SIAM J. Applied Dynamical Systems
Vol. 12, No. 2, pp. 1031–1084

Dynamics of Stochastically Blinking Systems. Part II: Asymptotic Properties*

Martin Hasler†, Vladimir Belykh‡, and Igor Belykh§

Abstract. We study stochastically blinking dynamical systems as in the companion paper (Part I). We analyze the asymptotic properties of the blinking system as time goes to infinity. The trajectories of the averaged and blinking system cannot stick together forever, but the trajectories of the blinking system may converge to an attractor of the averaged system. There are four distinct classes of blinking dynamical systems. Two properties differentiate them: single or multiple attractors of the averaged system and their invariance or noninvariance under the dynamics of the blinking system. In the case of invariance, we prove that the trajectories of the blinking system converge to the attractor(s) of the averaged system with high probability if switching is fast. In the noninvariant single attractor case, the trajectories reach a neighborhood of the attractor rapidly and remain close most of the time with high probability when switching is fast. In the noninvariant multiple attractor case, the trajectory may escape to another attractor with small probability. Using the Lyapunov function method, we derive explicit bounds for these probabilities. Each of the four cases is illustrated by a specific example of a blinking dynamical system. From a probability theory perspective, our results are obtained by directly deriving large deviation bounds. They are more conservative than those derived by using the action functional approach, but they are explicit in the parameters of the blinking system.

Key words. blinking networks, stochastic switching, averaging, attractor

AMS subject classifications. 34D05, 34D06, 34D45, 37H10, 37H20, 34D10

DOI. 10.1137/120893410

1. Introduction.

This paper focuses on a largely unexplored area, namely, mathematical analysis and modeling of dynamical systems and networks whose coupling or internal parameters stochastically evolve over time. Networks of dynamical systems are common models for many systems in physics, engineering, chemistry, biology, and social sciences [1, 2, 3]. Recently, a great deal of attention has been paid to algebraic, statistical, and graph theoretical properties of networks and their relationship to the dynamical properties of the underlying network (see, for example, [1, 2, 3, 4, 5, 6, 7, 8, 9] and the references therein). In most studies, network connectivity is assumed to be static. However, in many realistic networks the coupling strength or the connection topology can vary in time, according to a dynamical rule, whether deterministic or stochastic. Researchers are only now beginning to investigate the

*Received by the editors October 1, 2012; accepted for publication (in revised form) by B. Sandstede March 29, 2013; published electronically June 20, 2013.
http://www.siam.org/journals/siads/12-2/89341.html

†School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland (martin.hasler@epfl.ch). This author's work was supported by the Swiss National Science Foundation.

‡Mathematics Department, Volga State Academy, Nizhny Novgorod, 603 600 Russia (belykh@unn.ac.ru). This author's work was supported by RFFI grant 12-01-00694.

§Corresponding author. Department of Mathematics and Statistics, Georgia State University, Atlanta, GA 30303 (ibelykh@gsu.edu). This author's work was supported by the National Science Foundation under grant DMS-1009744.

link between the time-evolving structure and the overall dynamics of a system or a network [10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20].

In many engineering and biological networks, the individual nodes that compose the network interact only sporadically via short on-off interactions. Packet switched networks such as the Internet are an important example. To model realistic networks with intermittent connections, we previously introduced a class of dynamical networks with fast on-off connections that were called "blinking" networks [11, 12]. These networks are composed of oscillatory dynamical systems with connections that switch on and off randomly; the switching time is fast, with respect to the characteristic time of the individual node dynamics. In [11], we proved that global synchronization occurs almost surely in a blinking network, provided that coupling strengths are strong enough and the switching time of blinking connections is fast. Similar results for synchronization in on-off fast switching networks were also obtained in [13, 14, 15].

In this paper, we go beyond network synchronization and consider a general stochastically switched dynamical system. We develop a general rigorous theory of stochastically switched dynamical systems and networks and apply rigorous mathematical techniques to investigate the interplay between overall system dynamics and the stochastic switching process.

As in our companion paper [21], we consider a class of dynamical systems whose parameters are switched within a discrete set of values at equal time intervals. Similarly to the blinking of the eye, switching is fast and occurs stochastically and independently for different time intervals. Such blinking systems have two characteristic times, namely the characteristic time of the individual dynamical system and the characteristic time scale of the fast stochastic process. When it comes to a network, this stochastic process defines stochastic switchings of network connections. If the stochastic switching is fast enough, we expect the blinking system to follow the averaged system where the dynamical law is given by the expectation of the stochastic variables.

The fact that the rapidly switched blinking system has the same behavior as the averaged system seems apparent; however, there are exceptions. Therefore a careful proof of the property is needed which shows what parameters the occurrence of the exceptions depends on. In fact, the assumption that a trajectory of the blinking system can follow that of the averaged systems can be true only for finite time, unless there is a mechanism that forces them to stick together. Such a mechanism is present when the solutions of the averaged system converge towards an attractor.

While averaging is a classical technique in the study of nonlinear oscillators [26, 27, 28, 29, 30, 31, 37], research in averaging for blinking systems is more recent [32, 33, 34, 35, 36]. While the application that these works implicitly address is dynamical systems perturbed by noise, we rather target randomly switched networks. However, the mathematical techniques that are used apply, of course, to both application areas. Generally speaking, the crucial discipline of probability theory involved is *large deviations*. It concerns the exponentially fast decay of rare events in stochastic processes. Using the notion of an action functional, it is in principle possible to determine the exact exponential rate of decay of such events. However, in the context of dynamical systems, this involves the solution of a variational problem [34, 35]. We pursue in this paper a simpler approach. It has the advantage that all results are perfectly explicit in the dynamical systems parameters, namely bounds and Lipschitz constants for the right-hand side of the state equations of the blinking and the averaged system, and the rate of decreasing of the Lyapunov function for the averaged system. Except, perhaps, for very special cases, such results are not obtainable by the action functional approach. On the other hand, our techniques do not yield the actual rates of exponential decay of the rare events but only lower bounds on them. We used such techniques for synchronization of blinking dynamical networks [11] and for the convergence of the blinking network to an attractor [12]. In this paper, we shall use a somewhat different approach.

In the companion paper [21] we have shown that if the blinking system switches fast enough, i.e., if the switching period is sufficiently short, then a solution of the blinking system closely follows the solution of the averaged system for a certain time and afterwards they usually drift apart. We derived explicit bounds that relate the probability, the switching frequency, the precision, and the length of the time interval to each other. We discovered the presence of a soft upper bound for the time interval beyond which it is almost impossible to keep the two trajectories together.

In this paper, we address the question of how the solutions of the blinking and the averaged systems are related asymptotically, when $t \to +\infty$. More precisely, we ask the question under what conditions a solution of the blinking system converges to an attractor of the averaged system. The answer contains various subtleties, and it turns out that essentially four cases have to be distinguished, depending on whether or not the attractor in the averaged system is unique and whether it is an invariant set, attracting or not, for all switching sequences. We introduce them through numerical analysis of four corresponding examples. After that, using the technique developed in [21] combined with the Lyapunov function method, we prove four general theorems that describe the behavior of the blinking system in the four cases. More specifically, in the case of invariance, where the attractor of the averaged system is invariant under the blinking system, we prove that the trajectories of the blinking system converge to the attractor of the averaged system with high probability if switching is fast. In the noninvariant single attractor case, the trajectories of the blinking system reach a neighborhood of the attractor rapidly and remain close most of the time with high probability when switching is fast. In this case, the attractor of the averaged system acts as a *ghost* attractor for the blinking system. In the noninvariant multiple attractor case, the trajectory may escape to another attractor with small probability.

In the literature, a number of works can be found that study systems similar to our blinking system. Typically, they are more general in one direction and more restrictive in another and the results are usually less detailed or of a different nature. We mention here just a few examples as illustrations. In [41], randomly switched dynamical systems are considered, but switching is not necessarily fast with respect to the time scale of the dynamical systems. In addition, switching times are also random and there is perturbation by noise. The origin in state space is always a solution and therefore also a solution of the averaged system. This is one of the cases we consider, but for a less general system. The global asymptotic stability of the origin is proved by a Lyapunov function that by hypothesis almost surely decreases along solutions, even when taking into account the noise. This is much more restrictive than what is required in this paper. In [42], systems similar to those of [41], but with deterministic time dependence and deterministic perturbations, are studied. In [43], fast time-varying deterministic systems with the zero equilibrium point are considered and it is shown that for sufficiently rapid time dependence, the exponential stability of the equilibrium point in the averaged system implies the exponential stability of the equilibrium point in the time-varying system. In [44], Benaïm et al. studied fast random switching with random switching times between two linear asymptotically stable two-dimensional systems when the averaged system is unstable. This loosely corresponds to our case where the attractor of the averaged system (infinity) is invariant under the switched system, and indeed it is shown that for fast enough switching almost all trajectories diverge to infinity. However, no estimate of the rate of convergence is given. In the present paper, the asymptotic behavior of randomly switched dynamical systems is described by convergence theorems, bounds on speed of convergence, and bounds on permanence times in the vicinity of an attractor. Another characterization is invariant probability densities. Two recent papers that address the existence of invariant densities and their properties are [45] for continuous time and [46] for discrete time.

The layout of this paper is as follows. First, in section 2, we briefly describe the blinking model and the corresponding averaged system. Then, in sections 3, 4, 5, and 6, we present and numerically study four examples for the four distinct classes of blinking systems. These are (i) a blinking network of coupled Lorenz systems where connections are stochastically switched on and off (section 3); (ii) two bistable systems coupled by a blinking connection (section 4); (iii) a stochastically switched power converter (section 5); (iv) an information processing cellular neural network with blinking shortcuts (section 6). Then, in section 7, we make basic assumptions on the dynamics of the blinking systems and its averaged analogue. Sections 8, 9, 10, and 11 present four main theorems and their proofs. Finally, a brief discussion of the obtained results is given.

2. The blinking model.

We introduce the system only briefly in this study. More details can be found in the companion paper [21]. The blinking system is described by N time-dependent ordinary differential equations of the form

$$\frac{dx}{dt} = F\left(x\left(t\right), s\left(t\right)\right), \qquad x \in \mathbb{R}^N, \quad F : \mathbb{R}^{N+M} \to \mathbb{R}^N, \quad s\left(t\right) \in \left\{0,1\right\}^M, \tag{2.1}$$

where the function $s(t)$ is piecewise constant, taking the constant binary vector value $s^k = \left(s_1^k, \ldots, s_{\mathrm{M}}^k\right)$ in the time interval $t \in [(k-1)\,\tau,\, k\tau)$. The sequence of binary vectors s^k, $k = 1, 2, \ldots$, is called the switching sequence, as each component s_i^k of s^k switches on ($s_i^k = 1$) or off ($s_i^k = 0$) during the kth time interval. The switching sequences are assumed to be the instances of the stochastic process S^k, $k = 1, 2, \ldots$, where all random vectors S^k are independent and identically distributed, taking the value $s \in \{0,1\}^M$ with probability p_s.

In this paper, we study the asymptotic behavior of solutions of (2.1) as time goes to infinity. System (2.1) inherently has two time scales, the switching period τ and the time scale of the dynamics of the nonswitched system where the vector s is kept constant. We limit our attention to the case where switching is fast with respect to the time scale of the nonswitched system dynamics. In this case, one can expect that the dynamics of the stochastic blinking system (2.1) is close to that of the averaged system where the dynamical law is simply averaged over the driving stochastic variables $s^k(t)$ at each time instant.

The averaged system associated with the blinking system (2.1) reads as

$$\frac{dx}{dt} \;=\; \Phi\left(x\left(t\right)\right), \tag{2.2}$$

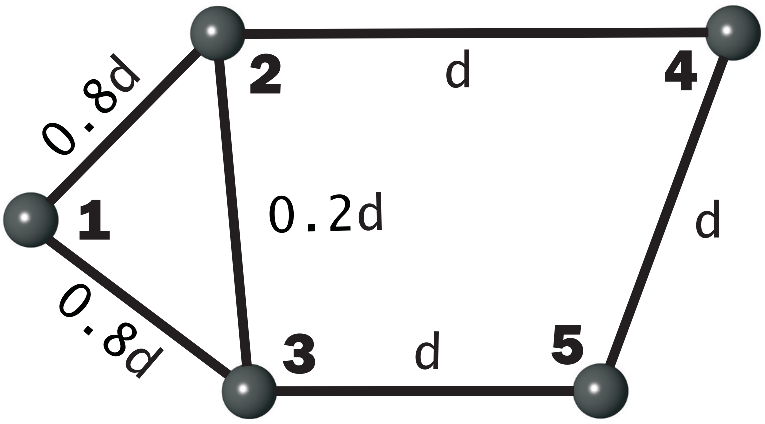

Figure 1. *First example: coupling graph of five coupled Lorenz systems.*

where

$$\Phi(x) = \mathrm{E}(F(x,S)) = \sum_{s\in\{0,1\}^{\mathrm{M}}} F(x,s)\, p_s \tag{2.3}$$

and $\mathrm{E}(F(x,S))$ is the expected value of $(F(x,S))$. In (2.3) we have omitted the upper index k for the switching variables since in each time interval they have the same probability distribution.

Blinking systems can be found in various applications. The four examples given below can give a first idea. It is worth noticing that these examples represent four distinct classes of blinking dynamical systems. Two properties differentiate them: single or multiple attractors of the averaged system and their invariance or noninvariance under the dynamics of the blinking system.

3. First example: Synchronization of chaotic systems.

Consider the network of five chaotic Lorenz systems, diffusively coupled with coupling strengths proportional to the constant $d > 0$ according to Figure 1. Instead of connecting the individual systems permanently, they are stochastically switched on and off, as described in section 2. The ith Lorenz system is described by the three ordinary differential equations

$$\begin{aligned} \dot{x}_i &= \sigma(y_i - x_i) &&= G_1(\mathbf{x}_i), \\ \dot{y}_i &= rx_i - y_i - x_i z_i &&= G_2(\mathbf{x}_i), \\ \dot{z}_i &= -bz_i + x_i y_i &&= G_3(\mathbf{x}_i), \end{aligned}$$

where $\mathbf{x}_i = (x_i, y_i, z_i)$. We choose the standard parameter values $b = 8/3$, $r = 28$, and $s = 10$, which guarantee chaotic behavior.

If we couple the first state variables of the five Lorenz systems in the blinking mode according to Figure 1, we obtain the system of equations

$$
\begin{aligned}
\dot{x}_1 &= G(x_1) + 0.8ds_1(t)P(x_2 - x_1) + 0.8ds_2(t)P(x_3 - x_1),\\
\dot{x}_2 &= G(x_2) + 0.8ds_1(t)P(x_1 - x_2) + 0.2ds_3(t)P(x_3 - x_2) + ds_4(t)P(x_4 - x_2),\\
\dot{x}_3 &= G(x_3) + 0.8ds_2(t)P(x_1 - x_3) + 0.2ds_3(t)P(x_2 - x_3) + ds_5(t)P(x_5 - x_3),\\
\dot{x}_4 &= G(x_4) + ds_4(t)P(x_2 - x_4) + ds_6(t)P(x_5 - x_4),\\
\dot{x}_5 &= G(x_5) + ds_5(t)P(x_3 - x_5) + ds_6(t)P(x_4 - x_5),
\end{aligned} \tag{3.1}
$$

where $G = \{G_1, G_2, G_3\}$ and P is the projection operator onto the first state variable

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

This blinking system has the form (2.1) with $N = 15$ and $M = 6$. During each time interval of length τ, each of the six edges (cf. Figure 1), defined by the switching variables $s_1(t), \dots, s_6(t)$, is turned on with probability p, independently of the switching on and off of the other edges, and independently of whether or not it has been turned on during the previous time interval.

The averaged system (2.2) associated with the blinking system (3.1) is obtained by replacing all switching variables, $s_1(t), \dots, s_6(t)$, by their mean value p. In Figure 1 this amounts to replacing d by pd.

The question we are interested in is whether the blinking system synchronizes when the averaged system does. Synchronization can be interpreted as convergence to the diagonal subspace $D = \{x_1 = x_2 = \cdots = x_5\}$. So the question can be reformulated as follows: if the solutions of the averaged system converge to the diagonal subspace D, is the same true for the solutions of the blinking system?

Applying the connection graph stability method [5], we can conclude that the averaged system synchronizes if the coupling coefficient pd is large enough. In this case, our previous analysis [11] guarantees that the blinking system synchronizes with high probability if the switching time is small enough. An explicit and rigorous upper bound on the switching time τ for synchronization in a blinking network of Lorenz systems was given in [11].

In order to illustrate these results, we introduce a measure of the synchronization error

$$V(x_1, \dots, x_5) = \sqrt{\frac{1}{30} \sum_{i=1}^{5} \sum_{j=i+1}^{5} \left((x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2 \right)},$$

i.e., the average deviation of the same component of a state in one Lorenz system from the same component of the state of another Lorenz system.

We now present the results of some numerical simulations. Figure 2 indicates, at least for the given switching sequence of the blinking system, that both the averaged and the blinking systems synchronize (cf. the upper panel of Figure 2). In other words, they converge to the same attracting set, the diagonal manifold D, but they are not close to each other. In the lower panel of Figure 2 the synchronization error V, as a function of time, is represented for the same solutions of the averaged system and the same instance of the blinking system as in the upper panel. Once again, this indicates that both synchronize. Furthermore, synchronization

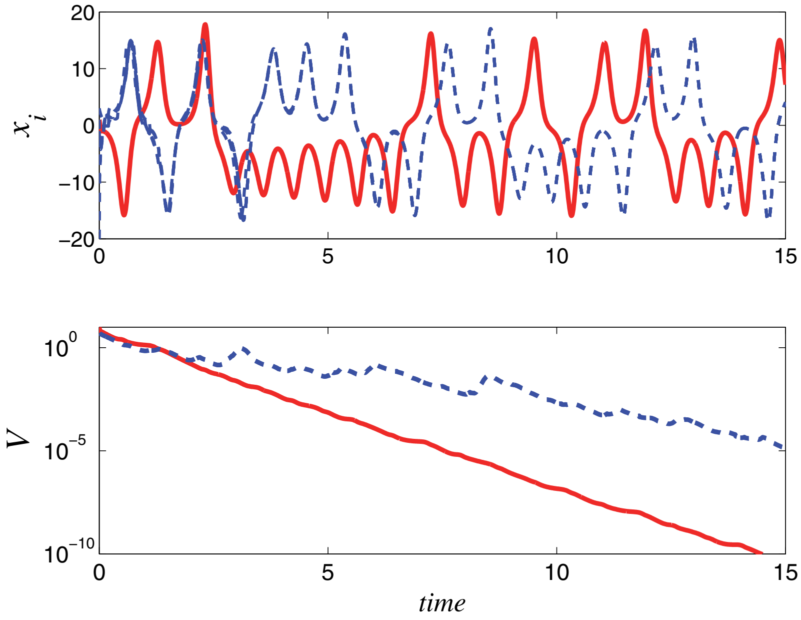

Figure 2. *Upper panel: x_1-coordinate of the first Lorenz system in the averaged system (red solid curve) and x_i-coordinates of all five blinking Lorenz systems (blue dashed curve) as functions of time. The blinking system and the averaged system start from the same nonhomogeneous initial conditions $x_1 \neq \cdots \neq x_5$. The x_i-components of the other Lorenz systems in the averaged system would be indistinguishable due to fast synchronization and are not shown. Lower panel: Synchronization error V as a function of time for the averaged system (red solid curve) and the blinking system (blue dashed curve). Synchronization is exponentially fast, but the convergence of the blinking system to the synchronized solution is slower than that of the averaged system. Parameters are the switching probability $p = 1/2$, switching time $\tau = 0.1$, and coupling parameter $d = 200$.*

appears to be exponentially fast, but the exponential speed of synchronization is smaller in the case of the blinking system as compared to the averaged system.

Let us remark that the diagonal subspace D is not really an attractor of the averaged system, but on the diagonal subspace all solutions are identical to the solutions of a single Lorenz system. Similarly, the network solutions of both the averaged and the blinking systems converge to the Lorenz attractor in the diagonal subspace.

Figure 3 indicates that the solution of the averaged system as well as the solution of the blinking system for $\tau = 0.1$ both synchronize exponentially fast, the blinking system being slower than the averaged system. This repeats the findings of Figure 2 for $\tau = 0.1$, but on a longer time scale. The deviation from synchronization of the solution of the blinking system, after having reached about 10^{-15}, irregularly oscillates between this value and about 10^{-10}, the precision of the numerical integration of the differential equations. This can be attributed to numerical errors.

For $\tau = 1$, after a much longer transient phase, the solution of the blinking system also appears to synchronize, whereas for $\tau = 5$ synchronization appears to be lost. This illustrates that somewhere there probably is a threshold for synchronization between $\tau = 1$ and $\tau = 5$.

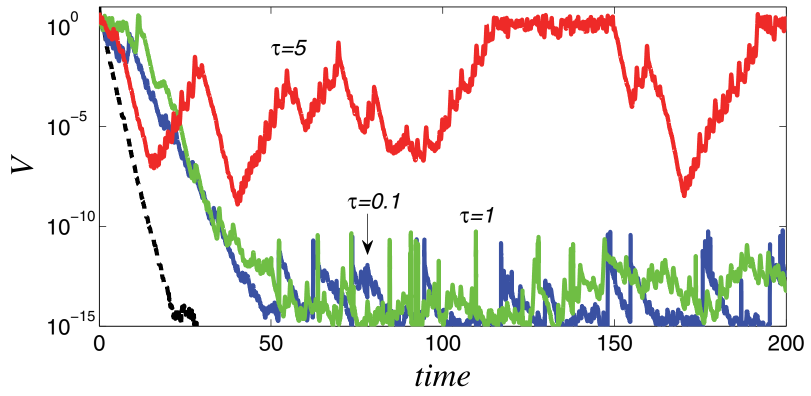

Figure 3. *Synchronization error as a function of time for the same solutions as in Figure 2 for the averaged system (black dashed curve) and the blinking system with switching times $\tau = 0.1$ (blue), $\tau = 1$ (green), and $\tau = 5$ (red). Synchronization takes place in the averaged system and in the blinking system for $\tau = 0.1$ and $\tau = 1$, but not for $\tau = 5$. Other parameters are the same as in Figure 2.*

It may be surprising that the blinking system synchronizes since at each time instant only about half of the connections between the Lorenz systems are active, which implies that most of the time the network is disconnected. Nevertheless, as our example shows, the time-varying interactions are sufficient to guarantee synchronization if they vary fast enough.

Another peculiarity of this example is that at each instant the synchronization subspace D is invariant under the dynamics of the blinking system. Contrary to the averaged system, this subspace is unstable most of the time. Later in the paper we shall generalize the results of the example to all blinking dynamical systems with an attractor of the averaged system that is an *invariant* set of the blinking system and make them more quantitative.

4. Second example: Coupled bistable systems.

Consider the system of two bistable systems coupled by a blinking connection

$$\begin{aligned} \dot{x} &= f(x) + 1.6 s(t)(y - x), \\ \dot{y} &= 2f(y) + 1.6 s(t)(x - y), \end{aligned} \tag{4.1}$$

where $f(x) = x(1 - x^2)$ and $s(t)$ is a binary switching function as represented in the general blinking system (2.1). We suppose that the switch is closed with probability $p = 0.5$.

The isolated bistable systems are described by

$$\dot{x} = f(x) \quad \text{and} \quad \dot{y} = 2f(y).$$

They both have two stable equilibrium points $x = 1$ and $x = -1$, and an unstable equilibrium point $x = 0$.

When the switch is open ($s(t) = 0$), the two bistable systems do not interact and the combined system has four stable and five unstable equilibrium points, namely all combinations of $x = -1, 0, 1$ and $y = -1, 0, 1$. When the switch is closed, there remain only two stable ($x = y = -1, 1$) and one unstable equilibrium ($x = y = 0$) points. The same is true for the averaged system obtained from (4.1) by replacing the switching variable $s(t)$ with its mean $p = 0.5$:

$$\dot{\xi} = f(\xi) + 0.8(\eta - \xi),$$
$$\dot{\eta} = 2f(\eta) + 0.8(\xi - \eta).$$

It is a gradient system. Its potential function is

$$V(\xi, \eta) = \frac{\xi^4}{4} + \frac{\eta^4}{2} - \frac{\xi^2}{10} - \frac{3\eta^2}{5} - \frac{4\xi\eta}{5} + \frac{3}{2}.$$

We have chosen the free constants in such a way that at the two minima, $(1, 1)$ and $(-1, -1)$, the potential takes the value 0. As in any gradient system, along the solutions of the averaged system the potential function decreases monotonically, except at the equilibrium points, where it is constant. The potential function is therefore a Lyapunov function of the averaged system. The two minima of V are, of course, asymptotically stable equilibria of the averaged system. They are also asymptotically stable equilibrium points of the blinking system. Therefore, we have an averaged system with two attractors which are also invariant sets of the blinking system. This enables the convergence of the trajectories of the blinking system to the attractors of the averaged system for fast switching.

It is not difficult to see that the square with the corners $(1, 1)$, $(1, -1)$, $(-1, -1)$, $(-1, 1)$ is forward invariant under both the averaged and the blinking systems. We consequently shall limit our attention to the dynamics in this square.

The blinking system has the two stable equilibrium points in common with the averaged system. Therefore, the solution of the blinking system can converge to one of these equilibrium points. If so, the question remains of whether it converges to the same equilibrium as the averaged system when starting from the same initial state. In Figure 4 an initial state is chosen close to the boundary between the attraction basins of the two stable equilibrium points of the averaged system. Starting from this state, a solution of the blinking system and the solution of the averaged system are shown. As expected, they converge to the same equilibrium point. However, there is a small probability that for another switching sequence it converges to the other equilibrium point (Figure 5). Practically, this happens only when the initial state is close to the attraction basin boundary, as is the case in Figure 5. The faster the switching, and the farther away from the basin boundary of the equilibrium points, the smaller is this probability.

On the other hand, the convergence to the equilibrium point is exponentially fast not only for the solutions of the averaged system but also for the solutions of the blinking system. This is illustrated in Figure 6 by representing the potential function V as a function of time in a semilogarithmic scale. Of course, when starting close to the attraction basin boundary, initially convergence is slow, but then it picks up its asymptotic exponential speed. Remarkably, the solution of the blinking system appears to have the same asymptotic exponential speed of convergence, whereas, in general, we would expect a slower exponential speed.

In both examples of chaotic Lorenz systems and coupled bistable systems discussed so far, the salient feature of the dynamics is the invariance of the averaged system's attractor(s) under the blinking system. In the following two examples this invariance property does not hold anymore, and the trajectories of the blinking system cannot converge to the attractor of

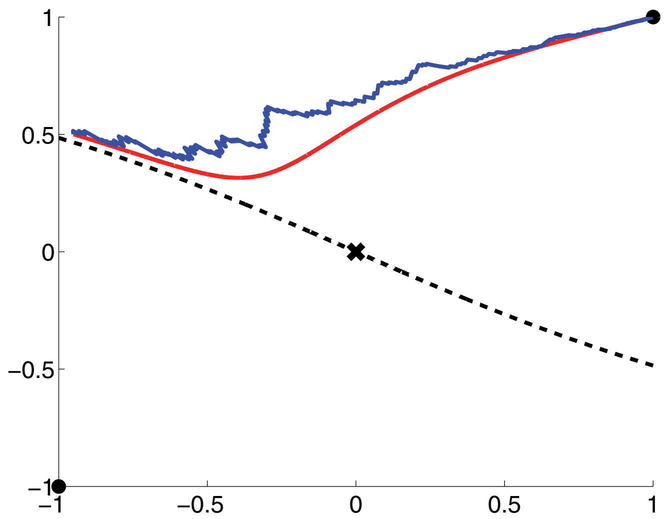

Figure 4. *Trajectories of the averaged system (red smooth curve) and of the blinking system (blue irregular curve) start from the same initial state and converge to the same equilibrium point* $(1, 1)$*. The other asymptotically stable equilibrium point at* $(-1, -1)$ *is also marked by a solid circle. The third equilibrium point* $(0, 0)$ *is a saddle, marked by a cross. Its stable manifold in the averaged system is drawn with a black dashed line. It is also the separatrix between the attraction basin boundaries of the two stable equilibrium points. The switching period is* $\tau = 0.01$*.*

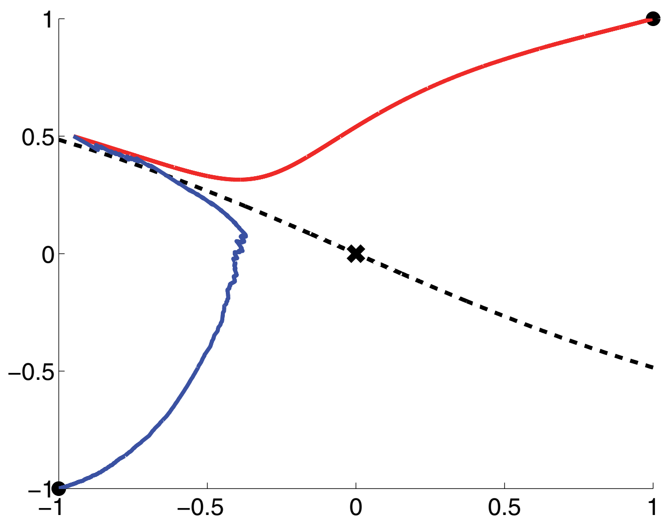

Figure 5. *Same representation as in Figure 4 but for another switching sequence of the blinking system. In this case the trajectory of the blinking system converges to the other stable equilibrium point. Note that the initial state, which is identical for all trajectories in this figure and Figure 4, is close to the attraction basin boundary.*

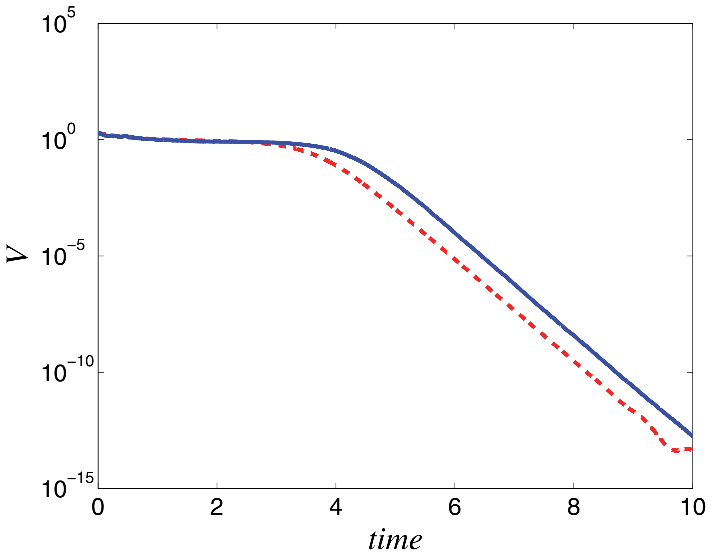

Figure 6. *Potential function V along a solution of the averaged system (red dashed curve) and the blinking system (blue solid curve) as a function of time in a semilogarithmic plot. The initial state is the same as in Figures 4 and 5. Since it is close to the attraction basin boundary, the potential function decreases only slowly in the beginning. Both trajectories appear to have the same asymptotic exponential speed of convergence, but the blinking trajectory is somewhat delayed with respect to the trajectory of the averaged system.*

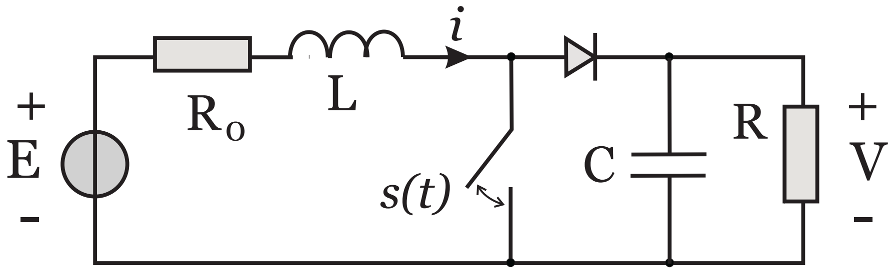

Figure 7. *Switched boost power converter as an example of a blinking system.*

the averaged system and can reach only a small neighborhood of it. In this case, the attractor of the averaged system may be viewed as a ghost attractor for the blinking system.

5. Third example: Switching power converter. Consider the circuit of Figure 7. Its function is to convert DC power at voltage E to DC power across the resistor R [22, 23]. The switch is usually operated periodically at a rather high frequency. This frequency and its harmonics are filtered as much as possible. However, some parasitic frequency components remain and pollute the network. Another possibility is to operate the switch stochastically. The advantage is that the power of the parasitic components is distributed over a whole frequency range, and therefore it is less disturbing than the narrow band parasitics in the case of periodic switching. This is discussed in [22, 23], where, however, somewhat different random switching schemes are used.

Let the switching variable be $s = 1$ when the switch is closed and $s = 0$ when the switch is open. It can be seen that the diode is open for $s = 1$ and is a short circuit for $s = 0$. The circuit equations are then

$$\frac{di}{dt} = \frac{E}{L} - \frac{R_0}{L} i - (1 - s(t)) \frac{v}{L},$$
$$\frac{dv}{dt} = (1 - s(t)) \frac{i}{C} - \frac{v}{RC}.$$

The corresponding averaged equation is

$$\frac{di}{dt}\frac{d\xi}{dt} = \frac{E}{L} - \frac{R_0}{L}\xi - (1 - p) \frac{\eta}{L},$$
$$\frac{d\eta}{dt} = (1 - p) \frac{\xi}{C} - \frac{\eta}{RC}.$$

The averaged system is asymptotically stable. The averaged DC-output voltage at its equilibrium point is $\eta = \frac{(1-p)R}{(1-p)^2 R + R_0} E$ and is adjustable by choosing p. It is larger than the input voltage if R_0 is sufficiently small and p is sufficiently large (this is the reason for calling the circuit a boost converter). The mechanism that keeps the solution of the blinking and the averaged system more or less together is the convergence to the equilibrium point in the averaged system. Note that this equilibrium is not shared by the blinking system. In fact, when both $s = 0$ (open switch) and $s = 1$ (closed switch), the system has a different asymptotically stable equilibrium point ($v = \frac{R}{R+R_0} E$ and $v = 0$, respectively). The consequence is that the solution of the blinking system cannot converge to the equilibrium point of the averaged system but approaches and then fluctuates around it (Figure 8). The fluctuations diminish in size as the switching period τ decreases.

The next example represents a class of blinking dynamical systems possessing *multiple* attractors that are not shared by the averaged and blinking systems.

6. Fourth example: Cellular neural network with blinking shortcuts.

Consider a two-dimensional array of locally coupled first-order linear systems with a piecewise linear output function, known under the name of *cellular neural networks* (CNNs) [24]. Such networks can perform many signal processing computations using their intrinsic nonlinear dynamics. One way to implement this is to insert data as initial values of the states and to let the states converge to an equilibrium point of the (multistable) network. The mapping from the initial to the final states is the function performed by the network.

However, certain functions cannot be obtained directly using only local connections. This is the case of the "winner-take-all" function, where the maximum among a given set of numbers has to be determined. In a wider context, this task amounts to detecting a brightest target spot, based on the given visual picture that can be represented as a matrix. The following globally coupled network of one-dimensional systems realizes this function for a suitable choice of the parameters a, d, and k [25]:

$$\dot{x}_i = -x_i + (1 - \delta) y_i - \alpha \sum_{j=1}^{N} y_j + \kappa,$$
$$y_i = f(x_i) = \begin{cases} 1 & \text{for} \quad x_i > 1, \\ x_i & \text{for} \quad -1 \le x_i \le 1, \\ -1 & \text{for} \quad x_i < -1. \end{cases} \tag{6.1}$$

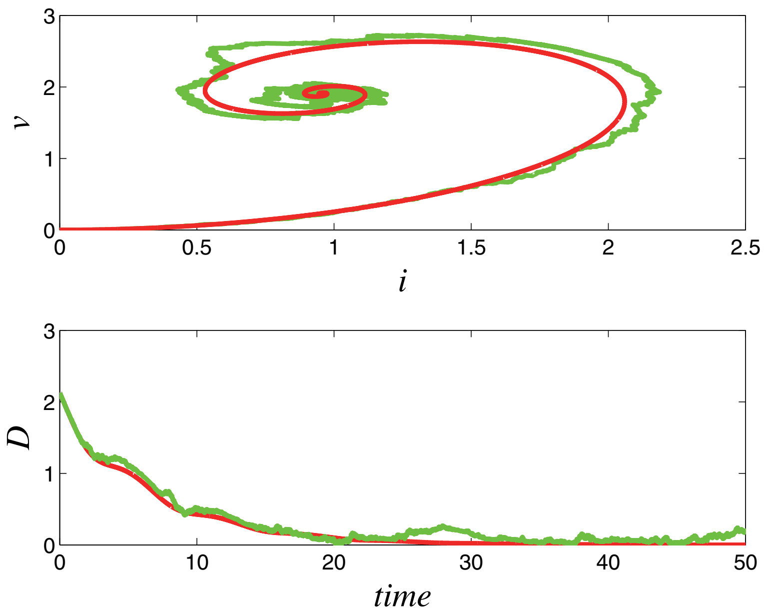

Figure 8. *Upper panel: Trajectories in state space of the averaged (red smooth curve) and of an instance of the blinking system (green irregular curve), starting from the origin. Lower panel: Distance D from the equilibrium point as a function of time for the same trajectories of the averaged (red) and the blinking (green) systems. Parameters are $E = 1$, $R_0 = 0.05$, $R = 4$, $L = C = 1$, $p = 0.5$, and $\tau = 0.005$.*

Note that α is the coupling coefficient between any two individual systems i and j of the network. Of course, instead of using an all-to-all coupling, computing the sum of all output signals y_i as an intermediate step would be a more efficient procedure. However, for our purposes, we start with the CNN with all-to-all coupling.

The correct functioning of this network is the following. At time t_0, the N given numbers are loaded as initial conditions $x_i(0)$, $i = 1, \dots, N$. Suppose that the largest among these numbers is $x_m(0)$. Then the state vector x evolves in time according to (6.1) and converges to the equilibrium point $\bar{x}$ such that $\bar{x}_m \geq 1$ and $\bar{x}_i \leq -1$ for $i \neq m$. In terms of the outputs, this means $\bar{y}_m = f(\bar{x}_m) = 1$ and $\bar{y}_i = f(\bar{x}_i) = -1$ for $i \neq m$. Hence, the whole system must have N asymptotically stable equilibrium points, one for each value of m, the index of the state with the largest initial value. The state space is divided into the N basins of attraction of these equilibrium points.

It is not difficult to see that it is not possible to design a "winner-take-all" CNN with only local connections. In fact, suppose that the initial state of a locally connected CNN has two local maxima, at cell i and at cell j, and these maxima are sufficiently far apart. Suppose that at cell i the maximum is also global. If this network performs the "winner-take-all" function correctly, there must be a stable equilibrium for which the output of the ith cell is $+1$ and all other outputs are -1. However, when all cells are in saturation, the jth cell and the ith cell do not interact. Then there will be another stable equilibrium where, in addition to the ith cell, the jth cell has output $+1$, and again all other cells have output -1. Such an equilibrium point is not compatible with the "winner-take-all" function.

On the other hand, local connections have an evident advantage when it comes to the realization of the CNN as an integrated circuit. A way out of this dilemma is to use switched connections instead as hardwired nonlocal connections and realize them by sending packets on a communication network that is associated with the CNN. Such a communication network has to be present anyway in order to charge the initial conditions and to read out the results. Thus, we consider the blinking system [12]

$$\dot{x}_i = -x_i + (1-\delta)\, y_i - \alpha \sum_{\substack{j \text{ nearest} \\ \text{neighbor of } i}}^{N} y_j + \frac{\alpha}{p} \sum_{\substack{j \text{ not nearest} \\ \text{neighbor of } i}}^{N} s_{ij}(t) y_j + \kappa,$$
$$y_i = f\,(x_i),$$

where in each time interval $(k-1)\,\tau \le t \le k\tau$ $s_{ij}(t)$ is constant, of value 1 with probability p and of value 0 with probability $1-p$. Its corresponding averaged system is system (2.2) as the switching variables s_{ij} are simply replaced by their mean value p.

As a numerical example, we consider a blinking 4×4 CNN ($N = 16$) that has the task of determining the largest among the 16 numbers:

$$(6.2)\qquad \begin{array}{rrrr} 0.3644 & 0.3958 & 0.1871 & 0.2898 \\ -0.3945 & -0.2433 & -0.0069 & 0.6359 \\ 0.0833 & 0.7200 & 0.7995 & 0.3205 \\ -0.6983 & 0.7073 & 0.6433 & -0.3161 \end{array}$$

In Figure 9, $x_{3,3}(t)$ and $x_{4,1}(t)$ are represented, for the blinking system, for the averaged system and for two instances of the blinking system. Here, we use the double indices for x to indicate the location of a given cell on the grid. The trajectory of one instance of the blinking system follows the trajectory of the averaged system and approaches the corresponding equilibrium point. The state $x_{3,3}$ increases beyond the value 1, as it should, since the element $(3,3)$ in (6.2) is the largest, whereas the state $x_{4,1}$ decreases below -1. Thus, both the averaged system with its all-to-all connections and this instance of the blinking system with its fixed local and the switched nonlocal connections perform the "winner-take-all" function correctly. In the case of the other instance of the blinking system the blinking trajectory converges to a wrong equilibrium point.

The peculiarity of this system is that it has several attractors (stable equilibrium points), but practically all the time, none of them is an equilibrium point of the blinking system. Therefore, the trajectory of the blinking system cannot converge to an equilibrium point of the averaged system, but it can only get close to it, as in the switched power converter (cf. Figure 10).

Furthermore, there is a nonzero probability that it will approach the wrong equilibrium point, as illustrated in Figure 10. This depends on how close the second largest initial state is and how small τ is. Reducing τ, the solution of the blinking system initially follows the solution of the averaged system more closely and thereby has a lower probability of approaching the wrong equilibrium point. It also remains closer to the equilibrium point at larger times. This point is illustrated in Figures 11 and 12, where τ is 10 times smaller than in Figure 9.

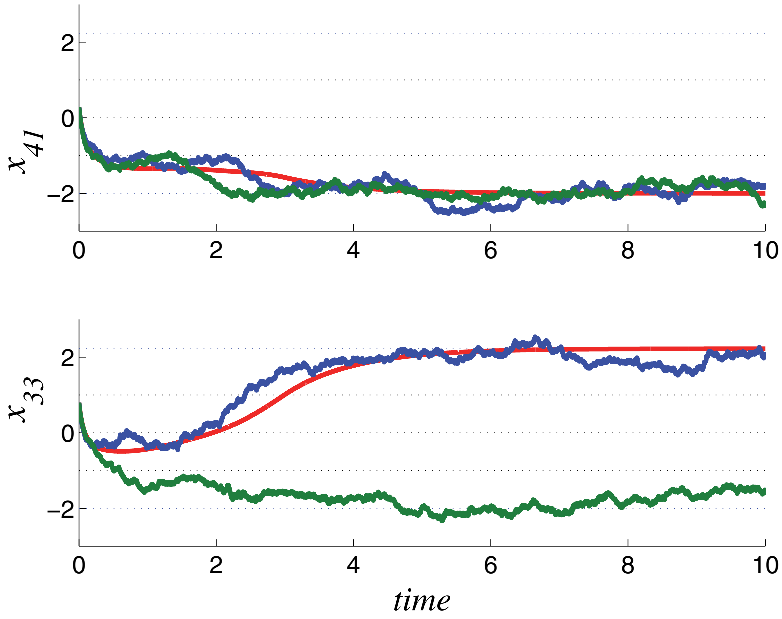

Figure 9. *Trajectories for two instances of the 4×4 "winner-take-all" blinking CNN, with $\alpha = 1$, $\delta = 1.11$, $\kappa = -13.89$, $\tau = 0.001$, $p = 0.1$ (irregular blue and green curves), together with the trajectory of the averaged system (smooth red curve). The trajectory of the averaged system always approaches the correct equilibrium point, whereas the trajectories of the blinking system may or may not reach it, depending on the instance of the switching process. Upper panel: Trajectories corresponding to the cell $(4, 1)$ whose initial condition does not have the maximal value. Lower panel: Trajectories corresponding to the cell $(3, 3)$ whose initial condition has the maximal value.*

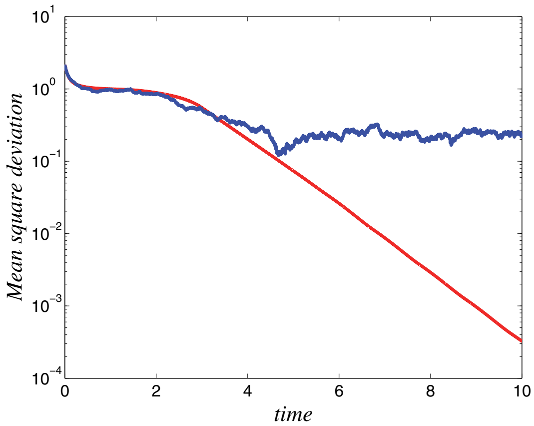

Figure 10. *Mean square deviation from the correct equilibrium point as a function of time. Smooth red curve: averaged system trajectory converges exponentially fast to the correct equilibrium point. Irregular blue curve: blinking system trajectory approaches the correct equilibrium point but does not converge to it. Parameters are as in Figure 9.*

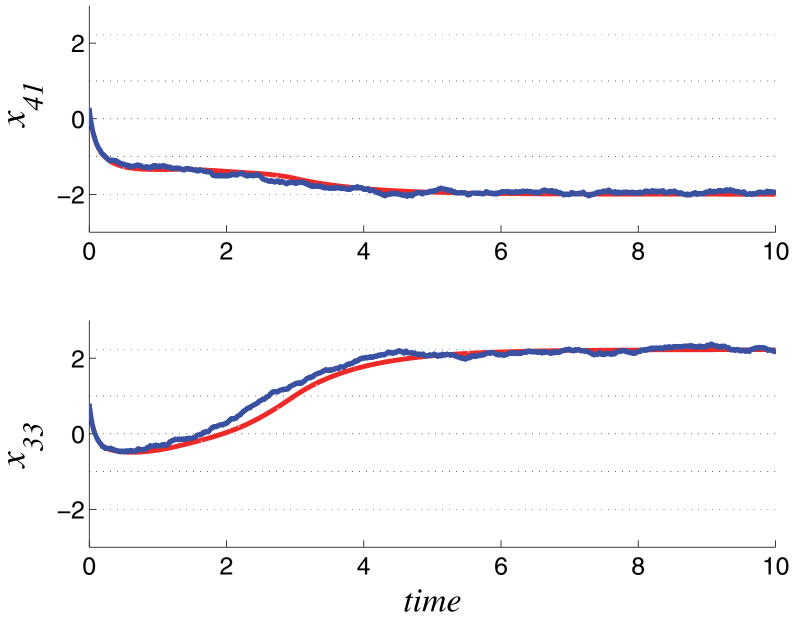

Figure 11. *Same as Figure 9 for a single trajectory of the blinking system that approaches the correct equilibrium point. Main difference:* $\tau = 0.0001$*, i.e.,* 10 *times smaller than for Figure 9.*

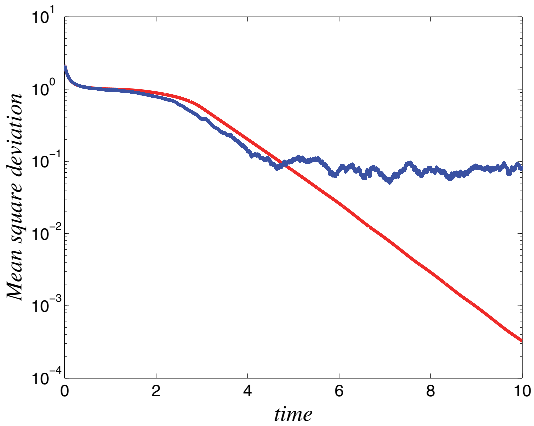

Figure 12. *Mean square deviation from the correct equilibrium point as a function of time for the trajectory of Figure 11.*

There is even a nonzero probability that after having approached the correct equilibrium point it will escape towards another equilibrium point of the averaged system. However, this last probability is much smaller than an initial approach to the wrong equilibrium point, so that it can be neglected in practice. In fact, in practice, once the trajectory of the blinking system is sufficiently close to a stable equilibrium point of the averaged system, a decision is taken and the system is stopped.

7. Analysis of the blinking system: Basic assumptions. We now return to the analysis of the asymptotic behavior of a general blinking system (2.1) and its relation to the solution of the averaged system (2.2), starting from the same initial state.

We will prove four general theorems regarding the asymptotic dynamical behavior of the blinking system. Each of the four above examples illustrates one of the theorems. The order of the theorems is opposite to the order of examples, as it is more convenient to start from the most general case (Theorem 8.3, the fourth example) and to proceed to the most constrained case (Theorem 11.1, the first example).

As in the companion paper [21], we make the following, not very restrictive, hypotheses.

Hypothesis 1.

(a) The function F that defines the blinking system (2.1) is locally Lipschitz continuous in x (the first N arguments) and continuous in s (the last M arguments).

(b) For any switching signal $s(t)$ and any state x_0 there exists exactly one trajectory $x(t)$ of the blinking system with $x(0) = x_0$, defined for $0 \le t < \infty$. Similarly, there exists a unique trajectory $\xi(t)$ of the averaged system

$$\frac{d\xi}{dt} = \Phi\left(\xi(t)\right),$$

defined for $0 \le t < \infty$ for a given initial state $\xi(0) = \xi_0$.

(c) There is a connected and compact, i.e., closed and bounded, region R in $\mathbb{R}^N$ such that (i) all trajectories of the blinking and averaged systems starting in R remain in R; (ii) all trajectories of the blinking and averaged systems starting outside of R reach R.

Thus, all interesting dynamics take place in R and, in particular, all attractors lie in R. In what follows, we shall restrict our attention to trajectories in R. The continuity of F implies that F and Φ are both bounded on R.

8. General case (Case 1): Multiple attractors are possible; their invariance for the blinking system is not required. The information processing CNN (the fourth example) is a case in point.

We now consider an attractor of the averaged system (2.2)–(2.3) and a solution of the averaged system that converges to it. The question is in what sense and under what conditions a solution of the blinking system (2.1) that starts from the same initial conditions converges to the attractor. Instead of an attractor, we shall use the less restrictive notion of an attracting set.

Definition 8.1. *An attracting set A of a dynamical system is a compact connected set such that*

- *any trajectory starting in A remains in A, and*
- *any trajectory starting sufficiently close to A converges to A.*

Let us remark that an attractor has to satisfy the additional constraint that it contains a dense trajectory.

Hypothesis 2. The averaged system (2.2) has an attracting set A with a corresponding Lyapunov function W. More precisely, we assume that there is a twice continuously differentiable function $W : \mathbb{R}^N \to \mathbb{R}$ such that the following hold:

(a) There is a constant V_1 and a connected component C_1 of the level set $\{x \mid W(x) \le V_1\}$, contained in R and containing A, such that for any $x \in C_1$

$$W(x) \geq 0 \text{ and } x \in A \Leftrightarrow W(x) = 0.$$

(b) For any $x \in C_1$ we have

$$\frac{\partial W}{\partial x}(x)\,\Phi(x) \leq 0$$

and

$$\frac{\partial W}{\partial x}(x)\,\Phi(x) = 0 \quad \Leftrightarrow \quad x \in A.$$

(c) The sets

$$C_1 \cap \left\{ x \in \mathbb{R}^N \mid W(x) \leq V \right\}$$

are compact and connected for $0 \leq V \leq V_1$.

Remark 8.1. The requirement that the Lyapunov function should be 0 exclusively on A may seem not realizable in many situations. In particular, if A is a chaotic attractor, W cannot be 0 on A and positive elsewhere. In this case, a larger attracting set A, which contains the attractor of the chaotic system, has to be chosen. In the first example, the trajectories of the averaged system converge to the chaotic attractor located in the diagonal subspace, but as an attracting set, the diagonal subspace, or rather its intersection with R, has to be chosen.

Note that for any solution $\xi(t)$ of the averaged system

$$\frac{d}{dt} W(\xi(t)) = \frac{\partial W}{\partial x}(\xi(t))\,\Phi(\xi(t))\,. \tag{8.1}$$

Thus, within the region C_1 the function W decreases strictly along any solution of the averaged system, except in the attracting set A, where it remains constant. This implies, in particular, that all solutions of the averaged system starting in C_1 converge to A. The question is then what happens to the solution of the blinking system. Note that if the level set $\{x \mid W(x) \leq V_1\}$ is not connected, then in connected components other than C_1 there are other attractors, or even solutions, diverging to infinity.

We proceed in two steps. The aim of the first step is to show that the Lyapunov function also decreases along solutions of the blinking system. Actually, because of the stochastic nature of switching, this is not always true. The Lyapunov function may increase temporarily, but the general tendency is to decrease. This can be expressed by showing that after a certain time Δt the Lyapunov function decreases with high probability. In the second step, we analyze the behavior of the blinking system for large times. We show that W decreases either to 0 or to a small value with high probability.

For the purpose of formulating Theorem 8.3, it is convenient to introduce, in addition to the functions F and Φ that give the time derivative of the states of the blinking and the averaged systems, the four functions $D_F W : \mathbb{R}^{N+M} \to \mathbb{R}$, $D_F^2 W : \mathbb{R}^{N+2M} \to \mathbb{R}$, and $D_\Phi W$, $D_\Phi^2 W : \mathbb{R}^N \to \mathbb{R}$ that give the first time derivative and a kind of second time derivative of the Lyapunov function W along solutions of the blinking and the averaged systems.

Definition 8.2. *Define the functions* $D_F W$, $D_F^2 W$, $D_\Phi W$, D_Φ^2 *by*

$$D_F W(x,s) = \sum_{i=1}^{N} \frac{\partial W}{\partial x_i}(x)\, F_i(x,s),$$
$$D_F^2 W(x,\tilde{s},s) = \sum_{i,j=1}^{N} \left[\frac{\partial^2 W}{\partial x_i \partial x_j}(x)\, F_i(x,\tilde{s})\, F_j(x,s) + \frac{\partial W}{\partial x_i}(x)\, \frac{\partial F_i}{\partial x_j}(x,\tilde{s})\, F_j(x,s)\right],$$
$$D_\Phi W(x,s) = \sum_{i=1}^{N} \frac{\partial W}{\partial x_i}(x)\, \Phi_i(x),$$
$$D_\Phi^2 W(x) = \sum_{i,j=1}^{N} \left[\frac{\partial^2 W}{\partial x_i \partial x_j}(x)\, \Phi_i(x)\, \Phi_j(x) + \frac{\partial W}{\partial x_i}(x)\, \frac{\partial \Phi_i}{\partial x_j}(x)\Phi_j(x)\right],$$

and introduce their bounds on R

$$\begin{aligned} B_{W\Phi} &= \max_{x\in R} |D_\Phi W(x)|, \\ LB_{W\Phi} &= \max_{x\in R} \left|D_\Phi^2 W(x)\right|, \\ B_{WF} &= \max_{s\in\{0,1\}^M} \max_{x\in R} |D_F W(x,s)|, \\ LB_{WF} &= \max_{s,\tilde{s}\in\{0,1\}^M} \max_{x\in R} \left|D_F^2 W(x,\tilde{s},s)\right|. \end{aligned} \tag{8.2}$$

Note that these constants can be explicitly formulated via the parameters of the general blinking system (2.1).

The following theorem expresses the behavior of the blinking system in the most general case where the averaged system (2.2) may have multiple attractors that are in general noninvariant under the blinking system.

Theorem 8.3 (Case 1: possible multiple attractors/noninvariance). *Suppose Hypothesis* 1 *is satisfied and the averaged system* (2.2) *has an attracting set* A *with a corresponding Lyapunov function* W*, satisfying Hypothesis* 2*. Choose* V_0 *such that* $0 < V_0 < V_1$*, and let*

$$\begin{aligned} -\gamma &= \max_{x\in C_1,\, V_0\le W(x)\le V_1} D_\Phi W(x), \\ \Delta t &= \frac{\gamma}{2(LB_{WF}+LB_{W\Phi})}, \\ \alpha &= B_{WF} + B_{W\Phi}, \\ c &= \frac{1}{64(LB_{WF}+LB_{W\Phi})B_{WF}^2}, \end{aligned} \tag{8.3}$$

where the various expressions are defined in Definition 8.2*. Suppose that*

$$V_1 - V_0 \ge \frac{3\gamma^2}{4\left(LB_{WF} + LB_{W\Phi}\right)} \tag{8.4}$$

and that τ *is sufficiently small, such that*

$$2e\frac{2\alpha+\gamma}{\gamma}\exp\left(-\frac{c\gamma^3}{\tau}\right) \le 1 - \sqrt{\frac{e}{3}}. \tag{8.5}$$

Consider the two open regions (see Figure 13)

$$U_0 = \{x \mid W(x) < V_0 + \gamma\Delta t\} \cap C_1, \quad U_\infty = \{x \mid W(x) > V_1 + \alpha\Delta t\}. \tag{8.6}$$

Consider a solution $x(t)$ of the blinking system with $W(x(0)) = V_1$ and $x(0) \in C_1$ (the connected component of the level set $\{x \mid W(x) \le V_1\}$ containing the attracting set A). Then the following hold:

(a) *The probability P^{direct}_{escape} that the solution $x(t)$ of the blinking system reaches U_∞ before reaching U_0 is bounded by*

$$P^{direct}_{escape} \le \frac{36\,\alpha^2}{\gamma^2} \exp\left(-\frac{c\gamma^3}{\tau}\right). \tag{8.7}$$

Inversely, the probability that the solution $x(t)$ of the blinking system reaches U_0 before reaching U_∞ is at least

$$P^{direct}_{attraction} \ge 1 - \frac{72\,\alpha^2}{\gamma^2} \exp\left(-\frac{c\gamma^3}{\tau}\right). \tag{8.8}$$

(b) *Let $T_{attraction}$ be the time for the solution of the blinking system to enter U_0 through its boundary and T_{remain} be the time it remains in*

$$\bar{U}_{0+} = \left\{x \,\middle|\, W(x) \le V_0 + \left(\frac{3}{2}\gamma + \alpha\right)\Delta t\right\} \tag{8.9}$$

after reaching U_0. These times are random variables with the properties

$$P\left(T_{attraction} \le 2\frac{V_1 - V_0}{\gamma}\right) > 1 - 8(V_1 - V_0)\frac{(LB_{WF} + LB_{W\Phi})}{\gamma^2}\exp\left(-\frac{c\gamma^3}{\tau}\right) \tag{8.10}$$

and

$$P(T_{remain} > T) > 1 - 4T\frac{(LB_{WF} + LB_{W\Phi})}{\gamma}\exp\left(-\frac{c\gamma^3}{\tau}\right). \tag{8.11}$$

Proof. The complete proof of Theorem 8.3 is given in the appendix. Here, we give a short narrative that underlies the main ideas of the proof. For the convenience of the reader, we refer to the corresponding parts of the complete proof.

As mentioned above, the proof is divided into two steps. In the first step, we consider the trajectories of the blinking and averaged systems, starting from the same initial condition at $t = 0$. Function W, being a Lyapunov function for the averaged system, decreases strictly along the trajectory of the averaged system. This is not necessarily true along the trajectory of the blinking system. However, if we can give a sufficiently small bound for the difference between the values of W along the two trajectories, we can ensure that W also decreases in the blinking system. Along this line, we give two bounds on their difference after some time Δt that can be thought of as an intermediate scale: large with respect to the switching time and small with respect to the time it takes the averaged system to get close to its attractor. The more conservative bound (cf. Lemma 13.2(a) in the appendix) is always valid, whereas

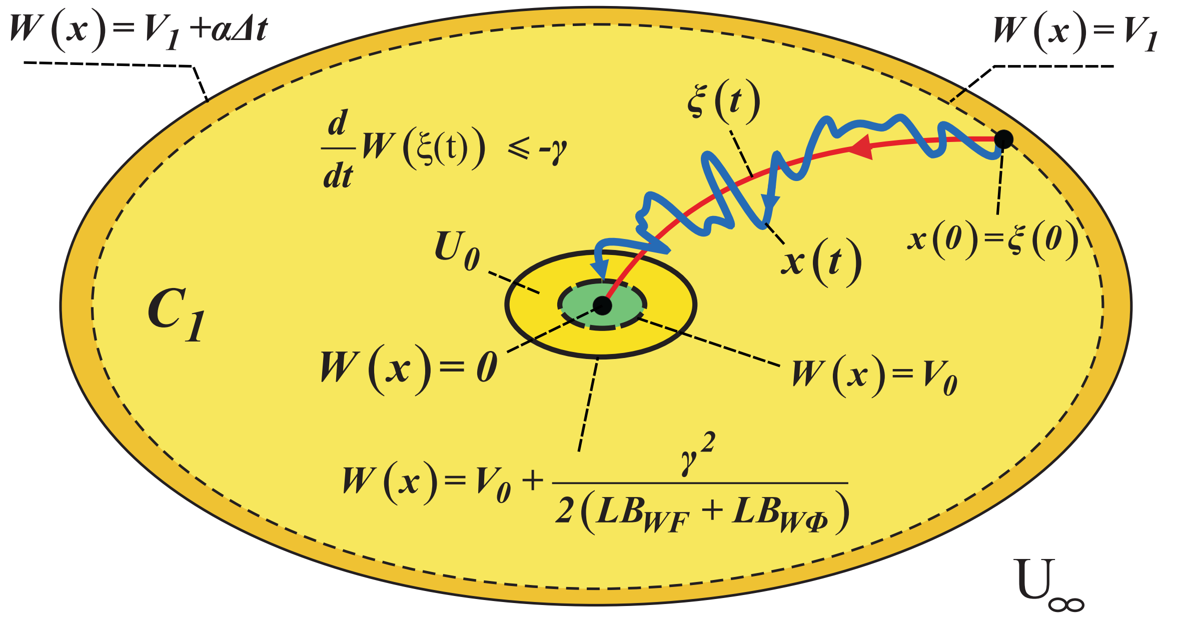

Figure 13. *Illustration of Theorem 8.3. Trajectory of the averaged system (regular red line) reaches its attractor at which $W(x) = 0$. The attractor of the averaged system is neither unique nor invariant under the blinking system. This attractor acts as a ghost attractor for the blinking system whose trajectory (irregular blue line) reaches a small neighborhood U_0 of the ghost attractor in time $T_{attraction}$ with probability $P^{direct}_{attraction}$. Subsequently, it may, after some time T_{remain}, go far away from the ghost attractor, but the probability that it happens in a given lapse of time can be made arbitrarily small by decreasing the switching period τ. Initially, the trajectory of the blinking system may also escape from the attraction basin C_1 right away and move toward another attractor in U_∞ (not shown) with probability P^{direct}_{escape}. This probability approaches 0 when $\tau \to 0$.*

the closer bound (cf. Lemma 13.2(b) in the appendix) holds only with a certain probability. A lower bound for this probability is given in Lemma 13.1 (see the appendix) and is based on the Hoeffding bound for large deviations of sums of independent random variables. The closer bound allows proving that W decreases in the blinking system for most switching sequences, whereas the larger bound limits its increase for the exceptions (cf. Lemma 13.3 in the appendix). In order to make it precise, we need W to decrease uniformly per unit time in the averaged system ($\dot{W}(\xi) \le -\gamma < 0$). This is possible only outside of a neighborhood of the attractor. At this point, we have a freedom to choose this neighborhood as a level set of W ($W(\xi) < V_0$)), where V_0 can be chosen arbitrarily within $0 < V_0 < V_1$. This choice determines constant γ. A suitable choice of Δt, taken in Lemma 13.3 (see the appendix), guarantees the decreasing of W in the blinking system due to the closer bound.

In the second step, we consider the trajectory of the blinking system $x(t)$ starting from region C_1 (cf. Hypothesis 1). We decompose the time axis into intervals of length Δt. Depending on the switching sequences, either function W decreases from $q\Delta t$ to $(q+1)\Delta t$ or it may increase in accordance with the bounds derived in the first step. Combining the bounds across the different intervals, we can prove that with high probability the trajectory of the blinking system reaches neighborhood U_0 of the averaged system attractor, and with low probability it escapes from region U_∞. This proof is not straightforward, as the events to which the closer bound can be applied in various intervals are not independent. To cope with this, we introduce the auxiliary random variables $Z_q(s) = W(x(q\Delta t))$, $q = 0, 1, 2, \ldots$. Even though the trajectory of the blinking system evaluated at multiples of Δt, $x(q\Delta t)$ is a vector-valued Markov chain, function W, applied to it to obtain stochastic process Z_q, destroys the Markov property. For this reason, a more detailed analysis of process Z_q is performed in Lemmas 13.4 and 13.5.

In this way, we obtain the lower bound for the probability of reaching neighborhood U_0 of the attractor and the upper bound for the probability of escaping to region U_∞ from domain C_1 (cf. Theorem 8.3(a)).

The proof of Theorem 8.3(b) also relies on the detailed analysis of stochastic process Z_q. In this way we obtain a probabilistic upper bound on the time to reach U_0. Once the trajectory of the blinking has reached U_0, it wiggles around the attractor of the averaged system (the ghost attractor). However, it remains in neighborhood $\bar{U}_{0+}$, slightly larger than U_0 for a long time. Indeed, a probabilistic lower bound for this time is obtained. ■

Remark 8.2.

(a) This general theorem is in all four cases applicable. However, when the attractor is an invariant set of the blinking system, a stronger theorem will be proved, which guarantees actual convergence to the attractor.

(b) By decreasing the switching period τ, the probability of escaping from the region C_1 before reaching U_0 can be made arbitrarily small.

(c) If the averaged system has no other attractor than A, or if all attractors of the averaged system lie in the attracting set A, eventually almost every solution reaches U_0. Typically, in such a case the Lyapunov function W is defined in the whole space, the level set $\{x \mid W(x) \le V_1\}$ is connected and thus identical to C_1, and Theorem 8.3 gives information only on the time needed to reach U_0. This case will be treated in more detail in Theorem 9.1.

(d) In the case of multistability, i.e., when there is an attractor outside of A, there is always a nonzero probability of reaching a neighborhood of this attractor before reaching U_0. It can be made arbitrarily small by increasing the speed of switching. The region C_1 is necessarily contained in the basin of attraction of A in the averaged system. Actually, in general, it will be distinctly smaller than the basin of attraction. Nevertheless, we can show that by switching sufficiently fast, the solution of the blinking system that starts in the basin of attraction of A of the averaged system will again with high probability reach a small neighborhood of A without leaving the basin. In fact this result is obtained by combining Theorem 3.1 from the companion paper [21] and Theorem 8.3. According to Theorem 3.1 for finite time from [21], the solution of the blinking system will follow the solution of the averaged system for some time, getting to C_1 or at least closer to C_1. In the latter case, Theorem 3.1 from [21] can be applied repeatedly, until C_1 is reached. Then Theorem 8.3 can be invoked.

(e) Note that the upper bound $P^{\text{direct}}_{\text{escape}}$ and the lower bound $P^{\text{direct}}_{\text{attraction}}$ do not imply that the sum of the two probabilities sums up to 1, as one would expect. Therefore, the trajectories of the blinking system that never reach U_0 or U_∞ might have positive probability. However, this is just a technical consequence of the way of deriving the bounds. Furthermore, this probability could not be larger than $P^{\text{direct}}_{\text{escape}}$, and the case that this probability vanishes is compatible with the bounds. The same technical detail will reappear in Theorem 9.1.

(f) At first sight, it is not evident what the role of the constraint (8.4) is and whether it seriously limits the applicability of the theorem. A closer examination shows that this condition can always be satisfied by reducing the size of V_0, which will diminish the value of γ, which in turn needs a smaller switching time τ.

9. Case 2: Unique attracting set; not necessarily invariant for the blinking system.

We now strengthen Hypothesis 2 to adapt it to the case where A is the unique attractor or attracting set of the averaged system or where all attractors of the system are contained in the attracting set A. The switching power converter (the third example) is a case in point.

Hypothesis 3. The level set $\{x \,|\, W(x) \le V_1\}$ is connected and thus identical to C_1 introduced in Hypothesis 2. It contains the compact attracting region R introduced in Hypothesis 1; i.e., we have

$$x \in R \;\Rightarrow\; W(x) \le V_1.$$

This implies that any solution of the blinking and the averaged systems after a finite time T (depending on the solution) satisfies

$$W(x(t)) \le V_1 \quad \text{for } t \ge T.$$

Note that if Hypothesis 3 is not satisfied because a level set of the Lyapunov function W is not connected, there are necessarily several disjoint attracting sets and we are back to Case 1.

This more restrictive hypothesis allows us to formulate the stronger theorem.

Theorem 9.1 (Case 2: unique attractor/noninvariance). *Under Hypotheses 1, 2, and 3 consider any solution $x(t)$, $t \in [0,\infty)$, of the blinking system. Let W be the Lyapunov function introduced in Hypothesis 2 and V_1 be the positive constant introduced in Hypothesis 3. As in Theorem 8.3, choose V_0 such that $0 < V_0 < V_1$, and let*

$$(9.1)\qquad \begin{aligned} -\gamma &= \max_{V_0 \le W(x) \le V_1} D_\Phi W(x), \\ c &= \frac{1}{64(LB_{WF}+LB_{W\Phi})B_{WF}^2}, \\ U_0 &= \left\{ x \;\middle|\; W(x) < V_0 + \frac{\gamma^2}{2(LB_{WF}+LB_{W\Phi})} \right\}. \end{aligned}$$

Then the following properties hold:

(a) *If the switching time τ satisfies*

$$(9.2)\qquad \tau < \frac{c\gamma^3}{\ln\left[D\frac{(V_1-V_0)}{\gamma^2}\right]},$$

where

$$(9.3)\qquad D = 8\left(LB_{WF} + LB_{W\Phi}\right),$$

then the solution $x(t)$ almost surely reaches the neighborhood U_0 of A in finite time.

(b) *Let $x \in R$, and assume that (9.2) holds. Suppose that $W(x(0)) \leq V_1$, and let, for all natural numbers n, $P_1(n)$ be the probability that it takes at least time $2n\frac{V_1-V_0}{\gamma}$ to reach U_0. Then*

$$P_1(n) \leq \exp\left(-n\left[\frac{c\gamma^3}{\tau} - \ln\left[D\frac{(V_1-V_0)}{\gamma^2}\right]\right]\right). \tag{9.4}$$

(c) *Suppose that $x(0) \in R$. The probability that at time t the solution satisfies*

$$W(x(t)) > V_0 + \frac{8\gamma^2}{D} \tag{9.5}$$

is bounded by $\frac{D}{\gamma^2}V_1 \cdot \exp(-\frac{c\gamma^3}{\tau})$ for $t \geq \frac{2}{\gamma}V_1$.

Proof. The complete proof of Theorem 9.1 is given in the appendix. As before, we give a short narrative that refers to various parts of the complete proof. Here, we employ the same technique developed for the proof of Theorem 8.3. To prove that the trajectory of the blinking system for almost all switching sequences reaches neighborhood U_0 of the ghost attractor, we give an upper bound for the probability that the trajectory, starting from outside of U_0, will not reach U_0 at least until time $T + nQ\Delta t$ for an arbitrary integer n. Here, Δt is chosen as in the proof of Theorem 8.3 and Q is the minimum number of time steps Δt it takes the trajectory to get into U_0 in accordance with the stricter bound from Lemma 13.5 (see the appendix). This upper bound for the probability converges to 0 as $n \to \infty$ (cf. (13.85)). This implies that with zero probability the trajectory never reaches U_0, or, equivalently, with probability 1 it reaches U_0. This completes the proof of Theorem 9.1(a). The short proof of Theorem 9.1(b) immediately follows from the bound obtained in part (a).

To prove part (c), we exploit the same probabilistic bounds for the time the trajectory reaches U_0 (cf. part (b)) and the time it remains in neighborhood $\bar{U}_{0+}$ (cf. Theorem 8.3(b). Beyond the combined time interval, the trajectory escapes from $\bar{U}_{0+}$; however, in this case the trajectory does not get lost in a region for which we have no assumptions about its behavior, as opposed to the assumptions for Theorem 8.3 dealing with multiple ghost attractors. Therefore, we use the current state of the escaped trajectory outside $\bar{U}_{0+}$ as the initial condition and apply the bounds for reaching U_0 and remaining in $\bar{U}_{0+}$ again. Hence, for any given time we can give an estimate on the probability that the trajectory is in $\bar{U}_{0+}$. Actually, for technical reasons this argument is carried out not for $\bar{U}_{0+}$ itself but for a slightly different neighborhood of the ghost attractor (cf. (9.5)). This concludes the sketch of the proof for part (c). ■

Remark 9.1.

(a) Part (a) of Theorem 9.1 expresses the fact that almost all trajectories of the blinking system get arbitrarily close to the attractor of the averaged system, provided switching is fast enough. This actually implies that almost all trajectories of the blinking system visit infinitely often any neighborhood of the attractor. Part (b) says that with high probability the trajectories of the blinking system reach such a neighborhood rather fast. Of course, the smaller the neighborhood, the longer the time necessary to reach it.

(b) While trajectories of the blinking system get arbitrarily close to the attractor of the averaged system, they do not stay close forever. However, their excursions far from

the attractor are relatively rare. In fact, part (c) ensures that at any given time, with high probability the trajectory of the blinking system is close to the attractor. By switching faster, one can force the trajectories of the blinking system to stay with high probability even closer to the attractor. The quantitative aspect of this property is somewhat obscured in Theorem 9.1 by the implicit dependence of the two parameters V_0 and γ. Close to a hyperbolic attractor, V_0 and γ are proportional. Supposing proportionality, we get a statement of the following form:
If for V_0 close to zero $\gamma \sim V_0$, then there exists a constant K such that the probability that $W(x(t)) > V_0$ is bounded by $\frac{K}{V_0^2}\exp(-\frac{c\gamma^3}{\tau})$. This probability can be made as small as desired by increasing the speed of switching.

10. Case 3: The attractor or attracting set is invariant under the blinking system; there may be other attractors. In the special case when the attractor or attracting set A of the averaged system is an invariant set of the blinking system, convergence of the solutions of the blinking system to A is possible. In this section, we give the conditions and the precise formulation when this happens. The property that the set A is invariant under the dynamics of the blinking system implies that the Lyapunov function W vanishes along any trajectory of the blinking system within A. We strengthen this property somewhat in Hypothesis 4 by requiring not only that the derivative of W along any solution of the blinking system that approaches A converge to zero but that the derivative be bounded by W multiplied by a constant. In other words, we assume that the derivative of the logarithm $\ln W$ is bounded. The bistable system (the second example) is a case in point.

Hypothesis 4. Suppose the averaged system (2.2) has an attracting set A with a corresponding Lyapunov function W satisfying Hypothesis 2. Introduce the following functions, similar to Definition 8.2, but using instead of the Lyapunov function its logarithm. It is well defined as long as $W(x) \neq 0$. Introduce

$$D_F \ln W(x,s) = \sum_{i=1}^{N} \frac{\partial \ln W}{\partial x_i}(x)\, F_i(x,s),$$

$$D_F^2 \ln W(x,\tilde{s},s) = \sum_{i,j=1}^{N} \left[\frac{\partial^2 \ln W}{\partial x_i \partial x_j}(x)\, F_i(x,\tilde{s})\, F_j(x,s) \;+\; \frac{\partial \ln W}{\partial x_i}(x)\, \frac{\partial F_i}{\partial x_j}(x,\tilde{s})\, F_j(x,s)\right],$$

$$D_\Phi \ln W(x) = \sum_{i=1}^{N} \frac{\partial \ln W}{\partial x_i}(x)\, \Phi_i(x),$$

$$D_\Phi^2 \ln W(x) = \sum_{i,j=1}^{N} \left[\frac{\partial^2 \ln W}{\partial x_i \partial x_j}(x)\, \Phi_i(x)\, \Phi_j(x) \;+\; \frac{\partial \ln W}{\partial x_i}(x)\, \frac{\partial \Phi_i}{\partial x_j}(x)\, \Phi_j(x)\right].$$

These functions could diverge to infinity when approaching A. We require that this is not the case, and that the following constants are finite

$$B_{\ln W\Phi} = \sup_{x\in R,\, W(x)\neq 0} |D_\Phi \ln W(x)|,$$

$$LB_{\ln W\Phi} = \sup_{x\in R,\, W(x)\neq 0} \left|D_\Phi^2 \ln W(x)\right|, \tag{10.1}$$

$$B_{\ln WF} = \max_{s\in\{0,1\}^M} \sup_{x\in R,\, W(x)\neq 0} |D_F \ln W(x,s)|,$$
$$LB_{\ln WF} = \max_{s,\tilde{s}\in\{0,1\}^M} \sup_{x\in R,\, W(x)\neq 0} \left|D_F^2 \ln W(x,\tilde{s},s)\right|.$$

In addition, we require that there be a positive constant γ such that

$$\sup_{x\in C_1,\, 0<W(x)\leq V_1,\, x\notin A} D_\Phi \ln W(x) = -\gamma < 0, \tag{10.2}$$

where C_1 and V_1 are defined in Hypothesis 2.

The following theorem describes the behavior of the blinking system in the case of multiple attractors of the averaged system that are invariant under the blinking system.

Theorem 10.1 (Case 3: nonunique attractor/invariance). *Under Hypotheses 1, 2, 3, and 4 consider a solution $x(t)$ of the blinking system that starts at a point $x(0)\in C_1$ (the connected component of the level set $\{x \mid W(x)\leq V_1\}$ containing A). Let*

$$-\gamma = \sup_{x\in C_1,\, 0<W(x)\leq V_1,\, x\notin A} D_\Phi \ln W(x),$$
$$\alpha = B_{\ln WF} + B_{\ln W\Phi},$$
$$\Delta t = \tfrac{\gamma}{2(LB_{\ln WF}+LB_{\ln W\Phi})},$$
$$c = \tfrac{1}{64(LB_{\ln WF}+LB_{\ln W\Phi})B^2_{\ln WF}},$$

where the constants B and LB are defined in (10.1) and C_1 and V_1 are defined in Hypothesis 2. Suppose that τ is sufficiently small such that

$$4e\frac{2\alpha+\gamma}{\gamma}\exp\left(-\frac{c\gamma^3}{\tau}\right) \leq 1-\sqrt{\frac{e}{3}}. \tag{10.3}$$

Consider the open region

$$U_\infty = \{x \mid \ln W(x) > \ln V_1 + \alpha\Delta t\}.$$

Then the following hold:

(a) *The probability P_{escape} that the solution $x(t)$ of the blinking system reaches U_∞ is bounded by*

$$P_{escape} \leq \frac{36\,\alpha^2}{\gamma^2}\cdot\exp\left(-\frac{c\gamma^3}{\tau}\right). \tag{10.4}$$

(b) *The probability that the solution $x(t)$ of the blinking system converges to A is at least*

$$P_{convergence} \geq 1 - \frac{216\,\alpha^2}{\gamma^2}\cdot\exp\left(-\frac{c\gamma^3}{\tau}\right). \tag{10.5}$$

More precisely, with probability at least $1-\frac{216\,\alpha^2}{\gamma^2}\cdot\exp(-\frac{c\gamma^3}{\tau})$ the convergence is exponentially fast according to

$$W(x(t)) \leq K\,\exp\left(-\frac{\gamma}{4}t\right), \tag{10.6}$$

where $K = V_1\,\exp(\frac{(\alpha+\frac{\gamma}{4})\,\gamma}{2(LB_{\ln WF}+LB_{\ln W\Phi})})$.

Proof. The complete proof is given in the appendix. As in the previous sketches of the proofs, we give the main reasoning below.

The proof closely follows that of Theorem 8.3. It starts with Lemmas 13.6, 13.7, 13.8, 13.9, and 13.10, which are analogous to Lemmas 13.1, 13.2, 13.3, 13.4, and 13.5, except that the logarithm of the Lyapunov function $\ln W$ is used instead of W. As far as the value of γ is concerned, in this case it is imposed by Hypothesis 4. Furthermore, its meaning is that the logarithm $\ln W$ of the Lyapunov function of the averaged system decreases at least linearly in time with coefficient γ. This corresponds to an exponential decrease of W with rate at least γ. This contrasts with the assumption of Theorem 8.3 that γ expresses the linear rate of the decreasing of W along the trajectory of the averaged system outside neighborhood U_0.

The derivation of the bound on the probability that the trajectory escapes to region U_∞ is completely analogous to the corresponding proof for Theorem 8.3. This completes the proof of part (a).

The lower bound on the probability of convergence to the attractor is derived in a similar way to the bound on the probability of direct attraction (8.8) in Theorem 8.3. The difference between the assumptions made in the two theorems is that the probabilistic bound on the decrease of function W in the present case is valid along the entire trajectory of the blinking system, whereas in the case addressed by Theorem 8.3 it is valid only as long as the trajectory stays outside of the ghost attractor's neighborhood U_0. Nevertheless, the proof that in the present case function $\ln W(x(t))$ almost never becomes larger than $\ln W(x_0) - \frac{\gamma}{4}t$ is analogous to the corresponding proof that the trajectory of the blinking system almost never remains trapped in the region between U_0 and U_∞. This proof in the present case leads to the lower bound for the convergence to the attractor. This concludes part (b). ■

11. Case 4: Unique attractor; invariant for the blinking system.

We now suppose that Hypotheses 1, 2, 3, and 4 are valid; i.e., the attractor A is unique or all attractors are contained in the attracting set A. Furthermore, A is invariant under the blinking system. Finally, the compact absorbing region R is contained in the level set $\{x \,|\, W(x) \leq V_1\}$ of the Lyapunov function W. From this we can conclude that if the switching time is sufficiently short, almost all solutions of the blinking system converge exponentially fast to A. The synchronization of coupled Lorenz systems (the first example) is a case in point.

Theorem 11.1 (Case 4: unique attractor/invariance). *Under Hypotheses 1, 2, 3, and 4 consider any solution $x(t)$, $t \in [0, \infty)$, of the blinking system. Let W be the Lyapunov function introduced in Hypothesis 2 and V_1 be the positive constant introduced in Hypothesis 3. Let*

$$
\begin{aligned}
-\gamma &= \sup_{x \in C_1,\, 0 < W(x) \leq V_1,\, x \notin A} D_\Phi \ln W(x),\\
\alpha &= B_{\ln W F} + B_{\ln W \Phi},\\
c &= \frac{1}{64(LB_{\ln W F} + LB_{\ln W \Phi})B^2_{\ln W F}},
\end{aligned}
$$

where the constants B and LB are defined in (10.1). Suppose that τ is sufficiently small such that

$$
2e\frac{2\alpha+\gamma}{\gamma}\exp\left(-\frac{c\gamma^3}{\tau}\right) \leq 1 - \sqrt{\frac{e}{3}}.
$$

Then the solution almost surely converges to A exponentially fast with exponential speed at least $\frac{\gamma}{4}$; i.e., for almost all switching sequences there is a constant K such that for all $t \geq 0$

$$W(x(t)) \leq K \exp\left(-\frac{\gamma}{4}t\right). \tag{11.1}$$

If $W(x(0)) \leq V_1$, then with probability at least

$$P_{\text{const}} \geq 1 - \frac{180\,\alpha^2}{\gamma^2} \cdot \exp\left(-\frac{c\gamma^3}{\tau}\right) \tag{11.2}$$

the following constant K is sufficient for (11.1):

$$K = V_1 \exp\left(\frac{\left(\alpha + \frac{\gamma}{4}\right)\gamma}{2\left(LB_{\ln WF} + LB_{\ln W\Phi}\right)}\right). \tag{11.3}$$

Proof. The complete proof is given in the appendix. As in the proof of Theorem 10.1, we show that $\ln W(x(t))$ is almost never larger than $\ln W(x_0) - \frac{\gamma}{4}t$ for infinitely many times $t = Q\Delta t$, $Q = 1, 2, 3, \ldots$. In contrast to Theorem 10.1 the probabilistic bound on the decrease of $\ln W$ along the trajectory of the blinking system is valid everywhere since the attractor of the averaged system is unique. This implies the convergence to the attractor of the blinking system and the existence of constant K such that $W(x(t)) \leq K\, e^{-\frac{\gamma}{4}t}$ for almost all switching sequences (cf. (13.141) in the complete proof). In the same way as in the proof of Theorem 10.1(b) it is shown that the constant K with high probability is smaller than the value (11.3). The lower bounds on the probabilities in the two theorems are the same if one discounts for the probability of escaping to U_∞ that is present in Theorem 10.1 but absent in Theorem 11.1. ■

12. Conclusions.

We have studied the asymptotic dynamics of general blinking systems with identically distributed independent random switching variables. Four distinct classes of blinking dynamical systems have to be distinguished. Two properties differentiate them: single or multiple attractors of the averaged system and their invariance or noninvariance under the dynamics of the blinking system.

In the most constrained class (Case 4), the averaged system has a single global attractor or attracting set A which is invariant under the blinking system. We have proved that if the switching period τ is smaller than an explicitly given bound, the trajectories of the blinking system for almost all switching sequences converge to A. This bound depends on the exponential speed of convergence γ in the averaged system. In fact, the crucial ratio τ/γ^3 has to be small. Furthermore, the trajectories of the blinking system converge to the attractor also exponentially fast, and the speed of convergence is slower than γ but of the same order of magnitude. As an example, we have used a blinking network of diffusively coupled identical Lorenz systems where connections are stochastically switched on and off, similar to networks considered in our previous paper [11]. The averaged system is obtained by replacing each blinking connection with a static link, corresponding to the averaged blinking connection. If the connections of the averaged network are strong enough [5], the diagonal subspace corresponding to completely synchronized solutions is an attracting set. It is also an invariant set of the blinking system at each time instant. In our example, most of the time the blinking network is disconnected; nevertheless, synchronization in the fast blinking system takes place for almost all switching sequences.

In the case where the attractor or attracting set A is invariant under the blinking system but there may be other attractors (Case 3), we have proved the following properties of the blinking system. Its trajectories converge to A with a probability that can be made arbitrarily close to 1 by decreasing the switching period. Again, the crucial ratio is τ/γ^3, where γ is the exponential speed of convergence to A in the averaged system. Below a certain threshold for the switching period, the trajectories of the blinking system that converge to A almost surely converge exponentially fast. Again, the speed of convergence is slightly slower than γ. The exceptional trajectories that do not converge to A, in general, escape from the attraction basin of the attractor. As an illustrative example, we have considered two bistable systems coupled by a blinking connection. The blinking connection implies that the two systems are part of the time uncoupled, and part of the time are coupled with a certain coupling strength d. The averaged system yields a coupled system with static connections of lower coupling strength pd, where p is the probability that the coupling is turned on. As a consequence, the averaged system has two asymptotically stable equilibrium points as attractors. They are also attractors of the blinking system at each time instant; however, their basins of attraction are different. The faster the switching, the more unlikely it is that the blinking system and the averaged system converge to two different attractors.

In the case where the attractor of the averaged system is unique but not invariant under the blinking system (Case 2), we have proved the following theorem (Theorem 9.1). For any choice of a small neighborhood U_0 of the attractor, the trajectories of the averaged system approach the attractor with a minimum linear speed $\gamma > 0$. We have limited our analysis to the linear speed because in any case the trajectories of the blinking system cannot converge to the attractor and can reach only a small neighborhood of it. Therefore, the attractor of the averaged system acts as a ghost attractor for the blinking system. The linear speed goes to zero when approaching the attractor as opposed to the exponential speed. As the switching period is small enough, the trajectory of the blinking system almost surely reaches a certain neighborhood of the attractor in finite time. This threshold is essentially proportional to γ^3. With probability close to 1, the trajectories of the blinking system reach U_0 in a short time. They may leave U_0 from time to time, but the probability that at any given time they are far from the ghost attractor is very small. As an illustrative example, we have chosen a stochastically switched DC-DC power converter. The averaged system is linear with a globally asymptotically stable equilibrium point. The blinking system is switched between two linear systems that both have a globally asymptotically stable equilibrium point but different from the averaged system and different from each other. Therefore, the unique attractor of the averaged system is not invariant under the blinking system. Nevertheless, in accordance with Theorem 9.1 the trajectories of the blinking system rapidly approach a neighborhood of the equilibrium point of the averaged system (ghost attractor) and stay close to the ghost attractor while stochastically wiggling around.

Finally, in the case where the attractor of the averaged system is neither unique nor invariant under the blinking system (Case 1), we have derived the following results. As in the previous case, we choose a small neighborhood U_0 of the ghost attractor of the blinking system and a subregion of its basin of attraction such that the linear speed γ of convergence of the averaged system is uniformly positive. The probability that the trajectory of the blinking system reaches U_0 rapidly can be made arbitrarily close to 1 by decreasing the switching period. Subsequently, it may after some time go far away from the ghost attractor, but the probability that it happens in a given lapse of time can be made arbitrarily small. For both probabilities we find once more τ/γ^3 to be essential. An illustrative example for this case is an information processing CNN. It consists of a regular planar array of linear first-order dynamical systems with nonlinear output functions. This network has static connections between nearest neighbors and blinking connections between nodes that are spaced farther apart. It has many asymptotically stable equilibrium points. The information processing is performed by the dynamics of the network, more specifically by the time evolution from the initial network state to the corresponding equilibrium point. The network is designed to perform a winner-take-all function, i.e., to determine the maximum component of the initial network state. This function cannot be performed by solely static local connections, and the blinking connections are necessary. If the switching is fast, the trajectory of the blinking network reaches the correct ghost equilibrium point as long as the initial state components are sufficiently distinct; otherwise, the initial state vector is too close to the basin boundary between two (the correct and a wrong ghost) equilibria. Furthermore, the trajectory of the blinking system stays sufficiently long in a small neighborhood of the correct ghost equilibrium point such that the information can be read out.

Figure 14 summarizes the general results for all four cases. The comparison of the explicit thresholds for the switching period and various probabilities with the effective properties of the blinking systems in the case of the four examples will be published elsewhere. In particular, note that the exponential rate of convergence to the attractor we obtained, when it is invariant under the blinking system, is not the actual one but just a lower bound. The precise rate could in principle be computed by determining the action functional and solving a variational problem [34]. However, our lower bound is given as an explicit expression in easily determined system parameters, which seems to be out of reach for the action functional method.

We have chosen binary vector–valued identically distributed independent random variables that are constant in a small time interval as the driving stochastic process for the blinking system. Thus, we have deliberately restricted the blinking system to a very specific form, as far as its probabilistic nature is concerned; however, its dynamical systems nature is very general. This setting is adapted to a network of interconnected dynamical systems, where the connections are stochastically switched on and off. Our results can easily be extended to blinking systems driven by a Markov vector process in discrete or continuous time, instead of sequences of independent random vectors. Indeed, there are only two places where the specific nature of the driving stochastic process is used. The first is the probability of large deviation of the sum or integral of random variables from its mean (cf. (13.5), (13.7), (13.10), and (13.105), (13.107), (13.108)), where we have used Hoeffding's inequality [39]. Similar inequalities can be obtained for Markov processes in discrete or continuous time, and one simply has to substitute in the formulas the quantities $\mathrm{P}_{W\lambda}$ or $P_{\ln W\lambda}$ by the corresponding expressions. The second place where the nature of the stochastic process is used is when combining the results on the decreasing of the Lyapunov function W or its logarithm $\ln W$ during a time interval of length Δt for subsequent intervals on the time axis. Our proofs

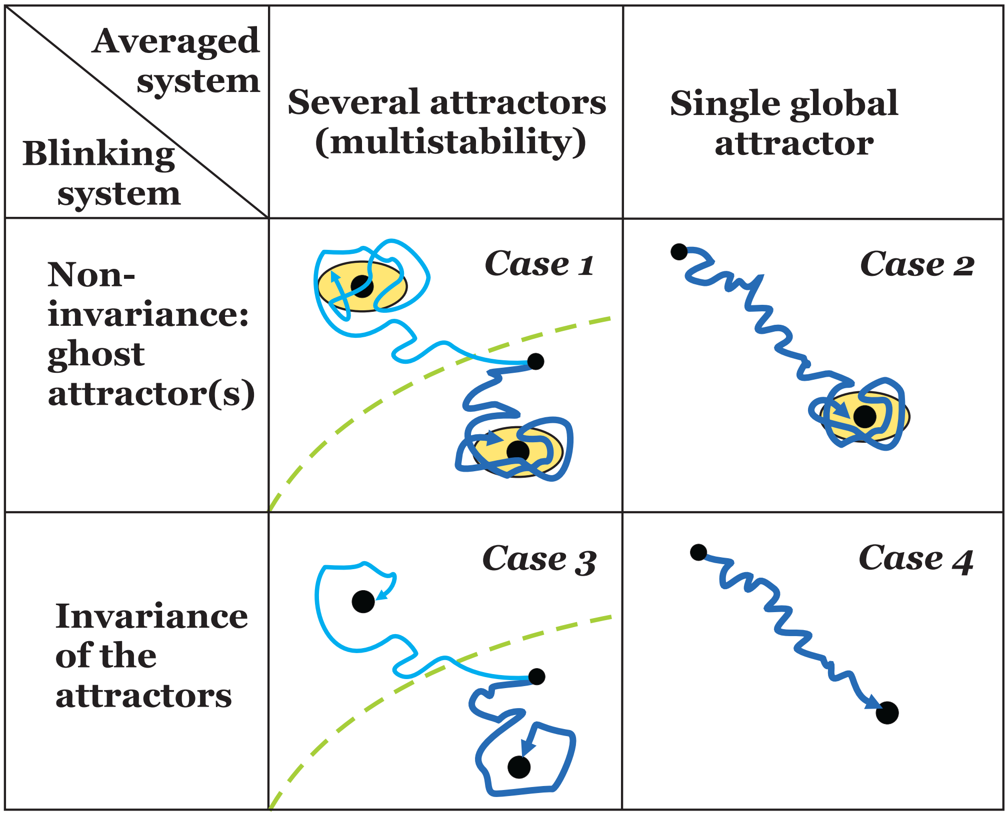

Figure 14. *Qualitative behavior of the blinking system's trajectories in the four cases. Upper row: Trajectories reach a neighborhood (the ghost attractor) of the same attractor as the averaged system. Lower row: Trajectories converge to the same attractor as the averaged system. Left column: Property holds with high probability; however, there is a probability of escape to another attractor. The dashed line separates attraction basins of two attractors in the averaged system. Right column: Property holds for almost all switching sequences. The bounds on the probabilities are given in the four theorems.*

remain valid as long as the solution $x(t)$ is a vector-valued Markov process. If the driving stochastic process has the Markov property, this is always the case.

13. Appendix. In this appendix, we give the details of the proofs for Theorems 8.3, 9.1, 10.1, and 11.1.

13.1. Theorem 8.3: Case 1, multiple ghost attractors.

13.1.1. Preliminary lemmas. We consider the trajectories of the blinking and averaged systems, starting from the same initial condition at $t = 0$. We assume that the Lyapunov function W of the averaged system (cf. Theorem 8.3) strictly decreases along the trajectory of the averaged system. As discussed in the sketch of the proof for Theorem 8.3, we complete the first step of the proof by deriving three lemmas to show that after a certain time Δt the Lyapunov function decreases along the trajectory of the blinking system either to 0 or to a small value with high probability.

First step. It follows that if $x(t)$ is a solution of the blinking and $\xi(t)$ a solution of the averaged system, then, in addition to (8.1),

$$\begin{aligned} \tfrac{d^2}{dt^2}W(\xi(t)) &= D_\Phi^2 W(\xi(t)), \\ \tfrac{d}{dt}W(x(t)) &= D_F W(x(t), s(t)), \\ \tfrac{d}{dt}D_F W(x(t), \tilde{s}) &= D_F^2 W(x(t), \tilde{s}, s(t)). \end{aligned} \tag{13.1}$$

Furthermore, the expectation

$$\mathrm{E}(D_F W(x, S)) = D_\Phi W(x), \tag{13.2}$$

and thus by the weak law of large numbers [38], for all $x(t)$ and $\lambda > 0$,

$$P\left\{\left|\frac{1}{K}\sum_{k=1}^{K} D_F W\left(x, S^k\right) - D_\Phi W(x)\right| > \lambda\right\} \underset{K\to\infty}{\to} 0. \tag{13.3}$$

Hence, we can define

$$\mathrm{P}_{W\lambda}(K) = \max_{x\in R} \mathrm{P}\left\{\left|\frac{1}{K}\sum_{k=1}^{K} D_F W\left(x, S^k\right) - D_\Phi W(x)\right| > \lambda\right\}, \tag{13.4}$$

which has the property

$$\mathrm{P}_{W\lambda}(K) \underset{K\to\infty}{\to} 0 \quad \text{for any } \lambda > 0, \tag{13.5}$$

and because of the stationarity of the stochastic process, we have also for each $k_0 \geq 1$

$$\mathrm{P}_{W\lambda}(K) = \max_{x\in R} \mathrm{P}\left\{\left|\frac{1}{K}\sum_{k=k_0+1}^{K+k_0} D_F W\left(x, S^k\right) - D_\Phi W(x)\right| > \lambda\right\}. \tag{13.6}$$

Applying the Hoeffding inequality [39, 21], we get the following lemma.

Lemma 13.1. *For any $\lambda > 0$, the following inequality holds:*

$$P_{W\lambda}(K) \leq 2\exp\left(-\frac{K\lambda^2}{2\,B_{WF}^2}\right). \tag{13.7}$$

Notice that

$$\int_t^{t+\Delta t} \left[D_F W(x, s(u)) - D_\Phi W(x)\right] du = \sum_{k=k_0+1}^{k_0+K} \tau\left[D_F W\left(x, s^k\right) - D_\Phi W(x)\right] \tag{13.8}$$

if $t = k_0\tau$ and $\Delta t = K\tau$ and inequality (13.7) can be applied to bound the left-hand side of (13.8). In order to be able to apply this bound to arbitrary positive real t and Δt, we again extend the definition of $\mathrm{P}_{W\lambda}(K)$ to noninteger K by

$$\begin{aligned} P_{W\lambda}(K) = \max_{0\leq\alpha\leq 1}\max_{x\in R} P\Bigg(\tfrac{1}{K}\Bigg|&(1-\alpha)D_F W\left(x, S^1\right) + \sum_{k=2}^{\lfloor K+\alpha\rfloor} D_F W\left(x, S^k\right) \\ &+ (K+\alpha-\lfloor K+\alpha\rfloor)D_F W\left(x, S^{\lfloor K+\alpha\rfloor+1}\right) - K D_\Phi W(x)\Bigg| > \lambda\Bigg). \end{aligned} \tag{13.9}$$

With this definition, on the one hand, (13.7) is valid for any real $K > 0$, and on the other hand, for any $t \geq 0,\ \Delta t \geq 0,\ \lambda > 0,\ x \in R$ we have

$$P\left(\left|\int_t^{t+\Delta t} \left[D_F W\left(x, s\left(u\right)\right) - D_\Phi W\left(x\right)\right] du\right| > \lambda \cdot \Delta t\right) \leq P_{W\lambda}\left(\frac{\Delta t}{\tau}\right). \tag{13.10}$$

The following lemma and its proof are similar to Lemma 6.2 in the appendix of the companion paper [21].

Lemma 13.2. *Consider a solution $x(t)$ of the blinking system. Choose a time $t \geq 0$ and the solution of the averaged system with $\xi\left(t\right) = x\left(t\right)$. Then the following hold:*

(a) *For any $\Delta t \geq 0$,*

$$\left|W\left(x\left(t + \Delta t\right)\right) - W\left(\xi\left(t + \Delta t\right)\right)\right| \leq \alpha \Delta t, \tag{13.11}$$

where $\alpha = B_{WF} + B_{W\Phi}$.

(b) *For any $\lambda > 0$ and $\Delta t \geq 0$, the conditional probability that*

$$\left|W\left(x\left(t + \Delta t\right)\right) - W\left(\xi\left(t + \Delta t\right)\right)\right| \leq \frac{LB_{WF} + LB_{W\Phi}}{2}\Delta t^2 + \lambda \Delta t \tag{13.12}$$

holds, given the value of $x(t)$, is at least $1 - P_{W\lambda}\left(\frac{\Delta t}{\tau}\right)$.

Proof.

(a) We first prove the first part of the lemma:

$$\begin{aligned} &\left|W\left(x\left(t + \Delta t\right)\right) - W\left(\xi\left(t + \Delta t\right)\right)\right| \\ &\quad\leq \left|\int_t^{t+\Delta t} \frac{d}{du} W\left(x\left(u\right)\right) du - \int_t^{t+\Delta t} \frac{d}{du} W\left(\xi\left(u\right)\right) du\right| \\ &\quad= \left|\int_t^{t+\Delta t} D_F W\left(x\left(u\right), s\left(u\right)\right) dt - \int_t^{t+\Delta t} D_\Phi W\left(\xi\left(u\right)\right) du\right| \\ &\quad\leq \left(B_{WF} + B_{W\Phi}\right) \Delta t. \end{aligned} \tag{13.13}$$

(b) Here, we prove the second part of the lemma:

$$\begin{aligned} &\left|W\left(x\left(t + \Delta t\right)\right) - W\left(\xi\left(t + \Delta t\right)\right)\right| \\ &\quad\leq \left|\int_t^{t+\Delta t} D_F W\left(x\left(u\right), s\left(u\right)\right) dt - \int_t^{t+\Delta t} D_\Phi W\left(\xi\left(u\right)\right) du\right| \\ &\quad\leq \left|\int_t^{t+\Delta t} \left[D_F W\left(x\left(u\right), s\left(u\right)\right) - D_F W\left(x\left(t\right), s\left(u\right)\right)\right] du\right| \\ &\qquad+ \left|\int_t^{t+\Delta t} \left[D_F W\left(x\left(t\right), s\left(u\right)\right) - D_\Phi W\left(x\left(t\right)\right)\right] du\right| \\ &\qquad+ \left|\int_t^{t+\Delta t} \left[D_\Phi W\left(\xi\left(u\right)\right) - D_\Phi W\left(\xi\left(t\right)\right)\right] du\right|. \end{aligned} \tag{13.14}$$

Using (13.1), the first and third terms can be rewritten as

(13.15)
$$\begin{aligned}
&\left|\int_t^{t+\Delta t}\left[D_F W\left(x\left(u\right),s\left(u\right)\right)-D_F W\left(x\left(t\right),s\left(u\right)\right)\right]du\right|\\
&\quad\leq\left|\int_t^{t+\Delta t}\left[\int_t^u \frac{d}{dv}D_F W\left(x\left(v\right),s\left(u\right)\right)dv\right]du\right|\\
&\quad=\left|\int_t^{t+\Delta t}\left[\int_t^u D_F^2 W\left(x\left(v\right),s\left(v\right),s\left(u\right)\right)dv\right]du\right|\leq LB_{WF}\frac{\left(\Delta t\right)^2}{2};\\
&\left|\int_t^{t+\Delta t}\left[D_\Phi W\left(\xi\left(u\right)\right)-D_\Phi W\left(\xi\left(t\right)\right)\right]dt\right|\leq\left|\int_t^{t+\Delta t}\left[\int_t^u \frac{d}{dv}D_\Phi W\left(\xi\left(v\right)\right)dv\right]du\right|\\
&=\left|\int_t^{t+\Delta t}\left[\int_t^u D_\Phi^2 W\left(\xi\left(v\right)\right)dv\right]du\right|\leq LB_{W\Phi}\frac{\left(\Delta t\right)^2}{2}.
\end{aligned}$$

For the second term we apply (13.10) to obtain that with probability at least $1-P_{W\lambda}\left(\frac{\Delta t}{\tau}\right)$

$$\left|\int_t^{t+\Delta t}\left[D_F W\left(x\left(t\right),s\left(u\right)\right)-D_\Phi W\left(x\left(t\right)\right)\right]du\right|\leq\lambda\Delta t. \tag{13.16}$$

Combining (13.15) and (13.16) proves the lemma. ■

Lemma 13.3. *For any V_0 with $0<V_0<V_1$, let*

$$\begin{aligned}
-\gamma &= \max_{x\in C_1,\, V_0\leq W(x)\leq V_1} D_\Phi W\left(x\right),\\
\Delta t &= \tfrac{\gamma}{2(LB_{WF}+LB_{W\Phi})},\\
\lambda &= \tfrac{\gamma}{4},\\
\tilde{V}_0 &= V_0+\gamma\Delta t.
\end{aligned}\tag{13.17}$$

Then for any $t\geq 0$ and for any solution $x(t)$ of the blinking system the following hold:

(a) *If $x\left(t\right)\in C_1$ and $\tilde{V}_0\leq W\left(x\left(t\right)\right)$, the conditional probability that*

$$W\left(x\left(t+\Delta t\right)\right)\leq W\left(x\left(t\right)\right)-\frac{\gamma}{2}\Delta t \tag{13.18}$$

holds, given $x(t)$, is at least $1-P_{W\lambda}\left(\frac{\Delta t}{\tau}\right)$.

(b) *In general, for $x\left(t\right)\in C_1$, the conditional probability that*

$$W\left(x\left(t+\Delta t\right)\right)\leq W\left(x\left(t\right)\right)+\frac{\gamma}{2}\Delta t \tag{13.19}$$

holds, given $x(t)$, is at least $1-P_{W\lambda}\left(\frac{\Delta t}{\tau}\right)$.

Proof.

(a) Consider a solution $\xi(t)$ of the averaged system with $\xi(t) = x(t)$. Then

$$W(\xi(t+\Delta t)) = W(\xi(t)) + \int_t^{t+\Delta t} D_\Phi W(\xi(u))\, du. \tag{13.20}$$

By (13.17), $D_\Phi W(\xi(u)) \le -\gamma$ as long as $W(\xi(u)) \ge V_0$. On the other hand, if, for some $t < u < t + \Delta t$, $W(\xi(u)) = V_0$, then $W(\xi(t+\Delta t)) < V_0$, and thus under the condition $V_0 + \gamma\,\Delta t \le W(x(t)) \le V_1$ in all cases

$$W(\xi(t+\Delta t)) \le W(\xi(t)) - \gamma\Delta t. \tag{13.21}$$

Hence by Lemma 13.2, with the choices of parameters (13.17), we get under the condition $\tilde{V}_0 \le W(x(t)) \le V_1$ with probability at least $1 - P_{W\lambda}\left(\frac{\Delta t}{\tau}\right)$

$$\begin{aligned} W(x(t+\Delta t)) &\le W(\xi(t+\Delta t)) + |W(x(t+\Delta t)) - W(\xi(t+\Delta t))| \\ &\le W(\xi(t)) - \gamma\,\Delta t + \tfrac{\gamma}{2}\Delta t = W(x(t)) - \tfrac{\gamma}{2}\Delta t. \end{aligned} \tag{13.22}$$

(b) If we do not require $\tilde{V}_0 \le W(x(t))$, instead of (13.21) we have only

$$W(\xi(t+\Delta t)) \le W(\xi(t)) \tag{13.23}$$

and thus instead of (13.22)

$$W(x(t+\Delta t)) \le W(x(t)) + \frac{\gamma}{2}\Delta t \tag{13.24}$$

with probability at least $1 - P_{W\lambda}\left(\frac{\Delta t}{\tau}\right)$. ■

Second step. For any choice of the initial state $x(0)$ with $0 < W(x(0)) \le V_1$ and $x(0) \in C_1$ (the connected component of the level set $\{x \,|\, W(x) \le V_1\}$ containing the attracting set A), for any choice of the constants $\tau > 0$ and V_0 with $0 < V_0 < V_1$, and for the constants $\gamma, \lambda, \Delta t, \tilde{V}_0$ given by (13.17), we consider the following sequence of random variables on the probability space of switching sequences:

$$Z_q(s) = W(x(q\Delta t)), \qquad q = 0, 1, 2, \ldots. \tag{13.25}$$

By Hypothesis 2, Z_0 is concentrated on a single value that is smaller than or equal to V_1. Note that here again, as in the companion paper [21], the scalar stochastic process $\{Z_q\}$ is not a Markov process, even though $x(q\Delta t)$ is a vector-valued Markov process, because the application of the function W destroys much of the information contained in the state $x(q\Delta t)$. However, according to Lemma 13.3 we have the deterministic bound

$$W(x(t+\Delta t)) \le W(x(t)) + \alpha\Delta t, \tag{13.26}$$

and under the condition $W(x(t)) \le V_1$, the probabilistic bound is

$$P\left(W(x(t+\Delta t)) \le W(x(t)) - \frac{\gamma}{2}\Delta t \,\middle|\, x(t), \tilde{V}_0 \le W(x(t)) \le V_1\right) > 1 - P_{W\lambda}\left(\frac{\Delta t}{\tau}\right), \tag{13.27}$$

and finally the probabilistic bound becomes

$$P\left(W\left(x\left(t+\Delta t\right)\right) \le W\left(x\left(t\right)\right) + \frac{\gamma}{2}\Delta t \,|x\left(t\right), 0 \le W\left(x\left(t\right)\right) \le V_1\right) > 1 - P_{W\lambda}\left(\frac{\Delta t}{\tau}\right). \tag{13.28}$$

As for the finite time analysis [21], it is convenient to introduce auxiliary binary random variables:
(13.29)

$$\theta_q = \begin{cases} 1 & \text{if } W\left(x\left(q\Delta t\right)\right) \le W\left(x\left(\left(q-1\right)\Delta t\right)\right) - \frac{\gamma}{2}\Delta t \text{ and } \tilde{V}_0 \le W\left(x\left(\left(q-1\right)\Delta t\right)\right), \\ 1 & \text{if } W\left(x\left(q\Delta t\right)\right) \le W\left(x\left(\left(q-1\right)\Delta t\right)\right) + \frac{\gamma}{2}\Delta t \text{ and } 0 \le W\left(x\left(\left(q-1\right)\Delta t\right)\right) < \tilde{V}_0, \\ 0 & \text{otherwise.} \end{cases}$$

The following results correspond to Lemma 6.3 in the companion paper [21], but it is complicated by the fact that the solution of the blinking system has to remain in the region $W\left(x\left(t\right)\right) \le V_1$, at least for the instants $t = q\Delta t$, $q = 0, 1, \ldots$, for the bound (13.21) or (13.24) to be applicable.

Lemma 13.4. *Suppose the various constants are chosen as in Lemma 13.3, and suppose that*

$$V_1 - V_0 \ge \frac{3}{2}\gamma\Delta t. \tag{13.30}$$

For any $Q \in \mathbb{N}$, let $\sigma = (\sigma_1, \ldots, \sigma_Q) \in \{0,1\}^Q$ be a binary vector of length Q and $m = Q - \sum_{q=1}^{Q} \sigma_q$ be the number of zeros in this vector. Then the following hold:

(a) *For $m = 0$,*

$$\begin{aligned} &P\left(\theta_q = 1,\ Z_{q-1} \le V_1 \text{ for } q = 1, \ldots, Q\right) \\ &\quad = P\left(\theta_q = 1 \text{ for } q = 1, \ldots, Q\right) > \left(1 - P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right)^Q. \end{aligned} \tag{13.31}$$

(b) *For $m > 0$,*

$$P\left(\theta_q = \sigma_q,\ \ Z_{q-1} \le V_1 \text{ for } q = 1, \ldots, Q\right) \le \left[P_{W\lambda}\left(\frac{\Delta t}{\tau}\right)\right]^m. \tag{13.32}$$

Proof.

(a) It follows from $Z_0 \le V_1$ and $\theta_q = 1$ for $q = 1, \ldots, Q$ that $Z_q \le V_1$ for $q = 1, \ldots, Q$. Indeed, from $\theta_q = 1$ it follows that $Z_q < Z_{q-1}$ unless $Z_{q-1} < \tilde{V}_0$. In this last case, using (13.30), we get $Z_q \le Z_{q-1} + \frac{\gamma}{2}\Delta t \le \tilde{V}_0 + \frac{\gamma}{2}\Delta t \le V_1$. Therefore,

$$P\left(\theta_q = 1,\ Z_{q-1} \le V_1 \text{ for } q = 1, \ldots, Q\right) = P\left(\theta_q = 1 \text{ for } q = 1, \ldots, Q\right). \tag{13.33}$$

Furthermore,

$$\begin{aligned} &P\left(\theta_q = 1 \text{ for } q = 1, \ldots, Q\right) \\ &\quad = \sum_{x((Q-1)\Delta t)} P\left(\theta_q = 1 \text{ for } q = 1, \ldots, Q,\ x\left(\left(Q-1\right)\Delta t\right)\right), \end{aligned} \tag{13.34}$$

where the summation is over (the finite number of) all possible values of $x\left((Q-1)\,\Delta t\right)$. Then

$$\begin{aligned} &P\left(\theta_q = 1 \,\text{for}\, q \;=\; 1,\ldots,Q\right)\\ &\quad= \sum_{x((Q-1)\Delta t)} P\left(\theta_Q = 1 \,|\, \theta_q = 1 \;\text{ for }\; q = 1,\ldots,Q-1,\; x\left((Q-1)\,\Delta t\right)\right)\\ &\qquad \times P\left(\theta_q = 1 \,\text{for}\, q \;=\; 1,\ldots,Q-1,\;\; x\left((Q-1)\,\Delta t\right)\right). \end{aligned} \tag{13.35}$$

Since θ_Q depends only on $x\left(Q\Delta t\right)$ and $x\left((Q-1)\,\Delta t\right)$, and since $x\left(q\Delta t\right),\; q = 1,2\ldots,$ is a vector-valued Markov process, we get

$$\begin{aligned} &P\left(\theta_q = 1 \;\text{ for }\; q \;=\; 1,\ldots,Q,\; x\left((Q-1)\,\Delta t\right)\right)\\ &= \sum_{x((Q-1)\Delta t)} P\left(\theta_Q = 1 \,|x\left((Q-1)\,\Delta t\right)\right)\\ &\times P\left(\theta_q = 1 \text{ for } q = 1,\ldots,Q-1, x\left((Q-1)\,\Delta t\right)\right)\\ &> \left(1 - P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right) \sum_{x((Q-1)\Delta t)} P\left(\theta_q = 1 \,\text{for}\, q = 1,\ldots,Q-1, x\left((Q-1)\,\Delta t\right)\right)\\ &= \left(1 - P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right) \cdot P\left(\theta_q = 1 \,\text{for}\, q \;=\; 1,\ldots,Q-1\right). \end{aligned} \tag{13.36}$$

Then inequality (13.31) follows by induction. Indeed, for $Q > 0$, inequality (13.36) is the induction step and for $Q = 1$, it becomes

$$P\left(\theta_1 = 1\right) \;>\; \left(1 - P_{W\lambda}\left(\frac{\Delta t}{\tau}\right)\right), \tag{13.37}$$

which holds by (13.36).

(b) If $\sigma_Q = 1$, we simply use

$$\begin{aligned} &P\left(\theta_q = \sigma_q,\; Z_{q-1} \le V_1 \;\text{ for }\; q = 1,\ldots,Q\right)\\ &\le\; P\;\left(\theta_q = \sigma_q,\;\; Z_{q-1} \le V_1 \text{ for }\; q = 1,\ldots,Q-1\right). \end{aligned} \tag{13.38}$$

If $\sigma_Q = 0$, we write

$$\begin{aligned} &P\left(\theta_q = \sigma_q,\;\; Z_{q-1} \le V_1 \text{ for }\; q = 1,\ldots,Q\right)\\ &= \sum_{x((Q-1)\Delta t)} P\left(\theta_q = \sigma_q,\; Z_{q-1} \le V_1 \;\text{ for }\; q = 1,\ldots,Q, x\left((Q-1)\,\Delta t\right)\right), \end{aligned} \tag{13.39}$$

where the summation is over (the finite number of) all possible values of $x\left((Q-1)\,\Delta t\right)$. The condition $Z_{Q-1} \le V_1$ is a restriction on this summation. Therefore,

$$\begin{aligned} &P\left(\theta_q = \sigma_q, Z_{q-1} \le V_1 \;\text{ for }\; q = 1,\ldots,Q\right)\\ &= \sum_{\substack{x((Q-1)\Delta t)\\ W(x((Q-1)\Delta t))\le V_1}}\\ &\left[P\left(\theta_Q = 0,\; \theta_q = \sigma_q,\;\; Z_{q-1} \le V_1 \text{ for }\; q = 1,\ldots,Q-1,\; x\left((Q-1)\,\Delta t\right)\right)\right]\\ &= \sum_{\substack{x((Q-1)\Delta t)\\ W(x((Q-1)\Delta t))\le V_1}} P\left(\theta_Q = 0\,|\theta_q = \sigma_q,\; Z_{q-1} \le V_1,\; q = 1,\ldots,Q-1\right)\\ &\quad \cdot P\left(\theta_q = \sigma_q,\; Z_{q-1} \le V_1 \;\text{ for }\; q = 1,\ldots,Q-1, x\left((Q-1)\,\Delta t\right)\right). \end{aligned} \tag{13.40}$$

Again, since θ_Q depends only on $x\,(Q\Delta t)$ and $x\,((Q-1)\,\Delta t)$, and since $x\,(q\Delta t)\,,\ q = 1, 2, \ldots,$ is a vector-valued Markov process, we get

$$
\begin{aligned}
&P\left(\theta_q = \sigma_q,\ \ Z_{q-1} \le V_1 \ \text{ for } \ q = 1, \ldots, Q\right) \\
&= \sum_{\substack{x((Q-1)\Delta t) \\ W(x((Q-1)\Delta t)) \le V_1}} P\left(\theta_Q = 0 \,|\, x\,((Q-1)\,\Delta t)\right) \\
&\quad \times P\left(\theta_q = \sigma_q,\ Z_{q-1} \le V_1 \ \text{ for } \ q = 1, \ldots, Q-1,\ x\,((Q-1)\,\Delta t)\right).
\end{aligned} \tag{13.41}
$$

Using for each term in the sum the complement of (13.27) or (13.28), depending on the value of $W\,(x\,((Q-1)\,\Delta t))$, we obtain

(13.42)
$$
\begin{aligned}
&P\left(\theta_q = \sigma_q,\ \ Z_{q-1} \le V_1 \ \text{ for } \ q = 1, \ldots, Q\right) \\
&\le \sum_{\substack{x((Q-1)\Delta t) \\ W(x((Q-1)\Delta t)) \le V_1}} P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right) \cdot P\left(\theta_q = \sigma_q,\ \ Z_{q-1} \le V_1,\ q = 1, \ldots, Q-1\right) \\
&\le P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right) P\left(\theta_q = \sigma_q,\ \ Z_{q-1} \le V_1 \ \text{ for } \ q = 1, \ldots, Q-1\right).
\end{aligned}
$$

Again by induction, applying (13.27) or (13.28), depending on the value of the corresponding σ_q, proves the lemma. ∎

Lemma 13.5. *For any $V \ge 0$ and any integer Q with $Q \ge (V_1 - V_2)/\Delta t$, we have*

$$
P\left(\tilde{V}_0 \le Z_{q-1} \le V_1 \ \text{ for } \ q = 1, \ldots, Q,\ \ Z_Q > V\right) \le \sum_{n=m}^{Q} \binom{Q}{n} \left[P_{W\lambda}\left(\frac{\Delta t}{\tau}\right)\right]^n, \tag{13.43}
$$

where

$$
m = \left\lfloor \frac{Q\gamma}{2\alpha + \gamma} - \frac{V_1 - V}{\Delta t} \cdot \frac{2}{2\alpha + \gamma} \right\rfloor + 1 \tag{13.44}
$$

and $\lfloor x \rfloor$ is the integer part of x.

Proof. Suppose that $\theta_Q = \sigma_Q, \ldots,\ \theta_1 = \sigma_1$, $m = Q - \sum_{q=1}^{Q} \sigma_q$, and $\tilde{V}_0 \le Z_{q-1} \le V_1$ for $q = 1, \ldots, Q$. Then

$$
Z_Q \le V_1 - (Q - m)\,\frac{\gamma}{2}\Delta t + m\alpha\Delta t. \tag{13.45}
$$

Hence, if Z_Q is to be larger than V, we must have

$$
m > \frac{Q\,\gamma}{2\alpha + \gamma} - \frac{V_1 - V}{\Delta t} \cdot \frac{2}{2\alpha + \gamma}. \tag{13.46}
$$

The smallest m satisfying (13.46) is (13.44). Therefore,

$$
\begin{aligned}
&P\left(\tilde{V}_0 \le Z_{q-1} \le V_1 \ \text{for} \ \ q=1,\dots,Q, \ \ Z_Q > V\right) \\
&\le \sum_{\substack{\sigma_1,\dots,\sigma_Q=0 \\ Q-\sum_{q=1}^{Q}\theta_q \ge m}}^{1} P\left(\theta_q=\sigma_q, \tilde{V}_0 \le Z_{q-1} \le V_1 \ \text{for} \ \ q=1,\dots,Q, Z_Q > V\right) \\
&\le \sum_{\substack{\sigma_1,\dots,\sigma_Q=0 \\ Q-\sum_{q=1}^{Q}\theta_q \ge m}}^{1} P\left(\theta_q=\sigma_q, \tilde{V}_0 \le Z_{q-1} \le V_1 \ \text{for} \ \ q=1,\dots,Q\right) \\
&\le \sum_{n=m}^{Q} \binom{Q}{n} \left[P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right]^n. \qquad \blacksquare
\end{aligned}
\tag{13.47}
$$

13.1.2. Proof of Theorem 8.3. Having derived the lemmas, we are ready to give the proof of Theorem 8.3.

Proof.

(a) We shall first prove the first part of the theorem. Let $\lambda = \frac{\gamma}{4}$. Consider the set S_{escape}^{direct} of switching sequences such that the solution of the blinking system reaches U_∞ before reaching U_0. For each such a switching sequence there must be an integer Q such that

$$\tilde{V}_0 \le Z_{q-1} \le V_1 \quad \text{for } q=1,\dots,Q \quad \text{and} \quad Z_Q > V_1, \tag{13.48}$$

where $\tilde{V}_0 = V_0 + \gamma\,\Delta t$. If this were not the case, then there would have to be a time t (that is not a multiple of Δt) such that $x(t) \in U_\infty$ but

$$\tilde{V}_0 \le Z_q \le V_1 \quad \text{for } q=0,1,2,\dots. \tag{13.49}$$

Let the integer q be such that $q\Delta t < t < (q+1)\Delta t$. Then

$$
\begin{aligned}
&W(x(t)) \le W(x(q\Delta t)) + |W(x(t)) - W(x(q\Delta t))| \\
&\le V_1 + \alpha(t - q\Delta t) < V_1 + \alpha\Delta t,
\end{aligned}
\tag{13.50}
$$

where we have used (13.11). But (13.50) is in contradiction with $x(t) \in U_\infty$. Hence, there is an integer Q such that (13.48) holds, and therefore

$$S_{escape}^{direct} \subseteq S_+ = \bigcup_{Q=1}^{\infty} \left\{ s \,\middle|\, \tilde{V}_0 \le Z_{q-1} \le V_1 \text{ for } \ q=1,\dots,Q \ \text{ and } \ Z_Q > V_1 \right\}, \tag{13.51}$$

and thus

$$P_{escape}^{direct} \le P(S_+) \le \sum_{Q=1}^{\infty} P\left(\tilde{V}_0 \le Z_q \le V_1 \ \text{for} \ \ q=0,\dots,Q-1 \ \text{ and } \ Z_Q > V_1\right). \tag{13.52}$$

Applying Lemma 13.5 for $V = V_1$, we get

$$P(S_+) \le \sum_{Q=1}^{\infty} \sum_{n=m}^{Q} \binom{Q}{n} \left[P_{W\lambda}\left(\frac{\Delta t}{\tau}\right)\right]^n, \tag{13.53}$$

where m is given by (13.44). Hence, the double sum goes over all integer pairs (Q, n) such that

$$Q \geq 1 \text{ and } \left\lfloor \frac{Q\gamma}{2\alpha+\gamma} \right\rfloor + 1 \leq n \leq Q \Leftrightarrow n \geq 1 \text{ and } n \leq Q < \frac{2\alpha+\gamma}{\gamma} n. \tag{13.54}$$

Therefore, using the upper bound on binomial coefficients [40]

$$\binom{Q}{n} \leq \left(\frac{Qe}{n}\right)^n, \tag{13.55}$$

we get

$$\begin{aligned} P(S_+) &\leq \sum_{n=1}^{\infty} \left[P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right]^n \sum_{Q=n}^{\frac{2\alpha+\gamma}{\gamma}n} \binom{Q}{n} \leq \sum_{n=1}^{\infty} \left[P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right]^n \sum_{Q=n}^{\frac{2\alpha+\gamma}{\gamma}n} \left(\tfrac{Qe}{n}\right)^n \\ &\leq \sum_{n=1}^{\infty} \left[P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right]^n \left(\tfrac{2\alpha+\gamma}{\gamma} - 1\right) n \left(\tfrac{2\alpha+\gamma}{\gamma} n\right)^n \left(\tfrac{e}{n}\right)^n \\ &\leq \tfrac{2\alpha}{\gamma} \sum_{n=1}^{\infty} \left[P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right]^n n \left(\tfrac{2\alpha+\gamma}{\gamma} e\right)^n = \tfrac{2\alpha}{\gamma^2} \frac{(2\alpha+\gamma)eP_{W\lambda}\left(\frac{\Delta t}{\tau}\right)}{\left(1 - \frac{(2\alpha+\gamma)}{\gamma} eP_{W\lambda}\left(\frac{\Delta t}{\tau}\right)\right)^2}. \end{aligned} \tag{13.56}$$

Application of (8.5) and (13.7) guarantees that the sum in (13.56) converges and leads to (8.7).

In order to prove (8.8), we reason as follows. The solution $x(t)$ of the blinking system either reaches U_∞ before reaching U_0 or reaches U_0 before reaching U_∞ or never reaches U_∞ nor U_0. The corresponding set of switching sequences is S_{escape}^{direct} in the first case. For the other two cases let us denote it by $S_{attraction}^{direct}$ and $S_{trapped}$, respectively. Thus, the set S of all switching sequences is decomposed into

$$S = S_{escape}^{direct} \cup S_{atttraction}^{direct} \cup S_{trapped}. \tag{13.57}$$

Another decomposition is

$$S = S_+ \cup S_- \cup S_0, \tag{13.58}$$

where S_+ is given by (13.51) and

$$\begin{aligned} S_- &= \bigcup_{Q=1}^{\infty} \left\{ s \,\middle|\, \tilde{V}_0 \leq Z_{q-1} \leq V_1 \text{ for } q = 1, \ldots, Q \text{ and } Z_Q < \tilde{V}_0 \right\}, \\ S_0 &= \left\{ s \,\middle|\, \tilde{V}_0 \leq Z_{q-1} \leq V_1 \text{ for } q = 1, 2, \ldots \right\}, \end{aligned} \tag{13.59}$$

Since the solution corresponding to a switching sequence in $S_{trapped}$ is constrained by $\tilde{V}_0 \leq W(x(t)) \leq V_1 + \alpha\Delta t$ rather than $\tilde{V}_0 \leq W(x(t)) \leq V_1$, we cannot claim $S_{trapped} \subseteq S_0$. However, clearly $S_{trapped} \cap S_- = \emptyset$, and therefore

$$S_{trapped} \subseteq S_0 \cup S_+. \tag{13.60}$$

Now, for any $Q \geq 1$,

$$S_0 \subseteq \left\{ s \,\middle|\, \tilde{V}_0 \leq Z_{q-1} \leq V_1 \text{ for } q = 1, \dots, Q \;\text{ and } Z_Q \geq \;\tilde{V}_0 \right\} \\ \subseteq \left\{ s \,\middle|\, \tilde{V}_0 \leq Z_{q-1} \leq V_1 \text{ for } q = 1, \dots, Q \;\text{ and } \mathrm{Z}_Q \;\geq 0 \right\}, \tag{13.61}$$

and therefore, applying Lemma 13.5, for any $Q \geq 1$,

$$P\left(S_0\right) \leq \sum_{n=m}^{Q} \binom{Q}{n} \left[P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right]^n, \\ m = \left\lfloor \tfrac{Q\gamma}{2\alpha+\gamma} - \tfrac{2V_1}{(2\alpha+\gamma)\Delta t} \right\rfloor + 1. \tag{13.62}$$

If we choose

$$Q \geq \frac{2V_1}{\gamma \Delta t}\sqrt{\frac{4}{e}}, \tag{13.63}$$

then

$$m > \tfrac{\gamma}{2\alpha+\gamma}\left(Q - \tfrac{2V_1}{\gamma\Delta t}\right) \geq \tfrac{\gamma Q}{2\alpha+\gamma}\left(1 - \sqrt{\tfrac{e}{4}}\right) = Q\beta, \text{ where } \beta = \tfrac{\gamma}{2\alpha+\gamma}\left(1 - \sqrt{\tfrac{e}{4}}\right). \tag{13.64}$$

Therefore, using again (8.5),

$$P\left(S_0\right) \leq \sum_{n=m}^{Q} \left(\tfrac{Qe}{n}\right)^n \left[P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right]^n \leq \sum_{n=m}^{Q} \left(\tfrac{Qe}{Q\beta}\right)^n \left[P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right]^n \\ \leq \sum_{n=m}^{\infty} \left(\tfrac{e}{\beta}\right)^n \left[P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right]^n = \left[\tfrac{e}{\beta} P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right]^m \cdot \tfrac{1}{1 - \frac{e}{\beta}P_{W\lambda}\left(\frac{\Delta t}{\tau}\right)}. \tag{13.65}$$

The last equality holds because, using (8.5),

$$\frac{e}{\beta} P_{W\lambda}\left(\frac{\Delta t}{\tau}\right) = \frac{2\alpha + \gamma}{\gamma} \cdot \frac{e}{1 - \sqrt{\frac{e}{4}}} P_{W\lambda}\left(\frac{\Delta t}{\tau}\right) \leq \frac{1 - \sqrt{\frac{e}{3}}}{1 - \sqrt{\frac{e}{4}}} < 1. \tag{13.66}$$

Since (13.65) holds for all Q satisfying (13.63) and when $Q \to \infty$, also $m \to \infty$ and the right-hand side of (13.65) converges to 0. Hence, $P\left(S_0\right) = 0$, and because of (13.60) and (13.56)

$$P\left(S_{trapped}\right) \leq \frac{36\,\alpha^2}{\gamma^2} \cdot \exp\left(-\frac{c\gamma^3}{\tau}\right) \tag{13.67}$$

and

$$P_{attraction}^{direct} = 1 - P_{trapped} - P_{escape}^{direct} \geq 1 - \frac{72\,\alpha^2}{\gamma^2} \cdot \exp\left(-\frac{c\gamma^3}{\tau}\right). \tag{13.68}$$

This completes the proof of the first statement in the theorem. ∎

(b) We shall now prove the second statement. Consider the set of switching sequences such that

$$\theta_q = 1 \text{ for } q = 1, \ldots, Q \quad \text{with} \quad Q = \left\lfloor \frac{2(V_1 - V_0)}{\gamma \Delta t} \right\rfloor - 1. \tag{13.69}$$

Then

$$Z_q \le Z_{q-1} - \frac{\gamma}{2}\Delta\, t \quad \text{as long as} \quad Z_{q-1} \ge \tilde{V}_0. \tag{13.70}$$

Now, if $Z_{Q-1} \ge \tilde{V}_0$, then

$$Z_Q \le V_1 - Q\frac{\gamma}{2}\Delta t \le V_1 - \left(\frac{2(V_1 - V_0)}{\gamma \Delta t} - 2\right)\frac{\gamma}{2}\Delta t = V_0 + \gamma \Delta t = \tilde{V}_0. \tag{13.71}$$

Hence, in any case,

$$T_{attraction} \le Q\Delta t \le \frac{2(V_1 - V_0)}{\gamma}. \tag{13.72}$$

Thus, the set of switching sequences where (13.69) is satisfied contains the set of switching sequences where (13.72) is satisfied. Hence,

$$\begin{aligned} &P\left(T_{attraction} \le \tfrac{2(V_1 - V_0)}{\gamma}\right) \ge P\left(\theta_q = 1 \text{ for } q = 1, \ldots, Q\right) \\ &\ge \left(1 - P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right)^Q \ge \left(1 - Q \cdot P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right) \ge 1 - \tfrac{4(V_1 - V_0)}{\gamma \Delta t}\exp\left(-\tfrac{c\gamma^3}{\tau}\right), \end{aligned} \tag{13.73}$$

where we have used Lemma 13.4 and inequality (13.7).
To get probabilistic bounds on T_{remain}, it is convenient to reset the time to 0 after reaching U_0. Thus, we suppose that $x(0) \in U_0$. Then, if

$$\theta_q = 1 \quad \text{for}\, q = 1, \ldots, Q, \tag{13.74}$$

it follows that

$$Z_q < \tilde{V}_0 + \frac{\gamma}{2}\Delta t \quad \text{for}\, q = 0, \ldots, Q \tag{13.75}$$

because

$$\begin{aligned} &Z_{q-1} < \tilde{V}_0 \Rightarrow Z_q \le Z_{q-1} + \tfrac{\gamma}{2}\Delta t \le \tilde{V}_0 + \tfrac{\gamma}{2}\Delta t, \\ &\tilde{V}_0 \le Z_{q-1} \le \tilde{V}_0 + \tfrac{\gamma}{2}\Delta t \Rightarrow Z_q \le Z_{q-1} - \tfrac{\gamma}{2}\Delta t \le \tilde{V}_0 + \tfrac{\gamma}{2}\Delta t. \end{aligned} \tag{13.76}$$

Using (13.11), we get

$$W(x(t)) < V_0 + \left(\frac{3}{2}\gamma + \alpha\right)\Delta t \quad \text{for } t \in [0,\ Q\Delta t], \tag{13.77}$$

and thus $T_{remain} \ge Q\Delta t$. It follows that

$$\begin{aligned} P\left(T_{remain} > T\right) &> P\left(\theta_q = 1 \text{ for } q = 1, \ldots, \left\lfloor \tfrac{T}{\Delta t} \right\rfloor\right) \\ &> \left(1 - P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right)^{\left\lfloor \frac{T}{\Delta t} \right\rfloor} > 1 - \tfrac{T}{\Delta t} 2\exp\left(-\tfrac{c\gamma^3}{\tau}\right), \end{aligned} \tag{13.78}$$

which is inequality (8.11) after the substitution (8.3). ■

13.2. Theorem 9.1: Case 2, single ghost attractor.

13.2.1. Proof. Define again

$$\Delta t = \frac{\gamma}{2\left(LB_{WF} + LB_{W\Phi}\right)}. \tag{13.79}$$

(a) By definition of the attracting region R in Hypothesis 1, there is a time $T \geq 0$ such that

$$x\left(T\right) \in R \quad \text{for } t \geq T, \tag{13.80}$$

and thus

$$W\left(x\left(t\right)\right) \leq V_1 \quad \text{for } t \geq T. \tag{13.81}$$

As in the proof of Theorem 8.3(b), define the integer Q to be

$$Q = \left\lfloor 2\frac{V_1 - V_0}{\gamma \Delta t} \right\rfloor - 1. \tag{13.82}$$

By (13.73), with probability at least $1 - \frac{4(V_1-V_0)}{\gamma\Delta t}\exp(-\frac{c\gamma^3}{\tau})$ the solution then reaches U_0 in a time not longer than $Q\Delta t$. Hence, the set of switching sequences, for which the trajectory has not yet reached U_0 in time $T + Q\Delta t$, has probability at most $\frac{4(V_1-V_0)}{\gamma\Delta t}\exp(-\frac{c\gamma^3}{\tau})$. Nevertheless, since the region R is invariant, these trajectories satisfy

$$W\left(x\left(T + Q\Delta t\right)\right) \leq V_1. \tag{13.83}$$

They again with probability at least $1 - \frac{4(V_1-V_0)}{\gamma\Delta t}e^{-\frac{c\gamma^3}{\tau}}$ reach U_0 in an additional time interval of length $Q\Delta t$, etc. More precisely, we can write, thanks to the Markov property of $x\left(t\right)$,

(13.84)
$$\begin{aligned}
&P\left(x\left(t\right) \notin U_0,\ t \in [0, T + nQ\Delta t]\right) \\
&= \sum_{x((n-1)Q\Delta t)} P\left(x\left(t\right) \notin U_0,\ t \in\ [0, T + nQ\Delta t]\,,\ x\left(\left(n-1\right)Q\Delta t\right)\right) \\
&= \sum_{x((n-1)Q\Delta t)} \begin{array}{l} P\left(x\left(t\right) \notin U_0,\ t \in\ [T + \left(n-1\right)Q\Delta t, T + nQ\Delta t] \,|\, x\left(t\right) \notin U_0,\right. \\ \qquad \left. t \in [0, T + \left(n-1\right)Q\Delta t]\,,\ \ x\left(\left(n-1\right)Q\Delta t\right)\right) \\ \cdot\, P\left(x\left(t\right) \notin U_0,\ t \in [0, T + \left(n-1\right)Q\Delta t]\,,\ \ x\left(\left(n-1\right)Q\Delta t\right)\right) \end{array} \\
&= \sum_{x((n-1)Q\Delta t)} \begin{array}{l} P\left(x\left(t\right) \notin U_0,\right. \\ \qquad \left. t \in\ [T + \left(n-1\right)Q\Delta t, T + nQ\Delta t] \,|\, x\left(\left(n-1\right)Q\Delta t\right)\right) \\ \cdot\, P\left(x\left(t\right) \notin U_0,\ t \in [0, T + \left(n-1\right)Q\Delta t]\,,\ \ x\left(\left(n-1\right)Q\Delta t\right)\right) \end{array} \\
&\leq \sum_{x((n-1)Q\Delta t)} \tfrac{4(V_1-V_0)}{\gamma\Delta t}\exp\left(-\tfrac{c\gamma^3}{\tau}\right) P\left(x\left(t\right) \notin U_0,\ t \in [0, T + \left(n-1\right)Q\Delta t]\right) \\
&\leq \tfrac{4(V_1-V_0)}{\gamma\Delta t}\exp\left(-\tfrac{c\gamma^3}{\tau}\right) \cdot P\left(x\left(t\right) \notin U_0,\ t \in [0, T + \left(n-1\right)Q\Delta t]\right).
\end{aligned}$$

By repeated application of (13.84) we get

$$P\left(x\left(t\right)\notin U_0 \quad \text{for } 0\le t\le T+nQ\Delta t\right)\le\left(\frac{4\left(V_1-V_0\right)}{\gamma\Delta t}\exp\left(-\frac{c\gamma^3}{\tau}\right)\right)^n. \tag{13.85}$$

Since the expression in the parentheses is smaller than 1 by (9.2), we obtain

$$\begin{aligned}&P\left(x\left(t\right)\notin U_0 \ \text{ for } \ 0\le t<\infty\right)\\ &\quad=\lim_{n\to\infty}P\left(x\left(t\right)\notin U_0 \quad \text{for } 0\le t\le T+nQ\Delta t\right)=0,\end{aligned} \tag{13.86}$$

which means that almost all trajectories reach U_0 in finite time for any initial state. This proves assertion (a).

(b) If the initial state already satisfies $W\left(x\left(0\right)\right)\le V_1$, we can set $T=0$ in (13.85), which implies (9.4). This completes the proof of assertion (b).

(c) For $q\ge 1$ and $\tilde V_0=V_0+\gamma\,\Delta t$,

$$\begin{aligned}P\left(Z_q\le V_1-\tfrac{m\gamma}{2}\Delta t\right)\ge P&\left(Z_q\le V_1-\tfrac{m\gamma}{2}\Delta t,\ \ \tilde V_0\le Z_{q-1}\le V_1\right)\\ &+P\left(Z_q\le\ V_1-\tfrac{m\gamma}{2}\Delta t,\ \ 0\le Z_{q-1}<\tilde V_0\right).\end{aligned} \tag{13.87}$$

Since

$$\left\{s\,|\theta_q=1 \ \text{ and } \ \tilde V_0\le Z_{q-1}\le V_1-\frac{(m-1)\,\gamma}{2}\Delta t\right\}\subseteq\left\{s\,|\ Z_q\le V_1-\frac{m\gamma}{2}\Delta t\right\} \tag{13.88}$$

for $1\le m\le Q$ (for $m>Q$ the result is useless),

(13.89)

$$\begin{aligned}&P\left(Z_q\le V_1-\tfrac{m\gamma}{2}\Delta t,\ \ \tilde V_0\le Z_{q-1}\le V_1\right)\\ &\ge P\left(\theta_q=1,\ \ Z_{q-1}\le V_1-\tfrac{(m-1)\gamma}{2}\Delta t,\ \ \tilde V_0\le Z_{q-1}\le V_1\right)\\ &=\sum_{x((q-1)\Delta t)}P\left(\theta_q=1,\ \ \tilde V_0\le Z_{q-1}\le V_1-\tfrac{(m-1)\gamma}{2}\Delta t,\ x\left(\left(q-1\right)\Delta t\right)\right)\\ &=\sum_{\tilde V_0\le W(x((q-1)\Delta t))\le V_1-\frac{(m-1)\gamma\Delta t}{2}}P\left(\theta_q=1\,|x\left(\left(q-1\right)\Delta t\right)\right)P\left(x\left(\left(q-1\right)\Delta t\right)\right).\end{aligned}$$

Applying (13.18), we get

$$\begin{aligned}&P\left(Z_q\le\ V_1-\tfrac{m\gamma}{2}\Delta t,\ \ \tilde V_0\le Z_{q-1}\le V_1\right)\\ &\ge\sum_{\tilde V_0\le W(x((q-1)\Delta t))\le\ V_1-\frac{(m-1)\gamma}{2}\Delta t}\left(1-P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right)\cdot P\left(x\left(\left(q-1\right)\Delta t\right)\right)\\ &=\left(1-P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right)\cdot P\left(\tilde V_0\le Z_{q-1}\le V_1-\tfrac{(m-1)\gamma}{2}\Delta t\right).\end{aligned} \tag{13.90}$$

Instead of (13.88), we can write

$$\left\{s\,|\theta_q=1 \ \text{ and } \ 0\le Z_{q-1}\le V_1-\frac{(m+1)\,\gamma}{2}\Delta t\right\}\subseteq\left\{s\,|\ Z_q\le V_1-\frac{m\gamma}{2}\Delta t\right\} \tag{13.91}$$

for any $m \geq 1$, and thus

$$\begin{aligned} &P\left(Z_q \leq V_1 - \tfrac{m\gamma}{2}\Delta t,\ \ 0 \leq Z_{q-1} < \tilde{V}_0\right) \\ &\quad \geq P\left(\theta_q = 1,\ \ Z_{q-1} \leq V_1 - \tfrac{(m+1)\gamma}{2}\Delta t,\ \ 0 \leq Z_{q-1} < \tilde{V}_0\right). \end{aligned} \tag{13.92}$$

For $1 \leq m \leq Q-2$,

$$V_1 - \frac{(m+1)\,\gamma}{2}\Delta t \geq \tilde{V}_0, \tag{13.93}$$

and therefore

$$\begin{aligned} &P\left(Z_q \leq V_1 - \tfrac{m\gamma}{2}\Delta t,\ \ 0 \leq Z_{q-1} < \tilde{V}_0\right) \geq P\left(\theta_q = 1,\ \ 0 \leq Z_{q-1} < \tilde{V}_0\right) \\ &= \sum_{x((q-1)\Delta t)} P\left(\theta_q = 1,\ \ 0 \leq Z_{q-1} < \tilde{V}_0,\ x\left((q-1)\,\Delta t\right)\right) \\ &= \sum_{0 \leq W(x((q-1)\Delta t)) \leq \tilde{V}_0} P\left(\theta_q = 1 \mid x\left((q-1)\,\Delta t\right)\right) \cdot P\left(x\left((q-1)\,\Delta t\right)\right) \\ &\geq \sum_{0 \leq W(x((q-1)\Delta t)) \leq \tilde{V}_0} \left(1 - P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right) \cdot P\left(x\left((q-1)\,\Delta t\right)\right) \\ &= \left(1 - P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right) \cdot P\left(0 \leq Z_{q-1} < \tilde{V}_0\right), \end{aligned} \tag{13.94}$$

where we have used (13.19).

Combining (13.90) and (13.94), the following inequality holds for $q \geq 1$ and $1 \leq m \leq Q-2$:

$$P\left(Z_q \leq V_1 - \frac{m\gamma}{2}\Delta t\right) \geq \left(1 - P_{W\lambda}\left(\frac{\Delta t}{\tau}\right)\right) \cdot P\left(Z_{q-1} \leq V_1 - \frac{(m-1)\,\gamma}{2}\Delta t\right). \tag{13.95}$$

Since we suppose $x(0) \in R$, R is invariant, and since by Hypothesis 3 $W(x(0)) \leq V_1$, we have

$$P\left(Z_q \leq V_1\right) = 1 \quad \text{for } q \geq 0. \tag{13.96}$$

Applying now (13.95) iteratively, we obtain

$$P\left(Z_q \leq V_1 - \frac{m\gamma}{2}\Delta t\right) \geq \left(1 - P_{W\lambda}\left(\frac{\Delta t}{\tau}\right)\right)^m \qquad \text{for } q \geq m \ \text{ and } \ 1 \leq m \leq Q-2. \tag{13.97}$$

We now rewrite this result. By the definition (13.82) of Q, for any n,

$$V_0 + \frac{n\gamma}{2}\Delta t > V_1 - \frac{(Q-n+2)\,\gamma}{2}\Delta t, \tag{13.98}$$

and therefore

$$\begin{aligned} &P\left(Z_q \leq V_0 + \tfrac{n\gamma}{2}\Delta t\right) \geq P\left(Z_q \leq V_1 - \tfrac{(Q-n+2)\gamma}{2}\Delta t\right) \\ &\geq \left(1 - P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right)^{Q-n+2} \quad \text{for } q \geq Q-n+2 \ \text{ and } \ 4 \leq n \leq Q+1. \end{aligned} \tag{13.99}$$

For $n = 4$,

$$P\left(Z_q > V_0 + 2\gamma\Delta t\right) = 1 - P\left(Z_q \le V_0 + 2\gamma\Delta t\right) \\ \le 1 - \left(1 - P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right)^{Q-2} \le (Q-2)\, P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right) \quad \text{for } q \ge Q-2 \\ \le \tfrac{2V_1}{\gamma\Delta t} P_{W\lambda}\left(\tfrac{\Delta t}{\tau}\right) \quad \text{for } q \ge Q-2. \tag{13.100}$$

This can be rewritten, using (13.7), as

$$P\left(W\left(x\left(q\Delta t\right)\right) > V_0 + \tfrac{\gamma^2}{LB_{WF}+LB_{W\Phi}}\right) \\ \le \tfrac{8(LB_{WF}+LB_{W\Phi})V_1}{\gamma^2} \exp\left(-\tfrac{c\gamma^3}{\tau}\right) \quad \text{for } q \ge Q-2. \tag{13.101}$$

To get a probabilistic bound on the deviation from the attractor not only for times that are multiples of Δt but for all positive t, we note that at any time, in particular in the time interval $(0, \Delta t)$, $W(t) \le V_1$, and therefore we can apply (13.101) to any time-shifted solution; i.e., we can set $x(t)$ for $0 < t < \Delta t$ 0 as a new initial condition and apply (13.101) to the new solution. From this we obtain assertion (c).

13.3. Theorem 10.1: Case 3, multiple invariant attractors.

13.3.1. Preliminary lemmas. The proof of Theorem 10.1 closely follows that of Theorem 8.3. It starts with preliminary lemmas that are analogous to Lemmas 13.1, 13.2, 13.3, 13.4, and 13.5, except that the logarithm of the Lyapunov function $\ln W$ is used instead of W. We start off with discussing some properties of the blinking and averaged systems in the case where the attractors of the averaged system are invariant under the blinking system.

Equation (10.1) implies that along any solution within R of the blinking system the following inequalities hold, as long as $W\left(x\left(t\right)\right) \ne 0$:

$$\left|\frac{d}{dt}\left[\ln W\left(x\left(t\right)\right)\right]\right| \le B_{\ln WF} \quad \text{and} \quad \left|\frac{d}{dt} D_F \ln W\left(x\left(t\right), \tilde{s}\right)\right| = LB_{\ln WF}. \tag{13.102}$$

Equations (10.1) and (10.2) imply that along any solution of the averaged system with $0 < W\left(x\left(0\right)\right) \le V_1$

$$-B_{\ln W\Phi} \le \frac{d}{dt}\left[\ln W\left(\xi\left(t\right)\right)\right] \le -\gamma \quad \text{and} \quad \left|\frac{d}{dt} D_\Phi \ln W\left(\xi\left(t\right), \tilde{s}\right)\right| \le LB_{\ln W\Phi}. \tag{13.103}$$

The first equations of (13.102) and (13.103) can be rewritten as

$$\left|\tfrac{d}{dt}\left[W\left(x\left(t\right)\right)\right]\right| \le B_{\ln WF} W\left(x\left(t\right)\right), \\ -B_{\ln W\Phi} W\left(\xi\left(t\right)\right) \le \tfrac{d}{dt}\left[W\left(\xi\left(t\right)\right)\right] \le -\gamma W\left(\xi\left(t\right)\right). \tag{13.104}$$

Note that the first inequality is possible only if A is an invariant set of the blinking system. On the other hand, if W is twice continuously differentiable with a positive definite Hessian and F is continuously differentiable with respect to x with a nonsingular Jacobian matrix on A, then Hypothesis 4 is satisfied.

Given Hypothesis 4, all properties of the time dependence of W previously proved during the first step now carry over to $\ln W$.

We still need to extend the probabilities of exceptional mean values from W to $\ln W$. We define for integer values of K

$$\mathrm{P}_{\ln W\lambda}(K) = \max_{x\in R} \mathrm{P}\left\{\left|\frac{1}{K}\sum_{k=1}^{K} D_F \ln W\left(x, S^k\right) - D_\Phi \ln W(x)\right| > \lambda\right\} \tag{13.105}$$

and for noninteger K

$$\begin{aligned}&\mathrm{P}_{\ln W\lambda}(K)\\ &= \max_{x\in R} \mathrm{P}\Bigg\{\Bigg|\tfrac{1}{K}\Bigg[\sum_{k=1}^{\lfloor K\rfloor} D_F \ln W\left(x, S^k\right) + (K-\lfloor K\rfloor)\, D_F \ln W\left(x, S^{\lfloor K\rfloor+1}\right)\Bigg]\\ &\qquad - D_\Phi \ln W(x)\Bigg| > \lambda\Bigg\}.\end{aligned} \tag{13.106}$$

Because of the stationarity of the process, by analogy to (13.10), for any $t \geq 0,\ \Delta t \geq 0,\ \lambda > 0,\ x \in R$ we have

$$P\left(\left|\int_t^{t+\Delta t}\left[D_F \ln W(x, s(u)) - D_\Phi \ln W(x)\right] du\right| > \lambda\cdot\Delta t\right) \leq P_{\ln W\lambda}\left(\frac{\Delta t}{\tau}\right). \tag{13.107}$$

Furthermore, Hoeffding's inequality [39] gives the following statement for this case.

Lemma 13.6. *For any* $\lambda > 0$,

$$P_{\ln W\lambda}(K) \leq 2\exp\left(-\frac{K\lambda^2}{2\,B^2_{\ln WF}}\right). \tag{13.108}$$

As before, we proceed in two steps. In the first step, we show that the Lyapunov function $\ln W$ decreases with high probability after a certain time Δt. In the second step, we analyze the behavior of the blinking system for large times.

First step.

Lemma 13.7. *Consider a solution $x(.)$ of the blinking system with switching period τ. Choose a time $t \geq 0$ and the solution of the averaged system with $\xi(t) = x(t) \notin A$. Then the following hold:*

(a) *For any $\Delta t \geq 0$,*

$$|\ln W(x(t+\Delta t)) - \ln W(\xi(t+\Delta t))| \leq (B_{\ln WF} + B_{\ln W\Phi})\,\Delta t. \tag{13.109}$$

(b) *For any $l > 0$ and $\Delta t \geq 0$, the conditional probability that*

$$|\ln W(x(t+\Delta t)) - \ln W(\xi(t+\Delta t))| \leq \frac{LB_{\ln WF} + LB_{\ln W\Phi}}{2}\Delta t^2 + \lambda\Delta t \tag{13.110}$$

holds, given the value of $x(t)$, is at least $1 - P_{\ln W\lambda}\left(\frac{\Delta t}{\tau}\right)$.

Proof. The proof is identical to the proof of Lemma 13.2, except that W has to be replaced everywhere by $\ln W$. ■

Lemma 13.8. *Let*

$$\Delta t = \frac{\gamma}{2\left(LB_{\ln WF} + LB_{\ln W\Phi}\right)} \quad \textit{and} \quad \lambda = \frac{\gamma}{4}; \tag{13.111}$$

then for any $t \geq 0$ *and for any solution* $x(.)$ *of the blinking system with* $0 < W(x(t)) \leq V_1$ *and switching period* τ*, the conditional probability that*

$$\ln W(x(t+\Delta t)) \leq \ln W(x(t)) - \frac{\gamma}{2}\Delta t \tag{13.112}$$

holds, given $x(t)$*, is at least* $1 - P_{\ln W\lambda}\left(\frac{\Delta t}{\tau}\right)$.

Proof. Consider a solution $\xi(\cdot)$ of the averaged system with $\xi(t) = x(t)$. Then, by (10.2), one obtains

$$\ln W(\xi(t+\Delta t)) = \ln W(\xi(t)) + \int_t^{t+\Delta t} D_\Phi W(\xi(u))\, du \leq \ln W(x(t)) - \gamma\Delta t. \tag{13.113}$$

Using (13.110) and (13.111), we get

$$\begin{aligned} \ln W(x(t+\Delta t)) &\leq \ln W(\xi(t+\Delta t)) + \left|\ln W(x(t+\Delta t)) - \ln W(\xi(t+\Delta t))\right| \\ &\leq \ln W(x(t)) - \gamma\Delta t + \tfrac{\gamma}{2}\Delta t. \end{aligned} \tag{13.114}$$

■

Second step. For any choice of the initial state $x(0)$ with $0 < W(x(0)) \leq V_1$, for any choice of the switching period $\tau > 0$, and for the constants γ, λ, and Δt given by (10.2) and (13.111), we consider the following sequence of random variables on the probability space of switching sequences:

$$Z_q(s) = \ln W(x(q\Delta t)), \quad q = 0, 1, 2, \ldots. \tag{13.115}$$

By hypothesis, Z_0 is concentrated on a single value that is smaller than or equal to $\ln V_1$. Again, $\{Z_q\}$ is not a Markov process. We again introduce the additional random variables

$$\theta_q = \begin{cases} 1 & \text{if} \quad \ln W(x(q\Delta t)) \leq \ln W(x((q-1)\Delta t)) - \frac{\gamma}{2}\Delta t \quad \text{and} \\ & \qquad 0 < W(x((q-1)\Delta t)) \leq V_1, \\ 0 & \qquad \text{otherwise}. \end{cases} \tag{13.116}$$

Lemma 13.9. *Suppose the various constants are chosen as in Lemma* 13.8. *Let* $\sigma = (\sigma_1, \ldots, \sigma_Q) \in \{0,1\}^Q$ *be a binary vector of length* Q *and* $m = Q - \sum_{q=1}^{Q} \sigma_q$ *be the number of zeros in this vector. Then, for* $m > 0$,

$$P\left(\theta_q = \sigma_q,\ \ Z_{q-1} \leq \ln V_1 \ \textit{for}\ q = 1, \ldots, Q\right) \leq \left[P_{\ln W\lambda}\left(\frac{\Delta t}{\tau}\right)\right]^m. \tag{13.117}$$

Proof. The proof is the same as the proof of Lemma 13.4(b), except that V_1, W, and $P_{W\lambda}$ have to be replaced by $\ln V_1$, $\ln W$, and $P_{\ln W\lambda}$ and that no lower bound constraint has to be observed for Z_q. ■

Lemma 13.10. *For any* $V \geq 0$ *and any* $Q \in \mathbb{N}$*, we have*

$$P\left(Z_{q-1} \leq \ln V_1 \ \textit{for}\ q = 1, \ldots, Q,\ \ Z_Q > \ln V\right) \leq \sum_{n=m}^{Q} \binom{Q}{n} \left[P_{\ln W\lambda}\left(\frac{\Delta t}{\tau}\right)\right]^n, \tag{13.118}$$

where

$$m = \left\lfloor \frac{Q\gamma}{2\alpha+\gamma} - \frac{\ln V_1 - \ln V}{\Delta t} \cdot \frac{2}{2\alpha+\gamma} \right\rfloor + 1 \quad (13.119)$$

and $\lfloor x \rfloor$ *is the integer part of* x *and* $\alpha = B_{\ln WF} + B_{\ln W\Phi}$.

Proof. The proof is the same as the proof of Lemma 13.5, except that V_1, V, and $P_{W\lambda}$ have to be replaced by $\ln V_1$, $\ln V$, and $P_{\ln W\lambda}$ and that no lower bound constraint has to be observed for Z_q. ■

13.3.2. Proof of Theorem 10.1.

Proof. Let $\lambda = \frac{\gamma}{4}$.

(a) Consider the set S_{escape} of switching sequences such that the solution of the blinking system reaches U_∞. By the same argument as in the proof of Theorem 8.3(a), substituting $\ln W$ for W and disregarding the lower bound on W, for each such switching sequence there must be an integer Q such that

$$Z_{q-1} \le \ln V_1 \quad \text{for } q = 1, \dots, Q \text{ and } Z_Q > \ln V_1. \quad (13.120)$$

Therefore

$$S_{escape} \subseteq S_+ = \bigcup_{Q=1}^{\infty} \{ s \mid Z_{q-1} \le \ln V_1 \quad \text{for } q = 1, \dots, Q \quad \text{and} \quad Z_Q > \ln V_1 \}, \quad (13.121)$$

and thus

(13.122)

$$P_{escape} \le P(S_+) = \sum_{Q=1}^{\infty} P(Z_q \le \ln V_1 \ \text{ for } \ q = 0, \dots, Q-1 \ \text{ and } \ Z_Q > \ln V_1).$$

Applying Lemma 13.10 for $V = V_1$, we get

$$P(S_+) \le \sum_{Q=1}^{\infty} \sum_{n=m}^{Q} \binom{Q}{n} \left[P_{\ln W\lambda} \left(\frac{\Delta t}{\tau} \right) \right]^n, \quad (13.123)$$

where

$$m = \left\lfloor \frac{Q\gamma}{2\alpha+\gamma} \right\rfloor + 1.$$

As in the proof of Theorem 8.3(a), it follows from (10.3) (which implies (8.5)) and (13.108) that the sum in (13.123) converges and is bounded by (10.4).

(b) We apply Lemma 13.10 for $V = V_1 \exp(-\frac{Q}{4}\gamma \Delta t)$. Then

$$\begin{aligned} &P\left(Z_{q-1} \le \ln V_1 \ \text{ for } \ q = 1, \dots, Q, \ \ Z_Q > \ln V_1 - \tfrac{Q}{4}\gamma\Delta t \right) \\ &\quad \le \sum_{n=m}^{Q} \binom{Q}{n} \left[P_{\ln W\lambda} \left(\tfrac{\Delta t}{\tau} \right) \right]^n, \end{aligned} \quad (13.124)$$

where

$$m = \left\lfloor \frac{Q\gamma}{2\,(2\alpha+\gamma)} \right\rfloor + 1. \tag{13.125}$$

Consider the set S_0 of switching sequences such that there exists a natural number Q such that $Z_{q-1} \leq \ln V_1$ for $q = 1, \dots, Q$ and $Z_Q > \ln V_1 - \frac{Q}{4}\gamma\Delta t$, i.e.,

$$S_0 = \bigcup_{Q=1}^{\infty} \left\{ s |\, Z_{q-1} \leq \ln V_1 \ \text{ for } \ q = 1, \dots, Q, \ \ Z_Q > \ln V_1 - \frac{Q}{4}\gamma\Delta t \right\}. \tag{13.126}$$

Then

$$\begin{aligned} P\left(S_0\right) \leq\ & \sum_{Q=1}^{\infty} P\left(Z_{q-1} \leq \ln V_1 \ \text{ for } \ q = 1, \dots, Q, \ Z_Q > \ln V_1 - \tfrac{Q}{4}\gamma\Delta t\right) \\ \leq\ & \sum_{Q=1}^{\infty} \sum_{n=m}^{Q} \binom{Q}{n} \left[P_{\ln W\lambda}\left(\tfrac{\Delta t}{\tau}\right)\right]^n, \end{aligned} \tag{13.127}$$

where m is given by (13.125).
The double summation over n and Q has the constraints

$$Q \geq 1 \text{ and } \left\lfloor \frac{Q\gamma}{2\,(2\alpha+\gamma)} \right\rfloor + 1 \leq n \leq Q \ \Leftrightarrow\ n \geq 1 \text{ and } n \leq Q < \frac{2\,(2\alpha+\gamma)}{\gamma} n. \tag{13.128}$$

Following the same path as in the proof of Theorem 8.3(a), we obtain the bound

$$P\left(S_0\right) \leq \frac{4\alpha+\gamma}{\gamma} \cdot \frac{\frac{2(2\alpha+\gamma)}{\gamma} e \cdot P_{\ln W\lambda}\left(\frac{\Delta t}{\tau}\right)}{\left(1 - \frac{2(2\alpha+\gamma)}{\gamma} e \cdot P_{\ln W\lambda}\left(\frac{\Delta t}{\tau}\right)\right)^2}. \tag{13.129}$$

Applying (10.3), (13.108), and $\gamma \leq \alpha$, we obtain

$$P\left(S_0\right) \leq \frac{180\,\alpha^2}{\gamma^2} \cdot \exp\left(-\frac{c\gamma^3}{\tau}\right). \tag{13.130}$$

Let S_- denote the following set of all switching sequences:

$$S_- = \left\{ s \,\middle|\, Z_Q \leq \ln V_1 - \frac{Q}{4}\gamma\Delta t \ \text{ for all } Q \in \mathbb{Z}_+ \right\}. \tag{13.131}$$

Then, if S denotes the set of all switching sequences,

$$S \backslash S_- = \bigcup_{Q=1}^{\infty} \left\{ s \,\middle|\, Z_Q > \ln V_1 - \frac{Q}{4}\gamma\,\Delta t \right\}. \tag{13.132}$$

If for a switching sequence s we have $Z_Q > \ln V_1 - \frac{Q}{4}\gamma\,\Delta t$, then either we have $Z_q \leq \ln V_1$ for $q = 0, \dots, Q-1$ and $Z_Q > \ln V_1 - \frac{Q}{4}\gamma\Delta t$, or for some $1 \leq q \leq Q-1$ we have $Z_q > \ln V_1$. This implies

$$S \backslash S_- \subseteq S_0 \cup S_+, \tag{13.133}$$

and thus

$$P\left(S\backslash S_-\right) \le P\left(S_0\right) + P\left(S_+\right) \le \tfrac{216\,\alpha^2}{\gamma^2} \cdot \exp\left(-\tfrac{c\gamma^3}{\tau}\right). \tag{13.134}$$

This is equivalent to

$$P\left(S_-\right) > 1 - \frac{216\,\alpha^2}{\gamma^2} \cdot \exp\left(-\frac{c\gamma^3}{\tau}\right). \tag{13.135}$$

Suppose now that a switching sequence $s \in S_-$. Then, for any $t > 0$, we can write $t = (q+\mu)\,\Delta t$, with $0 \le \mu < 1$, and thanks to (13.102) we get

$$\begin{aligned}\ln W\left(x\left(t\right)\right) &\le \ln W\left(x\left(q\Delta t\right)\right) + \alpha\mu\Delta t \le \ln V_1 - \tfrac{q}{4}\gamma\Delta t + \alpha\mu\Delta t\\ &= \ln V_1 - \tfrac{q+\mu}{4}\gamma\Delta t + \tfrac{\mu}{4}\gamma\Delta t + \alpha\mu\Delta t = \ln V_1 - \tfrac{\gamma}{4}t + \left(\alpha + \tfrac{\gamma}{4}\right)\Delta t.\end{aligned} \tag{13.136}$$

This, together with (13.135), implies part (b) of the theorem. ■

13.4. Theorem 11.1: Case 4, unique invariant attractor.

13.4.1. Proof of Theorem 11.1.

Proof. We again define

$$\Delta t = \frac{\gamma}{2\left(LB_{\ln WF} + LB_{\ln W\Phi}\right)}. \tag{13.137}$$

To prove the almost sure convergence with exponential speed to A, we suppose that $x\left(0\right) \in R$, and thus $x\left(t\right) \in R$ and $W\left(x\left(t\right)\right) \le V_1$ for all $t \ge 0$. This is no restriction of generality since if the initial state lies outside of R, the region R is reached in a finite time. This additional time does not change the exponential speed of convergence.

Consider the set of switching sequences for which the speed of convergence to A is not exponential with exponential speed at least $\frac{\gamma}{4}$:

$$S_{\text{exponential}} = \left\{s \mid \exists K \text{ such that for all } t \ge 0 \quad W\left(x\left(t\right)\right) \le K \exp\left(-\frac{\gamma}{4}t\right)\right\}. \tag{13.138}$$

Actually, it is sufficient to impose the bound on a sequence of equally spaced discrete time instants, e.g.,

$$S_{\text{exponential}} = \left\{s \mid \exists K \text{ such that for all Q} \ge 1 \quad W\left(x\left(Q\Delta t\right)\right) \le K \exp\left(-\frac{\gamma}{4}Q\Delta t\right)\right\}, \tag{13.139}$$

because due to (13.102) it follows from $W\left(x\left(q\Delta t\right)\right) \le K \exp\left(-\tfrac{\gamma}{4}q\Delta t\right)$ and $0 \le \mu < 1$ that

$$\begin{aligned}\ln W\left(\left(q+\mu\right)\Delta t\right) &\le \ln W\left(q\Delta t\right) + \alpha\mu\Delta t \le \ln K - \tfrac{\gamma}{4}q\Delta t + \alpha\mu\Delta t\\ &= \ln K - \tfrac{\gamma}{4}\left(q+\mu\right)\Delta t + \left(\alpha + \tfrac{\gamma}{4}\right)\Delta t,\end{aligned} \tag{13.140}$$

and thus for all $t \ge 0$

$$W\left(t\right) \le \tilde{K} \exp\left(-\frac{\gamma}{4}t\right), \tag{13.141}$$

where

$$\ln \tilde{K} = \ln K + \left(\alpha + \frac{\gamma}{4}\right) \Delta t. \tag{13.142}$$

Now we restrict ourselves to the switching sequences satisfying $W(x(Q\Delta t)) \le K \exp\left(-\frac{\gamma}{4} Q\Delta t\right)$ asymptotically for $Q \to \infty$ with the constant V_1:

$$\tilde{S}_{\text{exponential}} = \left\{ s \mid \exists M \text{ such that for all } Q \ge M \quad W(x(Q\Delta t)) \le V_1 e^{\frac{\gamma}{4} Q\Delta t} \right\}. \tag{13.143}$$

Clearly, $\tilde{S}_{\text{exponential}} \subseteq S_{\text{exponential}}$, and if S denotes the set of all switching sequences,

$$\begin{aligned} S \backslash \tilde{S}_{\text{exponential}} &= \{ s \mid \text{for all } M \ge 1\ \exists Q \ge M \text{ such that } W(x(Q\Delta t)) \} \\ &> V_1 \exp\left(-\tfrac{\gamma}{4} Q\Delta t\right) = \bigcap_{M=1}^{\infty} \bigcup_{Q=M}^{\infty} \left\{ s \mid W(x(Q\Delta t)) > V_1 \exp\left(-\tfrac{\gamma}{4} Q\Delta t\right) \right\}, \end{aligned} \tag{13.144}$$

and thus

$$\begin{aligned} P\left(S \backslash \tilde{S}_{\text{exponential}}\right) &= \lim_{M \to \infty} P\left(\bigcup_{Q=M}^{\infty} \left\{ s \mid W(x(Q\Delta t)) > V_1 \exp\left(-\tfrac{\gamma}{4} Q\Delta t\right) \right\} \right) \\ &\le \lim_{M \to \infty} \sum_{Q=M}^{\infty} P\left(\left\{ s \mid W(x(Q\Delta t)) > V_1 \exp\left(-\tfrac{\gamma}{4} Q\Delta t\right) \right\} \right). \end{aligned} \tag{13.145}$$

Since $W(x(t)) \le V_1$ for all $t \ge 0$,

$$\begin{aligned} P\left(S \backslash \tilde{S}_{\text{exponential}}\right) &\le \lim_{M \to \infty} \sum_{Q=M}^{\infty} P\left(Z_Q > \ln V_1 - \tfrac{Q}{4} \gamma \Delta t \right) \\ &= \lim_{M \to \infty} \sum_{Q=M}^{\infty} P\left(Z_{q-1} \le \ln V_1 \text{ for } q = 1, \ldots, Q,\ Z_Q > \ln V_1 - \tfrac{Q}{4} \gamma \Delta t \right). \end{aligned} \tag{13.146}$$

In the proof of Theorem 10.1 it has been shown that

$$\sum_{Q=1}^{\infty} P\left(Z_{q-1} \le \ln V_1 \text{ for } q = 1, \ldots, Q, \quad Z_Q > \ln V_1 - \frac{Q}{4} \gamma \Delta t \right) < \infty, \tag{13.147}$$

which implies that $P(S \backslash \tilde{S}_{\text{exponential}}) = 0$, and thus

$$P\left(S_{\text{exponential}}\right) \ge P\left(\tilde{S}_{\text{exponential}}\right) = 1 - P\left(S \backslash \tilde{S}_{\text{exponential}}\right) = 1. \tag{13.148}$$

This indeed proves that for almost all switching sequences the solution of the blinking system converges exponentially fast to A with exponential speed at least $\frac{\gamma}{4}$. The fact that for all $t \ge 0$

$$W(x(t)) \le K \exp\left(-\frac{\gamma}{4} t\right) \tag{13.149}$$

with K given by (11.3) and probability given by (11.2) has already been shown in the proof of Theorem 10.1. ■

REFERENCES

[1] S. H. Strogatz, *Exploring complex networks*, Nature, 410 (2001), pp. 268–276.

[2] A. L. Barabási and R. Albert, *Emergence of scaling in random networks*, Science, 286 (1999), pp. 509–512.

[3] M. E. J. Newman, *The structure and function of complex networks*, SIAM Rev., 45 (2003), pp. 167–256.

[4] L. M. Pecora and T. L. Carroll, *Master stability function for synchronized coupled systems*, Phys. Rev. Lett., 80 (1998), pp. 2109–2112.

[5] V. Belykh, I. Belykh, and M. Hasler, *Connection graph stability method for synchronized coupled chaotic systems*, Phys. D, 195 (2004), pp. 159–187.

[6] I. Belykh, V. Belykh, and M. Hasler, *Generalized connection graph method for synchronization in asymmetrical networks*, Phys. D, 224 (2006), pp. 42–51.

[7] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwanga, *Complex networks: Structure and dynamics*, Phys. Rep., 424 (2006), pp. 175–308.

[8] V. S. Afraimovich and L. A. Bunimovich, *Dynamical networks: Interplay of topology, interactions and local dynamics*, Nonlinearity, 20 (2007), pp. 1761–1771.

[9] T. Nishikawa and A. E. Motter, *Network synchronization landscape reveals compensatory structures, quantization, and the positive effect of negative interactions*, Proc. Natl. Acad. Sci. USA, 107 (2010), pp. 10342–10347.

[10] T. Stojanovsky, L. Kocarev, U. Parlitz, and R. Harris, *Sporadic driving of dynamical systems*, Phys. Rev. E (3), 55 (1997), pp. 4035–4048.

[11] I. Belykh, V. Belykh, and M. Hasler, *Blinking model and synchronization in small-world networks with a time-varying coupling*, Phys. D, 195 (2004), pp. 188–206.

[12] M. Hasler and I. Belykh, *Blinking long-range connections increase the functionality of locally connected networks*, IEICE Trans. Fund., E88-A (2005), pp. 2647–2655.

[13] J. D. Skufca and E. M. Bollt, *Communication and synchronization in disconnected networks with dynamic topology: Moving neighborhood networks*, Math. Biosci. Eng., 1 (2004), pp. 347–359.

[14] M. Porfiri, D. J. Stilwell, E. M. Bollt, and J. D. Skufca, *Random talk: Random walk and synchronizability in a moving neighborhood network*, Phys. D, 224 (2006), pp. 102–113.

[15] M. Porfiri and R. Pigliacampo, *Master-slave global stochastic synchronization of chaotic oscillators*, SIAM J. Appl. Dyn. Syst., 7 (2008), pp. 825–842.

[16] T. Gorochowski, M. Di Bernardo, and C. Groerson, *Evolving enhanced topologies for the synchronization of dynamical complex networks*, Phys. Rev. E (3), 81 (2010), 056212.

[17] P. De Lellis, M. Di Bernardo, F. Garofalo, and M. Porfiri, *Evolution of complex networks via edge snapping*, IEEE Trans. Circuits Syst. I Fundam. Theory Appl., 57 (2010), pp. 2132–2143.

[18] F. Sorrentino and E. Ott, *Adaptive synchronization of dynamics on evolving complex networks*, Phys. Rev. Lett., 100 (2008), 114101.

[19] P. So, B. Cotton, and E. Barreto, *Synchronization in interacting populations of heterogeneous oscillators with time-varying coupling*, Chaos, 18 (2008), 037114.

[20] P. So and E. Barreto, *Generating macroscopic chaos in a network of globally coupled phase oscillators*, Chaos, 21 (2011), 033127.

[21] M. Hasler, V. Belykh, and I. Belykh, *Dynamics of stochastically blinking systems. Part* I: *Finite time properties*, SIAM J. Appl. Dyn. Syst., 12 (2013), pp. 1007–1030.

[22] C. Tse and M. Di Bernardo, *Complex behavior in switching power converters*, Proc. IEEE, 90 (2002), pp. 768–781.

[23] A. Stankovic and H. Lev-Ari, *Randomized modulation in power electronic converters*, Proc. IEEE, 90 (2002), pp. 782–799.

[24] L. Chua and T. Roska, *Cellular Neural Networks and Visual Computing*, Cambridge University Press, Cambridge, UK, 2002.

[25] G. Seiler and J. Nossek, *Winner-take-all cellular neural network*, IEEE Trans. Circuits Syst. II, 40 (1993), pp. 184–190.

[26] J.-L. Lagrange, *Mécanique Analytique*, Vols. I and II, Albert Blanchard ed., Paris, 1788.

[27] H. Poincaré, *Les Méthodes Nouvelles de la Mécanique Céleste*, Vol. I, Gauthier-Villars, Paris, 1892.

[28] H. Poincaré, *Les Méthodes Nouvelles de la Mécanique Céleste*, Vol. II, Gauthier-Villars, Paris, 1893.

[29] N. N. Bogoliubov and Yu. A. Mitropolsky, *Asymptotic Methods in the Theory of Nonlinear Oscillations*, Gordon and Breach, New York, 1961.
[30] N. N. Krylov and N. N. Bogoliubov, *Introduction to Nonlinear Mechanics*, Izd. AN UkSSR, Kiev, 1937 (in Russian).
[31] J. Sanders, F. Verhulst, and J. Murdock, *Averaging Methods in Nonlinear Dynamical Systems*, Springer, New York, 2007.
[32] C. Beck, *Ergodic properties of a kicked damped particle*, Comm. Math. Phys., 130 (1990), pp. 51–60.
[33] A. Skorokhod, F. Hoppensteadt, and H. Salehi, *Random Perturbation Methods*, Springer-Verlag, New York, 2002.
[34] M. I. Freidlin and A. D. Wentzell, *Diffusion processes on graphs and the averaging principle*, Ann. Probab., 21 (1993), pp. 2215–2245.
[35] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, Springer-Verlag, New York, 1998.
[36] Yu. Kifer, *Large deviations and adiabatic transitions for dynamical systems and Markov processes in fully coupled averaging*, Mem. Amer. Math. Soc., 201 (2009), 944.
[37] R. Z. Khas'minskii, *A limit theorem for the solutions of differential equations with random right-hand sides*, Theory Probab. Appl., 11 (1966), pp. 390–406.
[38] M. Loève, *Probability Theory* 1, 4th ed., Springer-Verlag, New York, Heidelberg, 1977.
[39] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc., 58 (1963), pp. 13–30.
[40] B. Victor, *Aspects of Combinatorics*, Cambridge University Press, Cambridge, UK, 1993.
[41] X. Mao and C. Yuan, *Stochastic Differential Equations with Markovian Switching*, Imperial College Press, London, 2006.
[42] D. Nesic and A. R. Teel, *Input-to-state stability for nonlinear time-varying systems via averaging*, Math. Control Signals Systems, 14 (2001), pp. 257–280.
[43] D. Ayels and J. Peuteman, *Exponential stability of nonlinear time-varying systems*, Automatica, 35 (1999), pp. 1091–1100.
[44] M. Benaïm, S. Le Borgne, F. Malrieu, and P.-A. Zitt, *On the stability of planar randomly switched systems*, Ann. Appl. Probab., to appear.
[45] Y. Bakhtin and T. Hurth, *Invariant densities for dynamical systems with random switching*, Nonlinearity, 25 (2012), pp. 2937–2952.
[46] J. Marklof, Y. Tourigny, and L. Wolowski, *Explicit invariant measures for products of random matrices*, Trans. Amer. Math. Soc., 360 (2008), pp. 3391–3427.