k : \mathbf{t}_k \leqslant t\big\} \leqslant \max\Big\{k : \sum_{i=1}^k \mathbf{\Delta}_i \leqslant \int_0^t \lambda\big(\mathbf{x}(s)\big)ds\Big\} = \mathbf{N}\Big(\int_0^t \lambda(\mathbf{x}(s))ds\Big).$$

To prove that we actually have equality, assume by contradiction that

$$\max\big\{k : \mathbf{t}_k \leqslant t\big\} < \max\Big\{k : \sum_{i=1}^k \mathbf{\Delta}_i \leqslant \int_0^t \lambda\big(\mathbf{x}(s)\big)ds\Big\},$$

which means that there exists an index k^* such that $\mathbf{t}_{k^*} > t \geqslant \mathbf{t}_{k^*-1}$, but

$$\sum_{i=1}^{k^*} \mathbf{\Delta}_i = \int_0^{\mathbf{t}_{k^*}} \lambda\big(\mathbf{x}(s)\big)ds \leqslant \int_0^t \lambda\big(\mathbf{x}(s)\big)ds \quad \Leftrightarrow \quad \int_t^{\mathbf{t}_{k^*}} \lambda\big(\mathbf{x}(s)\big)ds \leqslant 0,$$

where we used (A.2). However, for $\mathbf{t}_{k^*} > t$ to be a jump time, because of (3.2) we must have

$$\int_{\mathbf{t}_{k^*-1}}^{\mathbf{t}_k^*} \lambda\big(\mathbf{x}(s)\big)dt = \int_{\mathbf{t}_{k^*-1}}^{t} \lambda\big(\mathbf{x}(s)\big)dt + \int_t^{\mathbf{t}_k^*} \lambda\big(\mathbf{x}(s)\big)dt = \mathbf{\Delta}_{k^*} \quad \Rightarrow \quad \int_{\mathbf{t}_{k^*-1}}^{t} \lambda\big(\mathbf{x}(s)\big)dt \geqslant \mathbf{\Delta}_{k^*},$$

which would mean that $\mathbf{t}_{k^*} \leqslant t$ and thus contradicts the assumption that $\mathbf{t}_{k^*} > t$. ∎

Given two scalars $a, b \in [-\infty, +\infty]$, in the proofs that follows we use the symbols $\wedge$ and $\vee$ to denote the min and max operations, as in

$$a \wedge b := \min\{a, b\}, \qquad\qquad a \vee b := \max\{a, b\}.$$

Proof of Lemma 5.1. To prove 1 and consequently that $x_{\mathcal{Q}}(t)$ is well defined on $[0, \infty)$, consider a generic outcome[17] $\omega \in \Omega_{\text{cadlag}}$ for which $\mathbf{T}_{\mathcal{Q}}(\omega) = \mathbf{T}_{\max}(\omega) < \infty$. We start by proving that the corresponding $\mathbf{x}(t;\omega)$ cannot have stopped at some time $\mathbf{T}_{\max}(\omega) = \mathbf{t}_k(\omega) = T_{\mathbf{t}_{k-1}(\omega), \mathbf{x}_{k-1}(\omega)}$ due to a finite escape from the open set $\mathcal{X}$ in step 2 in the construction of $\mathbf{x}(\omega)$. This is because, by continuity, in this case the solution should have exited the compact set $\mathcal{Q} \subset \mathcal{X}$ before $\mathbf{t}_k(\omega) = T_{\mathbf{t}_{k-1}(\omega), \mathbf{x}_{k-1}(\omega)}$, which contradicts the fact that $\mathbf{T}_{\mathcal{Q}}(\omega) = \mathbf{T}_{\max}(\omega)$.

[17] Here we will use the fact that $\mathrm{P}(A \cap \Omega_{\text{cadlag}}) = 0$ and $\mathrm{P}(\Omega_{\text{cadlag}}) = 1$ implies that $\mathrm{P}(A) = \mathrm{P}(A \cap \Omega_{\text{cadlag}}) + \mathrm{P}(A \cap \Omega \backslash \Omega_{\text{cadlag}}) \leqslant \mathrm{P}(A \cap \Omega_{\text{cadlag}}) + \mathrm{P}(\Omega \backslash \Omega_{\text{cadlag}}) = 0$.

We thus conclude that for such outcome ω, the maximal time $\mathbf{T}_{\max}(\omega)$ must be the finite accumulation point of an infinite set of jump times and therefore

$$\lim_{t\uparrow\mathbf{T}_{\max}(\omega)^-}\bar{\mathbf{N}}(t;\omega)=\infty,$$

which, because of (A.1), means that

$$\lim_{t\uparrow\mathbf{T}_{\max}(\omega)^-}\mathbf{N}\Big(\int_\tau^t\lambda(\mathbf{x}(s;\omega))ds;\omega\Big)=\infty. \tag{A.3}$$

Since we are assuming that $\mathbf{T}_{\mathcal{Q}}(\omega)=\mathbf{T}_{\max}(\omega)$, we have that $\mathbf{x}(s;\omega)\in\operatorname{Int}(\mathcal{Q})$, $\forall s\in\big[\tau,\mathbf{T}_{\max}(\omega)\big)$ and therefore

$$\lambda(\mathbf{x}(s;\omega))\leqslant\lambda_{\max}\coloneqq\sup_{x\in\mathcal{Q}}\lambda(x)<\infty,\quad \forall s\in\big[\tau,\mathbf{T}_{\max}(\omega)\big).$$

Therefore, for (A.3) to hold, we must necessarily have that

$$\infty=\lim_{t\uparrow\mathbf{T}_{\max}(\omega)^-}\mathbf{N}\Big(\int_\tau^t\lambda(\mathbf{x}(s;\omega))ds\Big)\leqslant\lim_{t\uparrow\mathbf{T}_{\max}(\omega)^-}\mathbf{N}(\lambda_{\max}(t-\tau);\omega)=\infty,$$

which means that the outcome ω requires the Poisson counter $\mathbf{N}$ to reach infinity in the finite time $\mathbf{T}_{\max}(\omega)<\infty$. Since the probability that a Poisson counter reaches infinity in finite time is equal to zero we conclude that

$$\mathrm{P}(T_{\mathcal{Q}}=\mathbf{T}_{\max}<\infty)=0.$$

This shows that $\mathbf{x}_{\mathcal{Q}}$ is well defined on $[0,\infty)$ with probability one. Moreover, since $\mathbf{x}(t)$ is cadlag with probability one, $\mathbf{x}_{\mathcal{Q}}$ is also cadlag with probability one, which also proves 3.

Finally, to prove 2 note that since there are no more jumps of $\bar{\mathbf{N}}_{\mathcal{Q}}$ after $\mathbf{T}_{\mathcal{Q}}$, we have that[18]

$$\begin{aligned}
\mathrm{E}\big[\bar{\mathbf{N}}_{\mathcal{Q}}(T)-\bar{\mathbf{N}}_{\mathcal{Q}}(\tau)\big]&=\mathrm{E}\big[\bar{\mathbf{N}}_{\mathcal{Q}}(T\wedge\mathbf{T}_{\mathcal{Q}})-\bar{\mathbf{N}}_{\mathcal{Q}}(\tau\wedge\mathbf{T}_{\mathcal{Q}})\big]\\
&=\mathrm{E}\big[\bar{\mathbf{N}}(T\wedge\mathbf{T}_{\mathcal{Q}})-\bar{\mathbf{N}}(\tau\wedge\mathbf{T}_{\mathcal{Q}})\big]=\mathrm{E}\Big[\mathbf{N}\Big(\int_0^{T\wedge\mathbf{T}_{\mathcal{Q}}}\lambda(\mathbf{x}(s))ds\Big)-\mathbf{N}\Big(\int_0^{\tau\wedge\mathbf{T}_{\mathcal{Q}}}\lambda(\mathbf{x}(s))ds\Big)\Big]\\
&=\mathrm{E}\Big[\mathbf{N}\Big(\int_0^{\tau\wedge\mathbf{T}_{\mathcal{Q}}}\lambda(\mathbf{x}(s))ds+\int_{\tau\wedge\mathbf{T}_{\mathcal{Q}}}^{T\wedge\mathbf{T}_{\mathcal{Q}}}\lambda(\mathbf{x}(s))ds\Big)-\mathbf{N}\Big(\int_0^{\tau\wedge\mathbf{T}_{\mathcal{Q}}}\lambda(\mathbf{x}(s))ds\Big)\Big]\\
&\leqslant\mathrm{E}\Big[\mathbf{N}\Big(\int_0^{\tau}\lambda(\mathbf{x}(s))ds+\lambda_{\max}(T-\tau)\Big)-\mathbf{N}\Big(\int_0^{\tau}\lambda(\mathbf{x}(s))ds\Big)\Big]=\lambda_{\max}(T-\tau).
\end{aligned}$$ ∎

While Theorem 6.2 could be proved by "patching" several results from [5], we provide a self-contained proof that only requires basic probability results and dispenses a background knowledge of Markov processes and stochastic integration. The following lemma is the key technical result needed to prove Theorem 6.2.

LEMMA A.1. *For a given compact set $\mathcal{Q}$ and times $0\leqslant\tau<T<\infty$, let $\mathbf{K}$ denote the (random) set of jump-time indices corresponding to the jumps of the stopped jump counter $\bar{\mathbf{N}}_{\mathcal{Q}}$ within the interval $(\tau,T]$, i.e., $\{\mathbf{t}_k : k\in\mathbf{K}\}$ is the set of discontinuities of $\bar{\mathbf{N}}_{\mathcal{Q}}$ in $(\tau,T]$. Given a function $h:\mathcal{X}\to\mathbb{R}$ such that, for every fixed $t_0\geqslant 0$, $x_0\in\mathcal{X}$ the map $t\mapsto h\big(\varphi(t;t_0,x_0)\big)$ is integrable on any finite interval $[t_0,T]$, $\forall T<T_{t_0,x_0}$, we have that*

$$\mathrm{E}\big[\mathbf{I}_{k\in\mathbf{K}}h\big(\mathbf{x}_{\mathcal{Q}}^-(\mathbf{t}_k)\big)\mid\mathcal{F}_\tau\big]=\mathrm{E}\Big[\int_{(\tau\vee\mathbf{t}_{k-1})\wedge T}^{T\wedge\mathbf{t}_k}(h\lambda)_{\mathcal{Q}}\big(\mathbf{x}_{\mathcal{Q}}(t)\big)dt\mid\mathcal{F}_\tau\Big], \tag{A.4}$$

where $\mathbf{I}_{k\in\mathbf{K}}$ denotes the indicator function of the event $k\in\mathbf{K}$ and

$$(h\lambda)_{\mathcal{Q}}(x)\coloneqq\begin{cases}h(x)\lambda(x) & x\in\operatorname{Int}(\mathcal{Q})\\ 0 & x\in\mathcal{X}\backslash\operatorname{Int}(\mathcal{Q})\end{cases}\qquad\forall x\in\mathcal{X}.$$ □

[18] Recall that if $\mathbf{N}$ is a standard Poisson process $\mathrm{E}\big[\mathbf{N}(b)-\mathbf{N}(a)\big]=b-a$, $\forall b\geqslant a\geqslant 0$.

Proof of Lemma A.1. To compute the expected values in (A.4), we need to determine the distribution of the jump times $\mathbf{t}_k$, $k \in \mathbf{K}$. From the construction of the process $\mathbf{x}$ and its jump times $\mathbf{t}_k$, we have that

$$\mathrm{P}(\mathbf{t}_k \leqslant t \mid \mathcal{F}_{\tau \vee \mathbf{t}_{k-1}}) = \mathrm{P}\left(\boldsymbol{\Delta}_{k-1} \leqslant \int_{\mathbf{t}_{k-1}}^{t} \lambda\big(\varphi_{s,\mathbf{t}_{k-1}}(\mathbf{x}(\mathbf{t}_{k-1}))\big) ds \mid \mathcal{F}_{\tau \vee \mathbf{t}_{k-1}}\right),$$

$\forall t \in (\tau \vee \mathbf{t}_{k-1}, T_{\mathbf{t}_{k-1},\mathbf{x}(\mathbf{t}_{k-1})})$, where the integral in the right-hand side can be decomposed as

$$\int_{\mathbf{t}_{k-1}}^{t} \lambda\big(\varphi_{s,\mathbf{t}_{k-1}}(\mathbf{x}(\mathbf{t}_{k-1}))\big) ds = \int_{\mathbf{t}_{k-1}}^{\tau \vee \mathbf{t}_{k-1}} \lambda\big(\mathbf{x}(s)\big) ds + \int_{\tau \vee \mathbf{t}_{k-1}}^{t} \lambda\Big(\varphi_{s,\tau \vee \mathbf{t}_{k-1}}\big(\mathbf{x}(\tau \vee \mathbf{t}_{k-1})\big)\Big) ds.$$

In case $\mathbf{t}_{k-1} < \tau$, the σ-algebra $\mathcal{F}_{\tau \vee \mathbf{t}_{k-1}} = \mathcal{F}_\tau$ provides some information about the random variable $\boldsymbol{\Delta}_{k-1}$. In fact, if $\mathbf{t}_{k-1} < \tau$ and $\mathbf{t}_k$ falls in $\big(\tau, T \wedge \mathbf{T}_{\mathcal{Q}}\big]$, we know that there was no jump in the interval $(\mathbf{t}_{k-1}, \tau]$ and therefore we have that

$$\int_{\mathbf{t}_{k-1}}^{\tau \vee \mathbf{t}_{k-1}} \lambda\big(\mathbf{x}(s)\big) ds < \boldsymbol{\Delta}_{k-1}. \tag{A.5}$$

However, when $\mathbf{t}_{k-1} \geqslant \tau$ the σ-algebra $\mathcal{F}_{\tau \vee \mathbf{t}_{k-1}} = \mathcal{F}_{\mathbf{t}_{k-1}}$ provides no information about $\boldsymbol{\Delta}_{k-1}$. Nevertheless, (A.5) still holds (trivially) with probability one, so we can always write

$$\begin{aligned}\mathrm{P}(\mathbf{t}_k \leqslant t \mid \mathcal{F}_{\tau \vee \mathbf{t}_{k-1}}) = \mathrm{P}_{\boldsymbol{\Delta}_{k-1}}\Bigg(\boldsymbol{\Delta}_{k-1} \leqslant \int_{\mathbf{t}_{k-1}}^{\tau \vee \mathbf{t}_{k-1}} \lambda\big(\mathbf{x}(s)\big) ds + \\ + \int_{\tau \vee \mathbf{t}_{k-1}}^{t} \lambda\Big(\varphi_{s,\tau \vee \mathbf{t}_{k-1}}\big(\mathbf{x}(\tau \vee \mathbf{t}_{k-1})\big)\Big) ds \;\Big|\; \int_{\mathbf{t}_{k-1}}^{\tau \vee \mathbf{t}_{k-1}} \lambda\big(\mathbf{x}(s)\big) ds < \boldsymbol{\Delta}_{k-1}\Bigg),\end{aligned}$$

where $\mathrm{P}_{\boldsymbol{\Delta}_{k-1}}(\cdot)$ denotes the measure of a standard exponential random variable. Therefore[19]

$$\mathrm{P}(\mathbf{t}_k \leqslant t \mid \mathcal{F}_{\tau \vee \mathbf{t}_{k-1}}) = 1 - e^{-\int_{\tau \vee \mathbf{t}_{k-1}}^{t} \lambda(\varphi_{s,\tau \vee \mathbf{t}_{k-1}}(\mathbf{x}(\tau \vee \mathbf{t}_{k-1})))ds},$$

$\forall t \in (\tau \vee \mathbf{t}_{k-1}, T_{\mathbf{t}_{k-1},\mathbf{x}(\mathbf{t}_{k-1})})$, from which we obtain the following (conditional) cummulative and probability density functions for $\mathbf{t}_k$:

$$F_k(t \mid \mathcal{F}_{\tau \vee \mathbf{t}_{k-1}}) = 1 - e^{-\int_{\tau \vee \mathbf{t}_{k-1}}^{t} \lambda(\varphi_{s,\tau \vee \mathbf{t}_{k-1}}(\mathbf{x}(\tau \vee \mathbf{t}_{k-1})))ds} \tag{A.6a}$$

$$\begin{aligned} f_k(t \mid \mathcal{F}_{\tau \vee \mathbf{t}_{k-1}}) &= \frac{dF_k(t \mid \mathcal{F}_{\tau \vee \mathbf{t}_{k-1}})}{dt} = \lambda\Big(\varphi_{t,\tau \vee \mathbf{t}_{k-1}}\big(\mathbf{x}(\tau \vee \mathbf{t}_{k-1})\big)\Big)\big(1 - F_k(t \mid \mathcal{F}_{\tau \vee \mathbf{t}_{k-1}}\big), \\ &= \lambda\Big(\varphi_{t,\tau \vee \mathbf{t}_{k-1}}\big(\mathbf{x}(\tau \vee \mathbf{t}_{k-1})\big)\Big)\, \mathrm{E}[\mathbf{I}_{\mathbf{t}_k > t} \mid \mathcal{F}_{\tau \vee \mathbf{t}_{k-1}}] \end{aligned} \tag{A.6b}$$

$\forall t \in (\tau \vee \mathbf{t}_{k-1}, T_{\mathbf{t}_{k-1},\mathbf{x}(\mathbf{t}_{k-1})})$, where $\mathbf{I}_{\mathbf{t}_k > t}$ denotes the indicator function of the event $\mathbf{t}_k > t$. We are now ready to compute $\mathrm{E}\big[\mathbf{I}_{k \in \mathbf{K}} h\big(\mathbf{x}^-_{\mathcal{Q}}(\mathbf{t}_k)\big) \mid \mathcal{F}_{\tau \vee \mathbf{t}_{k-1}}\big]$. To this effect, note that for any outcome $\omega \in \mathcal{F}_{\tau \vee \mathbf{t}_{k-1}}$,

$$\begin{aligned} k \in \mathbf{K}(\omega) = 1 \quad &\Leftrightarrow \quad \mathbf{x}_{\mathcal{Q}}(\tau \vee \mathbf{t}_{k-1}) \in \mathrm{Int}(\mathcal{Q}) \\ &\qquad\qquad \text{and } \tau < \mathbf{t}_k(\omega) \leqslant T \wedge T_{\mathcal{Q}}\big(\tau \vee \mathbf{t}_{k-1}, \mathbf{x}(\tau \vee \mathbf{t}_{k-1})\big) \\ k \in \mathbf{K}(\omega) = 1 \quad &\Rightarrow \quad \mathbf{x}^-_{\mathcal{Q}}(\mathbf{t}_k) = \mathbf{x}^-(\mathbf{t}_k) = \varphi_{\mathbf{t}_k,\tau \vee \mathbf{t}_{k-1}}\big(\mathbf{x}(\tau \vee \mathbf{t}_{k-1})\big), \end{aligned}$$

where $T_{\mathcal{Q}}(t_0, x_0)$ denotes the first time that the solution $\varphi(t; t_0, x_0)$ leaves the interior of $\mathcal{Q}$, i.e.,

$$\begin{aligned} T_{\mathcal{Q}}(t_0, x_0) &\coloneqq \inf\Big(\big\{t \in [0, T_{t_0,x_0}) : \varphi(t; t_0, x_0) \notin \mathrm{Int}(\mathcal{Q})\big\} \cup \{T_{t_0,x_0}\}\Big) \\ &= \sup\Big(\big\{T \in [0, T_{t_0,x_0}) : \varphi(t; t_0, x_0) \in \mathrm{Int}(\mathcal{Q}),\ \forall t \in [\tau, T)\big\} \cup \{0\}\Big), \end{aligned}$$

[19] Recall that if $\boldsymbol{\Delta}$ is a standard exponential random variable, then for every $b \geqslant a$, we have that $\mathrm{P}_{\boldsymbol{\Delta}}(\boldsymbol{\Delta} \leqslant b \mid \boldsymbol{\Delta} > a) = \frac{\mathrm{P}_{\boldsymbol{\Delta}}(a < \boldsymbol{\Delta} \leqslant b)}{\mathrm{P}_{\boldsymbol{\Delta}}(\boldsymbol{\Delta} > a)} = \frac{F_{\boldsymbol{\Delta}}(b) - F_{\boldsymbol{\Delta}}(a)}{1 - F_{\boldsymbol{\Delta}}(a)} = \frac{e^{-a} - e^{-b}}{e^{-a}} = 1 - e^{a-b}$, where $F_{\boldsymbol{\Delta}}(x) = 1 - e^{-x}$, $x \geqslant 0$ is the cummulative distribution function of $\boldsymbol{\Delta}$.

$\forall t_0 \geqslant 0,\ x_0 \in \mathcal{X}$. Therefore[20]

$$\mathrm{E}\left[\mathbf{I}_{k\in\mathbf{K}} h\big(\mathbf{x}_{\mathcal{Q}}^{-}(\mathbf{t}_k)\big) \mid \mathcal{F}_{\tau\vee\mathbf{t}_{k-1}}\right] = \mathbf{I}_{\mathbf{x}_{\mathcal{Q}}(\tau\vee\mathbf{t}_{k-1})\in\mathrm{Int}(\mathcal{Q})} \int_{(\tau\vee\mathbf{t}_{k-1})\wedge T}^{T\wedge T_{\mathcal{Q}}\left(\tau\vee\mathbf{t}_{k-1},\mathbf{x}(\tau\vee\mathbf{t}_{k-1})\right)} h\Big(\varphi_{t,\tau\vee\mathbf{t}_{k-1}}\big(\mathbf{x}(\tau\vee\mathbf{t}_{k-1})\big)\Big) f_k(t \mid \mathcal{F}_{\tau\vee\mathbf{t}_{k-1}})dt,$$

where $\mathbf{I}_{\mathbf{x}_{\mathcal{Q}}(\tau\vee\mathbf{t}_{k-1})\in\mathrm{Int}(\mathcal{Q})}$ denotes the indicator function of the event $\mathbf{x}_{\mathcal{Q}}(\tau\vee\mathbf{t}_{k-1})\in\mathrm{Int}(\mathcal{Q})$. In the expression above, we have used the fact that the map $t \mapsto h\big(\varphi(t;t_0,x_0)\big)$ is integrable on any finite interval $[t_0,T]$, $\forall T < T_{t_0,x_0}$ (cf. P4). Using the expression for $f_k(t \mid \mathcal{F}_{\tau\vee\mathbf{t}_{k-1}})$ given in (A.6b), we then conclude that[21]

$$\begin{aligned}
&\mathrm{E}\left[\mathbf{I}_{k\in\mathbf{K}} h\big(\mathbf{x}_{\mathcal{Q}}^{-}(\mathbf{t}_k)\big) \mid \mathcal{F}_{\tau\vee\mathbf{t}_{k-1}}\right] \\
&\quad= \mathbf{I}_{\mathbf{x}_{\mathcal{Q}}(\tau\vee\mathbf{t}_{k-1})\in\mathrm{Int}(\mathcal{Q})} \int_{(\tau\vee\mathbf{t}_{k-1})\wedge T}^{T\wedge T_{\mathcal{Q}}\left(\tau\vee\mathbf{t}_{k-1},\mathbf{x}(\tau\vee\mathbf{t}_{k-1})\right)} \\
&\qquad\qquad (h\lambda)\Big(\varphi_{t,\tau\vee\mathbf{t}_{k-1}}\big(\mathbf{x}(\tau\vee\mathbf{t}_{k-1})\big)\Big)\, \mathrm{E}[\mathbf{I}_{\mathbf{t}_k>t} \mid \mathcal{F}_{\tau\vee\mathbf{t}_{k-1}}]dt \\
&\quad= \mathrm{E}\Big[\mathbf{I}_{\mathbf{x}_{\mathcal{Q}}(\tau\vee\mathbf{t}_{k-1})\in\mathrm{Int}(\mathcal{Q})} \int_{(\tau\vee\mathbf{t}_{k-1})\wedge T}^{T\wedge T_{\mathcal{Q}}\left(\tau\vee\mathbf{t}_{k-1},\mathbf{x}(\tau\vee\mathbf{t}_{k-1})\right)} \\
&\qquad\qquad (h\lambda)\Big(\varphi_{t,\tau\vee\mathbf{t}_{k-1}}\big(\mathbf{x}(\tau\vee\mathbf{t}_{k-1})\big)\Big)\mathbf{I}_{\mathbf{t}_k>t}dt \mid \mathcal{F}_{\tau\vee\mathbf{t}_{k-1}}\Big] \\
&\quad= \mathrm{E}\Big[\mathbf{I}_{\mathbf{x}_{\mathcal{Q}}(\tau\vee\mathbf{t}_{k-1})\in\mathrm{Int}(\mathcal{Q})} \int_{(\tau\vee\mathbf{t}_{k-1})\wedge T}^{T\wedge\mathbf{t}_k\wedge T_{\mathcal{Q}}\left(\tau\vee\mathbf{t}_{k-1},\mathbf{x}(\tau\vee\mathbf{t}_{k-1})\right)} \\
&\qquad\qquad (h\lambda)\Big(\varphi_{t,\tau\vee\mathbf{t}_{k-1}}\big(\mathbf{x}(\tau\vee\mathbf{t}_{k-1})\big)\Big)dt \mid \mathcal{F}_{\tau\vee\mathbf{t}_{k-1}}\Big],
\end{aligned} \tag{A.7}$$

where $(h\lambda)(x) \coloneqq h(x)\lambda(x)$, $\forall x\in\mathcal{X}$. Since, $\forall t\in\big[\tau\vee\mathbf{t}_{k-1}, \mathbf{t}_k\wedge T_{\mathcal{Q}}\big(\tau\vee\mathbf{t}_{k-1},\mathbf{x}(\tau\vee\mathbf{t}_{k-1})\big)\big]$,

$$\mathbf{x}_{\mathcal{Q}}(t) = \begin{cases} \varphi_{t,\tau\vee\mathbf{t}_{k-1}}\big(\mathbf{x}(\tau\vee\mathbf{t}_{k-1})\big) & \mathbf{x}_{\mathcal{Q}}(\tau\vee\mathbf{t}_{k-1})\in\mathrm{Int}(\mathcal{Q}) \\ \mathbf{x}_{\mathcal{Q}}(\tau\vee\mathbf{t}_{k-1}) & \mathbf{x}_{\mathcal{Q}}(\tau\vee\mathbf{t}_{k-1})\notin\mathrm{Int}(\mathcal{Q}). \end{cases}$$

we can re-write the right-hand side of (A.7) compactly as

$$\mathrm{E}\left[\mathbf{I}_{k\in\mathbf{K}} h\big(\mathbf{x}_{\mathcal{Q}}^{-}(\mathbf{t}_k)\big) \mid \mathcal{F}_{\tau\vee\mathbf{t}_{k-1}}\right] = \mathrm{E}\Big[\int_{(\tau\vee\mathbf{t}_{k-1})\wedge T}^{T\wedge\mathbf{t}_k} (h\lambda)_{\mathcal{Q}}\big(\mathbf{x}_{\mathcal{Q}}(t)\big)dt \mid \mathcal{F}_{\tau\vee\mathbf{t}_{k-1}}\Big],$$

where

$$(h\lambda)_{\mathcal{Q}}(x) \coloneqq \begin{cases} h(x)\lambda(x) & x\in\mathrm{Int}(\mathcal{Q}) \\ 0 & x\in\mathcal{X}\backslash\mathrm{Int}(\mathcal{Q}) \end{cases} \qquad \forall x\in\mathcal{X}.$$

Since $\mathcal{F}_\tau \subset \mathcal{F}_{\tau\vee\mathbf{t}_{k-1}}$, equation (A.4) then follows from the Smoothing Property of the conditional expectations [16, p. 45]. ∎

Proof of Theorem 6.2. Let us denote by $\mathbf{K}$ the (random) set of jump-time indices corresponding to the jumps of the stopped jump counter $\bar{\mathbf{N}}_{\mathcal{Q}}$ within the interval $(\tau,T]$, i.e., $\{\mathbf{t}_k : k\in\mathbf{K}\}$ is the set of discontinuities of $\bar{\mathbf{N}}_{\mathcal{Q}}$ in $(\tau,T]$. Because of Lemma 5.1, we know that the set $\mathbf{K}$ has a finite (possibly zero) number of elements with probability one.

[20] The $\wedge T$ in the lower limit of integration, makes sure that we do get 0 when $\mathbf{t}_{k-1}\geqslant T$, since in this case $(\tau\vee\mathbf{t}_{k-1})\wedge T = T\wedge T_{\mathcal{Q}}\big(\tau\vee\mathbf{t}_{k-1},\mathbf{x}(\tau\vee\mathbf{t}_{k-1})\big) = T$.

[21] The exchange between integral and expected value is justified by Fubini's Theorem and the fact, on the time interval of interest, $\mathbf{x}=\mathbf{x}_{\mathcal{Q}}$ is bounded with probability one.

Since between jump times $\mathbf{x}$ is absolutely continuous and evolves according to the vector field f and, at each jump time $\mathbf{t}_k$, $k \in \mathbf{K}$ we have an instantaneous jump of $\mathbf{x}$ from $\mathbf{x}^-(\mathbf{t}_k)$ to $\mathbf{x}(\mathbf{t}_k)$, we conclude that

$$V\big(\mathbf{x}(T \wedge \mathbf{T}_{\mathcal{Q}})\big) = V\big(\mathbf{x}(\tau \wedge \mathbf{T}_{\mathcal{Q}})\big) + \int_{\tau \wedge \mathbf{T}_{\mathcal{Q}}}^{T \wedge \mathbf{T}_{\mathcal{Q}}} L_f V\big(\mathbf{x}(t)\big) dt + \sum_{k=1}^{\infty} \mathbf{I}_{k \in \mathbf{K}} \Big(V\big(\mathbf{x}(\mathbf{t}_k)\big) - V\big(\mathbf{x}^-(\mathbf{t}_k)\big)\Big),$$

where $\mathbf{I}_{k \in \mathbf{K}}$ denotes the indicator function of the event $k \in \mathbf{K}$. Because $\mathbf{x}$ is equal to the $\mathbf{x}_{\mathcal{Q}}$ in the closed interval $\big[\tau \wedge \mathbf{T}_{\mathcal{Q}}, T \wedge \mathbf{T}_{\mathcal{Q}}\big]$ and these processes are also equal at any discontinuity $\mathbf{t}_k$ of $\bar{\mathbf{N}}_{\mathcal{Q}}$, we further conclude that

$$V\big(\mathbf{x}_{\mathcal{Q}}(T \wedge \mathbf{T}_{\mathcal{Q}})\big) = V\big(\mathbf{x}_{\mathcal{Q}}(\tau \wedge \mathbf{T}_{\mathcal{Q}})\big) + \int_{\tau \wedge \mathbf{T}_{\mathcal{Q}}}^{T \wedge \mathbf{T}_{\mathcal{Q}}} L_f V\big(\mathbf{x}_{\mathcal{Q}}(t)\big) dt + \sum_{k=1}^{\infty} \mathbf{I}_{k \in \mathbf{K}} \Big(V\big(\mathbf{x}_{\mathcal{Q}}(\mathbf{t}_k)\big) - V\big(\mathbf{x}_{\mathcal{Q}}^-(\mathbf{t}_k)\big)\Big).$$

Moreover, $\mathbf{x}_{\mathcal{Q}}$ remains constant and outside $\mathrm{Int}(Q)$ on $[\tau \wedge \mathbf{T}_{\mathcal{Q}}, \tau]$ and on $[T \wedge \mathbf{T}_{\mathcal{Q}}, T]$ so we also have that

$$V\big(\mathbf{x}_{\mathcal{Q}}(T)\big) = V\big(\mathbf{x}_{\mathcal{Q}}(\tau)\big) + \int_{\tau}^{T} L_{f,\mathcal{Q}} V\big(\mathbf{x}_{\mathcal{Q}}(t)\big) dt + \sum_{k=1}^{\infty} \mathbf{I}_{k \in \mathbf{K}} \Big(V\big(\mathbf{x}_{\mathcal{Q}}(\mathbf{t}_k)\big) - V\big(\mathbf{x}_{\mathcal{Q}}^-(\mathbf{t}_k)\big)\Big). \tag{A.8}$$

where

$$L_{f,\mathcal{Q}}(x) := \begin{cases} L_f(x) & x \in \mathrm{Int}(\mathcal{Q}) \\ 0 & x \in \mathcal{X} \setminus \mathrm{Int}(\mathcal{Q}). \end{cases}$$

To proceed and eventually arrive at (6.2), we need to take expected values of both sides of (A.8), conditioned to the filtration $\mathcal{F}_\tau$. To perform this computation, we start by taking a conditional expectation of the summation in (A.8), which results in

$$\mathrm{E}\Big[\sum_{k=1}^{\infty} \mathbf{I}_{k \in \mathbf{K}} \Big(V\big(\mathbf{x}_{\mathcal{Q}}(\mathbf{t}_k)\big) - V\big(\mathbf{x}_{\mathcal{Q}}^-(\mathbf{t}_k)\big)\Big) \mid \mathcal{F}_\tau \Big] = \mathrm{E}\Big[\sum_{k=1}^{\infty} \mathbf{I}_{k \in \mathbf{K}} h\big(\mathbf{x}_{\mathcal{Q}}^-(\mathbf{t}_k)\big) \mid \mathcal{F}_\tau \Big], \tag{A.9}$$

where

$$h(x) := \int_{\mathcal{X}} V(z) \nu_x(dz) - V(x), \quad \forall x \in \mathcal{X},$$

because, by construction, the extraction of the jump point $\mathbf{x}_k = \mathbf{x}(\mathbf{t}_k) = \mathbf{x}_{\mathcal{Q}}(\mathbf{t}_k)$ is based on the measure $\nu_{\mathbf{x}^-(\mathbf{t}_k)} = \nu_{\mathbf{x}_{\mathcal{Q}}^-(\mathbf{t}_k)}$ and this extraction is independent of $\mathcal{F}_\tau$ for $\tau < \mathbf{t}_k$, $\forall k \in \mathbf{K}$.

We proceed by using the Dominated Convergence Theorem to move the series out of the expectation in (A.9). To this effect, note that since $\mathbf{x}_{\mathcal{Q}}(t) \in \mathrm{Int}(\mathcal{Q})$, $\forall t < \mathbf{t}_k$, $k \in \mathbf{K}$ and $\mathcal{Q}$ is compact, we have that $\mathbf{x}_{\mathcal{Q}}^-(\mathbf{t}_k) \overset{\text{wpo}}{\in} \mathcal{Q}$, $k \in \mathbf{K}$ and therefore

$$|\mathbf{I}_{k \in \mathbf{K}} h\big(\mathbf{x}_{\mathcal{Q}}^-(\mathbf{t}_k)\big)| \overset{\text{wpo}}{\leqslant} h_{\mathcal{Q}} := \sup_{x \in \mathcal{Q}} \Big| \int_{\mathcal{X}} V(z) \nu_x(dz) - V(x) \Big| < \infty, \quad \forall t \geqslant 0,$$

because $V \in \mathcal{D}$ (cf. P3). Since the number of elements in $\mathbf{K}$ is upper bounded by the total number of jumps of $\bar{\mathbf{N}}_{\mathcal{Q}}(T) - \bar{\mathbf{N}}_{\mathcal{Q}}(\tau)$ in the interval $(\tau, T]$, we then conclude that the finite sum $\sum_{k=1}^{N} \mathbf{I}_{k \in \mathbf{K}} h\big(\mathbf{x}_{\mathcal{Q}}^-(\mathbf{t}_k)\big)$ is dominated by

$$\Big| \sum_{k=1}^{N} \mathbf{I}_{k \in \mathbf{K}} h\big(\mathbf{x}_{\mathcal{Q}}^-(\mathbf{t}_k)\big) \Big| \leqslant \sum_{k=1}^{\infty} \mathbf{I}_{k \in \mathbf{K}} \left| h\big(\mathbf{x}_{\mathcal{Q}}^-(\mathbf{t}_k)\big) \right| \leqslant h_{\mathcal{Q}} \big(\bar{\mathbf{N}}_{\mathcal{Q}}(T) - \bar{\mathbf{N}}_{\mathcal{Q}}(\tau) \big), \quad \forall N \geqslant 1$$

and we can bound the expected value of the random variable on the right-hand side using Lemma 5.1 as follows

$$\mathrm{E}\left[h_{\mathcal{Q}} \big(\bar{\mathbf{N}}_{\mathcal{Q}}(T) - \bar{\mathbf{N}}_{\mathcal{Q}}(\tau) \big) \right] \leqslant h_{\mathcal{Q}} \lambda_{\max} (T - \tau) < \infty.$$

We therefore conclude from (A.9), the Dominated Convergence Theorem [16, p. 46], and Lemma A.1 that[22]

$$\begin{aligned}\mathrm{E}\Big[\sum_{k=1}^{\infty}\mathbf{I}_{k\in\mathbf{K}}\Big(V\big(\mathbf{x}_{\mathcal{Q}}(\mathbf{t}_k)\big)-V\big(\mathbf{x}_{\mathcal{Q}}^{-}(\mathbf{t}_k)\big)\Big)\mid\mathcal{F}_\tau\Big]&=\mathrm{E}\Big[\sum_{k=1}^{\infty}\mathbf{I}_{k\in\mathbf{K}}h\big(\mathbf{x}_{\mathcal{Q}}^{-}(\mathbf{t}_k)\big)\mid\mathcal{F}_\tau\Big]\\&=\sum_{k=1}^{\infty}\mathrm{E}\big[\mathbf{I}_{k\in\mathbf{K}}h\big(\mathbf{x}_{\mathcal{Q}}^{-}(\mathbf{t}_k)\big)\mid\mathcal{F}_\tau\big]=\sum_{k=1}^{\infty}\mathrm{E}\Big[\int_{(\tau\vee\mathbf{t}_{k-1})\wedge T}^{T\wedge\mathbf{t}_k}(h\lambda)_{\mathcal{Q}}\big(\mathbf{x}_{\mathcal{Q}}(t)\big)dt\mid\mathcal{F}_\tau\Big].\end{aligned}\tag{A.10}$$

To apply Lemma A.1, need the map $t\mapsto h\big(\varphi(t;t_0,x_0)\big)$ to be integrable on any finite interval $[t_0,T]$, $\forall T<T_{t_0,x_0}$, for every fixed $t_0\geqslant 0$, $x_0\in\mathcal{X}$, which is guaranteed by the fact that $V\in\mathcal{D}$, because of Property P4.

We now use the Dominated Convergence Theorem again to move the series back inside the expectation in (A.10). Since $\mathbf{t}_0=0\leqslant\tau$, the finite sum

$$\sum_{k=1}^{N}\int_{(\tau\vee\mathbf{t}_{k-1})\wedge T}^{T\wedge\mathbf{t}_k}(h\lambda)_{\mathcal{Q}}\big(\mathbf{x}_{\mathcal{Q}}(t)\big)dt$$

is dominated by

$$\Big|\sum_{k=1}^{N}\int_{(\tau\vee\mathbf{t}_{k-1})\wedge T}^{T\wedge\mathbf{t}_k}(h\lambda)_{\mathcal{Q}}\big(\mathbf{x}_{\mathcal{Q}}(t)\big)dt\Big|=\Big|\int_{\tau}^{T\wedge\mathbf{t}_N}(h\lambda)_{\mathcal{Q}}\big(\mathbf{x}_{\mathcal{Q}}(t)\big)dt\Big|\leqslant h_{\mathcal{Q}}\lambda_{\max}(T-\tau),$$

$\forall N\geqslant 1$. Moreover, since $\mathbf{t}_N$ eventually becomes larger than T [23], beyond which $\mathbf{I}_{(\tau\wedge\mathbf{T}_{\mathcal{Q}},T\wedge\mathbf{T}_{\mathcal{Q}}]}(s)h\big(\varphi_{s,\tau\vee\mathbf{t}_{k-1}}(\mathbf{x}(\tau\vee\mathbf{t}_{k-1}))\big)=0$, we have that

$$\sum_{k=1}^{\infty}\int_{(\tau\vee\mathbf{t}_{k-1})\wedge T}^{T\wedge\mathbf{t}_k}(h\lambda)_{\mathcal{Q}}\big(\mathbf{x}_{\mathcal{Q}}(t)\big)dt=\lim_{N\to\infty}\int_{\tau}^{T\wedge\mathbf{t}_N}(h\lambda)_{\mathcal{Q}}\big(\mathbf{x}_{\mathcal{Q}}(t)\big)dt\overset{\text{wpo}}{=}\int_{\tau}^{T}(h\lambda)_{\mathcal{Q}}\big(\mathbf{x}_{\mathcal{Q}}(t)\big)dt.$$

We can then use the Dominated Convergence Theorem [16, p. 46] to move the series in the right-hand side of (A.10) inside the expectation and conclude that

$$\begin{aligned}&\mathrm{E}\Big[\sum_{k=1}^{\infty}\mathbf{I}_{k\in\mathbf{K}}\Big(V\big(\mathbf{x}_{\mathcal{Q}}(\mathbf{t}_k)\big)-V\big(\mathbf{x}_{\mathcal{Q}}^{-}(\mathbf{t}_k)\big)\Big)\mid\mathcal{F}_\tau\Big]\\&\qquad=\mathrm{E}\Big[\sum_{k=1}^{\infty}\int_{(\tau\vee\mathbf{t}_{k-1})\wedge T}^{T\wedge\mathbf{t}_k}(h\lambda)_{\mathcal{Q}}\big(\mathbf{x}_{\mathcal{Q}}(t)\big)dt\mid\mathcal{F}_\tau\Big]=\mathrm{E}\Big[\int_{\tau}^{T}(h\lambda)_{\mathcal{Q}}\big(\mathbf{x}_{\mathcal{Q}}(t)\big)dt\mid\mathcal{F}_\tau\Big].\end{aligned}\tag{A.11}$$

If we now take expectations of (A.8) conditioned to the σ-algebra $\mathcal{F}_\tau$ and use (A.11), we obtain

$$\mathrm{E}\big[V\big(\mathbf{x}_{\mathcal{Q}}(T)\big)\mid\mathcal{F}_\tau\big]=V\big(\mathbf{x}_{\mathcal{Q}}(\tau)\big)+\mathrm{E}\Big[\int_{\tau}^{T}L_{f,\mathcal{Q}}V\big(\mathbf{x}_{\mathcal{Q}}(t)\big)dt+\int_{\tau}^{T}(h\lambda)_{\mathcal{Q}}\big(\mathbf{x}_{\mathcal{Q}}(t)\big)dt\mid\mathcal{F}_\tau\Big],$$

from which (6.2a) follows, using the definitions of $L_{f,\mathcal{Q}}V$, $(h\lambda)_{\mathcal{Q}}$, and $L_{\mathcal{Q}}V$. The second equality (6.2b) is then obtained using Fubini's Theorem to interchange the integration with respect to time and the expected value [16, p. 53].

The fact that the equations (6.2) also hold for the natural filtration $\{\bar{\mathcal{F}}_t : t\geqslant 0\}$ of the process $\mathbf{x}_{\mathcal{Q}}(t)$, $t\geqslant 0$ follows from the fact each $\bar{\mathcal{F}}_t\subset\mathcal{F}_t$, $\forall t\geqslant\tau$ and the Smoothing Property of the conditional expectations [16, p. 45]. ■

[22] While the sum inside the expectation in the left-hand side of (A.10) is finite with probability one, the sum outside the expectation in the right-hand side of (A.10) will generally be a series and therefore we do need to use the Dominated Convergence Theorem.

[23] Recall that when there was a finite number of jumps, we "completed" the sequence $\mathbf{t}_k$ with values larger than T

Proof of Lemma 7.3. Since

$$\{\omega \in \Omega \mid \exists t \in [\boldsymbol{\tau}, \infty) : \mathbf{M}(t) \geqslant m\} = \bigcup_{T \geqslant 0} \{\omega \in \Omega \mid \exists t \in [\boldsymbol{\tau}, \boldsymbol{\tau} \vee T] : \mathbf{M}(t) \geqslant m\},$$

with the sequence of sets inside the union monotonically increasing with T (with respect to the inclusion partial order), we conclude that[24]

$$\mathrm{P}\left(\exists t \in [\boldsymbol{\tau}, \infty) : \mathbf{M}(t) \geqslant m \mid \mathcal{F}_{\boldsymbol{\tau}}\right) \overset{\text{wpo}}{=} \lim_{T \to \infty} \mathrm{P}\left(\exists t \in [\boldsymbol{\tau}, \boldsymbol{\tau} \vee T] : \mathbf{M}(t) \geqslant m \mid \mathcal{F}_{\tau}\right). \tag{A.12}$$

To establish an upper bound for the right-hand side of the above equation, pick some time $T \geqslant 0$ and define the stopping time

$$\begin{aligned}\mathbf{T}_m &\coloneqq \inf\left(\{t \in [\boldsymbol{\tau}, \boldsymbol{\tau} \vee T) : \mathbf{M}(t) \geqslant m\} \cup \{\boldsymbol{\tau} \vee T\}\right) \\ &= \sup\left(\{s \in [\boldsymbol{\tau}, \boldsymbol{\tau} \vee T) : \mathbf{M}(t) < m,\ \forall t \in [\tau, s)\} \cup \{\boldsymbol{\tau}\}\right) \overset{\text{wpo}}{\geqslant} \boldsymbol{\tau},\end{aligned}$$

where $\boldsymbol{\tau} \vee T \coloneqq \max\{\boldsymbol{\tau}, T\}$. In essence, $\mathbf{T}_m$ is the first time in $[\boldsymbol{\tau}, T]$ at which $\mathbf{M}(t)$ becomes larger than or equal to m, with the understanding that $\mathbf{T}_m = \boldsymbol{\tau} \vee T$ if $\mathbf{M}(t)$ never gets this large. In view of this definition,

$$\mathrm{P}(\exists t \in [\boldsymbol{\tau}, \boldsymbol{\tau} \vee T] : \mathbf{M}(t) \geqslant m \mid \mathcal{F}_{\boldsymbol{\tau}}) \overset{\text{wpo}}{=} \mathrm{P}\left(\mathbf{M}(\mathbf{T}_m) \geqslant m \mid \mathcal{F}_{\boldsymbol{\tau}}\right) \overset{\text{wpo}}{\leqslant} \frac{\mathrm{E}\left[\mathbf{M}(\mathbf{T}_m) \mid \mathcal{F}_{\boldsymbol{\tau}}\right]}{m},$$

where the last inequality is a consequence of Markov's inequality. Since $\boldsymbol{\tau} \overset{\text{wpo}}{\leqslant} \mathbf{T}_m \leqslant \boldsymbol{\tau} \vee T \overset{\text{wpo}}{<} \infty$ and $\mathbf{M}$ is nonnegative, we can use Doob's stopping Theorem [14, Theorem 1.39, p. 10] to conclude that $\mathrm{E}\left[\mathbf{M}(\mathbf{T}_m) \mid \mathcal{F}_{\boldsymbol{\tau}}\right] \overset{\text{wpo}}{\leqslant} \mathbf{M}(\boldsymbol{\tau})$, from which we further conclude that

$$\mathrm{P}(\exists t \in [\boldsymbol{\tau}, \boldsymbol{\tau} \vee T] : \mathbf{M}(t) \geqslant m \mid \mathcal{F}_{\boldsymbol{\tau}}) \overset{\text{wpo}}{\leqslant} \frac{\mathbf{M}(\boldsymbol{\tau})}{m}, \qquad \forall T \geqslant 0.$$

The result then follows from using this upper bound in (A.12) and using the fact that $u(t) \leqslant u(\tau)$, $\forall t \geqslant \tau$. ■

The following result is inspired by Barbalat's Lemma [15].

Lemma A.2 (Comparison). *Given two non-negative scalar signals $u, v : [0, \infty) \to [0, \infty)$ such that*

$$u(T) \leqslant u(t) - \int_t^T v(r)dr, \qquad v(T) \geqslant v(t) - \int_t^T \big(au(s) + b\big)ds, \qquad \forall T \geqslant t \geqslant 0, \tag{A.13}$$

we have that $\lim_{t \to \infty} v(t) = 0$ and , for every $\epsilon > 0$,

$$v(t) \leqslant \Big(\frac{1}{2\epsilon} + a\,\epsilon\Big)u(\tau) + b\,\epsilon, \quad \forall t \geqslant \tau \geqslant 0. \tag{A.14}$$

□

Proof of Lemma A.2. From the right-hand side inequality in (A.13), we conclude that

$$\begin{aligned}\int_t^T v(r)dr &\geqslant \int_t^T \Big(v(t) - \int_t^r \big(au(s) + b\big)\Big)dr \geqslant \int_t^T \Big(v(t) - \int_t^r \big(au(t) + b\big)\Big)dr \\ &= v(t)(T - t) - \big(au(t) + b\big)\frac{(T - t)^2}{2}, \quad \forall T \geqslant t \geqslant 0,\end{aligned} \tag{A.15}$$

[24] We are using the fact that if $E_1 \subset E_2 \subset E_3 \subset \cdots$, then $\mathrm{P}(\cup_{i=1}^{\infty} E_i) = \lim_{i \to \infty} \mathrm{P}(E_i)$.

where the second inequality is a consequence of the fact that $u(s) \leqslant u(t)$, $\forall s \geqslant t$ because of the left-hand side inequality in (A.13) and the fact that v is non-negative. Replacing this integral in the left-hand side inequality in (A.13), we conclude that

$$u(T) \leqslant u(t) - v(t)(T-t) + \big(au(t)+b\big)\frac{(T-t)^2}{2},$$

which is equivalent to

$$v(t) \leqslant \frac{u(t)-u(T)}{T-t} + \big(au(t)+b\big)\frac{T-t}{2} \leqslant \Big(\frac{1}{T-t} + a\frac{T-t}{2}\Big)u(t) + \frac{b(T-t)}{2}.$$

Inequality (A.14) follows from specializing this inequality to $\epsilon = \frac{T-t}{2} \quad \Leftrightarrow \quad T-t = 2\epsilon$. To prove that $\lim_{t\to\infty} v(t) = 0$, we assume by contradiction that there exists some $\delta > 0$ and an unbounded monotonically increasing sequence to times $\{t_n : n \geqslant 0\}$ such that

$$v(t_n) \geqslant \delta, \quad \forall n \geqslant 0. \tag{A.16}$$

We therefore conclude from (A.15) that

$$\begin{aligned}\int_{t_n}^{T} v(r)dr &\geqslant v(t_n)(T-t_n) - \big(au(t_n)+b\big)\frac{(T-t_n)^2}{2}\\ &\geqslant \delta(T-t_n) - \big(au(0)+b\big)\frac{(T-t_n)^2}{2}, \qquad\qquad \forall n \geqslant 0, T \geqslant t_n,\end{aligned}$$

where the second inequality is a consequence of (A.16) and the fact that $u(t) \leqslant u(0)$, $\forall t \geqslant 0$. In particular, for $T = t_n + \frac{\delta}{au(0)+b}$, we conclude that

$$\int_{t_n}^{t_n+\frac{\delta}{au(0)+b}} v(r)dr \geqslant \frac{1}{2}\frac{\delta^2}{au(0)+b}.$$

Without loss of generality, we shall assume that the times in the sequence t_n are separated by more than $\frac{\delta}{au(0)+b}$, in which case

$$\int_0^\infty v(r)dr \geqslant \sum_{n=0}^{\infty}\int_{t_n}^{t_n+\frac{\delta}{au(0)+b}} v(r)dr \geqslant \sum_{n=0}^{\infty}\frac{1}{2}\frac{\delta^2}{au(0)+b} = +infty.$$

On the other hand, left-hand side inequality in (A.13) implies that

$$\lim_{T\to\infty} u(T) \leqslant u(0) - \int_0^\infty v(r)dr = -\infty,$$

which contradicts the fact that u is non-negative. ∎

References.

[1] M. L. Bujorianu and J. Lygeros. Towards a general theory of stochastic hybrid systems. In H. A. P. Blom and J. Lygeros, editors, *Stochastic Hybrid Systems*, volume 337 of *Lect. Notes in Contr. and Inform. Sci.*, pages 3–30. Springer-Verlag, Berlin, 2006. (cited in p. 1)

[2] C. G. Cassandras and J. Lygeros, editors. *Stochastic Hybrid Systems: Recent Developments and Research Trends*. Number 24 in Control Engineering Series. CRC Press, Boca Raton, Nov. 2006. (cited in p. 1)

[3] D. Chatterjee and D. Liberzon. Stabilizing randomly switched systems. *SIAM J. Contr. Optimization*, 2011. To appear. (cited in p. 5)

[4] Y. S. Chow and H. Teicher. *Probability Theory: independence, interchangeability, martingales*. Springer Texts in Statistics. Springer-Verlag, New York, 3rd edition, 1997. (cited in p. 11, 17, 22)

[5] M. H. A. Davis. *Markov models and optimization*. Monographs on statistics and applied probability. Chapman & Hall, London, UK, 1993. (cited in p. 1, 2, 9, 24)

[6] G. B. Folland. *Real Analysis: Modern Techniques and Their Applications*. Pure and Applied Mathematics. John Wiley & Sons, New York, 1984. (cited in p. 10, 11)

[7] M. K. Ghosh, A. Arapostathis, and S. Marcus. Ergodic control of switching diffusions. *SIAM J. Contr. Optimization*, 35(6):1952–1988, Nov. 1997. (cited in p. 1)

[8] R. Z. Hasminskii. *Stochastic stability of differential equations*. Number 7 in Monographs and textbooks on mechanics of solid and fluids. Sijthoff and Noordhoff, Alphen aan den Rijn, Netherlands, 1980. (cited in p. 1, 5, 6)

[9] J. P. Hespanha. Polynomial stochastic hybrid systems. In M. Morari and L. Thiele, editors, *Hybrid Systems: Computation and Control*, number 3414 in Lect. Notes in Comput. Science, pages 322–338. Springer-Verlag, Berlin, Mar. 2005. (cited in p. 1)

[10] J. P. Hespanha. A model for stochastic hybrid systems with application to communication networks. *Nonlinear Analysis,* Special Issue on Hybrid Systems, 62(8):1353–1383, Sept. 2005. (cited in p. 3)

[11] J. P. Hespanha. Modeling and analysis of stochastic hybrid systems. *IEE Proc. — Control Theory & Applications,* Special Issue on Hybrid Systems, 153(5):520–535, 2007. (cited in p. 1)

[12] J. P. Hespanha, D. Liberzon, and A. R. Teel. Lyapunov conditions for input-to-state stability for impulsive systems. *Automatica*, 44(11):2735–2744, Nov. 2008. (cited in p. 8)

[13] J. Hu, J. Lygeros, and S. Sastry. Towards a theory of stochastic hybrid systems. In N. A. Lynch and B. H. Krogh, editors, *Hybrid Systems: Computation and Control*, volume 1790 of *Lect. Notes in Comput. Science*, pages 160–173. Springer, 2000. (cited in p. 1)

[14] J. Jacod and A. N. Shiriaev. *Limit Theorems for Stochastic Processes*. Springer Verlag, 2nd edition, 2003. (cited in p. 11, 15, 20, 29)

[15] H. K. Khalil. *Nonlinear Systems*. Prentice Hall, Englewood Cliffs, NJ, 3rd edition, 2002. (cited in p. 29)

[16] F. C. Klebaner. *Introduction to Stochastic Calculus with Applications*. Imperial College Press, London, 2nd edition, 2005. (cited in p. 6, 12, 13, 14, 16, 18, 20, 26, 28)

[17] M. Krstic and H. Deng. *Stabilization of Nonlinear Uncertain Systems*. Springer, 1998. (cited in p. 1, 5)

[18] H. J. Kushner. *Stochastic Stability and Control*, volume 33 of *Mathematics in Science and Engineering*. Academic Press, New York, 1967. (cited in p. 2, 5, 6)

[19] X. Mao. *Stability of stochastic differential equations with respect to semimartingales*. Number 251 in Pitman Research Notes in Mathematics. Longman Scientific & Technical, 1991. (cited in p. 1, 5, 6)

[20] S. P. Meyn and R. L. Tweedie. Stability of markovian processes iii: Foster-lyapunov criteria for continuous-time processes. *Advances in Applied Probability*, 25(3):518–548, 1993 1993. (cited in p. 2, 10)

[21] E. D. Sontag. Comments on integral variants of ISS. *Syst. & Contr. Lett.*, 34:93–100, 1998. (cited in p. 20)

[22] A. R. Teel and L. Praly. On assigning the derivative of a disturbance attenuation control Lyapunov function. *Math. Contr., Signals, Syst.*, 13:95–124, 2000. (cited in p. 8)

[23] G. Yin and C. Zhu. *Hybrid Switching Diffusions: Properties and Applications*. Number 63 in Stochastic Modelling and Applied Probability. Springer, New York, 2000. (cited in p. 1, 5)