Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
J. Math. Anal. Appl. 309 (2005) 567–582
Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS
www.elsevier.com/locate/jmaa

Precision of neural timing: The small ε limit

Colleen C. Mitchell [1]

Received 21 March 2003
Available online 16 June 2005
Submitted by M. Iannelli

Abstract

We explore the precision of neural timing in a model neural system with n identical input neurons whose firing time in response to stimulation is chosen from a density f. These input neurons stimulate a target cell which fires when it receives m hits within ε msec. We prove that the density of the firing time of the target cell converges as $\varepsilon \to 0$ to the input density f raised to the mth and normalized. We give conditions for convergence of the density in L^1, pointwise, and uniformly as well as conditions for the convergence of the standard deviations.

Keywords: Coincidence detection; Order statistics; Neural timing

1. Introduction

Coincidence detection, in which a neuron (or group of neurons) fires only when it receives two or more inputs almost simultaneously, has long been thought to play an important role in the central nervous system [1,4,7–10] And, recently, coincidence detection has been proposed as the mechanism that creates "precise timers" in the auditory brainstem [1,2,5,12,13]. These cells fire a single action potential, if they fire at all, at a precise time delay after the onset of a sound. Under repeated trials with the same sound, the standard deviation of the time delay in these precise timers is typically 0.1 msec and can be as low as 0.03 msec. This is very surprising since all the information processed by these neurons

E-mail address: ccm@math.bu.edu.
[1] Supported by NSF grants nos. DMS-9709608 and DMS-9805761.

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2004.10.070

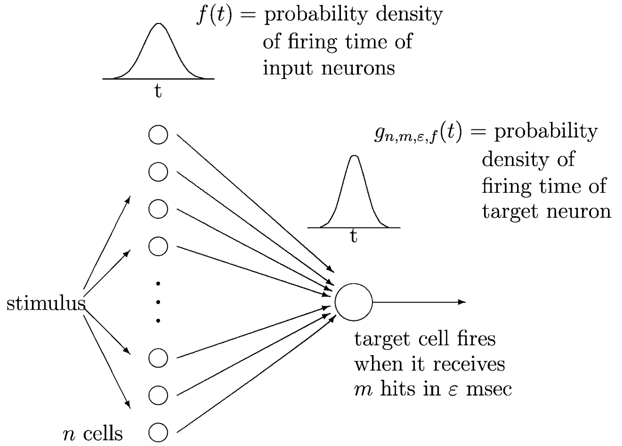

Fig. 1. Schematic of the model.

comes from the auditory nerve in which the time delays of individual fibers show standard deviations of approximately 1 msec under repeated trials. For further references and discussion of the biological background, see [14].

We formulate the question of the improvement of standard deviation by coincidence detection as follows. Imagine n identical input neurons each of which sends a projection of equal length to a target cell (see Fig. 1). In response to a stimulus each of the input neurons sends a signal after a time delay selected independently from a density f. The target cell fires, if it fires at all, at the first time that it received m inputs in the previous ε msec. We denote the random variable for the time of firing (conditioned on success) by $T_{m,n,\varepsilon,f}$, its density by $g_{m,n,\varepsilon,f}$ and its standard deviation by $\sigma_{m,n,\varepsilon,f}$. The mathematical question is to determine the behavior of $\sigma_{m,n,\varepsilon,f}$ (and $g_{m,n,\varepsilon,f}$) as a function of n, m, ε, and f.

In [14] it was shown using Monte Carlo simulations that the dependence of $\sigma_{n,m,\varepsilon,f}$ on ε and m is complex and often counter-intuitive. For example, one might expect that as ε increases, the timing would become less accurate, i.e., $\sigma_{n,m,\varepsilon,f}$ would be an increasing function of ε. In some cases, this is what was observed (for example, $n = 10$, $m = 2$, f is exponential). On the other hand, for the same f and n but with $m = 8$, $\sigma_{m,n,\varepsilon,f}$ is a decreasing function of ε and with $m = 5$, $\sigma_{n,m,\varepsilon,f}$ is non-monotone and has a peak at an intermediate value of ε. Similarly, one might expect that as m increases, $\sigma_{n,m,\varepsilon,f}$ would decrease. In fact, for most choices of parameters, $\sigma_{n,m,\varepsilon,f}$ is a non-monotone function of m. A scaling argument showed that it is sufficient to consider f with standard deviation equal to 1 msec.

This paper is devoted entirely to the mathematical issues involved in the small ε limit. Specifically, the purpose of this paper is to prove that the density $g_{m,n,\varepsilon,f}$ of $T_{m,n,\varepsilon,f}$ converges to the input density f raised to the mth power and normalized as $\varepsilon \to 0$. L^1 is the most natural type of convergence since the normalization requires division by a constant multiple of the L^1 norm of $g_{m,n,\varepsilon,f}$. We begin with the lemmas used in the L^1 proof in Section 2, then prove L^1 convergence in Section 3. Lastly, in Section 4, we address other types of convergence under additional hypothesis; most notably we give conditions for the convergence of the standard deviation.

2. Lemmas for the L^1 proof

We begin with five lemmas which will be used in the L^1 proof given in Section 3. The first lemma allows us to consider the relevant limit without normalizing the densities. Lemma 2 is essentially set theoretic and is used to estimate the key integrals which have as limits of integration a minimum of two variables. Lemmas 3–5 introduce and give properties of the bounded linear transformation J_ε, which appears repeatedly.

Lemma 1. *Let $\{f_\varepsilon\}$ be a parametrized family of non-negative L^1 functions with $\|f_\varepsilon\|_1 > 0$. Let $f_\varepsilon \to f$ in L^1 as $\varepsilon \to 0$ where f is also in L^1 with $\|f\|_1 > 0$. Then,*

$$\frac{f_\varepsilon}{\int f_\varepsilon} \xrightarrow{L^1} \frac{f}{\int f} \quad \text{as } \varepsilon \to 0.$$

Proof. Given γ, pick δ so that $\delta \leqslant \frac{\gamma \int f}{2+\gamma}$. Then, by the L^1 convergence of f_ε, we can choose α so that

$$\left|\int f_\varepsilon - \int f\right| \leqslant \int |f_\varepsilon - f| \leqslant \delta$$

for all $\varepsilon \leqslant \alpha$. Therefore,

$$\frac{f_\varepsilon}{\int f_\varepsilon} - \frac{f}{\int f} \leqslant \frac{f_\varepsilon\left(\frac{\int f - \delta}{\int f} + \frac{\delta}{\int f}\right)}{\int f - \delta} - \frac{f}{\int f} = \frac{f_\varepsilon - f}{\int f} + \frac{\delta f_\varepsilon}{\int f(\int f - \delta)}$$

and, similarly

$$\frac{f_\varepsilon}{\int f_\varepsilon} - \frac{f}{\int f} \geqslant \frac{f_\varepsilon - f}{\int f} - \frac{\delta f_\varepsilon}{\int f(\int f - \delta)}.$$

So

$$\left|\frac{f_\varepsilon}{\int f_\varepsilon} - \frac{f}{\int f}\right| \leqslant \frac{|f_\varepsilon - f|}{\int f} + \frac{\delta f_\varepsilon}{\int f(\int f - \delta)}.$$

Taking the integral of both sides gives

$$\left\|\frac{f_\varepsilon}{\int f_\varepsilon} - \frac{f}{\int f}\right\|_1 \leqslant \gamma,$$

for all $\varepsilon \leqslant \alpha$ which proves the lemma. □

Lemma 2. *We will denote the minimum of two numbers a and b by $a \vee b$. Let f be an integrable function on R^{k-l+1}, and let $k \geqslant l$, then*

$$\int_{-\infty}^{x_k}\int_{-\infty}^{x_{k-1}}\cdots\int_{-\infty}^{x_{l+1}}\int_{-\infty}^{x_l\vee x_k-\varepsilon} f(x_{l-1},\dots,x_{k-1})\,dx_{l-1}\cdots dx_{k-1}$$
$$=\int_{-\infty}^{x_k}\int_{-\infty}^{x_{k-1}}\cdots\int_{-\infty}^{x_{l+1}}\int_{-\infty}^{x_l} f(x_{l-1},\dots,x_{k-1})\,dx_{l-1}\cdots dx_{k-1}$$
$$-\int_{x_k-\varepsilon}^{x_k}\int_{x_k-\varepsilon}^{x_{k-1}}\cdots\int_{x_k-\varepsilon}^{x_{l+1}}\int_{x_k-\varepsilon}^{x_l} f(x_{l-1},\dots,x_{k-1})\,dx_{l-1}\cdots dx_{k-1}.$$

Proof. Let A, B, and C be the sets in R^{j-i+1}:

$$A=\{x_{l-1}<x_l<\cdots<x_{k-1}<x_k\},$$
$$B=\{x_{l-1}<x_l\vee x_k-\varepsilon, x_l<x_{l+1}<\cdots<x_{k-1}<x_k\},$$
$$C=\{x_k-\varepsilon\leqslant x_{l-1}<x_l<\cdots<x_{k-1}<x_k\}.$$

Then, the lemma can be rewritten

$$\int_B=\int_A-\int_C.$$

We therefore need only prove that $A=B\cup C$ and that B and C are disjoint.

$$A=A\cap\big(\{x_{l-1}<x_k-\varepsilon\}\cup\{x_{l-1}\geqslant x_k-\varepsilon\}\big)$$
$$=\big(A\cap\{x_{l-1}<x_k-\varepsilon\}\big)\cup\big(A\cap\{x_{l-1}\geqslant x_k-\varepsilon\}\big)$$
$$=B\cup C.$$

Since $\{x_l<x_k-\varepsilon\}$ and $\{x_l\geqslant x_k-\varepsilon\}$ are disjoint, their intersections with A are also disjoint. Therefore B and C are disjoint. □

Lemma 3. *Let f be a function in L^r, $1\leqslant r\leqslant\infty$, and let J_ε be the operator from L^r to L^r given by $(J_\varepsilon f)(x)=j_\varepsilon * f$ where j_ε is $\frac{1}{\varepsilon}$ for $0\leqslant x\leqslant\varepsilon$ and zero otherwise. Then for $1\leqslant i\leqslant r<\infty$,*

$$\|J_\varepsilon f\|_r\leqslant\|f\|_r\quad\textit{and}\quad\left|\int f(x)^i(J_\varepsilon f)(x)^{r-i}\,dx\right|\leqslant\|f\|_r^r.$$

Proof. Young's inequality [16] shows that $\|J_\varepsilon f\|_r\leqslant\|j_\varepsilon\|_1\|f\|_r\leqslant\|f\|_r$ and Hölder's inequality proves the second inequality. □

Lemma 4. *Let f be a density and F its cumulative distribution. Then, for $n\geqslant 0$, $J_\varepsilon(F^n f)\leqslant F^n J_\varepsilon f$.*

Proof. $(J_\varepsilon F^n f)(x)\leqslant\frac{1}{\varepsilon}\int_{x-\varepsilon}^{x}F^n(x)f(y)\,dy=F^n(x)J_\varepsilon f(x)$. □

Lemma 5. *Let f be a density. Then*:

(a) *If $f \in L^r$, $1 \leqslant r < \infty$, then $J_\varepsilon f \to f$ in L^r.*
(b) *If f is left continuous, then $J_\varepsilon f \to f$ pointwise.*
(c) *If f is uniformly continuous, the $J_\varepsilon f \to f$ uniformly.*

Proof. Parts (a) and (c) are proved in [6]. To prove (b), we let F be the cumulative distribution of f. Then by the mean value theorem, for each x,

$$(J_\varepsilon f)(x) = \frac{F(x) - F(x-\varepsilon)}{\varepsilon} = f\big(x - \tilde{\varepsilon}(x)\big),$$

where $0 < \tilde{\varepsilon}(x) < \varepsilon$. By left-continuity $f(x - \tilde{\varepsilon}(x)) \to f(x)$ pointwise as $\varepsilon \to 0$. □

3. L^1 convergence

Let $Y_1, Y_2, \ldots, Y_n$ be the independent, identically distributed firing times of the n input neurons, i.e., each of the Y_is has density f. It is more useful, however, to consider the ordered inputs known as order statistics. Let X_i be the ith order statistic, i.e., the random variable which is the ith smallest of the Y_is. A symmetry argument (for an introduction to order statistics, see [3]) shows that the joint density of the order statistics is given by

$$f_{1,2,\ldots,n:n}(x_1, x_2, \ldots, x_n) \equiv n! f(x_1) f(x_2) \ldots f(x_n) \chi_{\{x_1 < x_2 < \cdots < x_n\}}. \tag{1}$$

In terms of the order statistics, $T_{m,n,\varepsilon,f}$ is conditioned on success in triggering a response, i.e., on $A_\varepsilon := \{i\colon\ X_i - X_{i-m+1} \leqslant \varepsilon\} \neq \emptyset$. Therefore, we define

$$T_{m,n,\varepsilon,f} := \min_{i \in A_\varepsilon}(X_i).$$

In the course of the L^1 proof we take the additional time to find explicit bounds on the error terms because these will be helpful in proving the other types of convergence addressed in Section 4.

Theorem 1. *Let f be a density and let $f \in L^{m+1}$, then*

$$g_{m,n,\varepsilon,f} \xrightarrow{L^1} \frac{f^m}{\int f^m} \quad \text{as } \varepsilon \to 0.$$

Throughout the proof we will refer to the five lemmas stated and proved in Section 2. In this section, m, n, and f are fixed, so for simplicity of notation we will denote $T_{m,n,\varepsilon,f}$ by T_ε and $g_{m,n,\varepsilon,f}$ by g_ε. Note however that m, n, and f do play a strong role in the proof. In fact, it is the elimination of the dependence on n which makes the proof difficult and it is the specific characteristics of the density f which determine the type of convergence.

Proof of Theorem 1. Let $f_{\{X_i | T_\varepsilon = X_i\}}(x)$ denote the conditional density of X_i given that $T_\varepsilon = X_i$ and let P_i be the probability that $T_\varepsilon = X_i$. The density g_ε of T_ε is the normalized sum from $i = m$ to n of these conditional densities,

$$g_\varepsilon(x) = \frac{\sum_{i=m}^{n} f_{\{X_i|T_\varepsilon=X_i\}}(x) P_i}{P(\text{success})}, \tag{2}$$

where $P(\text{success}) = P(A_\varepsilon \neq \emptyset)$. Using the joint density of the X_is (1), we can compute $f_{\{X_i|T_\varepsilon=X_i\}}(x)$ by integrating over the appropriate event,

$$f_{\{X_i|T_\varepsilon=X_i\}}(x) = \frac{1}{P_i} \int_{\{T_\varepsilon=X_i=x\}} f_{1,2,\dots,n:n}(x_1, x_2, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)\, dx_1 \cdots dx_{i-1}\, dx_{i+1} \cdots dx_n. \tag{3}$$

If $T_\varepsilon = X_i = x$, then the first $i - m$ variables are sufficiently spread out to insure that i is the smallest element of A_ε. This means that for all j less than i (and of course $j \geqslant m$) $X_j - X_{j-m+1}$ must be greater than ε. If we let $k = j - m + 1$ we see that this is equivalent to the condition $X_k < X_{k+m-1} - \varepsilon$, where $k = 1, \dots, i - m$. The variables $x_{i-m+1}, \dots, x_{i-1}$ are the $m - 1$ inputs within ε of x and the last $n - i$ variables are greater than x. Then,

$$f_{\{X_i|T_\varepsilon=X_i\}}(x) = \frac{1}{P_i} n! f(x) \int_{\Omega_1} f(x_n) \cdots f(x_{i+1}) \int_{\Omega_2} f(x_{i-1}) \cdots f(x_{i-m+1}) \times \int_{\Omega_3} f(x_{i-m}) \cdots f(x_1)\, dx_1 \cdots dx_{i-1}\, dx_{i+1} \cdots dx_n, \tag{4}$$

where Ω_1, Ω_2, and Ω_3 are the sets

$$\Omega_1 = \{x < x_{i+1} < \cdots < x_n\},$$
$$\Omega_2 = \{x - \varepsilon < x_{i-m+1} < \cdots < x_{i-1} < x\},$$
$$\Omega_3 = \{x_1 < \cdots < x_{i-m} \text{ and } x_k < x_{k+m-1} - \varepsilon \text{ for } k = 1, \dots, i - m\}.$$

The upper limit of integration for x_k, where $1 \leqslant k \leqslant i - m$, is therefore the minimum of x_{k+1} and $x_{k+m-1} - \varepsilon$ which we denote $x_{k+1} \vee x_{k+m-1} - \varepsilon$ (we will call these mins). We can now write the integral over the first $i - m$ variables as follows:

$$h(x_{i-1}, \dots, x_{i-m+1}, \varepsilon) = \int_{\Omega_3} f(x_{i-m}) \cdots f(x_1)\, dx_1 \cdots dx_{i-m} = \int_{-\infty}^{x_{i-1}-\varepsilon} f(x_{i-m}) \int_{-\infty}^{x_{i-m} \vee x_{i-2}-\varepsilon} f(x_{i-m-1}) \cdots \times \int_{-\infty}^{x_2 \vee x_m - \varepsilon} f(x_1)\, dx_1 \cdots dx_{i-m}. \tag{5}$$

The upper limit on the first integral is simply $x_{i-1} - \varepsilon$ since x_{i-m+1} will be in $[x - \varepsilon, x]$ which implies that $x_{i-m+1} \vee x_{i-1} - \varepsilon = x_{i-1} - \varepsilon$. The integral over the next $m - 1$ variables is given by

$$I_i(x) = \int_{\Omega_2} f(x_{i-1}) \cdots f(x_{i-m+1}) h(x_{i-1}, \ldots, x_{i-m+1}, \varepsilon)\, dx_{i-m+1} \cdots dx_{i-1}$$
$$= \int_{x-\varepsilon}^{x} f(x_{i-1}) \cdots \int_{x-\varepsilon}^{x_{i-m+2}} f(x_{i-m+1}) h(x_{i-1}, \ldots, x_{i-m+1}, \varepsilon)\, dx_{i-m+1} \cdots dx_{i-1}. \tag{6}$$

Lastly, since $I_i(x)$ does not depend on $x_i + 1, \ldots, x_n$, it can be pulled out of the integral (4) to give

$$f_{\{X_i | T_\varepsilon = X_i\}}(x) = \frac{1}{P_i} n! f(x) I_i(x) \int_{\Omega_1} f(x_n) \cdots f(x_{i+1})\, dx_{i+1} \cdots dx_n$$
$$= \frac{1}{P_i} n! f(x) I_i(x) \frac{(1 - F(x))^{n-i}}{(n-i)!}, \tag{7}$$

where $F(x)$ is the cumulative distribution function of the Y_is. Using (7) and multiplying the numerator and denominator by the convenient factor $\frac{(m-1)!}{\varepsilon^{m-1}} \frac{(n-m)!}{n!}$, formula (2) becomes

$$g_\varepsilon = \frac{f(x) \sum_{i=m}^{n} (n-m)! \frac{(1-F(x))^{n-i}}{(n-i)!} \frac{(m-1)!}{\varepsilon^{m-1}} I_i(x)}{P_\varepsilon}, \tag{8}$$

where

$$P_\varepsilon = \frac{(m-1)!}{\varepsilon^{m-1}} (n-m)! P(\text{success})$$
$$= \frac{(m-1)!}{\varepsilon^{m-1}} (n-m)! \int f(x) \sum_{i=m}^{n} \frac{(1-F(x))^{n-i}}{(n-i)!} I_i(x)\, dx.$$

By Lemma 1, it suffices to show convergence of the un-normalized functions $P_\varepsilon g_\varepsilon \xrightarrow{L^1} f^m$.

If we could replace $\frac{(m-1)!}{\varepsilon^{m-1}} I_i(x)$ with $\frac{F(x)^{i-m}}{(i-m)!} f(x)^{m-1}$ in Eq. (8), then the binomial theorem would give us precisely $P_\varepsilon g_\varepsilon = f^m$. We therefore proceed to show that $\frac{(m-1)!}{\varepsilon^{m-1}} I_i(x)$ is approximately $\frac{F(x)^{i-m}}{(i-m)!} f(x)^{m-1}$ for each $m \leqslant i \leqslant n$ and to obtain explicit bounds on the errors. This is the heart of the proof since the integrals I_i and the errors depend on n but the limit does not.

The case $i = m$ is simple,

$$\frac{(m-1)!}{\varepsilon^{m-1}} I_m(x) = \frac{(m-1)!}{\varepsilon^{m-1}} \int_{x-\varepsilon}^{x} f(x_{i-1}) \cdots \int_{x-\varepsilon}^{x_{i-m+2}} f(x_{i-m+1})\, dx_{i-m+1} \cdots dx_{i-1}$$
$$= \frac{(F(x) - F(x-\varepsilon))^{m-1}}{\varepsilon^{m-1}} = (J_\varepsilon f)(x)^{m-1},$$

where $J_\varepsilon f = \frac{1}{\varepsilon}(F(x) - F(x-\varepsilon))$.

For $i > m$, we must find estimates of $h(x_{i-1}, \ldots, x_{i-m+1}, \varepsilon)$. Since each of the variables $x_{i-1}, \ldots, x_{i-m+1}$ in h is near x in the integral I_i, we need to show that $h(x_{i-1}, \ldots, x_{i-m+1}, \varepsilon)$ is approximately $h(x, \ldots, x, \varepsilon)$. Subtracting the integrals and using the positivity of f, for $x - \varepsilon \leqslant x_j \leqslant x$ we have

$$\begin{aligned}
&\left|h(x_{i-1},\dots,x,\dots,x_{i-m+1},\varepsilon)-h(x_{i-1},\dots,x_j,\dots,x_{i-m+1},\varepsilon)\right| \\
&\quad\leqslant \int_{-\infty}^{x_{i-1}-\varepsilon} f(x_{i-m})\cdots \int_{-\infty}^{x_{j-m+3}\vee x_{j+1}-\varepsilon} f(x_{j-m+2}) \\
&\qquad\times \int_{x_{j-m+2}\vee x_j-\varepsilon}^{x_{j-m+2}\vee x-\varepsilon} f(x_{j-m+1})\frac{F(x_{j-m+1})^{j-m}}{(j-m)!} \\
&\quad\leqslant \int_{-\infty}^{x_{i-1}-\varepsilon} f(x_{i-m})\cdots \int_{-\infty}^{x_{j-m+3}\vee x_{j+1}-\varepsilon} f(x_{j-m+2})\frac{F(x_{j-m+2})^{j-m}}{(j-m)!}\varepsilon(J_\varepsilon f)(x-\varepsilon) \\
&\quad=\varepsilon(J_\varepsilon f)(x-\varepsilon)\frac{F(x_{i-1}-\varepsilon)^{i-m-1}}{(i-m-1)!}.
\end{aligned}$$

Applying the triangle inequality with each of the variables $x_{i-m+1},\dots,x_{i-1}$ in $[x-\varepsilon,x]$ gives the estimate

$$\left|h(x,\dots,x,\varepsilon)-h(x_{i-1},\dots,x_{i-m+1},\varepsilon)\right|\leqslant\varepsilon(m-1)(J_\varepsilon f)(x-\varepsilon)\frac{F(x)^{i-m-1}}{(i-m-1)!}. \tag{9}$$

If $m<i<2m$, then we can compute $h(x,\dots,x,\varepsilon)$ explicitly:

$$\begin{aligned}
h(x,\dots,x,\varepsilon)&=\int_{-\infty}^{x-\varepsilon} f(x_{i-m})\cdots\int_{-\infty}^{x_3\vee x-\varepsilon} f(x_2)\int_{-\infty}^{x_2\vee x-\varepsilon} f(x_1)\,dx_1\cdots dx_{i-m} \\
&=\int_{-\infty}^{x-\varepsilon} f(x_{i-m})\cdots\int_{-\infty}^{x_3} f(x_2)\int_{-\infty}^{x_2} f(x_1)\,dx_1\cdots dx_{i-m} \\
&=\frac{F(x-\varepsilon)^{i-m}}{(i-m)!}.
\end{aligned} \tag{10}$$

If $i\geqslant 2m$, we shall show that $h(x,\dots,x,\varepsilon)$ is approximately $\frac{F(x-\varepsilon)^{i-m}}{(i-m)!}$. Above, we were able to remove all of the mins from the limits of integration. In this case there are more than $m-1$ variables so only the last $m-1$ mins can be removed yielding

$$\begin{aligned}
h(x,\dots,x,\varepsilon)&=\int_{-\infty}^{x-\varepsilon} f(x_{i-m})\cdots\int_{-\infty}^{x_{i-2m+3}} f(x_{i-2m+2})\int_{-\infty}^{x_{i-2m+2}\vee x_{i-m}-\varepsilon} f(x_{i-2m+1}) \\
&\quad\cdots\int_{-\infty}^{x_2\vee x_m-\varepsilon} f(x_1)\,dx_1\cdots dx_{i-m}.
\end{aligned}$$

Lemma 2 shows how to take an iterated integral with a min in the upper limit of the first variable and change it to a similar iterated integral without the min by subtracting an error term. Applying Lemma 2 repeatedly, we have

$$h(x,\dots,x,\varepsilon)=\frac{F(x-\varepsilon)^{i-m}}{(i-m)!}-\sum_{j=0}^{i-2m}e_j.$$

We must now estimate these error terms. In the formation of the jth error term, we have used Lemma 2 with $k=i-m-j$ and $l=i-2m-j+2$. Therefore, the jth error term is given by

$$e_j=\int_{-\infty}^{x-\varepsilon}f(x_{i-m})\cdots\int_{-\infty}^{x_{i-m-j+1}}f(x_{i-m-j})\int_{x_{i-m-j}-\varepsilon}^{x_{i-m-j}}f(x_{i-m-j-1})$$
$$\cdots\int_{x_{i-m-j}-\varepsilon}^{x_{i-2m+2-j}}f(x_{i-2m+1-j})\int_{-\infty}^{x_{i-2m+1-j}\vee x_{i-m-1-j}-\varepsilon}f(x_{i-2m-j})\cdots\int_{-\infty}^{x_2\vee x_m-\varepsilon}f(x_1).$$

Using the positivity of f, and applying Lemma 3 which bounds $\int f(J_\varepsilon f)^{m-1}$ by $\|f\|_m^m$, we have

$$\begin{aligned}
e_j&\leqslant\frac{F(x)^{i-2m-j}}{(i-2m-j)!}\frac{\varepsilon^{m-1}}{(m-1)!}\int_{-\infty}^{x-\varepsilon}f(x_{i-m})\cdots\int_{-\infty}^{x_{i-m-j+1}}f(x_{i-m-j})(J_\varepsilon f)(x_{i-m-j})^{m-1}\\
&\leqslant\frac{F(x)^{i-2m-j}}{(i-2m-j)!}\frac{\varepsilon^{m-1}}{(m-1)!}\int_{-\infty}^{x-\varepsilon}f(x_{i-m})\cdots\int_{-\infty}^{\infty}f(x_{i-m-j})(J_\varepsilon f)(x_{i-m-j})^{m-1}\\
&=\frac{F(x)^{i-2m}}{j!(i-2m-j)!}\frac{\varepsilon^{m-1}}{(m-1)!}\|f\|_m^m.
\end{aligned}$$

Note that this estimate depends on the fact that $f\in L^m$ (by interpolation, since $f\in L^{m+1}\cap L^1$ [15]). Summing these estimates and using the binomial theorem, we have that for $i\geqslant 2m$,

$$\left|h(x,\dots,x,\varepsilon)-\frac{F(x-\varepsilon)^{i-m}}{(i-m)!}\right|=\sum_{j=0}^{i-2m}e_j\leqslant\frac{\varepsilon^{m-1}}{(m-1)!}\|f\|_m^m\frac{(2F(x))^{i-2m}}{(i-2m)!}. \tag{11}$$

Now we can estimate $I_i(x)$. In the integral for $I_i(x)$ given in (6) we replace $h(x_{i-1},\dots,x_{i-m+1},\varepsilon)$ with

$$\frac{F(x)^{i-m}}{(i-m)!}+\big(h(x_{i-1},\dots,x_{i-m+1},\varepsilon)-h(x,\dots,x,\varepsilon)\big)$$
$$+\left(h(x,\dots,x,\varepsilon)-\frac{F(x-\varepsilon)^{i-m}}{(i-m)!}\right)+\left(\frac{F(x-\varepsilon)^{i-m}}{(i-m)!}-\frac{F(x)^{i-m}}{(i-m)!}\right),$$

giving rise to

$$\frac{(m-1)!}{\varepsilon^{m-1}} I_i(x) = \frac{F(x)^{i-m}}{(i-m)!}(J_\varepsilon f)(x)^{m-1} + E_1 + E_2 + E_3. \tag{12}$$

From (9),

$$|E_1| \leqslant \varepsilon(m-1)(J_\varepsilon f)(x-\varepsilon)\frac{F(x)^{i-m-1}}{(i-m-1)!}(J_\varepsilon f)(x)^{m-1}. \tag{13}$$

If $m < i < 2m$, $E_2 = 0$ by (10) and if $i \geqslant 2m$, we use (11) to obtain

$$|E_2| \leqslant \frac{\varepsilon^{m-1}}{(m-1)!}\|f\|_m^m \frac{(2F(x))^{i-2m}}{(i-2m)!}(J_\varepsilon f)(x)^{m-1}. \tag{14}$$

Finally, Lemma 4 shows that

$$|E_3| = \varepsilon\frac{(J_\varepsilon(F^{i-m-1}f))(x)}{(i-m-1)!}(J_\varepsilon f)(x)^{m-1} \leqslant \varepsilon\frac{F(x)^{i-m-1}}{(i-m-1)!}(J_\varepsilon f)(x)^m. \tag{15}$$

Summing (12) over i and using the binomial theorem yields

$$\begin{aligned} P_\varepsilon g_\varepsilon(x) &= f(x)\sum_{i=m}^{n}(n-m)!\frac{(1-F(x))^{n-i}}{(n-i)!} \\ &\quad\times\left(\frac{F(x)^{i-m}}{(i-m)!}(J_\varepsilon f)(x)^{m-1} + E_1 + E_2 + E_3\right) \\ &= f(x)(J_\varepsilon f)(x)^{m-1} + \tilde{E}_1 + \tilde{E}_2 + \tilde{E}_3, \end{aligned} \tag{16}$$

where

$$\begin{aligned} &|\tilde{E}_1| \leqslant \varepsilon(n-m)f(x)(m-1)(J_\varepsilon f)(x)^{m-1}(J_\varepsilon f)(x-\varepsilon), \\ &|\tilde{E}_2| \leqslant f(x)\frac{\varepsilon^{m-1}}{(m-1)!}(J_\varepsilon f)(x)^{m-1}\|f\|_m^m(n-m)!\frac{(2)^{n-2m}}{(n-2m)!}, \\ &|\tilde{E}_3| \leqslant \varepsilon(n-m)f(x)(J_\varepsilon f)(x)^m. \end{aligned} \tag{17}$$

We can now estimate

$$\begin{aligned} &\int\left|P_\varepsilon g_\varepsilon(x) - f(x)^m\right| \\ &\quad\leqslant \int f(x)\left|(J_\varepsilon f)(x)^{m-1} - f(x)^{m-1}\right| + \|\tilde{E}_1\|_1 + \|\tilde{E}_2\|_1 + \|\tilde{E}_3\|_1 \\ &\quad\leqslant (m-1)\|J_\varepsilon f - f\|_m\|f\|_m^{m-1} + \varepsilon m(n-m)\|f\|_{m+1}^{m+1} \\ &\qquad+ \frac{\varepsilon^{m-1}}{(m-1)!}(n-m)!\frac{2^{n-2m}}{(n-2m)!}\|f\|_m^{2m}, \end{aligned} \tag{18}$$

where we have used Lemma 3 and Hölder's inequality to bound the first term and Lemma 3 to bound the error terms. By Lemma 5, $J_\varepsilon f \to f$ in L^m. Therefore, Eq. (18) implies that $P_\varepsilon g_\varepsilon(x) \to f(x)^m$ in L^1 and the result follows from Lemma 1. □

4. Other types of convergence

In Section 3, we proved that the density $g_{m,n,\varepsilon,f}$ of the random variable $T_{m,n,\varepsilon,f}$ converges in L^1 to the limiting density $\frac{f^m}{\int f^m}$. In this section we extend Lemma 1 to include other types of convergence and state additional hypotheses on f required to obtain other forms of convergence of the densities. We then discuss convergence of the standard deviation $\sigma_{m,n,\varepsilon,f}$.

Lemma 6. *Let B be a Banach space of measurable functions. Let $\|.\|$ denote the norm on B. Let $\{f_\varepsilon\}$ be a parametrized family of non-negative functions in $L^1 \cap B$ with non-zero L^1 norm. Suppose that $f_\varepsilon \to f$ as $\varepsilon \to 0$ in L^1 and in the norm $\|.\|$ and that $\|f\| < \infty$ and $\|f\|_1 > 0$. Then,*

$$\frac{f_\varepsilon}{\int f_\varepsilon} \to \frac{f}{\int f}$$

as $\varepsilon \to 0$ in L^1 and in the norm $\|.\|$.

Proof. Given $\gamma > 0$, we can pick $\delta \leqslant \frac{\gamma(\int f)^2}{(1+\gamma)\int f + \|f\|}$. Then choose α_1 so that $|\int f - \int f_{\varepsilon_1}| \leqslant \delta$ for all $\varepsilon_1 \leqslant \alpha_1$ and choose α_2 so that $\|f - f_{\varepsilon_2}\| \leqslant \delta$ for all $\varepsilon_2 \leqslant \alpha_2$. Let $\alpha = \min(\alpha_1, \alpha_2)$. Then, using the same trick as in Lemma 1, for all $\varepsilon \leqslant \alpha$,

$$\left\| \frac{f_\varepsilon(x)}{\int f_\varepsilon} - \frac{f(x)}{\int f} \right\| \leqslant \frac{\delta}{\int f}\left(1 + \frac{\|f\| + \delta}{\int f - \delta}\right) \leqslant \gamma. \qquad \square$$

Corollary 1. *If, in addition to the hypothesis of Theorem 1, f is left-continuous, then*

$$g_{m,n,\varepsilon,f} \longrightarrow \frac{f^m}{\int f^m} \quad \text{pointwise as } \varepsilon \to 0.$$

Proof. Since f is left-continuous, Lemma 5(b) states that $J_\varepsilon f \to f$ pointwise. Therefore, Eqs. (16) and (17) imply that $P_\varepsilon g_{m,n,\varepsilon}(x) \to f(x)^m$ pointwise. For each x we let $\|f\|_x = |f(x)|$ be a norm on the set of equivalence classes of left-continuous L^1 functions where f is equivalent to g if $f(x) = g(x)$. Then Lemma 6 implies that for each x, $g_{m,n,\varepsilon,f}(x) \to \frac{f^m(x)}{\int f^m}$. $\square$

Corollary 2. *If, in addition to the hypothesis of Theorem 1, f is uniformly continuous, then*

$$g_{m,n,\varepsilon,f} \longrightarrow \frac{f^m}{\int f^m} \quad \text{uniformly as } \varepsilon \to 0.$$

Proof. Since f is uniformly continuous Lemma 5(c) states that $J_\varepsilon f \to f$ uniformly. Therefore, Eqs. (16) and (17) imply that $P_\varepsilon g_{m,n,\varepsilon}(x) \to f(x)^m$ uniformly. One can easily show that if f is uniformly continuous and integrable, then f is bounded and the hypotheses of Lemma 6 are satisfied. Therefore the convergence of $g_{m,n,\varepsilon,f}$ is also uniform. $\square$

We now turn our attention to the convergence of the standard deviation. If f and $\frac{f^m}{\int f^m}$ have finite standard deviation we wish to show that the standard deviation of $T_{m,n,\varepsilon,f}$ converges to the standard deviation of $\frac{f^m}{\int f^m}$. We use the weighted L^p spaces, $L^{p,k}$, where $f \in L^{p,k}$ if and only if

$$\|f\|_{p,k} \equiv \left(\int \left(1+x^2\right)^k \left|f(x)\right|^p \, dx\right)^{1/p} < \infty.$$

We begin with three lemmas. The first states that in order to have convergence of the standard deviations we need only show convergence in $L^{1,1}$. The second shows that J_ε is a bounded operator on $L^{p,1}$. And the third shows that $J_\varepsilon f \xrightarrow{L^{p,1}} f$.

Lemma 7. *Let f be a density in $L^{1,1}$ and let $\{f_n\}$ be a sequence of densities such that $f_n \to f$ in $L_{1,1}$. Then the standard deviations of the f_ns converge to the standard deviation of f.*

The proof of Lemma 7 is elementary and is omitted.

Lemma 8. *J_ε is a bounded linear operator on $L^{p,1}$, $1 \leqslant p < \infty$, with $\|J_\varepsilon\| \leqslant (1+\varepsilon)^{2/p}$.*

Proof. Let $f \in L^{p,1}$ and let $f_y(x) = f(x-y)$. Then, using the definition of J_ε and Hölder's inequality,

$$\|J_\varepsilon f\|_{p,1}^p = \int \left(1+x^2\right)(J_\varepsilon f)(x)^p \, dx \leqslant \int j_\varepsilon(y) \|J_\varepsilon f\|_{p,1}^{p-1} \|f_y\|_{p,1} \, dy. \tag{19}$$

Where we can estimate $\|f_y\|_{p,1}$, using the change of variables $t = x - y$, as follows:

$$\begin{aligned} \|f_y\|_{p,1}^p &= \int \left(1+t^2\right)\left|f(t)\right|^p dt + \int 2ty\left|f(t)\right|^p dt + \int y^2\left|f(t)\right|^p dt \\ &\leqslant (1+y)^2 \|f\|_{p,1}^p. \end{aligned} \tag{20}$$

Plugging back into Eq. (19) gives

$$\|J_\varepsilon f\|_{p,1}^p \leqslant \|J_\varepsilon f\|_{p,1}^{p-1} \|f\|_{p,1} \int (1+y)^{2/p} j_\varepsilon(y) \, dy \leqslant \|J_\varepsilon f\|_{p,1}^{p-1} \|f\|_{p,1} (1+\varepsilon)^{2/p}$$

and therefore $\|J_\varepsilon f\|_{p,1} \leqslant (1+\varepsilon)^{2/p} \|f\|_{p,1}$. □

Lemma 9. *If $f \in L^{p,1}$, then $J_\varepsilon f \xrightarrow{L^{p,1}} f$.*

Proof. Let $f_n(x) = f(x)$ if $|x| \leqslant n$ and 0 otherwise. Since $f \in L^{p,1}$, we can pick n so that $\|f - f_n\|_{p,1}$ is arbitrarily small and it is sufficient to show that $\|J_\varepsilon f_n - f_n\|_{p,1} \to 0$ as $\varepsilon \to 0$,

$$\|J_\varepsilon f_n - f_n\|_{p,1}^p = \int \left(1+x^2\right)|J_\varepsilon f_n - f_n|^p \leqslant \left(1+(n+\varepsilon)^2\right)\|J_\varepsilon f - f\|_p^p,$$

which can be made arbitrarily small by Lemma 5(a). □

Theorem 2. *If, in addition to the hypotheses of Theorem* 1, *we require that* $f \in L^{m+1,1}$, *then the standard deviation of* T_ε, *converges to the standard deviation of* $\frac{f^m}{\int f^m}$ *as* $\varepsilon \to 0$.

Proof. By Lemmas 6 and 7, it is sufficient to show that $P_\varepsilon g_\varepsilon \to f^m$ in $L^{1,1}$. By (16),

$$\left\| P_\varepsilon g_{m,n,\varepsilon,f} - f^m \right\|_{1,1} \leqslant \left\| f (J_\varepsilon f)^{m-1} - f^m \right\|_{1,1} + \|\tilde{E}_1\|_{1,1} + \|\tilde{E}_2\|_{1,1} + \|\tilde{E}_3\|_{1,1}.$$

We estimate the first term using Hölder's inequality and the bound on J_ε proven in Lemma 8,

$$\begin{aligned} \left\| f (J_\varepsilon f)^{m-1} - f^m \right\|_{1,1} &\leqslant \|f\|_{m,1} \|J_\varepsilon f - f\|_{m,1} \big(\|J_\varepsilon f\|_{m,1}^{m-2} + \|J_\varepsilon f\|_{m,1}^{m-3} \|f\|_{m,1} \\ &\qquad + \cdots + \|f\|_{m,1}^{m-2} \big) \\ &\leqslant (1+\varepsilon)^2 (m-1) \|f\|_{m,1}^{m-1} \|J_\varepsilon f - f\|_{m,1}. \end{aligned}$$

Since $f \in L^{m,1}$ by interpolation, Lemma 9 applies and $\|J_\varepsilon f - f\|_{m,1} \to 0$. Lastly we estimate $\|\tilde{E}_1\|_{1,1}$, $\|\tilde{E}_2\|_{1,1}$, and $\|\tilde{E}_3\|_{1,1}$ using the bounds on these error terms from Eq. (17) and the bound on J_ε proven in Lemma 8. For $\tilde{E}_1$ we must also use Eq. (20) with $y = \varepsilon$,

$$\begin{aligned} \|\tilde{E}_1\|_{1,1} &\leqslant \varepsilon (n-m)(m-1) \|f\|_{m+1,1} \|J_\varepsilon f\|_{m+1,1}^{m-1} \left\| (J_\varepsilon f)(x-\varepsilon) \right\|_{m+1,1} \\ &\leqslant \varepsilon (n-m)(m-1)(1+\varepsilon)^2 \|f\|_{m+1,1}^{m+1}, \\ \|\tilde{E}_2\|_{1,1} &\leqslant \frac{\varepsilon^{m-1}}{(m-1)!} (n-m)! \frac{2^{n-2m}}{(n-2m)!} (1+\varepsilon)^2 \|f\|_m^m \|f\|_{m,1}^m, \\ \|\tilde{E}_3\|_{1,1} &\leqslant \varepsilon (n-m)(1+\varepsilon)^2 \|f\|_{m+1,1}^{m+1}. \end{aligned}$$

Therefore $P_\varepsilon g_{m,n,\varepsilon,f} \to f^m$ in $L^{1,1}$. Lemma 6 implies that the normalized densities also converge in $L^{1,1}$, so Lemma 7 applies and the standard deviations converge. □

Note that the convergence of the standard deviation required a special type of convergence of the density functions and does not imply convergence in mean or in mean square of the underlying random variables. We have only discussed the convergence of the densities because each of the random variables, $T_{m,n,\varepsilon,f}$ is on a different probability space. Since the outputs are conditioned on the target cell firing, the sample space of $T_{m,n,\varepsilon,f}$ is the set of input firing times which will elicit a response. If we think of the input firing times as a vector in R^n, then the sample space is the subset of R^n with at least m of the entries within ε of each other. Therefore, although we have given conditions for several types of convergence of the densities, we have not claimed any type of convergence of the random variables. In fact, it only makes sense to talk about the convergence of the random variables in distribution which follows easily from Theorem 1.

5. Discussion

The limiting behavior for ε small is important because it gives a simple expression for the output density $g_{m,n,\varepsilon,f}$ in terms of the input density f. However, as $\varepsilon \to 0$ the probability that the target cell will fire also goes to 0. This means that the actual neural system can not operate at $\varepsilon = 0$. If ε is small, however, we can estimate $g_{m,n,\varepsilon,f}$ and $\sigma_{m,n,\varepsilon,f}$ by $g_{m,n,0,f}$ and $\sigma_{m,n,0,f}$ and the error can be bounded using the explicit bounds given. We can also compute the first asymptotic correction in ε for both the density $g_{m,n,\varepsilon,f}$ and its standard deviation $\sigma_{m,n,\varepsilon,f}$ [11].

As an example, we will compare g_ε and its small ε limit in the special case $n = 3$, $m = 2$ and the input density f is exponential. In this case we can compute the density g_ε and the standard deviation σ_ε explicitly. Using Eq. (8), we have

$$P_\varepsilon g_\varepsilon = \begin{cases} 0, & \text{for } x < 0, \\ \frac{6}{\varepsilon} e^{-2x}(1 - e^{-x}), & 0 \leqslant x < \varepsilon, \\ \frac{6}{\varepsilon}\left(\frac{1}{2} e^{-\varepsilon} e^{-x} - e^{-2x} + e^{-3x}\left(\frac{3}{2} e^{\varepsilon} - 1\right)\right), & \varepsilon \leqslant x < 2\varepsilon, \\ \frac{6}{\varepsilon}\left(e^{-2x}(e^{\varepsilon} - 1) + e^{-3x}\left(\frac{3}{2} e^{\varepsilon} - 1 - \frac{1}{2} e^{3\varepsilon}\right)\right), & x \geqslant 2\varepsilon. \end{cases}$$

Integrating this expression gives the value of P_ε,

$$P_\varepsilon = \frac{e^{3\varepsilon} - 1}{\varepsilon e^{3\varepsilon}}.$$

We can divide both sides of the equation by P_ε to get an explicit formula for the density $g_\varepsilon(x)$. Note that this value of P_ε corresponds to a probability of success of $(1 - e^{-3\varepsilon})$. Figure 2 shows this density for several values of ε. One can see that as ε decreases, the densities approach the limiting density.

We can further compute the mean and standard deviation in the standard way

$$\mu_\varepsilon = \frac{1}{(e^{3\varepsilon} - 1)}\left[\frac{5}{6} e^{3\varepsilon} + e^{\varepsilon} - \frac{11}{6} - 2\varepsilon\right]$$

and

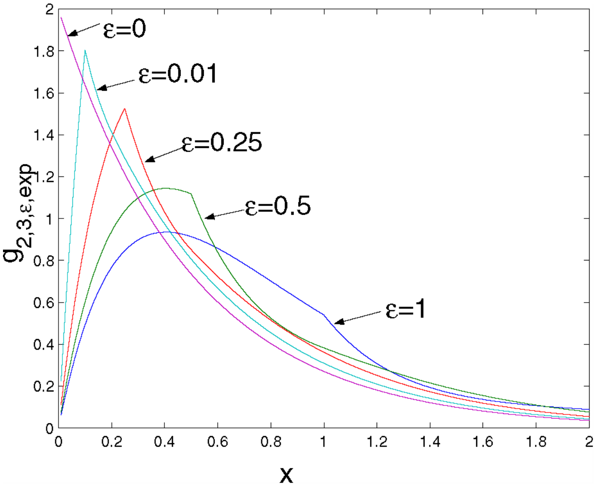

Fig. 2. The density g_ε for several values of ε. One can see the densities approaching the limiting density $2e^{-2x}$ (labeled as $\varepsilon = 0$).

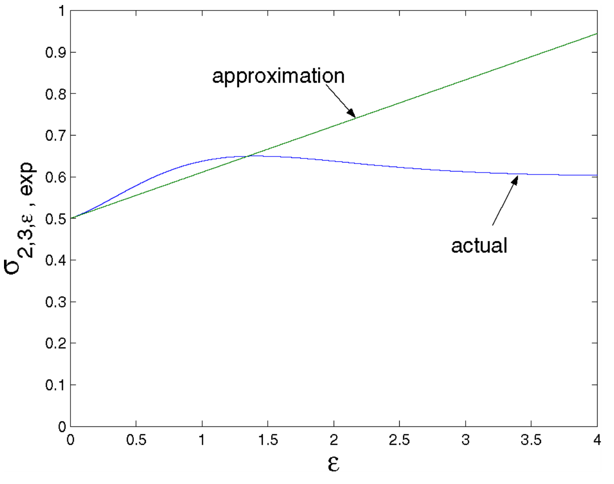

Fig. 3. The standard deviation σ_ε as a function of ε. The curve labeled "approximation" is the first asymptotic correction to the limit.

$$\sigma_\varepsilon^2 = \frac{1}{(e^{3\varepsilon}-1)^2}\left[\frac{13}{36}e^{6\varepsilon} + e^{4\varepsilon}(2+2\varepsilon) + e^{3\varepsilon}\left(-\frac{49}{18} - 4\varepsilon - 4\varepsilon^2\right) - e^{2\varepsilon} + 2\varepsilon e^{\varepsilon} + \frac{49}{36}\right]. \tag{21}$$

Taking a square root gives the standard deviation σ_ε. One can see that although explicit calculation of the density and standard deviation are possible in this simple case, it is still rather tedious. This is one reason that understanding the limits is important for cases where computing the density is either impossible or impractical. Figure 3 shows the standard deviation as a function of ε. In addition we have included the first asymptotic correction

$$\sigma_\varepsilon = \frac{1}{2} + \frac{1}{9}\varepsilon + O(\varepsilon^2),$$

which can be computed either from the formula given in [11] or from Eq. (21).

In [14], four example input densities were used: uniform, normal, exponential and hat (an upside down v with the peak at zero). All four of these densities satisfy the conditions for convergence in L^1 and for convergence of their standard deviations as $\varepsilon \to 0$. In addition, the four densities can all be made left-continuous to give pointwise convergence. However, only the normal and the hat are uniformly continuous and so only their densities converge uniformly.

Much work remains to be done in the mathematical question formulated in the introduction. We would like to prove theorems about the qualitative behavior of $\sigma_{m,n,\varepsilon,f}$ as a function of m and ε. For biological applications (in which n and m are often large), it would be useful to explore the limit $n \to \infty$, $m \to \infty$, with the ratio m/n fixed. Finally, it is also important to consider independent but non-identically distributed inputs.

Acknowledgments

This research is a portion of the author's PhD thesis at Duke University under the direction of Professor Michael Reed. The author thanks Dr. Reed for his support, guidance, and expertise. The author also thanks Kathy Bihari who prepared the figure.

References

[1] M. Abeles, Corticonics: Neural Circuits of the Cerebral Cortex, Cambridge Univ. Press, Cambridge, UK, 1991.
[2] H. Agmon-Snir, C.E. Carr, J. Rinzel, The role of dendrites in auditory coincidence detection, Nature 393 (1998) 268–272.
[3] B. Arnold, N. Balakrishnan, H. Nagaraja, A First Course in Order Statistics, Wiley Ser. Probab. Math. Statist., 1992.
[4] H.S. Colburn, Y. Han, C.P. Culotta, Coincidence model of MSO responses, Hear. Res. 49 (1990) 335–346.
[5] E. Covey, J. Casseday, The monaural nuclei of the lateral lemniscus of an echolocating bat: Parallel pathways for analyzing temporal features of sound, J. Neurosci. 11 (1991) 3456–3470.
[6] G. Folland, Introduction to Partial Differential Equations, Princeton Univ. Press, Princeton, NJ, 1995.
[7] D.R.F. Irvine, The Auditory Brainstem, Springer-Verlag, New York, 1986.
[8] R. Ivry, Cerebellar timing systems, Internat. Rev. Neurobiol. 41 (1997) 555–573.
[9] P. Marsalek, C. Koch, J. Maunsell, On the relationship between synaptic input and spike output jitter in individual neurons, Proc. Natl. Acad. Sci. USA 94 (1997) 735–740.
[10] M.S. Matell, W.H. Meck, Neurophysiological mechanisms of interval timing behavior, BioEssays 22 (2000) 94–103.
[11] C. Mitchell, Mathematical properties of time-windowing in neural systems, thesis, Duke University, 2003.
[12] D. Oertel, R. Bal, S. Gardner, P. Smith, P. Joris, Detection of synchrony in the activity of auditory nerve fibers by octopus cells of the mammalian cochlear nucleus, Proc. Natl. Acad. Sci. USA 97 (2000) 11773–11779.
[13] D. Oertel, The role of timing in the brainstem nuclei of vertebrates, Annu. Rev. Physiol. 61 (1999) 497–519.
[14] M. Reed, J. Blum, C. Mitchell, Precision of neural timing: effects of convergence and time-windowing, J. Comput. Neurosci., in press.
[15] M. Reed, B. Simon, Functional Analysis, Academic Press, San Diego, 1980.
[16] M. Reed, B. Simon, Fourier Analysis, Self-Adjointness, Academic Press, San Diego, 1975.