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 non-invariance 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 non-invariant 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 non-invariant 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.

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

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

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

In many engineering and biological networks, the individual nodes composing 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, and the switching time is fast, with respect to the characteristic

---

*This work was supported by the National Science Foundation under Grant DMS-1009744 (to I.B.); RFFI grant No. 09-01-00498 and Ministry of Education and Science of the Russian Federation under Grant 11.G34.31.0009 (to V.B.); Swiss National Science Foundation (to M.H.).

1 School of Computer and Communication Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Station 14, 1015 Lausanne, Switzerland.

2Mathematics Department, Volga State Academy, 5, Nesterov st., Nizhny Novgorod, 603 600 Russia.

3Department of Mathematics and Statistics, Georgia State University, 30 Pryor Street, Atlanta, GA 30303, USA. Corresponding author: ibelykh@gsu.edu
HASLER, BELYKH, AND BELYKH

time of the individual node dynamics. In [11], we proved that global synchronization occurs almost surely in blinking networks, 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 in any dynamical regime. 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 regular 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, the characteristic time of individual dynamical system and 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 stochastic variables.

The fact that the rapidly switched blinking system has the same behavior as the averaged system seems apparent; however, there are exceptions and therefore a careful proof of the property is needed which shows what parameters the occurrence of exceptions depends. In fact, the assumption that a trajectory of the blinking system can follow that of the averaged systems can only be true for finite time, unless there is a mechanism which forces them to stick together. Such a mechanism is present when 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, 32, 33], averaging for blinking systems needs special mathematical techniques for obtaining rigorous convergence proofs. We used such techniques for synchronization of blinking dynamical networks [11] and for convergence of the blinking network to an attractor [12]. In this paper, we will 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 blinking system follows closely the solution of the averaged system for a certain time and afterwards they usually drift apart. We derived explicit bounds that relate the probability, switching frequency, precision, and 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 how the solutions of the blinking and average 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 or not it is shared by the blinking system for all switching sequences. We introduce them through numerical analysis of four corresponding examples. After that, using the technique developed in [21] combined with Lyapunov function method, we prove four general theorems that describe the behavior of 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 trajectories of the blinking system converge to the attractor of the averaged
system with high probability if switching is fast. In the non-invariant 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 *ghost* attractor for the blinking system. In the non-invariant multiple attractor case, the trajectory may escape to another attractor with small probability.

The layout of this paper is as follows. First, in Sec. 2, we briefly describe the blinking model and corresponding averaged system. Then, in Secs. 3-6, we present and numerically study four examples for four distinct classes of blinking systems. These are (i) a blinking network of coupled Lorenz systems where connections are stochastically switched on and off (Sec. 3); (ii) two bistable systems coupled by a blinking connection (Sec. 4); (iii) stochastically switched power converter (Sec. 5); (iv) an information processing cellular neural network with blinking shortcuts (Sec. 6).

Then, in Sec. 7, we make basic assumptions on the dynamics of the blinking systems and its averaged analog. Sections 8-11 present four main theorems and their proofs. Finally, a brief discussion of the obtained results is given.

2. The blinking model. We only briefly introduce the system in the 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

\[
\frac{dx}{dt} = F(x(t), s(t)), \quad x \in \mathbb{R}^N, \quad F : \mathbb{R}^{N+M} \to \mathbb{R}^N, \quad s(t) \in \{0,1\}^M,
\]

where the function \( s(t) \) is piecewise constant, taking the constant binary vector value \( s^k = (s^k_1, \ldots, s^k_M) \) 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 \( s^k_i \) of \( s^k \) switches on \((s^k_i = 1)\) or off \((s^k_i = 0)\) during the \( k \)-th time interval. The switching sequences are assumed to be instances of the stochastic process \( S^k, \quad k = 1, 2, \ldots \), 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) has inherently two time scales, the switching period \( \tau \) and timescale of dynamics of the non-switched system where the vector \( s \) is kept constant. We limit our attention to the case where switching is fast with respect to the timescale of non-switched system dynamics. In this case, one can expect that 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

\[
\frac{dx}{dt} = \Phi(x(t)),
\]

where

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

and \( E(F(x, S)) \) 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 mass distribution.
Blinking systems can be found in various applications. The four examples given below illustrate a general 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 non-invariance 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 Fig. 3.1. Instead of connecting the individual systems permanently, they are stochastically switched on and off as described in Sec. 2. The $i$-th Lorenz system is described by the three ordinary differential equations

$$\begin{align*}
\dot{x}_i &= \sigma (y_i - x_i) = G_1(x_i) \\
\dot{y}_i &= r x_i - y_i - x_i z_i = G_2(x_i) \\
\dot{z}_i &= -b z_i + x_i y_i = G_3(x_i),
\end{align*}$$

where $x_i = (x_i, y_i, z_i)$. 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 blinking mode according to Fig. 3.1, we obtain the system of equations

$$\begin{align*}
\dot{x}_1 &= G(x_1) + 0.8 d s_1(t) P(x_2 - x_1) + 0.8 d s_2(t) P(x_3 - x_1) \\
\dot{x}_2 &= G(x_2) + 0.8 d s_3(t) P(x_1 - x_2) + 0.2 d s_3(t) P(x_3 - x_2) + 0.2 d s_3(t) P(x_4 - x_2) \\
\dot{x}_3 &= G(x_3) + 0.8 d s_4(t) P(x_1 - x_3) + 0.2 d s_4(t) P(x_2 - x_3) + 0.2 d s_4(t) P(x_5 - x_3) \\
\dot{x}_4 &= G(x_4) + d s_5(t) P(x_2 - x_4) + d s_6(t) P(x_5 - x_4) \\
\dot{x}_5 &= G(x_5) + d s_5(t) P(x_3 - x_5) + d s_6(t) P(x_4 - x_5),
\end{align*}$$

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

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (3.3)$$

This blinking system has the form (2.1) with $N = 15$ and $M = 6$. During each time interval of length $\tau$, each of the six edges (cf. Fig. 3.1), defined by the switching variables $s_1(t), \ldots, s_6(t)$, is turned on with probability $p$, independently of the switching.
on and off of 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 (2.1) is obtained
by replacing all switching variables, \( s_1(t), \ldots, s_6(t) \) by their mean value \( p \). In Fig. 3.1
this amounts to replacing \( d \) by \( pd \).

The question we are interested in is whether the blinking system synchronizes
when the average 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 solutions of the averaged system converge to the diagonal
subspace \( D \), is the same true for solutions of the blinking system?

Applying the connection graph stability method \([5]\), we can conclude that the
average 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 \( \tau \) for synchronization in a blinking network of Lorenz
systems was given in \([11]\).

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

\[
V(x_1, \cdots, x_5) = \frac{1}{30} \sum_{i=1}^{5} \sum_{j=i+1}^{5} ((x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2)
\]  

(3.4)

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

We now present the results of some numerical simulations. Figure 3.2 indicates,
at least for the given switching sequence of the blinking system, that both the average
and blinking systems synchronize, but not towards the same solution (cf. the upper
panel of Fig. 3.2). In other words, they converge to the same attracting set, the diag-
onal manifold \( D \), but they are not close to each other. In the lower panel of Fig. 3.2
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
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 solutions
of a single Lorenz system, and both network solutions of the averaged and blinking
system converge to the Lorenz attractor in the diagonal subspace.

Figure 3.3 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 slower than the averaged system. This repeats the findings of Fig. 3.2 for
\( \tau = 0.1 \), but on a longer time-scale. The deviation from synchronization of the solu-
tion for the blinking system, after having reached about \( 10^{-15} \), irregularly oscillates
between this value and about \( 10^{-10} \), the precision of numerical integration of differ-
ential 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 there probably is somewhere a threshold for synchronization
between \( \tau = 1 \) and \( \tau = 5 \).
Fig. 3.2. (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 as functions of time. The blinking system and averaged system start from the same non-homogeneous initial conditions $x_1 \neq \ldots \neq x_5$. The $x_i$-components of the other Lorenz systems in the averaged system would be indistinguishable due to fast synchronization and are not shown. Synchronization also takes place in the blinking system, but its synchronized solution is different from that of the average system. (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 blinking system is slower than the averaged system. Parameters are the switching probability $p = 1/2$, switching time $\tau = 0.1$ and coupling parameter $d = 200$.

Fig. 3.3. Synchronization error as a function of time for the same solutions as in Fig. 3.2, for the averaged system (black dashed curve) and blinking system with switching time $\tau = 0.1$ (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 Fig. 3.2.

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 will 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

$$
\dot{x} = f(x) + 1.6s(t)(y-x)
$$

$$
\dot{y} = 2f(y) + 1.6s(t)(x-y)
$$

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 $p = 0.5$

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

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

(4.3)

It is a gradient system. Its potential function is

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

(4.4)

We have chosen the free constants in such a way that at two minima, $(1, 1)$ and $(-1, -1)$, the potential takes the value 0. As in any gradient system, along 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 convergence of the trajectories for the blinking system to attractors of the averaged system, for fast switching.

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

The blinking system has 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 whether it converges to the same equilibrium as the averaged system when starting from the same initial state. In Fig. 4.1 an initial state is chosen close to the boundary between the attraction basins of two stable equilibrium points for the averaged system. Starting from this state, a solution of the blinking system and the solution of 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 (Fig. 4.2). Practically, this happens only, when the initial state is close to the attraction basin boundary, as is the case in Fig. 4.2. The faster the switching, and farther away from the basin boundary of equilibrium points, the smaller is this probability.

![Figure 4.1](image-url)  

**Fig. 4.1.** Trajectories of the averaged system (red smooth curve) and 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 by a black dashed line. It is also the separatrix between the attraction basin boundaries of the two stable equilibrium points. The switching period \(\tau = 0.01\).

On the other hand, the convergence to the equilibrium point is exponentially fast not only for solutions of the averaged system, but also for solutions of the blinking system. This is illustrated in Fig. 4.3 by representing the potential function \(V\) as a function of time in semi-logarithmic scale. Of course, when starting close to the attraction basin boundary, initially convergence is slow, but then 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 dynamics is the invariance of 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 blinking system cannot converge to the attractor of the averaged system and can only reach 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.
Fig. 4.2. Same representation as in Fig. 4.1, 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 Fig. 4.1, is close to the attraction basin boundary.

Fig. 4.3. 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 semi-logarithmic plot. The initial state is the same as in Figs. 4.1 and 4.2. 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.

5. Third example: Switching power converter. Consider the circuit of Fig. 5.1. 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 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 parasitic components is distributed over a whole frequency range and therefore it is
less disturbing than narrow band parasitics in the case of periodic switching. This is discussed in [22, 23], where 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 a short circuit for \( s = 0 \).

The circuit equations are then

\[
\frac{di}{dt} = \frac{E}{L} - \frac{R_0}{L} i - (1 - s(t)) \frac{v}{L} \tag{5.1}
\]

\[
\frac{dv}{dt} = (1 - s(t)) \frac{i}{C} - \frac{v}{RC}. \tag{5.2}
\]

The averaged system is asymptotically stable. The averaged DC-output voltage is at equilibrium point is

\[
\eta = \frac{(1 - p) R}{(1 - p)^2 R + R_0} E,
\]

and 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 averaged systems more or less together is convergence to the equilibrium point in the averaged system. Note that this equilibrium is not shared by the blinking system. In fact, both when \( s = 0 \) (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 it approaches it and then fluctuates around it (Fig. 5.2). The fluctuations diminish in size as the switching period \( \tau \) 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. Forth 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 (CNN) [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 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 “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

![Fig. 5.1. Switched boost power converter as an example of a blinking system.](image-url)
DYNAMICS OF BLINKING SYSTEMS. Part II

Fig. 5.2. (Upper panel): Trajectories in state space of the averaged (red smooth curve) and of an instance of the blinking system (blue 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 blinking (blue) system. Parameters $E = 1$, $R_0 = 0.05$, $R = 4$, $L = C = 1$, $p = 0.5$, and $\tau = 0.005$.

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].

\[
\begin{align*}
\dot{x}_i &= -x_i + (1 - \delta) y_i - \alpha \sum_{j=1}^{N} y_j + \kappa \\
y_i &= f(x_i) = \begin{cases} 
1 & \text{for } x_i > 1 \\
x_i & \text{for } -1 \leq x_i \leq 1 \\
-1 & \text{for } x_i < -1.
\end{cases}
\end{align*}
\] (6.1)

Note that $\alpha$ is the coupling coefficient between any two pair of individual systems $i$ and $j$ of the network. Of course, instead of using all-to-all coupling, computing the sum of all output signals $y_i$ as an intermediate step would be more efficient procedure. However, for our purpose, 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 $x_i(0)$, $i = 1, \ldots, N$. Suppose that the largest among these numbers is $x_m(0)$. Then, the state vector $\mathbf{x}$ 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 $m$, and the index of state with the largest initial value.

The state space is divided into $N$ basins of attraction for 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 cell $j$ such that 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 \(i\)-th cell is +1 and all other outputs are \(-1\). However, when all cells are in saturation, the \(j\)-the cell and \(i\)-th cell do not interact. Then there will be another stable equilibrium where, in addition to the \(i\)-th cell, the \(j\)-the 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 CNN as an integrated circuit. A way out of this dilemma is to use switched instead of hardwired non-local 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 read out the results. Thus, we consider the blinking system [12]

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

\[
y_i = f (x_i),
\]

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

As a numerical example, we consider a blinking \(4 \times 4\) CNN \((N = 16)\) that has the task to determine the largest among the 16 numbers.

\[
\begin{array}{cccc}
0.3644 & 0.3958 & 0.1871 & 0.2898 \\
-0.3945 & -0.2433 & -0.0069 & 0.6359 \\
0.0833 & 0.7200 & 0.7995 & 0.3205 \\
-0.6983 & 0.7073 & 0.6433 & -0.3161
\end{array}
\]

(6.3)

In Fig. 6.1, \(x_{3,3}(t)\) and \(x_{1,4}(t)\) are represented, for the blinking system, averaged system and 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 (6.3) is the largest, whereas the state \(x_{1,4}\) 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 non-local 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 can get close to it, as in the switched power converter (cf. Fig. 6.2).

Furthermore, there is a non-zero probability that it will approach the wrong equilibrium point, as illustrated in Fig. 6.2. This depends on how close the second largest initial state and small \(\tau\) is. Reducing \(\tau\), the solution of the blinking system follows initially more closely the solution of the averaged system and thereby has a lower probability to approach the wrong equilibrium point. It also remains closer to
Fig. 6.1. Trajectories for two instances of $4 \times 4$ “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 averaged system (smooth red curve). The trajectory of averaged system always approaches the correct equilibrium point, whereas the trajectories of 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 maximal value. (Lower panel): Trajectories corresponding to the cell $(3, 3)$ whose initial condition has maximal value.

Fig. 6.2. 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 as in Fig. 6.1.

There is even a non-zero 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 practice, once the trajectory of the blinking system is sufficiently close to a stable equilibrium point of
Fig. 6.3. Same as Fig. 6.1, 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 Fig. 6.1.

Fig. 6.4. Mean square deviation from the correct equilibrium point as a function of time, for the trajectory of Fig. 6.3.

the averaged system, a decision is taken and the system stops.

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 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 each theorem is opposite to the order of examples as it is more convenient to start from the most general case (Theorem 8.8, the forth example) and 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 7.1.**
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 $0 \leq t < \infty$. Similarly, there exists a unique trajectory $\xi(t)$ of the averaged system

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

(7.1)
defined for $0 \leq t < \infty$, 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 all trajectories of the blinking system and averaged system starting in $R$ remain in $R$.

Thus, all interesting dynamics take place in $R$ and in particular, all attractors lie in $R$. In the sequel, we shall restrict our attention to trajectories in $R$. The continuity of $F$ implies that $F$ and $\Phi$ 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 cellular neural network (the forth 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 does a solution of the blinking system (2.1), that starts from the same initial conditions, converge to the attractor. Instead of 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 $A$,
- 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 8.1.** The averaged system (2.2) has an attracting set $A$ with a corresponding Lyapunov function $W$. More precisely, we suppose that there is a twice continuously differentiable function $W : \mathbb{R}^N \to \mathbb{R}$ such that

a). There is a constant $V_1$ and a connected component $C_1$ of the level set

$$\{ x \mid W(x) \leq V_1 \} \text{, contained in } R \text{, such that for any } x \in C_1

W(x) \geq 0 \text{ and } x \in A \iff W(x) = 0 \quad (8.1)$$

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

$$\frac{\partial W}{\partial x} (x) \Phi(x) \leq 0 \quad (8.2)$$

and

$$\frac{\partial W}{\partial x} (x) \Phi(x) = 0 \iff x \in A \quad (8.3)$$

c). The sets

$$C_1 \cap \{ x \in \mathbb{R}^N \mid W(x) \leq V \} \quad (8.4)$$

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

Remark 8.2. 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 chosen.

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

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

Thus, within region $C_1$ the function $W$ decreases strictly along any solution of the averaged system except in 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, what happens to the solution of the blinking system. Note that if the level set $\{ x | W(x) \leq V_1 \}$ 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 $\Delta t$ 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.

**First step:**

For this purpose, it is convenient to introduce in addition to functions $F$ and $\Phi$ that give the time derivative of the states of the blinking and the averaged system, the four functions $D_F W : \mathbb{R}^{N+M} \to \mathbb{R}$, $D_2 W : \mathbb{R}^{N+2M} \to \mathbb{R}$, and $D_\Phi W$, $D_\Phi^2 W : \mathbb{R}^N \to \mathbb{R}$ that give the first and a kind of second time derivative of Lyapunov function $W$ along solutions of the blinking and averaged system.

**Definition 8.2.**

Define the functions $D_F W$, $D_2^2 W$, $D_\Phi W$, $D_\Phi^2 W$ by

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

$$
D_2^2 W (x, \tilde{s}, s) = \sum_{i,j=1}^{N} \left[ \frac{\partial^2 W}{\partial x_i \partial x_j} (x) F_i (x, \tilde{s}) F_j (x, s) + \frac{\partial W}{\partial x_i} (x) \frac{\partial F_j}{\partial x_j} (x, \tilde{s}) F_j (x, s) \right]
$$

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

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

and introduce their bounds on $R$

$$
B_{W, \Phi} = \max_{x \in R} |D_\Phi W (x)|
$$

$$
LB_{W, \Phi} = \max_{x \in R} |D_\Phi^2 W (x)|
$$

$$
B_{W, F} = \max_{s \in (0,1)^M} \max_{x \in R} |D_F W (x, s)|
$$

$$
LB_{W, F} = \max_{s, \tilde{s} \in (0,1)^M} \max_{x \in R} |D_2^2 W (x, \tilde{s}, s)|.
$$

(8.7)
It follows that if $x(t)$ is a solution of the blinking system and $\xi(t)$ a solution of the averaged system, then, in addition to (8.5)

$$
\frac{d^2}{dt^2} W(\xi(t)) = D^2_\Phi W(\xi(t)) \\
\frac{d}{dt} W(x(t)) = DFW(x(t), s(t)) \\
\frac{d}{dt} DFW(x(t), \tilde{s}) = D^2_\Phi W(x(t), s(t)).
$$

Furthermore, the expectation

$$
E(DFW(x, S)) = D\Phi W(x) \tag{8.9}
$$

and thus by the weak law of large numbers [34], for all $x(t)$ and $\lambda > 0$

$$
P \left\{ \left| \frac{1}{K} \sum_{k=1}^{K} DFW(x, S^k) - D\Phi W(x) \right| > \lambda \right\} \rightarrow 0 \quad \text{as} \quad K \rightarrow \infty. \tag{8.10}
$$

Hence, we can define

$$
P_{W\lambda}(K) = \max_{x \in \mathbb{R}} P \left\{ \left| \frac{1}{K} \sum_{k=1}^{K} DFW(x, S^k) - D\Phi W(x) \right| > \lambda \right\} \tag{8.11}
$$

which has the property

$$
P_{W\lambda}(K) \rightarrow 0 \quad \text{for any} \quad \lambda > 0 \quad \text{as} \quad K \rightarrow \infty \tag{8.12}
$$

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

$$
P_{W\lambda}(K) = \max_{x \in \mathbb{R}} P \left\{ \left| \frac{1}{K} \sum_{k=k_0+1}^{K+k_0} DFW(x, S^k) - D\Phi W(x) \right| > \lambda \right\}. \tag{8.13}
$$

Applying the Hoeffding inequality [35, 21], we get

**Lemma 8.3.** For any $\lambda > 0$, the following inequality holds

$$
P_{W\lambda}(K) \leq 2e^{\frac{-\kappa^2}{2D\Phi W}}. \tag{8.14}
$$

Notice that

$$
\int_{t}^{t+\Delta t} [DFW(x, s(u)) - D\Phi W(x)] du = \sum_{k=k_0+1}^{k_0+K} \tau [DFW(x, s^k) - D\Phi W(x)] \tag{8.15}
$$

if $t = k_0\tau$ and $\Delta t = K\tau$ and inequality (8.14) can be applied to bound the LHS of (8.15). In order to be able to apply this bound to arbitrary positive real $t$, $\Delta t$, we again extend the definition of $P_{W\lambda}(K)$ to non-integer $K$ by

$$
P_{W\lambda}(K) = \max_{0 \leq \alpha, \beta \leq 1} \max_{x \in \mathbb{R}} P \left( \frac{1}{K} |\alpha DFW(x, S^1) + \sum_{k=2}^{K'} DFW(x, S^k) + \beta DFW(x, S^{K'+1}) - KD\Phi W(x)| > \lambda \right). \tag{8.16}
$$
With this definition, on one hand, (8.14) is valid for any real \( K > 0 \) and on the other hand, for any \( t \geq 0, \Delta t \geq 0, \lambda > 0, x \in \mathbb{R} \) we have

\[
P \left( \left| \int_{t}^{t+\Delta t} \left[ D_F W(x, s(u)) - D_F W(x) \right] du \right| > \lambda \cdot \Delta t \right) \leq P_{W, \lambda} \left( \frac{\Delta t}{\tau} \right). \tag{8.17}
\]

The following lemma and its proof are similar to Lemma 4.1 in the companion paper [21].

**Lemma 8.4.** Consider a solution \( x(t) \) of the blinking system. Choose a time \( t \geq 0 \) and the solution of the averaged system with \( \xi(t) = x(t) \). Then

a). for any \( \Delta t \geq 0 \)

\[
|W(x(t+\Delta t)) - W(\xi(t+\Delta t))| \leq \alpha \Delta t, \tag{8.18}
\]

where \( \alpha = B_{WF} + B_{W\Phi} \);

b). for any \( \lambda > 0 \) and \( \Delta t \geq 0 \) the conditional probability that

\[
|W(x(t+\Delta t)) - W(\xi(t+\Delta t))| \leq \frac{LB_{WF} + LB_{W\Phi}}{2} \Delta t^2 + \lambda \Delta t \tag{8.19}
\]

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:

\[
|W(x(t+\Delta t)) - W(\xi(t+\Delta t))| \leq \left| \int_{t}^{t+\Delta t} \frac{d}{du} W(x(u)) \, du - \int_{t}^{t+\Delta t} \frac{d}{du} W(\xi(u)) \, du \right| = \left| \int_{t}^{t+\Delta t} D_F W(x(u), s(u)) \, du - \int_{t}^{t+\Delta t} D_F W(\xi(u)) \, du \right| \leq (B_{WF} + B_{W\Phi}) \Delta t. \tag{8.20}
\]

b). Here we prove the second part of the Lemma:

\[
|W(x(t+\Delta t)) - W(\xi(t+\Delta t))| \leq \left| \int_{t}^{t+\Delta t} D_F W(x(u), s(u)) \, du - \int_{t}^{t+\Delta t} D_F W(\xi(u)) \, du \right| \leq \left| \int_{t}^{t+\Delta t} [D_F W(x(u), s(u)) - D_F W(x(t), s(u))] \, du \right| + \left| \int_{t}^{t+\Delta t} [D_F W(x(t), s(u)) - D_F W(x(t))] \, du \right| + \left| \int_{t}^{t+\Delta t} [D_F W(\xi(u)) - D_F W(\xi(t))] \, du \right|. \tag{8.21}
\]
Using (8.8), the first and third term can be rewritten as

$$
\left| \int_{t}^{t+\Delta t} [D_F W(x(u), s(u)) - D_F W(x(t), s(u))] \, du \right| \leq \int_{t}^{t+\Delta t} \left[ D_F W(x(v), s(u)) \right] \, du = \int_{t}^{t+\Delta t} \left[ D_F W(x(v), s(v), s(u)) \right] \, du \leq LB_{WF} (\Delta t)^2; \\
\int_{t}^{t+\Delta t} [D_{\Phi} W(x(u)) - D_{\Phi} W(x(t))] \, du \leq \int_{t}^{t+\Delta t} \left[ D_{\Phi} W(x(v)) \right] \, du = \int_{t}^{t+\Delta t} \left[ D_{\Phi} W(x(v)) \right] \, du \leq LB_{WF} (\Delta t)^2.
$$

(8.22)

For the second term we apply (8.17) to obtain that with probability at least

$$
1 - P_{WF} (\Delta t)
$$

$$
\int_{t}^{t+\Delta t} \left[ D_{\Phi} W(x(t), s(u)) - D_{\Phi} W(x(t)) \right] \, du \leq \lambda \Delta t. \tag{8.23}
$$

Combining (8.22) and (8.23) proves the lemma.

\[\square\]

**Lemma 8.5.** For any $V_0$ with $0 < V_0 < V_1$, let

$$
-\gamma = \max_{x \in C_1, \, V_0 \leq W(x) \leq V_1} D_{\phi} W(x), \\
\Delta t = \frac{\gamma}{2(LB_{WF} + LB_{WF})}, \\
\lambda = \frac{\gamma}{2}, \\
\tilde{V}_0 = V_0 + \gamma \Delta t.
$$

(8.24)

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

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

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

given $x(t)$, is at least $1 - P_{WF} (\Delta t)$.

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

$$
W(x(t + \Delta t)) \leq W(x(t)) + \frac{\gamma}{2} \Delta t
$$

given $x(t)$, is at least $1 - P_{WF} (\Delta t)$.

**Proof.**

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

$$
W(\xi(t + \Delta t)) = W(\xi(t)) + \int_{t}^{t+\Delta t} D_{\phi} W(\xi(u)) \, du. \tag{8.27}
$$
By (8.24), \( D_\delta W (\xi (u)) \leq -\gamma \) as long as \( W (\xi (u)) \geq V_0 \). 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 \( V_0 + \gamma \Delta t \leq W (x (t)) \leq V_1 \) in all cases
\[
W (\xi (t + \Delta t)) \leq W (\xi (t)) - \gamma \Delta t. \tag{8.28}
\]

Hence by Lemma (8.4), with parameters (8.24), we obtain under the condition \( \tilde{V}_0 \leq W (x (t)) \leq V_1 \) with probability at least \( 1 - P_{W \lambda} (\frac{\Delta t}{\gamma}) \)
\[
W (x (t + \Delta t)) \leq W (\xi (t)) + |W (x (t + \Delta t)) - W (\xi (t + \Delta t))| \leq W (\xi (t)) - \gamma \Delta t + \frac{\gamma}{2} \Delta t = W (x (t)) - \frac{\gamma}{2} \Delta t. \tag{8.29}
\]
and thus instead of (8.29) we only have
\[
W (\xi (t + \Delta t)) \leq W (\xi (t)) \tag{8.30}
\]
with probability at least \( 1 - P_{W \lambda} (\frac{\Delta t}{\gamma}) \).

**Second step:**

For any choice of the initial state \( x (0) \) with \( 0 < W (x (0)) \leq V_1 \) and \( x (0) \in C_1 \) (the connected component of the level set \( \{ x \mid W (x) \leq V_1 \} \) containing attracting set \( A \) ), for any choice of constants \( \tau > 0 \) and \( V_0 \) with \( 0 < V_0 < V_1 \), and for constants \( \gamma, \lambda, \Delta t, \tilde{V}_0 \) given by (8.24), we consider the following sequence of random variables on the probability space of switching sequences:
\[
Z_q (s) = W (x (q \Delta t)), \quad q = 0, 1, 2, ... \tag{8.32}
\]

By Hypothesis 8.1, \( Z_0 \) is concentrated on a single value that is smaller 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 function \( W \) destroys much of the information contained in state \( x (q \Delta t) \). However, according to Lemma 8.5 we have the deterministic bound
\[
W (x (t + \Delta t)) \leq W (x (t)) + \alpha \Delta t \tag{8.33}
\]
and under the condition \( W (x (t)) \leq V_1 \), the probabilistic bound
\[
P \left( W (x (t + \Delta t)) \leq W (x (t)) - \frac{\gamma}{2} \Delta t \mid x (t), \tilde{V}_0 \leq W (x (t)) \leq V_1 \right) > 1 - P_{W \lambda} \left( \frac{\Delta t}{\gamma} \right) \tag{8.34}
\]
and finally the probabilistic bound
\[
P \left( W (x (t + \Delta t)) \leq W (x (t)) + \frac{\gamma}{2} \Delta t \mid x (t), 0 \leq W (x (t)) \leq V_1 \right) > 1 - P_{W \lambda} \left( \frac{\Delta t}{\gamma} \right). \tag{8.35}
\]

As for the finite time analysis [21], it is convenient to introduce auxiliary binary random variables:
\[
\theta_q = \begin{cases} 
1 & \text{if } W (x (q \Delta t)) \leq W (x ((q - 1) \Delta t)) - \frac{\gamma}{2} \Delta t \text{ and } \tilde{V}_0 \leq W (x ((q - 1) \Delta t)), \\
1 & \text{if } W (x (q \Delta t)) \leq W (x ((q - 1) \Delta t)) + \frac{\gamma}{2} \Delta t \text{ and } 0 \leq W (x ((q - 1) \Delta t)) < \tilde{V}_0, \\
0 & \text{otherwise}.
\end{cases} \tag{8.36}
\]
The following results correspond to Lemma 4.2 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 \( W(x(t)) \leq V_1 \), at least for instants \( t = q\Delta t \), \( q = 0,1,\cdots \) for bound (8.28) or (8.31) to be applicable.

**Lemma 8.6.** Suppose the various constants are chosen as in Lemma 8.5 and suppose that

\[ V_1 - V_0 \geq \frac{3}{2} \gamma \Delta t. \]  

(8.37)

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

a). for \( m = 0 \)

\[ P(\theta_q = 1, Z_{q-1} \leq V_1 \text{ for } q = 1,\cdots,Q) = P(\theta_q = 1 \text{ for } q = 1,\cdots,Q) (1 - P_{W^\lambda} (\Delta t/\tau))^Q. \]  

(8.38)

b). for \( m > 0 \)

\[ P(\theta_q = \sigma_q, Z_{q-1} \leq V_1 \text{ for } q = 1,\cdots,Q) \leq \left(P_{W^\lambda} (\Delta t/\tau)\right)^m. \]  

(8.39)

**Proof.**

a). It follows from \( Z_0 \leq V_1 \) and \( \theta_q = 1 \) for \( q = 1,\cdots,Q \) that \( Z_q \leq V_1 \) for \( q = 1,\cdots,Q \). Indeed, from \( \theta_q = 1 \) it follows that \( Z_q < Z_{q-1} \) unless \( Z_{q-1} \leq \tilde{V}_0 \). In this last case, using (8.37), \( Z_q \leq Z_{q-1} + \frac{2}{3} \Delta t \leq \tilde{V}_0 + \frac{2}{3} \Delta t \leq V_1 \). Therefore,

\[ P(\theta_q = 1, Z_{q-1} \leq V_1 \text{ for } q = 1,\cdots,Q) = P(\theta_q = 1 \text{ for } q = 1,\cdots,Q). \]  

(8.40)

Furthermore,

\[ P(\theta_q = 1 \text{ for } q = 1,\cdots,Q) = \sum_{x((Q-1)\Delta t)} P(\theta_q = 1 \text{ for } q = 1,\cdots,Q, x((Q-1)\Delta t)), \]  

(8.41)

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

\[ P(\theta_q = 1 \text{ for } q = 1,\cdots,Q) = \sum_{x((Q-1)\Delta t)} P(\theta_Q = 1 | \theta_q = 1 \text{ for } q = 1,\cdots,Q-1, x((Q-1)\Delta t)) \times P(\theta_q = 1 \text{ for } q = 1,\cdots,Q-1, x((Q-1)\Delta t)). \]  

(8.42)

Since \( \theta_Q \) depends only on \( x(Q\Delta t) \) and \( x((Q-1)\Delta t) \), and \( x(q\Delta t), q = 1,2,\cdots \) is a vector valued Markov process, we obtain

\[ P(\theta_q = 1 \text{ for } q = 1,\cdots,Q, x((Q-1)\Delta t)) = \sum_{x((Q-1)\Delta t)} P(\theta_Q = 1 | x((Q-1)\Delta t)) \times \]  

\[ \times P(\theta_q = 1 \text{ for } q = 1,\cdots,Q-1, x((Q-1)\Delta t)) > (1 - P_{W^\lambda} (\Delta t/\tau)) \sum_{x((Q-1)\Delta t)} P(\theta_q = 1 \text{ for } q = 1,\cdots,Q-1, x((Q-1)\Delta t)) = \]  

\[ = (1 - P_{W^\lambda} (\Delta t/\tau)) \cdot P(\theta_q = 1 \text{ for } q = 1,\cdots,Q-1). \]  

(8.43)
Then, inequality (8.38) follows by induction. Indeed, for \( Q > 0 \), inequality (8.43) is the induction step and for \( Q = 1 \), it becomes

\[
P (\theta_1 = 1) > \left( 1 - P_{W\lambda} \left( \frac{\Delta t}{\tau} \right) \right)
\]  

(8.44)

which holds by (8.43).

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

\[
P (\theta_q = \sigma_q, Z_{q-1} \leq V_1 \text{ for } q = 1, \cdots, Q) \leq P (\theta_q = \sigma_q, Z_{q-1} \leq V_1 \text{ for } q = 1, \cdots, Q - 1).
\]

(8.45)

If \( \sigma_Q = 1 \) we write

\[
P (\theta_q = \sigma_q, Z_{q-1} \leq V_1 \text{ for } q = 1, \cdots, Q) = \sum_{x=((Q-1)\Delta t)} x \cdot W (x ((Q-1)\Delta t))
\]

(8.46)

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

The condition \( Z_{Q-1} \leq V_1 \) is a restriction on this summation. Therefore,

\[
P (\theta_q = \sigma_q, Z_{q-1} \leq V_1 \text{ for } q = 1, \cdots, Q) = \sum_{x=((Q-1)\Delta t)} x \cdot W (x ((Q-1)\Delta t)) \leq \sum_{x=((Q-1)\Delta t)} x \cdot P (\theta_q = 0 | \theta_q = \sigma_q, Z_{q-1} \leq V_1, q = 1, \cdots, Q - 1).
\]

(8.47)

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

\[
P (\theta_q = \sigma_q, Z_{q-1} \leq V_1 \text{ for } q = 1, \cdots, Q) = \sum_{x=((Q-1)\Delta t)} x \cdot W (x ((Q-1)\Delta t)) \times P (\theta_q = 0 | x ((Q-1)\Delta t)) \leq \sum_{x=((Q-1)\Delta t)} P (\theta_q = \sigma_q, Z_{q-1} \leq V_1, q = 1, \cdots, Q - 1) \cdot P (\theta_q = 0 | x ((Q-1)\Delta t)).
\]

(8.48)

Using for each term in the sum the complement of (8.34) or (8.35), depending on the value of \( W (x ((Q-1)\Delta t)) \), we obtain

\[
P (\theta_q = \sigma_q, Z_{q-1} \leq V_1 \text{ for } q = 1, \cdots, Q) \leq \sum_{x=((Q-1)\Delta t)} P_{W\lambda} (\frac{\Delta t}{\tau}) \cdot P (\theta_q = \sigma_q, Z_{q-1} \leq V_1, q = 1, \cdots, Q - 1)
\]

(8.49)

Again by induction, applying (8.34) or (8.35), depending on the value of the corresponding \( \sigma_q \), proves the Lemma.
Lemma 8.7. For any $V \geq 0$ and any integer $Q$ with $Q \geq (V_1 - V_2)/\Delta t$ we have

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

where

$$m = \left\lfloor \frac{Q\gamma}{2\alpha + \gamma} - \frac{V_1 - V}{\Delta t} \cdot \frac{2}{2\alpha + \gamma} \right\rfloor + 1 \quad (8.51)$$

and $[x]$ is the integer part of $x$.

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

$$Z_Q \leq V_1 - (Q - m)\frac{\gamma}{2} \Delta t + ma \Delta t.$$  \hspace{1cm} (8.52)

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

$$m > \frac{Q\gamma}{2\alpha + \gamma} - \frac{V_1 - V}{\Delta t} \cdot \frac{2}{2\alpha + \gamma}.$$  \hspace{1cm} (8.53)

The smallest $m$ satisfying (8.53) is (8.51). Therefore,

$$P \left( \tilde{V}_0 \leq Z_{q-1} \leq V_1 \text{ for } q = 1, \cdots, Q, \ Z_Q > V \right) \leq \sum_{\sigma_1, \cdots, \sigma_Q = 0}^{1} \sum_{\sigma_1, \cdots, \sigma_Q = 0}^{Q - \sum_{q=1}^{Q} \sigma_q \geq m} \sum_{\sigma_1, \cdots, \sigma_Q = 0}^{Q - \sum_{q=1}^{Q} \sigma_q \geq m} \sum_{\sigma_1, \cdots, \sigma_Q = 0}^{Q} \binom{Q}{n} \left[ P_{W\lambda} \left( \frac{\Delta t}{\tau} \right) \right]^n.$$  \hspace{1cm} (8.54)

Theorem 8.8. [Case 1: possible multiple attractors/non-invariance].

Suppose Hypothesis 7.1 is satisfied and the averaged system (2.2) has an attracting set $A$ with a corresponding Lyapunov function $W$, satisfying Hypothesis 8.1. Choose $V_0$ such that $0 < V_0 < V_1$ and let

$$-\gamma = \max_{x \in C_1, \ V_0 \leq W(x) \leq V_1} D_{\Phi}W(x)$$

$$\Delta t = \frac{2(2LB_{WF} + LB_{W\Phi})}{B_{WF} + B_{W\Phi}}$$

$$\alpha = \frac{B_{WF} + B_{W\Phi}}{2(2LB_{WF} + LB_{W\Phi})}$$

$$c = \frac{1}{64(2LB_{WF} + LB_{W\Phi} + LB_{WF})}.$$  \hspace{1cm} (8.55)
where the various expressions are defined in Definition 8.2. Suppose that
\[ V_1 - V_0 \geq \frac{3\gamma^2}{4(LB_{WF} + LB_{W\Phi})} \] (8.56)
and \( \tau \) is sufficiently small, such that
\[ 2e^{\frac{2a + \gamma}{\gamma}} e^{-\frac{c^3}{\tau}} \leq 1 - \sqrt{\frac{e}{3}} \] (8.57)

Consider the two open regions
\[ U_0 = \{ x \mid W(x) < V_0 + \gamma \Delta t \} \cap C_1, \quad U_\infty = \{ x \mid W(x) > V_1 + \alpha \Delta t \}. \] (8.58)

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

a). The probability \( P_{\text{direct escape}} \) that the solution \( x(t) \) of the blinking system reaches \( U_\infty \) before reaching \( U_0 \) is bounded by
\[ P_{\text{direct escape}} \leq \frac{36a^2}{\gamma^2} e^{-\frac{c^3}{\tau}}. \] (8.59)

Inversely, the probability that the solution \( x(t) \) of the blinking system reaches \( U_0 \) before reaching \( U_\infty \) is at least
\[ P_{\text{direct attraction}} \geq 1 - \frac{72a^2}{\gamma^2} e^{-\frac{c^3}{\tau}}. \] (8.60)

b). Let \( T_{\text{attraction}} \) be the time for the solution of the blinking system to enter \( U_0 \) through its boundary and \( T_{\text{remain}} \) be the time it remains in
\[ \bar{U}_0 = \{ x \mid W(x) \leq V_0 + \left( \frac{3}{2} \gamma + a \right) \Delta t \} \] (8.61)
after reaching \( U_0 \). These times are random variables with the following properties
\[ P \left( T_{\text{attraction}} \leq 2 \frac{V_1 - V_0}{\gamma} \right) > 1 - 8 \frac{(V_1 - V_0)(LB_{WF} + LB_{W\Phi})}{\gamma^2} e^{-\frac{c^3}{\tau}} \] (8.62)
and
\[ P \left( T_{\text{remain}} > T \right) > 1 - 4T \frac{(LB_{WF} + LB_{W\Phi})}{\gamma} e^{-\frac{c^3}{\tau}}. \] (8.63)

Proof.

a). We shall first prove the first part of Theorem 8.8. Let \( \lambda = \frac{3}{2} \). Consider the set \( S_{\text{direct escape}} \) of switching sequences such that the solution of the blinking system reaches \( U_\infty \) before reaching \( U_0 \). For each, such a switching sequence there must be an integer \( Q \) such that
\[ \tilde{V}_0 \leq Z_{q-1} \leq V_1 \quad \text{for} \ q = 1, \ldots, Q \quad \text{and} \quad Z_Q > V_1, \] (8.64)
Fig. 8.1. Illustration of Theorem 8.8. 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 $P_{direct \text{ escape}}$. 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 $\tau$. Initially, the trajectory of the blinking system may also escape from the attraction basin $C_1$ right away and move toward another attractor in $U_\infty$ (not shown) with probability $P_{direct \text{ escape}}$. This probability approaches 0 when $\tau \to 0$.

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

$$\tilde{V}_0 \leq Z_q \leq V_1 \text{ for } q = 0, 1, 2 \cdots.$$  \hfill (8.65)

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

$$W(x(t)) \leq W(x(q\Delta t)) + |W(x(t)) - W(x(q\Delta t))| \leq V_1 + \alpha(t - q\Delta t) < V_1 + \alpha\Delta t,$$  \hfill (8.66)

where we have used (8.18). But (8.66) is in contradiction with $x(t) \in U_\infty$. Hence, there is an integer $Q$ such that (8.64) holds and therefore

$$S_{escape}^{direct} \subseteq S_+ = \bigcup_{Q=1}^{\infty} \left\{ s \mid \tilde{V}_0 \leq Z_{q-1} \leq V_1 \text{ for } q = 1, \cdots, Q \text{ and } Z_Q > V_1 \right\}$$  \hfill (8.67)

and thus

$$P_{direct \text{ escape}} \leq P(S_+) \leq \sum_{Q=1}^{\infty} P\left(\tilde{V}_0 \leq Z_q \leq V_1 \text{ for } q = 0, \cdots, Q - 1 \text{ and } Z_Q > V_1\right).$$  \hfill (8.68)

Applying Lemma 8.7 for $V = V_1$, we get

$$P(S_+) \leq \sum_{Q=1}^{\infty} \sum_{n=0}^{Q} \left( \begin{array}{c} Q \\ n \end{array} \right) \left[ P_{W\lambda} \left( \frac{\Delta t}{\tau} \right) \right]^n,$$  \hfill (8.69)
where \( m \) is given by (8.51). Hence, the double sum goes over all integer pairs \((Q,n)\) such that
\[
Q \geq 1 \quad \text{and} \quad \left\lfloor \frac{Q\gamma}{2\alpha + \gamma} \right\rfloor + 1 \leq n \leq Q \iff n \geq 1 \quad \text{and} \quad n \leq Q + \frac{2\alpha + \gamma}{\gamma}.
\] (8.70)

Therefore, using the upper bound on binomial coefficients \([36]\)
\[
\binom{Q}{n} \leq \left( \frac{Qe}{n} \right)^n,
\] (8.71)
we get
\[
P(S_+) \leq \sum_{n=1}^{\infty} [P_{W\lambda} \left( \frac{\Delta t}{\gamma} \right)]^n n^2 \sum_{Q=1}^{\infty} \left( \frac{Qe}{n} \right)^n \leq \sum_{n=1}^{\infty} [P_{W\lambda} \left( \frac{\Delta t}{\gamma} \right)]^n \sum_{Q=1}^{\infty} \left( \frac{Qe}{n} \right)^n \leq \frac{2\alpha}{\gamma} \sum_{n=1}^{\infty} [P_{W\lambda} \left( \frac{\Delta t}{\gamma} \right)]^n n^2 \left( \frac{2\alpha + \gamma}{\gamma} \right)^n \leq \frac{2\alpha e}{\gamma} \left( \frac{eP_{W\lambda} \left( \frac{\Delta t}{\gamma} \right)}{1 - \frac{2\alpha + \gamma}{\gamma} eP_{W\lambda} \left( \frac{\Delta t}{\gamma} \right)} \right).
\] (8.72)

Application of (8.57) and (8.14) guarantee that the sum in (8.72) converges and lead to (8.59).

In order to prove (8.60), we reason as follows. The solution \( x(t) \) of the blinking system either reaches \( U_\infty \) before reaching \( U_0 \) or reaches \( U_0 \) before reaching \( U_\infty \) or never reaches \( U_\infty \) or \( U_0 \). The corresponding set of switching sequences is \( S_{\text{direct escape}} \) in the first case. For the other two cases let us denote it by \( S_{\text{direct attraction}} \) and \( S_{\text{trapped}} \), respectively. Thus, the set \( S \) of all switching sequences is decomposed into
\[
S = S_{\text{direct escape}} \cup S_{\text{direct attraction}} \cup S_{\text{trapped}}.
\] (8.73)

Another decomposition is
\[
S = S_+ \cup S_- \cup S_0,
\] (8.74)
where \( S_+ \) is given by (8.67) and
\[
S_- = \bigcup_{Q=1}^{\infty} \left\{ s \mid \tilde{V}_0 \leq Z_{q-1} \leq V_1 \quad \text{for} \ q = 1, \cdots , Q \quad \text{and} \ Z_Q < \tilde{V}_0 \right\},
\] (8.75)
\[
S_0 = \left\{ s \mid \tilde{V}_0 \leq Z_{q-1} \leq V_1 \quad \text{for} \ q = 1, 2, \cdots \right\}.
\]

Since the solution corresponding to a switching sequence in \( S_{\text{trapped}} \) 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 \( S_{\text{trapped}} \subseteq S_0 \). However, clearly \( S_{\text{trapped}} \cap S_- = \emptyset \) and therefore
\[
S_{\text{trapped}} \subseteq S_0 \cup S_+.
\] (8.76)

Now, for any \( Q \geq 1 \)
\[
S_0 \subseteq \left\{ s \mid \tilde{V}_0 \leq Z_{q-1} \leq V_1 \quad \text{for} \ q = 1, \cdots , Q \quad \text{and} \ Z_Q \geq \tilde{V}_0 \right\} \cup \left\{ s \mid \tilde{V}_0 \leq Z_{q-1} \leq V_1 \quad \text{for} \ q = 1, \cdots , Q \quad \text{and} \ Z_Q \geq 0 \right\}
\] (8.77)
and therefore, applying Lemma 8.7, for any \( Q \geq 1 \)

\[
P(S_0) \leq \sum_{n=m}^{Q} \left( \frac{n}{P(W_\lambda(\frac{\Delta}{\tau}))} \right)^n, \quad m = \left\lfloor \frac{Q^2}{2\alpha + \gamma} - \frac{2V_1}{(2\alpha + \gamma) \Delta t} \right\rfloor + 1. \tag{8.78}
\]

If we choose

\[
Q \geq \frac{2V_1}{\gamma \Delta t \sqrt{\mathcal{P}}} \tag{8.79}
\]

then

\[
m > \frac{\gamma}{2\alpha + \gamma} \left( Q - \frac{2V_1}{\gamma \Delta t} \right) \geq \frac{\gamma Q}{2\alpha + \gamma} (1 - \sqrt{\mathcal{P}}) = Q \beta, \quad \text{where} \quad \beta = \frac{\gamma}{2\alpha + \gamma} (1 - \sqrt{\mathcal{P}}). \tag{8.80}
\]

Therefore, using again (8.57)

\[
P(S_0) \leq \sum_{n=m}^{Q} \left( \frac{Q^n}{P(W_\lambda(\frac{\Delta}{\tau}))} \right)^n \leq \sum_{n=m}^{Q} \left( \frac{Q^n}{P(W_\lambda(\frac{\Delta}{\tau}))} \right)^n \leq \sum_{n=m}^{Q} \left( \frac{n}{P(W_\lambda(\frac{\Delta}{\tau}))} \right)^n = \left[ \frac{1}{\mathcal{P}} P(W_\lambda(\frac{\Delta}{\tau})) \right]^m \cdot \frac{1}{1 - \mathcal{P} P(W_\lambda(\frac{\Delta}{\tau}))}. \tag{8.81}
\]

The last equality holds, because, using (8.57) gives

\[
\frac{e}{\beta} P(W_\lambda(\frac{\Delta}{\tau})) = \frac{2\alpha + \gamma}{\gamma} \cdot \frac{e}{1 - \sqrt{\mathcal{P}}} \leq 1 - \frac{\gamma}{\gamma + 1} < 1. \tag{8.82}
\]

Since (8.81) holds for all \( Q \) satisfying (8.79) and when \( Q \to \infty \) also \( m \to \infty \) and the RHS of (8.81) converges to 0. Hence, \( P(S_0) = 0 \) and because of (8.76) and (8.72) we obtain

\[
P(S_{\text{trapped}}) \leq \frac{36 \alpha^2}{\gamma} \cdot e^{-c_3^3} \tag{8.83}
\]

and

\[
P_{\text{direct}} = 1 - P_{\text{trapped}} - P_{\text{direct escape}} \geq 1 - \frac{72 \alpha^2}{\gamma} \cdot e^{-c_3^3}. \tag{8.84}
\]

This completes the proof of the first statement in Theorem 8.8.

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

\[
\theta_q = 1 \quad \text{for} \quad q = 1, \cdots, Q \quad \text{with} \quad Q = \left\lfloor \frac{2(V_1 - V_0)}{\gamma \Delta t} \right\rfloor - 1. \tag{8.85}
\]

Then,

\[
Z_q \leq Z_{q-1} - \frac{\gamma}{2} \Delta t \quad \text{as long as} \quad Z_{q-1} \geq \tilde{V}_0. \tag{8.86}
\]

Now, if \( Z_{Q-1} \geq \tilde{V}_0 \), then

\[
Z_Q \leq V_1 - Q \frac{\gamma}{2} \Delta t \leq V_1 - \left( \frac{2(V_1 - V_0)}{\gamma \Delta t} - 2 \right) \frac{\gamma}{2} \Delta t = V_0 + \gamma \Delta t = \tilde{V}_0. \tag{8.87}
\]


Hence, in any case,
\[ T_{\text{attraction}} \leq Q \Delta t \leq \frac{2(V_1 - V_0)}{\gamma}. \] (8.88)

Thus, the set of switching sequences where (8.85) is satisfied contains the set of switching sequences where (8.88) is satisfied. Hence,
\[ P\left(T_{\text{attraction}} \leq \frac{2(V_1 - V_0)}{\gamma}\right) \geq P\left(\theta_q = 1 \text{ for } q = 1, \ldots, Q\right) \geq (1 - P_{W,\gamma}(\Delta t))^Q \geq (1 - Q \cdot P_{W,\gamma}(\Delta t)) \geq 1 - \frac{4(V_1 - V_0)e^{-\frac{\gamma^2}{2}}}{\gamma^2 \Delta t}, \] (8.89)

where we have used Lemma 8.6 and (8.14).

To get probabilistic bounds on \( T_{\text{remain}} \), it is convenient to reset the time to 0 after reaching \( U_0 \). Thus, we suppose that \( x(0) \in U_0 \). Then, if
\[ \theta_q = 1 \text{ for } q = 1, \ldots, Q, \] (8.90)

it follows that
\[ Z_q < \tilde{V}_0 + \frac{\gamma}{2} \Delta t \text{ for } q = 0, \ldots, Q \] (8.91)

because
\[ Z_{q-1} < \tilde{V}_0 \Rightarrow Z_q \leq Z_{q-1} + \frac{\gamma}{2} \Delta t \leq \tilde{V}_0 + \frac{\gamma}{2} \Delta t, \]
\[ \tilde{V}_0 \leq Z_{q-1} \leq \tilde{V}_0 + \frac{\gamma}{2} \Delta t \Rightarrow Z_q \leq Z_{q-1} - \frac{\gamma}{2} \Delta t \leq \tilde{V}_0 + \frac{\gamma}{2} \Delta t. \] (8.92)

Using (8.18), we get
\[ W(x(t)) < V_0 + \left(\frac{3}{2} \gamma + \alpha\right) \Delta t \text{ for } t \in [0, Q \Delta t] \] (8.93)

and thus \( T_{\text{remain}} \geq Q \Delta t \). It follows that
\[ P(T_{\text{remain}} > T) > P(\theta_q = 1 \text{ for } q = 1, \ldots, \left\lfloor \frac{T}{\Delta t} \right\rfloor) > (1 - P_{W,\gamma}(\Delta t))^\left\lfloor \frac{T}{\Delta t} \right\rfloor > 1 - \frac{4(V_1 - V_0)e^{-\frac{\gamma^2}{2}}}{\gamma^2 \Delta t}, \] (8.94)

which is inequality (8.63) after the substitution (8.55).

\[ \square \]

**Remark 8.3.**

a). This general theorem is applicable in all four cases. 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 \( \tau \), the probability to escape from 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 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 \( \{x \mid W(x) \leq V_1\} \) is connected and thus identical to \( C_1 \) and Theorem 8.8 only gives information 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 to reach 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 with high probability, reaches a small neighborhood of \( A \) without leaving the basin. In fact, this result is obtained by combining Theorem 4.6 from the companion paper [21] and Theorem 8.8. According to Theorem 4.6 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 4.6 from [21] can be applied repeatedly, until \( C_1 \) is reached. Then, Theorem 8.8 can be invoked.

e). Note that the upper bound \( P_{\text{direct escape}} \) and lower bound \( P_{\text{direct attraction}} \) do not imply that the sum of the two probabilities sum up to 1, as one would expect. Therefore, the trajectories of the blinking system that never reach \( U_0 \) or \( U_\infty \) might have positive probability. However, this just a technical consequence of the way to derive the bounds. Furthermore, this probability could not be larger than \( P_{\text{direct escape}} \) 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 (8.56) 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 \( \gamma \), in turn needs a smaller switching time \( \tau \).

9. Case 2: unique attracting set; not necessarily invariant for the blinking system. We now strengthen Hypothesis 8.1 to adapt it to the case where \( A \) is the unique attractor of the averaged system or where all attractors of the system are contained in attracting set \( A \). The switching power converter (the third example) is a case in point.

**Hypothesis 9.1.** The level set \( \{ x \mid W(x) \leq V_1 \} \) is connected and thus identical to \( C_1 \) introduced in Hypothesis 8.1. It contains the compact attracting region \( R \) introduced in Hypothesis 1, i.e., we have

\[
x \in R \implies W(x) \leq V_1.
\]

(9.1)

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

\[
W(x(t)) \leq V_1 \text{ for } t \geq T.
\]

(9.2)

Now we can formulate the stronger theorem.

**Theorem 9.1.** [Case 2: unique attractor/non-invariance]. Under Hypothesis 7.1, Hypothesis 8.1, and Hypothesis 9.1 consider any solution \( x(t), t \in [0, \infty) \) of the blinking system. Let \( W \) be the Lyapunov function introduced in Hypothesis 8.1 and \( V_1 \) be the positive constant introduced in Hypothesis 9.1. As in Theorem 8.8, choose \( V_0 \)
such that $0 < V_0 < V_1$ and let

\[
\begin{align*}
-\gamma &= \max_{V_0 \leq W(x) \leq V_1} D_\Phi W(x), \\
c &= \frac{8(\text{LB}_{WF} + \text{LB}_{W\Phi})}{\text{D}_{WF}}, \\
U_0 &= \left\{ x \mid W(x) < V_0 + \frac{c^2}{2(\text{LB}_{WF} + \text{LB}_{W\Phi})} \right\}.
\end{align*}
\]  

(9.3)

Then the following properties hold:

a). If the switching time $\tau$ satisfies

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

(9.4)

where

\[
D = 8(\text{LB}_{WF} + \text{LB}_{W\Phi})
\]

(9.5)

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

b). Assume that (9.4) 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

\[
P_1(n) \leq e^{-n\left[ \frac{c^3}{\gamma^2} - \ln\left( \frac{D(V_1 - V_0)}{\gamma^2} \right) \right]},
\]

(9.6)

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

\[
W(x(t)) > V_0 + \frac{8\gamma^2}{D}
\]

(9.7)

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

Proof. Define again

\[
\Delta t = \frac{\gamma}{2(\text{LB}_{WF} + \text{LB}_{W\Phi})}.
\]

(9.8)

a). By definition of the attracting region $R$ in Hypothesis 7.1, there is a time $T \geq 0$ such that

\[
x(T) \in R \text{ for } t \geq T
\]

(9.9)

and thus

\[
W(x(t)) \leq V_1 \text{ for } t \geq T.
\]

(9.10)

As in the proof of Theorem 8.8b), define the integer $Q$ to be

\[
Q = \left\lfloor \frac{2V_1 - V_0}{\gamma \Delta t} \right\rfloor - 1.
\]

(9.11)

By (8.89), with probability at least $1 - \frac{4(V_1 - V_0)c^3}{\gamma^2}$, 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 \(4(V_1 - V_0) e^{-\frac{c^2}{\gamma \Delta t}}\). Nevertheless, since the region \(R\) is invariant, these trajectories satisfy

\[
W(x(T + Q\Delta t)) \leq V_1. \tag{9.12}
\]

They again with probability at least \(1 - 4(V_1 - V_0) e^{-\frac{c^2}{\gamma \Delta t}}\) 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 \(x(t)\),

\[
P(x(t) \notin U_0, t \in [0, T + nQ\Delta t]) = \\
\sum_{x((n-1)Q\Delta t)} P(x(t) \notin U_0, t \in [0, T + nQ\Delta t], x((n-1)Q\Delta t))
\]

\[
= \\
\sum_{x((n-1)Q\Delta t)} P(x(t) \notin U_0, t \in [T + (n-1)Q\Delta t, T + nQ\Delta t] | x(t) \notin U_0, x((n-1)Q\Delta t))
\]

\[
= \\
\sum_{x((n-1)Q\Delta t)} P(x(t) \notin U_0, t \in [T + (n-1)Q\Delta t, T + nQ\Delta t] | x(t) \notin U_0, x((n-1)Q\Delta t)) \cdot P(x(t) \notin U_0, t \in [0, T + (n-1)Q\Delta t], x((n-1)Q\Delta t))
\]

\[
\leq \\
\sum_{x((n-1)Q\Delta t)} \frac{4(V_1 - V_0)}{\gamma \Delta t} e^{-\frac{c^2}{\gamma \Delta t}} P(x(t) \notin U_0, t \in [0, T + (n-1)Q\Delta t])
\]

\[
\leq \frac{4(V_1 - V_0)}{\gamma \Delta t} e^{-\frac{c^2}{\gamma \Delta t}} \cdot P(x(t) \notin U_0, t \in [0, T + (n-1)Q\Delta t]) \tag{9.13}
\]

By repeated application of (9.13) we get

\[
P(x(t) \notin U_0 \text{ for } 0 \leq t \leq T + nQ\Delta t) \leq \left(\frac{4(V_1 - V_0)}{\gamma \Delta t} e^{-\frac{c^2}{\gamma \Delta t}}\right)^n. \tag{9.14}
\]

Since the expression in the parenthesis is smaller than 1 by (9.4), we obtain

\[
P(x(t) \notin U_0 \text{ for } 0 \leq t < \infty) = \\
\lim_{n \to \infty} P(x(t) \notin U_0 \text{ for } 0 \leq t \leq T + nQ\Delta t) = 0 \tag{9.15}
\]

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 satisfies already \(W(x(0)) \leq V_1\), we can set \(T = 0\) in (9.14) which implies (9.6). This completes the proof of assertion b).

c). For \(q \geq 1\) and \(\tilde{V}_0 = V_0 + \gamma \Delta t\)

\[
P(Z_q \leq V_1 - \frac{m \gamma}{2} \Delta t) \geq P\left(Z_q \leq V_1 - \frac{m \gamma}{2} \Delta t, \tilde{V}_0 \leq Z_{q-1} \leq V_1\right) + \\
P\left(Z_q \leq V_1 - \frac{m \gamma}{2} \Delta t, 0 \leq Z_{q-1} < \tilde{V}_0\right). \tag{9.16}
\]

Since

\[
\left\{ s | \theta_q = 1 \text{ and } \tilde{V}_0 \leq Z_{q-1} \leq V_1 - \frac{(m-1) \gamma}{2} \Delta t \right\} \subseteq \left\{ s | Z_q \leq V_1 - \frac{m \gamma}{2} \Delta t \right\} \tag{9.17}
\]
for $1 \leq m \leq Q$ (for $m > Q$ the result is useless)

$$
P \left( Z_q \leq V_1 - \frac{m\gamma}{2} \Delta t, \ V_0 \leq Z_{q-1} \leq V_1 \right) \geq \\
\geq P \left( \theta_q = 1, \ Z_{q-1} \leq V_1 - \frac{(m-1)\gamma}{2} \Delta t, \ V_0 \leq Z_{q-1} \leq V_1 \right) \\
= \sum_{x((q-1)\Delta t)} P \left( \theta_q = 1, \ V_0 \leq Z_{q-1} \leq V_1 - \frac{(m-1)\gamma}{2} \Delta t, \ x ((q-1) \Delta t) \right) \\
= \sum_{v_0 \leq W(x((q-1)\Delta t)) \leq V_1 - \frac{(m-1)\gamma}{2} \Delta t} P \left( \theta_q = 1 | x ((q-1) \Delta t) \right) P \left( x ((q-1) \Delta t) \right) \\
(9.18)
$$

Applying (8.25), we get

$$
P \left( Z_q \leq V_1 - \frac{m\gamma}{2} \Delta t, \ V_0 \leq Z_{q-1} \leq V_1 \right) \geq \\
\geq \sum_{v_0 \leq W(x((q-1)\Delta t)) \leq V_1 - \frac{(m-1)\gamma}{2} \Delta t} \left( 1 - P_{W \lambda} \left( \frac{\Delta t}{\tau} \right) \right) \cdot P \left( x ((q-1) \Delta t) \right) \\
= \left( 1 - P_{W \lambda} \left( \frac{\Delta t}{\tau} \right) \right) \cdot P \left( V_0 \leq Z_{q-1} \leq V_1 - \frac{(m-1)\gamma}{2} \Delta t \right) \\
(9.19)
$$

Instead of (9.17), we can write

$$
\left\{ s | \theta_q = 1 \text{ and } 0 \leq Z_{q-1} \leq V_1 - \frac{(m+1)\gamma}{2} \Delta t \right\} \subseteq \left\{ s | Z_q \leq V_1 - \frac{m\gamma}{2} \Delta t \right\} \\
(9.20)
$$

for any $m \geq 1$, and thus

$$
P \left( Z_q \leq V_1 - \frac{m\gamma}{2} \Delta t, \ 0 \leq Z_{q-1} < V_0 \right) \geq \\
\geq P \left( \theta_q = 1, \ Z_{q-1} \leq V_1 - \frac{(m+1)\gamma}{2} \Delta t, \ 0 \leq Z_{q-1} < V_0 \right) \\
(9.21)
$$

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

$$
V_1 - \frac{(m+1)\gamma}{2} \Delta t \geq V_0 \\
(9.22)
$$

and therefore

$$
P \left( Z_q \leq V_1 - \frac{m\gamma}{2} \Delta t, \ 0 \leq Z_{q-1} < V_0 \right) \geq P \left( \theta_q = 1, \ 0 \leq Z_{q-1} < V_0 \right) \\
= \sum_{x((q-1)\Delta t)} P \left( \theta_q = 1, \ 0 \leq Z_{q-1} < V_0, \ x ((q-1) \Delta t) \right) \\
= \sum_{0 \leq W(x((q-1)\Delta t)) \leq V_0} P \left( \theta_q = 1 | x ((q-1) \Delta t) \right) \cdot P \left( x ((q-1) \Delta t) \right) \\
\geq \sum_{0 \leq W(x((q-1)\Delta t)) \leq V_0} \left( 1 - P_{W \lambda} \left( \frac{\Delta t}{\tau} \right) \right) \cdot P \left( x ((q-1) \Delta t) \right) \\
= \left( 1 - P_{W \lambda} \left( \frac{\Delta t}{\tau} \right) \right) \cdot P \left( 0 \leq Z_{q-1} < V_0 \right) \\
(9.23)
$$

where we have used (8.26).

Combining (9.19) and (9.23), the following inequality holds for $q \geq 1$ and $1 \leq m \leq Q - 2$:

$$
P \left( Z_q \leq V_1 - \frac{m\gamma}{2} \Delta t \right) \geq \left( 1 - P_{W \lambda} \left( \frac{\Delta t}{\tau} \right) \right) \cdot P \left( Z_{q-1} \leq V_1 - \frac{(m-1)\gamma}{2} \Delta t \right) \\
(9.24)
$$
Since we suppose $x(0) \in R$, $R$ is invariant and since by Hypothesis 9.1 $W(x(0)) \leq V_1$, we have

$$P(Z_q \leq V_1) = 1 \text{ for } q \geq 0.$$  \hfill (9.25)

Applying (9.24) iteratively, we obtain

$$P\left(Z_q \leq V_1 - \frac{m\gamma}{2} \Delta t\right) \leq \left(1 - P_{W\lambda}\left(\frac{\Delta t}{\tau}\right)\right)^m \text{ for } q \geq m \text{ and } 1 \leq m \leq Q - 2.$$ \hfill (9.26)

We now rewrite this result. By the definition (9.11) of $Q$, for any $n$

$$V_0 + \frac{n\gamma}{2} \Delta t > V_1 - \frac{(Q-n+2)\gamma}{2} \Delta t$$ \hfill (9.27)

and therefore

$$P\left(Z_q \leq V_0 + \frac{n\gamma}{2} \Delta t\right) \leq P\left(Z_q \leq V_1 - \frac{(Q-n+2)\gamma}{2} \Delta t\right) \geq (1 - P_{W\lambda}\left(\frac{\Delta t}{\tau}\right))^{Q-n+2} \text{ for } q \geq Q - n + 2 \text{ and } 4 \leq n \leq Q + 1.$$ \hfill (9.28)

For $n = 4$:

$$P(Z_q > V_0 + 2\gamma \Delta t) = 1 - P(Z_q \leq V_0 + 2\gamma \Delta t) \leq 1 - (1 - P_{W\lambda}\left(\frac{\Delta t}{\tau}\right))^{Q-2} \leq (Q - 2) P_{W\lambda}\left(\frac{\Delta t}{\tau}\right) \text{ for } q \geq Q - 2.$$ \hfill (9.29)

This can be rewritten, using (8.14), as

$$P\left(W(x(q\Delta t)) > V_0 + \frac{\gamma^2}{LB_{W*} + LB_{W*}}\right) \leq \frac{8(LB_{W*} + LB_{W*})V_1}{\gamma^2} e^{-\frac{\gamma^2}{2}} \text{ for } q \geq Q - 2.$$ \hfill (9.30)

To get a probabilistic bound on the deviation from the attractor not only for times that are multiples of $\Delta t$, but for all positive $t$, we note that at any time, in particular the time interval $(0, \Delta t)$, $W(t) \leq V_1$ and therefore, we can apply (9.30) to any time-shifted solution, i.e. we can set $x(t)$ for $0 \leq t < \Delta t$ as a new initial condition and apply (9.30) to the new solution. From this we obtain assertion c).

\[\square\]

**Remark 9.2.**

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 it takes the necessary time 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) assures 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 two parameters $V_0$ and $\gamma$. Close to a hyperbolic attractor, $V_0$ and $\gamma$ are proportional. Assuming 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 $K V_0 e^{-\frac{\gamma^2}{V_0}}$. 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 solutions for the blinking system to $A$ is possible. In this section, we give the conditions and precise formulation when this happens. The property that set $A$ is invariant under the dynamics of the blinking system implies that Lyapunov function $W$ vanishes along any trajectory of the blinking system within $A$. We strengthen this property somewhat in Hypothesis 10.1 by requiring that not only the derivative of $W$ along any solution of the blinking system that approaches $A$ converges to zero but that the derivative is bounded by $W$ multiplied by a constant. The bistable system (second example) is a case in point.

**Hypothesis 10.1.** Suppose the averaged system (2.2) has an attracting set $A$ with a corresponding Lyapunov function $W$ satisfying Hypothesis 8.1. Introduce the following functions, similar to Definition 8.2, but using instead of the Lyapunov function $W$ its logarithm. It is well-defined as long as $\Phi_i(W(x)) \neq 0$. Introduce

\begin{align*}
D_F \ln W(x, s) &= \sum_{i=1}^{N} \frac{\partial \ln W(x)}{\partial x_i} F_i(x, s), \\
D_F^2 \ln W(x, \bar{s}, s) &= \sum_{i,j=1}^{N} \left[ \frac{\partial^2 \ln W(x)}{\partial x_i \partial x_j} F_i(x, \bar{s}) F_j(x, s) + \frac{\partial \ln W(x)}{\partial x_i} \frac{\partial F_i}{\partial x_j}(x, \bar{s}) F_j(x, s) \right], \\
D_\Phi \ln W(x) &= \sum_{i=1}^{N} \frac{\partial \ln W(x)}{\partial x_i} \Phi_i(x), \\
D_\Phi^2 \ln W(x) &= \sum_{i,j=1}^{N} \left[ \frac{\partial^2 \ln W(x)}{\partial x_i \partial x_j} \Phi_i(x) \Phi_j(x) + \frac{\partial \ln W(x)}{\partial x_i} \frac{\partial \Phi_i}{\partial x_j}(x, \bar{s}) \Phi_j(x) \right].
\end{align*}

(10.1)

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

\begin{align*}
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} |D_\Phi^2 \ln W(x)|, \\
B_{\ln W F} &= \max_{s \in (0,1)} \sup_{x \in R, W(x) \neq 0} |D_F \ln W(x, s)|, \\
LB_{\ln W F} &= \max_{s, \bar{s} \in (0,1)} \sup_{x \in R, W(x) \neq 0} |D_F^2 \ln W(x, \bar{s}, s)|.
\end{align*}

(10.2)

In addition, we assume that there is a positive constant $\gamma$ such that

\begin{align*}
\sup_{x \in C_1, 0 < W(x) \leq V_1, x \notin A} D_\Phi \ln W(x) = -\gamma < 0
\end{align*}

(10.3)

where $C_1$ and $V_1$ are defined in Hypothesis 8.1.
Remark 10.2.
a). Equation (10.2) implies that along any solution within $R$ of the blinking system the following inequalities hold, as long as $W(x(t)) \neq 0$:

$$\frac{d}{dt} \ln W(x(t)) \leq B \ln W \quad \text{and} \quad \frac{d}{dt} D_F \ln W(x(t), \tilde{s}) = LB \ln W.$$  \hfill (10.4)

Equations (10.2) and (10.3) imply that along any solution of the averaged system with $0 < W(x(0)) \leq V_1$

$$-B \ln W \leq \frac{d}{dt} \ln W(x(t)) \leq -\gamma \quad \text{and} \quad \frac{d}{dt} D_F \ln W(x(t), \tilde{s}) \leq LB \ln W.$$ \hfill (10.5)

The first equations of (10.4) and (10.5) can be rewritten as

$$\frac{d}{dt} \left[ W(x(t)) \right] \leq B \ln W \quad \text{and} \quad \frac{d}{dt} \left[ D_F \ln W(x) \right] \leq -B \ln W.$$

Note that the first inequality is only possible, 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 10.1 is satisfied.

b). Given Hypothesis 10.1, all properties for 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 $K$

$$P_{\ln W}(K) = \max_{x \in R} \left\{ \left[ \frac{1}{K} \sum_{k=1}^{K} D_F \ln W(x, S^k) - D_F \ln W(x) \right] > \lambda \right\}.$$  \hfill (10.7)

and for non-integer $K$

$$P_{\ln W}(K) = \max_{x \in R} \left\{ \left[ \frac{1}{K} \sum_{k=1}^{K} D_F \ln W(x, S^k) + (K - \lfloor K \rfloor) D_F \ln W(x, S^{\lfloor K \rfloor + 1}) \right] - D_F \ln W(x) \right\} > \lambda \right\}.$$ \hfill (10.8)

Because of stationarity of the process, by analogy to (8.17), for any $t \geq 0$, $\Delta t \geq 0$, $\lambda > 0$, $x \in R$ we have

$$P \left( \int_{t}^{t+\Delta t} \left[ D_F \ln W(x, s(u)) - D_F \ln W(x) \right] du \right) \geq \lambda \cdot \Delta t \right\} \leq P_{\ln W}(\frac{\lambda \cdot \Delta t}{\gamma}).$$ \hfill (10.9)

Furthermore, Hoeffding’s inequality [35] gives for this case the following statement.

Lemma 10.1. For any $\lambda > 0$

$$P_{\ln W}(K) \leq 2e^{-\frac{K^2}{2 B \ln W}}.$$ \hfill (10.10)

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

**Lemma 10.2.** Consider a solution $x(\cdot)$ of the blinking system with switching period $\tau$. Choose a time $t \geq 0$ and solution of the averaged system with $\xi(t) = x(t) \notin A$. Then

a). for any $\Delta t \geq 0$

$$|\ln W(x(t + \Delta t)) - \ln W(\xi(t + \Delta t))| \leq (B_{\ln W} F + B_{\ln W} \Phi) \Delta t; \quad (10.11)$$

b). for any $l > 0$ and $\Delta t \geq 0$ the conditional probability that

$$|\ln W(x(t + \Delta t)) - \ln W(\xi(t + \Delta t))| \leq \frac{L B_{\ln W} F + L B_{\ln W} \Phi}{2} \Delta t^2 + \lambda \Delta t \quad (10.12)$$

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

**Proof.** The proof is identical to the proof of Lemma 8.4, except that $W$ has to be replaced everywhere by $\ln W$. \[ \square \]

**Lemma 10.3.** Let

$$\Delta t = \frac{\gamma}{2 (L B_{\ln W} F + L B_{\ln W} \Phi)} \quad \text{and} \quad \lambda = \frac{\gamma}{4}, \quad (10.13)$$

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

$$\ln W(x(t + \Delta t)) \leq \ln W(x(t)) - \frac{\gamma}{2} \Delta t \quad (10.14)$$

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

**Proof.** Consider a solution $\xi(\cdot)$ of the averaged system with $\xi(t) = x(t)$. Then, by (10.3), one obtains

$$\ln W(\xi(t + \Delta t)) = \ln W(\xi(t)) + \int_t^{t+\Delta t} D_{\Phi} W(\xi(u)) du \leq \ln W(x(t)) - \gamma \Delta t. \quad (10.15)$$

Using (10.12) and (10.13), we get

$$\ln W(x(t + \Delta t)) \leq \ln W(\xi(t + \Delta t)) + |\ln W(x(t + \Delta t)) - \ln W(\xi(t + \Delta t))|$$

$$\leq \ln W(x(t)) - \gamma \Delta t + \frac{\gamma}{2} \Delta t. \quad (10.16)$$

\[ \square \]

**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 constants $\gamma$, $\lambda$, and $\Delta t$ given by (10.3) and (10.13), we consider the following sequence of random variables on the probability space of switching sequences:

$$Z_q(s) = \ln W(x(q \Delta t)), \quad q = 0, 1, 2, \ldots \quad (10.17)$$

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

$$\theta_q = \begin{cases} 1 & \text{if } \ln W(x(q \Delta t)) \leq \ln W(x((q - 1) \Delta t)) - \frac{\gamma}{2} \Delta t \quad \text{and} \\ 0 & \text{otherwise.} \end{cases} \quad (10.18)$$
**Lemma 10.4.** Suppose the various constants are chosen as in Lemma 10.3. Let \( \sigma = (\sigma_1, \ldots, \sigma_Q) \in \{0, 1\}^Q \) be a binary vector of length \( Q \) and \( m = Q - \sum_{q=1}^{Q} \sigma_q \) the number of zeros in this vector. Then, for \( m > 0 \)

\[
P(\theta_q = \sigma_q, \ Z_{q-1} \leq \ln V_1 \text{ for } q = 1, \cdots, Q) \leq \left[ P_{\ln W\lambda}\left(\frac{\Delta t}{\tau}\right)\right]^m. \tag{10.19}
\]

**Proof.** The proof is the same as the proof of Lemma 8.6b, except \( V_1, W, P_W \) have to be replaced by \( \ln V_1, \ln W, P_{\ln W\lambda} \) and no lower bound constraint has to be observed for \( Z_q \).

**Lemma 10.5.** For any \( V \geq 0 \) and any \( Q \in \mathbb{N} \) we have

\[
P\left( Z_{q-1} \leq \ln V_1 \text{ for } q = 1, \cdots, Q, Z_Q > \ln V \right) \leq \sum_{n=m}^{Q} \binom{Q}{n} \left[ P_{\ln W\lambda}\left(\frac{\Delta t}{\tau}\right)\right]^n, \tag{10.20}
\]

where

\[
m = \left\lfloor \frac{Q\gamma}{2\alpha + \gamma} - \frac{\ln V_1 - \ln V}{\Delta t} \cdot \frac{2}{2\alpha + \gamma} \right\rfloor + 1 \tag{10.21}
\]

and \( \lfloor x \rfloor \) is the integer part of \( x \) and \( \alpha = B_{\ln W\lambda} + B_{\ln W\Phi} \).

**Proof.** The proof is the same as the proof of Lemma 8.7, except \( V_1, V, P_W \) have to be replaced by \( \ln V_1, \ln V, P_{\ln W\lambda} \) and no lower bound constraint has to be observed for \( Z_q \).

**Theorem 10.6.** [Case 3: non-unique attractor/invariance]. Under Hypothesis 7.1, Hypothesis 8.1, Hypothesis 9.1, and Hypothesis 10.1 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

\[
-\gamma = \sup_{x \in C_1, 0 < W(x) \leq V_1, x \notin A} D_S \ln W(x)
\]

\[
\alpha = \frac{B_{\ln W\lambda} + B_{\ln W\Phi}}{4(L B_{\ln W\lambda} + L B_{\ln W\Phi})}
\]

\[
c = \frac{B_{\ln W\lambda} + B_{\ln W\Phi}}{4(L B_{\ln W\lambda} + L B_{\ln W\Phi})},
\]

where constants \( B \) and \( LB \) are defined in (10.2) and \( C_1 \) and \( V_1 \) are defined in Hypothesis 8.1. Suppose that \( \tau \) is sufficiently small such that

\[
4e \frac{2\alpha + \gamma}{\gamma} e^{-\frac{\alpha^2}{2}} \leq 1 - \sqrt{\frac{e}{3}}. \tag{10.23}
\]

Consider the open region

\[
U_\infty = \{ x \mid \ln W(x) > \ln V_1 + \alpha \Delta t \}. \tag{10.24}
\]

Then

a). the probability \( P_{\text{escape}} \) that solution \( x(t) \) of the blinking system reaches \( U_\infty \) is bounded by

\[
P_{\text{escape}} \leq \frac{36 \alpha^2}{\gamma^2} e^{-\frac{\alpha^2}{2}}; \tag{10.25}
\]
b). the probability that solution $x(t)$ of the blinking system converges to $A$ is at least

$$P_{\text{convergence}} \geq 1 - \frac{216 \alpha^2}{\gamma^2} \cdot e^{-\frac{c_3^2}{t}}. \quad (10.26)$$

More precisely, with probability at least $1 - \frac{216 \alpha^2}{\gamma^2} \cdot e^{-\frac{c_3^2}{t}}$, the convergence is exponentially fast according to

$$W(x(t)) \leq Ke^{-\frac{4}{\tau}t}, \quad (10.27)$$

where

$$K = V_1 e^{\frac{2}{LB} \ln W_{\Phi} + \frac{2}{LB} \ln W_{F}}.$$

**Proof.** Let $\lambda = \frac{2}{\tau}$.

a). Consider set $S_{\text{escape}}$ of switching sequences such that the solution of the blinking system reaches $U_{\infty}$. By the same argument as in the proof of Theorem 8.8a), substituting $\ln W$ for $W$ and disregarding the lower bound on $W$, for each such a switching sequence, there must be an integer $Q$ such that

$$Z_{q-1} \leq \ln V_1 \quad \text{for } q = 1, \cdots, Q \quad \text{and} \quad Z_Q > \ln V_1. \quad (10.28)$$

Therefore

$$S_{\text{escape}} \subseteq S_+ = \bigcup_{Q=1}^{\infty} \{s \mid Z_{q-1} \leq \ln V_1 \quad \text{for } q = 1, \cdots, Q \quad \text{and} \quad Z_Q > \ln V_1\} \quad (10.29)$$

and thus

$$P_{\text{escape}} \leq P(S_+) = \sum_{Q=1}^{\infty} P(Z_q \leq \ln V_1 \quad \text{for } q = 0, \cdots, Q-1 \quad \text{and} \quad Z_Q > \ln V_1). \quad (10.30)$$

Applying Lemma 10.5 for $V = V_1$, we get

$$P(S_+) \leq \sum_{Q=1}^{\infty} \sum_{n=m}^{Q} \frac{Q^n}{n} \left[ P_{\ln W_{\lambda}} \left( \frac{\Delta t}{\tau} \right) \right]^n, \quad (10.31)$$

where

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

As in the proof of Theorem 8.8a), it follows from (10.23) (which implies (8.57)) and (10.10) that the sum in (10.31) converges and is bounded by (10.25).

b). We apply Lemma 10.5 for $V = V_1 e^{-\frac{\Delta t}{4} \gamma}$. Then

$$P \left( Z_{q-1} \leq \ln V_1 \quad \text{for } q = 1, \cdots, Q, \quad Z_Q > \ln V_1 - \frac{Q}{4} \Delta t \right) \leq \sum_{n=m}^{Q} \frac{Q^n}{n} \left[ P_{\ln W_{\lambda}} \left( \frac{\Delta t}{\tau} \right) \right]^n. \quad (10.32)$$
where

\[ m = \left\lfloor \frac{Q\gamma}{2(2\alpha + \gamma)} \right\rfloor + 1. \]  

(10.33)

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

\[ S_0 = \bigcup_{Q=1}^{\infty} \left\{ s \mid Z_{q-1} \leq \ln V_1 \quad \text{for} \quad q = 1, \ldots, Q, \quad Z_Q > \ln V_1 - \frac{Q}{4} \gamma \Delta t \right\}. \]

(10.34)

Then

\[ P(S_0) \leq \sum_{Q=1}^{\infty} P \left( Z_{q-1} \leq \ln V_1 \quad \text{for} \quad q = 1, \ldots, Q, \quad Z_Q > \ln V_1 - \frac{Q}{4} \gamma \Delta t \right) \]

\[ \leq \sum_{Q=1}^{\infty} \sum_{n=m}^{Q} \left( \frac{Q}{n} \right) \left[ P_{ln W_\lambda \left( \frac{\Delta t}{n} \right)} \right]^n, \]

(10.35)

where \( m \) is given by (10.33).

The double summation over \( n \) and \( Q \) has the constraints

\[ Q \geq 1 \quad \text{and} \quad \left\lfloor \frac{Q\gamma}{2(2\alpha + \gamma)} \right\rfloor + 1 \leq n \leq Q \quad \Leftrightarrow \quad n \geq 1 \quad \text{and} \quad n \leq Q < \frac{2(2\alpha + \gamma)}{\gamma} n. \]

(10.36)

Following the same path as in the proof of Theorem 8.8a), we obtain the bound

\[ P(S_0) \leq \frac{4\alpha + \gamma}{\gamma} \cdot \frac{2(2\alpha + \gamma)}{\gamma} e \cdot P_{ln W_\lambda \left( \frac{\Delta t}{n} \right)} \]

\[ \cdot \left( 1 - \frac{2(2\alpha + \gamma)}{\gamma} e \cdot P_{ln W_\lambda \left( \frac{\Delta t}{n} \right)} \right)^2. \]

(10.37)

Applying (10.23), (10.10) and \( \gamma \leq \alpha \), we obtain

\[ P(S_0) \leq \frac{180\alpha^2}{\gamma^2} e^{-\frac{\gamma^2}{2\alpha}}. \]

(10.38)

Let \( S_- \) denote the following set of all switching sequences

\[ S_- = \left\{ s \mid Z_Q \leq \ln V_1 - \frac{Q}{4} \gamma \Delta t \quad \text{for all} \quad Q \in \mathbb{Z}^+ \right\}. \]

(10.39)

Then, if \( S \) denotes the set of all switching sequences

\[ S\setminus S_- = \bigcup_{Q=1}^{\infty} \left\{ s \mid Z_Q > \ln V_1 - \frac{Q}{4} \gamma \Delta t \right\}. \]

(10.40)

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 \( q = 0, \ldots, Q-1 \) 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

\[ S\setminus S_- \subseteq S_0 \cup S_+ \]

(10.41)
and thus
\[ P(S\setminus S_-) \leq P(S_0) + P(S_+) \leq \frac{216 \alpha^2}{\gamma^2} e^{-\frac{c_3}{4}}. \] (10.42)

This is equivalent to
\[ P(S_-) > 1 - \frac{216 \alpha^2}{\gamma^2} e^{-\frac{c_3}{4}}. \] (10.43)

Suppose now that a switching sequence \( s \in S_- \). Then, for any \( t > 0 \) we can write \( t = (q + \mu) \Delta t \), with \( 0 \leq \mu < 1 \), and thanks to (10.4) we obtain
\[
\ln W(x(t)) \leq \ln W(x(q \Delta t)) + \alpha \mu \Delta t \leq \ln V_1 + \frac{2}{4} \gamma \Delta t + \alpha \mu \Delta t = \ln V_1 + \frac{2}{4} \gamma \Delta t + \alpha \mu \Delta t = \ln V_1 - \frac{\tau}{4} + \left( \alpha + \frac{\gamma}{4} \right) \Delta t.
\] (10.44)

This, together with (10.43) implies part b) of the theorem.

11. Case 4: unique attractor; invariant for the blinking system. We now suppose that Hypotheses 7.1, 8.1, 9.1, 10.1 are valid, i.e., 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 level set \( \{ x \mid W(x) \leq V_1 \} \) of 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 (first example) is a case in point.

**Theorem 11.1.** [Case 4: unique attractor/invariance]. Under Hypotheses 7.1, 8.1, 9.1, 10.1, consider any solution \( x(t) \), \( t \in [0, \infty) \) of the blinking system. Let \( W \) be the Lyapunov function introduced in Hypothesis 8.1 and \( V_1 \) be the positive constant introduced in Hypothesis 9.1. Let
\[
\gamma = \sup_{x \in C_1, \ 0 < W(x) \leq V_1} D_{\Phi} \ln W(x), \quad \alpha = B_{\ln W^F} + B_{\ln W^\Phi}, \quad c = \frac{1}{64(LB_{\ln W^F} + LB_{\ln W^\Phi}) B_{\ln W^F}},
\] (11.1)

where constants \( B \) and \( LB \) are defined in (10.2). Suppose that \( \tau \) is sufficiently small such that
\[
2e^{\frac{2\alpha + \gamma}{\gamma}} e^{-\frac{c_3}{4}} \leq 1 - \sqrt{3}. \] (11.2)

Then the solution almost surely converges to \( A \) exponentially fast with exponential speed at least \( \frac{\tau}{4} \), i.e., for almost all switching sequences there is a constant \( K \) such that for all \( t \geq 0 \)
\[ W(x(t)) \leq K e^{-\frac{\tau}{4} t}. \] (11.3)

If \( W(x(0)) \leq V_1 \), then with probability at least
\[ P_{\text{const}} \geq 1 - \frac{180 \alpha^2}{\gamma^2} e^{-\frac{c_3}{4}} \] (11.4)
the following constant $K$ is sufficient for (11.3):

$$K = V_1 e^{\frac{(\gamma + \frac{\gamma}{4})}{2(LB_{lnWF} + LB_{lnW\Phi})}}.$$  

(11.5)

Proof. We again define

$$\Delta t = \frac{\gamma}{2(LB_{lnWF} + LB_{lnW\Phi})}.$$  

(11.6)

To prove the almost sure convergence with exponential speed to $A$, we assume that $x(0) \in R$, and thus $x(t) \in R$ and $W(x(t)) \leq V_1$ for all $t \geq 0$. This is no restriction of generality, since if the initial state lies outside of $R$, in a finite time region $R$ is reached. 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 $\frac{3}{4}$

$$S_{\text{exponential}} = \left\{ s \mid \exists K \text{ such that for all } t \geq 0 \ W(x(t)) \leq Ke^{-\frac{3}{4}t} \right\}.$$  

(11.7)

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

$$S_{\text{exponential}} = \left\{ s \mid \exists K \text{ such that for all } Q \geq 1 \ W(x(Q\Delta t)) \leq Ke^{-\frac{3}{4}Q\Delta t} \right\}$$  

(11.8)

because thanks to (10.4) it follows from $W(x(q\Delta t)) \leq Ke^{-\frac{3}{4}Q\Delta t}$ and $0 \leq \mu < 1$ that

$$\ln W((q + \mu)\Delta t) \leq \ln W(q\Delta t) + \alpha \mu \Delta t \leq \ln K - \frac{3}{4}q\Delta t + \alpha \mu \Delta t = \ln K - \frac{3}{4}(q + \mu)\Delta t + \left(\alpha + \frac{\gamma}{4}\right)\Delta t.$$  

(11.9)

and thus for all $t \geq 0$

$$W(t) \leq Ke^{-\frac{3}{4}t},$$  

(11.10)

where

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

(11.11)

Now we restrict to switching sequences that satisfy $W(x(Q\Delta t)) \leq Ke^{-\frac{3}{4}Q\Delta t}$ asymptotically for $Q \to \infty$ with the constant $V_1$:

$$\tilde{S}_{\text{exponential}} = \left\{ s \mid \exists M \text{ such that for } \forall Q \geq M \ W(x(Q\Delta t)) \leq V_1e^{-\frac{3}{4}Q\Delta t} \right\}.$$  

(11.12)

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

$$S \setminus \tilde{S}_{\text{exponential}} = \left\{ s \mid \forall M \geq 1 \ \exists Q \geq M \text{ such that } W(x(Q\Delta t)) > V_1e^{-\frac{3}{4}Q\Delta t} \right\} = \bigcap_{M=1}^{\infty} \bigcup_{Q=M}^{\infty} \left\{ s \mid W(x(Q\Delta t)) > V_1e^{-\frac{3}{4}Q\Delta t} \right\}.$$  

(11.13)
and thus
\[
P\left( S \setminus \tilde{S}_{\text{exponential}} \right) = \lim_{M \to \infty} \frac{1}{M} \sum_{Q=1}^{M} P\left\{ \left( \sum_{n=1}^{Q} e^{-\frac{4}{Q} \gamma t} \right) > \ln V_1 - \frac{Q}{4} \gamma \Delta t \right\}
\]
Since \( W(x(t)) \leq V_1 \) for all \( t \geq 0 \)
\[
P\left( S \setminus \tilde{S}_{\text{exponential}} \right) \leq \lim_{M \to \infty} \frac{1}{M} \sum_{Q=1}^{M} P\left( Z_Q > \ln V_1 - \frac{Q}{4} \gamma \Delta t \right)
\]
\[
= \lim_{M \to \infty} \frac{1}{M} \sum_{Q=1}^{M} P\left( Z_{Q-1} \leq \ln V_1 \text{ for } q = 1, \cdots, Q, Z_Q > \ln V_1 - \frac{Q}{4} \gamma \Delta t \right).
\]
In the proof of Theorem 10.6 it has been shown that
\[
\sum_{Q=1}^{\infty} P\left( Z_{Q-1} \leq \ln V_1 \text{ for } q = 1, \cdots, Q, Z_Q > \ln V_1 - \frac{Q}{4} \gamma \Delta t \right) < \infty
\]
which implies that \( P\left( S \setminus \tilde{S}_{\text{exponential}} \right) = 0 \) and thus
\[
P(S_{\text{exponential}}) \geq P\left( \tilde{S}_{\text{exponential}} \right) = 1 - P\left( S \setminus \tilde{S}_{\text{exponential}} \right) = 1.
\]
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 \( \gamma \). The fact that for all \( t \geq 0 \)
\[
W(x(t)) \leq Ke^{-\frac{4}{Q} \gamma t}
\]
with \( K \) given by (11.5) and probability given by (11.4) has already been shown in the proof of Theorem 10.6. \( \Box \)

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 non-invariance under 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 switching period \( \tau \) is smaller than an explicitly given bound, trajectories of the blinking system for almost all switching sequences converge to \( A \). This bound depends on exponential speed of convergence \( \gamma \) in the averaged system. In fact, the crucial ratio \( \tau/\gamma^3 \) has to be small. Furthermore, trajectories of the blinking system converge to the attractor also exponentially fast, and speed of convergence is slower than \( \gamma \) 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 average blinking connection. If 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 blink-
ing 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 \( \tau / \gamma^3 \), where \( \gamma \) 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 \( \gamma \). The exceptional trajectories
that do not converge to \( A \) in general escape from the attraction basin of the attrac-
tor. 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 time uncoupled, and part of time are coupled with a certain coupling strength \( d \).
The averaged system yields a coupled system with static connections of lower cou-
pling 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 un-
likely 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 (Theo-
rem 9.1). For any choice of a small neighborhood \( U_0 \) of the attractor, trajectories of
the averaged system approach the attractor with a minimum linear speed \( \gamma > 0 \). We
have limited our analysis to linear speed because in any case trajectories of blinking
system cannot converge to the attractor and can only reach 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 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 \( \gamma^3 \). 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 stochasti-
cally 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
approach rapidly 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 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 $\gamma$ 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 $\tau/\gamma^3$ to be essential. An illustrative example for this case is an information processing cellular neural network. 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 dynamics of the network, more specifically by 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 blinking connections are necessary. If 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 basin boundary between two (correct and a wrong ghost) equilibriums. 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 12.1 summarizes the general results for all four cases.

The comparison of explicit thresholds for the switching period and various probabilities with effective properties of the blinking systems in the case of four examples will be published elsewhere.

Our results can easily be extended to blinking systems driven by a Markov vector process instead of sequences of independent random vectors.

REFERENCES

Fig. 12.1. 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 probabilities are given in four Theorems.

<table>
<thead>
<tr>
<th>Averaged system</th>
<th>Blinking system</th>
<th>Several attractors (multistability)</th>
<th>Single global attractor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-invariance: ghost attractor(s)</td>
<td>Case 1</td>
<td>Case 2</td>
<td></td>
</tr>
<tr>
<td>Invariance of the attractors</td>
<td>Case 3</td>
<td>Case 4</td>
<td></td>
</tr>
</tbody>
</table>


