

14

Spike Train Analysis

14.1 INTRODUCTION

In the world of neural signals, spike trains take a special position. Spike trains play a crucial role in communication between cells in the nervous system. One might argue that knowing all communication in a system is the same, or comes close to being the same, as knowing a system. We must, however, realize that this is an optimistic view, requiring integration of function across scales similar to (but more complex than) reconstructing the application that is running on a computer (e.g., a game or a word processor) from the signals on all its buses.

While time series of action potentials can be measured intracellularly, they are often recorded extracellularly and are then referred to as spike trains. The intracellular recordings show all aspects of the membrane potential fluctuations, whereas the extracellular spike trains mainly reflect the timing of the occurrence of an action potential. Figure 14.1 shows how an action potential in a nerve fiber acts as a generator for extracellular current. This current can be measured by a biological amplifier. Depending on the positions of the recording electrodes relative to the nerve fiber, such recordings reflect the ‘vertical’ or ‘horizontal’ currents; in Fig. 14.1 it can be seen that these current components result in a tri- or biphasic wave (spike) for each action potential. Although the spikes in spike trains have similar shapes, they are often slightly different. These differences in wave morphology are due to differences in relative position and impedance between electrodes and different neurons; spikes originating from different cells differ in amplitude and waveform.

14.1.1 Deterministic versus Probabilistic Approach

A spike train may be considered as a list of the times t_i where spikes have occurred. If one considers the Hodgkin and Huxley equations (Hodgkin and Huxley, 1952) as the underlying principle for the generation of action

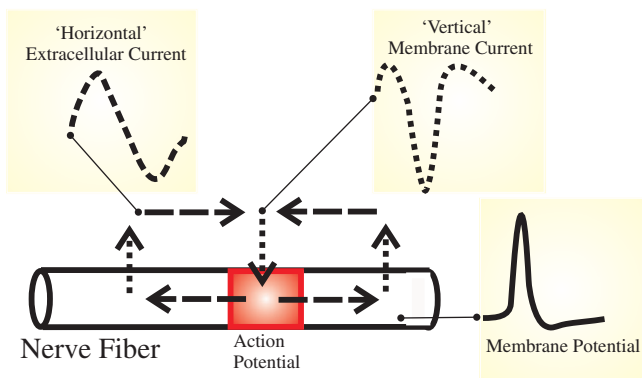


Figure 14.1 Schematic representation of intracellular, membrane, and extracellular currents associated with an action potential.

potentials, one must come to the conclusion that generation of a spike train $\{t_i\}$ is a fundamentally deterministic process. In other words, the response to a given stimulus is reproducible and fully determined by the underlying equations. However, all experimental neurophysiologists know that neural responses to the same stimuli s in repeated trials are seldom completely identical. To deal with this variability, Rieke and coworkers (1999) have successfully applied a probabilistic approach to the analysis of spike train data. They relate response to a stimulus in a probabilistic fashion with $P(\{t_i\} | s)$ denoting the probability of observing spike train $\{t_i\}$ given that stimulus s occurred. A unique aspect of their approach is that they not only consider the response but also the stimulus to be drawn from a probability density function. Although this is a somewhat unusual experimental approach where the stimulus is determined by the investigator, with a little bit of imagination it is easy to see that this is analogous to what the brain must do to interpret incoming spike trains and link these to external stimuli. Looking at neural action potential activity from a probabilistic perspective allows the use of Bayes's rule to link the probability $P(\{t_i\} | s)$ of observing a response $\{t_i\}$ to a given stimulus s with the probability that stimulus s occurred when response $\{t_i\}$ is observed, $P(s | \{t_i\})$ (Appendix 14.1).

An example of the latter approach where one attempts to determine the stimulus based on a recorded spike train is shown in Figure 14.2. In this example, the stimulus signal occurring before each spike is averaged to find the underlying signal evoking the spike event. The assumption here is that the external stimulus evoking the spike is masked by a random (noise) component that can be reduced by averaging (Chapter 4). The more conventional approach where one determines spike activity evoked by a given stimulus is shown in Figure 14.3.

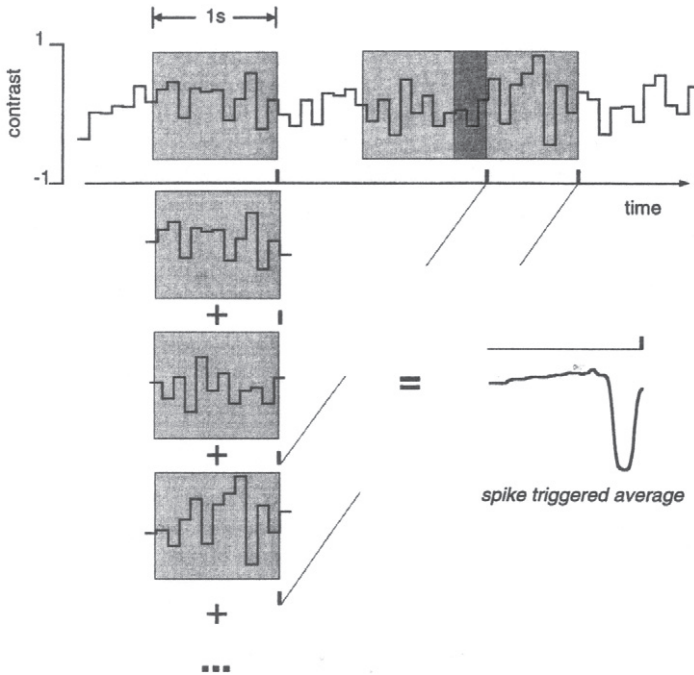


Figure 14.2 Interpretation of a spike train by averaging a prespike window of the stimulus. Top: trace stimulus; second trace shows a spike train with three spikes. The gray boxes represent the prespike windows that can be averaged to estimate a spike-triggered average. (From Rieke et al., 1999.)

14.1.2 The δ Function

In terms of signals, one may think of a spike in a spike train as an all-or-nothing process. In a general sense, a train of action potentials is a series of events occurring in time, at any given time, an event is either absent (off) or present (on). In many papers in which analyses of spike trains play a role, the activity is therefore presented in so-called raster plots (e.g., the top panel in Fig. 14.3): a time axis with each spike represented by a dot (or a short vertical line) on this axis. Implicitly one has now reduced the action potential to an event on a time line, with an event duration of zero (the dot/line). Interestingly, this approach can also be used to derive a formal representation of a spike train. Considering an epoch on the time line with an interval of size 1 and located between $-\frac{1}{2}$ and $\frac{1}{2}$, we can define a spike count function f_s for this epoch such that

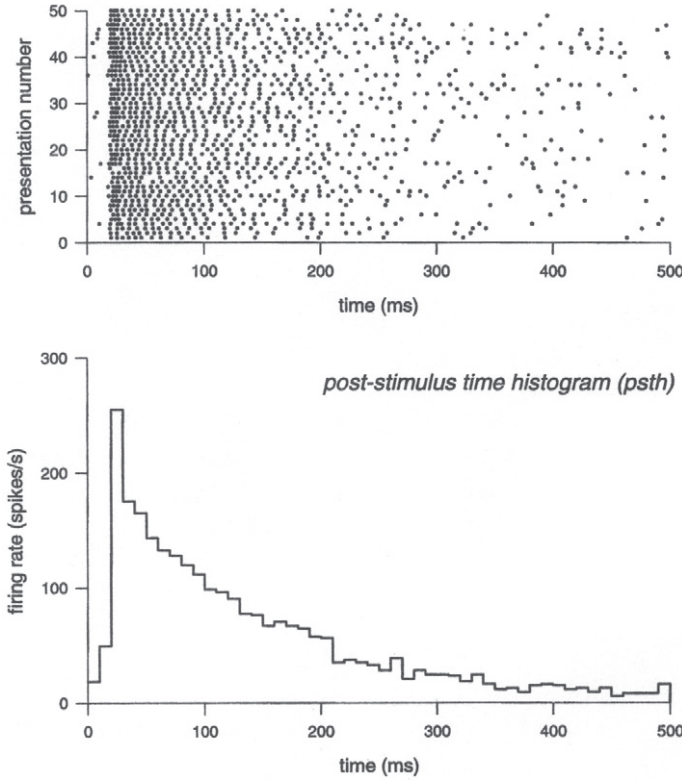


Figure 14.3 Top: Raster plots showing spikes in subsequent responses to the same stimulus. Each row is a single response plotted against time. Bottom: The raster plot data are used to plot the average spike count $\langle N \rangle$ in 10-ms bins. This is the so-called post-stimulus time histogram. (From Rieke et al., 1999.)

$$f_s[\tau] = 1 \quad \text{if} \quad -\frac{1}{2} \leq \tau \leq \frac{1}{2} \quad (14.1)$$

$$f_s[\tau] = 0 \quad \text{otherwise}$$

We can use this function to evaluate an epoch Δ around a particular time t_i by evaluating $f_s[(t - t_i)/\Delta]$. If there is a spike at t_i , the function generates a 1, so if we evaluate the sum of this function for a spike train including all times where a spike is found, we increment by 1 for each spike and obtain a spike count N :

$$N = \sum_i f_s[(t - t_i)/\Delta] \quad (14.2)$$

In real spike trains the neural response is typically variable and usually characterized by the average of a series of responses to an identical stimulus.

In the scenario shown in Figure 14.3, the spike rate in each bin can be calculated as $\langle N \rangle / \Delta$, where $\langle N \rangle$ is the average count over the trials, just as in the bottom panel in Figure 14.3. The instantaneous rate $r(t)$ can be found by letting $\Delta \rightarrow 0$:

$$r(t) = \lim_{\Delta \rightarrow 0} \frac{\langle N \rangle}{\Delta} = \left\langle \lim_{\Delta \rightarrow 0} \frac{1}{\Delta} \sum_i f_s[(t - t_i)/\Delta] \right\rangle = \left\langle \sum_i \delta(t - t_i) \right\rangle \quad (14.3)$$

In Equation (14.3), $\langle \cdot \rangle$ indicates the (vertical) average over each bin in the subsequent trials. The limiting case as $\Delta \rightarrow 0$ allows one to express the instantaneous rate as the average of the Dirac delta functions that sift out the spike timing in the responses (Chapter 2).

Notes:

1. The division of f_s , which according to Equation (14.1) has an amplitude of 1, by a factor Δ (dimension of time) is explicitly included in the definition in Equation (14.3). Therefore, the expression for rate $r(t)$ formally has the dimension s^{-1} . In the introduction of the Dirac in Chapter 2 (see Fig. 2.4), δ did not have the s^{-1} dimension because we included a dimensionless $1/\tau$ amplitude factor in the definition.
2. Because spikes actually have a finite duration, the derivation of the Dirac by letting $\Delta \rightarrow 0$ is not strictly appropriate for this application; in practice, one tries to determine Δ so that (for a single trace) the bin contains either one or no spikes. In spite of this difficulty, the δ function is often used to formally represent a spike in a spike train because it allows development of mathematical expressions for spike train analysis such as convolution, correlation, and so forth.

14.2 POISSON PROCESSES AND POISSON DISTRIBUTIONS

An important statistical model for understanding spike trains is the so-called Poisson process, which was explored most fruitfully in the context of a branch of statistics called renewal theory (for a great introduction in this field, see Cox, 1962). The major challenge in renewal theory is to understand component failure and its associated statistics. In this sense

there is a similarity between the component failure events and the occurrences of spikes in a spike train. The simplest model of the occurrence of an event is to assume a *constant probability ρ of a component failing, given that it has not failed yet*. Imagine you are managing the lightbulbs in a building where all lights are always on (broken bulbs are replaced immediately). We may consider ρ to be the probability that a functional bulb will fail. In this analogy, the process leading to a spike event of a single unit can be compared to observing a new bulb failing in a single light fixture. Similarly, the process of resetting the membrane potential after the spike is analogous to replacing a bulb that has failed. Because a bulb cannot fail twice (a broken bulb stays broken), the conditional probability (also called the age-specific failure rate) described by PDF $f(t)$ can be interpreted as the product of ρ and the survival function $\mathcal{F}(t) = 1 - F(t)$, where $F(t)$ is the cumulative distribution function (Chapter 3). Here we show that the PDF for the time of occurrence of a failure given by $f(t) = \rho e^{-\rho t}$ (with $f(t) = 0$ for $t < 0$) satisfies the previous condition:

$$\begin{aligned} \text{Survival } \mathcal{F}(t) &= 1 - F(t) = \int_t^{\infty} f(x) dx = [-e^{-\rho t}]_t^{\infty} = 0 - [-e^{-\rho t}] = e^{-\rho t} \\ d\mathcal{F}(t)/dt &= \frac{de^{-\rho t}}{dt} = -\rho e^{-\rho t} \end{aligned} \quad (14.4)$$

$$\text{also } d\mathcal{F}(t)/dt = d[1 - F(t)]/dt = -dF(t)/dt = -f(t) \quad (14.5)$$

Combining Equations (14.4) and (14.5), we get $f(t) = \rho e^{-\rho t}$ and $\mathcal{F}(t) = e^{-\rho t}$, which is consistent with our initial assumption in that this PDF embodies a constant failure (event) probability for a component that has not previously failed. This process satisfying $f(t) = \rho e^{-\rho t}$ is the so-called Poisson process, and because this process does not have a specific aging component (i.e., given the absence of previous failure, there is a constant probability that a failure will occur), it can be classified as memoryless. Graphs associated with the Poisson process are depicted in Figure 14.4. One important statistical feature of the PDF of the *Poisson process is that it is characterized by an equal mean and standard deviation* (Appendix 14.2). Accordingly, in spike trains this property can be evaluated by calculating mean and standard deviation of the interspike intervals.

Using the approach in Equations (14.4) and (14.5) is equivalent to considering the spike train in continuous time. In Section 14.1.2, we considered the spike train in discrete time (i.e., as a set of events in a binned trace). Some of the bins will contain spikes, others will be empty. In the following we consider an epoch of stationary spike activity in n bins of

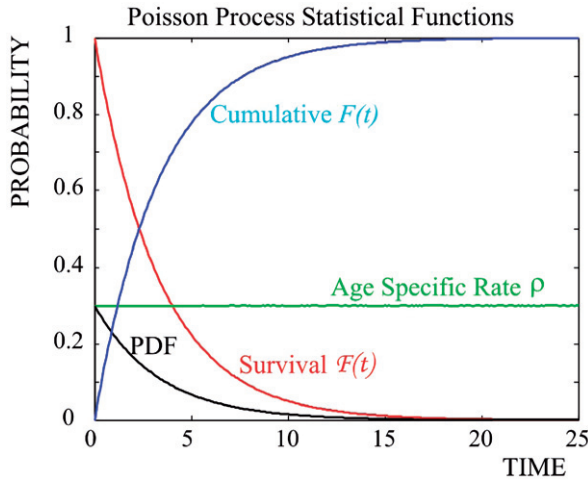


Figure 14.4 Overview of statistical functions associated with the Poisson process.

duration Δt . Assuming that the bins are about the duration of a single action potential, we might get a sequence that looks as follows:

0 0 1 0 1 1 0 0 0 1 0 0

In this epoch, we count four spikes and eight empty bins. From a distribution point of view this can be described by the binomial distribution, where the probability of a hit (event) occurring is p and the probability of a failure is $q = 1 - p$. Here we have $p = \rho\Delta t + 0(\Delta t)$, with $\rho\Delta t$ representing the probability that one event occurs in Δt and $0(\Delta t)$ for the probability that more than one event occurs in Δt . By selecting a suitably small value for Δt , we can ignore $0(\Delta t)$ because there will not be more than one spike per interval. This is a bit tricky because Δt cannot approach 0 either because of the finite duration of a spike. If the occurrence of hits and failures is truly independent, we could begin by saying that the probability of counting four spikes in the preceding sequence of 12 bins is p^4q^8 . This probability then needs to be corrected for all the other arrangements that could also lead to a total count of four spikes, such as,

1 1 1 1 0 0 0 0 0 0 0 0
 0 1 1 1 1 0 0 0 0 0 0 0
 0 0 1 1 1 1 0 0 0 0 0 0
 0 0 0 1 1 1 1 0 0 0 0 0
 ... and so on

In this example, for $n = 12$ trials we have $i = 4$ hits and $n - i = 8$ failures. The four hits can be arranged in $\frac{n!}{(n-i)!i!} = \frac{12!}{8!4!} = 12 \times 11 \times 10 \times 9 / 4 \times 3 \times 2 \times 1 = 495$ different ways over

the 12 bins (an alternative more compact notation for $\frac{n!}{(n-i)!i!}$ is $\binom{n}{i}$; note that this is *not* a fraction). This reasoning underlies the derivation of the binomial distribution, which links the probability p of a single hit occurring in one trial to the probability $P(i)$ of encountering i hits in n trials as

$$P(i) = \frac{n!}{(n-i)!i!} p^i (1-p)^{n-i} \quad (14.6)$$

In real spike trains, the intervals can be considered very small (a few ms) compared to the complete time series length (usually 1 to several seconds), in which case, with typical spike rates, most intervals do not contain spikes. We can therefore consider the preceding probability function for very large n and small values of p . Suppose that the average number of hits in n observations is λ ; then we could define p as λ/n . Using this to rewrite Equation (14.6),

$$\begin{aligned} P(i) &= \frac{n!}{(n-i)!i!} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i} \\ &= \frac{n(n-1)(n-2)\dots(n-i+1)}{i!} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i} \\ &= \frac{n(n-1)(n-2)\dots(n-i+1)}{\left(1 - \frac{\lambda}{n}\right)^i i!} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^n \\ &= \frac{n(n-1)(n-2)\dots(n-i+1)}{n^i \left(1 - \frac{\lambda}{n}\right)^i i!} \lambda^i \left(1 - \frac{\lambda}{n}\right)^n \\ &= \frac{1(1-1/n)(1-2/n)\dots(1-(i-1)/n)}{\left(1 - \frac{\lambda}{n}\right)^i i!} \lambda^i \left(1 - \frac{\lambda}{n}\right)^n \end{aligned} \quad (14.7)$$

For large n , all terms in the first part that contain a division by n will approach 0:

$$\frac{1(1-1/n)(1-2/n)\dots(1-(i-1)/n)}{\left(1 - \frac{\lambda}{n}\right)^i i!} \lambda^i \rightarrow \frac{\lambda^i}{i!}$$

and the second term is a power series: $\left(1 - \frac{\lambda}{n}\right)^n \rightarrow e^{-\lambda}$

Combining these results, we obtain the equation for the Poisson distribution:

$$P(i) = \frac{\lambda^i}{i!} e^{-\lambda} \quad \text{or} \quad P(i) = \frac{(\rho t)^i}{i!} e^{-\rho t} \quad (14.8)$$

In the second version of Equation (14.8), we used $\lambda = n\Delta t\rho = t\rho$. Equation (14.8) represents a probability density function (PDF) in which the sum of probabilities of all possible outcomes $\sum_{i=1}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda}$ equals 1. This can be seen when substituting the series $\sum_{i=1}^{\infty} \frac{\lambda^i}{i!}$ by exponential e^{λ} :

$$e^{\lambda} = 1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots$$

Comparing Equation (14.8) with the *Poisson process* $f(t) = \rho e^{-\rho t}$, we showed that the number of events in a fixed interval (e.g. spike counts in a $\frac{1}{2}$ s epoch) satisfy a *Poisson distribution*. Further we can show that the *mean and variance of the Poisson distribution are both equal to λ* . The mean value can be obtained from $E\{i\} = \sum_{i=0}^{\infty} i \frac{\lambda^i}{i!} e^{-\lambda}$. Using $i\lambda^i = \lambda \frac{\partial}{\partial \lambda} \lambda^i$, taking the exponential and other non- i -related factors out of the summation, using the preceding power series for e^{λ} , and taking into account that $\partial e^{\lambda} / \partial \lambda = e^{\lambda}$, we get

$$E\{i\} = \sum_{i=0}^{\infty} i \frac{\lambda^i}{i!} e^{-\lambda} = e^{-\lambda} \sum_{i=0}^{\infty} \frac{1}{i!} \lambda \frac{\partial}{\partial \lambda} \lambda^i = e^{-\lambda} \lambda \frac{\partial}{\partial \lambda} \sum_{i=0}^{\infty} \frac{1}{i!} \lambda^i = e^{-\lambda} \lambda \frac{\partial}{\partial \lambda} e^{\lambda} = e^{-\lambda} \lambda e^{\lambda} = \lambda$$

In a similar way, by using the second derivative, one can show that the variance is equal to the mean $E\{i^2\} - E\{i\}^2 = \lambda$. For a Poisson distribution, the so-called *Fano factor*, which is the ratio between variance/mean, is 1. The Fano factor is indeed ~ 1 for short spike trains (order of second(s)), but for larger epochs the Fano factor usually becomes >1 .

In van Drongelen et al. (1978), the Poisson process was used to study sensitivity effects of convergence of receptor cells in the olfactory system. In this system, ~ 1000 peripheral sensory neurons project onto a single mitral cell in the olfactory bulb. At threshold levels of stimulation, the sensory neurons show probabilistic firing patterns, so by convention a particular unit's threshold is arbitrarily defined as the stimulus level that evokes a response in 50% of the presentations of that stimulus. Because

the mitral cells receive input from a thousand neurons at the same time, we might predict that their threshold levels occur at significantly lower concentrations of odorant as compared to the peripheral threshold (i.e., the signal is amplified by the convergence of sensory neurons).

We can estimate the amplification effect if we analyze the preceding system in a simplified model. Assume absence of spontaneous activity and a sensory cell's firing probability upon stimulation equal to ρ . Under these assumptions, we can determine the probability that a stimulated sensory neuron fires at least once in observation interval T as

$$1 - (\text{probability that the sensory neuron does not fire}) = 1 - e^{-\rho T}$$

Consider a single mitral cell observing 1000 of these sensory cells. Further assume that the mitral cell is rather sensitive so that it spikes upon a spike in any of its connected sensory cell population; therefore, the probability that the mitral cell fires in the same interval T is

$$1 - (\text{probability that none of the 1000 sensory neurons fires})$$

If we consider the activity of the sensory neurons independent, which is not a bad assumption at low levels of odorant diffusing in the olfactory mucosa, we can express the probability that none of the 1000 sensors fires as the product of the individual probabilities $(e^{-\rho T})^{1000}$; therefore we can predict that the probability that the mitral cell fires is $1 - e^{-1000\rho T}$, much higher than the probability that a single sensory cell fires $1 - e^{-\rho T}$. Indeed, this prediction was experimentally confirmed (Duchamp-Viret et al., 1989).

14.3 ENTROPY AND INFORMATION

For most of us, "information" is a familiar concept because sending and receiving messages (i.e., information exchange) is an important part of our daily life. However, for a study of information transfer in the nervous system, we need a formal definition and preferably a metric quantifying information content. Shannon's communication system provides such a framework for analyzing the transmission of a message from a source to a destination (Shannon and Weaver, 1949). Considering a set of messages to be transmitted, Shannon's idea was to use the so-called entropy of this set to quantify its potential information content. It is important to realize that this entropy measure differs from what may be considered the everyday sense of information content. The entropy concept does not capture the information in a single message but merely quantifies the potential

information of the ensemble of messages. At first sight this may seem somewhat strange, but consider the example where we continuously transmit the same message. In this case, one might conclude that there is no need for information to be transmitted, or one might even argue that (from a data transmission standpoint) no information is sent at all. Similarly, if the received message is always the same, there is no reason to receive the message and one might state that in this case there is no information either. Therefore, when using Shannon's entropy concept, the context of the message in its ensemble and variability is critical for quantifying information content. As mentioned in Section 14.1.1, in the analysis of spike trains it is not uncommon to consider both the stimulus and the neural response to be drawn from a probability density function (i.e., Rieke et al., 1999), thus making it possible that variability is associated with information. This idea will be explored further in the following paragraphs.

Let X be a message consisting of two statistically independent random variables X_1 and X_2 . The probability of observing X_1 and X_2 is the product of the individual probabilities: $p[X_1] \times p[X_2]$. In other words, the uncertainty of X or the information conveyed by revealing that message X has taken on the values X_1 and X_2 depends on the probability density functions of X_1 and X_2 . Let's use the functional S to denote the entropy associated with the observations X_1 and X_2 , such that the information gained from observation X_1 is $S\{p[X_1]\}$ and the information associated with observation X_2 is $S\{p[X_2]\}$. If we now receive the information associated with both X_1 and X_2 , the information gained from each should add to represent the combined information:

$$S\{p[X]\} = S\{p[X_1, X_2]\} = S\{p[X_1]\} + S\{p[X_2]\} \quad (14.9)$$

Equation (14.9) shows that the product $p[X_1] \times p[X_2]$ is converted into a sum of entropies. Therefore the entropy behaves as a logarithm of the distribution.

Although this is not an in depth description of Shannon's approach, the intuitive notion is that the entropy of a system is proportional with the logarithm of the number of its possible states. Let X be a discrete variable, representing a spike train response to a stimulus, taking a finite number of possible values x_1, x_2, \dots, x_N . First, if we assume that each state is equally likely to occur (i.e., $p = 1/N$), the entropy is proportional to $\log(N)$ or $-\log(1/N)$ (note the $-$ sign). If the states are not equally probable but occur with probabilities p_1, p_2, \dots, p_N (such that $\sum_{i=1}^N p_i = 1$), the entropy is proportional

to $-\sum_{i=1}^N p_i \log p_i$ (note the $-$ sign). To obtain an entropy measure that can

be expressed in bits, it is common practice to use a base 2 logarithm:

$$S = -\sum_{i=1}^N p_i \log_2 p_i \text{ bits} \quad (14.10)$$

This discussion of the entropy-information quantification approach makes it clear that this information content can only be established if the PDF associated with a message is known.

In spite of the quantitative nature of Shannon's approach, it leaves quite a bit open to interpretation when applied to spike trains, a reflection of our ignorance of neural coding. For instance, if we record a particular response with N_1 spikes, do we consider the number N_1 as a number drawn from a PDF or not? If so, what PDF do we assume? Do we assume that the timing of each spike is important, or is it just the number N_1 that captures the essence of the message? It is difficult to answer these questions without knowing how the nervous system processes this particular message. It is, however, critical to think about these issues since the answers to these questions directly determine the assumed PDF from which our observation is drawn and therefore determines the entropy. Let's analyze an example where we observe N_1 spikes in a trace with N bins, similar to the examples we discussed in the previous section. There are several possible approaches:

1. We have N bins, so we can support $N + 1$ spike counts (in case of 3 bins we have 0, 1, 2, 3 as possible spike count observations, Fig. 14.5B). Assuming that each response is equally likely, the resulting entropy is $\log_2(N + 1)$ bits. This example is mentioned to be complete but is somewhat silly because its result depends on the number of bins (which we set arbitrarily). In addition, it is unlikely that all responses (0 – N spikes) are equal. On the other hand, there may be scenarios where, for a certain choice of interval, the number of spikes is the important message.
2. We have N_1 spikes, assuming that timing is important, and we have $N!/(N_1!N_0!)$ possible arrangements over the N bins, with $N_0 = N - N_1$. If we now assume that each arrangement is equally likely, we have an entropy of $\log_2[N!/(N_1!N_0!)]$. This example is a little less silly, but it still does not account for the fact that the observation of N_1 spikes varies (i.e., we consider the observation of N_1 spikes as deterministic). Within the given number of spikes, we allow different distributions over the available bins (i.e., we do consider the effect of timing). For instance, if we observe two spikes in a trace with three bins ($N_1 = 2$), we have $3!/(2!1!) = 3$ possibilities. Assuming these all equally likely, we have $3 \times \frac{1}{3} \log_2(3) \approx 1.6$ bits of total entropy (Fig. 14.5C).
3. This naturally leads to the third example where the observed spike train is drawn from a PDF such as a Poisson or a binomial distribution. In this case, the number of spikes is not fixed, and unlike in the pre-

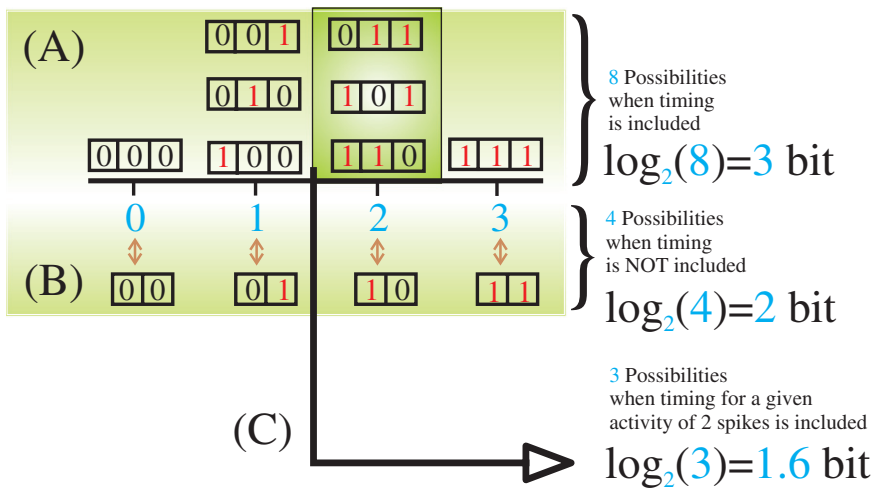


Figure 14.5 Simplified example of a spike train of three bins. The bin size is selected such that it can either contain one or no spikes. All eight possibilities having zero, one, two, or three spikes in the set of three bins are shown in (A). This results in an entropy value of 3 bits. This comes as no surprise since we have three bins that could contain zero or one. (B) If we consider the number of spikes, without paying attention to the timing, there are only four possibilities resulting in a 2-bit value for the entropy. Decimal values are shown in blue; binary values are black (0) or red (1). When taking the activity level as a given, for instance, two spikes in a three bin observation span, the timing of those spikes determines the entropy. In example (C) we have three possible arrangements, leading to a value of 1.6 bits.

vious example where we observed two spikes, we also allow other observations (e.g., the number of spikes = 0, 1, 2, or 3 in Fig. 14.5A). In this example, the number of spikes and their timing are both important.

In all of these examples, it is clear that bin size affects the outcome of the entropy measure. Therefore one needs an approach to obtain an objective and a reasonable bin size for the particular problem at hand. A commonly used approach is to find a distribution or bin size that maximizes the entropy, considering at minimum a bin equal to the duration of a spike plus the absolute refractory period. The guiding principle here is that this method determines the upper bound of the information of a measured spike train.

A few numerical examples of how one could calculate the entropy of a short spike train with three bins are shown in Figure 14.5. We stay with the example of short bins where each may or may not contain a single spike (i.e., each bin can only contain a 1 or a 0). Therefore, if we select the

first scenario described earlier, the three bins can code for a range of numbers from 0 to 3, and $S = \log_2(4) = 2$ bits. If instead we consider the third scenario described, we assume a binomial distribution (Equation (14.6)) in which the occurrence of a 1 or 0 in each bin is equally likely ($p = 0.5$). We have the following possibilities (Fig. 14.5):

$$\begin{aligned}
 \text{Count} = 0 & \rightarrow \frac{3!}{3!0!} \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^3 = \frac{1}{8} && \text{one out of eight cases} \\
 \text{Count} = 1 & \rightarrow \frac{3!}{2!1!} \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^2 = \frac{3}{8} && \text{three out of eight cases} \\
 \text{Count} = 2 & \rightarrow \frac{3!}{1!2!} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^1 = \frac{3}{8} && \text{three out of eight cases} \\
 \text{Count} = 3 & \rightarrow \frac{3!}{0!3!} \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^0 = \frac{1}{8} && \text{one out of eight cases}
 \end{aligned}$$

Note: In the preceding example we used the definition $0! \equiv 1$.

We can see that in this case there are eight arrangements that are equally likely, using Equation (14.10) the total entropy in this case is

$$S = -\sum_{i=1}^8 p_i \log_2 p_i = -8 \times \frac{1}{8} \log_2 \frac{1}{8} = \log_2 8 = 3 \text{ bits}$$

In other words, by revealing a single, particular sequence in this scenario, we provide total number of bits \times the probability of that single scenario: $3 \times 1/8$ bits of information. These numerical examples are fairly artificial, but they illustrate the point that the context in which a message is considered is critical for the associated entropy value.

In studies of spike trains, the probability of spike occurrence is often estimated from the recorded time series. In addition, the bin size is usually selected to generate a maximum entropy value (e.g., van Drongelen et al., 2003). In this study the spike activity of single cortical neurons was associated with bursting activity of the surrounding cortical network. The network bursts were used to align the spike trains, a similar procedure as one would follow when a stimulus is used as the trigger (e.g., Fig. 14.3). The aligned spike trains were binned to allow estimation of each spike train's entropy.

An illustration of the application of these entropy calculations is shown in Figure 14.6. Four spike train trials are triggered (aligned) by a population burst; each trial is divided into 9 bins resulting in a total of $4 \times 9 = 36$ bins. We can use our spike train statistics to estimate the probability asso-

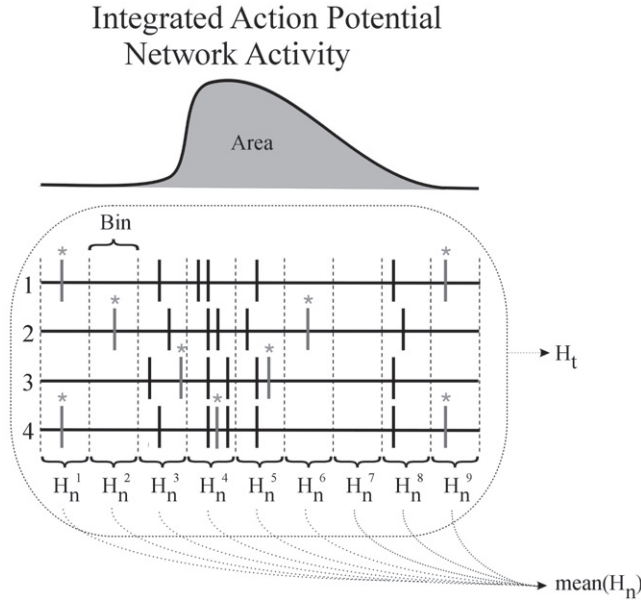


Figure 14.6 Integrated network activity-triggered average of a single neuron's spike train. In this theoretical example, the spikes marked with * are not burst related. This is only for the sake of the example; in a real measurement, the origin of the spikes is, of course, not known. Modified from van Drongelen et al., 2003.

ciated with each observation. In all of the 36 bins, we observe 14 bins with zero spikes (i.e., $p_0 = 14/36$); 16 bins with 1 spike (i.e., $p_1 = 16/36$); 5 bins with 2 spikes (i.e., $p_2 = 5/36$); and 1 bin with 3 spikes (i.e., $p_3 = 1/36$). These values can be used to estimate the total entropy H_t :

$$-(14/36) \times \log_2(14/36) - (16/36) \times \log_2(16/36) - (5/36) \times \log_2(5/36) - (1/36) \times \log_2(1/36) \approx 1.6 \text{ bits}$$

Normally one would correct the estimate for bias; here, in order to keep the example as simple as possible, we omit the correction. Now if the spiking activity was only due to the population activity, the number of spikes across the vertical bins should be identical. In other words, the variability across the vertical bins represents activity that is not associated with the burst (the trigger event) and therefore can be considered as noise. Applying this to the example in Figure 14.6 we get $-(1/2) \times \log_2(1/2) - (1/2) \times \log_2(1/2)$ for the first column and $-(2/3) \times \log_2(2/3) - (1/3) \times \log_2(1/3)$ for the second column, and so forth. Finally, the average noise entropy for all nine columns $\text{mean}(H_n)$ can be subtracted from the total

entropy in all traces (1.6 bits in this example) to estimate the information that is associated with the burst H_{burst} :

$$H_{burst} = H_t - \text{mean}(H_n) \quad (14.11)$$

14.4 THE AUTOCORRELATION FUNCTION

Due to the specific properties of the spike trains, applications of signal processing techniques such as correlation differ somewhat from the conventional approaches (Fig. 14.7).

In Chapter 8, we defined autocorrelation R_{xx} of signal x as $R_{xx}(t_1, t_2) = E\{x(t_1)x(t_2)\}$. In the case of a spike train, we do not have a continuous signal as we did in Chapter 8, but we may use the instantaneous rate function $r(t)$ as a proxy. Consequently, we can define the autocorrelation $R_{rr}(t_1, t_2)$ of r as

$$R_{rr}(t_1, t_2) = E\{r(t_1)r(t_2)\} \quad (14.12)$$

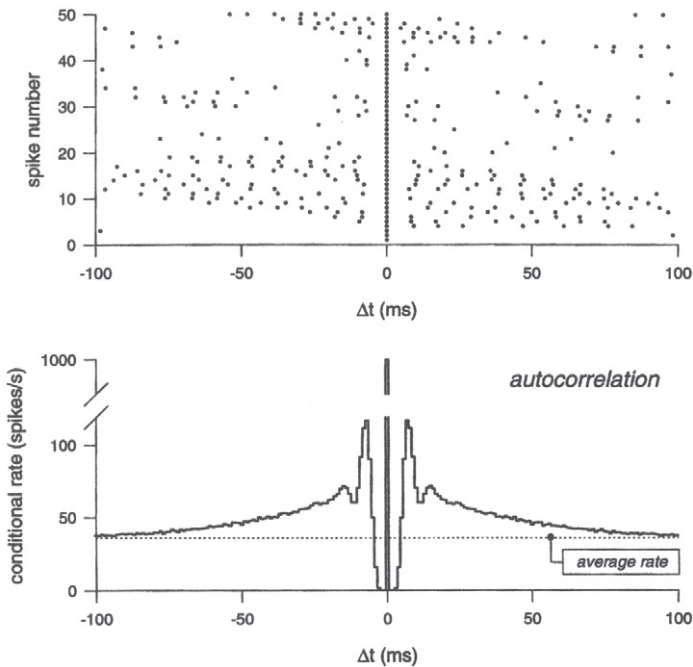


Figure 14.7 Autocorrelation of a spike train. Top: Raster plot of spike activity; each row is a plot of spike occurrence that is aligned with the middle spike in the top row. Bottom: The average of the spikes in each bin as a function of delay. (From Rieke et al., 1999.)

Assuming we may use a time average (denoted by $\langle \cdot \rangle$), we get

$$R_{rr}(t_1, t_2) = E\{r(t_1)r(t_2)\} = \langle r(t_1)r(t_2) \rangle \quad (14.13)$$

In Section 14.1.2 (Equation (14.3)), the rate function was defined as the spike average over a large number of epochs. This definition is practical because the average can be easily determined directly from experimental observations, such as in the example shown in Figure 14.3. More theoretically, one may relate the rate to the probability of the occurrence of a spike at a given time. This follows directly from the definition of the rate as the average occurrence of spikes across a large set of trials. Consequently, the two extremes for this average are that (1) there is always a spike in the observed epochs, or (2) we never observe any action potential in these epochs. In these two extreme cases, the resulting average for the rate (in terms of spiking probability) is 1 or 0, respectively, and in all other cases, the average value representing the instantaneous rate will be between 0 and 1. Note that in this example we wanted to estimate spike probability and therefore we did not divide by the bin width Δ . If we had, we would have obtained values expressed in spikes/s instead of probabilities between 0 and 1; an alternative interpretation of the instantaneous rate (not weighed by Δ) is the probability of the occurrence of an action potential. Therefore the autocorrelation of a single spike train $R_{rr}(t_1, t_2)$ is proportional with the probability of observing spikes at t_1 and t_2 :

$$P(\text{spike at } t_1 \text{ \& spike at } t_2) = \underbrace{P(\text{spike at } t_1 | \text{spike at } t_2)}_{\text{conditional rate}} \times \underbrace{P(\text{spike at } t_2)}_{-r(t_2)} \quad (14.14)$$

The conditional rate, the first term in Equation (14.14), is usually called the autocorrelation of the spike train. Just as we determined the probability of spiking after a stimulus by using a time average, it is intuitive to estimate the probability of a spike at t_1 in a similar fashion. In this case, we use a time average of traces where spike occurrence instead of stimulus occurrence is used to align the traces. An example is shown in Figure 14.7. Instead of shifting the time series by a small regular interval dt , the correlation function of the spike train is obtained from shifts that bring each subsequent occurrence of a spike to time 0 (top panel in Fig. 14.7); by following this procedure, we satisfy $P(\text{spike at } t_2) = 1$ in Equation (14.14). This process is repeated several times and then the average of the binned traces is used as an estimate of the autocorrelation function (e.g., autocorrelation in the bottom panel in Fig. 14.7). As shown in this figure, it is not uncommon to divide the outcome by the bin width in order to obtain a value in spikes/s. Also note that in the example in Figure 14.7, each spike in the train is used in the average (i.e., the process is assumed to be stationary so that the autocorrelation only depends on the difference $\tau = t_2 - t_1$).

An alternative procedure we used for calculating the autocorrelation also follows from the definition of autocorrelation discussed in Chapter 8 and the representation of the spike train as a series of unit impulses. If the spike train represents a stationary process, the underlying distributions are invariant and only the difference $\tau = t_2 - t_1$ is relevant:

$$R_{rr}(\tau) = E\{r(t)r(t+\tau)\} \quad (14.15)$$

Assuming ergodicity, we can use a time average:

$$R_{rr}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T r(t)r(t+\tau)dt \quad (14.16)$$

Using the fact that T becomes very large, we can use the definition from Equation (8.20):

$$R_{rr}(\tau) = \int_{-\infty}^{\infty} r(t)r(t+\tau)dt \quad (14.17)$$

So for each value of τ , we correlate the spike train with itself. Since a single instantiation of a spike train can be represented by a series of Diracs $\sum_{j=1}^N \delta(t-t_j)$, for each τ , this series of Diracs may be correlated with itself:

$$R_{rr}(\tau) = \int_{-\infty}^{\infty} \sum_{i=1}^N \delta(t-t_i) \sum_{j=1}^N \delta(t-t_j+\tau)dt \quad (14.18)$$

Assuming that we may interchange the integration and summation procedures,

$$R_{rr}(\tau) = \sum_{i=1}^N \sum_{j=1}^N \int_{-\infty}^{\infty} \delta(t-t_i) \delta(t-t_j+\tau)dt \quad (14.19)$$

Equation (14.19) can be simplified by using the sifting property of the δ function.

In the following we illustrate a simple example of a spike train with three spikes at times $t_1 = 0$, $t_2 = 1$, and $t_3 = 4$, and no other spike activity. Since we have two delta functions in the integral in Equation (14.19), we can sift the first with the second delta function or sift the second with the first one.

For the first spike:

$i = 1$ and $j = 1$:

$$\int_{-\infty}^{\infty} \delta(t-t_1) \delta(t-t_1+\tau)dt = \delta(t_1-t_1+\tau) = \delta(\tau)$$

$$\text{or} \quad \int_{-\infty}^{\infty} \delta(t-t_1) \delta(t-t_1+\tau) dt = \delta(t_1-\tau-t_1) = \delta(-\tau)$$

$$i = 1 \text{ and } j = 2: \quad \int_{-\infty}^{\infty} \delta(t-t_1) \delta(t-t_2+\tau) dt = \delta(t_1-t_2-\tau) = \delta(\tau-1)$$

$$\text{or} \quad \int_{-\infty}^{\infty} \delta(t-t_1) \delta(t-t_2+\tau) dt = \delta(t_2-\tau-t_1) = \delta(-\tau+1)$$

$$i = 1 \text{ and } j = 3: \quad \int_{-\infty}^{\infty} \delta(t-t_1) \delta(t-t_3+\tau) dt = \delta(t_1-t_3-\tau) = \delta(\tau-4)$$

$$\text{or} \quad \int_{-\infty}^{\infty} \delta(t-t_1) \delta(t-t_3+\tau) dt = \delta(t_3-\tau-t_1) = \delta(-\tau+4)$$

The preceding shows that the results for sifting one delta with the other (and vice versa) generate a pair of mirror shifts (i.e., τ and $-\tau$; $\tau-1$ and $-\tau+1$; $\tau-4$ and $-\tau+4$). Thus, because the autocorrelation is an even function (e.g., $R_{rr}(\tau) = R_{rr}(-\tau)$, Appendix 5.2), we only need to consider shifting in one direction.

For the second spike (in the following we omit the equations for the mirror cases),

$$i = 2 \text{ and } j = 1: \quad \int_{-\infty}^{\infty} \delta(t-t_2) \delta(t-t_1+\tau) dt = \delta(t_2-t_1+\tau) = \delta(\tau+1)$$

$$i = 2 \text{ and } j = 2: \quad \int_{-\infty}^{\infty} \delta(t-t_2) \delta(t-t_2+\tau) dt = \delta(t_2-t_2+\tau) = \delta(\tau)$$

$$i = 2 \text{ and } j = 3: \quad \int_{-\infty}^{\infty} \delta(t-t_2) \delta(t-t_3+\tau) dt = \delta(t_2-t_3+\tau) = \delta(\tau-3)$$

For the third spike:

$$i = 3 \text{ and } j = 1: \quad \int_{-\infty}^{\infty} \delta(t-t_3) \delta(t-t_1+\tau) dt = \delta(t_3-t_1+\tau) = \delta(\tau+4)$$

$$i = 3 \text{ and } j = 2: \quad \int_{-\infty}^{\infty} \delta(t-t_3) \delta(t-t_2+\tau) dt = \delta(t_3-t_2+\tau) = \delta(\tau+3)$$

$$i = 3 \text{ and } j = 3: \quad \int_{-\infty}^{\infty} \delta(t-t_3) \delta(t-t_3+\tau) dt = \delta(t_3-t_3+\tau) = \delta(\tau)$$

Summing the results we obtained for i and j from 1 to 3 (the preceding boxed equations), Equation (14.19) evaluates to

$$R_{rr}(\tau) = \sum_{i=1}^3 \sum_{j=1}^3 \int_{-\infty}^{\infty} \delta(t-t_i) \delta(t-t_j+\tau) dt = \sum_{i=1}^3 \sum_{j=1}^3 \delta(t_i-t_j+\tau) \quad (14.20)$$

$$= [\delta(\tau+4) + \delta(\tau+3) + \delta(\tau+1) + 3 \times \delta(\tau) + \delta(\tau-1) + \delta(\tau-3) + \delta(\tau-4)]$$

In Figure 14.8 we determine the autocorrelation for our example of a train of three spikes; in panel A we obtain the autocorrelation by shifting the spike train for each spike relative to one of the spikes, followed by a

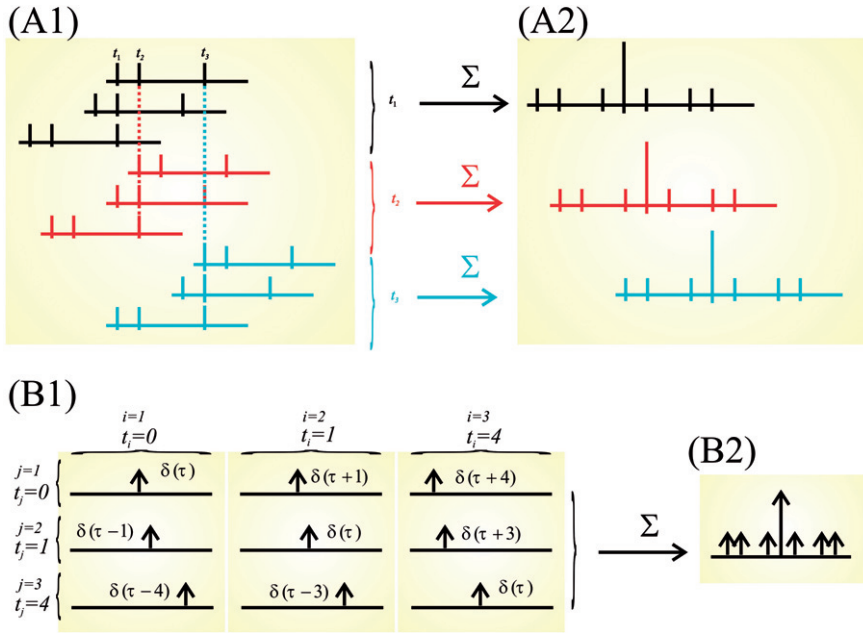


Figure 14.8 Example of a series of three spikes (A1) and the autocorrelation function (A2) based on the first spike (black), the second spike (red), and the third one (blue). The Σ symbol indicates the summation of the three traces of the associated color. In this example, there are few spikes so that each event can be indicated individually. In real spike trains, the traces are binned and the number of spikes per bin is determined because there can be many closely spaced spikes (see Fig. 14.7). In addition, it is customary to scale the ordinate to represent spikes/s or the correlation coefficient. Panels (B1) and (B2) show the correlation function based on Equations (14.19) and (14.20). Comparing the autocorrelation in (A2) and (B2), note that the different approaches generate the same result.

summation (Σ in Fig. 14.8A, representing the nonscaled average). In this example, we demonstrate this principle for each of the spikes at $t_1 = 0$, $t_2 = 1$, and $t_3 = 4$ in black, red, and blue, respectively. This procedure is similar to the one shown in Figure 14.7. In Figure 14.8B we use a different approach and demonstrate the summation underlying the result in Equation (14.20). Note that all results in Figure 14.8 are identical. In none of the examples were we worried about normalizing our result, but note that at the autocorrelation function at $\tau = 0$ always contains the summation of $N = 3$ spikes; therefore we obtain a scaled correlation equal to 1 at $\tau = 0$ by dividing the summed result by N .

14.5 CROSS-CORRELATION

Not surprisingly, cross-correlation between two spike trains follows a similar procedure to that discussed earlier for autocorrelation. The two spike trains are shifted relative to each other and the spike-triggered average represents an estimate of the cross-correlation. Another interesting application of cross-correlation is between a spike train and a continuous signal such as the stimulus $s(t)$ evoking the spike train $\{t_i\}$. Examples of such a correlation are shown in Figure 14.9. A cross-correlation between a stimulus signal $s(t)$ and a train with a single spike at time t_i considered over an interval T can be defined as

$$R(\tau)_{s(t), \{t_i\}} = \int_T s(t) \underbrace{\delta(t - t_i + \tau)}_{\delta(t - (t_i - \tau))} dt = s(t_i - \tau) \quad (14.21)$$

Unlike the autocorrelation, the cross-correlation of two signals is usually not an even function. For $\tau < 0$, the correlation of the signal at $s(t_i - \tau)$ suggests that the spike predicts the stimulus, a fairly unrealistic assumption because it violates causality. In the following we consider only positive values of τ in $s(t_i - \tau)$ indicating that we are looking from the spike time t_i backward (reverse-correlation function). If we now consider a spike train with N spikes over interval T , we obtain the reverse-correlation function as

$$R(\tau)_{s(t), \{t_i\}} = \int_T s(t) \left[\sum_{i=1}^N \delta(t - t_i + \tau) \right] dt \quad (14.22)$$

Interchange of the integration and summation gives

$$R(\tau)_{s(t), \{t_i\}} = \sum_{i=1}^N \int_T s(t) \delta(t - t_i + \tau) dt = \sum_{i=1}^N s(t_i - \tau) \quad (14.23)$$

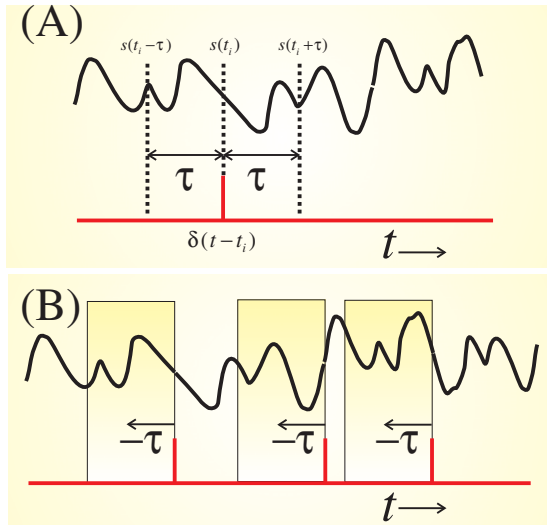


Figure 14.9 Cross-correlation between a spike train (red) and a continuous signal (black) representing the stimulus evoking the train. (A) shows the relationship between a single spike and its correlation. Because of the causal relationship between stimulus and spike, we are only interested in $s(t)$ preceding the spike, the so-called reverse-correlation function. (B) shows a spike train with three spikes and the associated preceding correlation windows. When the signal in these windows is averaged, we obtain the spike-triggered average (i.e., the estimated reverse correlation), as in the example shown in Figure 14.2.

In other words, the cross-correlation can be estimated by the sum of the signal epochs preceding each evoked spike, as shown in Figure 14.9. In the equations we did not bother with normalizing the expression for the cross-correlation. A common normalization is to divide by N so that Equation (14.23) becomes $\frac{1}{N} \sum_{i=1}^N s(t_i - \tau)$, a spike-triggered signal average of the stimulus preceding the spike event, such as in the example of Figure 14.2. The use of reverse correlation is directly related to the fact that we are considering the stimulus signal and that the spiking neuron obeys causality. Of course, this reasoning would reverse if the signal $s(t)$ represented a movement and the spike train was from a motoneuron steering this movement.

APPENDIX 14.1

Bayes's Rule

In Rieke et al. (1999), both the spike trains and stimuli are considered as events drawn from a probability density function. The following explana-

tion of the application of Bayes's rule in the field of spike train analysis is largely based on their reasoning.

Assuming that both the spikes and the stimuli are probabilistic, we can define the joint probability P of observing a specific stimulus s and a specific spike train $\{t_i\}$ as $P(\{t_i\}, s)$. Further it seems reasonable to assume that the PDFs for spike trains and stimuli are linked such that when stimulus s occurs, there is an associated probability of observing a specific spike train $\{t_i\}$; this can be defined as $P(\{t_i\}|s)$. This assumption characterizes the situation in which an experimenter provides a set of stimuli to a neuron while recording its spiking activity. If we now multiply the probability of observing $\{t_i\}$ given that s occurred with the probability that s indeed occurs $P(s)$, we obtain the probability to observe both s and $\{t_i\}$:

$$P(\{t_i\}, s) = P(\{t_i\}|s) \times P(s) \quad (\text{A14.1-1})$$

Equation (A14.1-1) relates to the observation that an experimenter studying a neural response may make. Note that in Equation (A14.1-1), the expressions $P(\dots)$ and $P(\dots|\dots)$ are defined as the probability of a combined event and the conditional probability, respectively. However, from the brain's standpoint, we might reverse the reasoning we followed above. For example, the brain only "sees" spike trains $\{t_i\}$ coming in from the sensors and it must link these to a stimulus s . In other words, the brain must evaluate the probability that s occurred given that $\{t_i\}$ is generated by the sensor; this probability is $P(s|\{t_i\})$. From the brain's point of view, the probability that both s and $\{t_i\}$ occur is $P(\{t_i\}|s)$ multiplied by the probability that we observe $\{t_i\}$:

$$P(\{t_i\}, s) = P(s|\{t_i\}) \times P(\{t_i\}) \quad (\text{A14.1-2})$$

Combining Equations (A14.1-1) and (A14.1-2), we have

$$\begin{aligned} P(\{t_i\}, s) &= P(s|\{t_i\}) \times P(\{t_i\}) = P(\{t_i\}|s) \times P(s) \\ &\rightarrow \boxed{P(s|\{t_i\}) = P(\{t_i\}|s) \times \frac{P(s)}{P(\{t_i\})}} \end{aligned} \quad (\text{A14.1-3})$$

Equation (A14.1-3), linking both conditional probabilities, is Bayes's rule.

APPENDIX 14.2

Poisson Process

The mean μ and standard deviation σ of a Poisson process (PDF $f(t) = \rho e^{-\rho t}$, with $f(t) = 0$ for $t < 0$) are equal. This can be shown by using the Laplace transform approach (see Appendix 3.4) or directly using Equations (3.9) to (3.11). For the mean, we get (using the integration limits from $0 \rightarrow \infty$ because $f(t) = 0$ for $t < 0$):

$$\begin{aligned}
 E(t) &= \int_0^{\infty} t \rho e^{-\rho t} dt = [-te^{-\rho t}]_0^{\infty} - \int_0^{\infty} -e^{-\rho t} dt = [-te^{-\rho t}]_0^{\infty} + \int_0^{\infty} e^{-\rho t} dt \\
 &= [-te^{-\rho t}]_0^{\infty} - \frac{1}{\rho} [e^{-\rho t}]_0^{\infty} \\
 &= [0 - 0] - \left[0 - \frac{1}{\rho} \right] = \frac{1}{\rho}
 \end{aligned} \tag{A14.2-1}$$

In this equation we used integration by parts to evaluate the integral; note that $[-te^{-\rho t}]_0^{\infty}$ evaluates to zero because $\left[-\frac{t}{e^{\rho t}}\right] = 0$ for $t = 0$ and also for $t = \infty$. The latter can be seen replacing the exponential with a power series, or if you prefer, by using l'Hôpital's rule.

For the expectation of t^2 :

$$\begin{aligned}
 E(t^2) &= \int_0^{\infty} t^2 \rho e^{-\rho t} dt = [-t^2 e^{-\rho t}]_0^{\infty} - \int_0^{\infty} -2te^{-\rho t} dt = [-t^2 e^{-\rho t}]_0^{\infty} + 2 \int_0^{\infty} te^{-\rho t} dt \\
 &= \underbrace{[-t^2 e^{-\rho t}]_0^{\infty}}_0 + \frac{2}{\rho} \underbrace{\int_0^{\infty} t \rho e^{-\rho t} dt}_{E(t) = \frac{1}{\rho}} = \frac{2}{\rho^2}
 \end{aligned} \tag{A14.2-2}$$

Similar to the approach we took in Equation (A14.2-1), we used integration by parts to evaluate the integral. The expression $[-t^2 e^{-\rho t}]_0^{\infty}$ evaluates to zero because $\left[-\frac{t^2}{e^{\rho t}}\right] = 0$ for $t = 0$ and after substituting a power series for the exponential it can be seen that the expression is also 0 for $t = \infty$. The variance σ^2 of the Poisson process is

$$\sigma^2 = E(t^2) - E(t)^2 = \frac{2}{\rho^2} - \frac{1}{\rho^2} = \frac{1}{\rho^2} \rightarrow \sigma = \frac{1}{\rho} \quad (\text{A14.2-3})$$

Combining Equations (A14.2-1) and (A14.2-3), we find that the mean and standard deviations are both equal to $\frac{1}{\rho}$.