

Article

Stochastic Thermodynamics: A Dynamical Systems Approach

Tanmay Rajpurohit [†] and Wassim M. Haddad ^{*,†}

School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0150, USA;
tanmay.rajpurohit@gatech.edu

* Correspondence: wm.haddad@aerospace.gatech.edu; Tel.: +1-404-894-0178

[†] These authors contributed equally to this work.

Received: 16 October 2017; Accepted: 13 December 2017; Published: 17 December 2017

Abstract: In this paper, we develop an energy-based, large-scale dynamical system model driven by Markov diffusion processes to present a unified framework for statistical thermodynamics predicated on a stochastic dynamical systems formalism. Specifically, using a stochastic state space formulation, we develop a nonlinear stochastic compartmental dynamical system model characterized by energy conservation laws that is consistent with statistical thermodynamic principles. In particular, we show that the difference between the average supplied system energy and the average stored system energy for our stochastic thermodynamic model is a martingale with respect to the system filtration. In addition, we show that the average stored system energy is equal to the mean energy that can be extracted from the system and the mean energy that can be delivered to the system in order to transfer it from a zero energy level to an arbitrary nonempty subset in the state space over a finite stopping time.

Keywords: statistical thermodynamics; dynamical thermodynamics; entropy; stochastic dynamical systems theory; stochastic semistability; energy equipartition; Markov processes

1. Introduction

In an attempt to generalize classical thermodynamics to irreversible nonequilibrium thermodynamics, a relatively new framework has been developed that combines stochasticity and nonequilibrium dynamics. This framework is known as stochastic thermodynamics [1–5] and goes beyond linear irreversible thermodynamics addressing transport properties and entropy production in terms of forces and fluxes via linear system response theory [6–9]. Stochastic thermodynamics is applicable to nonequilibrium systems extending the validity of the laws of thermodynamics beyond the linear response regime by providing a system thermodynamic paradigm formulated on the level of individual system state realizations that are arbitrarily far from equilibrium. The thermodynamic variables of heat, work, and entropy, along with the concomitant first and second laws of thermodynamics, are formulated on the level of individual dynamical system trajectories using stochastic differential equations.

The nonequilibrium conditions in stochastic thermodynamics are imposed by an exogenous stochastic disturbance or an initial system state that is far from the system equilibrium resulting in an open (i.e., driven) or relaxation dynamical process. More specifically, the exogenous disturbance is modeled as an independent standard Wiener process (i.e., Brownian motion) defined on a complete filtered probability space wherein the current state is only dependent on the most recent event. The stochastic system dynamics are described by an overdamped Langevin equation [2,3] in which fluctuation and dissipation forces obey the Einstein relation expressing that diffusion is a result of both thermal fluctuations and frictional dissipation [10].

Brownian motion refers to the irregular movement of microscopic particles suspended in a liquid and was discovered [11,12] by the botanist Robert Brown [13]. This random motion is explained as the result of collisions between the suspended particles (i.e., Brownian particles) and the molecules of the liquid. Einstein was the first to formulate the theory of Brownian motion by assuming that the particles suspended in the liquid contribute to the thermal fluctuations of the medium and, in accordance with the principle of equipartition of energy [14], the average translational kinetic energy of each particle [10]. Thus, Brownian motion results from collisions by molecules of the fluid, wherein the suspended particles acquire the same average kinetic energy as the molecules of the fluid. This theory suggested that all matter consists of atoms (or molecules) and heat is the energy of motion (i.e., kinetic energy) of the atoms.

The use of statistical methods in developing a general molecular theory of heat predicated on random motions of Newtonian atoms led to the connection between the dynamics of heat flow and the behavior of electromagnetic radiation. A year after Einstein published his theory on Brownian motion, Smoluchovski [15] confirmed the relation between friction and diffusion. In an attempt to simplify Einstein's theory of Brownian motion, Langevin [16] was the first to model the effect of Brownian motion using a stochastic differential equation (now known as a Langevin equation) wherein spherical particles are suspended in a medium and acted upon by external forces.

In stochastic thermodynamics, the Langevin equation captures the coupling between the system particle damping and the energy input to the particles via thermal effects. Namely, the frictional forces extract the particle kinetic energy, which in turn is injected back to the particles in the form of thermal fluctuations. This captures the phenomenological behavior of a Brownian particle suspended in a fluid medium which can be modeled as a continuous Markov process [17]. Specifically, since collisions between the fluid molecules and a Brownian particle are more inelastic at higher viscosities, and temperature decreases with increasing viscosity in a fluid, additional heat is transferred to the fluid to maintain its temperature in accordance with the equipartition theorem. This heat is transferred to the Brownian particle through an increased disturbance intensity by the fluid molecules. These collisions between the Brownian particle and fluid molecules result in the observed persistent irregular and random motion of the particles.

The balance between damping (i.e., deceleration) of the particles due to frictional effects resulting in local heating of the fluid, and consequently entropy production, and the energy injection of the particles due to thermal fluctuations resulting in local cooling of the fluid, and consequently entropy consumption, is quantified by fluctuation theorems [18–28]. Thus, even though, on average, the entropy is positive (i.e., entropy production), there exist sample paths wherein the entropy decreases, albeit with an exponentially lower probability than that of entropy production. In other words, a stochastic thermodynamic system exhibits a symmetry in the probability distribution of the entropy production in the asymptotic nonequilibrium process.

Fluctuation theorems give a precise prediction for the cases in which entropy decreases in stochastic thermodynamic systems and provide a key relation between entropy production and irreversibility. Specifically, the entropy production of individual sample path trajectories of a stochastic thermodynamic system described by a Markov process is not restricted by the second law, but rather the average entropy production is determined to be positive. Furthermore, the notions of heat and work in stochastic thermodynamic systems allow for a formulation of the first law of thermodynamics on the level of individual sample path trajectories with microscopic states (i.e., positions and velocities) governed by a stochastic Langevin equation and macroscopic states governed by a Fokker–Planck equation [29] (or a Kolmogorov forward equation, depending on context) describing the evolution of the probability density function of the microscopic (stochastic) states.

In this paper, we combine our large-scale thermodynamic system model developed in [30] with stochastic thermodynamics to develop a stochastic dynamical systems framework of thermodynamics. Specifically, we develop a large-scale dynamical system model driven by Markov diffusion processes to present a unified framework for statistical thermodynamics predicated on a stochastic dynamical

systems formalism. In particular, using a stochastic state space formulation, we develop a nonlinear stochastic compartmental dynamical system model characterized by energy conservation laws that is consistent with statistical thermodynamic principles. Moreover, we show that the difference between the average supplied system energy and the average stored system energy for our stochastic thermodynamic model is a martingale with respect to the system filtration. In addition, we show that the average stored system energy is equal to the mean energy that can be extracted from the system and the mean energy that can be delivered to the system in order to transfer it from a zero energy level to an arbitrary nonempty subset in the state space over a finite stopping time.

Finally, using the system entropy [30] as a Lyapunov function candidate, we show that in the absence of energy exchange with the environment the proposed stochastic thermodynamic model is stochastically semistable in the sense that all sample path trajectories converge almost surely to a set of equilibrium solutions, wherein every equilibrium solution in the set is almost surely Lyapunov stable. In addition, we show that the steady-state distribution of the large-scale sample path system energies is uniform, leading to system energy equipartitioning corresponding to a maximum entropy equilibrium state.

2. Stochastic Dynamical Systems

To extend the dynamical thermodynamic formulation of [30] to stochastic thermodynamics we need to establish some notation, definitions, and mathematical preliminaries. A review of some basic results on nonlinear stochastic dynamical systems is given in [31–35]. Recall that given a sample space Ω , a σ -algebra $\mathcal{F}$ on Ω is a collection of subsets of Ω such that $\emptyset \in \mathcal{F}$, if $F \in \mathcal{F}$, then $\Omega \setminus F \in \mathcal{F}$, and if $F_1, F_2, \dots \in \mathcal{F}$, then $\bigcup_{i=1}^{\infty} F_i \in \mathcal{F}$ and $\bigcap_{i=1}^{\infty} F_i \in \mathcal{F}$. The pair $(\Omega, \mathcal{F})$ is called a *measurable space* and the *probability measure* $\mathbb{P}$ defined on $(\Omega, \mathcal{F})$ is a function $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ such that $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$, and if $F_1, F_2, \dots \in \mathcal{F}$ and $F_i \cap F_j = \emptyset$, $i \neq j$, then $\mathbb{P}(\bigcup_{i=1}^{\infty} F_i) = \sum_{i=1}^{\infty} \mathbb{P}(F_i)$. The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *probability space* if $\mathcal{F}$ contains all subsets of Ω with $\mathbb{P}$ -outer measure [36] zero [33].

The subsets $\mathcal{F}$ of Ω belonging to $\mathcal{F}$ are called *$\mathcal{F}$ -measurable sets*. If $\Omega = \mathbb{R}^n$ and $\mathfrak{B}^n$ is the family of all open sets in $\mathbb{R}^n$, then $\mathfrak{B}^n$ is called the *Borel σ -algebra* and the elements $\mathcal{B}$ of $\mathfrak{B}^n$ are called *Borel sets*. If $(\Omega, \mathcal{F}, \mathbb{P})$ is a given probability space, then the real valued function (random variable) $x : \Omega \rightarrow \mathbb{R}$ is *$\mathcal{F}$ -measurable* if $\{\omega \in \Omega : x(\omega) \in \mathcal{B}\} \in \mathcal{F}$ for all Borel sets $\mathcal{B} \subset \mathbb{R}^n$. Given the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a *filtration* is a family $\{\mathcal{F}_t\}_{t \geq 0}$ of σ -algebras $\mathcal{F}_t \subset \mathcal{F}$ such that $\mathcal{F}_t \subset \mathcal{F}_s$ for all $0 \leq t < s < \infty$.

In this paper, we use the notation and terminology as established in [37]. Specifically, define a complete probability space as $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω denotes the sample space, $\mathcal{F}$ denotes a σ -algebra, and $\mathbb{P}$ defines a probability measure on the σ -algebra $\mathcal{F}$; that is, $\mathbb{P}$ is a nonnegative countably additive set function on $\mathcal{F}$ such that $\mathbb{P}(\Omega) = 1$ [32]. Furthermore, we assume that $w(\cdot)$ is a standard d -dimensional Wiener process defined by $(w(\cdot), \Omega, \mathcal{F}, \mathbb{P}^{w_0})$, where $\mathbb{P}^{w_0}$ is the classical Wiener measure ([33], p. 10), with a continuous-time filtration $\{\mathcal{F}_t\}_{t \geq 0}$ generated by the Wiener process $w(t)$ up to time t .

We denote a stochastic dynamical system by $\mathcal{G}$ generating a filtration $\{\mathcal{F}_t\}_{t \geq 0}$ adapted to the stochastic process $x : \mathbb{R}_+ \times \Omega \rightarrow \mathcal{D}$ on $(\Omega, \mathcal{F}, \mathbb{P}^{x_0})$ satisfying $\mathcal{F}_\tau \subset \mathcal{F}_t$, $0 \leq \tau < t$, such that $\{\omega \in \Omega : x(t, \omega) \in \mathcal{B}\} \in \mathcal{F}_t$, $t \geq 0$, for all Borel sets $\mathcal{B} \subset \mathbb{R}^n$ contained in the Borel σ -algebra $\mathfrak{B}^n$. We say that the stochastic process $x : \mathbb{R}_+ \times \Omega \rightarrow \mathcal{D}$ is *$\mathcal{F}_t$ -adapted* if $x(t)$ is $\mathcal{F}_t$ -measurable for every $t \geq 0$. Furthermore, we say that $\mathcal{G}$ satisfies the *Markov property* if the conditional probability distribution of the future states of the stochastic process generated by $\mathcal{G}$ only depends on the present state. In this case, $\mathcal{G}$ generates a *Markov process* which results in a decoupling of the past from the future in the sense that the present state of $\mathcal{G}$ contains sufficient information so as to encapsulate the effects of the past system inputs. Here we use the notation $x(t)$ to represent the stochastic process $x(t, \omega)$ omitting its dependence on ω . Furthermore, $\mathfrak{B}^n$ denotes the σ -algebra of Borel sets in $\mathcal{D} \subseteq \mathbb{R}^n$ and $\mathfrak{S}$ denotes a σ -algebra generated on a set $\mathcal{S} \subseteq \mathbb{R}^n$.

We denote the set of equivalence classes of measurable, integrable, and square-integrable $\mathbb{R}^n$ or $\mathbb{R}^{n \times m}$ (depending on context) valued random processes on $(\Omega, \mathcal{F}, \mathbb{P})$ over the semi-infinite parameter space $[0, \infty)$ by $\mathcal{L}^0(\Omega, \mathcal{F}, \mathbb{P})$, $\mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, and $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$, respectively, where the equivalence relation

is the one induced by $\mathbb{P}$ -almost-sure equality. In particular, elements of $\mathcal{L}^0(\Omega, \mathcal{F}, \mathbb{P})$ take finite values $\mathbb{P}$ -almost surely (a.s.) or with probability one. Hence, depending on the context, $\mathbb{R}^n$ will denote either the set of $n \times 1$ real variables or the subspace of $\mathcal{L}^0(\Omega, \mathcal{F}, \mathbb{P})$ comprising of $\mathbb{R}^n$ random processes that are constant almost surely. All inequalities and equalities involving random processes on $(\Omega, \mathcal{F}, \mathbb{P})$ are to be understood to hold $\mathbb{P}$ -almost surely. Furthermore, $\mathbb{E}[\cdot]$ and $\mathbb{E}^{x_0}[\cdot]$ denote, respectively, the expectation with respect to the probability measure $\mathbb{P}$ and with respect to the classical Wiener measure $\mathbb{P}^{x_0}$.

Given $x \in \mathcal{L}^0(\Omega, \mathcal{F}, \mathbb{P})$, $\{x = 0\}$ denotes the set $\{\omega \in \Omega : x(t, \omega) = 0\}$, and so on. Given $x \in \mathcal{L}^0(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{E} \in \mathcal{F}$, we say x is nonzero on $\mathcal{E}$ if $\mathbb{P}(\{x = 0\} \cap \mathcal{E}) = 0$. Furthermore, given $x \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ and a σ -algebra $\mathcal{E} \subseteq \mathcal{F}$, $\mathbb{E}^{\mathbb{P}}[x]$ and $\mathbb{E}^{\mathbb{P}}[x|\mathcal{E}]$ denote, respectively, the expectation of the random variable x and the conditional expectation of x given $\mathcal{E}$, with all moments taken under the measure $\mathbb{P}$. In formulations wherein it is clear from context which measure is used, we omit the symbol $\mathbb{P}$ in denoting expectation, and similarly for conditional expectation. Specifically, in such cases we denote the expectation with respect to the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ by $\mathbb{E}[\cdot]$, and similarly for conditional expectation.

A stochastic process $x : \bar{\mathbb{R}}_+ \times \Omega \rightarrow \mathcal{D}$ on $(\Omega, \mathcal{F}, \mathbb{P}^{x_0})$ is called a *martingale* with respect to the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ if and only if $x(t)$ is a $\mathcal{F}_t$ -measurable random vector for all $t \geq 0$, $\mathbb{E}[x(t)] < \infty$, and $x(\tau) = \mathbb{E}[x(t)|\mathcal{F}_\tau]$ for all $t \geq \tau \geq 0$. Thus, a martingale has the property that the expectation of the next value of the martingale is equal to its current value given all previous values of the dynamical process. If we replace the equality in $x(\tau) = \mathbb{E}[x(t)|\mathcal{F}_\tau]$ with " $\leq$ " (respectively, " $\geq$ "), then $x(\cdot)$ is a *supermartingale* (respectively, *submartingale*). Note that every martingale is both a submartingale and supermartingale.

A random variable $\tau : \Omega \rightarrow [0, \infty]$ is called a *stopping time* with respect to $\mathcal{F}_t$ if and only if $\{\omega \in \Omega : \tau(\omega) \leq t\} \in \mathcal{F}_t, t \geq 0$. Thus, the set of all $\omega \in \Omega$ such that $\tau(\omega) \leq t$ is a $\mathcal{F}_t$ -measurable set. Note that $\tau(\omega)$ can take on finite as well as infinite values and characterizes whether at each time t an event at time $\tau(\omega) < t$ has occurred using only the information in $\mathcal{F}_t$.

Finally, we write $\|\cdot\|$ for the Euclidean vector norm, $\text{row}_i(A)$ and $\text{col}_j(A)$ for the i -th row and j -th column of a matrix $A \in \mathbb{R}^{p \times q}$, $\text{tr}(\cdot)$ for the trace operator, $(\cdot)^{-1}$ for the inverse operator, $V'(x) \triangleq \frac{\partial V(x)}{\partial x}$ for the Fréchet derivative of V at x , $V''(x) \triangleq \frac{\partial^2 V(x)}{\partial x^2}$ for the Hessian of V at x , and $\mathcal{H}_n$ for the Hilbert space of random vectors $x \in \mathbb{R}^n$ with finite average power, that is, $\mathcal{H}_n \triangleq \{x : \Omega \rightarrow \mathbb{R}^n : \mathbb{E}[x^T x] < \infty\}$. For an open set $\mathcal{D} \subseteq \mathbb{R}^n$, $\mathcal{H}_n^{\mathcal{D}} \triangleq \{x \in \mathcal{H}_n : x : \Omega \rightarrow \mathcal{D}\}$ denotes the set of all the random vectors in $\mathcal{H}_n$ induced by $\mathcal{D}$. Similarly, for every $x_0 \in \mathbb{R}^n$, $\mathcal{H}_n^{x_0} \triangleq \{x \in \mathcal{H}_n : x \stackrel{\text{a.s.}}{=} x_0\}$. Moreover, $\bar{\mathbb{R}}_+^n$ and $\mathbb{R}_+^n$ denote the nonnegative and positive orthants of $\mathbb{R}^n$, that is, if $x \in \mathbb{R}^n$, then $x \in \bar{\mathbb{R}}_+^n$ and $x \in \mathbb{R}_+^n$ are equivalent, respectively, to $x \geq 0$ and $x > 0$, where $x \geq 0$ (respectively, $x > 0$) indicates that every component of x is nonnegative (respectively, positive). Furthermore, C^2 denotes the space of real-valued functions $V : \mathcal{D} \rightarrow \mathbb{R}$ that are two-times continuously differentiable with respect to $x \in \mathcal{D} \subseteq \mathbb{R}^n$. Finally, we write $x(t) \rightarrow \mathcal{M}$ as $t \rightarrow \infty$ to denote that $x(t)$ approaches the set $\mathcal{M}$, that is, for every $\varepsilon > 0$ there exists $T > 0$ such that $\text{dist}(x(t), \mathcal{M}) < \varepsilon$ for all $t > T$, where $\text{dist}(p, \mathcal{M}) \triangleq \inf_{x \in \mathcal{M}} \|p - x\|$.

Definition 1. Let $(S, \mathfrak{S})$ and $(T, \mathfrak{T})$ be measurable spaces, and let $\mu : S \times \mathfrak{T} \rightarrow \bar{\mathbb{R}}_+$. If the function $\mu(s, \mathcal{B})$ is $\mathfrak{S}$ -measurable in $s \in S$ for a fixed $\mathcal{B} \in \mathfrak{T}$ and $\mu(s, \mathcal{B})$ is a probability measure in $\mathcal{B} \in \mathfrak{T}$ for a fixed $s \in S$, then μ is called a (probability) kernel from S to T . Furthermore, for $s \leq t$, the function $\mu_{s,t} : S \times \mathfrak{S} \rightarrow \mathbb{R}$ is called a regular conditional probability measure if $\mu_{s,t}(\cdot, \mathfrak{S})$ is measurable, $\mu_{s,t}(S, \cdot)$ is a probability measure, and $\mu_{s,t}(\cdot, \cdot)$ satisfies

$$\mu_{s,t}(x(s), \mathcal{B}) = \mathbb{P}(x(t) \in \mathcal{B} | x(s)) = \mathbb{P}(x(t) \in \mathcal{B} | \mathcal{F}_s), \quad x(\cdot) \in \mathcal{H}_n, \quad (1)$$

where $\mathbb{P}(x(t) \in \mathcal{B} | x(s)) = \mathbb{P}(0, x, t, \mathcal{B})$, $x \in \mathbb{R}^n$, and $\mathbb{P}(s, x, t, \mathcal{B})$, $t \geq s$, is the transition probability of the point $x \in \mathbb{R}^n$ at time instant s into all Borel subsets $\mathcal{B} \subset \mathbb{R}^n$ at time instant t .

Any family of regular conditional probability measures $\{\mu_{s,t}\}_{s \leq t}$ satisfying the Chapman–Kolmogorov equation ([32])

$$\mathbb{P}(s, x, t, \mathcal{B}) = \int_{\mathbb{R}^n} \mathbb{P}(s, x, \sigma, dz) \mathbb{P}(s, z, t, \mathcal{B}), \quad (2)$$

or, equivalently,

$$\mu_{s,t}(x, \mathcal{B}) = \int_{\mathbb{R}^n} \mu_{s,\sigma}(x, dz) \mu_{\sigma,t}(z, \mathcal{B}), \quad (3)$$

where $0 \leq s \leq \sigma \leq t < \infty$, $x, z \in \mathbb{R}^n$, and $\mathcal{B} \in \mathfrak{B}^n$, is called a *semigroup of Markov kernels*. The Markov kernels are called *time homogeneous* if and only if $\mu_{s,t} = \mu_{0,t-s}$ holds for all $s \leq t$.

Consider the nonlinear stochastic dynamical system $\mathcal{G}$ given by

$$dx(t) = f(x(t))dt + D(x(t))dw(t), \quad x(0) \stackrel{\text{a.s.}}{=} x_0, \quad t \in \mathcal{I}_{x(0)}, \quad (4)$$

where, for every $t \in \mathcal{I}_{x_0}$, $x(t) \in \mathcal{H}_n^{\mathcal{D}}$ is a $\mathcal{F}_t$ -measurable random state vector, $x(0) \in \mathcal{H}_n^{x_0}$, $\mathcal{D} \subseteq \mathbb{R}^n$ is relatively open set with $0 \in \mathcal{D}$, $w(\cdot)$ is a d -dimensional independent standard Wiener process (i.e., Brownian motion) defined on a complete filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$, $x(0)$ is independent of $(w(t) - w(0)), t \geq 0$, $f : \mathcal{D} \rightarrow \mathbb{R}^n$ and $D : \mathcal{D} \rightarrow \mathbb{R}^{n \times d}$ are continuous, $\mathcal{E} \triangleq f^{-1}(0) \cap D^{-1}(0) \triangleq \{x \in \mathcal{D} : f(x) = 0 \text{ and } D(x) = 0\}$ is nonempty, and $\mathcal{I}_{x(0)} = [0, \tau_{x(0)})$, $0 \leq \tau_{x(0)} \leq \infty$, is the maximal interval of existence for the solution $x(\cdot)$ of (4).

An *equilibrium point* of (4) is a point $x_e \in \mathbb{R}^n$ such that $f(x_e) = 0$ and $D(x_e) = 0$. It is easy to see that x_e is an equilibrium point of (4) if and only if the constant stochastic process $x(\cdot) \stackrel{\text{a.s.}}{=} x_e$ is a solution of (4). We denote the set of equilibrium points of (4) by $\mathcal{E} \triangleq \{\omega \in \Omega : x(t, \omega) = x_e\} = \{x_e \in \mathcal{D} : f(x_e) = 0 \text{ and } D(x_e) = 0\}$.

The filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ is clearly a real vector space with addition and scalar multiplication defined componentwise and pointwise. A $\mathbb{R}^n$ -valued stochastic process $x : [0, \tau] \times \Omega \rightarrow \mathcal{D}$ is said to be a *solution* of (4) on the time interval $[0, \tau]$ with initial condition $x(0) \stackrel{\text{a.s.}}{=} x_0$ if $x(\cdot)$ is *progressively measurable* (i.e., $x(\cdot)$ is nonanticipating and measurable in t and ω) with respect to the filtration $\{\mathcal{F}_t\}_{t \geq 0}$, $f \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, $D \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$, and

$$x(t) = x_0 + \int_0^t f(x(\sigma))d\sigma + \int_0^t D(x(\sigma))dw(\sigma) \text{ a.s., } t \in [0, \tau], \quad (5)$$

where the integrals in (5) are Itô integrals [38]. If the map $t \rightarrow w(t, \omega)$, $\omega \in \Omega$, had a bounded variation, then the natural definition for the integrals in (5) would be the Lebesgue–Stieltjes integral where ω is viewed as a parameter. However, since sample Wiener paths are nowhere differentiable and not of bounded variation for almost all $\omega \in \Omega$ the integrals in (5) need to be defined as Itô integrals [39,40].

Note that for each fixed $t \geq 0$, the random variable $\omega \mapsto x(t, \omega)$ assigns a vector $x(\omega)$ to every outcome $\omega \in \Omega$ of an experiment, and for each fixed $\omega \in \Omega$, the mapping $t \mapsto x(t, \omega)$ is the *sample path* of the stochastic process $x(t)$, $t \geq 0$. A path-wise solution $t \mapsto x(t)$ of (4) in $(\Omega, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P}^{x_0})$ is said to be *right maximally* defined if x cannot be extended (either uniquely or non-uniquely) forward in time. We assume that all right maximal path-wise solutions to (4) in $(\Omega, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P}^{x_0})$ exist on $[0, \infty)$, and hence, we assume that (4) is *forward complete*. Sufficient conditions for forward completeness or *global solutions* of (4) are given in [34,38].

Furthermore, we assume that $f : \mathcal{D} \rightarrow \mathbb{R}^n$ and $D : \mathcal{D} \rightarrow \mathbb{R}^{n \times d}$ satisfy the uniform Lipschitz continuity condition

$$\|f(x) - f(y)\| + \|D(x) - D(y)\|_F \leq L\|x - y\|, \quad x, y \in \mathcal{D} \setminus \{0\}, \quad (6)$$

and the growth restriction condition

$$\|f(x)\|^2 + \|D(x)\|_F^2 \leq L^2(1 + \|x\|^2), \quad x \in \mathcal{D} \setminus \{0\}, \quad (7)$$

for some Lipschitz constant $L > 0$, and hence, since $x(0) \in \mathcal{H}_n^D$ and $x(0)$ is independent of $(w(t) - w(0)), t \geq 0$, it follows that there exists a unique solution $x \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ of (4) forward in time for all initial conditions in the following sense. For every $x \in \mathcal{H}_n^D \setminus \{0\}$ there exists $\tau_x > 0$ such that if $x_1 : [0, \tau_1] \times \Omega \rightarrow \mathcal{D}$ and $x_2 : [0, \tau_2] \times \Omega \rightarrow \mathcal{D}$ are two solutions of (4); that is, if $x_1, x_2 \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ with continuous sample paths almost surely solve (4), then $\tau_x \leq \min\{\tau_1, \tau_2\}$ and $\mathbb{P}(x_1(t) = x_2(t), 0 \leq t \leq \tau_x) = 1$.

The uniform Lipschitz continuity condition (6) guarantees uniqueness of solutions, whereas the linear growth condition (7) rules out finite escape times. A weaker sufficient condition for the existence of a unique solution to (4) using a notion of (finite or infinite) escape time under the local Lipschitz continuity condition (6) without the growth condition (7) is given in [41]. Alternatively, existence and uniqueness of solutions even when the uniform Lipschitz continuity condition (6) does not hold are given in ([38], p. 152).

The unique solution to (4) determines a $\mathbb{R}^n$ -valued, time homogeneous Feller continuous Markov process $x(\cdot)$, and hence, its stationary Feller transition probability function is given by ([31], Theorem 3.4), ([32], Theorem 9.2.8))

$$\mathbb{P}(x(t) \in \mathcal{B} | x(t_0) \stackrel{\text{a.s.}}{=} x_0) = \mathbb{P}(0, x_0, t - t_0, \mathcal{B}), \quad x_0 \in \mathbb{R}^n, \quad (8)$$

for all $t \geq t_0$ and all Borel subsets $\mathcal{B}$ of $\mathbb{R}^n$, where $\mathbb{P}(\sigma, x, t, \mathcal{B}), t \geq \sigma$, denotes the probability of transition of the point $x \in \mathbb{R}^n$ at time instant s into the set $\mathcal{B} \subset \mathbb{R}^n$ at time instant t . Recall that every continuous process with Feller transition probability function is also a strong Markov process ([31] p. 101). Finally, we say that the dynamical system (4) is *convergent in probability* with respect to the closed set $\mathcal{H}_n^{D_c} \subseteq \mathcal{H}_n^D$ if and only if the pointwise $\lim_{t \rightarrow \infty} s(t, x, \omega)$ exists for every $x \in \mathcal{D}_c \subseteq \mathbb{R}^n$ and $\omega \in \Omega$.

Here, the measurable map $s : [0, \tau_x) \times \mathcal{D} \times \Omega \rightarrow \mathcal{D}$ is the *dynamic* or *flow* of the stochastic dynamical system (2) and, for all $t, \tau \in [0, \tau_x)$, satisfies the *cocycle* property $s(\tau, s(t, x), \omega) = s(t + \tau, x, \omega)$ and the *identity* (on $\mathcal{D}$) property $s(0, x, \omega) = x$ for all $x \in \mathcal{D}$ and $\omega \in \Omega$. The measurable map $s_t \triangleq s(t, \cdot, \omega) : \mathcal{D} \rightarrow \mathcal{D}$ is continuously differentiable for all $t \in [0, \tau_x)$ outside a $\mathbb{P}$ -null set and the sample path trajectory $s^x \triangleq s(\cdot, x, \omega) : [0, \tau_x) \rightarrow \mathcal{D}$ is continuous in $\mathcal{D}$ for all $t \in [0, \tau_x)$. Thus, for every $x \in \mathcal{D}$, there exists a trajectory of measures defined for all $t \in [0, \tau_x)$ satisfying the dynamical processes (4) with initial condition $x(0) \stackrel{\text{a.s.}}{=} x_0$. For simplicity of exposition we write $s(t, x)$ for $s(t, x, \omega)$ omitting its dependence on ω .

Definition 2. A point $p \in \mathcal{D}$ is a limit point of the trajectory $s(\cdot, x)$ of (4) if there exists a monotonic sequence $\{t_n\}_{n=0}^\infty$ of positive numbers, with $t_n \rightarrow \infty$ as $n \rightarrow \infty$, such that $s(t_n, x) \xrightarrow{\text{a.s.}} p$ as $n \rightarrow \infty$. The set of all limit points of $s(t, x), t \geq 0$, is the limit set $\omega(x)$ of $s(\cdot, x)$ of (4).

It is important to note that the ω -limit set of a stochastic dynamical system is a ω -limit set of a trajectory of measures, that is, $p \in \omega(x)$ is a weak limit of a sequence of measures taken along every sample continuous bounded trajectory of (4). It can be shown that the ω -limit set of a stationary stochastic dynamical system attracts bounded sets and is measurable with respect to the σ -algebra of invariant sets. Thus, the measures of the stochastic process $x(\cdot)$ tend to an invariant set of measures and $x(t)$ asymptotically tends to the closure of the support set (i.e., kernel) of this set of measures almost surely.

However, unlike deterministic dynamical systems, wherein ω -limit sets serve as global attractors, in stochastic dynamical systems stochastic invariance (see Definition 4) leads to ω -limit sets being defined for each fixed sample $\omega \in \Omega$ of the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and hence, are path-wise attractors. This is due to the fact that a cocycle property rather than a semigroup property holds for stochastic dynamical systems. For details see [42–44].

Definition 3 ([33], Def. 7.7). Let $x(\cdot)$ be a time-homogeneous Markov process in $\mathcal{H}_n^D$ and let $V : \mathcal{D} \rightarrow \mathbb{R}$. Then the infinitesimal generator $\mathcal{L}$ of $x(t)$, $t \geq 0$, with $x(0) \stackrel{\text{a.s.}}{=} x_0$, is defined by

$$\mathcal{L}V(x_0) \triangleq \lim_{t \rightarrow 0^+} \frac{\mathbb{E}^{x_0}[V(x(t))] - V(x_0)}{t}, \quad x_0 \in \mathcal{D}, \quad (9)$$

where $\mathbb{E}^{x_0}$ denotes the expectation with respect to the transition probability measure $\mathbb{P}^{x_0}(x(t) \in \mathcal{B}) \triangleq \mathbb{P}(0, x_0, t, \mathcal{B})$.

If $V \in C^2$ and has a compact support [45], and $x(t)$, $t \geq 0$, satisfies (4), then the limit in (9) exists for all $x \in \mathcal{D}$ and the infinitesimal generator $\mathcal{L}$ of $x(t)$, $t \geq 0$, can be characterized by the system drift and diffusion functions $f(x)$ and $D(x)$ defining the stochastic dynamical system (4) and is given by ([33], Theorem 7.9)

$$\mathcal{L}V(x) \triangleq \frac{\partial V(x)}{\partial x} f(x) + \frac{1}{2} \text{tr} D^T(x) \frac{\partial^2 V(x)}{\partial x^2} D(x), \quad x \in \mathcal{D}. \quad (10)$$

Next, we extend Proposition 2.1 of [30] to stochastic dynamical systems. First, however, the following definitions on stochastic invariance and essentially nonnegative vector fields are needed.

Definition 4. A relatively open set $\mathcal{D} \subset \mathbb{R}^n$ is invariant with respect to (4) if $\mathcal{D}$ is Borel and, for all $x_0 \in \mathcal{D}$, $\mathbb{P}^{x_0}(x(t) \in \mathcal{D}) = 1$, $t \geq 0$.

Definition 5. Let $f = [f_1, \dots, f_n]^T : \mathcal{D} \subseteq \mathbb{R}_+^n \rightarrow \mathbb{R}^n$. Then f is essentially nonnegative if $f_i(x) \geq 0$ for all $i = 1, \dots, n$ and $x \in \mathbb{R}_+^n$ such that $x_i = 0$, where x_i denotes the i -th component of x .

Proposition 1. Suppose $\mathbb{R}_+^n \subset \mathcal{D}$. Then $\mathbb{R}_+^n$ is an invariant set with respect to (4) if and only if $f : \mathcal{D} \rightarrow \mathbb{R}^n$ is essentially nonnegative and $D_{(i,j)}(x) = 0$, $j = 1, \dots, d$, whenever $x_i = 0$, $i = 1, \dots, n$.

Proof. Define $\text{dist}(x, \mathbb{R}_+^n) \triangleq \inf_{y \in \mathbb{R}_+^n} \|x - y\|$, $x \in \mathbb{R}^n$. Now, suppose $f : \mathcal{D} \rightarrow \mathbb{R}^n$ is essentially nonnegative and let $x \in \mathbb{R}_+^n$. For every $i \in \{1, \dots, q\}$, if $x_i = 0$, then $x_i + hf_i(x) + \text{row}_i(D(x))[w(h, \omega) - w(0, \omega)] = hf_i(x) \geq 0$ for all $h \geq 0$ and all $\omega \in \Omega$, whereas, if $x_i > 0$, then it follows from the continuity of $D(\cdot)$ and the sample continuity of $w(\cdot)$ that $x_i + hf_i(x) + \text{row}_i(D(x))[w(h, \omega) - w(0, \omega)] \geq 0$ for all $|h|$ sufficiently small and all $\omega \in \Omega$. Thus, $x + hf(x) + \text{row}_i(D(x))[w(h, \omega) - w(0, \omega)] \in \mathbb{R}_+^n$ for all sufficiently small $h > 0$ and all $\omega \in \Omega$, and hence, $\lim_{h \rightarrow 0^+} \text{dist}(x + hf(x) + \text{row}_i(D(x))[w(h, \omega) - w(0, \omega)], \mathbb{R}_+^n)/h = 0$. It now follows from Lemma 2.1 of [46], with $x(0) \stackrel{\text{a.s.}}{=} x_0$, that $\mathbb{P}^{x_0}(x(t) \in \mathbb{R}_+^n) = 1$ for all $t \in [0, \tau_{x_0})$.

Conversely, suppose that $\mathbb{R}_+^n$ is invariant with respect to (4), let $\mathbb{P}^{x_0}(x(0) \in \mathbb{R}_+^n) = 1$, and suppose, ad absurdum, x is such that there exists $i \in \{1, \dots, q\}$ such that $x_i(0) \stackrel{\text{a.s.}}{=} 0$ and $f_i(x(0))h + \text{row}_i(D(x))[w(h, \omega) - w(0, \omega)] < 0$ for all $\omega \in \Omega$. Then, since f and D are continuous and a Wiener process $w(\cdot)$ can be positive or negative with equal probability, there exists sufficiently small $h > 0$ such that $\mathbb{P}^{x_0}(f_i(x(t))dt + \text{row}_i(D(x(t)))dw(t) < 0) \neq 0$ for all $t \in [0, h)$, where $x(t)$ is the solution to (4). Hence, $x_i(t)$ is strictly decreasing on $[0, h)$ with nonzero probability, and thus, $\mathbb{P}^{x_0}(x(t) \in \mathbb{R}_+^n) \neq 1$ for all $t \in (0, h)$, which leads to a contradiction. $\square$

It follows from Proposition 1 that if $x_0 \geq 0$, then $x(t) \stackrel{\text{a.s.}}{\geq} 0$, $t \geq 0$, if and only if f is essentially nonnegative and $D_{(i,j)}(x) = 0$, $j = 1, \dots, d$, whenever $x_i = 0$, $i = 1, \dots, n$. In this case, we say that (4) is a stochastic nonnegative dynamical system. Henceforth, we assume that f and D are such that the nonlinear stochastic dynamical system (4) is a stochastic nonnegative dynamical system.

3. Stability Theory for Stochastic Nonnegative Dynamical Systems

In this section, we establish key stability results in probability for stochastic nonnegative dynamical systems. The mathematical machinery used is supermartingale theory and ergodic theory of Markov processes [31]. Specifically, deterministic stability theory is extended to stochastic dynamical systems by establishing supermartingale properties of Lyapunov functions. The following definition introduces several notions of stability in probability for the equilibrium solution $x(t) \stackrel{\text{a.s.}}{=} x_e \in \bar{\mathbb{R}}_+^n$ of the stochastic nonnegative dynamical system (4) for $\mathcal{I}_{x(0)} = [0, \infty)$.

Definition 6. (i) The equilibrium solution $x(t) \stackrel{\text{a.s.}}{=} x_e \in \bar{\mathbb{R}}_+^n$ to (4) is Lyapunov stable in probability with respect to $\bar{\mathbb{R}}_+^n$ if, for every $\varepsilon > 0$,

$$\lim_{x_0 \rightarrow x_e} \mathbb{P}^{x_0} \left(\sup_{t \geq 0} \|x(t) - x_e\| > \varepsilon \right) = 0. \quad (11)$$

Equivalently, the equilibrium solution $x(t) \stackrel{\text{a.s.}}{=} x_e \in \bar{\mathbb{R}}_+^n$ to (4) is Lyapunov stable in probability with respect to $\bar{\mathbb{R}}_+^n$ if, for every $\varepsilon > 0$ and $\rho \in (0, 1)$, there exists $\delta = \delta(\rho, \varepsilon) > 0$ such that, for all $x_0 \in \mathcal{B}_\delta(x_e) \cap \bar{\mathbb{R}}_+^n$,

$$\mathbb{P}^{x_0} \left(\sup_{t \geq 0} \|x(t) - x_e\| > \varepsilon \right) \leq \rho. \quad (12)$$

(ii) The equilibrium solution $x(t) \stackrel{\text{a.s.}}{=} x_e \in \bar{\mathbb{R}}_+^n$ to (4) is asymptotically stable in probability with respect to $\bar{\mathbb{R}}_+^n$ if it is Lyapunov stable in probability with respect to $\bar{\mathbb{R}}_+^n$ and

$$\lim_{x_0 \rightarrow x_e} \mathbb{P}^{x_0} \left(\lim_{t \rightarrow \infty} \|x(t) - x_e\| = 0 \right) = 1. \quad (13)$$

Equivalently, the equilibrium solution $x(t) \stackrel{\text{a.s.}}{=} x_e \in \bar{\mathbb{R}}_+^n$ to (4) is asymptotically stable in probability with respect to $\bar{\mathbb{R}}_+^n$ if it is Lyapunov stable in probability with respect to $\bar{\mathbb{R}}_+^n$ and, for every $\rho \in (0, 1)$, there exists $\delta = \delta(\rho) > 0$ such that if $x_0 \in \mathcal{B}_\delta(x_e) \cap \bar{\mathbb{R}}_+^n$, then

$$\mathbb{P}^{x_0} \left(\lim_{t \rightarrow \infty} \|x(t) - x_e\| = 0 \right) \geq 1 - \rho. \quad (14)$$

(iii) The equilibrium solution $x(t) \stackrel{\text{a.s.}}{=} x_e \in \bar{\mathbb{R}}_+^n$ to (4) is globally asymptotically stable in probability with respect to $\bar{\mathbb{R}}_+^n$ if it is Lyapunov stable in probability with respect to $\bar{\mathbb{R}}_+^n$ and, for all $x_0 \in \bar{\mathbb{R}}_+^n$,

$$\mathbb{P}^{x_0} \left(\lim_{t \rightarrow \infty} \|x(t) - x_e\| = 0 \right) = 1. \quad (15)$$

As in deterministic stability theory, for a given $\varepsilon > 0$ the subset $\mathcal{B}_\varepsilon(x_e) \cap \bar{\mathbb{R}}_+^n$ defines a cylindrical region in the (t, x) -space wherein the trajectory $x(t)$, $t \geq 0$, belongs to. However, in stochastic stability theory, for every $x_0 \in \mathcal{B}_\delta(x_e) \cap \bar{\mathbb{R}}_+^n$, there exists a probability of less than or equal to ρ that the system solution $s(t, x_0)$ leaves the subset $\mathcal{B}_\varepsilon(x_e) \cap \bar{\mathbb{R}}_+^n$; and for $x_0 = x_e$ this probability is zero. In other words, the probability of escape is continuous at $x_0 = x_e$ with small deviations from the equilibrium implying a small probability of escape.

The following lemma gives an equivalent characterization of Lyapunov and asymptotic stability in probability with respect to $\bar{\mathbb{R}}_+^n$ in terms of class $\mathcal{K}$, $\mathcal{K}_\infty$, and $\mathcal{KL}$ functions ([47], p. 162).

Lemma 1. (i) The equilibrium solution $x(t) \stackrel{\text{a.s.}}{=} x_e$ to (4) is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$ if and only if for every $\rho > 0$ there exist a class $\mathcal{K}$ function $\alpha_\rho(\cdot)$ and a constant $c = c(\rho) > 0$ such that, for all $x_0 \in \mathcal{B}_c(x_e) \cap \overline{\mathbb{R}}_+^n$,

$$\mathbb{P}^{x_0} (\|x(t) - x_e\| > \alpha_\rho(\|x_0 - x_e\|)) \leq \rho, \quad t \geq 0. \quad (16)$$

(ii) The equilibrium solution $x(t) \stackrel{\text{a.s.}}{=} x_e$ to (4) is asymptotically stable in probability with respect to $\overline{\mathbb{R}}_+^n$ if and only if for every $\rho > 0$ there exist a class $\mathcal{KL}$ function $\beta_\rho(\cdot, \cdot)$ and a constant $c = c(\rho) > 0$ such that, for all $x_0 \in \mathcal{B}_c(x_e) \cap \overline{\mathbb{R}}_+^n$,

$$\mathbb{P}^{x_0} (\|x(t) - x_e\| > \beta_\rho(\|x_0 - x_e\|, t)) \leq \rho, \quad t \geq 0. \quad (17)$$

Proof. (i) Suppose that there exist a class $\mathcal{K}$ function $\alpha_\rho(\cdot)$ and a constant $c = c(\rho) > 0$ such that, for every $\rho > 0$ and $x_0 \in \mathcal{B}_c(x_e) \cap \overline{\mathbb{R}}_+^n$,

$$\mathbb{P}^{x_0} (\|x(t) - x_e\| > \alpha_\rho(\|x_0 - x_e\|)) \leq \rho, \quad t \geq 0. \quad (18)$$

Now, given $\varepsilon > 0$, let $\delta(\rho, \varepsilon) = \min\{c(\rho), \alpha_\rho^{-1}(\varepsilon)\}$. Then, for $x_0 \in \mathcal{B}_\delta(x_e) \cap \overline{\mathbb{R}}_+^n$ and $t \geq 0$,

$$\begin{aligned} \mathbb{P}^{x_0} (\|x(t) - x_e\| > \alpha_\rho(\|x_0 - x_e\|)) &\geq \mathbb{P}^{x_0} (\|x(t) - x_e\| > \alpha_\rho(\delta)) \\ &\geq \mathbb{P}^{x_0} (\|x(t) - x_e\| > \alpha_\rho(\alpha_\rho^{-1}(\varepsilon))) \\ &\geq \mathbb{P}^{x_0} (\|x(t) - x_e\| > \varepsilon). \end{aligned}$$

Therefore, for every given $\varepsilon > 0$ and $\rho > 0$, there exists $\delta > 0$ such that, for all $x_0 \in \mathcal{B}_\delta(x_e) \cap \overline{\mathbb{R}}_+^n$,

$$\mathbb{P}^{x_0} \left(\sup_{t \geq 0} \|x(t) - x_e\| > \varepsilon \right) \leq \rho,$$

which proves that the equilibrium solution $x(t) \stackrel{\text{a.s.}}{=} x_e$ is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$.

Conversely, for every given ε and ρ , let $\bar{\delta}(\varepsilon, \rho)$ be the supremum of all admissible $\delta(\varepsilon, \rho)$. Note that the function $\delta(\cdot, \cdot)$ is positive and nondecreasing in its first argument, but not necessarily continuous. For every $\rho > 0$ choose a class $\mathcal{K}$ function $\gamma_\rho(r)$ such that $\gamma_\rho(r) \leq k\bar{\delta}(r, \rho)$, $0 < k < 1$. Let $c(\rho) = \lim_{r \rightarrow \infty} \gamma_\rho(r)$ and $\alpha_\rho(r) = \gamma_\rho^{-1}(r)$, and note that $\alpha_\rho(\cdot)$ is class $\mathcal{K}$ ([48], Lemma 4.2). Next, for every $\rho > 0$ and $x_0 \in \mathcal{B}_{c(\rho)}(x_e) \cap \overline{\mathbb{R}}_+^n$, let $\varepsilon = \alpha_\rho(\|x_0 - x_e\|)$. Then, $\|x_0 - x_e\| < \bar{\delta}(\varepsilon, \rho)$ and

$$\mathbb{P}^{x_0} \left(\sup_{t \geq 0} \|x(t) - x_e\| > \varepsilon \right) \leq \rho \quad (19)$$

imply

$$\mathbb{P}^{x_0} (\|x(t) - x_e\| > \alpha_\rho(\|x_0 - x_e\|)) \leq \rho, \quad t \geq 0. \quad (20)$$

(ii) Suppose that there exists a class $\mathcal{KL}$ function $\beta(r, s)$ such that (17) is satisfied. Then,

$$\mathbb{P}^{x_0} (\|x(t) - x_e\| > \beta_\rho(\|x_0 - x_e\|, 0)) \leq \rho, \quad t \geq 0,$$

which implies that equilibrium solution $x(t) \stackrel{\text{a.s.}}{=} x_e$ is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$. Moreover, for $x_0 \in \mathcal{B}_{c(\rho)}(x_e) \cap \overline{\mathbb{R}}_+^n$, the solution to (4) satisfies

$$\mathbb{P}^{x_0} (\|x(t) - x_e\| > \beta_\rho(c(\rho), t)) \leq \rho, \quad t \geq 0.$$

Now, letting $t \rightarrow \infty$ yields $\mathbb{P}^{x_0}(\lim_{t \rightarrow \infty} \|x(t) - x_e\| > 0) \leq \rho$ for every $\rho > 0$, and hence, $\mathbb{P}^{x_0}(\lim_{t \rightarrow \infty} \|x(t) - x_e\| = 0) \geq 1 - \rho$, which implies that the equilibrium solution $x(t) \stackrel{\text{a.s.}}{=} x_e$ is asymptotically stable in probability with respect to $\mathbb{R}_+^n$.

Conversely, suppose that the equilibrium solution $x(t) \stackrel{\text{a.s.}}{=} x_e$ is asymptotically stable in probability with respect to $\mathbb{R}_+^n$. In this case, for every $\rho > 0$ there exist a constant $c(\rho) > 0$ and a class $\mathcal{K}$ function $\alpha_\rho(\cdot)$ such that, for every $r \in (0, c(\rho)]$, the solution $x(t)$, $t \geq 0$, to (4) satisfies

$$\mathbb{P}^{x_0} \left(\sup_{t \geq 0} \|x(t) - x_e\| > \alpha_\rho(r) \right) \leq \mathbb{P}^{x_0} \left(\sup_{t \geq 0} \|x(t) - x_e\| > \alpha_\rho(\|x_0 - x_e\|) \right) \leq \rho \quad (21)$$

for all $\|x_0 - x_e\| < r$. Moreover, given $\eta > 0$ there exists $T = T_\rho(\eta, r) \geq 0$ such that

$$\mathbb{P}^{x_0} \left(\sup_{t \geq T_\rho(\eta, r)} \|x(t) - x_e\| > \eta \right) \leq \rho.$$

Let $\bar{T}_\rho(\eta, r)$ be the infimum of all admissible $T_\rho(\eta, r)$ and note that $\bar{T}_\rho(\eta, r)$ is nonnegative and nonincreasing in η , nondecreasing in r , and $\bar{T}_\rho(\eta, r) = 0$ for all $\eta \geq \alpha(r)$. Now, let

$$W_{r,\rho}(\eta) = \frac{2}{\eta} \int_{\frac{\eta}{2}}^{\eta} \bar{T}_\rho(s, r) ds + \frac{r}{\eta} \geq \bar{T}_\rho(\eta, r) + \frac{r}{\eta}$$

and note that $W_{r,\rho}(\eta)$ is positive and has the following properties: (i) For every fixed r and ρ , $W_{r,\rho}(\eta)$ is continuous, strictly decreasing, and $\lim_{\eta \rightarrow \infty} W_{r,\rho}(\eta) = 0$; and (ii) for every fixed η and ρ , $W_{r,\rho}(\eta)$ is strictly increasing in r .

Next, let $U_{r,\rho} = W_{r,\rho}^{-1}$ and note that $U_{r,\rho}$ satisfies properties (i) and (ii) of $W_{r,\rho}$, and $\bar{T}_\rho(U_{r,\rho}(\sigma), r) < W_{r,\rho}(U_{r,\rho}(\sigma)) = \sigma$. Therefore,

$$\mathbb{P}^{x_0} (\|x(t) - x_e\| > U_{r,\rho}(t)) \leq \rho, \quad t \geq 0, \quad (22)$$

for all $\|x_0 - x_e\| < r$. Now, using (21) and (22) it follows that

$$\mathbb{P}^{x_0} \left(\|x(t) - x_e\| > \sqrt{\alpha_\rho(\|x_0 - x_e\|) U_{c(\rho),\rho}(t)} \right) \leq \rho, \quad \|x_0 - x_e\| < c(\rho), \quad t \geq 0.$$

Thus, inequality (17) is satisfied with $\beta_\rho(\|x_0 - x_e\|, t) = \sqrt{\alpha_\rho(\|x_0 - x_e\|) U_{c(\rho),\rho}(t)}$. $\square$

Next, we present sufficient conditions for Lyapunov and asymptotic stability in probability for nonlinear stochastic nonnegative dynamical systems. First, however, the following definition of a recurrent process relative to a domain $\mathcal{D}_r$ is needed.

Definition 7. A Markov process $x(\cdot)$ in $\mathcal{H}_n^{\mathcal{D}}$ is recurrent relative to the domain $\mathcal{D}_r$ or, equivalently, $\mathcal{D}_r \subset \mathcal{D}$ is recurrent in $\mathcal{D}$, if there exists a finite-time $t > 0$ such that

$$\mathbb{P}^x (x(t) \in \mathcal{D}_r) = 1. \quad (23)$$

In addition, $\mathcal{D}_r$ is positive recurrent if

$$\sup_{x \in \mathcal{D}_c} \mathbb{E}^x \inf\{t \geq 0 : x(t) \in \mathcal{D}_r\} < \infty \quad (24)$$

for every compact set $\mathcal{D}_c \subset \mathcal{D}$.

Theorem 1. Let $\mathcal{D}$ be an open subset relative to $\overline{\mathbb{R}}_+^n$ that contains x_e . Consider the nonlinear stochastic dynamical system (4) where f is essentially nonnegative and $f(x_e) = 0$, $D_{(i,j)}(x) = 0$, $j = 1, \dots, d$, whenever $x_i = 0$, $i = 1, \dots, n$, and $D(x_e) = 0$. Assume that there exists a two-times continuously differentiable function $V : \mathcal{D} \rightarrow \mathbb{R}$ such that

$$V(x_e) = 0, \quad (25)$$

$$V(x) > 0, \quad x \in \mathcal{D}, \quad x \neq x_e, \quad (26)$$

$$\frac{\partial V(x)}{\partial x} f(x) + \frac{1}{2} \text{tr} D^T(x) \frac{\partial^2 V(x)}{\partial x^2} D(x) \leq 0, \quad x \in \mathcal{D}. \quad (27)$$

Then the equilibrium solution $x(t) \equiv x_e$ to (4) is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$. If, in addition,

$$\frac{\partial V(x)}{\partial x} f(x) + \frac{1}{2} \text{tr} D^T(x) \frac{\partial^2 V(x)}{\partial x^2} D(x) < 0, \quad x \in \mathcal{D}, \quad x \neq x_e, \quad (28)$$

then the equilibrium solution $x(t) \equiv x_e$ to (4) is asymptotically stable in probability with respect to $\overline{\mathbb{R}}_+^n$. Finally, if $\mathcal{D} = \overline{\mathbb{R}}_+^n$ and $V(\cdot)$ is radially unbounded, then the equilibrium solution $x(t) \equiv x_e$ to (4) is globally asymptotically stable in probability with respect to $\overline{\mathbb{R}}_+^n$.

Proof. Let $\delta > 0$ be such that $\mathcal{B}_\delta(x_e) \cap \overline{\mathbb{R}}_+^n \subseteq \mathcal{D}$, define $V_\delta \triangleq \inf_{x \in \mathcal{D} \setminus \mathcal{B}_\delta(x_e) \cap \overline{\mathbb{R}}_+^n} V(x) > 0$, and let τ_δ be the stopping time wherein the trajectory $x(t)$, $t \geq 0$, of (4) exits the bounded domain $\mathcal{B}_\delta(x_e) \cap \overline{\mathbb{R}}_+^n \subseteq \mathcal{D}$ with $\tau_\delta(t) \triangleq \min\{t, \tau_\delta\}$. Since $V(\cdot)$ is two-times continuously differentiable and (27) holds it follows from Lemma 5.4 of [31] that

$$\mathbb{E}^x [V(x(\tau_\delta(t)))] \leq V(x) \quad (29)$$

for all $x \in \mathcal{B}_\delta(x_e) \cap \overline{\mathbb{R}}_+^n$ and $t \geq 0$. Now, using Chebyshev's inequality ([31], p. 29) yields

$$\mathbb{P}^x \left(\sup_{0 \leq s \leq t} \|x(s) - x_e\| > \delta \right) \leq \frac{\mathbb{E}^x [V(x(\tau_\delta(t)))]}{V_\delta} \leq \frac{V(x)}{V_\delta}. \quad (30)$$

Next, taking the limit as $t \rightarrow \infty$, (30) yields

$$\mathbb{P}^x \left(\sup_{s \geq 0} \|x(s) - x_e\| > \delta \right) \leq \frac{V(x)}{V_\delta}, \quad (31)$$

and hence, Lyapunov stability in probability with respect to $\overline{\mathbb{R}}_+^n$ follows from the continuity of $V(\cdot)$ and (25).

To prove asymptotic stability in probability with respect to $\overline{\mathbb{R}}_+^n$, note that the stochastic process $V(x(\tau_\delta(t)))$ is a supermartingale ([31], Lemma 5.4), and hence, it follows from Theorem 5.1 of [31] that

$$\lim_{t \rightarrow \infty} V(x(\tau_\delta(t))) \stackrel{\text{a.s.}}{=} \nu. \quad (32)$$

Let $\mathfrak{B}^x$ denote the set of all sample trajectories of (4) starting from $x \in \overline{\mathbb{R}}_+^n$ for which $\tau_\delta = \infty$. Since the equilibrium solution $x(t) \equiv x_e$ to (4) is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$, it follows that

$$\lim_{x \rightarrow x_e} \mathbb{P}^x(\mathfrak{B}^x) = 1. \quad (33)$$

Next, it follows from Theorem 3.9 of [31] and (28) that all sample trajectories contained in $\mathfrak{B}^x$, except for a set of trajectories with measure zero, satisfy $\inf_{t > 0} \|x(t) - x_e\| = 0$. Moreover, it follows from Lemma 5.3 of [31] that

$$\liminf_{t \rightarrow \infty} \|x(t) - x_e\| = 0, \quad (34)$$

and hence, using (25), $\liminf_{t \rightarrow \infty} V(x(t)) = 0$. Now, (32) implies

$$\lim_{t \rightarrow \infty} V(x(\tau_\delta(t))) = \lim_{t \rightarrow \infty} V(x(t)) \quad (35)$$

for almost all sample trajectories in $\mathfrak{B}^x$, and hence,

$$\lim_{t \rightarrow \infty} V(x(t)) = \liminf_{t \rightarrow \infty} V(x(t)) = 0, \quad (36)$$

which, using (25) and (26), further implies, that

$$\lim_{t \rightarrow \infty} \|x(t) - x_e\| = 0. \quad (37)$$

Now, asymptotic stability in probability with respect to $\overline{\mathbb{R}}_+^n$ is direct consequence of (33) and (37).

Finally, to prove global asymptotic stability in probability with respect to $\overline{\mathbb{R}}_+^n$ note that it follows from Lyapunov stability in probability with respect to $\overline{\mathbb{R}}_+^n$ that, for every $\varepsilon > 0$ and $\rho = \varepsilon$, there exists $\delta > 0$ such that, for all $x \in \mathcal{B}_\delta(x_e) \cap \overline{\mathbb{R}}_+^n$,

$$\mathbb{P}^x \left(\sup_{t > 0} \|x(t) - x_e\| > \varepsilon \right) < \varepsilon. \quad (38)$$

Moreover, it follows from Lemma 3.9, Theorem 3.9 of [31], and the radial unboundedness of $V(\cdot)$ that the solution $x(t)$, $t \geq 0$, of (4) is recurrent relative to the domain $\mathcal{B}_\varepsilon(x_e) \cap \overline{\mathbb{R}}_+^n$ for every $\varepsilon > 0$. Thus, $\tilde{\tau}_\delta \stackrel{\text{a.s.}}{<} \infty$, where $\tilde{\tau}_\delta$ is the first hitting time of the trajectories starting from the set $\overline{\mathbb{R}}_+^n \setminus \mathcal{B}_\delta(x_e)$ and transitioning into the set $\mathcal{B}_\delta(x_e) \cap \overline{\mathbb{R}}_+^n$.

Now, using the strong Markov property of solutions and choosing $\delta > 0$ such that $x \in \overline{\mathbb{R}}_+^n \setminus \mathcal{B}_\delta(x_e)$ yields

$$\begin{aligned} & \mathbb{P}^x \left(\limsup_{t \rightarrow \infty} \|x(t) - x_e\| > \varepsilon \right) \\ &= \int_{\sigma=0}^{\infty} \int_{y \in \partial \mathcal{B}_\delta(x_e) \cap \overline{\mathbb{R}}_+^n} \mathbb{P}(\tilde{\tau}_\delta \in d\sigma, x(\tilde{\tau}_\delta) \in dy) \mathbb{P}^y \left(\limsup_{t \rightarrow \infty} \|x(t) - x_e\| > \varepsilon \right) \\ &= \int_{\sigma=0}^{\infty} \int_{y \in \partial \mathcal{B}_\delta(x_e) \cap \overline{\mathbb{R}}_+^n} \mathbb{P}(\tilde{\tau}_\delta \in d\sigma, x(\tilde{\tau}_\delta) \in dy) \mathbb{P}^y \left(\sup_{t > 0} \|x(t) - x_e\| > \varepsilon \right) \\ &\leq \varepsilon, \end{aligned} \quad (39)$$

which proves global asymptotic stability in probability with respect to $\overline{\mathbb{R}}_+^n$. $\square$

As noted in [37], a more general stochastic stability notion can also be introduced here involving stochastic stability and convergence to an invariant (stationary) distribution. In this case, state convergence is not to an equilibrium point but rather to a stationary distribution. This framework can relax the vanishing perturbation assumption $D(x_e) = 0$ at the equilibrium point x_e and requires a more involved analysis framework showing stability of the underlying Markov semigroup [49].

As in nonlinear stochastic dynamical system theory [31], converse Lyapunov theorems for Lyapunov and asymptotic stability in probability for stochastic nonnegative dynamical systems can also be established. However, in this case, a non-degeneracy condition on $D(x)$, $x \in \mathcal{D}$, is required [31].

Finally, we establish a stochastic version of the Krasovskii-LaSalle stability theorem for nonnegative dynamical systems. For nonlinear stochastic dynamical systems this result is due to Mao [50].

Theorem 2. Consider the nonlinear stochastic nonnegative dynamical system (4). Let $\mathcal{D} \subseteq \overline{\mathbb{R}}_+^n$ be an invariant set with respect to (4) and assume that there exists a two-times continuously differentiable function $V : \mathcal{D} \rightarrow \overline{\mathbb{R}}_+$ and a continuous function $\eta : \overline{\mathbb{R}}_+ \rightarrow \overline{\mathbb{R}}_+$ such that

$$\frac{\partial V(x)}{\partial x} f(x) + \frac{1}{2} \text{tr } D^T(x) \frac{\partial^2 V(x)}{\partial x^2} D(x) \leq -\eta(V(x)), \quad x \in \mathcal{D}. \quad (40)$$

Then, for every $x_0 \in \mathcal{D}$, $\lim_{t \rightarrow \infty} V(x(t))$ exists and is finite almost surely, and

$$\lim_{t \rightarrow \infty} \eta(V(x(t))) \stackrel{\text{a.s.}}{=} 0. \quad (41)$$

Proof. Since $\mathcal{D} \subseteq \overline{\mathbb{R}}_+^n$ is invariant with respect to (4), it follows that, for all $x_0 \in \mathcal{D}$, $\mathbb{P}^{x_0}(x(t) \in \mathcal{D}) = 1$, $t \geq 0$. Furthermore, using Itô's chain rule formula and (40) we have

$$\begin{aligned} V(x(t)) &= V(x_0) + \int_0^t \mathcal{L}V(x(\sigma)) d\sigma + \int_0^t \frac{\partial V(x(\sigma))}{\partial x} D(x(\sigma)) dw(\sigma) \\ &\leq V(x_0) - \int_0^t \eta(V(x(\sigma))) d\sigma + \int_0^t \frac{\partial V(x(\sigma))}{\partial x} D(x(\sigma)) dw(\sigma). \end{aligned} \quad (42)$$

Now, it follows from Theorem 7 of ([51], p. 139) that $\lim_{t \rightarrow \infty} V(x(t))$ exists and is finite almost surely, and

$$\lim_{t \rightarrow \infty} \int_0^t \eta(V(x(\sigma))) d\sigma \stackrel{\text{a.s.}}{<} \infty. \quad (43)$$

To show that $\lim_{t \rightarrow \infty} \eta(V(x(t))) \stackrel{\text{a.s.}}{=} 0$ suppose, ad absurdum, that there exists a sample space $\bar{\Omega} \subset \Omega$ such that $\mathbb{P}(\bar{\Omega}) > 0$ and

$$\limsup_{t \rightarrow \infty} \eta(V(x(t, \omega))) > 0, \quad \omega \in \bar{\Omega}. \quad (44)$$

Let $\{t_n\}_{n=0}^\infty$, $n \in \mathbb{Z}_+$, be a monotonic sequence with $t_n + 1 < t_{n+1}$ and note that there exist $\varepsilon > 0$ and $N \in \mathbb{Z}_+$ such that

$$\eta(V(x(t_n, \omega))) > \varepsilon, \quad n \geq N. \quad (45)$$

Now, it follows from the continuity of $\eta(\cdot)$ and $V(\cdot)$, and the sample continuity of $x(\cdot)$ that there exist $\delta > 0$, $\delta_1 > 0$, and $\delta_2 > 0$, such that if $|V(x(t_n, \omega)) - V(x(t, \omega))| \leq \delta_2$, then

$$|\eta(V(x(t_n, \omega))) - \eta(V(x(t, \omega)))| \leq \frac{\varepsilon}{2}, \quad (46)$$

if $\|x(t_n, \omega) - x(t, \omega)\| \leq \delta_1$, then

$$|V(x(t_n, \omega)) - V(x(t, \omega))| \leq \delta_2, \quad (47)$$

and if $|t_n - t| \leq \delta$, then

$$\|x(t_n, \omega) - x(t, \omega)\| \leq \delta_1. \quad (48)$$

Thus, using (45) and (46) it follows that, for all $|t_n - t| \leq \delta$ and $n \geq N$,

$$\eta(V(x(t, \omega))) \geq \eta(V(x(t_n, \omega))) - |\eta(V(x(t_n, \omega))) - \eta(V(x(t, \omega)))| > \frac{\varepsilon}{2}, \quad (49)$$

and hence,

$$\lim_{t \rightarrow \infty} \int_0^t \eta(V(x(\sigma))) d\sigma \geq \sum_{n=N}^\infty \int_{t_n}^{t_n+\delta} \eta(V(x(\sigma))) d\sigma \geq \sum_{n=N}^\infty \frac{\varepsilon \delta}{2} = \infty, \quad (50)$$

which contradicts (43). Thus, $\lim_{t \rightarrow \infty} \eta(V(x(t))) \stackrel{\text{a.s.}}{=} 0$. $\square$

Note that if we define $\mathcal{D}_\eta \triangleq \{v \geq 0 : \eta(v) = 0\}$, then it can be shown that $\mathcal{D}_\eta \neq \emptyset$ and (41) implies

$$\lim_{t \rightarrow \infty} \text{dist}(V(x(t)), \mathcal{D}_\eta) \stackrel{\text{a.s.}}{=} 0, \quad (51)$$

that is, $V(x(t))$ will asymptotically approach the set $\mathcal{D}_\eta$ with probability one. Thus, if $\eta(V(x)) = x_e$ if and only if $x = x_e$ and $V(\cdot)$ is positive definite with respect to x_e , then it follows from Theorems 1 and 2 that the equilibrium solution $x(t) \stackrel{\text{a.s.}}{=} x_e$ to (4) is asymptotically stable in probability with respect to $\overline{\mathbb{R}}_+^n$.

4. Semistability of Stochastic Nonnegative Dynamical Systems

As shown in [30], thermodynamic systems give rise to systems that possess a continuum of equilibria. In this section, we develop a stability analysis framework for stochastic systems having a continuum of equilibria. Since, as noted in [37,52], every neighborhood of a non-isolated equilibrium contains another equilibrium, a non-isolated equilibrium cannot be asymptotically stable. Hence, asymptotic stability is not the appropriate notion of stability for systems having a continuum of equilibria. Two notions that are of particular relevance to such systems are *convergence* and *semistability*. Convergence is the property whereby every system solution converges to a limit point that may depend on the system initial condition. Semistability is the additional requirement that all solutions converge to limit points that are Lyapunov stable. Semistability for an equilibrium thus implies Lyapunov stability, and is implied by asymptotic stability.

In this section, we present necessary and sufficient conditions for stochastic semistability. It is important to note that stochastic semistability theory was also developed in [37] for a stronger set of stability in probability definitions. The results in this section, though parallel the results in [37], are predicated on a weaker set of stability in probability definitions, and hence, provide a stronger set of semistability results. First, we present several key propositions. The following proposition gives a sufficient condition for a trajectory of (4) to converge to a limit point. For this result, $\mathcal{D}_c \subseteq \mathcal{D} \subseteq \overline{\mathbb{R}}_+^n$ denotes a positively invariant set with respect to (4) and $s_t(\mathcal{H}_n^{\mathcal{D}_c})$ denotes the image of $\mathcal{H}_n^{\mathcal{D}_c} \subset \mathcal{H}_n^{\mathcal{D}}$ under the flow $s_t : \mathcal{H}_n^{\mathcal{D}_c} \rightarrow \mathcal{H}_n^{\mathcal{D}}$, that is, $s_t(\mathcal{H}_n^{\mathcal{D}_c}) \triangleq \{y : y = s_t(x_0) \text{ for some } x_0 \in \mathcal{H}_n^{\mathcal{D}_c}\} \stackrel{\text{a.s.}}{=} \{y : y = s_t(x_0) \text{ for some } x_0 \in \mathcal{H}_n^{\mathcal{D}_c}\}$.

Proposition 2. Consider the nonlinear stochastic nonnegative dynamical system (4) and let $x \in \mathcal{D}_c$. If the limit set $\omega(x)$ of (4) contains a Lyapunov stable in probability (with respect to $\overline{\mathbb{R}}_+^n$) equilibrium point y , then $\lim_{x \rightarrow y} \mathbb{P}^x(\|\lim_{t \rightarrow \infty} s(t, x) - y\| = 0) = 1$, that is, $\omega(x) \stackrel{\text{a.s.}}{=} \{y\}$ as $x \rightarrow y$.

Proof. Suppose $y \in \omega(x)$ is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$ and let $\mathcal{N}_\epsilon \subseteq \mathcal{D}_c$ be a relatively open neighborhood of y . Since y is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$, there exists a relatively open neighborhood $\mathcal{N}_\delta \subset \mathcal{D}_c$ of y such that $s_t(\mathcal{H}_n^{\mathcal{N}_\delta}) \subseteq \mathcal{H}_n^{\mathcal{N}_\epsilon}$ as $x \rightarrow y$ for every $t \geq 0$. Now, since $y \in \omega(x)$, it follows that there exists $\tau \geq 0$ such that $s(\tau, x) \in \mathcal{H}_n^{\mathcal{N}_\delta}$. Hence, $s(t + \tau, x) = s_t(s(\tau, x)) \in s_t(\mathcal{H}_n^{\mathcal{N}_\delta}) \subseteq \mathcal{H}_n^{\mathcal{N}_\epsilon}$ for every $t > 0$. Since $\mathcal{N}_\epsilon \subseteq \mathcal{D}_c$ is arbitrary, it follows that $y \stackrel{\text{a.s.}}{=} \lim_{t \rightarrow \infty} s(t, x)$. Thus, $\lim_{n \rightarrow \infty} s(t_n, x) \stackrel{\text{a.s.}}{=} y$ as $x \rightarrow y$ for every sequence $\{t_n\}_{n=1}^\infty$, and hence, $\omega(x) \stackrel{\text{a.s.}}{=} \{y\}$ as $x \rightarrow y$. $\square$

The following definition introduces the notion of stochastic semistability.

Definition 8. An equilibrium solution $x(t) \stackrel{\text{a.s.}}{=} x_e \in \mathcal{E}$ of (4) is stochastically semistable with respect to $\overline{\mathbb{R}}_+^n$ if the following statements hold.

- (i) For every $\epsilon > 0$, $\lim_{x_0 \rightarrow x_e} \mathbb{P}^{x_0}(\sup_{0 \leq t < \infty} \|x(t) - x_e\| > \epsilon) = 0$. Equivalently, for every $\epsilon > 0$ and $\rho \in (0, 1)$, there exist $\delta = \delta(\epsilon, \rho) > 0$ such that, for all $x_0 \in \mathcal{B}_\delta(x_e) \cap \overline{\mathbb{R}}_+^n$,

$$\mathbb{P}^{x_0} \left(\sup_{0 \leq t < \infty} \|x(t) - x_e\| > \epsilon \right) \leq \rho.$$

(ii) $\lim_{\text{dist}(x_0, \mathcal{E}) \rightarrow 0} \mathbb{P}^{x_0} (\lim_{t \rightarrow \infty} \text{dist}(x(t), \mathcal{E}) = 0) = 1$. Equivalently, for every $\rho \in (0, 1)$, there exist $\delta = \delta(\rho) > 0$ such that if $\text{dist}(x_0, \mathcal{E}) \leq \delta$, then $\mathbb{P}^{x_0} (\lim_{t \rightarrow \infty} \text{dist}(x(t), \mathcal{E}) = 0) \geq 1 - \rho$.

The dynamical system (4) is stochastically semistable with respect to $\overline{\mathbb{R}}_+^n$ if every equilibrium solution of (4) is stochastically semistable with respect to $\overline{\mathbb{R}}_+^n$. Finally, the dynamical system (4) is globally stochastically semistable with respect to $\overline{\mathbb{R}}_+^n$ if (i) holds and $\mathbb{P}^{x_0} (\lim_{t \rightarrow \infty} \text{dist}(x(t), \mathcal{E}) = 0) = 1$ for all $x_0 \in \overline{\mathbb{R}}_+^n$.

Note that if $x(t) \stackrel{\text{a.s.}}{=} x_e \in \mathcal{E}$ only satisfies (i) in Definition 8, then the equilibrium solution $x(t) \stackrel{\text{a.s.}}{=} x_e \in \mathcal{E}$ of (4) is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$.

Definition 9. For a given $\rho \in (0, 1)$, the ρ -domain of semistability with respect to $\overline{\mathbb{R}}_+^n$ is the set of points $x_0 \in \mathcal{D} \subseteq \overline{\mathbb{R}}_+^n$ such that if $x(t)$, $t \geq 0$, is a solution to (4) with $x(0) \stackrel{\text{a.s.}}{=} x_0$, then $x(t)$ converges to a Lyapunov stable (with respect to $\overline{\mathbb{R}}_+^n$) in probability equilibrium point in $\mathcal{D}$ with probability greater than or equal to $1 - \rho$.

Note that if (4) is stochastically semistable, then its ρ -domain of semistability contains the set of equilibria in its interior.

Next, we present alternative equivalent characterizations for stochastic semistability of (4). This result is an extension of Proposition 2.2 of [37] to the more general semistability definition presented in this paper.

Proposition 3. Consider the nonlinear stochastic nonnegative dynamical system $\mathcal{G}$ given by (4). Then the following statements are equivalent:

- (i) $\mathcal{G}$ is stochastically semistable with respect to $\overline{\mathbb{R}}_+^n$.
- (ii) For every $x_e \in \mathcal{E}$ and $\rho > 0$, there exist class $\mathcal{K}$ and $\mathcal{L}$ functions $\alpha_\rho(\cdot)$ and $\beta_\rho(\cdot)$, respectively, and $\delta = \delta(x_e, \rho) > 0$ such that, if $x_0 \in \mathcal{B}_\delta(x_e) \cap \overline{\mathbb{R}}_+^n$, then

$$\mathbb{P}^{x_0} (\|x(t) - x_e\| > \alpha_\rho(\|x_0 - x_e\|)) \leq \rho, \quad t \geq 0,$$

and $\mathbb{P}^{x_0} (\text{dist}(x(t), \mathcal{E}) > \beta_\rho(t)) \leq \rho$, $t \geq 0$.

- (iii) For every $x_e \in \mathcal{E}$ and $\rho > 0$, there exist class $\mathcal{K}$ functions $\alpha_{1\rho}(\cdot)$ and $\alpha_{2\rho}(\cdot)$, a class $\mathcal{L}$ function $\beta_\rho(\cdot)$, and $\delta = \delta(x_e, \rho) > 0$ such that, if $x_0 \in \mathcal{B}_\delta(x_e) \cap \overline{\mathbb{R}}_+^n$, then

$$\begin{aligned} \mathbb{P}^{x_0} (\text{dist}(x(t), \mathcal{E}) > \alpha_{2\rho}(\|x_0 - x_e\|)\beta_\rho(t)) \\ \leq \mathbb{P}^{x_0} (\alpha_{1\rho}(\|x(t) - x_e\|) > \alpha_{2\rho}(\|x_0 - x_e\|)) \leq \rho, \quad t \geq 0. \end{aligned}$$

Proof. To show that (i) implies (ii), suppose (4) is stochastically semistable with respect to $\overline{\mathbb{R}}_+^n$ and let $x_e \in \mathcal{E}$. It follows from Lemma 1 that for every $\rho > 0$ there exists $\delta = \delta(x_e, \rho) > 0$ and a class $\mathcal{K}$ function $\alpha_\rho(\cdot)$ such that if $\|x_0 - x_e\| \leq \delta$, then $\mathbb{P}^{x_0} (\|x(t) - x_e\| > \alpha_\rho(\|x_0 - x_e\|)) \leq \rho$, $t \geq 0$. Without loss of generality, we can assume that δ is such that $\overline{\mathcal{B}_\delta(x_e)} \cap \overline{\mathbb{R}}_+^n$ is contained in the ρ -domain of semistability of (4). Hence, for every $x_0 \in \overline{\mathcal{B}_\delta(x_e)} \cap \overline{\mathbb{R}}_+^n$, $\lim_{t \rightarrow \infty} x(t) \stackrel{\text{a.s.}}{=} x^* \in \mathcal{E}$ and, consequently, $\mathbb{P}^{x_0} (\lim_{t \rightarrow \infty} \text{dist}(x(t), \mathcal{E}) = 0) = 1$.

For every $\varepsilon > 0$, $\rho > 0$, and $x_0 \in \overline{\mathcal{B}_\delta(x_e)} \cap \overline{\mathbb{R}}_+^n$, define $T_{x_0}(\varepsilon, \rho)$ to be the infimum of T with the property that $\mathbb{P}^{x_0} (\sup_{t \geq T} \text{dist}(x(t), \mathcal{E}) > \varepsilon) \leq \rho$, that is,

$$T_{x_0}(\varepsilon, \rho) \triangleq \inf \left\{ T : \mathbb{P}^{x_0} \left(\sup_{t \geq T} \text{dist}(x(t), \mathcal{E}) > \varepsilon \right) \leq \rho \right\}.$$

For each $x_0 \in \overline{\mathcal{B}_\delta(x_e)} \cap \overline{\mathbb{R}}_+^n$ and ρ , the function $T_{x_0}(\varepsilon, \rho)$ is nonnegative and nonincreasing in ε , and $T_{x_0}(\varepsilon, \rho) = 0$ for sufficiently large ε .

Next, let $T(\varepsilon, \rho) \triangleq \sup\{T_{x_0}(\varepsilon, \rho) : x_0 \in \overline{\mathcal{B}_\delta(x_e)} \cap \overline{\mathbb{R}_+^n}\}$. We claim that T is well defined. To show this, consider $\varepsilon > 0$, $\rho > 0$, and $x_0 \in \overline{\mathcal{B}_\delta(x_e)} \cap \overline{\mathbb{R}_+^n}$. Since $\mathbb{P}^{x_0} \left(\sup_{t \geq T_{x_0}(\varepsilon, \rho)} \text{dist}(x(t), \mathcal{E}) > \varepsilon \right) \leq \rho$, it follows from the sample continuity of s that, for every $\varepsilon > 0$ and $\rho > 0$, there exists an open neighborhood $\mathcal{U}$ of x_0 such that $\mathbb{P}^{x_0} \left(\sup_{t \geq T_z(\varepsilon, \rho)} \text{dist}(s(t, z), \mathcal{E}) > \varepsilon \right) \leq \rho$ for every $z \in \mathcal{U}$. Hence, $\limsup_{z \rightarrow x_0} T_z(\varepsilon, \rho) \leq T_{x_0}(\varepsilon, \rho)$ implying that the function $x_0 \mapsto T_{x_0}(\varepsilon, \rho)$ is upper semicontinuous at the arbitrarily chosen point x_0 , and hence on $\overline{\mathcal{B}_\delta(x_e)} \cap \overline{\mathbb{R}_+^n}$. Since an upper semicontinuous function defined on a compact set achieves its supremum, it follows that $T(\varepsilon, \rho)$ is well defined. The function $T(\cdot)$ is the pointwise supremum of a collection of nonnegative and nonincreasing functions, and hence is nonnegative and nonincreasing. Moreover, $T(\varepsilon, \rho) = 0$ for every $\varepsilon > \max\{\alpha_\rho(\|x_0 - x_e\|) : x_0 \in \overline{\mathcal{B}_\delta(x_e)} \cap \overline{\mathbb{R}_+^n}\}$.

Let $\psi_\rho(\varepsilon) \triangleq \frac{2}{\varepsilon} \int_{\varepsilon/2}^\varepsilon T(\sigma, \rho) d\sigma + \frac{1}{\varepsilon} \geq T(\varepsilon, \rho) + \frac{1}{\varepsilon}$. The function $\psi_\rho(\varepsilon)$ is positive, continuous, strictly decreasing, and $\psi_\rho(\varepsilon) \rightarrow 0$ as $\varepsilon \rightarrow \infty$. Choose $\beta_\rho(\cdot) = \psi^{-1}(\cdot)$. Then $\beta_\rho(\cdot)$ is positive, continuous, strictly decreasing, and $\lim_{\sigma \rightarrow \infty} \beta_\rho(\sigma) = 0$. Furthermore, $T(\beta_\rho(\sigma), \rho) < \psi_\rho(\beta_\rho(\sigma)) = \sigma$. Hence, $\mathbb{P}^{x_0}(\text{dist}(x(t), \mathcal{E}) > \beta_\rho(t)) \leq \rho$, $t \geq 0$.

Next, to show that (ii) implies (iii), suppose (ii) holds and let $x_e \in \mathcal{E}$. Then it follows from (i) of Lemma 1 that x_e is Lyapunov stable in probability with respect to $\overline{\mathbb{R}_+^n}$. For every $\rho > 0$, choosing x_0 sufficiently close to x_e , it follows from the inequality $\mathbb{P}^{x_0}(\|x(t) - x_e\| > \alpha_\rho(\|x_0 - x_e\|)) \leq \rho$, $t \geq 0$, that trajectories of (4) starting sufficiently close to x_e are bounded, and hence, the positive limit set of (4) is nonempty. Since $\mathbb{P}^{x_0}(\lim_{t \rightarrow \infty} \text{dist}(x(t), \mathcal{E}) = 0) = 1$ as $\text{dist}(x_0, \mathcal{E}) \rightarrow 0$, it follows that the positive limit set is contained in $\mathcal{E}$ as $\text{dist}(x_0, \mathcal{E}) \rightarrow 0$.

Now, since every point in $\mathcal{E}$ is Lyapunov stable in probability with respect to $\overline{\mathbb{R}_+^n}$, it follows from Proposition 2 that $\lim_{t \rightarrow \infty} x(t) \stackrel{\text{a.s.}}{=} x^*$ as $x_0 \rightarrow x^*$, where $x^* \in \mathcal{E}$ is Lyapunov stable in probability with respect to $\overline{\mathbb{R}_+^n}$. If $x^* = x_e$, then it follows using similar arguments as above that there exists a class $\mathcal{L}$ function $\hat{\beta}_\rho(\cdot)$ such that

$$\mathbb{P}^{x_0}(\text{dist}(x(t), \mathcal{E}) > \hat{\beta}_\rho(t)) \leq \mathbb{P}^{x_0}(\|x(t) - x_e\| > \hat{\beta}_\rho(t)) \leq \rho$$

for every x_0 satisfying $\|x_0 - x_e\| < \delta$ and $t \geq 0$. Hence,

$$\mathbb{P}^{x_0} \left(\text{dist}(x(t), \mathcal{E}) > \sqrt{\|x(t) - x_e\|} \sqrt{\hat{\beta}_\rho(t)} \right) \leq \rho, \quad t \geq 0.$$

Next, consider the case where $x^* \neq x_e$ and let $\alpha_{1\rho}(\cdot)$ be a class $\mathcal{K}$ function. In this case, note that

$$\mathbb{P}^{x_0} \left(\lim_{t \rightarrow \infty} \text{dist}(x(t), \mathcal{E}) / \alpha_{1\rho}(\|x(t) - x_e\|) = 0 \right) \geq 1 - \rho,$$

and hence, it follows using similar arguments as above that there exists a class $\mathcal{L}$ function $\beta_\rho(\cdot)$ such that

$$\mathbb{P}^{x_0}(\text{dist}(x(t), \mathcal{E}) > \alpha_{1\rho}(\|x(t) - x_e\|)\beta_\rho(t)) \leq \rho, \quad t \geq 0.$$

Now, note that $\alpha_{1\rho} \circ \alpha_\rho$ is of class $\mathcal{K}$ (by [48]), Lemma 4.2), and hence, (iii) follows immediately.

Finally, to show that (iii) implies (i), suppose (iii) holds and let $x_e \in \mathcal{E}$. Then it follows that for every $\rho > 0$,

$$\mathbb{P}^{x_0}(\alpha_{1\rho}(\|x(t) - x_e\|) > \alpha_{2\rho}(\|x(0) - x_e\|)) \leq \rho, \quad t \geq 0,$$

that is, $\mathbb{P}^{x_0}[\|x(t) - x_e\| > \alpha_\rho(\|x(0) - x_e\|)] \leq \rho$, where $t \geq 0$ and $\alpha_\rho = \alpha_{1\rho}^{-1} \circ \alpha_{2\rho}$ is of class $\mathcal{K}$ (by [48], Lemma 4.2). It now follows from (i) of Lemma 1 that x_e is Lyapunov stable in probability with respect to $\overline{\mathbb{R}_+^n}$. Since x_e was chosen arbitrarily, it follows that every equilibrium point is Lyapunov stable in probability with respect to $\overline{\mathbb{R}_+^n}$. Furthermore, $\mathbb{P}^{x_0}(\lim_{t \rightarrow \infty} \text{dist}(x(t), \mathcal{E}) = 0) \geq 1 - \rho$.

Choosing x_0 sufficiently close to x_e , it follows from the inequality

$$\mathbb{P}^{x_0}(\|x(t) - x_e\| > \alpha_\rho(\|x_0 - x_e\|)) \leq \rho, \quad t \geq 0,$$

that trajectories of (4) are almost surely bounded as $x_0 \rightarrow x_e$, and hence, the positive limit set of (4) is nonempty as $x_0 \rightarrow x_e$. Since every point in $\mathcal{E}$ is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$, it follows from Proposition 2 that $\lim_{t \rightarrow \infty} x(t) \stackrel{\text{a.s.}}{=} x^*$ as $x_0 \rightarrow x^*$, where $x^* \in \mathcal{E}$ is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$. Hence, by Definition 8, (4) is stochastically semistable with respect to $\overline{\mathbb{R}}_+^n$. $\square$

Next, we develop necessary and sufficient conditions for stochastic semistability. First, we present sufficient conditions for stochastic semistability. The following theorems generalize Theorems 3.1 and 3.2 of [37].

Theorem 3. Consider the nonlinear stochastic nonnegative dynamical system (4). Let $\mathcal{Q} \subseteq \overline{\mathbb{R}}_+^n$ be a relatively open neighborhood of $\mathcal{E}$ and assume that there exists a two-times continuously differentiable function $V : \mathcal{Q} \rightarrow \overline{\mathbb{R}}_+$ such that

$$V'(x)f(x) + \frac{1}{2}\text{tr } D^T(x)V''(x)D(x) < 0, \quad x \in \mathcal{Q} \setminus \mathcal{E}. \quad (52)$$

If every equilibrium point of (4) is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$, then (4) is stochastically semistable with respect to $\overline{\mathbb{R}}_+^n$. Moreover, if $\mathcal{Q} = \overline{\mathbb{R}}_+^n$ and $V(x) \rightarrow \infty$ as $\|x\| \rightarrow \infty$, then (4) is globally stochastically semistable with respect to $\overline{\mathbb{R}}_+^n$.

Proof. Since every equilibrium point of (4) is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$ by assumption, for every $z \in \mathcal{E}$, there exists a relatively open neighborhood $\mathcal{V}_z$ of z such that $s([0, \infty) \times \mathcal{V}_z \cap \mathcal{B}_\varepsilon(z))$, $\varepsilon > 0$, is bounded and contained in $\mathcal{Q}$ as $\varepsilon \rightarrow 0$. The set $\mathcal{V}_\varepsilon \triangleq \bigcup_{z \in \mathcal{E}} \mathcal{V}_z \cap \mathcal{B}_\varepsilon(z)$, $\varepsilon > 0$, is a relatively open neighborhood of $\mathcal{E}$ contained in $\mathcal{Q}$. Consider $x \in \mathcal{V}_\varepsilon$ so that there exists $z \in \mathcal{E}$ such that $x \in \mathcal{V}_z \cap \mathcal{B}_\varepsilon(z)$ and $s(t, x) \in \mathcal{H}_n^{\mathcal{V}_z \cap \mathcal{B}_\varepsilon(z)}$, $t \geq 0$, as $\varepsilon \rightarrow 0$. Since $\mathcal{V}_z \cap \mathcal{B}_\varepsilon(z)$ is bounded and invariant with respect to the solution of (4) as $\varepsilon \rightarrow 0$, it follows that $\mathcal{V}_\varepsilon$ is invariant with respect to the solution of (4) as $\varepsilon \rightarrow 0$. Furthermore, it follows from (52) that $\mathcal{L}V(s(t, x)) < 0$, $t \geq 0$, and hence, since $\mathcal{V}_\varepsilon$ is bounded it follows from Theorem 2 that $\lim_{t \rightarrow \infty} \mathcal{L}V(s(t, x)) \stackrel{\text{a.s.}}{=} 0$ as $\varepsilon \rightarrow 0$.

It is easy to see that $\mathcal{L}V(x) \neq 0$ by assumption and $\mathcal{L}V(x_e) = 0$, $x_e \in \mathcal{E}$. Therefore, $s(t, x) \xrightarrow{\text{a.s.}} \mathcal{E}$ as $t \rightarrow \infty$ and $\varepsilon \rightarrow 0$, which implies that $\lim_{\text{dist}(x, \mathcal{E}) \rightarrow 0} \mathbb{P}^x(\lim_{t \rightarrow \infty} \text{dist}(s(t, x), \mathcal{E}) = 0) = 1$. Finally, since every point in $\mathcal{E}$ is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$, it follows from Proposition 2 that $\lim_{t \rightarrow \infty} s(t, x) \stackrel{\text{a.s.}}{=} x^*$ as $x \rightarrow x^*$, where $x^* \in \mathcal{E}$ is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$. Hence, by Definition 8, (4) is semistable. For $\mathcal{Q} = \overline{\mathbb{R}}_+^n$ global stochastic semistability with respect to $\overline{\mathbb{R}}_+^n$ follows from identical arguments using the radially unbounded condition on $V(\cdot)$. $\square$

Next, we present a slightly more general theorem for stochastic semistability wherein we do not assume that all points in $\mathcal{L}V^{-1}(0)$ are Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$ but rather we assume that all points in $(\eta \circ V)^{-1}(0)$ are Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$ for some continuous function $\eta : \overline{\mathbb{R}}_+ \rightarrow \overline{\mathbb{R}}_+$.

Theorem 4. Consider the nonlinear stochastic nonnegative dynamical system (4) and let $\mathcal{Q} \subseteq \overline{\mathbb{R}}_+^n$ be a relatively open neighborhood of $\mathcal{E}$. Assume that there exist a two-times continuously differentiable function $V : \mathcal{Q} \rightarrow \overline{\mathbb{R}}_+$ and a continuous function $\eta : \overline{\mathbb{R}}_+ \rightarrow \overline{\mathbb{R}}_+$ such that

$$V'(x)f(x) + \frac{1}{2}\text{tr } D^T(x)V''(x)D(x) \leq -\eta(V(x)), \quad x \in \mathcal{Q}. \quad (53)$$

If every point in the set $\mathcal{M} \triangleq \{x \in \mathcal{Q} : \eta(V(x)) = 0\}$ is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$, then (4) is stochastically semistable with respect to $\overline{\mathbb{R}}_+^n$. Moreover, if $\mathcal{Q} = \overline{\mathbb{R}}_+^n$ and $V(x) \rightarrow \infty$ as $\|x\| \rightarrow \infty$, then (4) is globally stochastically semistable with respect to $\overline{\mathbb{R}}_+^n$.

Proof. Since, by assumption, (4) is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$ for all $z \in \mathcal{M}$, there exists a relatively open neighborhood $\mathcal{V}_z$ of z such that $s([0, \infty) \times \mathcal{V}_z \cap \mathcal{B}_\varepsilon(z))$, $\varepsilon > 0$, is bounded and contained in $\mathcal{Q}$ as $\varepsilon \rightarrow 0$. The set $\mathcal{V}_\varepsilon \triangleq \bigcup_{z \in \mathcal{M}} \mathcal{V}_z \cap \mathcal{B}_\varepsilon(z)$ is a relatively open neighborhood of $\mathcal{M}$ contained in $\mathcal{Q}$. Consider $x \in \mathcal{V}_\varepsilon$ so that there exists $z \in \mathcal{M}$ such that $x \in \mathcal{V}_z \cap \mathcal{B}_\varepsilon(z)$ and $s(t, x) \in \mathcal{H}_n^{\mathcal{V}_z \cap \mathcal{B}_\varepsilon(z)}$, $t \geq 0$, as $\varepsilon \rightarrow 0$. Since $\mathcal{V}_z$ is bounded it follows that $\mathcal{V}_\varepsilon$ is invariant with respect to the solution of (4) as $\varepsilon \rightarrow 0$. Furthermore, it follows from (53) that $\mathcal{L}V(s(t, x)) \leq -\eta(V(s(t, x)))$, $t \geq 0$, and hence, since $\mathcal{V}_\varepsilon$ is bounded and invariant with respect to the solution of (4) as $\varepsilon \rightarrow 0$, it follows from Theorem 2 that $\lim_{t \rightarrow \infty} \eta(V(s(t, x))) \stackrel{\text{a.s.}}{=} 0$ as $\varepsilon \rightarrow 0$. Therefore, $s(t, x) \xrightarrow{\text{a.s.}} \mathcal{M}$ as $t \rightarrow \infty$ and $\varepsilon \rightarrow 0$, which implies that $\lim_{\text{dist}(x, \mathcal{M}) \rightarrow 0} \mathbb{P}^x(\lim_{t \rightarrow \infty} \text{dist}(s(t, x), \mathcal{M}) = 0) = 1$.

Finally, since every point in $\mathcal{M}$ is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$, it follows from Proposition 2 that $\lim_{t \rightarrow \infty} s(t, x) \stackrel{\text{a.s.}}{=} x^*$ as $x \rightarrow x^*$, where $x^* \in \mathcal{M}$ is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$. Hence, by definition, (4) is semistable. For $\mathcal{Q} = \overline{\mathbb{R}}_+^n$ global stochastic semistability with respect to $\overline{\mathbb{R}}_+^n$ follows from identical arguments using the radially unbounded condition on $V(\cdot)$. $\square$

Example 1. Consider the nonlinear stochastic nonnegative dynamical system on $\mathcal{H}_2$ given by ([37])

$$dx_1(t) = [\sigma_{12}(x_2(t)) - \sigma_{21}(x_1(t))]dt + \gamma(x_2(t) - x_1(t))dw(t), \quad x_1(0) \stackrel{\text{a.s.}}{=} x_{10}, \quad t \geq 0, \quad (54)$$

$$dx_2(t) = [\sigma_{21}(x_1(t)) - \sigma_{12}(x_2(t))]dt + \gamma(x_1(t) - x_2(t))dw(t), \quad x_2(0) \stackrel{\text{a.s.}}{=} x_{20}, \quad (55)$$

where $\sigma_{ij}(\cdot)$, $i, j = 1, 2$, $i \neq j$, are Lipschitz continuous and $\gamma > 0$. Equations (54) and (55) represent the collective dynamics of two subsystems which interact by exchanging energy. The energy states of the subsystems are described by the scalar random variables x_1 and x_2 . The unity coefficients scaling $\sigma_{ij}(\cdot)$, $i, j \in \{1, 2\}$, $i \neq j$, appearing in (54) and (55) represent the topology of the energy exchange between the subsystems. More specifically, given $i, j \in \{1, 2\}$, $i \neq j$, a coefficient of 1 denotes that subsystem j receives energy from subsystem i , and a coefficient of zero denotes that subsystem i and j are disconnected, and hence, cannot exchange energies.

The connectivity between the subsystems can be represented by a graph $\mathfrak{G}$ having two nodes such that $\mathfrak{G}$ has a directed edge from node i to node j if and only if subsystem j can receive energy from subsystem i . Since the coefficients scaling $\sigma_{ij}(\cdot)$, $i, j \in \{1, 2\}$, $i \neq j$, are constants, the graph topology is fixed. Furthermore, note that the directed graph $\mathfrak{G}$ is weakly connected since the underlying undirected graph is connected; that is, every subsystem receives energy from, or delivers energy to, at least one other subsystem.

Note that (54) and (55) can be cast in the form of (4) with

$$f(x) = \begin{bmatrix} \sigma_{12}(x_2) - \sigma_{21}(x_1) \\ \sigma_{21}(x_1) - \sigma_{12}(x_2) \end{bmatrix}, \quad D(x) = \begin{bmatrix} \gamma(x_2 - x_1) \\ \gamma(x_1 - x_2) \end{bmatrix},$$

where the stochastic term $D(x)dw$ represents probabilistic variations in the energy transfer between the two subsystems. Furthermore, note that since

$$e_2^T dx(t) = e_2^T f(x(t))dt + e_2^T D(x(t))dw(t) = 0, \quad x(0) \stackrel{\text{a.s.}}{=} x_0, \quad t \geq 0,$$

where $e_2 \triangleq [1 \ 1]^T$, it follows that $dx_1(t) + dx_2(t) = 0$, which implies that the total system energy is conserved.

In this example, we use Theorem 3 to analyze the collective behavior of (54) and (55). Specifically, we are interested in the energy equipartitioning behavior of the subsystems. For this purpose, we make the assumptions $\sigma_{ij}(x_j) - \sigma_{ji}(x_i) = 0$ if and only if $x_i = x_j$, $i \neq j$, and $(x_i - x_j)[\sigma_{ij}(x_j) - \sigma_{ji}(x_i)] \leq -\gamma^2(x_1 - x_2)$ for $i, j \in \{1, 2\}$.

The first assumption implies that if the energies in the connected subsystems i and j are equal, then energy exchange between the subsystems is not possible. This statement is reminiscent of the zeroth law of thermodynamics, which postulates that temperature equality is a necessary and sufficient condition for thermal equilibrium. The second assumption implies that energy flows from more energetic subsystems to less energetic subsystems and is reminiscent of the second law of thermodynamics, which states that heat (energy) must flow in the direction of lower temperatures. It is important to note here that due to the stochastic term $D(x)dw$ capturing probabilistic variations in the energy transfer between the subsystems, the second assumption requires that the scaled net energy flow $(x_i - x_j)[\sigma_{ij}(x_j) - \sigma_{ji}(x_i)]$ is bounded by the negative intensity of the diffusion coefficient given by $\frac{1}{2} \text{tr } D(x)D^T(x)$.

To show that (54) and (55) is stochastically semistable with respect to $\overline{\mathbb{R}}_+^2$, note that $\mathcal{E} \triangleq f^{-1}(0) \cap D^{-1}(0) = \{(x_1, x_2) \in \overline{\mathbb{R}}_+^2 : x_1 = x_2 = \alpha, \alpha \in \overline{\mathbb{R}}_+\}$ and consider the Lyapunov function candidate $V(x_1, x_2) = \frac{1}{2}(x_1 - \alpha)^2 + \frac{1}{2}(x_2 - \alpha)^2$, where $\alpha \in \overline{\mathbb{R}}_+$. Now, it follows that

$$\begin{aligned} \mathcal{L}V(x_1, x_2) &= (x_1 - \alpha)[\sigma_{12}(x_2) - \sigma_{21}(x_1)] + (x_2 - \alpha)[\sigma_{21}(x_1) - \sigma_{12}(x_2)] \\ &\quad + \frac{1}{2}[(\gamma(x_2 - x_1))^2 + (\gamma(x_1 - x_2))^2] \\ &= x_1[\sigma_{12}(x_2) - \sigma_{21}(x_1)] + x_2[\sigma_{21}(x_1) - \sigma_{12}(x_2)] + (\gamma(x_1 - x_2))^2 \\ &= (x_1 - x_2)[\sigma_{12}(x_2) - \sigma_{21}(x_1) + \gamma^2(x_1 - x_2)] \\ &\leq 0, \quad (x_1, x_2) \in \overline{\mathbb{R}}_+ \times \overline{\mathbb{R}}_+, \end{aligned} \quad (56)$$

which implies that $x_1 = x_2 = \alpha$ is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^2$.

Next, it is easy to see that $\mathcal{L}V(x_1, x_2) \neq 0$ when $x_1 \neq x_2$, and hence, $\mathcal{L}V(x_1, x_2) < 0$, $(x_1, x_2) \in \overline{\mathbb{R}}_+^2 \setminus \mathcal{E}$. Therefore, it follows from Theorem 3 that $x_1 = x_2 = \alpha$ is stochastically semistable with respect to $\overline{\mathbb{R}}_+^2$ for all $\alpha \in \overline{\mathbb{R}}_+$. Furthermore, note that $e_2^T dx(t) \stackrel{a.s.}{=} 0$, $t \geq 0$, implies

$$x(t) \stackrel{a.s.}{\rightarrow} \frac{1}{2}e_2e_2^T x(0) \stackrel{a.s.}{=} \frac{1}{2}[x_1(0) + x_2(0)]e_2 \quad \text{as } t \rightarrow \infty.$$

Note that an identical assertion holds for the collective dynamics of n subsystems with a connected undirected energy graph topology. $\triangle$

Finally, we extend Theorem 3.3 of [37] to provide a converse Lyapunov theorem for stochastic semistability. For this result, recall that $\mathcal{L}V(x_e) = 0$ for every $x_e \in \mathcal{E}$. Also note that it follows from (9) that $\mathcal{L}V(x) = \mathcal{L}V(s(0, x))$.

Theorem 5. Consider the nonlinear stochastic nonnegative dynamical system (4). Suppose (4) is stochastically semistable with a ρ -domain of semistability $\mathcal{D}_0$. Then there exist a continuous nonnegative function $V : \mathcal{D}_0 \rightarrow \overline{\mathbb{R}}_+$ and a class $\mathcal{K}_\infty$ function $\alpha(\cdot)$ such that (i) $V(x) = 0$, $x \in \mathcal{E}$, (ii) $V(x) \geq \alpha(\text{dist}(x, \mathcal{E}))$, $x \in \mathcal{D}_0$, and (iii) $\mathcal{L}V(x) < 0$, $x \in \mathcal{D}_0 \setminus \mathcal{E}$.

Proof. Let $\mathfrak{B}^{x_0}$ denote the set of all sample trajectories of (4) for which $\lim_{t \rightarrow \infty} \text{dist}(x(t, \omega), \mathcal{E}) = 0$ and $x(\{t \geq 0\}, \omega) \in \mathfrak{B}^{x_0}$, $\omega \in \Omega$, and let $\mathbb{1}_{\mathfrak{B}^{x_0}}(\omega)$, $\omega \in \Omega$, denote the indicator function defined on the set $\mathfrak{B}^{x_0}$, that is,

$$\mathbb{1}_{\mathfrak{B}^{x_0}}(\omega) \triangleq \begin{cases} 1, & \text{if } x(\{t \geq 0\}, \omega) \in \mathfrak{B}^{x_0}, \\ 0, & \text{otherwise.} \end{cases}$$

Note that by definition $\mathbb{P}^{x_0}(\mathfrak{B}^{x_0}) \geq 1 - \rho$ for all $x_0 \in \mathcal{D}_0$. Define the function $V : \mathcal{D}_0 \rightarrow \overline{\mathbb{R}}_+$ by

$$V(x) \triangleq \sup_{t \geq 0} \left\{ \frac{1 + 2t}{1 + t} \mathbb{E}[\text{dist}(s(t, x), \mathcal{E}) \mathbb{1}_{\mathfrak{B}^x}(\omega)] \right\}, \quad x \in \mathcal{D}_0, \quad (57)$$

and note that $V(\cdot)$ is well defined since (4) is stochastically semistable with respect to $\overline{\mathbb{R}}_+^n$. Clearly, (i) holds. Furthermore, since $V(x) \geq \text{dist}(x, \mathcal{E})$, $x \in \mathcal{D}_0$, it follows that (ii) holds with $a(r) = r$.

To show that $V(\cdot)$ is continuous on $\mathcal{D}_0 \setminus \mathcal{E}$, define $T : \mathcal{D}_0 \setminus \mathcal{E} \rightarrow [0, \infty)$ by $T(z) \triangleq \inf\{h : \mathbb{E}[\text{dist}(s(h, z), \mathcal{E}) \mathbb{1}_{\mathfrak{B}^z}(\omega)] < \text{dist}(z, \mathcal{E})/2 \text{ for all } t \geq h > 0\}$, and denote

$$\mathcal{W}_\varepsilon \triangleq \left\{ x \in \mathcal{D}_0 : \mathbb{P}^x \left(\sup_{t \geq 0} \text{dist}(s(t, x), \mathcal{E}) \leq \varepsilon \right) \geq 1 - \rho \right\}. \quad (58)$$

Note that $\mathcal{W}_\varepsilon \supset \mathcal{E}$ is open and contains an open neighborhood of $\mathcal{E}$. Consider $z \in \mathcal{D}_0 \setminus \mathcal{E}$ and define $\lambda \triangleq \text{dist}(z, \mathcal{E}) > 0$. Then it follows from stochastic semistability of (4) that there exists $h > 0$ such that $\mathbb{P}^z(s(h, z) \in \mathcal{W}_{\lambda/2}) \geq 1 - \rho$. Consequently, $\mathbb{P}^z(s(h + t, z) \in \mathcal{W}_{\lambda/2}) \geq 1 - \rho$ for all $t \geq 0$, and hence, it follows that $T(z)$ is well defined. Since $\mathcal{W}_{\lambda/2}$ is open, there exists a neighborhood $\mathcal{B}_\sigma(s(T(z), z))$ such that $\mathbb{P}^z(\mathcal{B}_\sigma(s(T(z), z)) \subset \mathcal{W}_{\lambda/2}) \geq 1 - \rho$. Hence, $\mathcal{N} \subset \mathcal{D}_0$ is a neighborhood of z such that $s_{T(z)}(\mathcal{H}_n^N) \triangleq \mathcal{B}_\sigma(s(T(z), z))$.

Next, choose $\eta > 0$ such that $\eta < \lambda/2$ and $\mathcal{B}_\eta(z) \subset \mathcal{N}$. Then, for every $t > T(z)$ and $y \in \mathcal{B}_\eta(z)$,

$$[(1 + 2t)/(1 + t)] \mathbb{E}[\text{dist}(s(t, y), \mathcal{E}) \mathbb{1}_{\mathfrak{B}^y}(\omega)] \leq 2 \mathbb{E}[\text{dist}(s(t, y), \mathcal{E}) \mathbb{1}_{\mathfrak{B}^y}(\omega)] \leq \lambda.$$

Therefore, for every $y \in \mathcal{B}_\eta(z)$,

$$\begin{aligned} V(z) - V(y) &= \sup_{t \geq 0} \left\{ \frac{1 + 2t}{1 + t} \mathbb{E}[\text{dist}(s(t, z), \mathcal{E}) \mathbb{1}_{\mathfrak{B}^z}(\omega)] \right\} \\ &\quad - \sup_{t \geq 0} \left\{ \frac{1 + 2t}{1 + t} \mathbb{E}[\text{dist}(s(t, y), \mathcal{E}) \mathbb{1}_{\mathfrak{B}^y}(\omega)] \right\} \\ &= \sup_{0 \leq t \leq T(z)} \left\{ \frac{1 + 2t}{1 + t} \mathbb{E}[\text{dist}(s(t, z), \mathcal{E}) \mathbb{1}_{\mathfrak{B}^z}(\omega)] \right\} \\ &\quad - \sup_{0 \leq t \leq T(z)} \left\{ \frac{1 + 2t}{1 + t} \mathbb{E}[\text{dist}(s(t, y), \mathcal{E}) \mathbb{1}_{\mathfrak{B}^y}(\omega)] \right\}. \end{aligned} \quad (59)$$

Hence,

$$|V(z) - V(y)| \leq \sup_{0 \leq t \leq T(z)} \left| \frac{1 + 2t}{1 + t} (\mathbb{E}[\text{dist}(s(t, z), \mathcal{E}) \mathbb{1}_{\mathfrak{B}^z}(\omega)] \right. \quad (60)$$

$$\begin{aligned} &\quad \left. - \mathbb{E}[\text{dist}(s(t, y), \mathcal{E}) \mathbb{1}_{\mathfrak{B}^y}(\omega)] \right| \\ &\leq 2 \sup_{0 \leq t \leq T(z)} |\mathbb{E}[\text{dist}(s(t, z), \mathcal{E}) \mathbb{1}_{\mathfrak{B}^z}(\omega)] - \mathbb{E}[\text{dist}(s(t, y), \mathcal{E}) \mathbb{1}_{\mathfrak{B}^y}(\omega)]| \\ &\leq 2 \sup_{0 \leq t \leq T(z)} \mathbb{E}[\text{dist}(s(t, z), s(t, y))], \quad z \in \mathcal{D}_0 \setminus \mathcal{E}, \quad y \in \mathcal{B}_\eta(z). \end{aligned} \quad (61)$$

Now, since $f(\cdot)$ and $D(\cdot)$ satisfy (6) and (7), it follows from continuous dependence of solutions $s(\cdot, \cdot)$ on system initial conditions ([32], Theorem 7.3.1) and (61) that $V(\cdot)$ is continuous on $\mathcal{D}_0 \setminus \mathcal{E}$.

To show that $V(\cdot)$ is continuous on $\mathcal{E}$, consider $x_e \in \mathcal{E}$. Let $\{x_n\}_{n=1}^\infty$ be a sequence in $\mathcal{D}_0 \setminus \mathcal{E}$ that converges to x_e . Since x_e is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$, it follows that $x(t) \xrightarrow{\text{a.s.}} x_e$ is the unique solution to (4) with $x(0) \xrightarrow{\text{a.s.}} x_e$. By continuous dependence of solutions $s(\cdot, \cdot)$ on system initial conditions ([32], Theorem 7.3.1), $s(t, x_n) \xrightarrow{\text{a.s.}} s(t, x_e) \xrightarrow{\text{a.s.}} x_e$ as $n \rightarrow \infty$, $t \geq 0$.

Let $\varepsilon > 0$ and note that it follows from (ii) of Proposition 3 that there exists $\delta = \delta(x_e) > 0$ such that for every solution of (4) in $\mathcal{B}_\delta(x_e)$ there exists $\hat{T} = \hat{T}(x_e, \varepsilon) > 0$ such that $\mathbb{P}(s_t(\mathcal{H}_n^{\mathcal{B}_\delta(x_e)}) \subset \mathcal{W}_\varepsilon) \geq 1 - \rho$

for all $t \geq \hat{T}$. Next, note that there exists a positive integer N_1 such that $x_n \in \mathcal{B}_\delta(x_e)$ for all $n \geq N_1$. Now, it follows from (57) that

$$V(x_n) \leq 2 \sup_{0 \leq t \leq \hat{T}} \mathbb{E}[\text{dist}(s(t, x_n), \mathcal{E}) \mathbb{1}_{\mathfrak{B}^{x_n}}(\omega)] + 2\varepsilon, \quad n \geq N_1. \quad (62)$$

Next, it follows from ([32], Theorem 7.3.1) that $\mathbb{E}[|s(\cdot, x_n)|]$ converges to $\mathbb{E}[|s(\cdot, x_e)|]$ uniformly on $[0, \hat{T}]$. Hence,

$$\begin{aligned} \lim_{n \rightarrow \infty} \sup_{0 \leq t \leq \hat{T}} \mathbb{E}[\text{dist}(s(t, x_n), \mathcal{E}) \mathbb{1}_{\mathfrak{B}^{x_n}}(\omega)] &= \sup_{0 \leq t \leq \hat{T}} \mathbb{E} \left[\lim_{n \rightarrow \infty} \text{dist}(s(t, x_n), \mathcal{E}) \mathbb{1}_{\mathfrak{B}^{x_n}}(\omega) \right] \\ &\leq \sup_{0 \leq t \leq \hat{T}} \text{dist}(x_e, \mathcal{E}) \\ &= 0, \end{aligned} \quad (63)$$

which implies that there exists a positive integer $N_2 = N_2(x_e, \varepsilon) \geq N_1$ such that

$$\sup_{0 \leq t \leq \hat{T}} \mathbb{E}[\text{dist}(s(t, x_n), \mathcal{E}) \mathbb{1}_{\mathfrak{B}^{x_n}}(\omega)] < \varepsilon$$

for all $n \geq N_2$. Combining (62) with the above result yields $V(x_n) < 4\varepsilon$ for all $n \geq N_2$, which implies that $\lim_{n \rightarrow \infty} V(x_n) = 0 = V(x_e)$.

Finally, we show that $\mathcal{L}V(x(t))$ is negative along the solution of (4) on $\mathcal{D}_0 \setminus \mathcal{E}$. Note that for every $x \in \mathcal{D}_0 \setminus \mathcal{E}$ and $0 < h \leq 1/2$ such that $\mathbb{P}(s(h, x) \in \mathcal{D}_0 \setminus \mathcal{E}) \geq 1 - \rho$, it follows from the definition of $T(\cdot)$ that $\mathbb{E}[V(s(h, x))]$ is reached at some time $\hat{t}$ such that $0 \leq \hat{t} \leq T(x)$. Hence, it follows from the law of iterated expectation that

$$\begin{aligned} \mathbb{E}[V(s(h, x))] &= \mathbb{E} \left[\mathbb{E}[\text{dist}(s(\hat{t} + h, x), \mathcal{E}) \mathbb{1}_{\mathfrak{B}^{s(h, x)}}(\omega)] \frac{1 + 2\hat{t}}{1 + \hat{t}} \right] \\ &= \mathbb{E}[\text{dist}(s(\hat{t} + h, x), \mathcal{E}) \mathbb{1}_{\mathfrak{B}^x}(\omega)] \frac{1 + 2\hat{t} + 2h}{1 + \hat{t} + h} \left[1 - \frac{h}{(1 + 2\hat{t} + 2h)(1 + \hat{t})} \right] \\ &\leq V(x) \left[1 - \frac{h}{2(1 + T(x))^2} \right], \end{aligned} \quad (64)$$

which implies that

$$\mathcal{L}V(x) = \lim_{h \rightarrow 0^+} \frac{\mathbb{E}[V(s(h, x))] - V(x)}{h} \leq -\frac{1}{2}V(x)(1 + T(x))^{-2} < 0, \quad x \in \mathcal{D}_0 \setminus \mathcal{E},$$

and hence, (iii) holds. $\square$

5. Conservation of Energy and the First Law of Thermodynamics: A Stochastic Perspective

In this section, we extend the thermodynamic model proposed in [30] to include probabilistic variations in the instantaneous rate of energy dissipation as well as probabilistic variations in the energy transfer between the subsystems. Even though the treatment in this and the next two sections closely parallels that of [30] for deterministic thermodynamics, the thermodynamic models and proofs of our results are rendered more difficult due to the inclusion of stochastic disturbances. To formulate our state space stochastic thermodynamic model, we consider the large-scale stochastic dynamical system $\mathcal{G}$ shown in Figure 1 involving energy exchange between q interconnected subsystems and use the notation developed in [30].

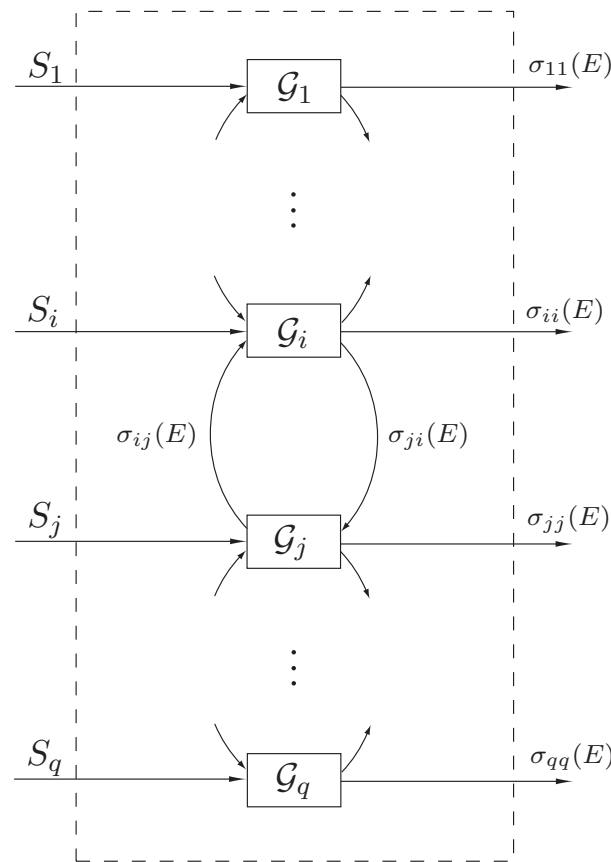


Figure 1. Large-scale dynamical system $\mathcal{G}$ with $D(E) = 0$ and $J(E) = 0$.

Specifically, $E_i : [0, \infty) \rightarrow \bar{\mathbb{R}}_+$ denotes the energy (and hence a nonnegative quantity) of the i -th subsystem, $S_i : [0, \infty) \rightarrow \mathbb{R}$ denotes the external power (heat flux) supplied to (or extracted from) the i -th subsystem, $\sigma_{ij} : \bar{\mathbb{R}}_+^q \rightarrow \bar{\mathbb{R}}_+, i \neq j, i, j = 1, \dots, q$, denotes the instantaneous rate of energy (heat) flow from the j -th subsystem to the i -th subsystem, $J_{(i,k)} : \bar{\mathbb{R}}_+^q \rightarrow \bar{\mathbb{R}}_+, i = 1, \dots, q, k = 1, \dots, d_1$, denotes the instantaneous rate of energy (heat) received or delivered to the i -th subsystem from all other subsystems due to the stochastic disturbance $w_{1k}(\cdot)$, $\sigma_{ii} : \bar{\mathbb{R}}_+^q \rightarrow \bar{\mathbb{R}}_+, i = 1, \dots, q$, denotes the instantaneous rate of energy (heat) dissipation from the i -th subsystem to the environment, and $D_{(i,l)} : \bar{\mathbb{R}}_+^q \rightarrow \bar{\mathbb{R}}_+, i = 1, \dots, q, l = 1, \dots, d_2$, denotes the instantaneous rate of energy (heat) dissipation from the i -th subsystem to the environment due to the stochastic disturbance $w_{2l}(\cdot)$. Here we assume that $\sigma_{ij} : \bar{\mathbb{R}}_+^q \rightarrow \bar{\mathbb{R}}_+, i, j = 1, \dots, q, J_{(i,k)} : \bar{\mathbb{R}}_+^q \rightarrow \bar{\mathbb{R}}_+, i = 1, \dots, q, k = 1, \dots, d_1$, and $D_{(i,l)} : \bar{\mathbb{R}}_+^q \rightarrow \bar{\mathbb{R}}_+, i = 1, \dots, q, l = 1, \dots, d_2$, are locally Lipschitz continuous on $\bar{\mathbb{R}}_+^q$ and satisfy a linear growth condition, and $S_i : [0, \infty) \rightarrow \mathbb{R}, i = 1, \dots, q$, are bounded piecewise continuous functions of time.

An energy balance for the i -th subsystem yields

$$\begin{aligned}
 E_i(T) = & E_i(t_0) + \sum_{j=1, j \neq i}^q \int_{t_0}^T [\sigma_{ij}(E(t)) - \sigma_{ji}(E(t))] dt \\
 & + \int_{t_0}^T \text{row}_i(J(E(t))) dw_1(t) - \int_{t_0}^T \sigma_{ii}(E(t)) dt \\
 & - \int_{t_0}^T \text{row}_i(D(E(t))) dw_2(t) + \int_{t_0}^T S_i(t) dt, \quad T \geq t_0,
 \end{aligned} \tag{65}$$

or, equivalently, in vector form,

$$\begin{aligned} E(T) = & E(t_0) + \int_{t_0}^T f(E(t))dt + \int_{t_0}^T J(E(t))dw_1(t) - \int_{t_0}^T d(E(t))dt \\ & - \int_{t_0}^T D(E(t))dw_2(t) + \int_{t_0}^T S(t)dt, \quad T \geq t_0, \end{aligned} \quad (66)$$

where $E(t) \triangleq [E_1(t), \dots, E_q(t)]^T$, $w_1(\cdot)$ and $w_2(\cdot)$ are, respectively, a d_1 -dimensional and d_2 -dimensional independent standard Wiener process (i.e., Brownian motion) defined on a complete filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq t_0}, \mathbb{P})$, $E(t_0)$ is independent of $(w_1(t) - w_1(t_0)), t \geq t_0$, and $(w_2(t) - w_2(t_0)), t \geq t_0$,

$$\begin{aligned} d(E(t)) &\triangleq [\sigma_{11}(E(t)), \dots, \sigma_{qq}(E(t))]^T, \\ S(t) &\triangleq [S_1(t), \dots, S_q(t)]^T, \\ f(E) &= [f_1(E), \dots, f_q(E)]^T : \mathbb{R}_+^q \rightarrow \mathbb{R}^q, \\ J(E) &= [\text{row}_1(J(E)), \dots, \text{row}_q(J(E))]^T : \mathbb{R}_+^q \rightarrow \mathbb{R}^q \times \mathbb{R}^{d_1}, \\ D(E) &= [\text{row}_1(D(E)), \dots, \text{row}_q(D(E))]^T : \mathbb{R}_+^q \rightarrow \mathbb{R}^q \times \mathbb{R}^{d_2}. \end{aligned}$$

Here, the stochastic disturbance $J(E)dw_1$ in (66) captures probabilistic variations in the energy transfer rates between compartments and the stochastic disturbance $D(E)dw_2$ captures probabilistic variations in the instantaneous rate of energy dissipation.

Equivalently, (65) can be rewritten as

$$\begin{aligned} dE_i(t) = & \sum_{j=1, j \neq i}^q [\sigma_{ij}(E(t)) - \sigma_{ji}(E(t))]dt + \text{row}_i(J(E(t)))dw_1(t) - \sigma_{ii}(E(t))dt \\ & - \text{row}_i(D(E(t)))dw_2(t) + S_i(t)dt, \quad E_i(t_0) \stackrel{\text{a.s.}}{=} E_{i0}, \quad t \geq t_0, \end{aligned} \quad (67)$$

or, in vector form,

$$dE(t) = f(E(t))dt + J(E(t))dw_1(t) - d(E(t))dt - D(E(t))dw_2(t) + S(t)dt, \quad E(t_0) \stackrel{\text{a.s.}}{=} E_0, \quad t \geq t_0, \quad (68)$$

where $E_0 \triangleq [E_{10}, \dots, E_{q0}]^T$, yielding a *differential energy balance* equation that characterizes energy flow between subsystems of the large-scale stochastic dynamical system $\mathcal{G}$. Here we assume that $S(\cdot)$ satisfies sufficient regularity conditions such that (68) has a unique solution forward in time. Specifically, we assume that the external power (heat flux) $S(\cdot)$ supplied to the large-scale stochastic dynamical system $\mathcal{G}$ consists of measurable functions $S(\cdot)$ adapted to the filtration $\{\mathcal{F}_t\}_{t \geq t_0}$ such that $S(t) \in \mathcal{H}_q$, $t \geq t_0$, for all $t \geq s$, $w(t) - w(s)$ is independent of $S(\tau)$, $w(\tau)$, $\tau \leq s$, and $E(t_0)$, where $w(t) \triangleq [w_1^T(t), w_2^T(t)]^T$, and hence, $S(\cdot)$ is non-anticipative. Furthermore, we assume that $S(\cdot)$ takes values in a compact metrizable set. In this case, it follows from Theorem 2.2.4 of [53] that there exists a path-wise unique solution to (68) in $(\Omega, \{\mathcal{F}_t\}_{t \geq t_0}, \mathbb{P}^{E_0})$.

Equation (66) or, equivalently, (68) is a statement of the *first law for stochastic thermodynamics* as applied to *isochoric transformations* (i.e., constant subsystem volume transformations) for each of the subsystems $\mathcal{G}_i$, $i = 1, \dots, q$. To see this, let the total energy in the large-scale stochastic dynamical system $\mathcal{G}$ be given by $U \triangleq \mathbf{e}^T E$, where $\mathbf{e}^T \triangleq [1, \dots, 1]$ and $E \in \mathbb{R}_+^q$, and let the net energy received by the large-scale dynamical system $\mathcal{G}$ over the time interval $[t_1, t_2]$ be given by

$$Q \triangleq \int_{t_1}^{t_2} \mathbf{e}^T [S(t) - d(E(t))]dt - \int_{t_1}^{t_2} \mathbf{e}^T D(E(t))dw_2(t), \quad (69)$$

where $E(t)$, $t \geq t_0$, is the solution to (68). Then, premultiplying (66) by $\mathbf{e}^T$ and using the fact that $\mathbf{e}^T f(E) \equiv 0$ and $\mathbf{e}^T J(E) \equiv 0$, it follows that

$$\Delta U = Q, \quad (70)$$

where $\Delta U \triangleq U(t_2) - U(t_1)$ denotes the variation in the total energy of the large-scale stochastic dynamical system $\mathcal{G}$ over the time interval $[t_1, t_2]$.

For our large-scale stochastic dynamical system model $\mathcal{G}$, we assume that $\sigma_{ij}(E) = 0$, $E \in \bar{\mathbb{R}}_+^q$, $\sigma_{jj}(E) = 0$, $E \in \bar{\mathbb{R}}_+^q$, $J_{(j,k)}(E) = 0$, $E \in \bar{\mathbb{R}}_+^q$, $k = 1, \dots, d_1$, and $D_{(j,l)}(E) = 0$, $E \in \bar{\mathbb{R}}_+^q$, $l = 1, \dots, d_2$, whenever $E_j = 0$, $j = 1, \dots, q$. In this case, $f(E) - d(E)$, $E \in \bar{\mathbb{R}}_+^q$, is essentially nonnegative. The above constraint implies that if the energy of the j -th subsystem of $\mathcal{G}$ is zero, then this subsystem cannot supply any energy to its surroundings nor dissipate energy to the environment. Moreover, we assume that $S_i(t) \geq 0$ whenever $E_i(t) = 0$, $t \geq t_0$, $i = 1, \dots, q$, which implies that when the energy of the i -th subsystem is zero, then no energy can be extracted from this subsystem.

The following proposition is needed for the main results of this paper.

Proposition 4. Consider the large-scale stochastic dynamical system $\mathcal{G}$ with differential energy balance equation given by (68). Suppose $\sigma_{ij}(E) = 0$, $E \in \bar{\mathbb{R}}_+^q$, $J_{(j,k)}(E) = 0$, $E \in \bar{\mathbb{R}}_+^q$, $k = 1, \dots, d_1$, and $D_{(j,l)}(E) = 0$, $E \in \bar{\mathbb{R}}_+^q$, $l = 1, \dots, d_2$, whenever $E_j = 0$, $j = 1, \dots, q$, and $S_i(t) \geq 0$ whenever $E_i(t) = 0$, $t \geq t_0$, $i = 1, \dots, q$. Then the solution $E(t)$, $t \geq t_0$, to (68) is nonnegative for all nonnegative initial conditions $E_0 \in \bar{\mathbb{R}}_+^q$.

Proof. First note that $f(E) - d(E)$, $E \in \bar{\mathbb{R}}_+^q$, is essentially nonnegative, $J_{(j,k)}(E) = 0$, $E \in \bar{\mathbb{R}}_+^q$, $k = 1, \dots, d_1$, and $D_{(j,l)}(E) = 0$, $E \in \bar{\mathbb{R}}_+^q$, $l = 1, \dots, d_2$, whenever $E_j = 0$, $j = 1, \dots, q$. Next, since $S_i(t) \geq 0$ whenever $E_i(t) = 0$, $t \geq t_0$, $i = 1, \dots, q$, it follows that $dE_i(t) \geq 0$ for all $t \geq t_0$ and $i = 1, \dots, q$ whenever $E_i(t) = 0$ and $E_j(t) \geq 0$ for all $j \neq i$ and $t \geq t_0$. This implies that for all nonnegative initial conditions $E_0 \in \bar{\mathbb{R}}_+^q$, every sample trajectory of $\mathcal{G}$ is directed towards the interior of the nonnegative orthant $\bar{\mathbb{R}}_+^q$ whenever $E_i(t) = 0$, $i = 1, \dots, q$, and hence, remains nonnegative almost surely for all $t \geq t_0$. $\square$

Next, premultiplying (66) by $\mathbf{e}^T$, using Proposition 4, and using the fact that $\mathbf{e}^T f(E) \equiv 0$ and $\mathbf{e}^T J(E) \equiv 0$, it follows that

$$\mathbf{e}^T E(T) = \mathbf{e}^T E(t_0) + \int_{t_0}^T \mathbf{e}^T S(t) dt - \int_{t_0}^T \mathbf{e}^T d(E(t)) dt - \int_{t_0}^T \mathbf{e}^T D(E(t)) dw_2(t), \quad T \geq t_0. \quad (71)$$

Now, for the large-scale stochastic dynamical system $\mathcal{G}$, define the input $u(t) \triangleq S(t)$ and the output $y(t) \triangleq d(E(t))$. Hence, it follows from (71) that for any two $\mathcal{F}_t$ -stopping times τ_1 and τ_2 such that $\tau_1 \geq \tau_2$ almost surely,

$$\begin{aligned} \mathbb{E} \left[\mathbf{e}^T E(\tau_2) | \mathcal{F}_{\tau_1} \right] &= \mathbf{e}^T E(\tau_1) + \mathbb{E} \left[\int_{\tau_1}^{\tau_2} \mathbf{e}^T S(t) dt | \mathcal{F}_{\tau_1} \right] - \mathbb{E} \left[\int_{\tau_1}^{\tau_2} \mathbf{e}^T d(E(t)) dt | \mathcal{F}_{\tau_1} \right] \\ &\quad - \mathbb{E} \left[\int_{\tau_1}^{\tau_2} \mathbf{e}^T D(E(t)) dw_2(t) | \mathcal{F}_{\tau_1} \right] \\ &= \mathbf{e}^T E(\tau_1) + \mathbb{E} \left[\int_{\tau_1}^{\tau_2} \left[\mathbf{e}^T S(t) - \mathbf{e}^T d(E(t)) \right] dt | \mathcal{F}_{\tau_1} \right]. \end{aligned} \quad (72)$$

Thus, the large-scale stochastic dynamical system $\mathcal{G}$ is stochastically lossless [54] with respect to the energy supply rate $r(u, y) \triangleq \mathbf{e}^T u - \mathbf{e}^T y$ and with the energy storage function $U(E) \triangleq \mathbf{e}^T E$, $E \in \bar{\mathbb{R}}_+^q$. In other words, the difference between the supplied system energy and the stored system energy is a martingale with respect to the differential energy balance system filtration.

The following lemma is required for our next result.

Lemma 2. Consider the large-scale stochastic dynamical system $\mathcal{G}$ with differential energy balance Equation (68). Then, for every equilibrium state $E_e \in \mathcal{H}_q^+$ and every $\varepsilon > 0$ and $\tau \xrightarrow{\text{a.s.}} 0$, there exist $S_e \in \mathcal{H}_q$, $\alpha > 0$, and $\tau \xrightarrow{\text{a.s.}} \hat{\tau} \xrightarrow{\text{a.s.}} 0$ such that, for every $\hat{E} \in \mathcal{H}_q^+$ with $\|\hat{E} - E_e\| \leq \alpha\tau$, there exists $S : \mathbb{R}_+ \rightarrow \mathcal{H}_q$ such that $\|S(t) - S_e\| \leq \varepsilon$, $t \in [0, \hat{\tau}]$, and $E(t) = E_e + \frac{(\hat{E} - E_e)}{\hat{\tau}}t$, $t \in [0, \hat{\tau}]$.

Proof. Note that with $S_e \triangleq d(E_e) - f(E_e)$, the state $E_e \in \mathbb{R}_+^q$ is an equilibrium state of (68). Let $\theta > 0$ and $\tau \xrightarrow{\text{a.s.}} 0$, and define

$$M(\theta, \tau) \triangleq \sup_{E \in \mathcal{H}_q^{\bar{B}_1(0)}, t \in [0, \tau]} \|f(E_e + \theta t E) - d(E_e + \theta t E) + S_e\|, \quad (73)$$

$$M_J(\theta, \tau) \triangleq \sup_{E \in \mathcal{H}_q^{\bar{B}_1(0)}, t \in [0, \tau]} \left\| J \left(E_e + \frac{(\hat{E} - E_e)}{\|\hat{E} - E_e\|} \alpha t \right) \right\|, \quad (74)$$

$$M_D(\theta, \tau) \triangleq \sup_{E \in \mathcal{H}_q^{\bar{B}_1(0)}, t \in [0, \tau]} \left\| D \left(E_e + \frac{(\hat{E} - E_e)}{\|\hat{E} - E_e\|} \alpha t \right) \right\|. \quad (75)$$

Note that for every $\tau \xrightarrow{\text{a.s.}} 0$, $\lim_{\theta \rightarrow 0^+} M(\theta, \tau) \xrightarrow{\text{a.s.}} 0$, $\lim_{\theta \rightarrow 0^+} M_J(\theta, \tau) \xrightarrow{\text{a.s.}} 0$, and $\lim_{\theta \rightarrow 0^+} M_D(\theta, \tau) \xrightarrow{\text{a.s.}} 0$, and for every $\theta > 0$, $\lim_{\tau \rightarrow 0^+} M(\theta, \tau) \xrightarrow{\text{a.s.}} 0$, $\lim_{\tau \rightarrow 0^+} M_J(\theta, \tau) \xrightarrow{\text{a.s.}} 0$, and $\lim_{\tau \rightarrow 0^+} M_D(\theta, \tau) \xrightarrow{\text{a.s.}} 0$. Moreover, it follows from Lévy's modulus of continuity theorem [55] that for sufficiently small $dt > 0$, $\|dw_1(t)\| \leq M_W(dt)dt$ and $\|dw_2(t)\| \leq M_W(dt)dt$, where $M_W(dt) \triangleq \sqrt{2dt \log_e \left(\frac{1}{dt} \right)}$.

Next, let $\varepsilon > 0$ and $\tau \xrightarrow{\text{a.s.}} 0$ be given and, for sufficiently small $dt > 0$, let $\alpha > 0$ be such that

$$M(\alpha, \tau) + \alpha + M_J(\alpha, \tau)M_W(dt) + M_D(\alpha, \tau)M_W(dt) \xrightarrow{\text{a.s.}} \varepsilon.$$

(The existence of such an α is guaranteed since $M(\alpha, \tau) \xrightarrow{\text{a.s.}} 0$, $M_J(\alpha, \tau) \xrightarrow{\text{a.s.}} 0$, and $M_D(\alpha, \tau) \xrightarrow{\text{a.s.}} 0$ as $\alpha \rightarrow 0^+$.) Now, let $\hat{E} \in \mathcal{H}_q^+$ be such that $\|\hat{E} - E_e\| \leq \alpha\tau$. With $\hat{\tau} \triangleq \frac{\|\hat{E} - E_e\|}{\alpha} \leq \tau$ and

$$S(t)dt = \left[-f(E(t)) + d(E(t)) + \alpha \frac{(\hat{E} - E_e)}{\|\hat{E} - E_e\|} \right] dt - J(E(t))dw_1(t) + D(E(t))dw_2(t), \quad t \in [0, \hat{\tau}], \quad (76)$$

it follows that

$$E(t) = E_e + \frac{(\hat{E} - E_e)}{\|\hat{E} - E_e\|} \alpha t, \quad t \in [0, \hat{\tau}], \quad (77)$$

is a solution to (68).

The result is now immediate by noting that $E(\hat{\tau}) \xrightarrow{\text{a.s.}} \hat{E}$ and

$$\begin{aligned} \|S(t) - S_e\|dt &\leq \left\| f \left(E_e + \frac{(\hat{E} - E_e)}{\|\hat{E} - E_e\|} \alpha t \right) - d \left(E_e + \frac{(\hat{E} - E_e)}{\|\hat{E} - E_e\|} \alpha t \right) + S_e \right\| dt \\ &\quad + \alpha dt + \left\| J \left(E_e + \frac{(\hat{E} - E_e)}{\|\hat{E} - E_e\|} \alpha t \right) \right\| |dw_1(t)| \\ &\quad + \left\| D \left(E_e + \frac{(\hat{E} - E_e)}{\|\hat{E} - E_e\|} \alpha t \right) \right\| |dw_2(t)| \\ &\leq [M(\alpha, \tau) + \alpha + M_J(\alpha, \tau)M_W(dt) + M_D(\alpha, \tau)M_W(dt)]dt, \quad t \in [0, \hat{\tau}], \end{aligned} \quad (78)$$

and hence,

$$\|S(t) - S_e\| \stackrel{\text{a.s.}}{\leq} M(\alpha, \tau) + \alpha + M_J(\alpha, \tau)M_W(dt) + M_D(\alpha, \tau)M_W(dt) \stackrel{\text{a.s.}}{\leq} \varepsilon,$$

which proves the result. $\square$

It follows from Lemma 2 that the large-scale stochastic dynamical system $\mathcal{G}$ with the differential energy balance Equation (68) is *stochastically reachable* from and *stochastically controllable* to the origin in $\mathbb{R}_+^q$. Recall from [54] that the large-scale stochastic dynamical system $\mathcal{G}$ with the differential energy balance Equation (68) is stochastically reachable from the origin in $\mathbb{R}_+^q$ if, for all $E_0 \in \mathbb{R}_+^q$ and $\varepsilon > 0$, there exist a finite random variable $\tau_{\mathcal{B}_\varepsilon(E_0)} \stackrel{\text{a.s.}}{\geq} t_0$, called the *first hitting time*, defined by

$$\tau_{\mathcal{B}_\varepsilon(E_0)}(\omega) \triangleq \inf\{t \geq t_0 : E(t, \omega) \in \mathcal{B}_\varepsilon(E_0)\},$$

and a $\mathcal{F}_t$ -adapted square integrable input $S(\cdot)$ defined on $[t_0, \tau_{\mathcal{B}_\varepsilon(E_0)}]$ such that the state $E(t)$, $t \geq t_0$, can be driven from $E(t_0) \stackrel{\text{a.s.}}{=} 0$ to $E(\tau_{\mathcal{B}_\varepsilon(E_0)})$ and $\mathbb{E}[\tau_{E_0}] < \infty$, where $\tau_{E_0} \triangleq \sup_{\varepsilon > 0} \tau_{\mathcal{B}_\varepsilon(E_0)}$ and the supremum is taken pointwise. Alternatively, $\mathcal{G}$ is stochastically controllable to the origin in $\mathbb{R}_+^q$ if, for all $E(t_0) \stackrel{\text{a.s.}}{=} E_0$, $E_0 \in \mathbb{R}_+^q$, there exists a finite random variable $\tilde{\tau}_{\mathcal{B}_\varepsilon(E_0)} \stackrel{\text{a.s.}}{\geq} t_0$ defined by

$$\tilde{\tau}_{\mathcal{B}_\varepsilon(E_0)}(\omega) \triangleq \inf\{t \geq t_0 : E(t, \omega) \in \mathcal{B}_\varepsilon(0)\},$$

and a $\mathcal{F}_t$ -adapted square integrable input $S(\cdot)$ defined on $[t_0, \tilde{\tau}_{\mathcal{B}_\varepsilon(E_0)}]$ such that the state $E(t)$, $t \geq t_0$, can be driven from $E(t_0) \stackrel{\text{a.s.}}{=} E_0$ to $E(\tilde{\tau}_{\mathcal{B}_\varepsilon(E_0)}) \in \mathcal{B}_\varepsilon(0)$ and $\tilde{\tau}_{E_0} \triangleq \sup_{\varepsilon > 0} \tilde{\tau}_{\mathcal{B}_\varepsilon(E_0)}$ with a pointwise supremum.

We let $\mathcal{U}_r$ denote the set of measurable bounded $\mathcal{H}_q^+$ -valued stochastic processes on the semi-infinite interval $[t_0, \infty)$ consisting of power inputs (heat fluxes) to the large-scale stochastic dynamical system $\mathcal{G}$ such that for every $\tau_{E_0} \stackrel{\text{a.s.}}{\geq} t_0$ the system energy state can be driven from $E(t_0) \stackrel{\text{a.s.}}{=} 0$ to $E(\tau_{E_0})$ by $S(\cdot) \in \mathcal{U}_r$. Furthermore, we let $\mathcal{U}_c$ denote the set of measurable bounded $\mathcal{H}_q^+$ -valued stochastic processes on the semi-infinite interval $[t_0, \infty)$ consisting of power inputs (heat fluxes) to the large-scale stochastic dynamical system $\mathcal{G}$ such that the system energy state can be driven from $E(t_0) \stackrel{\text{a.s.}}{=} E_0$, $E_0 \in \mathbb{R}_+^q$ to $E(\tilde{\tau}_{E_0})$ by $S(\cdot) \in \mathcal{U}_c$. Finally, let $\mathcal{U}$ be an input space that is a subset of measurable bounded $\mathcal{H}_q^+$ -valued stochastic processes on $\mathbb{R}$. The spaces $\mathcal{U}_r$, $\mathcal{U}_c$, and $\mathcal{U}$ are assumed to be closed under the shift operator, that is, if $S(\cdot) \in \mathcal{U}$ (respectively, $\mathcal{U}_c$ or $\mathcal{U}_r$), then the function S_T defined by $S_T(t) \triangleq S(t + T)$ is contained in $\mathcal{U}$ (respectively, $\mathcal{U}_c$ or $\mathcal{U}_r$) for all $T \geq 0$.

The next result establishes the uniqueness of the internal energy function $U(E)$, $E \in \mathbb{R}_+^q$, for our large-scale stochastic dynamical system $\mathcal{G}$. For this result define the *available energy* of the large-scale stochastic dynamical system $\mathcal{G}$ by

$$U_a(E_0) \triangleq - \inf_{u(\cdot) \in \mathcal{U}, \tau \stackrel{\text{a.s.}}{\geq} t_0} \mathbb{E} \left[\mathbb{E} \left[\int_{t_0}^{\tau} [\mathbf{e}^T u(t) - \mathbf{e}^T y(t)] dt | E(t_0) \stackrel{\text{a.s.}}{=} E_0 \right] \right], \quad E_0 \in \mathbb{R}_+^q, \quad (79)$$

where $E(t)$, $t \geq t_0$, is the solution to (68) with $E(t) \stackrel{\text{a.s.}}{=} E_0$ and admissible inputs $S(\cdot) \in \mathcal{U}$. The infimum in (79) is taken over all $\mathcal{F}_t$ -measurable inputs $S(\cdot)$, all finite $\mathcal{F}_t$ -stopping times $\tau \stackrel{\text{a.s.}}{\geq} 0$, and all system sample paths with initial value $E(t_0) \stackrel{\text{a.s.}}{=} E_0$ and terminal value left free. Furthermore, define the *required energy supply* of the large-scale stochastic dynamical system $\mathcal{G}$ by

$$U_r(E_0) \triangleq \inf_{u(\cdot) \in \mathcal{U}_r, \tau_{E_0} \stackrel{\text{a.s.}}{\geq} 0} \mathbb{E} \left[\mathbb{E} \left[\int_0^{\tau_{E_0}} [\mathbf{e}^T u(t) - \mathbf{e}^T y(t)] dt | E(0) \stackrel{\text{a.s.}}{=} 0 \right] \right], \quad E_0 \in \mathbb{R}_+^q. \quad (80)$$

The infimum in (80) is taken over all system sample paths starting from $E(t_0) \stackrel{\text{a.s.}}{=} 0$ and ending at $E(\tau_{E_0}) \stackrel{\text{a.s.}}{=} E_0$ at time $t = \tau_{E_0}$, and all times $t \geq t_0$.

Note that the available energy $U_a(E)$ is the maximum amount of stored energy (net heat) that can be extracted from the large-scale stochastic dynamical system $\mathcal{G}$ at any finite stopping time τ , and the required energy supply $U_r(E)$ is the minimum amount of energy (net heat) that can be delivered to the large-scale stochastic dynamical system $\mathcal{G}$ such that, for all $\varepsilon > 0$, $\mathbb{P}^0 \left(\lim_{t \rightarrow \tau_{E_0}} E(t) \in \mathcal{B}_\varepsilon(E_0) \right) = 1$.

Theorem 6. Consider the large-scale stochastic dynamical system $\mathcal{G}$ with differential energy balance equation given by (68). Then $\mathcal{G}$ is stochastically lossless with respect to the energy supply rate $r(u, y) = \mathbf{e}^T u - \mathbf{e}^T y$, where $u(t) \equiv S(t)$ and $y(t) \equiv d(E(t))$, and with the unique energy storage function corresponding to the total energy of the system $\mathcal{G}$ given by

$$\begin{aligned} U(E_0) &= \mathbf{e}^T E_0 \\ &= -\mathbb{E} \left[\mathbb{E} \left[\int_0^{\tau_0} [\mathbf{e}^T u(t) - \mathbf{e}^T y(t)] dt \middle| E(0) \stackrel{\text{a.s.}}{=} E_0 \right] \right] \\ &= \mathbb{E} \left[\mathbb{E} \left[\int_0^{\tau_{E_0}} [\mathbf{e}^T u(t) - \mathbf{e}^T y(t)] dt \middle| E(0) \stackrel{\text{a.s.}}{=} 0 \right] \right], \quad E_0 \in \overline{\mathbb{R}}_+^q, \end{aligned} \quad (81)$$

where $E(t)$, $t \geq t_0$, is the solution to (68) with admissible input $u(\cdot) \in \mathcal{U}$, $E(\tau_0) \stackrel{\text{a.s.}}{=} 0$, and $E(\tau_{E_0}) \stackrel{\text{a.s.}}{=} E_0 \in \overline{\mathbb{R}}_+^q$. Furthermore,

$$0 \leq U_a(E_0) = U(E_0) = U_r(E_0) < \infty, \quad E_0 \in \overline{\mathbb{R}}_+^q. \quad (82)$$

Proof. Note that it follows from (71) that $\mathcal{G}$ is stochastically lossless with respect to the energy supply rate $r(u, y) = \mathbf{e}^T u - \mathbf{e}^T y$ and with the energy storage function $U(E) = \mathbf{e}^T E$, $E \in \overline{\mathbb{R}}_+^q$. Since, by Lemma 2, $\mathcal{G}$ is reachable from and controllable to the origin in $\overline{\mathbb{R}}_+^q$, it follows from (71), with $E(t_0) \stackrel{\text{a.s.}}{=} E_0 \in \overline{\mathbb{R}}_+^q$ and $E(\tau_+) \stackrel{\text{a.s.}}{=} 0$ for some $\tau_+ \geq t_0$ and $u(\cdot) \in \mathcal{U}$, that

$$\begin{aligned} \mathbf{e}^T E_0 &= -\mathbb{E} \left[\mathbb{E} \left[\int_{t_0}^{\tau_+} [\mathbf{e}^T u(t) - \mathbf{e}^T y(t)] dt \middle| E(t_0) \stackrel{\text{a.s.}}{=} E_0 \right] \right] \\ &\leq \sup_{u(\cdot) \in \mathcal{U}, \tau_+ \geq t_0} -\mathbb{E} \left[\mathbb{E} \left[\int_{t_0}^{\tau_+} [\mathbf{e}^T u(t) - \mathbf{e}^T y(t)] dt \middle| E(t_0) \stackrel{\text{a.s.}}{=} E_0 \right] \right] \\ &= -\inf_{u(\cdot) \in \mathcal{U}, \tau_+ \geq t_0} \mathbb{E} \left[\mathbb{E} \left[\int_{t_0}^{\tau_+} [\mathbf{e}^T u(t) - \mathbf{e}^T y(t)] dt \middle| E(t_0) \stackrel{\text{a.s.}}{=} E_0 \right] \right] \\ &= U_a(E_0), \quad E_0 \in \overline{\mathbb{R}}_+^q. \end{aligned} \quad (83)$$

Alternatively, it follows from (71), with $E(0) \stackrel{\text{a.s.}}{=} 0$ for some $\tau_- \geq 0$ and $u(\cdot) \in \mathcal{U}_r$, that

$$\begin{aligned} \mathbf{e}^T E_0 &= \mathbb{E} \left[\mathbb{E} \left[\int_0^{\tau_-} [\mathbf{e}^T u(t) - \mathbf{e}^T y(t)] dt \middle| E(0) \stackrel{\text{a.s.}}{=} 0 \right] \right] \\ &\geq \inf_{u(\cdot) \in \mathcal{U}_r, \tau_- \geq 0} \mathbb{E} \left[\mathbb{E} \left[\int_0^{\tau_-} [\mathbf{e}^T u(t) - \mathbf{e}^T y(t)] dt \middle| E(0) \stackrel{\text{a.s.}}{=} 0 \right] \right] \\ &= U_r(E_0), \quad E_0 \in \overline{\mathbb{R}}_+^q. \end{aligned} \quad (84)$$

Thus, (83) and (84) imply that (81) is satisfied and

$$U_r(E_0) \leq \mathbf{e}^T E_0 \leq U_a(E_0), \quad E_0 \in \overline{\mathbb{R}}_+^q. \quad (85)$$

Conversely, it follows from (71) and the fact that $U(E) = \mathbf{e}^T E \geq 0$, $E \in \overline{\mathbb{R}}_+^q$, that, for all $\tau \stackrel{\text{a.s.}}{\geq} t_0$ and $u(\cdot) \in \mathcal{U}$,

$$\mathbf{e}^T E_0 \geq -\mathbb{E} \left[\mathbb{E} \left[\int_{t_0}^{\tau} [\mathbf{e}^T u(t) - \mathbf{e}^T y(t)] dt | E(t_0) \stackrel{\text{a.s.}}{=} E_0 \right] \right], \quad E_0 \in \overline{\mathbb{R}}_+^q, \quad (86)$$

which implies that

$$\begin{aligned} \mathbf{e}^T E(t_0) &\geq \sup_{u(\cdot) \in \mathcal{U}, \tau \stackrel{\text{a.s.}}{\geq} t_0} -\mathbb{E} \left[\mathbb{E} \left[\int_{t_0}^{\tau} [\mathbf{e}^T u(t) - \mathbf{e}^T y(t)] dt | E(t_0) \stackrel{\text{a.s.}}{=} E_0 \right] \right] \\ &= - \inf_{u(\cdot) \in \mathcal{U}, \tau \stackrel{\text{a.s.}}{\geq} t_0} \mathbb{E} \left[\mathbb{E} \left[\int_{t_0}^{\tau} [\mathbf{e}^T u(t) - \mathbf{e}^T y(t)] dt | E(t_0) \stackrel{\text{a.s.}}{=} E_0 \right] \right] \\ &= U_a(E_0), \quad E_0 \in \overline{\mathbb{R}}_+^q. \end{aligned} \quad (87)$$

Furthermore, it follows from the definition of $U_a(\cdot)$ that $U_a(E) \geq 0$, $E \in \overline{\mathbb{R}}_+^q$, since the infimum in (79) is taken over the set of values containing the zero value ($\tau \stackrel{\text{a.s.}}{=} t_0$).

Next, note that it follows from (71), with $E(0) \stackrel{\text{a.s.}}{=} 0$ and $E(\tau) \stackrel{\text{a.s.}}{=} E_0$, $E_0 \in \overline{\mathbb{R}}_+^q$, for all $\tau \stackrel{\text{a.s.}}{\geq} 0$ and $u(\cdot) \in \mathcal{U}_r$, that

$$\begin{aligned} \mathbf{e}^T E_0 &= \mathbb{E} \left[\mathbb{E} \left[\int_0^{\tau} [\mathbf{e}^T u(t) - \mathbf{e}^T y(t)] dt | E(0) \stackrel{\text{a.s.}}{=} 0 \right] \right] \\ &= \inf_{u(\cdot) \in \mathcal{U}_r, \tau \stackrel{\text{a.s.}}{\geq} 0} \mathbb{E} \left[\mathbb{E} \left[\int_0^{\tau} [\mathbf{e}^T u(t) - \mathbf{e}^T y(t)] dt | E(0) \stackrel{\text{a.s.}}{=} 0 \right] \right] \\ &= U_r(E_0), \quad E_0 \in \overline{\mathbb{R}}_+^q. \end{aligned} \quad (88)$$

Moreover, since the system $\mathcal{G}$ is reachable from the origin, it follows that for every $E_0 \in \overline{\mathbb{R}}_+^q$, there exists $\tau \stackrel{\text{a.s.}}{\geq} 0$ and $u(\cdot) \in \mathcal{U}_r$ such that

$$\mathbb{E} \left[\mathbb{E} \left[\int_0^{\tau} [\mathbf{e}^T u(t) - \mathbf{e}^T y(t)] dt | E(0) \stackrel{\text{a.s.}}{=} 0 \right] \right] \quad (89)$$

is finite, and hence, $U_r(E_0) < \infty$, $E_0 \in \overline{\mathbb{R}}_+^q$. Finally, combining (85), (87), and (88), it follows that (82) holds. $\square$

It follows from (82) and the definitions of available energy $U_a(E_0)$ and the required energy supply $U_r(E_0)$, $E_0 \in \overline{\mathbb{R}}_+^q$, that the large-scale stochastic dynamical system $\mathcal{G}$ can deliver to its surroundings all of its stored subsystem energies and can store all of the work done to all of its subsystems. This is in essence a statement of the first law of stochastic thermodynamics and places no limitation on the possibility of transforming heat into work or work into heat. In the case where $S(t) \equiv 0$, it follows from (71) and the fact that $\sigma_{ii}(E) \geq 0$, $E \in \overline{\mathbb{R}}_+^q$, $i = 1, \dots, q$, that the zero solution $E(t) \equiv 0$ of the large-scale stochastic dynamical system $\mathcal{G}$ with the differential energy balance Equation (68) is Lyapunov stable in probability with respect to $\overline{\mathbb{R}}_+^n$ with Lyapunov function $U(E)$ corresponding to the total energy in the system.

6. Entropy and the Second Law of Thermodynamics

As for the deterministic dynamical thermodynamic model presented in [30], the nonlinear differential energy balance Equation (68) can exhibit a full range of nonlinear behavior, including bifurcations, limit cycles, and even chaos. However, a thermodynamically consistent energy flow model should ensure that the evolution of the system energy is diffusive in character with convergent subsystem energies. As established in [30], such a system model would guarantee the absence of the

Poincaré recurrence phenomenon [56]. To ensure a thermodynamically consistent energy flow model, we require the following axioms [57]. For the statement of these axioms, we first recall the following graph-theoretic notions.

Definition 10 ([58]). A directed graph $\mathfrak{G}(\mathcal{C})$ associated with the connectivity matrix $\mathcal{C} \in \mathbb{R}^{q \times q}$ has vertices $\{1, 2, \dots, q\}$ and an arc from vertex i to vertex j , $i \neq j$, if and only if $\mathcal{C}_{(j,i)} \neq 0$. A graph $\mathfrak{G}(\mathcal{C})$ associated with the connectivity matrix $\mathcal{C} \in \mathbb{R}^{q \times q}$ is a directed graph for which the arc set is symmetric, that is, $\mathcal{C} = \mathcal{C}^T$. We say that $\mathfrak{G}(\mathcal{C})$ is strongly connected if for any ordered pair of vertices (i, j) , $i \neq j$, there exists a path (i.e., a sequence of arcs) leading from i to j .

Recall that the connectivity matrix $\mathcal{C} \in \mathbb{R}^{q \times q}$ is irreducible, that is, there does not exist a permutation matrix such that $\mathcal{C}$ is cogredient to a lower-block triangular matrix, if and only if $\mathfrak{G}(\mathcal{C})$ is strongly connected (see Theorem 2.7 of [58]). Let $\phi_{ij}(E) \triangleq \sigma_{ij}(E) - \sigma_{ji}(E)$, $E \in \mathbb{R}_+^q$, denote the net energy flow from the j -th subsystem $\mathcal{G}_j$ to the i -th subsystem $\mathcal{G}_i$ of the large-scale stochastic dynamical system $\mathcal{G}$.

Axiom (i): For the connectivity matrix $\mathcal{C} \in \mathbb{R}^{q \times q}$ associated with the large-scale stochastic dynamical system $\mathcal{G}$ defined by

$$\mathcal{C}_{(i,j)} \triangleq \begin{cases} 0, & \text{if } \phi_{ij}(E) \equiv 0, \\ 1, & \text{otherwise,} \end{cases} \quad i \neq j, \quad i, j = 1, \dots, q, \quad (90)$$

and

$$\mathcal{C}_{(i,i)} \triangleq - \sum_{k=1, k \neq i}^q \mathcal{C}_{(k,i)}, \quad i = j, \quad i = 1, \dots, q, \quad (91)$$

$\text{rank } \mathcal{C} = q - 1$, and for $\mathcal{C}_{(i,j)} = 1$, $i \neq j$, $\phi_{ij}(E) = 0$ if and only if $E_i = E_j$.

Axiom (ii): For $i, j = 1, \dots, q$, $(E_i - E_j)\phi_{ij}(E) \stackrel{\text{a.s.}}{\leq} 0$, $E \in \mathbb{R}_+^q$, and, for all $c > 0$,

$$\sum_{j=1, j \neq i}^q (E_i - E_j)\phi_{ij}(E) \frac{c + E_i}{c + E_j} \leq -\text{row}_i(J(E))\text{row}_i^T(J(E)), \quad i = 1, \dots, q.$$

As discussed in [30] for the deterministic thermodynamic problem, the fact that $\phi_{ij}(E) = 0$ if and only if $E_i = E_j$, $i \neq j$, implies that subsystems $\mathcal{G}_i$ and $\mathcal{G}_j$ of $\mathcal{G}$ are *connected*; alternatively, $\phi_{ij}(E) \equiv 0$ implies that $\mathcal{G}_i$ and $\mathcal{G}_j$ are *disconnected*. Axiom (i) implies that if the energies in the connected subsystems $\mathcal{G}_i$ and $\mathcal{G}_j$ are equal, then energy exchange between these subsystems is not possible. This statement is consistent with the *zeroth law of thermodynamics*, which postulates that temperature equality is a necessary and sufficient condition for thermal equilibrium. Furthermore, it follows from the fact that $\mathcal{C} = \mathcal{C}^T$ and $\text{rank } \mathcal{C} = q - 1$ that the connectivity matrix $\mathcal{C}$ is irreducible, which implies that for any pair of subsystems $\mathcal{G}_i$ and $\mathcal{G}_j$, $i \neq j$, of $\mathcal{G}$ there exists a sequence of connectors (arcs) of $\mathcal{G}$ that connect $\mathcal{G}_i$ and $\mathcal{G}_j$.

Axiom (ii) implies that energy flows from more energetic subsystems to less energetic subsystems and is consistent with the *second law of thermodynamics*, which states that heat (energy) must flow in the direction of lower temperatures [59]. Furthermore, note that $\phi_{ij}(E) = -\phi_{ji}(E)$, $E \in \mathbb{R}_+^q$, $i \neq j$, $i, j = 1, \dots, q$, which implies conservation of energy between lossless subsystems. With $S(t) \equiv 0$ and $J(E) = 0$, Axioms (i) and (ii) along with the fact that $\phi_{ij}(E) = -\phi_{ji}(E)$, $E \in \mathbb{R}_+^q$, $i \neq j$, $i, j = 1, \dots, q$, imply that at a given instant of time, energy can only be transported, stored, or dissipated but not created, and the maximum amount of energy that can be transported and/or dissipated from a subsystem cannot exceed the energy in the subsystem. Finally, it is important to note here that due to the stochastic disturbance term $J(E)dw_1$ capturing probabilistic variations in heat transfer between the

subsystems, Axiom (ii) requires that the scaled net energy flow between the subsystems is bounded by the negative intensity of the system diffusion.

Next, we show that the classical Clausius equality and inequality for reversible and irreversible thermodynamics over cyclic motions are satisfied for our stochastic thermodynamically consistent energy flow model. For this result $\oint$ denotes a cyclic integral evaluated along an arbitrary closed path of (68) in $\overline{\mathbb{R}}_+^q$; that is, $\oint \triangleq \int_{t_0}^{\tau_f}$ with $\tau_f \geq t_0$ and $S(\cdot) \in \mathcal{U}$ such that $E(\tau_f) \stackrel{\text{a.s.}}{=} E(t_0) \stackrel{\text{a.s.}}{=} E_0 \in \overline{\mathbb{R}}_+^q$.

Proposition 5. Consider the large-scale stochastic dynamical system $\mathcal{G}$ with differential energy balance Equation (68), and assume that Axioms (i) and (ii) hold. Then, for all $E_0 \in \overline{\mathbb{R}}_+^q$, $\tau_f \geq t_0$, and $S(\cdot) \in \mathcal{U}$ such that $E(\tau_f) \stackrel{\text{a.s.}}{=} E(t_0) \stackrel{\text{a.s.}}{=} E_0$,

$$\begin{aligned} \mathbb{E}^{E_0} \left[\int_{t_0}^{\tau_f} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} dt \right] \right] &= \mathbb{E}^{E_0} \left[\oint \sum_{i=1}^q \frac{dQ_i(t)}{c + E_i(t)} \right] \\ &\leq \mathbb{E}^{E_0} \left[\int_{t_0}^{\tau_f} \sum_{i=1}^q \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} dt \right], \end{aligned} \quad (92)$$

where $c > 0$, $dQ_i(t) \triangleq [S_i(t) - \sigma_{ii}(E(t))]dt$, $i = 1, \dots, q$, is the amount of net energy (heat) received by the i -th subsystem over the infinitesimal time interval dt , and $E(t)$, $t \geq t_0$, is the solution to (68) with initial condition $E(t_0) \stackrel{\text{a.s.}}{=} E_0$. Furthermore,

$$\mathbb{E}^{E_0} \left[\oint \sum_{i=1}^q \frac{dQ_i(t)}{c + E_i(t)} \right] = \mathbb{E}^{E_0} \left[\int_{t_0}^{\tau_f} \sum_{i=1}^q \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} dt \right] \quad (93)$$

if and only if there exists a continuous function $\alpha : [t_0, t_f] \rightarrow \overline{\mathbb{R}}_+$ such that $E(t) = \alpha(t)\mathbf{e}$, $t \in [t_0, t_f]$.

Proof. Since, by Proposition 4, $E(t) \geq 0$, $t \geq t_0$, and $\phi_{ij}(E) = -\phi_{ji}(E)$, $E \in \overline{\mathbb{R}}_+^q$, $i \neq j$, $i, j = 1, \dots, q$, it follows from (68), Ito's lemma, and Axiom (ii) that, for all $\tau_f \stackrel{\text{a.s.}}{\geq} t_0$,

$$\begin{aligned} \mathbb{E}^{E_0} \left[\oint \sum_{i=1}^q \frac{dQ_i(t)}{c + E_i(t)} \right] &= \mathbb{E}^{E_0} \left[\int_{t_0}^{\tau_f} \sum_{i=1}^q \frac{dE_i(t) - \sum_{j=1, j \neq i}^q \phi_{ij}(E(t)) dt}{c + E_i(t)} \right] \\ &= \mathbb{E}^{E_0} \left[\int_{t_0}^{\tau_f} \sum_{i=1}^q \frac{dE_i(t)}{c + E_i(t)} \right] - \mathbb{E}^{E_0} \left[\int_{t_0}^{\tau_f} \sum_{i=1}^q \frac{\sum_{j=1, j \neq i}^q \phi_{ij}(E(t)) dt}{c + E_i(t)} \right] \\ &= \mathbb{E}^{E_0} \left[\int_{t_0}^{\tau_f} \sum_{i=1}^q \left[d \log_e (c + E_i(t)) + \frac{1}{2} \frac{\text{row}_i(J(E(t))) \text{row}_i^T(J(E(t)))}{(c + E_i(t))^2} dt \right. \right. \\ &\quad \left. \left. + \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} dt \right] \right] \\ &\quad - \mathbb{E}^{E_0} \left[\int_{t_0}^{\tau_f} \sum_{i=1}^q \sum_{j=1, j \neq i}^q \frac{1}{2} \left(\frac{\phi_{ij}(E(t))}{c + E_i(t)} - \frac{\phi_{ij}(E(t))}{c + E_j(t)} \right) dt \right] \\ &= \mathbb{E}^{E_0} \left[\sum_{i=1}^q \log_e \left(\frac{c + E_i(\tau_f)}{c + E_i(t_0)} \right) \right] \\ &\quad - \mathbb{E}^{E_0} \left[\int_{t_0}^{\tau_f} \sum_{i=1}^q \sum_{j=1, j \neq i}^q \frac{1}{2} \frac{\phi_{ij}(E(t)) [E_j(t) - E_i(t)]}{(c + E_i(t))(c + E_j(t))} dt \right] \\ &\quad + \mathbb{E}^{E_0} \left[\int_{t_0}^{\tau_f} \sum_{i=1}^q \frac{1}{2} \frac{\text{row}_i(J(E(t))) \text{row}_i^T(J(E(t)))}{(c + E_i(t))^2} dt \right] \end{aligned} \quad (94)$$

$$\begin{aligned}
 & + \mathbb{E}^{E_0} \left[\int_{t_0}^{\tau_f} \sum_{i=1}^q \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} dt \right] \\
 & = \mathbb{E}^{E_0} \left[\int_{t_0}^{\tau_f} \sum_{i=1}^q \frac{1}{2} \frac{1}{(c + E_i(t))^2} \left[\sum_{j=1, j \neq i}^q \phi_{ij}(E(t)) [E_i(t) - E_j(t)] \frac{c + E_i(t)}{c + E_j(t)} \right. \right. \\
 & \quad \left. \left. + \text{row}_i(J(E(t))) \text{row}_i^T(J(E(t))) \right] dt \right] \\
 & + \mathbb{E}^{E_0} \left[\int_{t_0}^{\tau_f} \sum_{i=1}^q \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} dt \right] \\
 & \leq \mathbb{E}^{E_0} \left[\int_{t_0}^{\tau_f} \sum_{i=1}^q \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} dt \right],
 \end{aligned}$$

which proves (92).

To show (93), note that it follows from (94), Axiom (i), and Axiom (ii) that (93) holds if and only if $E_i(t) \stackrel{\text{a.s.}}{=} E_j(t)$, $t \in [t_0, \tau_f]$, $i \neq j$, $i, j = 1, \dots, q$, or, equivalently, there exists a continuous function $\alpha : [t_0, \tau_f] \rightarrow \mathbb{R}_+$ such that $E(t) \stackrel{\text{a.s.}}{=} \alpha(t) \mathbf{e}$, $t \in [t_0, \tau_f]$. $\square$

Inequality (92) is a generalization of Clausius' inequality for reversible and irreversible thermodynamics as applied to large-scale stochastic dynamical systems and restricts the manner in which the system dissipates (scaled) heat over cyclic motions. Note that the Clausius inequality (92) for the stochastic thermodynamic model is stronger than the Clausius inequality for the deterministic model presented in [30].

It follows from Axiom (i) and (68) that for the *adiabatically isolated* large-scale stochastic dynamical system $\mathcal{G}$ (that is, $S(t) \equiv 0$ and $D(E(t)) \equiv 0$), the energy states given by $E_e = \alpha \mathbf{e}$, $\alpha \geq 0$, correspond to the equilibrium energy states of $\mathcal{G}$. Thus, as in classical thermodynamics, we can define an equilibrium process as a process in which the trajectory of the large-scale stochastic dynamical system $\mathcal{G}$ moves along the equilibrium manifold $\mathcal{M}_e \triangleq \{E \in \mathbb{R}_+^q : E = \alpha \mathbf{e}, \alpha \geq 0\}$ corresponding to the set of equilibria of the isolated [60] system $\mathcal{G}$. The power input that can generate such a trajectory can be given by $S(t) = d(E(t)) + u(t)$, $t \geq t_0$, where $u(\cdot) \in \mathcal{U}$ is such that $u_i(t) \equiv u_j(t)$, $i \neq j$, $i, j = 1, \dots, q$. Our definition of an equilibrium transformation involves a continuous succession of intermediate states that differ by infinitesimals from equilibrium system states and thus can only connect initial and final states, which are states of equilibrium. This process need not be slowly varying, and hence, equilibrium and quasistatic processes are not synonymous in this paper. Alternatively, a *nonequilibrium process* is a process that does not lie on the equilibrium manifold $\mathcal{M}_e$. Hence, it follows from Axiom (i) that for an equilibrium process $\phi_{ij}(E(t)) = 0$, $t \geq t_0$, $i \neq j$, $i, j = 1, \dots, q$, and thus, by Proposition 5, inequality (92) is satisfied as an equality. Alternatively, for a nonequilibrium process it follows from Axioms (i) and (ii) that (92) is satisfied as a strict inequality.

Next, we give a stochastic definition of entropy for the large-scale stochastic dynamical system $\mathcal{G}$ that is consistent with the classical thermodynamic definition of entropy.

Definition 11. For the large-scale stochastic dynamical system $\mathcal{G}$ with differential energy balance Equation (68), a function $\mathcal{S} : \mathbb{R}_+^q \rightarrow \mathbb{R}$ satisfying

$$\begin{aligned}
 \mathbb{E}[\mathcal{S}(E(\tau_2)) | \mathcal{F}_{\tau_1}] & \geq \mathcal{S}(E(\tau_1)) + \mathbb{E} \left[\int_{\tau_1}^{\tau_2} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} \right. \right. \\
 & \quad \left. \left. - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | \mathcal{F}_{\tau_1} \right]
 \end{aligned} \tag{95}$$

for every $\mathcal{F}_t$ -stopping times $\tau_2 \stackrel{\text{a.s.}}{\geq} \tau_1 \stackrel{\text{a.s.}}{\geq} t_0$ and $S(\cdot) \in \mathcal{U}$ is called the entropy function of $\mathcal{G}$.

Note that it follows from Definition 11 that the difference between the system entropy production and the stored system entropy is a submartingale with respect to the differential energy balance filtration.

Next, we show that (92) guarantees the existence of an entropy function for $\mathcal{G}$. For this result define the *available entropy* of the large-scale stochastic dynamical system $\mathcal{G}$ by

$$\mathcal{S}_a(E_0) \triangleq - \sup_{S(\cdot) \in \mathcal{U}_c, \tau_0 \stackrel{\text{a.s.}}{\geq} t_0} \mathbb{E} \left[\mathbb{E} \left[\int_{t_0}^{\tau_0} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | E(t_0) \stackrel{\text{a.s.}}{=} E_0 \right] \right], \quad (96)$$

where $E_0 \in \mathbb{R}_+^q$ and $E(\tau_0) \stackrel{\text{a.s.}}{=} 0$, and define the *required entropy supply* of the large-scale stochastic dynamical system $\mathcal{G}$ by

$$\mathcal{S}_r(E_0) \triangleq \sup_{S(\cdot) \in \mathcal{U}_r, \tau_{E_0} \stackrel{\text{a.s.}}{\geq} t_0} \mathbb{E} \left[\mathbb{E} \left[\int_{t_0}^{\tau_{E_0}} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | E(t_0) \stackrel{\text{a.s.}}{=} 0 \right] \right], \quad (97)$$

where $E(\tau_{E_0}) \stackrel{\text{a.s.}}{=} E_0 \in \mathbb{R}_+^q$. Note that the available entropy $\mathcal{S}_a(E_0)$ is the minimum amount of scaled heat (entropy) that can be extracted from the large-scale stochastic dynamical system $\mathcal{G}$ in order to transfer it from an initial state $E(t_0) = E_0$ to $E(T) = 0$. Alternatively, the required entropy supply $\mathcal{S}_r(E_0)$ is the maximum amount of scaled heat (entropy) that can be delivered to $\mathcal{G}$ to transfer it from the origin to a given subset in the state space containing the initial state $E(t_0) = E_0$ over a finite stopping time. For further details, see [54].

Theorem 7. Consider the large-scale stochastic dynamical system $\mathcal{G}$ with differential energy balance Equation (68), and assume that Axiom (ii) holds. Then there exists an entropy function for $\mathcal{G}$. Moreover, $\mathcal{S}_a(E)$, $E \in \mathbb{R}_+^q$, and $\mathcal{S}_r(E)$, $E \in \mathbb{R}_+^q$, are possible entropy functions for $\mathcal{G}$ with $\mathcal{S}_a(0) = \mathcal{S}_r(0) = 0$. Finally, all entropy functions $\mathcal{S}(E)$, $E \in \mathbb{R}_+^q$, for $\mathcal{G}$ satisfy

$$\mathcal{S}_r(E) \leq \mathcal{S}(E) - \mathcal{S}(0) \leq \mathcal{S}_a(E), \quad E \in \mathbb{R}_+^q. \quad (98)$$

Proof. Since, by Lemma 2, $\mathcal{G}$ is stochastically controllable to and stochastically reachable from the origin in $\mathbb{R}_+^q$, it follows from (96) and (97) that $\mathcal{S}_a(E_0) < \infty$, $E_0 \in \mathbb{R}_+^q$, and $\mathcal{S}_r(E_0) > -\infty$, $E_0 \in \mathbb{R}_+^q$, respectively. Next, let $E_0 \in \mathbb{R}_+^q$, and let $S(\cdot) \in \mathcal{U}$ be such that $E(\tau_i) \stackrel{\text{a.s.}}{=} E(\tau_f) \stackrel{\text{a.s.}}{=} 0$ and $E(\tau_0) \stackrel{\text{a.s.}}{=} E_0$, where $\tau_i \stackrel{\text{a.s.}}{<} \tau_0 \stackrel{\text{a.s.}}{<} \tau_f$. In this case, it follows from (92) that, for all $\tau_i \stackrel{\text{a.s.}}{<} \tau_0 \stackrel{\text{a.s.}}{<} \tau_f$,

$$\mathbb{E} \left[\mathbb{E} \left[\int_{\tau_i}^{\tau_f} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | E(\tau_i) \stackrel{\text{a.s.}}{=} 0 \right] \right] \leq 0. \quad (99)$$

Next, using the strong Markov property we have

$$\begin{aligned} & \mathbb{E} \left[\mathbb{E} \left[\int_{\tau_i}^{\tau_f} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | E(\tau_i) \right] \right] \\ &= \mathbb{E} \left[\mathbb{E} \left[\int_{\tau_i}^{\tau_0} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt \right. \right. \\ & \quad \left. \left. + \int_{\tau_0}^{\tau_f} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | E(\tau_i) \right] \right] \end{aligned}$$

$$\begin{aligned}
&= \mathbb{E} \left[\mathbb{E} \left[\int_{\tau_1}^{\tau_0} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | E(\tau_1) \right] \right] \\
&\quad + \mathbb{E} \left[\mathbb{E} \left[\int_{\tau_0}^{\tau_f} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | \mathcal{F}_{\tau_0} \right] \right] \\
&= \mathbb{E} \left[\mathbb{E} \left[\int_{\tau_1}^{\tau_0} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | E(\tau_1) \right] \right] \\
&\quad + \mathbb{E} \left[\mathbb{E} \left[\int_{\tau_0}^{\tau_f} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | E(\tau_0) \right] \right],
\end{aligned} \tag{100}$$

and hence, (99) implies

$$\begin{aligned}
&\mathbb{E} \left[\mathbb{E} \left[\int_{\tau_1}^{\tau_0} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | E(\tau_1) \right] \right] \\
&\leq -\mathbb{E} \left[\mathbb{E} \left[\int_{\tau_0}^{\tau_f} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | E(\tau_0) \right] \right].
\end{aligned} \tag{101}$$

Now, taking the supremum on both sides of (101) over all $S(\cdot) \in \mathcal{U}_r$ and $\tau_1 \leq \tau_0$ ^{a.s.} yields

$$\begin{aligned}
\mathcal{S}_r(E_0) &= \sup_{S(\cdot) \in \mathcal{U}_r, \tau_1 \leq \tau_0} \mathbb{E} \left[\mathbb{E} \left[\int_{\tau_1}^{\tau_0} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} \right. \right. \right. \\
&\quad \left. \left. \left. - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | E(\tau_1) \right] \right] \\
&\leq -\mathbb{E} \left[\mathbb{E} \left[\int_{\tau_0}^{\tau_f} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | E(\tau_0) \right] \right].
\end{aligned} \tag{102}$$

Next, taking the infimum on both sides of (102) over all $S(\cdot) \in \mathcal{U}_c$ and $\tau_f \geq \tau_0$ ^{a.s.}, we obtain $\mathcal{S}_r(E_0) \leq \mathcal{S}_a(E_0)$, $E_0 \in \mathbb{R}_+^q$, which implies that $-\infty < \mathcal{S}_r(E_0) \leq \mathcal{S}_a(E_0) < \infty$, $E_0 \in \mathbb{R}_+^q$. Hence, the functions $\mathcal{S}_a(\cdot)$ and $\mathcal{S}_r(\cdot)$ are well defined.

Next, it follows from the definition of $\mathcal{S}_a(\cdot)$, the law of iterated expectation, and the strong Markov property that for every stopping time $\mathcal{T} \geq \tau_1$ ^{a.s.} and $S(\cdot) \in \mathcal{U}_c$ such that $E(\tau_1) \in \overline{\mathcal{H}}_q^+$ and $E(\mathcal{T}) \stackrel{\text{a.s.}}{=} 0$,

$$\begin{aligned}
&-\mathcal{S}_a(E(\tau_1)) \\
&= \sup_{S(\cdot) \in \mathcal{U}_c, T \geq \tau_1} \mathbb{E} \left[\int_{\tau_1}^{\mathcal{T}} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | E(\tau_1) \right], \\
&\geq \mathbb{E} \left[\int_{\tau_1}^{\tau_2} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | \mathcal{F}_{\tau_1} \right] \\
&\quad + \sup_{S(\cdot) \in \mathcal{U}_c, T \geq \tau_2} \mathbb{E} \left[\int_{\tau_2}^{\mathcal{T}} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | \mathcal{F}_{\tau_1} \right], \\
&= \mathbb{E} \left[\int_{\tau_1}^{\tau_2} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | \mathcal{F}_{\tau_1} \right] \\
&\quad + \mathbb{E} \left[\sup_{S(\cdot) \in \mathcal{U}_c, T \geq \tau_2} \mathbb{E} \left[\int_{\tau_2}^{\mathcal{T}} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} \right. \right. \right. \\
&\quad \left. \left. \left. - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | \mathcal{F}_{\tau_2} \right] | \mathcal{F}_{\tau_1} \right],
\end{aligned} \tag{103}$$

$$= \mathbb{E} \left[\int_{\tau_1}^{\tau_2} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | \mathcal{F}_{\tau_1} \right] \\ - \mathbb{E} [\mathcal{S}_a(E(\tau_2)) | \mathcal{F}_{\tau_1}], \quad \tau_1 \stackrel{\text{a.s.}}{\leq} \tau_2 \stackrel{\text{a.s.}}{\leq} \mathcal{T},$$

which implies that $\mathcal{S}_a(E)$, $E \in \overline{\mathbb{R}}_+^q$, satisfies (95). Thus, $\mathcal{S}_a(E)$, $E \in \overline{\mathbb{R}}_+^q$, is a possible entropy function for $\mathcal{G}$. Note that with $E(\tau_0) \stackrel{\text{a.s.}}{=} E(\mathcal{T}) \stackrel{\text{a.s.}}{=} 0$ it follows from (92) that the supremum in (96) is taken over the set of negative semidefinite values with one of the values being zero for $S(t) \stackrel{\text{a.s.}}{=} 0$. Thus, $\mathcal{S}_a(0) = 0$.

Similarly, it follows from the definition of $\mathcal{S}_r(\cdot)$ that for every stopping time $\mathcal{T} \leq \tau_2$ and $S(\cdot) \in \mathcal{U}_r$ such that $E(\tau_2) \in \overline{\mathcal{H}}_q^+$ and $E(\mathcal{T}) \stackrel{\text{a.s.}}{=} 0$,

$$\begin{aligned} \mathcal{S}_r(E(\tau_2)) &= \sup_{S(\cdot) \in \mathcal{U}_r, \mathcal{T} \stackrel{\text{a.s.}}{\leq} \tau_2} \int_{\mathcal{T}}^{\tau_2} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt, \\ &\geq \sup_{S(\cdot) \in \mathcal{U}_r, \mathcal{T} \stackrel{\text{a.s.}}{\leq} \tau_1} \int_{\mathcal{T}}^{\tau_1} \left[\sum_{i=1}^q \frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt \\ &\quad + \int_{\tau_1}^{\tau_2} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt, \quad (104) \\ &= \mathcal{S}_r(E(\tau_1)) + \int_{\tau_1}^{\tau_2} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt, \\ &\quad \mathcal{T} \stackrel{\text{a.s.}}{\leq} \tau_1 \stackrel{\text{a.s.}}{\leq} \tau_2, \end{aligned}$$

which implies that $\mathcal{S}_r(E)$, $E \in \overline{\mathbb{R}}_+^q$, satisfies (95). Thus, $\mathcal{S}_r(E)$, $E \in \overline{\mathbb{R}}_+^q$, is a possible entropy function for $\mathcal{G}$. Note that with $E(t_0) \stackrel{\text{a.s.}}{=} E(\mathcal{T}) \stackrel{\text{a.s.}}{=} 0$ it follows from (92) that the supremum in (97) is taken over the set of negative semidefinite values with one of the values being zero for $S(t) \stackrel{\text{a.s.}}{=} 0$. Thus, $\mathcal{S}_r(0) = 0$.

Next, suppose there exists an entropy function $\mathcal{S} : \overline{\mathbb{R}}_+^q \rightarrow \mathbb{R}$ for $\mathcal{G}$, and let $E(\tau_2) \stackrel{\text{a.s.}}{=} 0$ in (95). Then it follows from (95) that

$$\mathcal{S}(E(\tau_1)) - \mathcal{S}(0) \leq -\mathbb{E} \left[\int_{\tau_1}^{\tau_2} \sum_{i=1}^q \frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} dt | \mathcal{F}_{\tau_1} \right] \quad (105)$$

for all $\tau_2 \stackrel{\text{a.s.}}{\geq} \tau_1$ and $S(\cdot) \in \mathcal{U}_c$, which implies that

$$\begin{aligned} &\mathcal{S}(E(\tau_1)) - \mathcal{S}(0) \\ &\leq \inf_{S(\cdot) \in \mathcal{U}_c, \tau_2 \stackrel{\text{a.s.}}{\geq} \tau_1} \left[-\mathbb{E} \left[\int_{\tau_1}^{\tau_2} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | \mathcal{F}_{\tau_1} \right] \right] \\ &= - \sup_{S(\cdot) \in \mathcal{U}_c, \tau_2 \stackrel{\text{a.s.}}{\geq} \tau_1} \mathbb{E} \left[\int_{\tau_1}^{\tau_2} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | \mathcal{F}_{\tau_1} \right] \quad (106) \\ &= \mathcal{S}_a(E(\tau_1)). \end{aligned}$$

Since $E(\tau_1)$ is arbitrary, it follows that $\mathcal{S}(E) - \mathcal{S}(0) \leq \mathcal{S}_a(E)$, $E \in \overline{\mathbb{R}}_+^q$.

Alternatively, let $E(\tau_1) \stackrel{\text{a.s.}}{=} 0$ in (95). Then it follows from (95) that

$$\begin{aligned} \mathcal{S}(E(\tau_2)) - \mathcal{S}(0) &\geq \mathbb{E} \left[\int_{\tau_1}^{\tau_2} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} \right. \right. \\ &\quad \left. \left. - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | \mathcal{F}_{\tau_1} \right] \quad (107) \end{aligned}$$

for all $\tau_1 \leq \tau_2$ and $S(\cdot) \in \mathcal{U}_r$. Hence,

$$\begin{aligned} \mathcal{S}(E(\tau_2)) - \mathcal{S}(0) &\geq \sup_{S(\cdot) \in \mathcal{U}_r, \tau_1 \leq \tau_2} \mathbb{E} \left[\int_{\tau_1}^{\tau_2} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} \right. \right. \\ &\quad \left. \left. - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt \middle| \mathcal{F}_{\tau_1} \right] \\ &= \mathcal{S}_r(E(\tau_2)), \end{aligned} \quad (108)$$

which, since $E(\tau_2)$ is arbitrary, implies that $\mathcal{S}_r(E) \leq \mathcal{S}(E) - \mathcal{S}(0)$, $E \in \overline{\mathbb{R}}_+^q$. Thus, all entropy functions for $\mathcal{G}$ satisfy (98). $\square$

It is important to note that inequality (92) is equivalent to the existence of an entropy function for $\mathcal{G}$. Sufficiency is simply a statement of Theorem 7, while necessity follows from (95) with $E(t_2) \stackrel{\text{a.s.}}{=} E(t_1)$. This definition of entropy leads to the second law of stochastic thermodynamics being viewed as an axiom in the context of stochastic (anti)cyclo-dissipative dynamical systems [54].

The next result shows that all entropy functions for $\mathcal{G}$ are continuous on $\overline{\mathbb{R}}_+^q$.

Theorem 8. Consider the large-scale stochastic dynamical system $\mathcal{G}$ with differential energy balance Equation (68), and let $\mathcal{S} : \overline{\mathbb{R}}_+^q \rightarrow \mathbb{R}$ be an entropy function of $\mathcal{G}$. Then $\mathcal{S}(\cdot)$ is continuous on $\overline{\mathbb{R}}_+^q$.

Proof. Let $E_e \in \overline{\mathbb{R}}_+^q$ and $S_e \in \mathbb{R}^q$ be such that $S_e = d(E_e) - f(E_e)$. Note that with $S(t) \stackrel{\text{a.s.}}{=} S_e$, E_e is an equilibrium point of the differential energy balance Equation (68). Next, it follows from Lemma 2 that $\mathcal{G}$ is locally stochastically controllable, that is, for every $\tau > 0$ and $\varepsilon > 0$, the set of points that can be reached from and to E_e in time T using admissible inputs $S : [0, \tau] \rightarrow \mathcal{H}_q$, satisfying $\|S(t) - S_e\| \stackrel{\text{a.s.}}{<} \varepsilon$, contains a neighborhood of E_e .

Next, let $\delta > 0$ and note that it follows from the continuity of $f(\cdot)$, $d(\cdot)$, $J(\cdot)$, and $D(\cdot)$ that there exist $\tau > 0$ and $\varepsilon > 0$ such that for every $S : [0, \tau] \rightarrow \mathbb{R}^q$ and $\|S(t) - S_e\| \stackrel{\text{a.s.}}{<} \varepsilon$, $\|E(t) - E_e\| \stackrel{\text{a.s.}}{<} \delta$, $t \in [0, \tau]$, where $S(\cdot) \in \mathcal{U}$ and $E(t)$, $t \in [0, \tau]$, denotes the solution to (68) with the initial condition E_e . Furthermore, it follows from the local controllability of $\mathcal{G}$ that for every $\hat{\tau} \in (0, \tau]$, there exists a strictly increasing, continuous function $\gamma : \overline{\mathbb{R}}_+^q \rightarrow \overline{\mathbb{R}}_+^q$ such that $\gamma(0) = 0$, and for every $E_0 \in \mathcal{H}_q^+$ such that $\|E_0 - E_e\| \stackrel{\text{a.s.}}{\leq} \gamma(\hat{\tau})$, there exists $0 \leq \hat{\tau} \leq \tau$ and an input $S : [0, \hat{\tau}] \rightarrow \mathcal{H}_q$ such that $\|S(t) - S_e\| < \varepsilon$, $t \in [0, \hat{\tau}]$, and $E(\hat{\tau}) \stackrel{\text{a.s.}}{=} E_0$. Hence, there exists $\beta > 0$ such that for every $E_0 \in \mathcal{H}_q^+$ such that $\|E_0 - E_e\| \stackrel{\text{a.s.}}{\leq} \beta$, there exists $0 \leq \hat{\tau} \leq \gamma^{-1}(\|E_0 - E_e\|)$ and an input $S : [0, \hat{\tau}] \rightarrow \mathcal{H}_q$ such that $\|S(t) - S_e\| \stackrel{\text{a.s.}}{<} \varepsilon$, $t \in [0, \hat{\tau}]$, and $E(\hat{\tau}) \stackrel{\text{a.s.}}{=} E_0$. In addition, it follows from Lemma 2 that $S : [0, \hat{\tau}] \rightarrow \mathcal{H}_q$ is such that $E(t) \stackrel{\text{a.s.}}{\geq} 0$, $t \in [0, \hat{\tau}]$.

Next, since $\sigma_{ii}(\cdot)$, $i = 1, \dots, q$, is continuous, it follows that there exists $M \in \mathcal{H}_1^+$ such that

$$\sup_{\|E - E_e\| \stackrel{\text{a.s.}}{<} \delta, \|S - S_e\| \stackrel{\text{a.s.}}{<} \varepsilon} \left| \sum_{i=1}^q \left[\frac{S_i - \sigma_{ii}(E)}{c + E_i} - \frac{1}{2} \frac{\text{row}_i(D(E)) \text{row}_i^T(D(E))}{(c + E_i)^2} \right] \right| = M. \quad (109)$$

Hence, it follows that

$$\begin{aligned} &\left| \int_0^{\hat{\tau}} \sum_{i=1}^q \left[\frac{S_i(\sigma) - \sigma_{ii}(E(\sigma))}{c + E_i(\sigma)} - \frac{1}{2} \frac{\text{row}_i(D(E(\sigma))) \text{row}_i^T(D(E(\sigma))))}{(c + E_i(\sigma))^2} \right] d\sigma \right| \\ &\stackrel{\text{a.s.}}{\leq} \int_0^{\hat{\tau}} \left| \sum_{i=1}^q \left[\frac{S_i(\sigma) - \sigma_{ii}(E(\sigma))}{c + E_i(\sigma)} - \frac{1}{2} \frac{\text{row}_i(D(E(\sigma))) \text{row}_i^T(D(E(\sigma))))}{(c + E_i(\sigma))^2} \right] \right| d\sigma \end{aligned}$$

$$\begin{aligned} & \stackrel{\text{a.s.}}{\leq} M\hat{\tau} \\ & \stackrel{\text{a.s.}}{\leq} M\gamma^{-1}(\|E_0 - E_e\|). \end{aligned} \quad (110)$$

Now, if $\mathcal{S}(\cdot)$ is an entropy function of $\mathcal{G}$, then

$$\begin{aligned} \mathbb{E}[\mathcal{S}(E(\hat{\tau}))|\mathcal{F}_0] & \stackrel{\text{a.s.}}{\geq} \mathcal{S}(E_e) + \mathbb{E}\left[\int_0^{\hat{\tau}} \sum_{i=1}^q \left[\frac{S_i(\sigma) - \sigma_{ii}(E(\sigma))}{c + E_i(\sigma)} \right. \right. \\ & \quad \left. \left. - \frac{1}{2} \frac{\text{row}_i(D(E(\sigma)))\text{row}_i^T(D(E(\sigma)))}{(c + E_i(\sigma))^2} \right] d\sigma | \mathcal{F}_0\right] \end{aligned} \quad (111)$$

or, equivalently,

$$\begin{aligned} -\mathbb{E}\left[\int_0^{\hat{\tau}} \sum_{i=1}^q \left[\frac{S_i(\sigma) - \sigma_{ii}(E(\sigma))}{c + E_i(\sigma)} - \frac{1}{2} \frac{\text{row}_i(D(E(\sigma)))\text{row}_i^T(D(E(\sigma)))}{(c + E_i(\sigma))^2} \right] d\sigma | \mathcal{F}_0\right] \\ \stackrel{\text{a.s.}}{\geq} \mathcal{S}(E_e) - \mathbb{E}[\mathcal{S}(E(\hat{\tau}))|\mathcal{F}_0]. \end{aligned} \quad (112)$$

If $\mathcal{S}(E_e) \stackrel{\text{a.s.}}{\geq} \mathcal{S}(E(\hat{\tau}))$, then combining (110) and (112) yields

$$|\mathcal{S}(E_e) - \mathbb{E}[\mathcal{S}(E(\hat{\tau}))|\mathcal{F}_0]| \stackrel{\text{a.s.}}{\leq} \mathbb{E}[M\gamma^{-1}(\|E_0 - E_e\|)|\mathcal{F}_0]. \quad (113)$$

Alternatively, if $\mathcal{S}(E(\hat{\tau})) \stackrel{\text{a.s.}}{\geq} \mathcal{S}(E_e)$, then (113) can be derived by reversing the roles of E_e and $E(\hat{\tau})$. Specifically, for $E_0 \in \overline{\mathbb{R}}_+^q$ and $E(\hat{\tau}) \stackrel{\text{a.s.}}{=} E_0$, (113) becomes

$$|\mathcal{S}(E_0) - \mathcal{S}(E_e)| \leq \mathbb{E}[M]\gamma^{-1}(\|E_0 - E_e\|).$$

Hence, since $\gamma(\cdot)$ is continuous and $E(\hat{\tau})$ is arbitrary, it follows that $\mathcal{S}(\cdot)$ is continuous on $\overline{\mathbb{R}}_+^q$. $\square$

Next, as a direct consequence of Theorem 7, we show that all possible entropy functions of $\mathcal{G}$ form a convex set, and hence, there exists a continuum of possible entropy functions for $\mathcal{G}$ ranging from the required entropy supply $\mathcal{S}_r(E)$ to the available entropy $\mathcal{S}_a(E)$.

Proposition 6. Consider the large-scale stochastic dynamical system $\mathcal{G}$ with differential energy balance Equation (68), and assume that Axioms (i) and (ii) hold. Then

$$\mathcal{S}(E) \triangleq \alpha \mathcal{S}_r(E) + (1 - \alpha) \mathcal{S}_a(E), \quad \alpha \in [0, 1], \quad (114)$$

is an entropy function for $\mathcal{G}$.

Proof. The result is a direct consequence of the reachability of $\mathcal{G}$ along with inequality (95) by noting that if $\mathcal{S}_r(E)$ and $\mathcal{S}_a(E)$ satisfy (95), then $\mathcal{S}(E)$ satisfies (95). $\square$

It follows from Proposition 6 that Definition 11 does not provide enough information to define the entropy uniquely for nonequilibrium thermodynamic systems with differential energy balance Equation (68). This difficulty has long been pointed out in [61]. Two particular entropy functions for $\mathcal{G}$ can be computed a priori via the variational problems given by (96) and (97). For equilibrium thermodynamics, however, uniqueness is not an issue, as shown in the next proposition.

Proposition 7. Consider the large-scale stochastic dynamical system $\mathcal{G}$ with differential energy balance Equation (68), and assume that Axioms (i) and (ii) hold. Then at every equilibrium state $E = E_e$ of the isolated system $\mathcal{G}$, the entropy $\mathcal{S}(E)$, $E \in \mathbb{R}_+^q$, of $\mathcal{G}$ is unique (modulo a constant of integration) and is given by

$$\mathcal{S}(E) - \mathcal{S}(0) = \mathcal{S}_a(E) = \mathcal{S}_r(E) = \mathbf{e}^T \mathbf{log}_e(\mathbf{ce} + E) - q \log_e c, \quad (115)$$

where $E = E_e$ and $\mathbf{log}_e(\mathbf{ce} + E)$ denotes the vector natural logarithm given by $[\log_e(c + E_1), \dots, \log_e(c + E_q)]^T$.

Proof. It follows from Axiom (i) and Axiom (ii) that for an equilibrium process $\phi_{ij}(E(t)) \stackrel{\text{a.s.}}{=} 0$, $i \neq j$, $i, j = 1, \dots, q$, $D(E(t)) \stackrel{\text{a.s.}}{=} 0$, and $J(E(t)) \stackrel{\text{a.s.}}{=} 0$. Consider the entropy function $\mathcal{S}_a(\cdot)$ given by (96), and let $E_0 = E_e$ for some equilibrium state E_e . Then it follows from (68) that

$$\begin{aligned} \mathcal{S}_a(E_0) &= - \sup_{S(\cdot) \in \mathcal{U}_c, \mathcal{T} \geq t_0} \mathbb{E} \left[\mathbb{E} \left[\int_{t_0}^{\mathcal{T}} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} \right. \right. \right. \\ &\quad \left. \left. \left. - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt \middle| E(t_0) \stackrel{\text{a.s.}}{=} E_0 \right] \right] \\ &= - \sup_{S(\cdot) \in \mathcal{U}_c, \mathcal{T} \geq t_0} \mathbb{E} \left[\mathbb{E} \left[\int_{t_0}^{\mathcal{T}} \sum_{i=1}^q \left[\frac{dE_i(t) - \sum_{j=1, j \neq i}^q \phi_{ij}(E(t)) dt}{c + E_i(t)} \right. \right. \right. \\ &\quad \left. \left. \left. - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} dt \right] \middle| E(t_0) \stackrel{\text{a.s.}}{=} E_0 \right] \right] \\ &= - \sup_{S(\cdot) \in \mathcal{U}_c, \mathcal{T} \geq t_0} \mathbb{E} \left[\mathbb{E} \left[\sum_{i=1}^q \log_e \left(\frac{c}{c + E_{i0}} \right) + \int_{t_0}^{\mathcal{T}} \sum_{i=1}^q \frac{1}{2} \frac{\text{row}_i(J(E(t))) \text{row}_i^T(J(E(t)))}{(c + E_i(t))^2} dt \right. \right. \right. \\ &\quad \left. \left. \left. - \int_{t_0}^{\mathcal{T}} \sum_{i=1}^q \sum_{j=1, j \neq i}^q \frac{\phi_{ij}(E(t))}{c + E_i(t)} dt \middle| E(t_0) \stackrel{\text{a.s.}}{=} E_0 \right] \right] \quad (116) \\ &= - \sup_{S(\cdot) \in \mathcal{U}_c, \mathcal{T} \geq t_0} \mathbb{E} \left[\mathbb{E} \left[\sum_{i=1}^q \log_e \left(\frac{c}{c + E_{i0}} \right) + \int_{t_0}^{\mathcal{T}} \sum_{i=1}^q \frac{1}{2} \frac{\text{row}_i(J(E(t))) \text{row}_i^T(J(E(t)))}{(c + E_i(t))^2} dt \right. \right. \right. \\ &\quad \left. \left. \left. - \int_{t_0}^{\mathcal{T}} \sum_{i=1}^q \sum_{j=1, j \neq i}^q \frac{1}{2} \left(\frac{\phi_{ij}(E(t))}{c + E_i(t)} - \frac{\phi_{ij}(E(t))}{c + E_j(t)} \right) dt \middle| E(t_0) \stackrel{\text{a.s.}}{=} E_0 \right] \right] \\ &= \sum_{i=1}^q \log_e \left(\frac{c + E_{i0}}{c} \right) + \inf_{S(\cdot) \in \mathcal{U}_c, \mathcal{T} \geq t_0} \mathbb{E} \left[\mathbb{E} \left[\int_{t_0}^{\mathcal{T}} \sum_{i=1}^q -\frac{1}{2} \frac{1}{(c + E_i(t))^2} \left[\sum_{j=1, j \neq i}^q \phi_{ij}(E(t)) \right. \right. \right. \right. \\ &\quad \left. \left. \left. \cdot [E_i(t) - E_j(t)] \frac{c + E_i(t)}{c + E_j(t)} + \text{row}_i(J(E(t))) \text{row}_i^T(J(E(t))) \right] dt \middle| E(t_0) \stackrel{\text{a.s.}}{=} E_0 \right] \right]. \end{aligned}$$

Since the solution $E(t)$, $t \geq t_0$, to (68) is nonnegative for all nonnegative initial conditions, it follows from Axiom (ii) that the infimum in (116) is taken over the set of nonnegative values. However, the zero value of the infimum is achieved on an equilibrium process for which $\phi_{ij}(E(t)) \stackrel{\text{a.s.}}{=} 0$, $i \neq j$, $i, j = 1, \dots, q$. Thus,

$$\mathcal{S}_a(E_0) = \mathbf{e}^T \mathbf{log}_e(\mathbf{ce} + E_0) - q \log_e c, \quad E_0 = E_e. \quad (117)$$

Similarly, consider the entropy function $\mathcal{S}_r(\cdot)$ given by (97). Then, it follows from (68) that, for $E_0 = E_e$,

$$\begin{aligned}
 \mathcal{S}_r(E_0) &= \sup_{S(\cdot) \in \mathcal{U}_r, \mathcal{T} \geq t_0}^{\text{a.s.}} \mathbb{E} \left[\mathbb{E} \left[\int_{t_0}^{\mathcal{T}} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} \right. \right. \right. \\
 &\quad \left. \left. \left. - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt \middle| E(t_0) \stackrel{\text{a.s.}}{=} 0 \right] \right] \\
 &= \sup_{S(\cdot) \in \mathcal{U}_r, \mathcal{T} \geq t_0}^{\text{a.s.}} \mathbb{E} \left[\mathbb{E} \left[\int_{t_0}^{\mathcal{T}} \sum_{i=1}^q \left[\frac{dE_i(t) - \sum_{j=1, j \neq i}^q \phi_{ij}(E(t)) dt}{c + E_i(t)} \right. \right. \right. \\
 &\quad \left. \left. \left. - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} dt \right] \middle| E(t_0) \stackrel{\text{a.s.}}{=} 0 \right] \right] \\
 &= \sup_{S(\cdot) \in \mathcal{U}_r, \mathcal{T} \geq t_0}^{\text{a.s.}} \mathbb{E} \left[\mathbb{E} \left[\sum_{i=1}^q \log_e \left(\frac{c + E_{i0}}{c} \right) + \int_{t_0}^{\mathcal{T}} \sum_{i=1}^q \frac{1}{2} \frac{\text{row}_i(J(E(t))) \text{row}_i^T(J(E(t)))}{(c + E_i(t))^2} dt \right. \right. \\
 &\quad \left. \left. - \int_{t_0}^{\mathcal{T}} \sum_{i=1}^q \sum_{j=1, j \neq i}^q \frac{\phi_{ij}(E(t))}{c + E_i(t)} dt \middle| E(t_0) \stackrel{\text{a.s.}}{=} 0 \right] \right] \\
 &= \sum_{i=1}^q \log_e \left(\frac{c + E_{i0}}{c} \right) + \sup_{S(\cdot) \in \mathcal{U}_r, \mathcal{T} \geq t_0}^{\text{a.s.}} \mathbb{E} \left[\mathbb{E} \left[\int_{t_0}^{\mathcal{T}} \sum_{i=1}^q \frac{1}{2} \frac{1}{(c + E_i(t))^2} \left[\sum_{j=1, j \neq i}^q \phi_{ij}(E(t)) \right. \right. \right. \\
 &\quad \left. \left. \left. \cdot [E_i(t) - E_j(t)] \frac{c + E_i(t)}{c + E_j(t)} + \text{row}_i(J(E(t))) \text{row}_i^T(J(E(t))) \right] dt \middle| E(t_0) \stackrel{\text{a.s.}}{=} 0 \right] \right].
 \end{aligned} \tag{118}$$

Now, it follows from Axioms (i) and (ii) that the zero value of the supremum in (118) is achieved on an equilibrium process and thus

$$\mathcal{S}_r(E_0) = \mathbf{e}^T \mathbf{log}_e(c\mathbf{e} + E_0) - q \log_e c, \quad E_0 = E_e. \tag{119}$$

Finally, it follows from (98) that (115) holds. $\square$

The next proposition shows that if (95) holds as an equality for some transformation starting and ending at an equilibrium point of the isolated dynamical system $\mathcal{G}$, then this transformation must lie on the equilibrium manifold $\mathcal{M}_e$.

Proposition 8. Consider the large-scale stochastic dynamical system $\mathcal{G}$ with differential energy balance Equation (68), and assume that Axioms (i) and (ii) hold. Let $\mathcal{S}(\cdot)$ denote an entropy of $\mathcal{G}$, and let $E : [t_0, t_1] \rightarrow \overline{\mathbb{R}}_+^q$ denote the solution to (68) with $E(t_0) \stackrel{\text{a.s.}}{=} \alpha_0 \mathbf{e}$ and $E(t_1) \stackrel{\text{a.s.}}{=} \alpha_1 \mathbf{e}$, where $\alpha_0, \alpha_1 \geq 0$. Then

$$\begin{aligned}
 \mathbb{E}[\mathcal{S}(E(t_1)) | \mathcal{F}_{t_0}] &= \mathcal{S}(E(t_0)) + \mathbb{E} \left[\int_{t_0}^{t_1} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} \right. \right. \\
 &\quad \left. \left. - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt \middle| \mathcal{F}_{t_0} \right]
 \end{aligned} \tag{120}$$

if and only if there exists a continuous function $\alpha : [t_0, t_1] \rightarrow \overline{\mathbb{R}}_+$ such that $\alpha(t_0) = \alpha_0$, $\alpha(t_1) = \alpha_1$, and $E(t) \stackrel{\text{a.s.}}{=} \alpha(t) \mathbf{e}$, $t \in [t_0, t_1]$.

Proof. Since $E(t_0)$ and $E(t_1)$ are equilibrium states of the isolated dynamical system $\mathcal{G}$, it follows from Proposition 7 that

$$\mathbb{E}[\mathcal{S}(E(t_1)) | \mathcal{F}_{t_0}] - \mathcal{S}(E(t_0)) \stackrel{\text{a.s.}}{=} q \log_e(c + \alpha_1) - q \log_e(c + \alpha_0). \tag{121}$$

Furthermore, it follows from (68) that

$$\begin{aligned} & \mathbb{E} \left[\int_{t_0}^{t_1} \left[\sum_{i=1}^q \frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | \mathcal{F}_{t_0} \right] \\ &= \mathbb{E} \left[\int_{t_0}^{t_1} \sum_{i=1}^q \frac{dE_i(t) - \sum_{j=1, j \neq i}^q \phi_{ij}(E(t)) dt}{c + E_i(t)} | \mathcal{F}_{t_0} \right] \\ &\quad - \mathbb{E} \left[\int_{t_0}^{t_1} \sum_{i=1}^q \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} dt | \mathcal{F}_{t_0} \right] \\ &= q \log_e \left(\frac{c + \alpha_1}{c + \alpha_0} \right) + \mathbb{E} \left[\int_{t_0}^{t_1} \sum_{i=1}^q \frac{1}{2} \frac{1}{(c + E_i(t))^2} \left[\sum_{j=1, j \neq i}^q \phi_{ij}(E(t)) [E_i(t) - E_j(t)] \right. \right. \\ &\quad \left. \left. \cdot \frac{c + E_i(t)}{c + E_j(t)} + \text{row}_i(J(E(t))) \text{row}_i^T(J(E(t))) \right] dt | \mathcal{F}_{t_0} \right]. \end{aligned} \quad (122)$$

Now, it follows from Axioms (i) and (ii) that (120) holds if and only if $E_i(t) = E_j(t)$, $t \in [t_0, t_1]$, $i \neq j$, $i, j = 1, \dots, q$, or, equivalently, there exists a continuous function $\alpha : [t_0, t_1] \rightarrow \mathbb{R}_+$ such that $E(t) \stackrel{\text{a.s.}}{=} \alpha(t)\mathbf{e}$, $t \in [t_0, t_1]$, $\alpha(t_0) = \alpha_0$, and $\alpha(t_1) = \alpha_1$. $\square$

Even though it follows from Proposition 6 that Definition 11 does not provide a unique continuous entropy function for nonequilibrium systems, the next theorem gives a *unique, two-times continuously differentiable* entropy function for $\mathcal{G}$ for equilibrium and nonequilibrium processes. This result answers the long-standing question of how the entropy of a nonequilibrium state of a dynamical process should be defined [61,62], and establishes its global existence and uniqueness.

Theorem 9. Consider the large-scale stochastic dynamical system $\mathcal{G}$ with differential energy balance Equation (68), and assume that Axioms (i) and (ii) hold. Then the function $\mathcal{S} : \mathbb{R}_+^q \rightarrow \mathbb{R}_+^q$ given by

$$\mathcal{S}(E) = \mathbf{e}^T \log_e(\mathbf{c}\mathbf{e} + E) - q \log_e c, \quad E \in \mathbb{R}_+^q, \quad (123)$$

where $c > 0$, is a unique (modulo a constant of integration), two-times continuously differentiable entropy function of $\mathcal{G}$. Furthermore, for $E(t) \notin \mathcal{H}_q^{\mathcal{M}_e}$, $t \geq t_0$, where $E(t)$, $t \geq t_0$, denotes the solution to (68) and $\mathcal{M}_e = \{E \in \mathbb{R}_+^q : E = \alpha\mathbf{e}, \alpha \geq 0\}$, (123) satisfies

$$\begin{aligned} \mathbb{E} [\mathcal{S}(E(t_2)) | \mathcal{F}_{t_1}] &> \mathcal{S}(E(t_1)) + \mathbb{E} \left[\int_{t_1}^{t_2} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} \right. \right. \\ &\quad \left. \left. - \frac{1}{2} \frac{\text{row}_i(D(E(t))) \text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | \mathcal{F}_{t_1} \right] \end{aligned} \quad (124)$$

for every $t_2 \geq t_1 \geq t_0$ and $S(\cdot) \in \mathcal{U}$.

Proof. Since, by Proposition 1, $E(t) \geq 0$, $t \geq t_0$, and $\phi_{ij}(E) = -\phi_{ji}(E)$, $E \in \mathbb{R}_+^q$, $i \neq j$, $i, j = 1, \dots, q$, it follows that

$$\begin{aligned} & \mathbb{E} [\mathcal{S}(E(t_2)) | \mathcal{F}_{t_1}] - \mathcal{S}(E(t_1)) \\ &= \mathbb{E} \left[\int_{t_1}^{t_2} d\mathcal{S}(E(t)) | \mathcal{F}_{t_1} \right] \\ &= \mathbb{E} \left[\int_{t_1}^{t_2} \sum_{i=1}^q \frac{dE_i(t)}{c + E_i(t)} - \frac{1}{2} \left[\frac{\text{row}_i(J(E(t))) \text{row}_i^T(J(E(t)))}{(c + E_i(t))^2} \right] \right] \end{aligned}$$

$$\begin{aligned}
& + \frac{\text{row}_i(D(E(t)))\text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \Big] dt | \mathcal{F}_{t_1} \Big] \\
& = \mathbb{E} \left[\int_{t_1}^{t_2} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t)))\text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | \mathcal{F}_{t_1} \right] \\
& \quad + \mathbb{E} \left[\int_{t_1}^{t_2} \sum_{i=1}^q \left[\sum_{j=1, j \neq i}^q \frac{\phi_{ij}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(J(E(t)))\text{row}_i^T(J(E(t)))}{(c + E_i(t))^2} \right] dt | \mathcal{F}_{t_1} \right] \quad (125) \\
& = \mathbb{E} \left[\int_{t_1}^{t_2} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t)))\text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | \mathcal{F}_{t_1} \right] \\
& \quad - \mathbb{E} \left[\int_{t_1}^{t_2} \sum_{i=1}^q \frac{1}{2} \frac{1}{(c + E_i(t))^2} \left[\sum_{j=1, j \neq i}^q \phi_{ij}(E(t))(E_i(t) - E_j(t)) \frac{c + E_i(t)}{c + E_j(t)} \right. \right. \\
& \quad \left. \left. + \text{row}_i(J(E(t)))\text{row}_i^T(J(E(t))) \right] dt | \mathcal{F}_{t_1} \right] \\
& \geq \mathbb{E} \left[\int_{t_1}^{t_2} \sum_{i=1}^q \left[\frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t)))\text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right] dt | \mathcal{F}_{t_1} \right], \quad t \geq t_0.
\end{aligned}$$

Furthermore, in the case where $E(t) \notin \mathcal{H}_q^{\mathcal{M}_e}$, $t \geq t_0$, it follows from Axiom (i), Axiom (ii), and (125) that (124) holds.

To show that (123) is a unique, two-times continuously differentiable entropy function of $\mathcal{G}$, let $S(E)$ be a two-times continuously differentiable entropy function of $\mathcal{G}$ so that $S(E)$ satisfies (95) or, equivalently,

$$\mathfrak{L}S(E) \geq \mu_1^T(E)[S - d(E)] - \frac{1}{2} \text{tr} \mu_2(E)D(E)D^T(E), \quad E \in \overline{\mathbb{R}}_+^q, \quad S \in \mathbb{R}^q, \quad (126)$$

where $\mu_1^T(E) = [\frac{1}{c+E_1}, \dots, \frac{1}{c+E_q}]$ and $\mu_2(E) = \text{diag}[\frac{1}{(c+E_1)^2}, \dots, \frac{1}{(c+E_q)^2}]$, $E \in \overline{\mathbb{R}}_+^q$, $E(t)$, $t \geq t_0$, denotes the solution to the differential energy balance Equation (68), and $\mathfrak{L}S(E(t))$ denotes the infinitesimal generator of $S(E)$ along the solution $E(t)$, $t \geq t_0$. Hence, it follows from (126) that

$$\begin{aligned}
& S'(E)[f(E) - d(E) + S] + \frac{1}{2} \text{tr} S''(E)[J(E)J^T(E) + D(E)D^T(E)] \\
& \geq \mu_1^T(E)[S - d(E)] - \frac{1}{2} \text{tr} \mu_2(E)D(E)D^T(E), \quad E \in \overline{\mathbb{R}}_+^q, \quad S \in \mathbb{R}^q, \quad (127)
\end{aligned}$$

which implies that there exist continuous functions $\ell : \overline{\mathbb{R}}_+^q \rightarrow \mathbb{R}^p$ and $\mathcal{W} : \overline{\mathbb{R}}_+^q \rightarrow \mathbb{R}^{p \times q}$ such that

$$\begin{aligned}
0 & = S'(E)[f(E) - d(E) + S] + \frac{1}{2} \text{tr} S''(E)[J(E)J^T(E) + D(E)D^T(E)] \\
& \quad - \mu_1^T(E)[S - d(E)] + \frac{1}{2} \text{tr} \mu_2(E)D(E)D^T(E) \\
& \quad - [\ell(E) + \mathcal{W}(E)S]^T[\ell(E) + \mathcal{W}(E)S], \quad E \in \overline{\mathbb{R}}_+^q, \quad S \in \mathbb{R}^q. \quad (128)
\end{aligned}$$

Now, equating coefficients of equal powers (of S and D), it follows that $\mathcal{W}(E) \equiv 0$, $S'(E) = \mu_1^T(E)$, $S''(E) = -\mu_2(E)$, $E \in \overline{\mathbb{R}}_+^q$, and

$$0 = S'(E)f(E) + \frac{1}{2} \text{tr} S''(E)J(E)J^T(E) - \ell^T(E)\ell(E), \quad E \in \overline{\mathbb{R}}_+^q. \quad (129)$$

Hence, $\mathcal{S}(E) = \mathbf{e}^T \log_e(c\mathbf{e} + E) - q \log_e c$, $E \in \overline{\mathbb{R}}_+^q$, and

$$0 = \mu_1^T(E)f(E) - \frac{1}{2} \text{tr} \mu_2(E)J(E)J^T(E) - \ell(E)\ell^T(E), \quad E \in \overline{\mathbb{R}}_+^q. \quad (130)$$

Thus, (123) is a unique, two-times continuously differentiable entropy function for $\mathcal{G}$. $\square$

Note that it follows from Axiom (i), Axiom (ii), and the last equality in (125) that the entropy function given by (123) satisfies (95) as an equality for an equilibrium process and as a strict inequality for a nonequilibrium process. For any entropy function of $\mathcal{G}$, it follows from Proposition 8 that if (95) holds as an equality for some transformation starting and ending at equilibrium points of the isolated system $\mathcal{G}$, then this transformation must lie on the equilibrium manifold $\mathcal{M}_e$. However, (95) may hold as an equality for nonequilibrium processes starting and ending at nonequilibrium states.

The entropy expression given by (123) is identical in form to the Boltzmann entropy for statistical thermodynamics. Due to the fact that the entropy given by (123) is indeterminate to the extent of an additive constant, we can place the constant of integration $q \log_e c$ to zero by taking $c = 1$. Since $\mathcal{S}(E)$ given by (123) achieves a maximum when all the subsystem energies E_i , $i = 1, \dots, q$, are equal, the entropy of $\mathcal{G}$ can be thought of as a measure of the tendency of a system to lose the ability to do useful work, lose order, and settle to a more homogenous state. For further details see [30].

Recalling that $\mathbb{E}[dQ_i(t)|\mathcal{F}_t] = [S_i(t) - \sigma_{ii}(E(t))]dt$, $i = 1, \dots, q$, is the infinitesimal amount of the net heat received or dissipated by the i -th subsystem of $\mathcal{G}$ over the infinitesimal time interval dt , it follows from (95) that

$$\mathbb{E}[d\mathcal{S}(E(t))|\mathcal{F}_t] \geq \sum_{i=1}^q \left[\frac{dQ_i(t)}{c + E_i(t)} - \frac{1}{2} \frac{\text{row}_i(D(E(t)))\text{row}_i^T(D(E(t)))}{(c + E_i(t))^2} \right], \quad t \geq t_0. \quad (131)$$

Inequality (131) is analogous to the classical thermodynamic inequality for the variation of entropy during an infinitesimal irreversible transformation with the shifted subsystem energies $c + E_i$ playing the role of the i -th subsystem thermodynamic (absolute) temperatures. Specifically, note that since $\frac{dS_i}{dE_i} = \frac{1}{c+E_i}$, where $S_i = \log_e(c + E_i) - \log_e c$ denotes the unique continuously differentiable i -th subsystem entropy, it follows that $\frac{dS_i}{dE_i}$, $i = 1, \dots, q$, defines the reciprocal of the subsystem thermodynamic temperatures. That is,

$$\frac{1}{T_i} \triangleq \frac{dS_i}{dE_i} \quad (132)$$

and $T_i > 0$, $i = 1, \dots, q$. Hence, in our formulation, temperature is a function derived from entropy and does not involve the primitive subjective notions of hotness and coldness.

It is important to note that in this paper we view subsystem temperatures to be synonymous with subsystem energies. Even though this does not limit the generality of our theory from a mathematical perspective, it can be physically limiting since it does not allow for the consideration of two subsystems of $\mathcal{G}$ having the same stored energy with one of the subsystems being at a higher temperature (i.e., hotter) than the other. This, however, can be easily addressed by assigning different specific heats (i.e., thermal capacities) for each of the compartments of the large-scale system $\mathcal{G}$ as shown in [30].

7. Stochastic Semistability and Energy Equipartition

For the (adiabatically) isolated large-scale stochastic dynamical system $\mathcal{G}$, (95) yields the fundamental inequality

$$\mathbb{E}[\mathcal{S}(E(\tau_2))|\mathcal{F}_{\tau_1}] \geq \mathcal{S}(E(\tau_1)), \quad \tau_2 \stackrel{\text{a.s.}}{\geq} \tau_1. \quad (133)$$

Inequality (133) implies that, for any dynamical change in an adiabatically isolated large-scale stochastic dynamical system $\mathcal{G}$, the entropy of the final state can never be less than the entropy of the initial state and is a generalization of Clausius' version of the entropy principle, which states that for every irreversible (*nicht umkehrbar*) process in an adiabatically isolated system beginning and ending at an equilibrium state, the entropy of the final state is greater than or equal to the entropy of the initial state. Inequality (133) is often identified with the second law of thermodynamics for stochastic systems and gives as a statement about entropy increase. It is important to stress that this result holds for an adiabatically isolated dynamical system. It is, however, possible with power (heat flux) supplied from an external system to reduce the entropy of the dynamical system $\mathcal{G}$. The entropy of both systems taken together, however, cannot decrease.

As for the deterministic thermodynamic problem [30], this observation implies that when the isolated large-scale dynamical system $\mathcal{G}$ with thermodynamically consistent energy flow characteristics (i.e., Axioms (i) and (ii) hold) is at a state of maximum entropy consistent with its energy, it cannot be subject to any further dynamical change since any such change would result in a decrease of entropy. This of course implies that the state of *maximum entropy* is the stable state of an isolated system, and this equilibrium state has to be stochastically semistable. The following theorem generalizes Theorem 3.9 of [30] to the stochastic setting.

Theorem 10. Consider the large-scale stochastic dynamical system $\mathcal{G}$ with differential energy balance Equation (68) with $S(t) \stackrel{\text{a.s.}}{=} 0$, $D(E(t)) \stackrel{\text{a.s.}}{=} 0$, and $d(E(t)) \stackrel{\text{a.s.}}{=} 0$, and assume that Axioms (i) and (ii) hold. Then, for every $\alpha \geq 0$, $\alpha \mathbf{e}$ is a stochastic semistable equilibrium state of (68). Furthermore, $E(t) \xrightarrow{\text{a.s.}} \frac{1}{q} \mathbf{e} \mathbf{e}^T E(t_0)$ as $t \rightarrow \infty$ and $\frac{1}{q} \mathbf{e} \mathbf{e}^T E(t_0)$ is a semistable equilibrium state. Finally, if for some $k \in \{1, \dots, q\}$, $\sigma_{kk}(E) \geq 0$, $E \in \overline{\mathbb{R}}_+^q$, and $\sigma_{kk}(E) = 0$ if and only if $E_k = 0$ [63], then the zero solution $E(t) \equiv 0$ to (68) is a globally asymptotically stable in probability equilibrium state of (68).

Proof. It follows from Axiom (i) and (ii) that $\alpha \mathbf{e} \in \overline{\mathbb{R}}_+^q$, $\alpha \geq 0$, is an equilibrium state of (68). To show Lyapunov stability of the equilibrium state $\alpha \mathbf{e}$, consider $V(E) = \frac{1}{2}(E - \alpha \mathbf{e})^T(E - \alpha \mathbf{e})$ as a Lyapunov function candidate. Note that for $c \gg \max\{E_i, E_j\}$, $i \neq j$, $i, j = 1, \dots, q$,

$$\frac{c + E_i}{c + E_j} = \frac{1 + E_i/c}{1 + E_j/c} \approx 1. \quad (134)$$

Since, Axiom (ii) holds for all $c > 0$, we have

$$\sum_{j=1, j \neq i}^q (E_i - E_j) \phi_{ij}(E) \leq -\text{row}_i(J(E)) \text{row}_i^T(J(E)), \quad i = 1, \dots, q. \quad (135)$$

Now, since $\phi_{ij}(E) = -\phi_{ji}(E)$, $E \in \overline{\mathbb{R}}_+^q$, $i \neq j$, $i, j = 1, \dots, q$, $\mathbf{e}^T f(E) = 0$, $E \in \overline{\mathbb{R}}_+^q$, and $\mathbf{e}^T J(E) = 0$, $E \in \overline{\mathbb{R}}_+^q$, it follows from (135) that

$$\begin{aligned} \mathcal{L}V(E) &= (E - \alpha \mathbf{e})^T f(E) + \frac{1}{2} \text{tr}(J(E) J^T(E)) \\ &= E^T f(E) + \frac{1}{2} J(E) J^T(E) \\ &= \sum_{i=1}^q E_i \left[\sum_{j=1, j \neq i}^q \phi_{ij}(E) \right] + \frac{1}{2} \sum_{i=1}^q \text{row}_i(J(E)) \text{row}_i^T(J(E)) \\ &= \frac{1}{2} \sum_{i=1}^q \left[\sum_{j=1, j \neq i}^q (E_i - E_j) \phi_{ij}(E) + \text{row}_i(J(E)) \text{row}_i^T(J(E)) \right] \\ &\leq 0, \quad E \in \overline{\mathbb{R}}_+^q, \end{aligned} \quad (136)$$

which establishes Lyapunov stability in probability of the equilibrium state αe .

To show that αe is stochastically semistable, let $\mathcal{R} \triangleq \{E \in \mathbb{R}_+^q : \mathcal{L}V(E) = 0\}$. Now, by Axiom (i) and (ii) the directed graph associated with the connectivity matrix $\mathcal{C}$ for the large-scale dynamical system $\mathcal{G}$ is strongly connected, which implies that $\mathcal{R} = \{E \in \mathbb{R}_+^q : E_1 = \dots = E_q\}$. Since the $\mathbb{R}_+^q$ is an invariant set and $V(E)$ is radially unbounded, it follows from the Theorem 2 that for every initial condition $E(t_0) \in \mathbb{R}_+^q$, $E(t) \xrightarrow{\text{a.s.}} \mathcal{R}$ as $t \rightarrow \infty$, and hence, αe is a stochastic semistable equilibrium state of (68). Next, note that since $e^T E(t) = e^T E(t_0)$ and $E(t) \xrightarrow{\text{a.s.}} \mathcal{R}$ as $t \rightarrow \infty$, it follows that $E(t) \xrightarrow{\text{a.s.}} \frac{1}{q} ee^T E(t_0)$ as $t \rightarrow \infty$. Hence, with $\alpha = \frac{1}{q} e^T E(t_0)$, $\alpha e = \frac{1}{q} ee^T E(t_0)$ is a semistable equilibrium state of (68).

To show that in the case where for some $k \in \{1, \dots, q\}$, $\sigma_{kk}(E) \geq 0$, $E \in \mathbb{R}_+^q$, and $\sigma_{kk}(E) = 0$ if and only if $E_k = 0$, the zero solution $E(t) \equiv 0$ to (68) is globally asymptotically stable in probability, consider $V(E) = \frac{1}{2} E^T E$, $E \in \mathbb{R}_+^q$, as a candidate Lyapunov function. Note that $V(0) = 0$, $V(E) > 0$, $E \in \mathbb{R}_+^q$, $E \neq 0$, and $V(E)$ is radially unbounded. Now, the infinitesimal generator of Lyapunov function along the system energy trajectories of (68) is given by

$$\begin{aligned} \mathcal{L}V(E) &= E^T[f(E) - d(E)] + \frac{1}{2} \text{tr}J(E)J^T(E) \\ &= E^T f(E) + \frac{1}{2} \text{tr}J(E)J^T(E) - E_k \sigma_{kk}(E) \\ &= \sum_{i=1}^q E_i \left[\sum_{j=1, j \neq i}^q \phi_{ij}(E) \right] + \frac{1}{2} \sum_{i=1}^q \text{row}_i(J(E)) \text{row}_i^T(J(E)) - E_k \sigma_{kk}(E) \\ &= \frac{1}{2} \sum_{i=1}^q \left[\sum_{j=1, j \neq i}^q (E_i - E_j) \phi_{ij}(E) + \text{row}_i(J(E)) \text{row}_i^T(J(E)) \right] - E_k \sigma_{kk}(E) \\ &\leq 0, \quad E \in \mathbb{R}_+^q, \end{aligned} \quad (137)$$

which shows that the zero solution $E(t) \xrightarrow{\text{a.s.}} 0$ to (68) is Lyapunov stable in probability.

Finally, to show global asymptotic stability in probability of the zero equilibrium state, let $\mathcal{R} \triangleq \{E \in \mathbb{R}_+^q : \mathcal{L}V(E) = 0\}$. Now, since Axiom (i) holds and $\sigma_{kk}(E) = 0$ if and only if $E_k = 0$, it follows that $\mathcal{R} = \{E \in \mathbb{R}_+^q : E_k = 0, k \in \{1, \dots, q\}\} \cap \{E \in \mathbb{R}_+^q : E_1 = E_2 = \dots = E_q\} = \{0\}$. Hence, it follows from Theorem 2 that for every initial condition $E(t_0) \in \mathbb{R}_+^q$, $E(t) \xrightarrow{\text{a.s.}} \mathcal{R} = \{0\}$ as $t \rightarrow \infty$, which proves global asymptotic stability in probability of the zero equilibrium state of (68). $\square$

Theorem 10 shows that the isolated (i.e., $S(t) \xrightarrow{\text{a.s.}} 0$, $d(E) \xrightarrow{\text{a.s.}} 0$, and $D(E) \xrightarrow{\text{a.s.}} 0$) large-scale stochastic dynamical system $\mathcal{G}$ is stochastically semistable. In Theorem 10 we used the energy Lyapunov function to show that for the isolated (i.e., $S(t) \xrightarrow{\text{a.s.}} 0$, $d(E) \xrightarrow{\text{a.s.}} 0$, and $D(E) \xrightarrow{\text{a.s.}} 0$) large-scale stochastic dynamical system $\mathcal{G}$, $E(t) \xrightarrow{\text{a.s.}} \frac{1}{q} ee^T E(t_0)$ as $t \rightarrow \infty$ and $\frac{1}{q} ee^T E(t_0)$ is a stochastic semistable equilibrium state. This result can also be arrived at using the system entropy.

Specifically, using the system entropy given by (123), we can show attraction of the system trajectories to stochastic Lyapunov stable equilibrium points αe , $\alpha \geq 0$, and hence show stochastic semistability of these equilibrium states. To see this, note that since $e^T f(E) = 0$, $E \in \mathbb{R}_+^q$ and $e^T J(E) = 0$, $E \in \mathbb{R}_+^q$, it follows that $e^T dE(t) = 0$, $t \geq t_0$. Hence, $e^T E(t) \xrightarrow{\text{a.s.}} e^T E(t_0)$, $t \geq t_0$. Furthermore, since $E(t) \geq 0$, $t \geq t_0$, it follows that $0 \leq E(t) \leq ee^T E(t_0)$ a.s., $t \geq t_0$, which implies that all solutions to (68) are almost surely bounded.

Next, since by (125) the function $-S(E(t))$, $t \geq t_0$, is a supermartingale and $E(t)$, $t \geq t_0$, is bounded, it follows from Theorem 2 that for every initial condition $E(t_0) \in \mathbb{R}_+^q$, $E(t) \xrightarrow{\text{a.s.}} \mathcal{R}$ as $t \rightarrow \infty$, where $\mathcal{R} \triangleq \{E \in \mathbb{R}_+^q : \mathcal{L}S(E) = 0\}$. It now follows from the last inequality of (125) that $\mathcal{R} = \{E \in \mathbb{R}_+^q : (E_i - E_j) \phi_{ij}(E) = 0, i = 1, \dots, q, j \in \mathcal{K}_i\}$, which, since the directed graph associated with the connectivity matrix $\mathcal{C}$ for the large-scale dynamical system $\mathcal{G}$ is strongly connected, implies that $\mathcal{R} = \{E \in \mathbb{R}_+^q : E_1 = \dots = E_q\}$. Since the set $\mathcal{R}$ consists of the equilibrium states of (68),

it follows that $\mathcal{M} = \mathcal{R}$, which, along with (136), establishes stochastic semistability of the equilibrium states $\alpha \mathbf{e}$, $\alpha \geq 0$.

Theorem 10 implies that the steady-state value of the energy in each subsystem $\mathcal{G}_i$ of the isolated stochastic large-scale dynamical system $\mathcal{G}$ is equal, that is, the steady-state energy of the isolated large-scale stochastic dynamical system $\mathcal{G}$ given by

$$E_\infty \stackrel{\text{a.s.}}{=} \frac{1}{q} \mathbf{e} \mathbf{e}^T E(t_0) \stackrel{\text{a.s.}}{=} \left[\frac{1}{q} \sum_{i=1}^q E_i(t_0) \right] \mathbf{e} \quad (138)$$

is uniformly distributed over all subsystems of $\mathcal{G}$. This phenomenon is known as *equipartition of energy* [64–68] and is an emergent behavior in thermodynamic systems [30].

Example 2. In this example, we apply Theorem 10 to the five-compartment thermodynamic system shown in Figure 2. Specifically, consider

$$dE_1(t) = [E_2(t) - E_1(t)]dt + \gamma[E_2(t) - E_1(t)]dw(t), \quad E_1(0) \stackrel{\text{a.s.}}{=} E_{10}, \quad t \geq 0, \quad (139)$$

$$dE_2(t) = [E_1(t) - E_2(t) + E_3(t) - E_2(t) + E_5(t) - E_2(t)]dt + \gamma[E_1(t) - E_2(t) + E_3(t) - E_2(t) + E_5(t) - E_2(t)]dw(t), \quad E_2(0) \stackrel{\text{a.s.}}{=} E_{20}, \quad (140)$$

$$dE_3(t) = [E_2(t) - E_3(t) + E_4(t) - E_3(t)]dt + \gamma[E_2(t) - E_3(t) + E_4(t) - E_3(t)]dw(t), \quad E_3(0) \stackrel{\text{a.s.}}{=} E_{30}, \quad (141)$$

$$dE_4(t) = [E_3(t) - E_4(t)]dt + \gamma[E_3(t) - E_4(t)]dw(t), \quad E_4(0) \stackrel{\text{a.s.}}{=} E_{40}, \quad (142)$$

$$dE_5(t) = [E_2(t) - E_5(t)]dt + \gamma[E_2(t) - E_5(t)]dw(t), \quad E_5(0) \stackrel{\text{a.s.}}{=} E_{50}. \quad (143)$$

Note that (139)–(143) can be cast in the form of (68) with $E \triangleq [E_1, E_2, E_3, E_4, E_5]^T$, $d(E) = 0$, $D(E) = 0$, $S(t) \equiv 0$, $w_1 = w$,

$$f(E) = \begin{bmatrix} E_2 - E_1 \\ E_1 - E_2 + E_3 - E_2 + E_5 - E_2 \\ E_2 - E_3 + E_4 - E_3 \\ E_3 - E_4 \\ E_2 - E_5 \end{bmatrix}, \quad J(E) = \gamma \begin{bmatrix} E_2 - E_1 \\ E_1 - E_2 + E_3 - E_2 + E_5 - E_2 \\ E_2 - E_3 + E_4 - E_3 \\ E_3 - E_4 \\ E_2 - E_5 \end{bmatrix}.$$

It follows from Theorem 10 that the thermodynamic heat flow model (139)–(143) is stochastically semistable with respect to $\mathbb{R}_+^5$ and achieves energy equipartition. To see this, let $E_{10} = 0$, $E_{20} = 10$, $E_{30} = 20$, $E_{40} = 30$, $E_{50} = 40$, and $\gamma = 0.2$. Figure 3 shows the sample energy trajectories along with the standard deviation of the states of each thermodynamic compartment versus time for 10 sample paths.

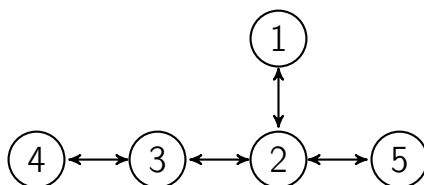


Figure 2. Thermodynamic model with undirected heat flow.

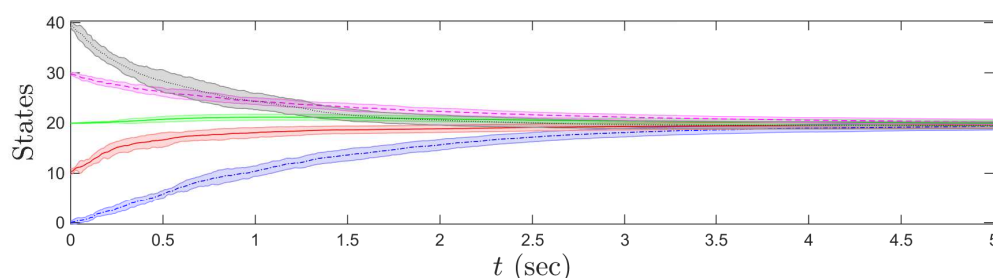


Figure 3. Sample average along with the sample standard deviation of the system energies versus time; $E_1(t)$ in blue, $E_2(t)$ in red, $E_3(t)$ in green, $E_4(t)$ in magenta, and $E_5(t)$ in black.

8. Conclusions

In this paper, we combined thermodynamics and stochastic dynamical system theory to provide a system-theoretic foundation of thermodynamics. The proposed dynamical systems framework of thermodynamics can potentially provide deeper insights into the constitutive mechanisms that explain fundamental thermodynamic processes and describe acute microcosms and macrocosms in the ever-elusive pursuit of unifying the subatomic and astronomical domains. In future research, we will use the realizations of each sample path of the stochastic energy variables characterized by the stochastic differential energy balance dynamical model to describe the probability density function of our large-scale stochastic thermodynamic model by a continuous-time and continuous-space Fokker–Planck evolution equation to give a thermodynamic interpretation between the stationary solution of the Fokker–Planck equation and the canonical thermodynamic equilibrium distribution. Furthermore, since our stochastic thermodynamic model does not restrict the second law on each individual sample path trajectory, we will explore the second law as a fluctuation theorem to give a precise prediction of the cases in which the system entropy decreases over a given time interval for our model.

Acknowledgments: This research was supported in part by the Air Force Office of Scientific Research under Grant FA9550-16-1-0100.

Author Contributions: Both authors contributed equally to this work.

Conflicts of Interest: The authors declare no conflict of interest.

References and Notes

1. Sekimoto, K. Kinetic characterization of heat bath and the energetics of thermal ratchet models. *J. Phys. Soc. Jpn.* **1997**, *66*, 1234–1237.
2. Sekimoto, K. Langevin equation and thermodynamics. *Prog. Theor. Phys. Supp.* **1998**, *130*, 17–27.
3. Sekimoto, K. *Stochastic Energetics*; Springer: Berlin, Germany, 2010.
4. Seifert, U. Stochastic thermodynamics: Principles and perspectives. *Eur. Phys. J. B* **2008**, *64*, 423–431.
5. Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* **2012**, *75*, 1–58.
6. Onsager, L. Reciprocal relations in irreversible processes, I. *Phys. Rev.* **1931**, *37*, 405–426.
7. Onsager, L. Reciprocal relations in irreversible processes, II. *Phys. Rev.* **1932**, *38*, 2265–2279.
8. De Groot, S.R. *Thermodynamics of Irreversible Processes*; North-Holland: Amsterdam, The Netherlands, 1951.
9. Prigogine, I. *Thermodynamics of Irreversible Processes*; Interscience: New York, NY, USA, 1955.
10. Einstein, A. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. Phys.* **1905**, *322*, 549–560.
11. Even though Robert Brown [13] is credited with the discovery of Brownian motion, the chaotic motion of atoms was first conjectured by Leukippos—the ancient Greek philosopher to first develop the theory of atomism. In Book II of his poem *De Rerum Natura* (*On the Nature of the Universe*) Lucretius (99–55 B.C.) attributes the observed disordered motion of dust lit by a sunray to Leukippos. He goes on to state that

- Leukippos asserted that the irregular and extremely fast motion of atoms is the cause for the slower motion of the larger dust particles ([12], p. 23).
12. Russo, L. *The Forgotten Revolution: How Science was Born in 300 B.C. and Why it Had to be Reborn*; Springer: Berlin, Germany, 2004.
 13. Brown, R. A brief account of microscopical observations made in the months of June, July, and August, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *Philos. Mag.* **1827**, *4*, 161–173.
 14. Jüttner, F. Das Maxwellsche Gesetz der Geschwindigkeitsverteilung in der Relativtheorie. *Ann. Phys.* **1911**, *34*, 856–882.
 15. Smoluchowski, M. Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen. *Ann. Phys.* **1906**, *21*, 756–780.
 16. Langevin, P. Sur la théorie du mouvement Brownien. *C. R. Acad. Sci. Paris* **1908**, *146*, 530–533.
 17. It is important to note here that Brownian motion is not a Markov process on arbitrary time scales since the Markov assumption (i.e., the conditional probability distribution of the future states only depend on the current state) does not hold for the detailed dynamics of the Brownian particle collisions. However, the outcome of each Brownian particle collision only depends on the initial condition of the collision, and hence, on the most recent collision, which is precisely the Markov assumption. Thus, on a time scale of the same order of magnitude as the mean time between particle collisions (i.e., the Markov-Einstein time scale), Brownian motion can be assumed to be Markovian.
 18. Bochkov, G.N.; Kuzovlev, Y.E. General theory of thermal fluctuations in nonlinear systems. *J. Exp. Theor. Phys.* **1977**, *45*, 125–130.
 19. Bochkov, G.N.; Kuzovlev, Y.E. Fluctuation-dissipation relations for nonequilibrium processes in open systems. *J. Exp. Theor. Phys.* **1979**, *49*, 543–551.
 20. Gallavotti, G.; Cohen, E.G.D. Dynamical ensembles in nonequilibrium statistical mechanics. *Phys. Rev. Lett.* **1995**, *74*, 2694–2697.
 21. Kurchan, J. Fluctuation theorem for stochastic dynamics. *J. Phys. A* **1998**, *31*, 3719–3729.
 22. Lebowitz, J.L.; Spohn, H. A Gallavotti-Cohen-type symmetry in the large deviation functional for stochastic dynamics. *J. Stat. Phys.* **1999**, *95*, 333–365.
 23. Evans, D.J.; Searles, D.J. Equilibrium microstates which generate second law violating steady states. *Phys. Rev. E* **1994**, *50*, 1645–1648.
 24. Jarzynski, C. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **1997**, *78*, 2690–2693.
 25. Jarzynski, C. Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. *Phys. Rev. E* **1997**, *56*, 5018–5035.
 26. Crooks, G.E. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E* **1999**, *60*, 2721–2726.
 27. Crooks, G.E. Path-ensemble averages in systems driven far from equilibrium. *Phys. Rev. E* **2000**, *61*, 2361–2366.
 28. Hummer, G.; Szabo, A. Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 3658–3661.
 29. Van Kampen, N.G. *Stochastic Processes in Physics and Chemistry*; Elsevier: Amsterdam, The Netherlands, 1992.
 30. Haddad, W.M.; Chellaboina, V.; Nersisov, S.G. *Thermodynamics: A Dynamical Systems Approach*; Princeton University Press: Princeton, NJ, USA, 2005.
 31. Khasminskii, R.Z. *Stochastic Stability of Differential Equations*; Springer: Berlin, Germany, 2012.
 32. Arnold, L. *Stochastic Differential Equations: Theory and Applications*; Wiley-Interscience: New York, NY, USA, 1974.
 33. Oksendal, B. *Stochastic Differential Equations: An Introduction with Applications*; Springer: Berlin, Germany, 1995.
 34. Mao, X.; Yuan, C. *Stochastic Differential Equations with Markovian Switching*; Imperial College Press: London, UK, 2006.
 35. Mao, X. *Stochastic Differential Equations and Applications*; Harwood: New York, NY, USA, 1997.
 36. The $\mathbb{P}$ -outer measure on a set Ω is an isotone (i.e., order-preserving), countably additive, extended real-valued set function defined for all subsets of Ω with $\mathbb{P}(\emptyset) = 0$.
 37. Rajpurohit, T.; Haddad, W.M. Lyapunov and converse Lyapunov theorems for stochastic semistability. *Syst. Control Lett.* **2016**, *97*, 83–90.
 38. Gard, T.C. *Introduction to Stochastic Differential Equations*; Marcel Dekker: New York, NY, USA, 1988.
 39. Itô, K. Differential equations determining Markov processes. *Zenkoku Shijo Sugaku Danwaki* **1942**, *1077*, 1352–1400.

40. Itô, K. Stochastic integral. *Proc. Imp. Acad. Tokyo* **1944**, *20*, 519–524.
41. Wu, Z.J.; Xie, X.J.; Shi, P.; Xia, Y.Q. Backstepping controller design for a class of stochastic nonlinear systems with Markovian switching. *Automatica* **2009**, *45*, 997–1004.
42. Brzezniak, Z.; Capinski, M.; Flandoli, F. Pathwise global attractors for stationary random dynamical systems. *Prob. Theory Relat. Fields* **1993**, *95*, 87–102.
43. Crauel, H.; Flandoli, F. Attractors of random dynamical systems. *Prob. Theory Relat. Fields* **1994**, *100*, 365–393.
44. Crauel, H.; Debussche, A.; Flandoli, F. Random attractors. *J. Dyn. Differ. Equ.* **1997**, *9*, 307–341.
45. The support of a function is the unique smallest closed set for which the complement of the set has measure zero. A function has a compact support if its support is a compact set.
46. Haddad, W.M.; Chellaboina, V.; Hui, Q. *Nonnegative and Compartmental Dynamical Systems*; Princeton University Press: Princeton, NJ, USA, 2010.
47. Haddad, W.M.; Chellaboina, V. *Nonlinear Dynamical Systems and Control: A Lyapunov-Based Approach*; Princeton University Press: Princeton, NJ, USA, 2008.
48. Khalil, H.K. *Nonlinear Systems*, 3rd ed.; Prentice-Hall: Upper Saddle River, NJ, USA, 2002.
49. Meyn, S.P.; Tweedie, R.L. *Markov Chains and Stochastic Stability*; Springer: London, UK, 1993.
50. Mao, X. Stochastic versions of the LaSalle theorem. *J. Differ. Equ.* **1999**, *153*, 175–195.
51. Liptser, R.S.; Shiriyayev, A.N. *Theory of Martingales*; Kluwer Academic Publishers: Norwell, MA, USA, 1989.
52. Rajpurohit, T.; Haddad, W.M. Lyapunov and converse Lyapunov theorems for stochastic semistability. In Proceedings of the 55th IEEE Conference on Decision and Control, Las Vegas, NV, USA, 10–13 December 2016; pp. 83–90.
53. Arapostathis, A.; Borkar, V.S.; Ghosh, M.K. *Ergodic Control of Diffusion Processes*; Cambridge University Press: Cambridge, UK, 2012.
54. Rajpurohit, T.; Haddad, W.M. Dissipativity theory for nonlinear stochastic dynamical systems. *IEEE Trans. Autom. Control* **2017**, *62*, 1684–1699.
55. Lévy, P.P. *Théorie de L'addition des Variables Aléatoires*; Gauthier-Villars: Paris, France, 1937.
56. Arnold, V.I. *Mathematical Models of Classical Mechanics*; Springer: New York, NY, USA, 1989.
57. It can be argued here that a more appropriate terminology is *assumptions* rather than *axioms* since, as will be seen, these are statements taken to be true and used as premises in order to infer certain results, but may not otherwise be accepted. However, as we will see, these statements are equivalent (within our formulation) to the stipulated postulates of the zeroth and second laws of thermodynamics involving transitivity of a thermal equilibrium and heat flowing from *hotter* to *colder* bodies, and as such we refer to them as *axioms*.
58. Berman, A.; Plemmons, R.J. *Nonnegative Matrices in the Mathematical Sciences*; Academic: New York, NY, USA, 1979.
59. It is important to note that our formulation of the second law of thermodynamics as given by Axiom (ii) does not require the mentioning of temperature nor the more primitive subjective notions of hotness or coldness. As we will see later, temperature is defined in terms of the system entropy after we establish the existence of a unique, two-times continuously differentiable entropy function for $\mathcal{G}$.
60. Since in this section we are not considering work performed by and on the system, the notions of an *isolated* system and an *adiabatically isolated* system are equivalent.
61. Meixner, J. On the foundation of thermodynamics of processes. In *A Critical Review of Thermodynamics*; Stuart, E.B., Gal-Or, B., Brainard, A.J., Eds.; Mono Book Corp.: Baltimore, MD, USA, 1970; pp. 37–47.
62. Lavenda, B. *Thermodynamics of Irreversible Processes*; Macmillan: London, UK, 1978; Dover: New York, NY, USA, 1993.
63. The assumption $\sigma_{kk}(E) \geq 0$, $E \in \mathbb{R}_+^q$, and $\sigma_{kk}(E) = 0$ if and only if $E_k = 0$ for some $k \in \{1, \dots, q\}$ implies that if the k -th subsystem possesses no energy, then this subsystem cannot dissipate energy to the environment. Conversely, if the k -th subsystem does not dissipate energy to the environment, then this subsystem has no energy.
64. Lyon, R.H. *Statistical Energy Analysis of Dynamical Systems: Theory and Applications*; MIT Press: Cambridge, MA, USA, 1975.
65. Bernstein, D.S.; Hyland, D.C. Compartmental modeling and second-moment analysis of state space systems. *SIAM J. Matrix Anal. Appl.* **1993**, *14*, 880–901.

66. Bernstein, D.S.; Bhat, S.P. Energy equipartition and the emergence of damping in lossless systems. In Proceedings of the 41st IEEE Conference on Decision and Control, Las Vegas, NV, USA, 10–13 December 2002; pp. 2913–2918.
67. Pearson, R.K.; Johnson, T.L. Energy equipartition and fluctuation-dissipation theorems for damped flexible structures. *Q. Appl. Math.* **1987**, *45*, 223–238.
68. Hall, S.R.; MacMartin, D.G.; Bernstein, D.S. Covariance averaging in the analysis of uncertain systems. *IEEE Trans. Autom. Control* **1992**, *38*, 1842–1859.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).