Biol. Cybern. 84, 31–40 (2001)

Biological Cybernetics

Estimating statistics of neuronal dynamics via Markov chains

Gary Froyland[1], **Kazuyuki Aihara**[2,3]

[1] Department of Mathematics, University of Paderborn, Warburgerstrasse 100, 33098 Paderborn, Germany
[2] Department of Mathematical Engineering, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[3] CREST, Japan Science and Technology Corporation (JST), 4-1-8 Hon-cho, Kawaguchi, Saitama 332-0012, Japan

Received: 26 January 2000 / Accepted in revised form: 9 June 2000

Abstract. We present an efficient computational method for estimating the mean and variance of interspike intervals defined by the timing of spikes in typical orbits of one-dimensional neuronal maps. This is equivalent to finding the mean and variance of return times of orbits to particular regions of phase space. Rather than computing estimates directly from time series, the system is modelled as a finite state Markov chain to extract stationary behaviour in the form of invariant measures and average absorption times. Ergodic-theoretic formulae are then applied to produce the estimates without the need to generate orbits directly. The approach may be applied to both deterministic and randomly forced systems.

1 Introduction

It is well known that spike timing and interspike intervals of neurons are highly variable and irregular (Softky and Koch 1993; Shadlen and Newsome 1998). It is an interesting problem to explore whether such variability and irregularity of neuronal firing in the brain (1) are just inevitable noise to be smoothed away to estimate the average firing rate and its temporal modulation (Shadlen and Newsome 1998), and/or (2) reflect essential properties inherent in underlying dynamics of single neurons and neural networks (Shinomoto et al. 1999), and/or (3) include spatio-temporal structures that can play significant roles for neural coding and processing (Fujii et al. 1997). Since sophisticated analyses with tools such as Joint-PSTH (Aertsen et al. 1989), information theory (Kitazawa et al. 1998), or a generalized linear model of firing probability (Kobayashi et al. 1998) are succeeding in extracting valuable information from "noisy" neuronal data, this issue is becoming a matter of vital importance both to experimentalists and to modellers.

Correspondence to: G. Froylond
(e-mail: froyland@upb.de, URL: www.upb.de/math/~froyland)

To theoretically study such variability and irregularity of neuronal firing, many stochastic models have been derived (Holden 1976; Tuckwell 1989; Ricciardi 1995), and statistical properties like the first passage time and the coefficient of variation of interspike intervals have been analysed both analytically and numerically with these stochastic models. In fact, such statistical properties are important indicators for considering not only possible roles of cortical neurons but also possible neural codes (Abeles 1982; Softky and Koch 1993; König et al. 1996; Fujii et al. 1997; Shadlen and Newsome 1998; Freeman 2000).

On the other hand, well-controlled electrophysiological experiments with real neurons have shown that their dynamical behavior can be approximately described by deterministic nonlinear maps on appropriate Poincaré sections, especially by one-dimensional nonlinear maps with possibly chaotic dynamics (Aihara et al. 1986; Segundo et al. 1991; Hayashi and Ishizuka 1992; Mees et al. 1992; Aihara 1995). On the basis of such neuronal properties, simple mathematical models of one-dimensional nonlinear maps on neuronal response have been proposed (Aihara et al. 1990; Ichinose et al. 1998) as models of neurons in neural networks to aid our understanding of complex neurodynamics. For example, a one-dimensional map to be introduced in Sect. 4 (Ichinose et al. 1998) can well represent characteristics observed by electrophysiological experiments with squid giant axons (Mees et al. 1992; Aihara 1995). These studies imply that variability and irregularity of neuronal firing may be explained to some extent from the viewpoint of dynamical system theory. In particular, since it is not easy to obtain a sufficiently large amount of physiological data for reliable calculation of statistical properties due to nonstationarity peculiar to biological systems, it is desirable to estimate such statistical properties of a deterministic map that is approximately derived on the basis of a smaller amount of possibly stationary data.

In this article, we provide theoretical techniques to calculate such statistics as the interspike intervals and the coefficient of variation that are generated by non-

linear neuronal maps stimulated by inputs from other neurons in neural networks.

2 Outline of approach

First, we outline the approach taken for a deterministic neuron model and in later sections progress to neuron models with iid inputs, where the neuronal firing set may depend on the input. We suppose that our neuron model is governed by a deterministic map $T : X \circlearrowleft$, where X is a compact subset of $\mathbb{R}^d$ (usually X will be an interval), and denote the *firing set* by $B \subset X$. Given an initial condition $x \in X$, one can generate a trajectory of T as $\{x, Tx, T^2x, \ldots\}$. Whenever the trajectory enters B, we say the neuron "fires" and we record the time; in this way we generate a sequence of times $s_1, s_2, \ldots \in \mathbb{Z}^+$, where $T^{s_k}x \in B$ for $k \geq 1$, (and $s_0 = 0$). These times represent the instants at which the neuron fires. We are interested in the statistical properties of the sequence $t_k = s_k - s_{k-1}$, $k = 1, 2, \ldots$; the t_k are the interspike intervals.

To avoid the inefficient direct computation of very long orbits, our approach is to directly use the *natural invariant*[1] *measure* μ of the deterministic dynamics. We will approximate a probability measure μ that satisfies

$$\mu(A) = \lim_{N\to\infty} \frac{\#\{0 \leq k \leq N-1 : T^k x \in A\}}{N} \tag{1}$$

for all measurable subsets $A \subset X$ and for Lebesgue almost all $x \in X$. We assume throughout that T possesses sufficient transitivity properties so that the definition of μ in (1) is indeed independent of x.

Example 2.1 (a natural invariant measure). Define a map $T : [0,1] \circlearrowleft$ by $Tx = 4x(1-x)$; this is the often-studied logistic mapping (May 1976); see Fig. 1a. It is well known (see Lasota and Mackey 1994, for example) that the natural invariant measure μ is defined by $\mu([a,b]) = \int_a^b \phi(x)\mathrm{d}x$, where $\phi(x) = 1/(\pi\sqrt{x(1-x)})$ is the density function for the measure μ; see Fig. 1b. Infinitely long orbits $\{x, Tx, T^2x, \ldots\}$ are distributed according to ϕ for Lebesgue almost all initial points x.

The natural invariant measure μ will be combined with the *return time* function $R : X \to \mathbb{Z}^+$ defined by

$$R(x) = \inf\{k \geq 1 : T^k x \in B\} \ . \tag{2}$$

We may use this function to define what is known as an *induced map* $\tilde{T} : B \circlearrowleft$ by setting $\tilde{T}x = T^{R(x)}x$. The map $\tilde{T}$ simply maps a point $x \in B$ into the point where the orbit of x next enters B, and the sequences $R(x), R(\tilde{T}x)$, $R(\tilde{T}^2x), \ldots$ and $t_1, t_2, t_3, \ldots$ are identical.

Suppose that μ is an invariant probability measure for T; it is easy to see that the probability measure on B defined by $\mu|_B(A) := \mu(A \cap B)/\mu(B)$ is invariant under $\tilde{T}$, and therefore one may apply the Birkhoff ergodic theorem to conclude that

[1] A probability measure μ is invariant under a map T if $\mu \circ T^{-1} = \mu$.

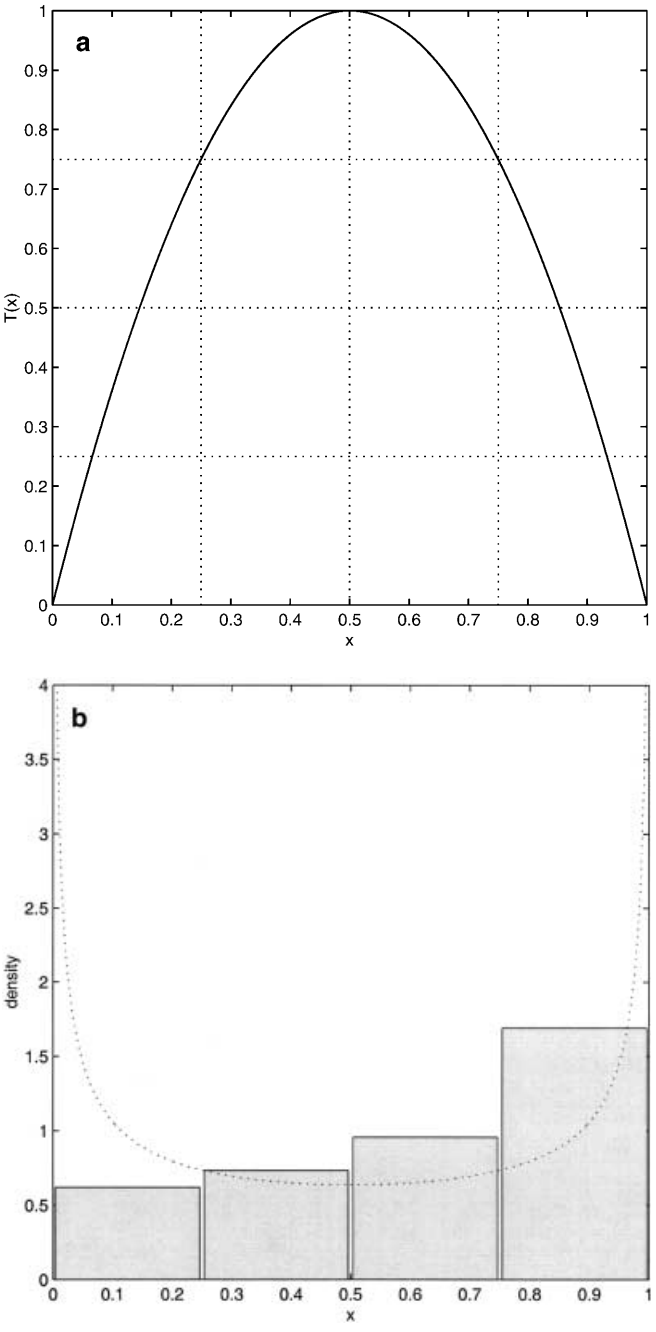

Fig. 1a,b. The logistic map. **a** Graph of the map $Tx = 4x(1-x)$. The partitioning according to the sets A_1, A_2, A_3, A_4 is indicated by the *dotted lines*. **b** Piecewise constant approximation of the density ϕ of the natural invariant measure μ. The true density ϕ is shown *dotted*

$$\mathbb{E}_{\mu|B}(R) := \int_B R(x)\mathrm{d}\mu|_B(x) = \lim_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} R(\tilde{T}^k x)$$
$$= \lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} t_k \tag{3}$$

for μ almost all $x \in B$, where $\mathbb{E}_{\mu|_B}(R)$ denotes the expectation of the function R with respect to the measure $\mu|_B$.

By Kac's theorem (Kac 1947; Petersen 1983), $\mathbb{E}_{\mu|_B}(R) = 1/\mu(B)$, a beautifully simple result. Thus the mean of the interspike intervals is given simply by the reciprocal of the natural invariant measure of the firing set.

What now about the variance of the interspike intervals? This can be defined as $\mathrm{var}_{\mu|_B}(R) = \mathbb{E}_{\mu|_B}(R^2) - \mathbb{E}_{\mu|_B}(R)^2$. We use the following result (Theorem 3 of Francke et al. 1985; see Blum and Rosenblatt 1967 for the original result; also Wolfowitz 1967):

Theorem 2.2. *Denote* $B^c = X \setminus B$. *Suppose that* μ *is ergodic*[2], *that* $0 < \mu(B) < 1$, *and define* $\mu|_{B^c}(A) = \mu(A \cap B^c)/\mu(B^c)$. *Then*

$$\mathrm{var}_{\mu|_B}(R) = \frac{1-\mu(B)}{\mu(B)}\left(2\mathbb{E}_{\mu|_{B^c}}(R) - 1/\mu(B)\right) \tag{4}$$

Thus,

1. Mean of interspike intervals $= \mathbb{E}_{\mu|_B}(R) = 1/\mu(B)$,
2. Variance of interspike intervals $= \mathbb{E}_{\mu|_B}(R^2) - \mathbb{E}_{\mu|_B}(R)^2 = \frac{1-\mu(B)}{\mu(B)}\left(2\mathbb{E}_{\mu|_{B^c}}(R) - 1/\mu(B)\right)$.

Our approach is therefore to estimate $\mu(B)$ and $\mathbb{E}_{\mu|_{B^c}}(R)$. This is achieved by approximating the piecewise smooth dynamics T as a finite state Markov chain. The unique stationary density of this Markov chain is used to approximate the natural invariant measure μ, and theory from absorbing Markov chains allows the estimation of $\mathbb{E}_{\mu|_{B^c}}(R)$.

3 Approximation by Markov chains: deterministic case

We approximate the deterministic dynamics of the neuron model as a finite state Markov chain governed by a stochastic matrix P. This is achieved by partitioning the phase space X into a finite number of partition sets $A_1, \ldots, A_n$ and identifying each state of the Markov chain with a subset of phase space; if there are n partition sets, then P is an $n \times n$ matrix. For a deterministic map T, we define

$$P_{ij} = \frac{m(A_i \cap T^{-1}A_j)}{m(A_i)} \tag{5}$$

where m denotes normalised Lebesgue measure on X. For a survey of this modelling approach, we refer the reader to Froyland (2000). One should think of P_{ij} as the probability of a trajectory being in the set A_j at time $t+1$, given that it was in A_i at time t; further details on this construction are given in the following section.

We assume that the transition matrix P has a unique invariant density p (p is a $1 \times n$ vector) satisfying $pP = p$, with $\sum_{i=1}^n p_i = 1$. We further assume that the stationary density p (normalised to sum to unity) of the Markov chain approximates the stationary distribution of orbits of the neuronal map, in the sense that

$$p_i \approx \#\{0 \le k \le N-1 : T^k x \in A_i\}/N$$

for Lebesgue almost all $x \in X$ and large N; there are good reasons for this assumption (Froyland 2000).

Example 3.1 (Modelling by Markov chains). We continue with Example 2.1 and partition $X = [0,1]$ into four equal sets $A_1 = [0, 1/4]$, $A_2 = [1/4, 1/2]$, $A_3 = [1/2, 3/4]$, $A_4 = [3/4, 1]$ (see Fig. 1a) and compute the matrix P from (5). This stochastic matrix (6) governs a Markov chain with four states, which we label $\{1,2,3,4\}$.

$$P = \begin{pmatrix} 0.2680 & 0.3178 & 0.4142 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0.2680 & 0.3178 & 0.4142 & 0 \end{pmatrix}. \tag{6}$$

The unique stationary density of P is the vector $p = [0.1547, 0.1835, 0.2391, 0.4226]$ (to four significant figures). This is a very rough piecewise constant approximation of the natural invariant measure μ with density $\phi(x) = 1/(\pi\sqrt{x(1-x)})$; see Fig. 1b.

The firing region B of the neuron will ideally correspond to a union of partition sets $B = \bigcup_{i \in F} A_i$, with $F \subsetneq \{1, \ldots, n\}$, and therefore to a collection of states (with indices in F) in the Markov chain. Denote by r the cardinality of F (with $1 \le r < n$). We approximate $\mu(A_i)$ by p_i for $i = 1, \ldots, n$, and in particular, we approximate $\mu(B)$ by $\sum_{i \in F} p_i$. Thus, we immediately arrive at the estimate.

$$\mathbb{E}_{\mu|_B}(R) = 1/\mu(B) \approx 1/\sum_{i \in F} p_i \ .$$

We have the following identifications:

Markov chain object	Dynamical system object
P (transition matrix)	T (map)
$i \in \{1, \ldots, n\}$ (state of Markov chain)	$A_i \subset X$ (partition set in phase space)
$F \subset \{1, \ldots, n\}$ (collection of "firing states")	$B = \bigcup_{i \in F} A_i$ (collection of "firing subsets")
$p = [p_1, \ldots, p_n]$ (invariant density, $pP = p$)	μ (invariant probability measure, $\mu \circ T^{-1} = \mu$)
p_i (weight given to state i by p)	$\mu(A_i)$ (weight given to partition set A_i by μ)
$p_F = \sum_{i \in F} p_i$ (weight given to "firing states")	$\mu(B)$ (weight given to "firing set")
$1/p_F$ (mean return time to F)	$1/\mu(B)$ (mean return time to B : $\mathbb{E}_{\mu|_B}(R)$)

3.1 Estimating absorption times

We now construct an estimate of $\mathbb{E}_{\mu|_{B^c}}(R)$. This quantity is the mean absorption time of points $x \in B^c$ into B, that is, the mean number of iterations required for a trajectory beginning at $x \in B^c$ to land in B (averaged over B^c with $\mu|_{B^c}$). Mean absorption times are exactly computable for finite state Markov chains; this is another reason for modelling our neuron system as such. We follow the exposition of Kemeny and Snell (1976). Write P in the block form:

[2] An invariant measure μ is ergodic if for all measurable sets $A \subset X$, $T^{-1}A = A \Rightarrow \mu(A) = 0$ or 1. Natural invariant measures are ergodic by definition.

$$P = \begin{array}{c} \\ F^c \\ F \end{array} \begin{array}{c} \begin{array}{cc} F^c & F \end{array} \\ \begin{pmatrix} Q & W \\ U & V \end{pmatrix} \end{array} . \tag{7}$$

Here we assume that $F^c = \{1, \ldots, n-r\}$ and $F = \{n-r+1, \ldots, n\}$; this can easily be accomplished by rearranging the indices of the chain.

Definition 3.2. Define $\mathscr{N}_j : F^c \to \mathbb{Z}^+$ by setting $\mathscr{N}_j(i)$ to equal the mean number of times that a trajectory of the Markov chain starting in state $i \in F^c$ visits the state $j \in F^c$ before the trajectory enters F for the first time (averaging over all trajectories starting at i). Further, define $\mathscr{R} : \{1, \ldots, n\} \to \mathbb{Z}^+$ by setting $\mathscr{R}(i)$ to equal the mean time required for a trajectory starting in state i to enter F for the first time. Note that $\mathscr{R}(i) \geq 1$ as we stipulate that trajectories starting in F must first leave F in order to reenter again due to refractoriness.

Clearly $\mathscr{R}|_{F^c} = \sum_{j \in F^c} \mathscr{N}_j$; for $i \in F^c$, $\mathscr{R}(i)$ is the mean number of iterations required for a trajectory starting at i to be absorbed into the collection of states F. One should think of $\mathscr{R} : \{1, \ldots, n\} \to \mathbb{Z}^+$ as being analogous to the function $R : X \to \mathbb{Z}^+$.

Theorem 3.3. *Assume that P has the form (7), and set $S = (I - Q)^{-1}$. Then $S_{ij} = \mathscr{N}_j(i)$.*

Proof. See Theorem 3.2.4 in Kemeny and Snell (1976). □

Corollary 3.4. *$\mathscr{R}(i) = \sum_{j \in F^c} S_{ij} = \tau_i$ for $i = 1, \ldots, n-r$, where the $(n-r) \times 1$-vector τ is the unique solution of the linear equation $(I - Q)\tau = \underbrace{(1, 1, \ldots, 1)}_{n-r}^{\top}$.*

Proof. The fact that $\mathscr{R}(i) = \sum_{j \in F^c} S_{ij}$ follows from Theorem 3.3. For the second claim, using the equality $\sum_{k=1}^{n-r} S_{ik}^{-1} S_{kj} = I_{ij}$, one has

$$\sum_{j=1}^{n-r} \left(\sum_{k=1}^{n-r} S_{ik}^{-1} S_{kj} \right) = 1 \quad \text{for all } i = 1, \ldots, n-r,$$

$$\Rightarrow \sum_{k=1}^{n-r} S_{ik}^{-1} \underbrace{\sum_{j=1}^{n-r} S_{kj}}_{\tau_k} = 1 \quad \text{for all } i = 1, \ldots, n-r.$$

Since $S^{-1} = (I - Q)$, we get the result. □

Definition 3.5. Define $p_i^F = p_i / p_F$ for $i \in F$; p^F is simply the (normalised) restriction of the density p to states in F. Similarly, define $p_i^{F^c} = p_i / (1 - p_F)$ for $i \in F^c$, where p^{F^c} is the restriction of the density p to states in F^c. Set τ to be the unique solution of the equation

$$(I - Q)\tau = (1, 1, \ldots, 1)^{\top} . \tag{8}$$

Finally, define $\mathbb{E}_{p^{F^c}}(\mathscr{R}) = \sum_{i=1}^{n-r} p_i^{F^c} \tau_i$.

$\mathbb{E}_{p^{F^c}}(\mathscr{R})$ is the expected absorption time into F of a typical trajectory starting in F^c and is analogous to the quantity $\mathbb{E}_{\mu|_{B^c}}(R)$; in fact, we use the former quantity as an approximation of the latter.

Our estimates of the mean and variance of the interspike intervals are now given by

1. mean of interspike intervals $= \mathbb{E}_{\mu|_B}(R) = 1/\mu(B) \approx 1/p_F$
2. variance of interspike intervals $= E_{\mu|_B}(R^2) - \mathbb{E}_{\mu|_B}(R)^2 \approx \frac{1-p_F}{p_F}(2\mathbb{E}_{p^{F^c}}(\mathscr{R}) - 1/p_F)$

Example 3.6 (Calculation of return time statistics). Continuing with Example 3.1, suppose that the firing set $B = [1/2, 1]$, so that $F = \{3, 4\}$ are the "firing" states. We may immediately then state that the mean return time to states in F (or the set B) is

$$1/p_F := 1/(p_3 + p_4) = 1/0.6618 = 1.5111 . \tag{9}$$

P is already in the block form (7), so we may immediately extract the upper left 2×2 block

$$Q = \begin{pmatrix} 0.2680 & 0.3178 \\ 0 & 0 \end{pmatrix} .$$

We obtain τ by solving the linear equation $(I_2 - Q)\tau = (1, 1)^{\top}$, where I_2 is the 2×2 identity matrix; numerically

$$\begin{pmatrix} 0.7320 & -0.3178 \\ 0 & 1 \end{pmatrix} \tau = \begin{pmatrix} 1 \\ 1 \end{pmatrix} ,$$

which has the unique solution $\tau = (1.8003, 1)^{\top}$. The density p^{F^c} on $F^c = \{1, 2\}$ is given by

$$\begin{aligned} p^{F^c} &= [p_1^{F^c} p_2^{F^c}] := [p_1 p_2]/(1 - p_F) \\ &= [0.1547, 0.1835]/(1 - 0.6618) \\ &= [0.4575, 0.5425] . \end{aligned}$$

We now compute the average time required for absorption into F:

$$\begin{aligned} \mathbb{E}_{p^{F^c}}(\mathscr{R}) &= p_1^{F^c} \tau_1 + p_2^{F^c} \tau_2 \\ &= 0.4575(1.8003) + 0.5425(1.0) \\ &= 1.3661 . \end{aligned} \tag{10}$$

Using item 2 above, with the values from (9) and (10), we obtain the variance of return times to the set F:

$$\begin{aligned} \mathrm{var}_{p^F}(\mathscr{R}) &= \frac{1 - 0.6618}{0.6618}(2(1.3661) - 1/0.6618) \\ &= 0.6241 . \end{aligned}$$

Thus a (very approximate) mean return time of orbits of the map T of Example 2.1 to the set $B = [1/2, 1]$ is 1.5111 time units, and the (again very approximate) variance of these return times is 0.6241.

4 Deterministic systems: numerical example

We study a neuron model that is derived by extending the concept of isochrons (Winfree 1974, 1980; Glass and Mackey 1988) from oscillatory to excitable systems for modelling excitable dynamics peculiar to biological

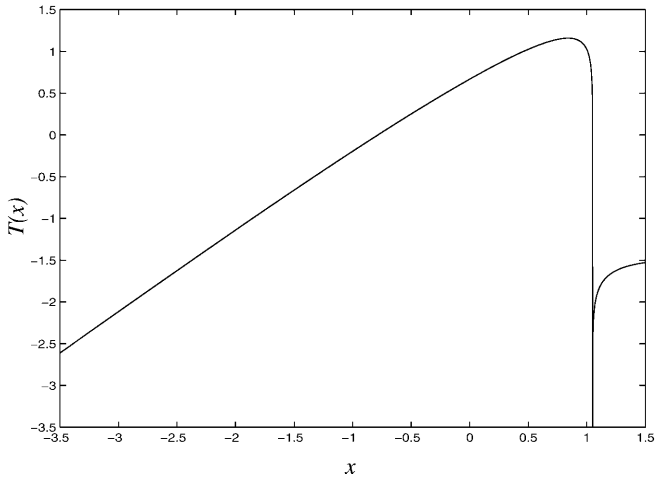

Fig. 2. Graph of T

neurons (Ichinose et al. 1998). A simple version of such a model with periodic stimulation composed of impulsive pulses is described by the following map $T : [-3.5, 1.5] \circlearrowleft$.

$$T(x) = \begin{cases} -\epsilon \log((1 - g(\gamma e^{-(x+L)}))/(s - g(\gamma e^{-(x+L)}))) \\ \quad + \log(\beta/(1 - (1-\beta)\gamma e^{-(x+L)})), \\ \qquad \text{if } s > g(\gamma e^{-(x+L)}), \\ (x + L) - \epsilon \log(g(\gamma e^{-(x+L)})/(g(\gamma e^{-(x+L)}) - s)), \\ \qquad \text{if } s < g(\gamma e^{-(x+L)}) \ , \end{cases} \tag{11}$$

where $\alpha = 0.15, \beta = 0.3, \epsilon = 0.2, \gamma = (1 + 1/\alpha)^{\epsilon}, s = 0.6$, $g(x) = \alpha + \beta + (1 - \beta)x$, and the value of L is chosen such that $s = g(\gamma e^{-(1.05+L)})$; this choice of L (namely $L = 0.8978$) fixes the position of the singularity in the graph of T at $x = 1.05$; see Fig. 2. This map represents the response of a neuron to periodic stimulation with constant strength s and period L. In other words, the map describes the charge of the isochronal coordinate of a neuron from x to $T(x)$ due to stimulation with the strength s, where x is the old isochronal coordinate a time L earlier and $T(x)$ is the new one just after the stimulation (see Judd and Aihara 1993; Ichinose et al. 1998 for details). The firing set $B = [1.05, 1.5]$; this is the set to the right of the singularity in the graph. We initially constructed $n \times n$ transition matrices using equipartitions of $[-3.5, 1.5]$ with $A_i = [-3.5 + 5(i-1)/n, -3.5 + 5i/n]$, $i = 1, \ldots, n$ for $n = 500, 1000, 2000, 5000, 10000$. The results of the computations[3] (performed with MATLAB and the GAIO[4] software package) are shown in Table 1.

[3] We give some indication of the computational effort required for these calculations. Using a Pentium III 450 MHz processor, to construct the 10000×10000 transition matrix P took 90 CPU seconds; see Sect. 6.1 for a brief description of how we numerically approximate the matrix P. To find the fixed left eigenvector p to machine precision took 5.2 CPU seconds (using the power method), and to calculate τ and $\mathrm{var}_{p^F}(\mathscr{R})$ took 0.66 CPU seconds. The storage of the 10000×10000 transition matrix required 400 kilobytes of memory.

[4] Available at `http : //www.upb.de/math/~agdellnitz/gaio`

Table 1. Approximations of $\mathbb{E}_{\mu|_B}(R)$ and $\mathrm{var}_{\mu|_B}(R)$ using equipartitions of $X = [-3.5, 1.5]$

n	Approximation of $\mathbb{E}_{\mu|_B}(R)$	Approximation of $\mathrm{var}_{\mu|_B}(R)$
500	6.4321	3.2174
1000	6.3178	2.4851
2000	6.2605	2.1717
5000	6.1923	1.7567
10000	6.1851	1.5862

A second means of partitioning the phase space $[-3.5, 1.5]$ (described in Dellnitz and Junge 1998) was also tested in order to obtain a faster rate of convergence. The algorithm is as follows:

Algorithm 1 (Dellnitz and Junge 1998)

0. Select an initial partition $\{A_1, \ldots, A_n\}$ of cardinality $n = n_0$ and an approximate cardinality n_{final} of the desired final partition with $n_0 \ll n_{\text{final}}$.
1. Calculate the transition matrix P_n and its normalised left fixed vector $p_n = [p_{n,1}, \ldots, p_{n,n}]$.
2. Define an index set $\mathscr{I} \subset \{1, \ldots, n\}$ by $i \in \mathscr{I} \iff p_{n,i} \geq 1/n$.
3. For every $i \in \mathscr{I}$, divide the partition set A_i in half to create two new partition sets; this will create a refined partition of X with cardinality $n' > n$.
4. If $n' < n_{\text{final}}$, set $n := n'$ and return to 1. Otherwise, keep the current partition and proceed to carry out the calculations of Sects. 2 and 3.

The reason for trying this alternative algorithm is that it is observed that the invariant density of T appears to be very singular in nature; see Fig. 3a,b.

Algorithm 1 attempts to place approximately equal mass in each partition set; thus, sets in the singular regions of the density are made smaller, and we focus more precisely on these regions, while ignoring the regions with little mass that give us correspondingly little information.

The computational results[5] using Algorithm 1 are shown in Table 2. We observe much better convergence results in this case. Compare these values with $\mathbb{E}_{\mu|_B}(R) \approx 6.2151632 \pm 0.0005108$ and $\mathrm{var}_{\mu|_B}(R) \approx 1.6020849 \pm 0.0023237$ [(mean) $\pm$ (one SD)] obtained by direct calculation[6] from 10 orbits of length 10^6.

An approximation of the function $R : B^c \to \mathbb{Z}^+$ is shown in Fig. 4. As n increases and the Markov chain more closely approximates the deterministic nature of the map T, the function R should take on only integer values. To calculate $\mathbb{E}_{\mu|_{B^c}}(R)$, we integrate the function shown in Fig. 4 with respect to the restriction of the density shown in Fig. 3 to the interval $[-3.5, 1.05]$; in terms of matrix calculations, this simply amounts to taking the dot product of the vectors τ and p^{F^c}.

[5] To construct the 9177×9177 transition matrix P took 65 CPU seconds, to find the fixed left eigenvector p took 4.3 CPU seconds, and to calculate τ and $\mathrm{var}_{p^F}(\mathscr{R})$ took 1.2 CPU seconds.

[6] Taking around 225 CPU seconds for *each* orbit

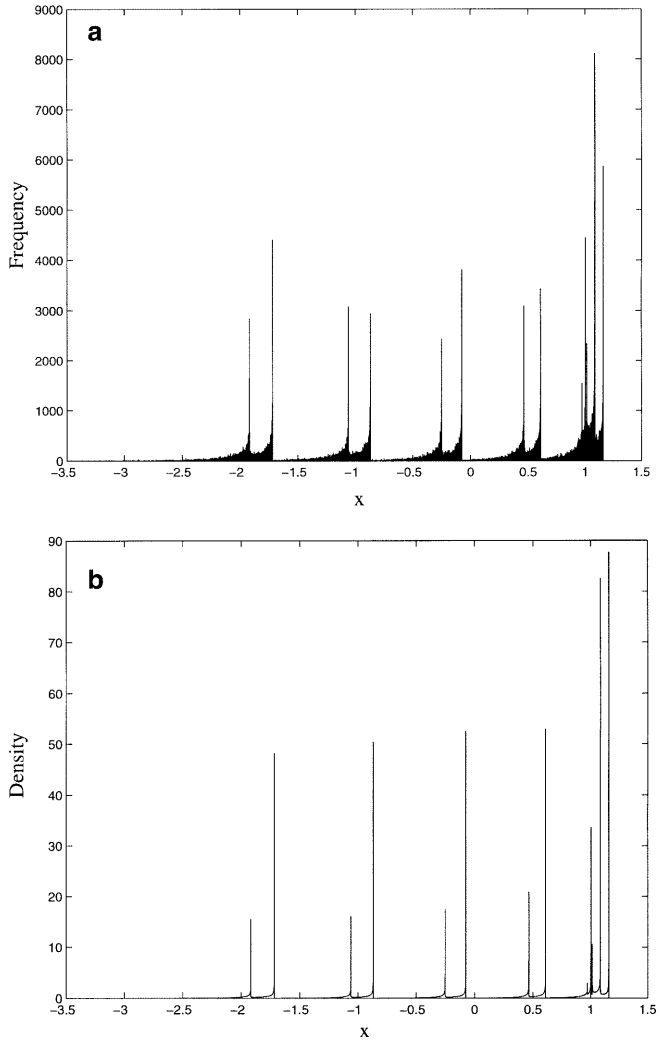

Fig. 3a,b. Comparison of densities. **a** Histogram from an orbit of length 10^6 binned onto 10^4 sets. **b** An approximate density of μ: Plot of the left fixed eigenvector p of a transition matrix P constructed from $n = 9167$ partition sets, using Algorithm 1 to partition the interval $[-3.5, 1.5]$

Table 2. Approximations of $\mathbb{E}_{\mu|_B}(R)$ and $\mathrm{var}_{\mu|_B}(R)$ using Algorithm 1 to select partitions of $[-3.5, 1.5]$

n	Approximation of $\mathbb{E}_{\mu\|_B}(R)$	Approximation of $\mathrm{var}_{\mu\|_B}(R)$
492	6.2262	1.9419
1019	6.1907	1.6208
2130	6.2033	1.5906
4416	6.2105	1.6022
6376	6.2123	1.6044
9177	6.2128	1.6047
13335	6.2115	1.5964

The values of the interspike interval (ISI) variance are more difficult to approximate than the mean interspike interval because they are a higher-order quantity. The severe nonlinearities of this map also make it difficult to approximate.

If our neuron system consists of a single deterministic mapping, we now know how to proceed. Note that the bulk of the time required by the computations is in computing the transition matrix. The evaluation of the invariant density and calculation of the mean and variance of the interspike intervals requires much less time. Therefore, if our system consists of random compositions of a collection of mappings, such as systems that arise with inputs arriving at random times, we can make a great saving on calculating the mean and variances of ISIs, especially if we wish to alter the random part of the system while keeping the same collection of maps. The next two sections will detail these extensions.

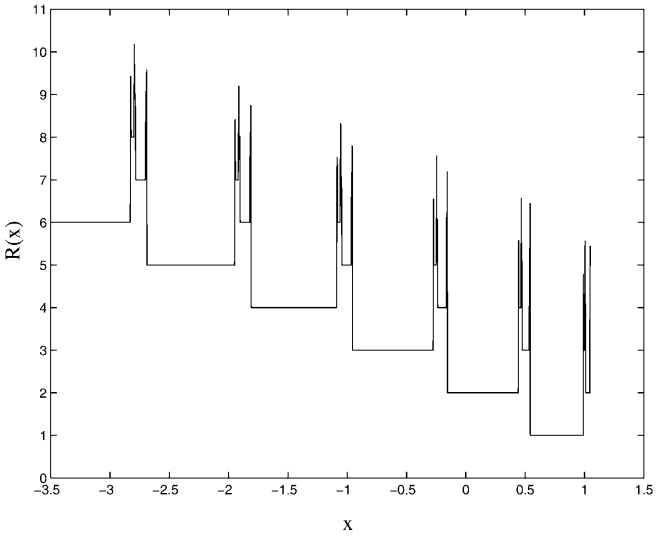

Fig. 4. Graph of an approximation of $R : B^c \to \mathbb{Z}^+$: Plot of τ constructed from a 5000×5000 transition matrix generated on an equipartition

5 Systems with discrete random iid inputs and a fixed firing set

We now allow for our neuron model to have random input. Suppose that at each time step, an input is selected in an independently, identically distributed (iid) fashion. This means that the value of the input is randomly selected according to a fixed probability distribution on the input space at each time step, independent of previously selected inputs. The effect of this input will be that possibly different dynamics occur at each time step. These different dynamics are represented by different maps $T_1, T_2, \ldots, T_r$, and the (random) orbits produced are of the form $\{x, T_{i_1}x, T_{i_2}(T_{i_1}x), T_{i_3}(T_{i_2}(T_{i_1}x)), \ldots\}$, where the values of the indices $i_1, i_2, \ldots$ are produced by an iid process. We assume that there are only finitely many different inputs allowed, meaning that we have a finite collection of maps (an extension to approximating continuous, rather than discrete, sets of inputs is described in the following section). An example of an input may be to add some (random) constant value to the map T of Sect. 4. Then we would define $T_k = T + e_k$, where e_k is some random value. Suppose that the input described by T_k occurs with probability w_k, and that the firing set B remains

fixed for all inputs (the next section also shows how to deal with firing sets that depend on inputs). We construct a stochastic matrix $P(k)$, as in (5) for each of the maps T_k and combine them to form a matrix

$$P = \sum_{k=1}^{r} w_k P(k) \tag{12}$$

The analysis described for the deterministic case may be applied directly to the matrix P to yield the mean and variance of the interspike intervals for the neuron system with iid random inputs.

6 Systems with continuous random iid inputs and variable firing sets: numerical example

We consider the map (11) where the value of L is now governed by a continuous Poisson process, rather than taking the fixed value $L = 0.8978$. The resulting system represents the response of a neuron to Poisson process stimulation with constant strength s.

More precisely, we consider random applications of the maps $T(x + L)$, where $L \in \mathbb{R}^+$ is drawn in an iid fashion from the exponential distribution with probability density[7] $e(\lambda, L) = \lambda \exp(-\lambda L)$. The parameter $\lambda > 0$ controls the "tightness" of the distribution about $L = 0$, with large λ giving distributions concentrated about $L = 0$, and small λ giving less concentrated densities.

We consider the orbit defined by $x_N = T(x_{N-1} + L_{N-1})$, where the sequence $L_0, L_1, \ldots$ is drawn in an iid fashion from the exponential distribution ($L_i \in \mathbb{R}^+$ for all i). See Fig. 5 for a plot of x_{i+1} versus x_i from a long random orbit; Figure 6 shows a histogram of the values of x_i in this long orbit.

The firing set B described in earlier sections now depends on the map $x \mapsto T(x + L)$ that is applied, since the "firing section" (with boundary given by the singularity of T) moves in phase space with L, being defined by

$$x \in B_L \iff s > g(\gamma e^{-(x+L)}) \tag{13}$$

[7] There is experimental evidence that interspike intervals of neuronal firing are nearly equivalent to a Poisson process (Softky and Koch 1993; Fujii et al. 1997; Shadlen and Newsome 1998; Freeman 2000). The one-dimensional map (11) has been derived under an assumption that intervals between successive inputs are sufficiently large (Ichinose et al. 1998). In spite of the fact that the exponential distribution generates higher probabilities for very short intervals, the approximation of the map is not so bad if most of such inputs are subthreshold because the neuronal state returns quickly to a lower slow manifold including the resting state after subthreshold response (Ichinose et al. 1998). In fact, the time required to return to the lower slow manifold after subthreshold response is given by (5) of Ichinose et al. (1998), which is the order of ϵ. The subthreshold inputs with interstimulus intervals less than this value can be a cause of errors by reducing the possibility of firing. A more precise description of neuronal firing with the two-dimensional Z-model (Ichinose et al. 1998) would increase a firing rate due to the temporal summation effect of subthreshold inputs.

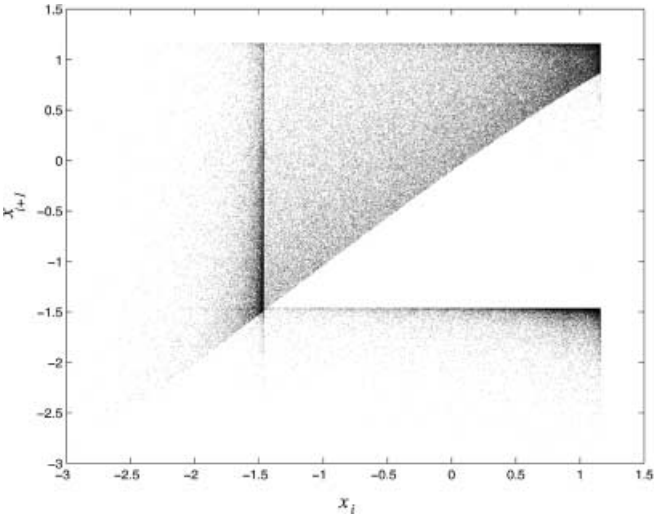

Fig. 5. Plot of x_{i+1} versus x_i for a random trajectory of length 10^5 using $\lambda = 1$

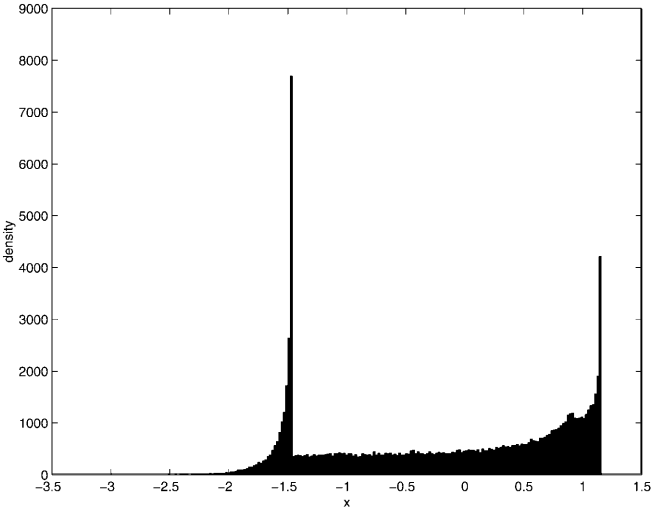

Fig. 6. Histogram of the orbit of Fig. 5 on 210 sets. For increasing λ, the peaks slowly disappear, and the mass is more evenly spread

Thus, if at time step i, we select the map $x \mapsto T(x + L_i)$, then we must check whether at this time, the current point x_i lies in B_{L_i} (if it does, then the neuron fires, and if not, the neuron does not fire).

6.1 The approximation

All of the interesting dynamics occur in the interval $[-3.5, 1.5]$, so we partition this interval into 256 equal sets, namely, $A_i = [-3.5 + 5(i-1)/256, -3.5 + 5i/256]$, $i = 1, \ldots, 256$. We choose an equipartition, as we expect the invariant density *not* to be singular in nature, due to the underlying smooth random process forcing the dynamics.

This example presents two new difficulties that we have not yet encountered. First, our random system not only uses infinitely many maps $x \mapsto T(x + L)$, $L \in \mathbb{R}^+$,

but the parameter range extends over all of $\mathbb{R}^+$. Second, as already mentioned, the firing set B_L depends on the map applied.

We get around the first problem by truncating and then discretising the exponential distribution $e(L)$. Precisely, for our computations, we use 100 maps, $T_k(x) := T(x + (k-1)/10)$, where $k = 1, \ldots, 100$. We will say that the map T_k is selected with probability

$$e'(\lambda, k) := \frac{\int_{(k-1)/10-0.05}^{(k-1)/10+0.05} \lambda \exp(-\lambda L) \mathrm{d}L}{\int_0^{9.95} \lambda \exp(-\lambda L) \mathrm{d}L}$$

with an appropriate modification for $k = 1$.

Remark 6.1. Our rationale behind the truncation of $\mathbb{R}^+$ to the interval $[0, 10]$ is that provided $\lambda \geq 0.5$, the probability of a map T_k being selected with $k > 100$ is less than 0.0064 (in other words, very small).

We have decided on a uniform discretisation of the interval $[0, 10]$ because the firing set depends linearly on L (and hence also on k), namely $B_k = [1.9478 - (k-1)/10, \infty)$ [calculated from (13)].

We now are in the situation where we have 100 maps T_k, and these maps are being composed in an iid fashion, where T_k has probability $e'(\lambda, k)$ of occurring. To produce a suitable Markov chain that will capture the essential elements of the random dynamics, we compute a matrix $P(k)$ for each map T_k via

$$P(k)_{ij} = \frac{m(A_i \cap T_k^{-1} A_j)}{m(A_i)} \quad . \tag{14}$$

Lebesgue measure m is simulated (using GAIO, or otherwise) by a uniform grid of 1000 test points within each A_i; these test points are mapped forward by T_k to numerically evaluate the entries of $P(k)$. The matrix $P(k)$ is a Markov chain approximation of the action of the map T_k. These matrices are then arranged in an array:

$$\mathbf{M}(\lambda) = \begin{pmatrix} e'(\lambda,1)P(1) & e'(\lambda,2)P(1) & \cdots & e'(\lambda,100)P(1) \\ e'(\lambda,1)P(2) & e'(\lambda,2)P(2) & \cdots & e'(\lambda,100)P(2) \\ \vdots & \vdots & \ddots & \vdots \\ e'(\lambda,1)P(100) & e'(\lambda,2)P(100) & \cdots & e'(\lambda,100)P(100) \end{pmatrix}$$

Notice that $\mathbf{M}$ (dropping λ dependence) is a stochastic matrix and thus defines a Markov chain. One should think of the (i,j)th entry $(i, j = 1, \ldots, n)$ of the (k,l)th block $(k, l = 1, \ldots, 100)$ of $\mathbf{M}$ as the probability that (1) T_l will be the next map to be used and (2) $T_k x \in A_j$, given that (1) T_k is the current map used and (2) $x \in A_i$. We will use $\mathbf{M}$ in place of the matrix P of Sect. 3. The first thing that we need to do is to identify the "firing set" F of Sect. 3. In the (k,l)th block of $\mathbf{M}$, we should label all states corresponding to sets in B_k as firing states and include them in F. Thus, from the kth n-block, (with labels $(k-1)n + 1, \ldots, kn$) we have a selection of state labels corresponding to the firing set B_k. For each $k = 1, \ldots, 100$ we add these state labels to the collection F.

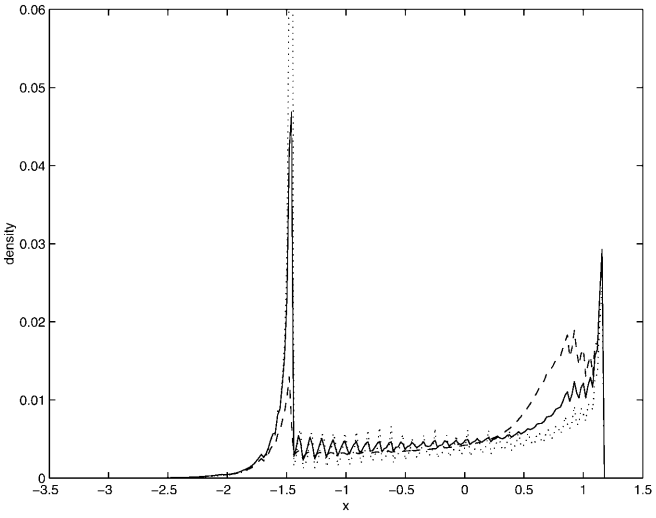

Fig. 7. Invariant density estimates for $\lambda = 0.5$ (*dotted*), $\lambda = 1.0$ (*solid*), and $\lambda = 2.0$ (*dashed*). The large peak at the left has been truncated for $\lambda = 0.5$. These are plots of the vector $\tilde{\mathbf{m}}(\lambda) := \sum_{k=1}^{100} \mathbf{m}^{(k)}$ where $\mathbf{m} = [\mathbf{m}^{(1)}|\mathbf{m}^{(2)}|\cdots|\mathbf{m}^{(100)}]$ has been decomposed into 100 blocks of length 256 ($\mathbf{m}^{(k)} = [\mathbf{m}_{256(k-1)+1}, \ldots, \mathbf{m}_{256k}]$). Compare the solid curve with Fig. 6

We now find the fixed left eigenvector $\mathbf{m}$ so that $\mathbf{mM} = \mathbf{m}$. This eigenvector may be used to approximate the invariant density of our random system; see Fig. 7.

Using the identifications $\mathbf{M} \leftrightarrow P$ and $\mathbf{m} \leftrightarrow p$, we use the approximation method described in Sect. 3 to compute the mean and variance of the return times to F for the Markov chain governed by $\mathbf{M}$. The return times to F correspond to firing times of the random system.

A final approximation enters in the identification of the firing states. As we have set our problem up, the firing set B_k is not exactly a union of partition sets A_i. Each B_k is an interval of the form $[1.9478 - (k-1)/10, \infty)$ as noted earlier. The left endpoint $1.9478 - (k-1)/10$ will most likely lie in the interior of one of the A_i. We therefore have the choice of including or excluding the single set A_i from the collection of firing sets. In our computations, we use both options and find that the answers produced are very similar.

6.2 Discussion of results

Figures 8 and 9 show the results of our computations for $\lambda = 0.1, 0.2, \ldots, 2.9, 3.0$. Note that the matrices $P(k)$ *do not need to be recomputed* for different λ values; they are fixed and independent of λ. For each λ, the $P(k)$ are combined in different ways to form $\mathbf{M}(\lambda)$. From the new $\mathbf{M}(\lambda)$, we compute the fixed left eigenvector $\mathbf{m}(\lambda)$ and solve (2) to compute the variance. Note also that the firing sets B_k are also fixed and independent of λ, so they do not need to be recalculated.

Figure 8 shows that both the mean and variance of interspike intervals are monotonically increasing with

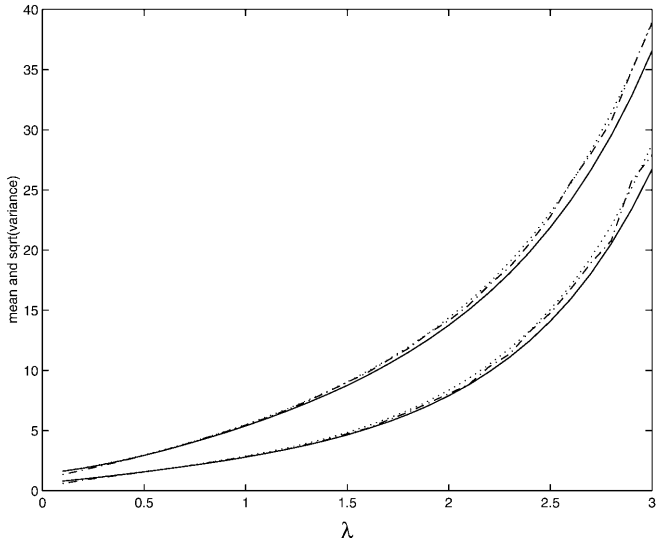

Fig. 8. Plot of the mean (*upper lines*) and $\sqrt{\text{variance}}$ (*lower lines*) using inclusion (*solid*), exclusion (*dotted*), and a typical orbit of length 10^5 (*dash-dot*)

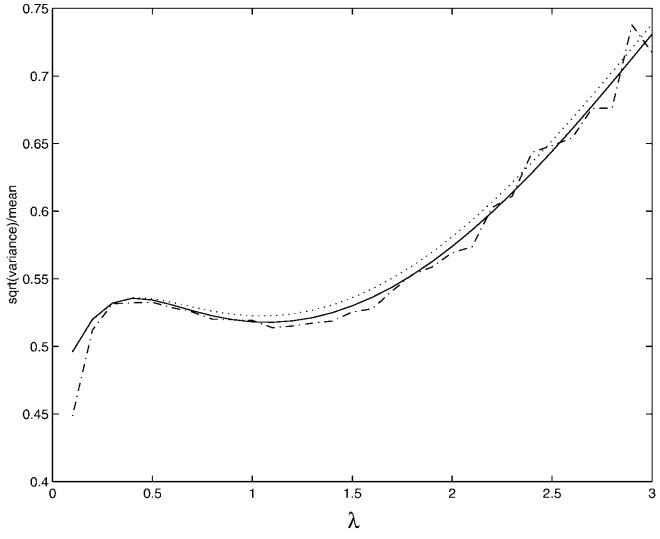

Fig. 9. Plot of the coefficient of variation $\sqrt{\text{variance}}/\text{mean}$ using inclusion (*solid*), exclusion (*dotted*), and a typical orbit of length 10^5 (*dash-dot*)

increasing λ. This makes sense, because firing is only possible[8] for maps $x \mapsto T(x+L)$ with $L > 1.9478 - \max_{x\in[-3.5,1.5]} T(x) \approx 0.789$ and with increasing λ such maps are applied less frequently. Figure 9 suggests that there is a minimum in the coefficient of variation $\sqrt{\text{var}(R)}/\mathbb{E}(R)$ (the ratio of the mean interspike interval to the standard deviation of the interspike intervals) around $\lambda = 1.0$ or 1.1. This minimum is difficult to discern from a direct calculation via a long orbit (see the dash-dot curve in Fig. 9) as there is significant statistical fluctuation. As mentioned in Remark 6.1, the solid and dotted curves should not be taken seriously for $\lambda < 0.5$ as our discrete approximation of the exponential distribution begins to break down in this region. Likewise, these curves should not be taken seriously for very large λ, as inputs with very short interstimulus intervals do not satisfy the assumption under which the map (11) is derived (see the remarks in Footnote 7).

There are many improvements that could have been made to our calculations. Obviously, we could have used finer partitions of the interval $[-3.5, 1.5]$, and also a finer discrete approximation of the exponential distribution (using more maps T_k). We could also have tried to match the map parameters k to the boundaries of our partition sets in $[-3.5, 1.5]$ to make our firing sets B_k exact unions of partition sets, and to avoid the "inclusion/exclusion" dilemma. Various alternative partitioning strategies could also have been tried, such as having smaller partition sets in regions where the invariant density is higher (the "peak" regions of Fig. 7, for example). Instead, we have taken a simple-minded approach, where a relatively coarse partition of $[-3.5, 1.5]$ has been used, with a relatively small number of maps T_k. We made no attempt to match partition set boundaries with the boundaries of the firing set, and no sophisticated partitioning methods were used. Still, we achieved useful and apparently accurate results (when compared with direct calculations from time series) with this straightforward approach, using no prior knowledge of the dynamics. Thus, these computational techniques are effective, even when applied relatively crudely. We emphasise that the savings in computing times using these techniques are considerable, especially when investigating a range of stochastic input processes.

In summary, we have demonstrated theoretical techniques to calculate statistics such as the mean interspike intervals and the coefficient of variation on the basis of nonlinear neuronal maps. Future problems are to apply these techniques not only to real physiological data (Mees et al. 1992; Aihara 1995) but also to neuron models incorporated in neural networks (Aihara et al. 1990; Judd and Aihara 1993, 2000).

Acknowledgements. The authors thank Michael Dellnitz, Oliver Junge and Natsuhiro Ichinose for helpful comments. G.F. was supported by a Japan Society for the Promotion of Science Postdoctoral Fellowship (research grant from the Ministry of Education, Science, and Culture No. 09-97004) and the Deutsche Forschungsgemeinschaft under Grant De 448/5-4.

References

Abeles M (1982) Role of the cortical neuron: integrator or coincidence detector? Isr J Med Sci 18: 83–92

Aertsen AMHJ, Gerstein GL, Habib MK, Palm G (1989) Dynamics of neuronal firing correlation: modulation of "effective connectivity". J Neurophysiol 61(5): 900–917

Aihara K (1995) Chaos in axons. In: Arbib MA (ed) The handbook of brain theory and neural networks. MIT Press, Cambridge, Mass., pp 183–185

Aihara K, Numajiri T, Matsumoto G, Kotani M (1986) Structure of attractors in periodically forced neural oscillators. Phys Lett A 116(7): 313–317

[8] To fire, one requires $s > g(\gamma e^{-(x+L)})$ [the upper case of (11)]. Solving this inequality for $x+L$ produces $x+L > -\log((s-\alpha-\beta)/((1-\beta)\gamma)) = 1.9478$, from which the constraint follows.

Aihara K, Takabe T, Toyoda M (1990) Chaotic neural networks. Phys Lett A 144(6/7): 333–340

Blum JR, Rosenblatt JL (1967) On the moments of recurrence time. J Math Sci 2: 1–6

Dellnitz M, Junge O (1998) An adaptive subdivision technique for the approximation of attractors and invariant measures. Comput Visual Sci 1: 63–68

Francke H, Plachky D, Thomsen W (1985) A finitely additive version of Poincaré's recurrence theorem. In: Christopeit N, Helmes K, Kohlmann M (eds) Stochastic differential systems – Proceedings of the 3rd Bad Honnef Conference. (Lecture Notes in Control and Information Science) Springer, Berlin Heidelberg New York

Freeman WJ (2000) Neurodynamics: an exploration in mesoscopic brain dynamics. Springer, London Berlin Heidelberg

Froyland G (2000) Extracting dynamical behaviour via Markov models. In: Mees A (ed) Nonlinear dynamics and statistics: Proceedings, Newton Institute. Birkhäuser, Basel, pp 283–324

Fujii H, Itoh H, Aihara K, Ichinose N, Tsukada M (1997) Dynamical cell assembly hypothesis – theoretical possibility of spatio-temporal coding in the cortex. Neural Netw 9: 1303–1350

Glass L, Mackey MC (1988) From clocks to chaos. Princeton University Press, Princeton, N.J.

Hayashi H, Ishizuka S (1992) Chaotic nature of bursting discharges in the Onchidium pacemaker neuron. J Theor Biol 156: 269–291

Holden AV (1976) Models of the stochastic activity of neurons. Springer, Berlin Heidelberg New York

Ichinose N, Aihara K, Judd K (1998) Extending the concept of isochrons from oscillatory to excitable systems for modeling an excitable neuron. Int J Bifurcation Chaos 8: 2375–2385

Judd K, Aihara K (1993) Pulse propagation networks: a neural network model that uses temporal coding by action potentials. Neural Netw 6(2): 203–215 Errata. Neural Netw 7: 1491, 1994

Judd K, Aihara K (2000) Generation, recognition and learning of recurrent signals by pulse propagation networks. Int J Bifurcation Chaos, in press

Kac M (1947) On the notion of recurrence in discrete stochastic processes. Bull Am Math Soc 53: 1002–1010

Kemeny JG, Snell JL (1976) Finite Markov chains. Springer, Berlin Heidelberg New York

Kitazawa S, Kimura T, Yin P-B (1998) Cerebellar complex spikes encode both destinations and errors in arm movement. Nature 392: 494–497

Kobayashi Y, Kawano K, Takemura A, Inoue Y, Kitama T, Gomi H, Kawato M (1998) Temporal firing patterns of purkinje cells in the cerebellar ventral paraflocculus during ocular following responses in monkeys II. Complex spikes. J Neurophysiol 80: 832–848

König P, Engel AK, Singer W (1996) Integrator or coincidence detector? The role of the cortical neuron revisited. Trends Neurosci 19: 130–137

Lasota A, Mackey MC (1994) Chaos, fractals, and noise. Stochastic aspects of dynamics, 2nd edn. (Applied Mathematical Sciences vol 97) Springer, Berlin Heidelberg New York

May RM (1976) Simple mathematical models with very complicated dynamics. Nature 261: 459–467

Mees A, Aihara K, Adachi M, Judd K, Ikeguchi T, Matsumoto G (1992) Deterministic prediction and chaos in squid axon response. Phys Lett A 169(1/2): 41–45

Petersen K (1983) Ergodic theory. (Cambridge Studies in Advanced Mathematics vol 2) Cambridge University Press, Cambridge

Ricciardi LM (1995) Diffusion models of neuron activity. In: Arbib MA (ed) The handbook of brain theory and neural networks. MIT Press, Cambridge, Mass., pp 299–304

Segundo JP, Altshuler E, Stiber M, Garfinkel A (1991) Periodic inhibition of living pacemaker neurons (I) locked, intermittent, messy, and hopping behaviours. Int J Bifurcation Chaos 1: 549–581

Shadlen MN, Newsome WT (1998) The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. J Neurosci 18: 3870–3896

Shinomoto S, Sakai Y, Funahashi S (1999) The Ornstein-Uhlenbeck process does not reproduce spiking statistics of neurons in prefrontal cortex. Neural Comput 11: 935–951

Softky WR, Koch C (1993) The highly irregular firing of cortical cells is inconsistent with temporal integration of random ESPSs. J Neurosci 13: 334–350

Tuckwell HC (1989) Stochastic processes in the neurosciences. SIAM, Philadelphia, Pa.

Winfree A (1974) Patterns of phase compromise in biological cycles. J Math Biol 1: 73–95

Winfree A (1980) The geometry of biological time. Springer, Berlin Heidelberg New York

Wolfowitz J (1967) The moments of recurrence time. Proc Am Math Soc 18: 613–614