

## Detecting and Estimating Signals in Noisy Cable Structures, II: Information Theoretical Analysis

Amit Manwani

Christof Koch

*Computation and Neural Systems Program, California Institute of Technology,  
Pasadena, CA 91125, U.S.A.*

This is the second in a series of articles that seek to recast classical single-neuron biophysics in information-theoretical terms. Classical cable theory focuses on analyzing the voltage or current attenuation of a synaptic signal as it propagates from its dendritic input location to the spike initiation zone. On the other hand, we are interested in analyzing the amount of information lost about the signal in this process due to the presence of various noise sources distributed throughout the neuronal membrane. We use a stochastic version of the linear one-dimensional cable equation to derive closed-form expressions for the second-order moments of the fluctuations of the membrane potential associated with different membrane current noise sources: thermal noise, noise due to the random opening and closing of sodium and potassium channels, and noise due to the presence of “spontaneous” synaptic input.

We consider two different scenarios. In the signal estimation paradigm, the time course of the membrane potential at a location on the cable is used to reconstruct the detailed time course of a random, band-limited current injected some distance away. Estimation performance is characterized in terms of the coding fraction and the mutual information. In the signal detection paradigm, the membrane potential is used to determine whether a distant synaptic event occurred within a given observation interval. In the light of our analytical results, we speculate that the length of weakly active apical dendrites might be limited by the information loss due to the accumulated noise between distal synaptic input sites and the soma and that the presence of dendritic nonlinearities probably serves to increase dendritic information transfer.

### 1 Introduction ---

The problem of neural coding, or how neural systems represent and process sensory information to make behavioral decisions crucial for the survival of the organism, is fundamental to understanding how brains work. Several strategies have been suggested as plausible candidates for the neural code (Perkel & Bullock, 1968; Theunissen & Miller, 1995). Currently, it is unclear

which, if any, is the most universal strategy. In fact, it is likely that different neural systems use different codes or maybe even a combination of different neural codes. Knowledge of the manner in which information is represented in the brain is crucial to the understanding of neural coding, since the efficacy of a code depends on the nature of the underlying representation. In the absence of a clear choice, it becomes necessary to compare the performance of neural systems under different representational paradigms. It is reasonable to assume that if neural systems were optimized to transmit information, the strategy yielding the highest information capacity is a likely candidate for the neural code used by the system.

In this article our goal is to quantify the information loss in linear cables due to three different sources of neuronal noise, under two different representational paradigms. The noise sources we shall consider have been modeled and characterized in the first part of this study, the previous article in this volume, henceforth referred to as M-K. The noise sources we consider are thermal noise due to the passive membrane resistance (*Johnson noise*), noise due to the stochastic channel openings and closings of membrane voltage-gated ion channels ( $K^+$  and  $Na^+$  here), and noise due to random background synaptic activity. Using results from M-K, we compare the relative magnitudes of the noise sources in linear cables. A list of mathematical symbols used in this article and the previous one is contained in the appendix of the previous article. For the purpose of this study, the cable is assumed to be infinite; however, the analysis can be easily generalized to accommodate other cable geometries.

Quantifying the magnitude of the membrane noise sources allows us to assess the efficacy of information transfer under two different paradigms. In the *signal estimation paradigm*, the goal is to estimate a random current waveform injected at a particular location from the membrane voltage at another location on the cable. We define a quantity called the *normalized coding fraction*,  $\xi$ , and use it to assess signal fidelity in the signal estimation task. In the *signal detection paradigm*, the objective is to detect the presence or absence of a presynaptic signal (a single spike) on observing the postsynaptic membrane voltage. The probability of detection error,  $P_e$ , is used to quantify performance in the signal detection task. Much of modern psychophysical research (Green & Swets, 1966) uses a signal-detection paradigm to assess performance. The framework used in the article is illustrated schematically in Figure 1. We derive expressions for the corresponding information-theoretical measures of signal efficacy (mean square error and information rate for signal estimation and, probability of error and mutual information for signal detection) and examine their dependence on different biophysical parameters.

The analysis should be viewed within the context of a long-term research program to reformulate one-dimensional linear and nonlinear cable theory in terms of an information-theoretical framework. Instead of adopting the classical approach pioneered by Rall (1959, 1969a, 1969b, 1989), which

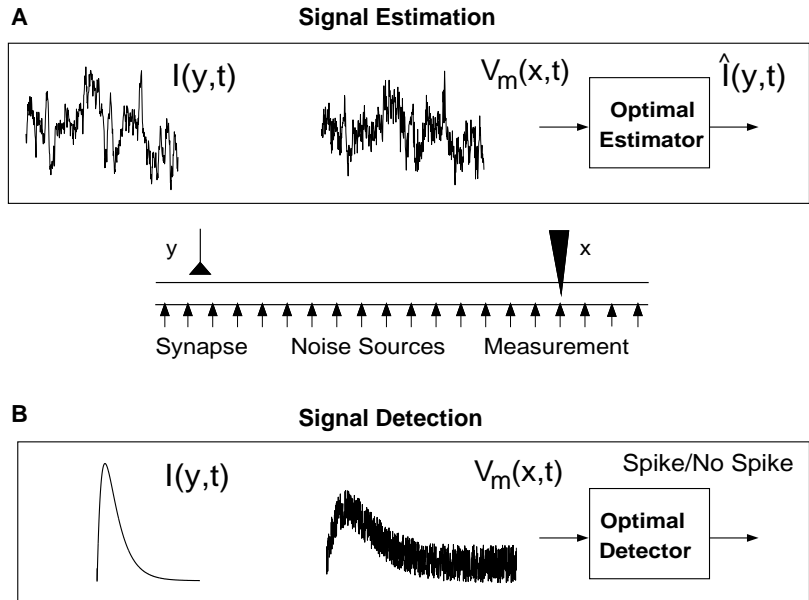


Figure 1: Channel model of a weakly active dendrite. The dendrite is modeled as a weakly active 1D cable with noise sources distributed along its length. By “weakly active,” we mean that the magnitude of the conductance fluctuations due to these sources is small compared to the baseline conductance of the membrane. Formally, this can be stated as  $\delta \ll 1$  (equation 2.10). These noise sources distort the synaptic signal as it propagates from its postsynaptic site  $y$  to a measurement (output) location  $x$ . Loss of fidelity is studied under two representational paradigms. (A) In signal estimation, the objective is to estimate optimally the input current  $I(y, t)$  from the membrane voltage  $V_m(x, t)$ . The normalized coding fraction  $\xi$  and the mutual information are used to quantify signal fidelity in the estimation task. (B) In signal detection, the objective is to detect optimally the presence of the synaptic input  $I(y, t)$  (in the form of a unitary synaptic event) on the basis of  $V_m(x, t)$ . The probability of error,  $P_e$ , and mutual information are used to quantify signal fidelity in the detection task.

focuses on the voltage change in response to single or multiple synaptic inputs, its effect on the cell body, the initiation and propagation of action potentials, and so on (Jack, Noble, & Tsien, 1975; Johnston & Wu, 1995; Koch, 1999), here we evaluate the ability of biophysical model systems to estimate, detect, and transmit information-bearing signals. We believe that like any other information processing system, neural systems need to be analyzed with both the (bio)physical and the information-theoretical aspects in mind. (For a related approach applied to electrical circuits, see Andreou & Furth, 1998.)

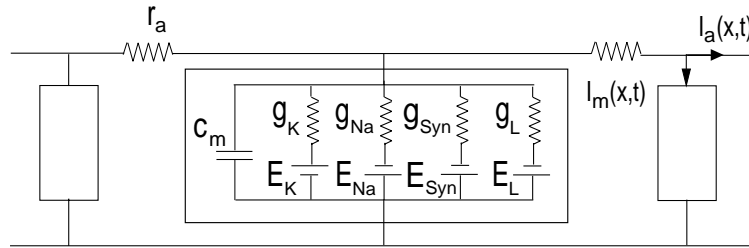


Figure 2: Equivalent circuit diagram of a dendritic 1D cable. The cable is modeled as an infinite ladder network.  $r_a$  ( $\Omega/\mu\text{m}$ ) denotes the longitudinal cytoplasmic resistance;  $c_m$  ( $\text{F}/\mu\text{m}$ ) and  $g_L$  ( $\text{S}/\mu\text{m}$ ) denote the transverse membrane capacitance and conductance (due to leak channels with reversal potential  $E_L$ ), respectively.  $I_a(x, t)$  denotes the longitudinal current, whereas,  $I_m(x, t)$  is the transverse membrane current. The membrane also contains active channels ( $\text{K}^+$ ,  $\text{Na}^+$ ) with conductances and reversal potentials denoted by  $(g_K, g_{Na})$  and  $(E_K, E_{Na})$  respectively, and fast, voltage-independent (AMPA-like) synapses with conductance  $g_{Syn}$  and reversal potential  $E_{Syn}$ . All conductances are in units of  $\text{S}/\mu\text{m}$ .

Most of the tools used for our work on membrane noise are contained within the excellent text by DeFelice (1981), which presents a thorough treatment of different sources of noise in biological membranes, along with an exhaustive review of relevant early research in the field. The two information-theoretical paradigms we use here, signal estimation and detection, are also well known. The novelty of our article, we believe, is that it combines classical cable theory with noise analysis and information theory in the context of neural coding. This will allow us to reinterpret a host of results from cable theory using information-theoretical measures.

## 2 The Cable Equation

We model the dendrite as the usual one-dimensional ladder network shown in Figure 2. (For assumptions underlying one-dimensional cable theory see Koch, 1999.)  $r_a$  represents the axial resistance of the intracellular cytoplasm.  $r_a$  (expressed in units of  $\Omega/\mu\text{m}$ ) can be obtained in terms of the more commonly used intracellular resistivity  $R_i$  as,

$$r_a = \frac{4 R_i}{\pi d^2}, \quad (2.1)$$

where  $d$  is the dendritic diameter (expressed in  $\mu\text{m}$ ).  $g_K$ ,  $g_{Na}$ , and  $g_L$  denote the transverse membrane conductances due to  $\text{K}^+$ ,  $\text{Na}^+$ , and leak channels distributed throughout the dendritic membrane. Recent research has established the existence of several types of active voltage-gated ion channels in

dendrites (Johnston, Magee, Colbert, & Cristie, 1996; Colbert & Johnston, 1996; Yuste & Tank, 1996; Magee, Hoffman, Colbert, & Johnston, 1998). The dendritic membrane also has an incidence of a large number of synapses from a vast multitude of other neurons. However, as in M-K, we restrict ourselves to fast voltage-independent synapses (AMPA-type) synapses here. Let  $g_{Syn}$  denote the transverse membrane conductance due to these fast AMPA-like synapses. All the conductances above are expressed in units of  $S/\mu\text{m}$ . The membrane capacitance due to the phospholipid bilayer is denoted by  $c_m$ . The units of  $c_m$  are  $\text{F}/\mu\text{m}$ , and it can be expressed in terms of the more commonly used specific capacitance  $C_m$  as

$$c_m = \pi d C_m. \quad (2.2)$$

The membrane voltage  $V_m$  satisfies the following partial differential equation,

$$\frac{\partial^2 V_m}{\partial x^2} = r_a \left[ c_m \frac{\partial V_m}{\partial t} + g_K(V_m - E_K) + g_{Na}(V_m - E_{Na}) + g_{Syn}(V_m - E_{Syn}) + g_L(V_m - E_L) + I_{inj} \right], \quad (2.3)$$

where  $I_{inj}(x, t)$  represents the current injected into the membrane from other sources that we have not explicitly considered here (thermal noise, synaptic input, stimulating electrode, etc.).

Since the conductances  $g_K$ ,  $g_{Na}$ ,  $g_{Syn}$  (and possibly even  $I_{inj}$ ) are stochastic processes, equation 2.3 denotes a highly nonlinear stochastic reaction-diffusion equation (Tuckwell, 1988b) since the ionic conductances are functions of  $V_m$  in themselves. However, it is more illustrative to express the random variables as deviations around some baseline values, as in M-K:

$$g_K = g_K^o + \tilde{g}_K, \quad (2.4)$$

$$g_{Na} = g_{Na}^o + \tilde{g}_{Na}, \quad (2.5)$$

$$g_{Syn} = g_{Syn}^o + \tilde{g}_{Syn}, \quad (2.6)$$

$$V_m = V^o + V. \quad (2.7)$$

$V^o$  is chosen such that it satisfies the equation

$$V^o = \frac{g_K^o E_K + g_{Na}^o E_{Na} + g_{Syn}^o E_{Syn} + g_L E_L}{G}, \quad (2.8)$$

where  $G = g_K^o + g_{Na}^o + g_{Syn}^o + g_L$  is the total input conductance, given by the sum of all the baseline conductances. Substituting for equations 2.4 through 2.7 in equation 2.3 gives

$$-\lambda^2 \frac{\partial^2 V}{\partial x^2} + \tau \frac{\partial V}{\partial t} + (1 + \delta)V = \frac{I_n}{G}, \quad (2.9)$$

where  $\lambda = 1/\sqrt{r_a G}$  is the characteristic length constant (in  $\mu\text{m}$ ) and  $\tau = c_m/G$  is the characteristic passive time constant (in msec) of the cable.  $\delta$  and  $I_n$  are random processes defined as

$$\delta = \frac{\tilde{g}_K + \tilde{g}_{Na} + \tilde{g}_{Syn}}{G}, \quad (2.10)$$

$$I_n = \tilde{g}_K(E_K - V^o) + \tilde{g}_{Na}(E_{Na} - V^o) + \tilde{g}_{Syn}(E_{Syn} - V^o) + \tilde{I}_{inj}. \quad (2.11)$$

$\delta$  corresponds to membrane conductance fluctuations due to synaptic and channel contributions and has a multiplicative effect on  $V$ . As in M-K,  $I_n$ , on the other hand, is a sum of the additive noise current due to these conductance fluctuations and the random component of the injected current  $\tilde{I}_{inj}$  (the expected value of  $I_{inj}$  is assumed to be zero). We assume that the conductance fluctuations are spatially white, zero-mean, wide-sense stationary (WSS) random processes, that is, the fluctuations at a location  $x$  are independent of those at another location  $y$ . It is plausible to assume that the individual conductance fluctuations are statistically independent since they have different origins. Thus,  $I_n$  is also a zero-mean WSS random process,  $\langle I_n(x, t) \rangle = 0$ .

We now make a simplifying assumption that  $\delta \ll 1$  and can be neglected in equation 2.9. We refer to this as the *weakly active* assumption. This allows us to reduce equation 2.9 to a linear, stochastic, partial differential equation. We shall also assume that the dynamics of components of the noise current  $I_n$  are given by their values at  $V_m = V^o$ . The steady-state (resting) solution of equation 2.9 (obtained by setting  $\delta$  and  $I_n$  to zero) is  $V = 0$ , which implies that we choose  $V^o = V_{rest}$ . Consequently,  $G$  is the resting membrane conductance. Similarly, the baseline conductances  $g_i^o$  satisfy  $g_i^o = g_i^\infty(V_{rest})$  where  $g_i^\infty(V_m)$  denotes the steady-state value of the conductance as a function of the membrane voltage. Thus, our assumptions are equivalent to saying that conductance fluctuations around  $V_{rest}$  are negligible compared to the resting conductance  $G$ . Additionally, the dynamics of the resulting current noise can be obtained from the dynamics of conductance fluctuations evaluated around  $V_{rest}$ . These assumptions need to be verified on a case-by-case basis. The simplest way to ensure their validity is to check for self-consistency of the solutions. Notice that equation 2.9 is an extension of the membrane patch analysis in M-K to a 1D cable.

Thus, our simplified version of equation 2.9 reads,

$$-\lambda^2 \frac{\partial^2 V}{\partial x^2} + \tau \frac{\partial V}{\partial t} + V = \frac{I_n}{G}, \quad (2.12)$$

and is in effect a stochastic version of the one-dimensional cable equation (Rall, 1969a; Tuckwell, 1988a, 1988b). Details of the derivation of the cable equation can be found in Rall (1969a) and Tuckwell (1988a). For the most part, our notation is similar to the one used in Tuckwell & Walsh (1983).

### 3 Noise in Linear Cables

The cable equation has a unique solution once the initial conditions and the boundary conditions are specified. For resting initial conditions ( $V = 0$  for  $t \leq 0$ ), the membrane fluctuations  $V$  are linearly related to the current input  $I_n$  and can be mathematically expressed as a convolution of  $I_n$  with the Green's function of the cable equation for the given boundary conditions. The Green's function of the cable, denoted by  $g(x, x', t, t')$ , specifies the voltage response of the cable at location  $x$  at time  $t$  to a current impulse  $\delta(x - x') \delta(t - t')$  injected at the location  $x'$  at time  $t'$ .  $g(x, x', t, t')$  has units of  $\mu\text{m}^{-1}\text{msec}^{-1}$ . By superposition,  $V(x, t)$  can be written as

$$V(x, t) = \frac{1}{G} \int_{-\infty}^{\infty} dx' \int_0^t dt' g(x, x', t, t') I_n(x', t'). \tag{3.1}$$

Since the system is time invariant,  $g(x, x', t, t') = g(x, x', t - t')$ . The exact form of  $g(x, x', t - t')$  depends on the nature of the boundary conditions of the partial differential equation. The expected value of  $V(x, t)$  is given by

$$\langle V(x, t) \rangle = \frac{1}{G} \int_{-\infty}^{\infty} dx' \int_0^t dt' g(x, x', t - t') \langle I_n(x', t') \rangle. \tag{3.2}$$

Since the current noise  $I_n$  is a zero-mean process,  $\langle V(x, t) \rangle = 0$ . Thus the variance of the membrane voltage fluctuations  $\sigma_V^2(x, t) = \langle V^2(x, t) \rangle$  is given by,

$$\begin{aligned} \sigma_V^2(x, t) = & \frac{1}{G^2} \int_{-\infty}^{\infty} dx' \int_{-\infty}^{\infty} dx'' \int_0^t dt' \int_0^t dt'' g(x, x', t - t') g(x, x'', t - t'') \\ & \langle I_n(x', t') I_n(x'', t'') \rangle. \end{aligned} \tag{3.3}$$

The quantity  $\langle I_n(x', t') I_n(x'', t'') \rangle$  represents the autocovariance of the current input, which we denote by  $C_n(x', x'', t', t'')$ . Since  $I_n(x, t)$  is a spatially white WSS process,  $C_n$  is of the form  $C_n(x', x'', t', t'') = C_n(t' - t'') \delta(x' - x'')$ , which simplifies equation 3.3 to

$$\begin{aligned} \sigma_V^2(x, t) = & \frac{1}{G^2} \int_{-\infty}^{\infty} dx' \int_0^t dt' \int_0^t dt'' \\ & \times g(x, x', t - t') g(x, x', t - t'') C_n(t' - t''). \end{aligned} \tag{3.4}$$

Since we assume that the cable starts receiving inputs at time  $t = 0$ , the membrane voltage fluctuations  $V$  cannot be a WSS process. This can be easily seen as  $\sigma_V^2$  depends on  $t$ . However, if we wait long enough for the transients associated with the initial condition to die out, at long timescales

the statistical properties of  $V(x, t)$  do not depend on  $t$ . In fact, it can be shown that  $V(x, t)$  is asymptotically ( $t \rightarrow \infty$ ) WSS (Tuckwell, 1988a). Another way to observe the same is by assuming that the system starts receiving its input at  $t = -\infty$ , in which case the dynamics stabilize by  $t$ . This can be observed by changing the limits of the time variable to  $(-\infty, t)$  in equation 3.3. The steady-state variance of  $V(x, t)$  is given by,

$$\sigma_V^2(x, \infty) = \frac{1}{G^2} \int_{-\infty}^{\infty} dx' \int_0^{\infty} dt \int_{-t'}^{\infty} dz g(x, x', t') g(x, x', t' + z) C_n(z). \quad (3.5)$$

When the autocovariance of the current noise  $C_n(z)$  decays much faster (has a much smaller support) than  $g(x, x', t')$ , one can approximate it by  $C_n(z) \approx C_0 \delta(z)$ , which allows equation 3.5 to be written as<sup>1</sup>

$$\sigma_V^2(x, \infty) \approx \frac{C_0}{G^2} \int_{-\infty}^{\infty} dx' \int_0^{\infty} dt g^2(x, x', t'). \quad (3.6)$$

This approximation holds when the membrane time constant  $\tau$ , which determines the temporal support of  $g(x, x', t')$ , is much larger than the time constants governing the dynamics of the noise sources. We call this approximation the *white noise approximation* (WNA), since we approximate the current noise covariance  $C_n$  by an impulse, the correlation function of a spectrally white stochastic process. The validity of this approximation can be verified easily by comparing the temporal width of  $C_n$  with the membrane time constant.

In general, the steady-state covariance  $C_V(x, s)$  of  $V(x, t)$  is given by

$$\begin{aligned} C_V(x, s) &= \lim_{t \rightarrow \infty} \langle V(x, t) V(x, t + s) \rangle, \\ &= \frac{1}{G^2} \int_{-\infty}^{\infty} dx' \int_0^{\infty} dt \int_{-t'}^{\infty} dz g(x, x', t') \\ &\quad \times g(x, x', t' + z) C_n(z - s). \end{aligned} \quad (3.7)$$

Notice that  $C_V(x, s)$  is of the form  $C_V(x, s) = \int_{-\infty}^{\infty} dx' g(x, x', s) * g(x, x', -s) * C_n(s)$ , where  $*$  denotes a convolution operation. Consequently, the voltage noise power spectrum is given by

$$S_V(x, f) = \mathcal{F}\{C_V(x, t)\} = \underbrace{\frac{S_n(f)}{G^2}}_{S_{F_n}} \underbrace{\int_{-\infty}^{\infty} dx' |\mathcal{G}(x, x', f)|^2}_{G_{F_n}}, \quad (3.8)$$

<sup>1</sup> By definition,  $C_0 = S_n(0)$  where  $S_n(f)$  is the Fourier transform of  $C_n$ , or equivalently, the power spectrum of the current noise.



where  $S_n(f) = \mathcal{F}\{C_n(s)\}$  is the power spectral density of the current noise and  $\mathcal{G}(x, x', f) = \mathcal{F}\{g(x, x', t)\}$  is the transfer function of the Green's function of the system.  $\mathcal{F}\{g(x)\}$  denotes the Fourier transform operation defined as  $\int_{-\infty}^{\infty} dx g(x) \exp(-i2\pi fx)$ .

Notice that we have expressed the voltage spectrum  $S_V(x, f)$  (in units of  $V^2/\text{Hz}$ ) in equation 3.8 as a product of two factors. The first factor,  $SF_n$  (source factor), represents the power spectral density of the current noise source scaled appropriately (by  $1/G^2$ ) to have the units of  $V^2\mu\text{m}/\text{Hz}$ .  $SF_n$  depends on the properties of the noise sources and the resting membrane conductance. The second factor,  $GF_n$  (geometry factor), characterizes the transformation of the current noise input by the cable into membrane voltage fluctuations and has units of  $\mu\text{m}^{-1}$ .  $GF_n$  depends on factors (geometry, boundary conditions, and so on) that determine the Green's function of the cable. This decomposition allows us to decouple the effects of cable geometry from those of the current noise sources. When the WNA holds,  $SF_n$  is a constant ( $SF_n \approx S_n(0)/G^2$ ), and in effect  $GF_n$  describes the spectral properties of  $V(x, t)$ .

**3.1 Special Case: The Infinite Cable.** Here we consider the simplistic case of an infinite cable. Although this theoretical idealization approximates reality only loosely, it offers significant insight into understanding more complicated scenarios. The analytical tractability of the infinite case allows us to derive closed-form expressions for the quantities of interest and use them to develop an intuitive understanding of some of the fundamental issues of the problem. Unfortunately, closed-form expressions for other cable geometries (semi-infinite cable with a sealed end, finite cable with sealed or killed ends) cannot be derived, and one has to take recourse to numerical techniques. Nevertheless, the Green's functions for these cable geometries have been derived in semiclosed form (Jack et al., 1975; Tuckwell, 1988a). Moreover, compartmental modeling of realistic dendritic trees (Segev & Burke, 1998) has become routine. Thus, using numerical approaches, it is relatively straightforward to extend the analysis to more complicated scenarios.

The Green's function for the infinite cable is given as (Jack et al., 1975),

$$g(x, x', t) = \frac{1}{\lambda \tau} \frac{e^{-T}}{\sqrt{4\pi T}} e^{-\frac{(x-x')^2}{4T}} \quad -\infty < x, x' < \infty, \quad 0 \leq t < \infty, \quad (3.9)$$

where  $X = x/\lambda, X' = x'/\lambda$ , and  $T = t/\tau$  are the corresponding dimensionless variables. It can be shown that the geometry factor corresponding to the voltage variance is given by (Tuckwell & Walsh, 1983),

$$\sigma_V^2(x, t) = \frac{1}{4\lambda \tau} \left[ 1 - \text{Erfc} \left( \sqrt{2t/\tau} \right) \right], \quad (3.10)$$

where  $\text{Erfc}(\cdot)$  is the complementary error function,

$$\text{Erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-y^2} dy. \quad (3.11)$$

Thus, in steady state, the voltage variance geometry factor is given by

$$\sigma_V^2(x) = \lim_{t \rightarrow \infty} \sigma_V^2(x, t) = \frac{1}{4\lambda\tau}. \quad (3.12)$$

Note that the voltage noise variance  $\sigma_V^2$  is independent of the measurement location  $x$ . This is also intuitively consistent with the inherent symmetry of the infinite cable. The expressions for the geometry factors for  $C_V(x, s)$  and  $S_V(x, f)$  are given as

$$C_V(x, s) = \frac{1}{4\lambda\tau} \text{Erfc}(\sqrt{s/\tau}), \quad (3.13)$$

$$S_V(x, f) = \frac{1}{2\lambda} \frac{\sin[\tan^{-1}(2\pi f\tau)/2]}{2\pi f\tau [1 + (2\pi f\tau)^2]^{1/4}}. \quad (3.14)$$

Notice that in the limit of high frequencies,

$$S_V(x, f) \sim \frac{1}{8\lambda(\pi f\tau)^{3/2}}. \quad (3.15)$$

Thus, for the infinite cable, the voltage noise spectrum decays asymptotically as  $f^{-3/2}$  with frequency. This holds for frequencies larger than  $f_m = 1/\tau$  but smaller than those for which  $S_n(f)$  can no longer be regarded as a constant (equal to its value at  $f = 0$ ,  $S_n(0)$ ). For very high frequencies,  $S_V(f)$  decays faster than  $f^{-3/2}$  due to the spectral profile of the current noise  $S_n(f)$ . The exact expression (after multiplying by  $SF_n$ ) for  $S_V(x, f)$  is given as

$$S_V(x, f) = \frac{S_n(f)}{2\lambda G^2} \frac{\sin[\tan^{-1}(2\pi f\tau)/2]}{2\pi f\tau [1 + (2\pi f\tau)^2]^{1/4}}. \quad (3.16)$$

#### 4 Signal Transmission in Linear Cables

Up to this point, we have addressed the problem of noise accumulation in a linear cable as a result of fluctuations due to different membrane conductances distributed along the dendritic length. We now analyze the attenuation of a synaptic signal, delivered at a particular dendritic location, as it propagates passively along the dendrite. Our approach is to exploit the linearity of the cable equation and decompose the voltage at a given location into signal and noise components. The input signal depends on

the paradigm we use. In the signal estimation paradigm, the input is in the form of a random current waveform  $I_s(t)$ , injected at a given dendritic location; in the signal detection paradigm, the input is a unitary, excitatory, postsynaptic current pulse (EPSC) delivered across a dendritic synapse at the given location.

In principle, a synaptic input should be treated as a conductance change triggered by a presynaptic action potential in parallel with a synaptic battery. However, in the signal estimation paradigm, where our goal is to assess how well continuous signals can be reconstructed from the membrane potential, we would need to invoke a mechanism that transforms a continuous signal into a spike train driving the synapse. For now, we bypass this problem and assume that the synaptic input corresponds to a continuous current that is directly injected into the cable. (We will return to the problem of linking a presynaptic spike train to the postsynaptic synaptic current in a future publication.)

We now use the appropriate Green's function  $g(x, y, t)$  for a given cable geometry to derive expressions for the voltage response  $V(x, y, t)$  due to a current  $I_s(t)$  injected at location  $y$ . By superposition,

$$V(x, y, t) = \frac{1}{G} \int_0^t dt' g(x, y, t - t') I_s(t'). \quad (4.1)$$

In the signal detection task,  $I_s(t)$  is a deterministic signal, which we model by the  $\alpha$  function, first introduced by Rall, 1967,  $I_s(t) = A t \exp(-t/t_{peak})$ , whereas in the signal estimation task,  $I_s(t)$  is a continuous random process. Consequently,  $V(x, y, t)$  is a (nonstationary) random process, which is asymptotically wide-sense stationary as  $t \rightarrow \infty$  (steady state). It is straightforward to derive expressions for the signal component (due to  $I_s(t)$ ) of the voltage power spectra  $S_V(x, y, f)$  and variance  $\sigma_V^2(x, y)$  as

$$S_V(x, y, f) = \frac{S_s(f)}{G^2} |\mathcal{G}(x, y, f)|^2, \quad (4.2)$$

$$\sigma_V^2(x, y) = \int_{-\infty}^{\infty} df S_V(f), \quad (4.3)$$

where  $S_s(f)$  is the power spectral density of the input  $I_s(t)$ . Thus, using equations 4.2 and 4.3, we can analyze how the signal component of the membrane voltage decreases as a function of the distance from the input location for different cable geometries.

**4.1 Special Case: The Infinite Cable.** As before, we restrict ourselves to the case of an infinite cable. The expression for the signal component

$S_V(x, y, f)$  for the infinite cable is given by

$$S_V(x, y, f) = \frac{S_s(f)}{4\lambda^2 G^2} \frac{\exp(-\rho|x-y|/\lambda)}{[1 + (2\pi f\tau)^2]^{1/2}}, \quad (4.4)$$

where

$$\rho = 2 \left[ 1 + (2\pi f\tau)^2 \right]^{1/4} \cos \left[ \tan^{-1}(2\pi f\tau)/2 \right]. \quad (4.5)$$

Notice that  $S_V(x, y, f)$  is symmetric with respect to  $x$  and  $y$  and depends on only the electrotonic distance  $X = |x - y|/\lambda$  between the input and the measurement location. For  $f \rightarrow \infty$ ,  $S_V(x, y, f)$  varies as

$$S_V(x, y, f) \sim \frac{S_s(f)}{4\lambda^2 G^2} \frac{\exp(-\sqrt{4\pi f\tau} X)}{2\pi f\tau}. \quad (4.6)$$

If  $S_s(f)$  is almost flat over the bandwidth of the cable, we can derive a simplified expression for the variance  $\sigma_V^2(X)$  as

$$\sigma_V^2(X) = \frac{S_s(0)}{\lambda^2 G^2 \tau} \frac{K_0(2X)}{2\pi}, \quad (4.7)$$

where  $K_0(\cdot)$  denotes the zeroth-order modified Bessel function of the second kind.  $K_0(u)$  has a singularity at the origin, and so the variance at the input location ( $x = y$ ) is unbounded. The asymptotic behavior of  $K_0(u)$  can be expressed as (Wan & Tuckwell, 1979)

$$K_0(u) \sim -\log(u) \quad (u \rightarrow 0) \quad (4.8)$$

$$K_0(u) \sim \sqrt{\frac{\pi}{2u}} e^{-u} \quad (u \rightarrow \infty). \quad (4.9)$$

Thus, the variance  $\sigma_V^2(X)$  has a logarithmic singularity at the origin and decays approximately exponentially with  $X$  for large  $X$ . The singularity is a result of the approximation of the autocorrelation of  $I_s(t)$  by a  $\delta$  function, in comparison to the Green's function of the cable. This approximation breaks down for  $X \approx 0$ , for which  $g(x, y, t)$  has a very small temporal support, comparable to or smaller than the correlation time of  $I_s(t)$ . This eliminates the singularity in  $\sigma_V^2$ .

More realistic models like the "cylinder with a lumped soma model" (Rall, 1960, 1969b), which includes the effect of the low somatic impedance, or compartmental models of neurons with extensive dendritic trees (Segev & Burke, 1998), are not amenable to closed-form analysis and can only be studied numerically. However, a knowledge of the Green's function of the cable enables us to determine the spectral properties of both the signal and noise contributions to the membrane voltage fluctuations. As we will see, knowledge of the signal and noise spectra is sufficient to quantify the information loss.

## 5 Signal Estimation

---

Consider the following problem of estimating a signal in the presence of noise. Let  $s(t)$  be a WSS random process (signal), filtered by a linear filter  $g(t)$  and additively corrupted by another WSS random process (noise)  $n(t)$  to give the observed process (measurement)  $m(t)$ ,

$$m(t) = g(t) * s(t) + n(t). \quad (5.1)$$

Our goal is to recover the signal  $s(t)$  from the noisy measurements  $m(t)$  in an optimal way. The criterion of optimality we adopt is the mean-square error between  $s(t)$  and our estimate of  $s(t)$  obtained using the measurements  $m(t)$ , denoted by  $\hat{s}(t)$ . Thus, we choose  $\hat{s}(t)$  such that the variance of the error between  $s(t)$  and  $\hat{s}(t)$  is minimized. For the sake of simplicity, we will restrict ourselves to linear estimates of the form

$$\hat{s}(t) = h(t) * m(t). \quad (5.2)$$

Since  $\hat{s}(t)$  is completely specified by the filter  $h(t)$ , the objective is to derive the optimal filter that minimizes the mean-square estimation error  $\mathcal{E}$ ,

$$\mathcal{E} = \langle (s(t) - \hat{s}(t))^2 \rangle = \langle s^2(t) \rangle + \langle \hat{s}^2(t) \rangle - 2\langle s(t)\hat{s}(t) \rangle. \quad (5.3)$$

This optimal linear estimation problem, first formulated and solved by Wiener (1949), led to the development of statistical communication theory and information theory (Shannon, 1949; Cover & Thomas, 1991). It has been modified by Bialek and colleagues (Bialek, Rieke, van Steveninck, & Warland, 1991; Bialek & Rieke, 1992; Rieke, Warland, van Steveninck, & Bialek, 1997) and successfully applied to quantify information processing in some peripheral biological systems (van Steveninck & Bialek, 1988, 1995; Rieke, Warland, & Bialek, 1993; Rieke, Bodnar, & Bialek, 1995; Rieke et al., 1997). This approach, called the *reconstruction approach*, has come to be an important tool in theoretical neuroscience (Theunissen & Miller, 1991; Wessel, Koch, & Gabbiani, 1996; Gabbiani, Metzner, Wessel, & Koch, 1996; Gabbiani, 1996). (For an extensive tutorial on the topic, see Gabbiani & Koch, 1998.)

Optimal linear estimators satisfy the orthogonality property (Gabbiani, 1996), which in our context can be expressed as

$$\langle (s(t_1) - \hat{s}(t_1)) m(t_2) \rangle = 0 \quad \forall t_1, t_2. \quad (5.4)$$

(For additional properties on optimal linear estimators, refer to Papoulis, 1991.)

If the constraint of causality is not imposed on the filter  $h(t)$ , the optimal filter can be obtained by substituting for  $\hat{s}(t)$  from equation 5.2 in eq. 5.4,

$$R_{sm}(t) = h(t) * R_{mm}(t), \quad (5.5)$$

where  $R_{sm}(t)$  is the cross-correlation between  $s(t)$  and  $m(t)$  and  $R_{mm}(t)$  is the autocorrelation of  $m(t)$ . Taking Fourier transforms on both sides of equation 5.5 gives the transfer function  $H(f)$  of the optimal filter in terms of the power spectrum of  $m(t)$  ( $S_{mm}(f)$ ) and the cross-spectrum between  $s(t)$  and  $m(t)$  ( $S_{sm}(f)$ ),

$$H(f) = \frac{S_{sm}(f)}{S_{mm}(f)}, \quad (5.6)$$

where

$$H(f) = \mathcal{F}\{h(t)\}, \quad S_{sm}(f) = \mathcal{F}\{R_{sm}(t)\} \text{ and } S_{mm}(f) = \mathcal{F}\{R_{mm}(t)\}. \quad (5.7)$$

Thus, we can use optimal linear estimation theory to analyze the problem of signal estimation in linear cables. We assume that information is encoded in the time variations of the input current  $I_s(t)$ , which is injected at a certain location along the cable. We are interested in quantifying how much information is lost due to electrotonic attenuation and the membrane noise sources as the signal corresponding to this input propagates passively down the cable. We estimate this by assessing how well we can recover  $I_s(t)$  from the voltage fluctuations  $V(x, t)$  as a function of distance from the input location. By analogy to the problem in equation 5.1,  $s(t)$  corresponds to  $I_s(t)$  and  $m(t)$  to  $V(x, t)$ . We can decompose  $V(x, t)$  into two components: a signal component,  $V_s(x, t)$ , due to  $I_s(t)$ , and a noise component,  $V_n(x, t)$ , reflecting the combined influence of all the noise sources that have been discussed in detail in M-K.  $g(t)$  corresponds to the Green's function of the cable for an input received at location  $y$ . Due to linearity,  $V(x, t) = V_s(x, t) + V_n(x, t)$ . Thus, the power spectrum of the signal component  $V_s(x, t)$  defined as  $S_V^s(x, y, f)$  can be written as

$$S_V^s(x, y, f) = \frac{S_s(f)}{G^2} |\mathcal{G}(x, y, f)|^2, \quad (5.8)$$

where  $S_s(f)$  denotes the power spectral density of  $I_s(t)$ ,  $\mathcal{G}(x, y, f)$  denotes the Fourier transform of the Green's function of the cable, and  $G$  is the input conductance.

Similarly, the power spectrum of the noise component,  $V_n(x, t)$ , defined as  $S_V^n(x, y, f)$  is given by

$$S_V^n(x, y, f) = \frac{S_n(f)}{G^2} \int_{-\infty}^{\infty} dy |\mathcal{G}(x, y, f)|^2. \quad (5.9)$$

We assume that the noise component  $V_n(x, t)$  and the signal component  $V_s(x, t)$  are uncorrelated with each other. Thus, the power spectrum of  $V(x, t)$

(denoted by  $S_{vv}(x, f)$ ) is

$$S_{vv}(x, f) = S_V^s(x, y, f) + S_V^n(x, f) \quad (5.10)$$

$$= \frac{S_s(f)}{G^2} |\mathcal{G}(x, y, f)|^2 + \frac{S_n(f)}{G^2} \int_{-\infty}^{\infty} dy |\mathcal{G}(x, y, f)|^2. \quad (5.11)$$

Similarly the cross-spectrum between  $I_s(t)$  and  $V(x, t)$  (denoted by  $S_{iv}(x, f)$ ) is

$$S_{iv}(x, f) = S_V^s(x, y, f) \quad (5.12)$$

$$= \frac{S_s(f)}{G^2} |\mathcal{G}(x, y, f)|^2. \quad (5.13)$$

Thus, using equation 5.6, the expression for the optimal filter can be derived in the frequency domain as

$$H(f) = \frac{S_V^s(x, y, f)}{S_V^s(x, y, f) + S_V^n(x, f)}. \quad (5.14)$$

Thus, the mean-square error  $\mathcal{E}$  in the signal estimation task is

$$\mathcal{E} = \int_{-\infty}^{\infty} df \frac{S_s(f) S_V^n(x, f)}{S_V^s(x, y, f) + S_V^n(x, f)}. \quad (5.15)$$

Notice that the computation of  $\mathcal{E}$  requires knowledge of only the signal and noise spectra ( $S_s(f)$  and  $S_n(f)$ , respectively) and the Green's function  $g(x, y, t)$  of the cable.

We assume that the input  $I_s(t)$  is a white, band-limited signal with bandwidth  $B_s$  and variance  $\sigma_s^2$ . This implies that the signal spectra  $S_s(f)$  is flat over the frequency range  $[-B_s, B_s]$  and zero elsewhere:

$$S_s(f) = \begin{cases} \frac{\sigma_s^2}{2B_s}, & |f| \leq B_s, \\ 0, & \text{otherwise.} \end{cases} \quad (5.16)$$

Substituting for equation 5.16 in equation 5.15 gives

$$\mathcal{E} = \frac{\sigma_s^2}{B_s} \int_0^{B_s} df \frac{S_V^n(x, f)}{S_V^s(x, y, f) + S_V^n(x, f)}. \quad (5.17)$$

As in Gabbiani (1996), we normalize  $\mathcal{E}$  with respect to the input variance,  $\sigma_s^2$ , to obtain a dimensionless quantity, called the coding fraction  $\xi$ ,

$$\xi = 1 - \frac{\mathcal{E}}{\sigma_s^2}, \quad 0 \leq \xi \leq 1. \quad (5.18)$$

The coding fraction  $\xi$  is an index of the efficacy in the signal estimation task;  $\xi = 1$  implies perfect reconstruction, whereas  $\xi = 0$  implies performance at chance.

We can also define a frequency-dependent signal-to-noise ratio  $SNR(x, y, f)$ ,

$$SNR(x, y, f) = \frac{S_V^s(x, y, f)}{S_V^n(x, f)}, \quad (5.19)$$

which is a ratio of the signal and noise power at frequency  $f$ . This allows us to express  $\xi$  as

$$\xi = \frac{1}{B_s} \int_0^{B_s} df \frac{SNR(x, y, f)}{1 + SNR(x, y, f)}. \quad (5.20)$$

If  $SNR(x, y, f)$  monotonically decreases with frequency, it can easily be seen that for a fixed amount of input power  $\sigma_s^2$ , the coding fraction  $\xi$  decreases with the input bandwidth  $B_s$ —that is, the reconstructions become poorer as the signal bandwidth increases. For the infinite cable, the signal component of the voltage fluctuations  $S_V^s(x, y, f)$  depends on only  $|x - y|$ . Thus,  $SNR(x, y, f)$  and  $\xi$  depend on only the relative electrotonic distance  $X$  between the input ( $y$ ) and measurement ( $x$ ) locations and not on their absolute values. Since the signal power attenuates with  $X$ , whereas noise power does not depend on  $X$ ,  $SNR(x, y, f)$  decreases monotonically with  $X$ . Consequently,  $\xi$  decreases monotonically with  $X$ .

Our analysis so far has remained independent of the probability distributions of the signal and the noise. Only a knowledge of the signal and noise power spectra (second-order statistics) was needed to compute  $\xi$ . This is because we restricted ourselves to the class of linear estimators. In order to derive more sophisticated nonlinear estimators, which would outperform linear estimators in general, we would need to make use of higher-order (greater than second-order) statistical information about the signal and noise processes. However, these nonlinear estimators are usually complicated to implement and difficult to analyze. Besides, it can be shown that if both the signal and noise are jointly gaussian, the optimal linear estimator is also the optimal estimator (over the class of all estimators). The gaussian assumption simplifies the analysis considerably and allows us to derive expressions for measures of signal fidelity other than the reconstruction error  $\mathcal{E}$ . Since the choice of the input  $I_s(t)$  lies with the experimenter, we can assume it to be gaussian by design. It can also be shown that under conditions for which the central limit theorem (Papoulis, 1991) holds,  $V(x, t)$  can be regarded as a gaussian process as well. Thus, henceforth we shall assume that both  $I_s(t)$  and  $V(x, t)$  are gaussian processes.



Information theory (Shannon, 1949; Cover & Thomas, 1991) allows us to quantify the amount of statistical information one random quantity conveys about another, given their joint probability distribution. It also provides a model-independent measure of the similarity between random covarying quantities  $a$  and  $b$ , called the mutual information (denoted by  $I(a; b)$ ) between  $a$  and  $b$ . For stochastic processes  $I_s(t)$  and  $V(t)$ ,  $I[I_s(t); V(t)]$  is called the information rate and is measured in units of bits per second. The information rate depends in general on the joint probability distribution of the two processes since gaussian processes are completely characterized by their second-order moments,  $I[I_s(t); V(t)]$  depends on only the joint spectral properties of  $I_s(t)$  and  $V(t)$ .

We can regard the signal estimation task as an effective continuous communication channel in the information-theoretical sense (see Figure 3A).  $I_s(t)$  denotes the input to the channel, whereas  $\hat{I}_s(t)$ , the optimal linear estimate obtained from  $V(x, t)$ , denotes its output. The effective additive noise added by the channel can be denoted by  $I_n(t)$ . This channel model allows us to compute the mutual information between  $I_s(t)$  and  $V(x, t)$ . If  $I_s(t)$  and  $V(t)$  (dropping the argument  $x$  for convenience) are jointly gaussian processes, the mutual information between them is given by (Shannon, 1949):

$$\begin{aligned} I[I_s(t); V(t)] &= \frac{1}{2} \int_{-\infty}^{\infty} df \log_2 \left[ \frac{S_{vv}(x, f)}{S_V^n(x, f)} \right], \\ &= \frac{1}{2} \int_{-\infty}^{\infty} df \log_2 \left[ 1 + \frac{S_V^s(x, y, f)}{S_V^n(f)} \right] \quad \text{bits/sec.} \end{aligned} \quad (5.21)$$

In terms of the signal-to-noise ratio  $SNR(x, y, f)$  and the bandwidth  $B_s$ , the mutual information can be expressed as

$$I[I_s(t); V(t)] = \frac{1}{2} \int_{-B_s}^{B_s} df \log_2 [1 + SNR(x, y, f)] \quad \text{bits/sec.} \quad (5.22)$$

The capacity of a communication channel is defined as the maximum amount of information that can be transmitted across it. If the noise properties of the system are given, we are left to vary only the properties of the input signal to achieve maximal information transfer. It is known that when the noise is additive and gaussian, the mutual information is maximized when the signal itself is gaussian (Cover & Thomas, 1991). Since a gaussian process is completely specified by its power spectral density, we need to find the optimal input power spectrum that maximizes  $I$ . This optimization is well defined only when we impose some constraints on the input spectra, since  $I$  can be made arbitrarily high by choosing an infinite power input signal. Thus, we assume that the input is both power and bandwidth limited, which is equivalent to saying that the input spectra satisfies the following

constraint:

$$\int_{-B_s}^{B_s} df S_s(f) = \sigma_s^2, \quad (5.23)$$

where  $\sigma_s^2$  is the input variance (power) and  $B_s$  denotes the input bandwidth. The capacity of the estimation channel can be formally defined as

$$C = \operatorname{argmax}_{S_s(f)} I[I_s(t); V(t)] \text{ such that } \int_{-B_s}^{B_s} df S_s(f) = \sigma_s^2. \quad (5.24)$$

This allows us to rewrite  $I[I_s(t); V(t)]$  from equation 5.22 as

$$I[I_s(t); V(t)] = \frac{1}{2} \int_{-\infty}^{\infty} df \log_2 \left[ 1 + \frac{S_s(f)}{S_{en}(f)} \right]. \quad (5.25)$$

Setting up the optimization problem as a Lagrange multiplier problem, we need to maximize the following functional,

$$F(S_s, \nu) = \frac{1}{2} \int_{-B_s}^{B_s} df \log_2 \left[ 1 + \frac{S_s(f)}{S_{en}(f)} \right] - \nu \int_{-B_s}^{B_s} df S_s(f), \quad (5.26)$$

where  $\nu$  is a Lagrange multiplier corresponding to the power constraint.

---

Figure 3: *Facing page*. Channel models for signal estimation and signal detection. (A) Effective communication channel model for the signal estimation task. The injected current  $I_s(t)$  represents the input to the channel, and the optimal linear estimate  $\hat{I}_s(t)$  derived from the membrane voltage,  $V(x, t)$ , represents the channel output.  $I_n(t) = \hat{I}_s(t) - I_s(t)$  is the equivalent additive noise introduced by the channel. The mutual information between  $I_s(t)$  and  $V(x, t)$  is bounded below by the information between  $I_s(t)$  and  $\hat{I}_s(t)$ . (B) Graphical demonstration of the “water-filling” algorithm used to compute the channel capacity for signal estimation.  $S_{en}(f)$  represents the effective current noise spectral density due to the membrane noise sources (referred back to the input),  $\nu$  represents the Lagrange multiplier (see equation 5.28), and  $S_s(f)$  represents the optimal signal power spectrum that maximizes channel capacity. For the given amount of signal power ( $\sigma_s^2$ ), the optimal strategy is to transmit higher power at frequencies where the noise power is low, and vice versa, such that, wherever possible, the sum of the signal power and noise power is a constant ( $1/\nu$ ). (C) Effective binary communication channel model for signal detection where the goal is to detect the presence of a synaptic input from the voltage  $V(x, t)$  at a distance  $X$  from the input location. Binary random variables  $M$  and  $D$  denote the input and output of the channel, respectively. False alarm  $P_F$  and miss error  $P_M$  rates of the optimal detector represent the crossover probabilities of the binary detection channel.

We express  $SNR(x, y, f)$  as a ratio of the input spectrum  $S_s(f)$  and an effective noise power spectral density denoted by  $S_{en}(f)$ ,

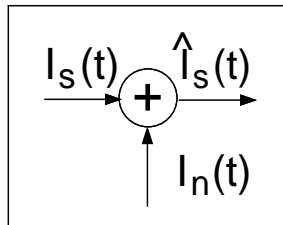
$$SNR(x, y, f) = \frac{S_s(f)}{S_{en}(f)},$$

$$\text{where } S_{en}(f) = \frac{S_n(f)}{|\mathcal{G}(x, y, f)|^2} \int_{-\infty}^{\infty} dy |\mathcal{G}(x, y, f)|^2. \quad (5.27)$$

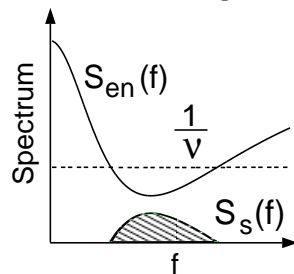
A simple exercise in calculus of variations (Courant & Hilbert, 1989) reveals that at the extrema of  $F(S_s, v)$ , the following equation is satisfied,

$$S_s(f) = \left[ \frac{1}{v} - S_{en}(f) \right]_+, \quad (5.28)$$

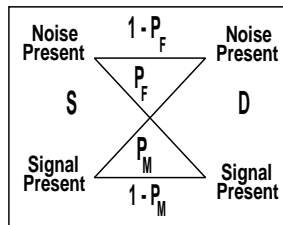
**A Estimation Channel**



**B "Water Filling"**



**C Detection Channel**



where

$$\lfloor x \rfloor_+ = \begin{cases} x, & \text{for } x \geq 0, \\ 0, & \text{for } x < 0. \end{cases} \quad (5.29)$$

The Lagrange multiplier  $\nu$  can be determined by solving

$$\int_{-B_s}^{B_s} df \left[ \frac{1}{\nu} - S_{en}(f) \right]_+ = \sigma_s^2. \quad (5.30)$$

The optimal way to distribute the available signal power is to transmit higher power at frequencies where the noise power is low and lesser or even zero power at frequencies for which the noise power is large. This procedure is graphically illustrated in Figure 3B. Thus, when the effective noise spectrum is low pass (high pass, respectively), the optimal input signal spectrum is high pass (low pass, respectively). Frequencies for which equation 5.28 can be satisfied without violating the power constraint (equation 5.23), the sum of the signal and noise power is constant. This is often referred to as the water-filling strategy (Cover & Thomas, 1991). By definition, the input power spectrum is nonnegative ( $S_s(f) \geq 0$ ), and so equation 5.28 cannot be satisfied for all frequencies in general, especially if the available input power  $\sigma_s^2$  is small. Let  $\Delta_s$  denote the set of frequencies,  $\{f \mid -B_s \leq f \leq B_s, 1/\nu - S_{en}(f) \geq 0\}$ , which is also referred to as the support of  $S_s(f)$ . The capacity of the estimation channel can be formally expressed as

$$C = \frac{1}{2} \int_{\Delta_s} df \log_2 \left[ \frac{\nu}{S_{en}(f)} \right] \quad \text{bits/sec.} \quad (5.31)$$

## 6 Signal Detection

---

In the signal estimation paradigm, both the signal and noise were continuous random processes. We now consider a different problem: detecting the presence of a known deterministic signal in noise. This scenario arises quite frequently in science and engineering (radar, communications, pattern recognition, psychophysics, etc.) and is commonly known as the *signal detection problem*. The goal in signal detection is to decide which member from a finite set of known signals was generated by a source, based on noisy measurements of its output. We restrict ourselves to the binary case, where the set has two elements: the signal (denoted by  $s(t)$ ) and the noise (denoted by  $n(t)$ ). We further assume that  $s(t)$  is filtered by a known filter  $g(t)$  and additively corrupted by  $n(t)$ , to give rise to the measured output (denoted by  $m(t)$ ). Our goal is to decide whether the observations  $m(t)$  (available over a period  $0 \leq t \leq T$ ) are due to noise  $n(t)$  (hypothesis  $H_0$ ) or a filtered, noisy

version of the signal  $s(t)$  (hypothesis  $H_1$ ). This can be formally expressed as

$$\begin{aligned} \mathbf{H}_0 : m(t) &= n(t), & 0 \leq t \leq T & \quad \text{Noise} \\ \mathbf{H}_1 : m(t) &= g(t) * s(t) + n(t), & 0 \leq t \leq T & \quad \text{Signal + noise.} \end{aligned} \quad (6.1)$$

Thus, a signal detection task involves making a decision about the presence or absence of a known signal  $s(t)$  buried in noise  $n(t)$  on the basis of the observations  $m(t)$ . In psychophysics, such a procedure is known as a yes/no task (Green & Swets, 1966). Within a neurobiological context, Newsome and his colleagues used a binary motion detection task to great effect (Newsome, Britten, & Movshon, 1989; Britten, Shadlen, Newsome, & Movshon, 1992; Shadlen & Newsome, 1998) to study the extent to which individual cortical neurons explain the performance of the monkey.

Decision errors are of two kinds. A false alarm (F) error occurs when we decide in favor of the signal ( $H_1$ ) when actually only noise was present ( $H_0$ ), and a miss (M) error occurs when we decide in favor of the noise ( $H_0$ ) when in fact the signal was present ( $H_1$ ). The probabilities of these errors are denoted as

$$\begin{aligned} P_F &= P[\text{Choose } H_1 \mid H_0 \text{ present}], \\ P_M &= P[\text{Choose } H_0 \mid H_1 \text{ present}]. \end{aligned}$$

The probability of detection error  $P_e$  is given by

$$P_e = p_0 P_F + p_1 P_M, \quad (6.2)$$

where  $p_0$  and  $p_1 = 1 - p_0$  are the prior probabilities of  $H_0$  and  $H_1$ , respectively. We define a likelihood ratio  $\Lambda(m)$  as,

$$\Lambda(m) = \frac{P[m \mid H_1]}{P[m \mid H_0]}, \quad (6.3)$$

where  $P[m \mid H_1]$  and  $P[m \mid H_0]$  denote the conditional probabilities of observing  $m(t)$  under the hypotheses  $H_1$  and  $H_0$ , respectively. Using Bayes' rule,  $\Lambda(m)$  can be expanded as

$$\Lambda(m) = \frac{P[H_1 \mid m] P[H_0]}{P[H_0 \mid m] P[H_1]}, \quad (6.4)$$

where  $P[H_1 \mid m]$  and  $P[H_0 \mid m]$  denote the posterior probabilities of the hypotheses conditioned on  $m(t)$ . The ratio  $\mathcal{L}(m) = P[H_1 \mid m]/P[H_0 \mid m]$  is commonly referred to as the posterior likelihood, whereas  $\mathcal{L}_0 = P[H_0]/P[H_1] = (1 - p_0)/p_0$  is called the prior likelihood. All the information needed to disambiguate between the two hypotheses using  $m(t)$  is contained in  $\mathcal{L}(m)$ .

The decision rule that minimizes  $P_e$  is given by Poor (1994),

$$\begin{aligned} &\text{Choose } H_1 \text{ for } \{m \mid \mathcal{L}(m) \geq 1\}, \\ &\text{Choose } H_0 \text{ for } \{m \mid \mathcal{L}(m) < 1\}, \end{aligned} \quad (6.5)$$

which can be compactly written as

$$\mathcal{L}(m) \underset{H_0}{\overset{H_1}{\geq}} 1 \quad \Rightarrow \quad \Lambda(m) \underset{H_0}{\overset{H_1}{\geq}} \mathcal{L}_0^{-1}. \quad (6.6)$$

If the noise  $n(t)$  is gaussian, the above decision rule reduces to convolving  $m(t)$  with a linear filter  $h_d(t)$  and comparing the sampled value of the filter output at  $t = T$  to a threshold. If the output exceeds the threshold, hypothesis  $H_1$  is chosen; otherwise  $H_0$  is chosen.  $h_d(t)$  is called a *matched filter* (Poor, 1994) and depends on the input signal  $s(t)$ , the filter  $g(t)$ , and the autocorrelation of noise  $n(t)$ . For finite  $T$ , deriving the exact form of  $h_d(t)$  involves solving an analytically intractable Fredholm integral equation (Helstrom, 1968) in general. However, in the limit  $T \rightarrow \infty$  (which means we can delay our decision indefinitely), we can derive a simple closed-form expression for  $h_d(t)$  in the frequency domain,

$$H_d(f) = \exp(-i2\pi f T) \frac{\mathcal{G}^*(f) S^*(f)}{S_n(f)}, \quad (6.7)$$

where  $\mathcal{G}(f) = \mathcal{F}\{g(t)\}$ ,  $S(f) = \mathcal{F}\{s(t)\}$ , and  $S_n(f)$  is the noise power spectral density. In our case, the measurement  $m(t)$  corresponds to the membrane voltage  $V(x, t)$ , the known signal  $s(t)$  corresponds to the EPSC waveform  $I_s(t) = A t e^{-t/t_{\text{peak}}}$ , the filter  $g(t)$  corresponds to the Green's function of the cable,  $g(x, y, t)/G$ , and  $S_n(f)$  corresponds to the noise component of the membrane voltage fluctuations  $S_V^n(x, f)$ .

Let us denote the sampled value of the output of the matched filter at  $t = T$  by the random variable  $r$ ,

$$r = (m(t) * h_d(t))(T) = \int_0^\infty dt m(t) h_d(-t). \quad (6.8)$$

Notice that  $r$  has the form of a correlation between the measurement  $m(t)$  and the time-reversed matched filter  $h_d(-t)$ . When  $S_n(f)$  has a flat spectrum ( $n(t)$  is band-limited white noise),  $h_d(t)$  is a shifted, time-reversed version of the excitatory postsynaptic potential (EPSP) at  $x$  in response to the EPSC  $I_s(t)$  at location  $y$ .  $r$  can be computed by correlating  $V(x, t)$  with the EPSP shape, which is given by  $g(x, y, t) * I_s(t)/G$ .

We can rewrite the optimal decision rule in equation 6.6 in terms of  $r$  as

$$r \underset{H_0}{\overset{H_1}{\geq}} Th, \quad (6.9)$$

where  $Th$  is the threshold chosen for optimal performance. Thus, the performance of the matched filter can be determined in terms of the statistical properties of the random variable  $r$ . Since  $h_d(t)$  is a linear filter,  $r$  is a gaussian random variable, and its conditional means and variances (under  $H_0$  and  $H_1$ ) specify it completely:

$$\begin{aligned}\mu_0 &= \langle r \mid H_0 \rangle, & \sigma_0 &= \langle r^2 \mid H_0 \rangle - \mu_0^2, \\ \mu_1 &= \langle r \mid H_1 \rangle, & \sigma_1 &= \langle r^2 \mid H_1 \rangle - \mu_1^2.\end{aligned}$$

It can be easily shown that

$$\mu_0 = 0; \quad \mu_1 = \int_{-\infty}^{\infty} df \frac{|N_{syn} \mathcal{G}(x, y, f) I_S(f)|^2}{G^2 S_V^n(f)}, \quad (6.10)$$

$$\sigma_0^2 = \sigma_1^2 = \sigma^2 = \int_{-\infty}^{\infty} df \frac{|N_{syn} \mathcal{G}(x, y, f) I_S(f)|^2}{G^2 S_V^n(f)}, \quad (6.11)$$

where  $I_S(f) = \mathcal{F}\{I_S(t)\}$  is the Fourier transform of the EPSC pulse and  $N_{syn}$  denotes the number of parallel synapses that are activated by a presynaptic action potential. Here we assume that the synaptic transmission is perfectly reliable and the synapses respond synchronously to the action potential. Thus, if there are  $N_{syn}$  synchronous synaptic connections between the dendrite and the presynaptic terminal, the current injected at the synaptic location due to a presynaptic action potential is scaled by a factor  $N_{syn}$ . (For an investigation of the information loss due to synaptic unreliability see Manwani & Koch, 1998.)

The error probabilities  $P_F$  and  $P_M$  can be computed as

$$P_F = \int_{Th}^{\infty} dr P[r \mid H_0] dr; \quad P_M = \int_{-\infty}^{Th} dr P[r \mid H_1]. \quad (6.12)$$

The optimal value of the threshold  $Th$  depends on the standard deviation  $\sigma$  and the prior probability  $p_0$ . However, for equiprobable hypotheses ( $p_0 = 1 - p_0 = 0.5$ ), the optimal threshold  $Th = (\mu_0 + \mu_1)/2 = \sigma^2/2$ . This gives

$$P_e = P_F = P_M = \frac{1}{2} \operatorname{Erfc}\left(\frac{\sigma}{2\sqrt{2}}\right). \quad (6.13)$$

The probability of error  $P_e$  ranges between  $P_e = 0$ , which implies perfect detection, and  $P_e = 0.5$ , which implies chance performance (pure guessing).  $P_e$  decreases monotonically as  $\sigma$  varies from  $\sigma = 0$  to  $\sigma = \infty$ . In the signal detection task,  $\sigma$  (equivalent to  $d'$  in psychophysics; (Green & Swets, 1966) plays the role that the quantity  $SNR$  does in the signal estimation task.

We can regard the overall decision system as an effective binary communication channel in the information-theoretical sense. We denote the input

and output of this channel by the binary random variables  $S$  and  $D$ , both of which assume values in the set  $\{H_0, H_1\}$ . The effective binary channel model corresponding to the detection task is shown in Figure 3C, with the errors  $P_F$  and  $P_M$  denoting the channel cross-over probabilities. In addition to  $P_e$ , the system performance can be assessed by computing the mutual information  $I(S; D)$  between  $S$  and  $D$ . For the binary detection channel,  $I(S; D)$  can be computed as in Cover & Thomas, 1991,

$$I(S; D) = \mathcal{H}(p_o(1 - P_M) + (1 - p_o)P_F) - p_o \mathcal{H}(P_M) - (1 - p_o) \mathcal{H}(P_F), \quad (6.14)$$

where  $\mathcal{H}(x)$  denotes the binary entropy function,

$$\mathcal{H}(x) = -[x \log_2(x) + (1 - x) \log_2(1 - x)], \quad 0 \leq x \leq 1. \quad (6.15)$$

For equiprobable hypotheses,

$$I(S; D) = 1 - \mathcal{H}(P_e) \text{ bits}. \quad (6.16)$$

Since  $S$  and  $D$  are binary random variables,  $0 \leq I(S; D) \leq 1$ . As before,  $I(S; D) = 1$  bit implies perfect detection with no information loss, whereas  $I(S; D) = 0$  implies chance performance.

## 7 Results

---

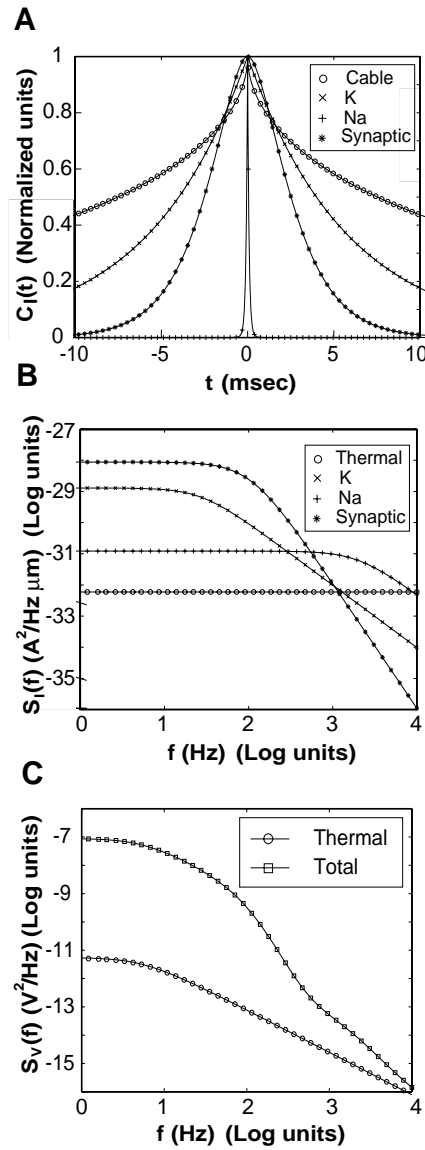
We now use the formalism developed above to assess the efficacy of information transfer in an infinite, 1D linear cable. As a first approximation, this can

---

Figure 4: *Facing page*. Membrane noise in dendritic cables. (A) Comparison of the normalized correlation functions  $C_I(t)/C_I(0)$  of the different noise sources with the autocorrelation of the Green's function of an infinite cable, for parameter values summarized below. (B) Comparison of current power spectra  $S_I(f)$  of the different membrane noise sources: thermal noise,  $K^+$  channel noise,  $Na^+$  channel noise, and synaptic background noise as a function of frequency (up to 10 kHz). (C) Voltage spectrum  $S_V(f)$  of the noise in a weakly active dendrite due to the influence of the above sources. Power spectrum of the voltage fluctuations due to thermal noise alone  $S_{V_{th}}(f)$  is also shown for comparison. Summary of the parameters adopted from Mainen & Sejnowski (1998) to mimic the apical dendrite of a layer V pyramidal neuron:  $R_m = 40 \text{ k}\Omega\text{cm}^2$ ,  $C_m = 0.75 \text{ }\mu\text{F}/\text{cm}^2$ ,  $r_i = 200 \text{ }\Omega\text{cm}$ ,  $d$  (dendritic diameter) =  $0.75 \text{ }\mu\text{m}$ ,  $\eta_K = 2.3$  channels per  $\mu\text{m}$ ,  $\eta_{Na} = 3$  channels per  $\mu\text{m}$ ,  $\eta_{syn} = 0.1$  synapse per  $\mu\text{m}$  with background activity modeled as a Poisson process with mean firing rate  $\lambda_n = 0.5 \text{ Hz}$ ,  $E_K = -95 \text{ mV}$ ,  $E_{Na} = 50 \text{ mV}$ ,  $E_{syn} = 0 \text{ mV}$ ,  $E_L = -70 \text{ mV}$ ,  $\gamma_K = \gamma_{Na} = 20 \text{ pS}$ . Refer to M-K for details.



be regarded as a model of a weakly active apical dendrite of a cortical pyramidal cell. Thus, the biophysical parameter values we shall use are obtained from the literature on pyramidal neuron models (Mainen & Sejnowski, 1998). In addition to estimating signal and noise magnitudes and studying the dependence of the different measures of signal fidelity ( $\xi, P_e, I$ ) on the electrotonic distance



X, we will also explore the effect of varying various biophysical parameters on these quantities.

**7.1 Noise in a Weakly Active Dendrite.** The membrane noise sources we consider are thermal noise (due to thermal agitation of charge carriers), channel noise (due to stochastic channel openings/closings of  $K^+$  and  $Na^+$  voltage-gated ionic channels), and synaptic noise (due to the spontaneous background firing activity of presynaptic neurons). A discussion of the origins of these noise sources and their characterization was carried out in M-K.

Here we only make use of the expressions of the power spectral densities of the current noise sources, referring the reader to M-K for details. For  $V_m \approx V_{rest}$ , power spectral densities of these channel noise sources ( $K^+$ ,  $Na^+$ ) are approximately Lorentzian ( $[1 + (f/f_c)^2]^{-1}$ ). When the EPSC is modeled as an  $\alpha$  function and the background activity assumed to be a homogeneous Poisson process, the power spectral density of the synaptic background noise is shaped like a double Lorentzian ( $[1 + (f/f_c)^2]^{-2}$ ).

Using biophysical values for the  $K^+$  and  $Na^+$  channel densities and kinetics, synaptic innervation density, EPSC parameters, and so on, obtained from the literature on the weakly active properties of apical neocortical dendrites (Mainen & Sejnowski, 1998) (parameter values are summarized in the caption of Figure 4), we computed the magnitudes of the different noise sources and quantified the corresponding voltage noise in a 1D infinite cable (see Figure 4). The normalized autocorrelation functions of the noise sources and the Green's function of the cable are compared in Figure 4A. Notice that the temporal spread of the noise sources is much smaller (except for  $K^+$  noise) than the Green's function of the cable. Thus, the noise spectra can be assumed to be approximately flat over the bandwidth of the cable, thereby justifying the white noise approximation. The noise spectra are compared in Figure 4B, and the standard deviations of the voltage noise  $\sigma_V$  due to different sources are compared in Table 1. For the parameter values considered, the magnitude of voltage fluctuations  $\sigma_V$  is on the order of 1.4 mV, which is small enough to justify the perturbative approximation. Thus, in general, the magnitude of the voltage fluctuations can be used to test the validity of the approximation. It can also be seen that synaptic background activity is the dominant source of membrane noise. Thermal noise is almost negligible (at least up to 1000 Hz) in comparison to the other sources. Experimentally these spectra can be computed by voltage clamping the dendrite around  $V_{rest}$  and using different pharmacological manipulations to isolate individual contributions, for example, TTX to eliminate  $Na^+$  noise TEA to eliminate  $K^+$  noise, and so on (Manwani, Segev, Yarom, & Koch, 1998). These spectra can then be compared with analytical expressions corresponding to different membrane noise sources (DeFelice, 1981; M-K).

The power spectral density of the voltage noise in an infinite cable due to these distributed sources (using equation 3.16) is shown in Figure 4C. The power spectral density of the contribution due to thermal noise alone is also shown alongside for comparison. Notice that the voltage noise spectrum is a monotonically decreasing function of frequency since the active membrane conductances

Table 1: Comparison of the Magnitudes of Voltage Noise Contributions Due to Different Membrane Noise Sources in Our Weakly Active Infinitely Long Dendrite.

Noise Type	$\sigma_V$
Thermal	0.012 mV
K <sup>+</sup>	0.459 mV
Na <sup>+</sup>	0.056 mV
Synaptic	1.316 mV
Total	1.395 mV

Note: For parameter values, see the Figure 4 caption.

are modeled as pure conductances. However, in general, the small-signal membrane impedance due to voltage and time-dependent conductances can exhibit resonance and give rise to bandpass voltage noise spectra (Koch, 1999).

**7.2 Signal Propagation in a Weakly Active Dendrite.** The filters responsible for shaping the synaptic input signal are scaled versions ( $1/G$ ) of the Green’s function of the infinite cable and are shown in Figure 5A. Notice how the filter gain and bandwidth change with distance. At small distances from the input location (since  $g(x, y, t)$  is symmetric, only the relative electrotonic distance  $X$  matters), the filter is sharply peaked and has a high gain. However, at larger distances, the filter becomes broader and has lower gain owing to the fact that some signal is lost due to leakage through the transmembrane resistance. The increase in temporal spread of the filter with distance is due to the increased capacitance that needs to be charged up as the measurement location moves farther away from the input location ( $X$  increases), causing the effective time constant of the filter to increase.

The voltage change due to a synaptic input (in the form of an EPSC pulse) is obtained by convolving the EPSC waveform (shown in the inset) with  $g(x, y, t)/G$ . The membrane voltage depolarizations (from  $V_{rest}$ ) due to the delivery of a unitary EPSC at different distances are shown in Figure 5B. The peak of the depolarization occurs at the synaptic location and is about 2.2 mV. Notice that at  $X = 0$ , the EPSP is almost identical in shape to the EPSC waveform, implying that the filtering due to the cable is minimal. However, at larger distances, the EPSP becomes smaller in magnitude, and its temporal spread increases. For both figures, distances are expressed in dimensionless electrotonic units, where  $\lambda$  is around 550  $\mu\text{m}$ .

We also examine the dependence of variance of the voltage fluctuations  $\sigma_V^2$  due to the injection of a random current input on the electrotonic distance  $X$ . The current  $I_s(t)$  is in the form of a gaussian random process of variance  $\sigma_s^2$ . Its power spectrum is assumed to be spectrally flat over a bandwidth  $B_s$  (see the inset of Figure 5C.) The standard deviations of the resulting voltage fluctuations  $\sigma_V$  as a function of  $X$ , for different values of  $B_s$ , are shown in Figure 5C. Notice that except for signals with small bandwidths (e.g., 10 Hz in Figure 5), where

the membrane voltage fluctuations might be strong enough to generate action potentials, our weakly active assumption is not violated for the most part. Thus, by measuring the magnitude of the resulting fluctuations for a set of biophysical parameters, one can easily verify the validity of our perturbative approximation on a case-by-case basis.

Like the peak of the EPSP above,  $\sigma_V$  also decreases monotonically with  $X$ , representative of the fact that the signal component of the voltage attenuates with distance from the input location. Since the cable acts like a low-pass filter, higher frequencies are transmitted less effectively, and so  $\sigma_V$  decreases with  $B_s$  (for a fixed  $\sigma_s$ ). This allows us to predict intuitively that the reconstructions of  $I_s(t)$  from  $V(t)$  should get poorer as  $B_s$  increases.

We are now equipped with all the information we need to estimate the information loss of the synaptic signal due to electrotonic attenuation and the membrane noise sources, under the two coding paradigms.

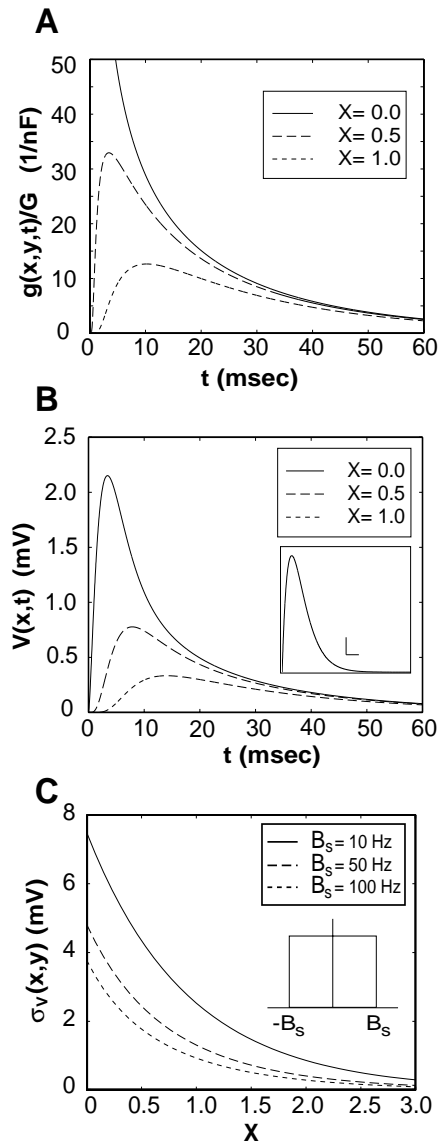
**7.3 Efficacy of Signal Estimation.** The effective communication channel corresponding to the estimation task is shown in Figure 3A. The channel input is the random current  $I_s(t)$ , and the channel output is the estimate  $\hat{I}_s(t)$ , obtained from  $V(x, t)$  after convolution with the optimal linear filter  $h(t)$ . The effective noise introduced by the channel is the difference,  $I_n(t) = \hat{I}_s(t) - I_s(t)$ . If we assume that  $I_s(t)$  is a gaussian process with variance  $\sigma_s^2$ , the channel reduces to the classical additive white band-limited gaussian noise channel (Cover & Thomas, 1991). It is straightforward to compute the mutual information and capacity for this channel model (see equation 5.22).

The coding fraction  $\xi$  and the mutual information  $I[I_s(t); V(t)]$  as functions of  $X$  are plotted in Figures 6A and 6B, respectively.  $\xi$  is close to one for short distances but falls rapidly as  $X$  increases because the signal attenuates with distance. Moreover, the rate of decay of  $\xi$  with respect to  $X$  depends on  $B_s$ . Additionally, if the signal-to-noise ratio is a monotonically decreasing function of frequency (equivalently, signal power spectrum decays faster than the noise spectrum),  $\xi$  also decreases with  $B_s$ . Similarly, the mutual information  $I$  decays monotonically with  $X$ . However its dependence on  $B_s$  is slightly more complicated; at small distances  $I$  increases with  $B_s$ , but this behavior reverses at larger distances.

---

Figure 5: *Facing page.* Signal propagation in dendritic cables. (A) Scaled version of the Green's function  $g(x, y, t)/G$  for an infinite linear cable corresponding to different electrotonic distances expressed in dimensionless ( $X = l/\lambda$ ) units. (B) Excitatory postsynaptic potentials in response to a unitary EPSC input (inset: calibration 1 pA, 2 msec) at different electrotonic distances from the synapse, obtained by convolving the EPSC with the filters in A. (C) The standard deviation of voltage fluctuations  $\sigma_V$  in response to a gaussian white band-limited current waveform  $I_s(t)$  of bandwidth  $B_s$  (inset: band-limited power spectrum of  $I_s(t)$ ) and standard deviation  $\sigma_s = 5$  pA plotted as a function of  $X$  for different values of  $B_s$ .

An intuitive explanation for this phenomenon is as follows. The mutual information  $I$  broadly depends on two quantities: the signal-to-noise ratio ( $SNR$ ) and the input bandwidth ( $B_s$ ). In general,  $SNR$  is a function of frequency, but for the moment let us assume that it is a frequency-independent constant. The expression for  $I$  in terms of  $SNR$  and  $B_s$  (a simplified version of equation 5.22) is



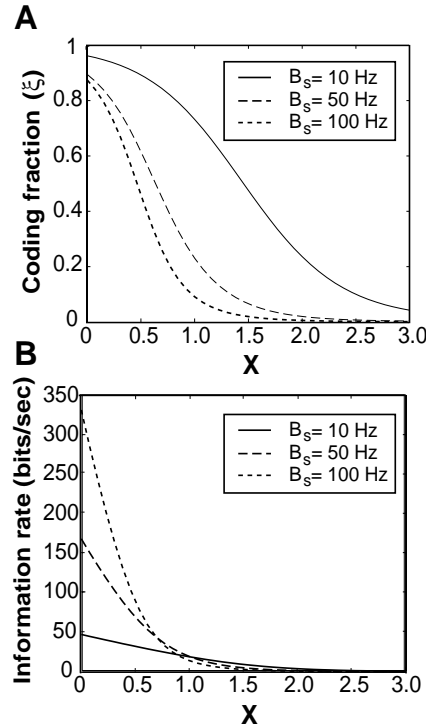


Figure 6: Efficacy of signal estimation. Coding fraction  $\xi$  (A) and mutual information  $I[I_s(t); V(t)]$  (B) for an infinite 1D cable as a function of electrotonic distance  $X$  from the input location for different values of the input bandwidth ( $\sigma_s = 5$  pA). Parameter values are identical to those in Figure 4.

given as  $I = B_s \log(1 + SNR)$ .  $SNR$  is inversely proportional to  $B_s$  ( $SNR = \kappa/B_s$ , where  $\kappa$  is the constant of proportionality) since for a fixed input power, if we increase  $B_s$ , the signal power per unit frequency (and thus  $SNR$ ) decreases. For small values of  $X$ , the signal power is possibly much larger than the noise power, and the  $SNR$  values for different  $B_s$  are large enough to lie in the saturating regime of the logarithm. Thus, for small  $X$ , the bandwidth component of the product ( $I = B_s \log(1 + SNR)$ ) dominates, and  $I$  increases with  $B_s$ . On the other hand, for large  $X$ , the magnitude of  $SNR$  is small, which implies  $B_s \log(1 + SNR) \approx B_s SNR = \kappa$ . Thus, one expects  $I$  to be independent of  $B_s$  for large  $X$ . This analysis is valid exactly when the  $SNR$  does not depend on  $f$  (signal and noise spectra vary with  $f$  in a similar manner), which is not true in our case since the signal and noise spectra have different shapes. In our case, for large  $X$ , the product is marginally larger for a lower value of  $B_s$  as opposed to a higher value. This causes the slight reversal in  $I$  for large  $X$ .

We also numerically compute the information capacity for signal estimation using the “water-filling” algorithm, maximizing  $I$  by choosing the optimal  $S_s(f)$  at each distance  $X$  (the procedure is illustrated in Figure 7). In reality this is biophysically unrealistic since the optimal  $S_s(f)$  depends on  $X$ —that is, the optimal signal power distribution is different for synaptic inputs received at different input locations. However, this allows us to compare performance for a particular choice of input spectra  $S_s(f)$  (white band-limited spectrum, in our case) against the best achievable performance. We find that for the parameter values we consider, the capacity  $C$  is not significantly different in magnitude from  $I$  computed using a white band-limited input spectrum.  $I$  is indistinguishable from  $C$  for small  $X$  (high SNR) and is not significantly different in absolute terms for large  $X$ . As an example, the maximum difference between  $C$  and  $I$  ( $\sigma_s = 5$  pA,  $B_s = 100$  Hz) is on the order of 8.5 bits per second for  $X \approx 1$ . However, the magnitudes of  $C$  and  $I$  for  $X \approx 1$  are about 22.4 bits per second and 13.9 bits per second, respectively, and so as a percentage, the capacity is about 60% higher.

**7.4 Efficacy of Signal Detection.** The effective binary communication channel corresponding to the detection task is shown in Figure 3C. The input to the channel is a random variable denoted by  $S$ , which corresponds to the binary nature of the presence or the absence of an EPSC. Since the goal in the detection task is to detect whether such an event occurred, the output of the channel corresponds to this binary decision denoted by  $D$ . The cross-over probabilities of this detection channel are given by  $P_F$  and  $P_M$ .

The probability of error  $P_e$  and the mutual information  $I(S; D)$  for the detection task are plotted in Figures 8A and 8B, respectively.  $P_e$  varies from  $P_e \approx 0$  (perfect detection) for  $X = 0$  to  $P_e = 0.5$  (pure guessing) as  $X \rightarrow \infty$ . Correspondingly,  $I(S; D)$  varies from 1 to 0 bits. We also vary the number of synchronous synapses  $N_{syn}$ , assumed to deliver EPSCs simultaneously in response to a presynaptic action potential. As can be seen from the figures, there is a critical distance before which an EPSC can be detected almost perfectly. However, once this threshold distance is exceeded, performance deteriorates considerably. This critical distance depends on the SNR of the detection task and increases with  $N_{syn}$ . This threshold behavior is due to the nonlinear threshold decision rule of the signal detection task.

Thus, we find that considerations of signal-to-noise limit the distance over which synaptic signals can be reliably transmitted in noisy, weaklyactive, dendritic cables. This is true for both paradigms we consider here, though the threshold behavior is more pronounced for detection.

**7.5 Comparing Cable Theory and Information Theory.** In order to analyze and characterize the role of neurons as information transmission and processing devices, we argue that the relevant metrics should not only be quantities like electrotonic length, attenuation of the EPSP peak or the charge delivered, and so on, which are motivated by physiology and an extensive application of cable theory over the past 40 years, but should also include information-

theoretical measures. As an application of our approach, we have considered different information-theoretical quantities like  $\xi$ ,  $P_e$ ,  $I$ , and so on and examined how they vary with  $X$ .

In order to contrast our approach with that of classical cable theory, we compared some of our metrics against physiologically relevant quantities. In Figure 9A, we plot the standard deviation of the voltage fluctuations  $\sigma_V$  in response to white band-limited noise injection as a function of  $X$  for different input bandwidths  $B_s$ . The standard deviations are normalized by their values at  $X = 0$  ( $\sigma_V(0)$ ) since their absolute values depend on  $B_s$ . The same procedure is carried out for the mutual information  $I[I_s(t); V(x, t)]$ , shown in Figure 9B. It is clear that for a given  $B_s$ ,  $I$  decays relatively faster with  $X$  than  $\sigma_V$ . Moreover, the rate of decay with respect to  $X$  depends on  $B_s$  and is higher for  $I$  than  $\sigma_V$ . Thus for small  $X$ , even though  $I$  is higher for higher bandwidths (as seen in Figure 6C), the rate of loss of information with distance is higher for signals with larger bandwidths. This can be intuitively expected since for large  $X$ , the cable bandwidth is small, and higher  $B_s$  signals have a greater portion of their power outside the bandwidth of the cable.

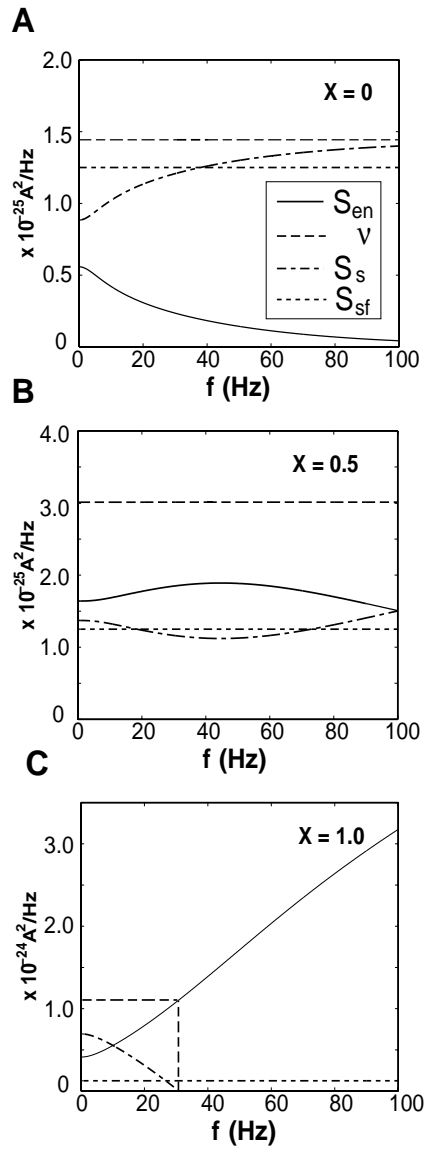
We also compared the mutual information  $I(S; D)$  in the binary detection task with the peak of the synaptic potential and the steady-state voltage attenuation ( $e^{-X}$ ) in response to DC current injection in Figure 9C. It is clear that for small distances,  $I(S; D)$  is almost constant even though the peak of the EPSP decays faster than  $e^{-X}$ . This is because the magnitude of the EPSP close to the postsy-

---

Figure 7: *Facing page*. Channel capacity using the water-filling algorithm. Graphical demonstration of the algorithm used to compute the channel capacity for the estimation task at three different electrotonic distances from the input location (see also Figure 3B). The solid line denotes the effective noise  $S_{en}(f)$ , the broadly dashed horizontal line represents the Lagrange multiplier  $1/\nu$  (see equation 5.28), the dot-dashed curve represents the optimal signal power spectrum that maximizes channel capacity  $S_s(f)$ , and the narrowly dashed line represents the flat band-limited spectrum  $S_{sf}(f)$  (see equation 5.16). (A) For  $X = 0$ ,  $S_{en}(f)$  is a low-pass spectrum, as there is negligible filtering of the input due to the cable. Correspondingly, the optimal  $S_s(f)$  is high pass and is nonzero over the entire available bandwidth ( $B_s = 100$  Hz) since there is sufficient input power available ( $\sigma_s = 5$  pA). In this case, both the channel capacity,  $C$ , and the mutual information,  $I$ , assuming that the input has a flat spectrum,  $S_{sf}$ , are equal to 328 bits per second. (B) For  $X = 0.5$  ( $C, I \approx 88$  bits per second), the bandpass nature of  $S_{en}(f)$  reflects attenuation (the effective noise level is much larger) and filtering of the input by the cable. The optimal  $S_s(f)$  has a complementary shape and is nonzero over the entire bandwidth. (C) For  $X = 1.0$  ( $C = 22.4$  bits per second,  $I = 13.9$  bits per second), the time constant of the cable filter is large, and signal power spectrum decays much faster than the noise spectrum.  $S_{en}(f)$  is high pass, and due to signal attenuation, the magnitude of the noise is large compared to  $\sigma_s^2$ , and equation 5.28 can only be satisfied over a limited portion of the available bandwidth. However,  $S_s(f)$  is much larger than  $S_{sf}(f)$  over this range (0–30 Hz). Parameter values are identical to those in Figure 4.



naptic location is large (around 2.2 mV at  $X = 0$ , Figure 5B) compared to the level of the ambient noise ( $\sigma_V = 1.395$  mV, Table 1) and can be detected almost perfectly. However, as soon as the EPSP becomes smaller than the noise, performance drops precipitously—much more steeply than the rate of decay of the peak postsynaptic potential. This threshold distance depends on the magnitude



of the EPSP at  $X = 0$  in comparison to the noise and is a measure of the SNR of the detection task. This threshold behavior is quite characteristic of nonlinear systems. The threshold nature of FM radio reception is a classic example of this phenomenon.

**7.6 Dependence on Biophysical Parameters.** There are several parameters in our analysis, and it is neither prudent nor necessary to consider the effect of varying all of them, in the multitude of different combinations possible, on the different variables of interest. Since the parameters belong to different equivalence classes (varying parameters within a class has the same effect on the variable of interest), it suffices to explore dependence with respect to the few abstract parameters characteristic of these classes instead of varying all the individual parameters. As a simple example, consider the expression for the steady-state synaptic conductance,

$$g_{Syn}^o = \eta_{Syn} \lambda_n g_{peak} e^{t_{peak}}, \quad (7.1)$$

where  $\eta_{Syn}$  is the synaptic density,  $\lambda_n$  is the background mean firing rate of presynaptic Poisson neurons,  $g_{peak}$  is the peak synaptic conductance of a unitary synaptic event (modeled by an  $\alpha$  function), and  $t_{peak}$  is the time when the peak is reached. Since  $g_{Syn}^o$  depends linearly on all the parameters in the product above, scaling the magnitude of any of the above parameters by a factor  $\eta$  causes  $g_{Syn}^o$  to be scaled by a corresponding factor  $\eta$ . Thus, these parameters belong to the same class (with respect to  $g_{Syn}^o$ ) and can be represented by an abstract scale factor  $\eta$ .

First we consider the effect of simultaneously varying different parameters on the resting properties of the dendrite:  $V_{rest}$ ,  $G$ ,  $\tau$ , and  $\lambda$ . We vary the abstract parameters corresponding to  $K^+$ ,  $Na^+$ , and synaptic conductances (except  $g_L$ ) by the same factor. We denote this scale parameter  $\eta$ . Thus,  $\eta = 0$  corresponds to a purely passive cable with only leak channels, whereas  $\eta = 1$  corresponds to the nominal values of the parameters, obtained from the literature, that we have used so far. The results of this exercise are summarized in Figure 10A. Instead of using absolute values for the quantities of interest, we normalize them with respect to their corresponding values at  $\eta = 0$ . Notice that  $V_{rest}$  changes (becomes more positive) by about 4%,  $\lambda$  changes (decreases) by about 9%, and  $\tau$  and  $G^{-1}$  change (decrease) by about 17%, as  $\eta$  is varied from 0 to 1. Despite the nonlinearities due to the active conductances  $K^+$  and  $Na^+$ , it is noteworthy that the quantities vary almost linearly with  $\eta$ . This further justifies our perturbative approximation.

The effects of parameter variation on the coding fraction  $\xi$  and the mutual information  $I[I_s(t); V(t)]$  are explored in Figures 10B and 10C, respectively. Here we allow parameters corresponding to the different noise sources to change individually ( $\eta$  goes from 0 to 1), while maintaining the others at their nominal values, in order to determine which noise source is dominant in determining performance. It is clear from the figures that the system performance is most sensitive to the synaptic noise parameters. The coding fraction  $\xi$  (for  $X = 0.18$ , corresponding to a distance of 100  $\mu m$  from the input location) drops from

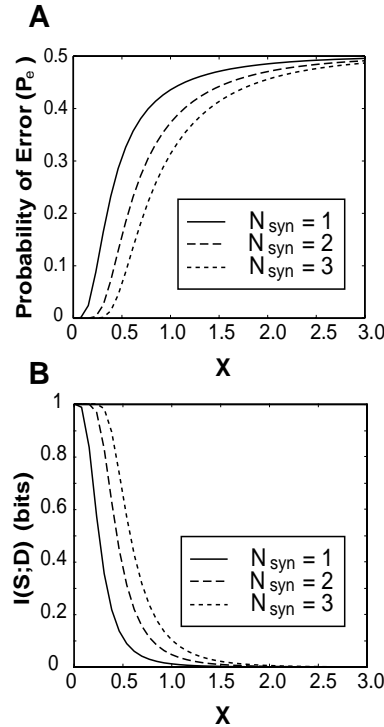


Figure 8: Efficacy of signal detection. Probability of error  $P_e$  (A) and mutual information  $I(S; D)$  (B) for an infinite cable as functions of the electrotonic distance  $X$  from the synaptic input. The number of synapses activated by a presynaptic action potential,  $N_{syn}$ , varies between one and three. The parameters associated with the EPSC are  $g_{peak} = 100$  pS,  $t_{peak} = 1.5$  msec, and  $E_{syn} = 0$  mV. Parameter values are identical to those in Figure 4.

around 0.96 in the absence of synaptic noise to around 0.78 when synaptic parameters are at their nominal values. This effect is even more dramatic for  $I$ , which drops from around 480 bits per second to around 225 bits per second. The sensitivity to parameters associated with potassium channels is small and is almost negligible for  $Na^+$  channel parameters.

## 8 Discussion

---

In this study, we investigated how neuronal membrane noise sources influence and ultimately limit the ability of one-dimensional dendritic cables to transmit information. In M-K, we characterized the dominant sources of membrane noise that could cause the loss of information as a signal spreads along neu-

ronal structures. By making the perturbative approximation that the conductances fluctuations (due to the noise sources) are small compared to the resting conductance of the membrane, we were able to derive a stochastic version of the cable equation satisfied by the membrane voltage fluctuations. We used this to derive analytical expressions for statistical properties of the voltage fluctuations (autocovariance, power spectrum) in weakly active dendrites in terms of the current noise spectra from M-K. Although we assumed a particular form for the sodium and potassium channel kinetics, our calculus can readily be adapted to investigate noise associated with any discrete-state Markov channel model.

We derived expressions for a few information-theoretical measures, quantifying the information loss under the estimation and detection paradigms. Earlier we made use of these paradigms to estimate the information capacity of an unreliable cortical synapse (Manwani & Koch, 1998). This study should be seen as a continuation of our efforts to understand the problem of neural coding in single neurons in terms of the distinct biophysical stages (synapse, dendritic tree, soma, axon and so on) constituting a neuronal link.

Our approach is different from some of the other paradigms addressing the problem of neural coding. Bialek and colleagues (Rieke et al., 1997) pioneered the reconstruction technique to quantify the information capacity and coding efficiency of spiking neurons and applied it to understand the nature of neural codes in various biological neural systems. Direct, model-independent methods to compute the information capacity of spiking neurons have also been developed recently (Deweese & Bialek, 1995; Stevens & Zador, 1996; Strong, Koberle, van Steveninck, & Bialek, 1998). In a more specific context, Zador, 1998 has investigated the influence of synaptic unreliability on the information transfer by spiking neurons.

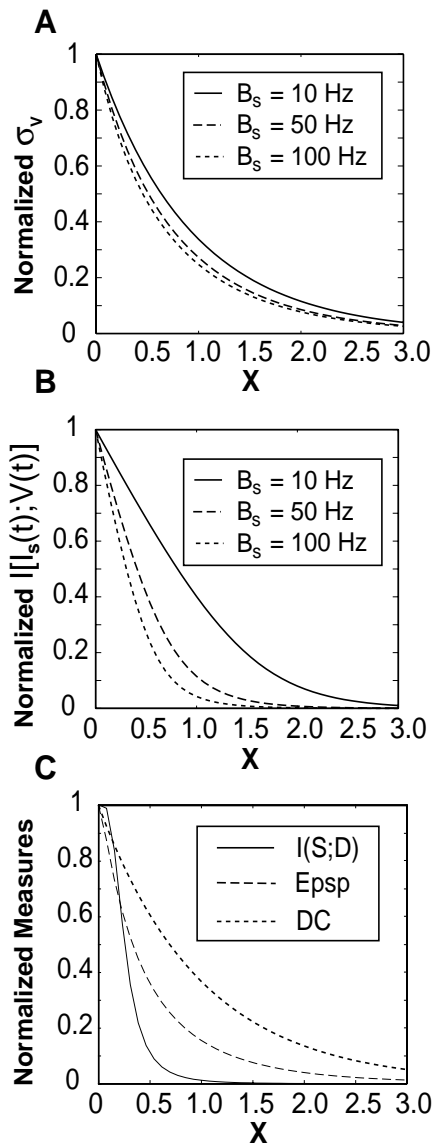
We are interested in deconstructing neuronal information transfer into its constituent biophysical components and assessing the role of each stage in this context rather than arriving at an accurate estimate of neuronal capacity. Ultimately our goal is to answer questions like, Is the length of the apical dendrite of a neocortical pyramidal cell limited by considerations of signal-to-noise? What influences the noise level in the dendritic tree of a real neuron endowed with voltage-dependent channels? How accurately can the time course of a synaptic signal be reconstructed from the voltage at the spike initiation zone? What is the channel capacity of an unreliable synapse onto a spine? and so on. Our

---

Figure 9: *Facing page*. Classical cable theory vs. information theory. Standard deviation of voltage fluctuations  $\sigma_V$  (A) and mutual information  $I[I_s(t); V(t)]$  (B) for the signal estimation paradigm as functions of the electrotonic distance  $X$  from the input location for different input bandwidths  $B_s$  ( $\sigma_s = 5$  pA). For ease of comparison, all curves are normalized with respect to their values at  $X = 0$ .  $I$  is much more sensitive to  $B_s$  than  $\sigma_V$ . (C) Comparison of the dependence of the normalized peak of the EPSP and the mutual information in the signal detection paradigm  $I(S; D)$  ( $N_{syn} = 1$ ) on  $X$ . The normalized steady-state electrotonic attenuation due to DC current injection is also shown. The detection performance is close to ideal for small  $X$ , but after a certain threshold distance, performance drops significantly.

research program is driven by the hypothesis that noise fundamentally limits the precision, speed, and accuracy of computation in the nervous system (Koch, 1999).

There exists a substantial experimental literature pertaining to the so-called lower envelope principle. It holds that the performance on psychophysical



threshold discrimination tasks is determined by single neurons (Parker & Newsome, 1998). The most dramatic illustration comes from recordings of single peripheral fibers in the median nerve of conscious human volunteers (Vallbo & Johansson, 1976; Vallbo, 1995). Remarkably, the occurrence of a single action potential in the fiber predicted the detection of the stimulus by the observer almost perfectly. This points to the functional utility for the system to be able to carry out a signal detection task of the type we study here.

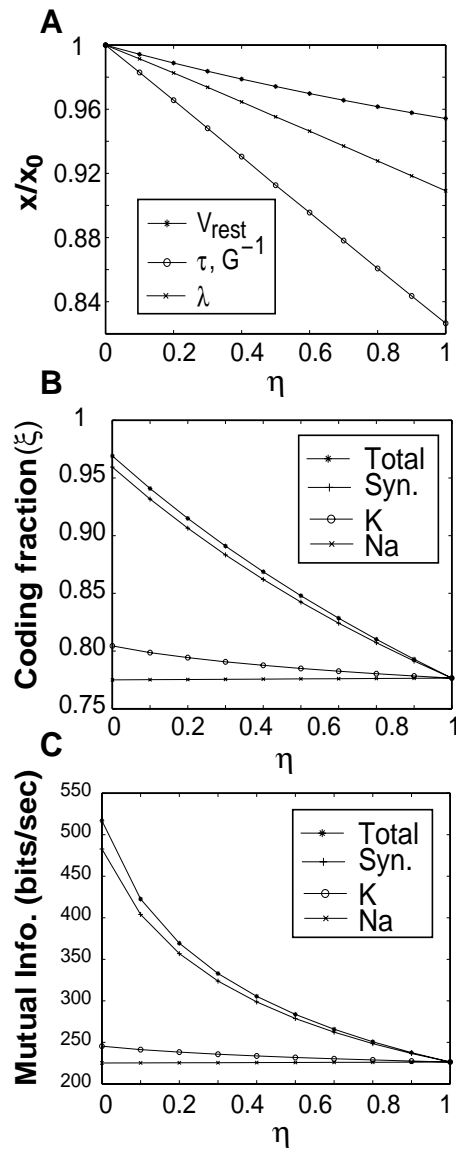
Our theoretical analyses for a simplified cable geometry reveal that signal transmission is indeed limited by considerations of signal-to-noise and that information cannot be transmitted passively along dendrites over long distances due to the presence of distributed membrane noise sources. Our argument needs to be qualified, however, since we still have to explore the effect of realistic dendritic geometries and neuronal parameters. Given the recent interest in determining the role of active channels in dendritic integration (Colbert & Johnston, 1996; Johnston et al., 1996; Yuste & Tank, 1996; Mainen & Sejnowski, 1998), it seems timely to apply an information-theoretical approach to study dendritic integration. The validity of our theoretical results needs to be assessed by comparison with experimental data from a well-characterized neurobiological system. We are currently engaged in such a quantitative comparison involving neocortical pyramidal neurons (Manwani et al., 1998).

Our analysis makes a strong point in favor of the presence of strongly active nonlinearities along apical dendrites for the sake of reliable information transfer. As evidenced in Figure 8, detecting the presence or absence of a synaptic signal more than roughly one space constant away becomes very difficult. While the various biophysical parameters used here need to be carefully compared against those of relevance to neocortical pyramidal cells, they do indicate that noise might limit the ability of extended apical dendrites to signal distal events reliably to the spike triggering zone and points out the need for “smart” amplifiers in the distal apical tuft that amplify the signal but not the noise (Bernander, Koch, & Douglas, 1994). Given the critical role of the apical dendrite in determining the thickness of the cortical sheet (Allman, 1990), it is possible that such noise consideration provided a fundamental constraint for the evolution of cortex.

---

Figure 10: *Facing page*. Influence of biophysical parameters. (A) Dependence of the passive membrane parameters ( $V_{rest}$ ,  $\tau$ ,  $\lambda$ ) on the channel and synaptic densities. The  $K^+$  and  $Na^+$  channel densities and the synaptic density are scaled by the same factor  $\eta$ , which varies from  $\eta = 0$ , corresponding to a completely passive system, to  $\eta = 1$ , which corresponds to the nominal weakly active parameter values used to generate Figure 4. The membrane parameters are expressed as a ratio of their values at  $\eta = 0$ . Effect of varying individual parameter values (the remaining parameters are maintained at their nominal values) on the coding fraction  $\xi$  (B) and the mutual information  $I$  (C) at a distance of  $X = 100 \mu\text{m}$  ( $X = 0.18$ ) from the input location. Thus, varying only the  $\eta$  associated with the synaptic background activity alone reduces both the coding fraction and the mutual information almost as much as changing the  $\eta$  associated with the synaptic and channel parameters.

Our analysis can readily be extended to deal with complicated dendritic geometries in a conceptually straightforward manner since we only require the Green's function corresponding to the geometry. Morphological reconstructions of biological neurons, followed by compartmental modeling, can be used to obtain realistic dendritic geometries. Analyzing dendritic morphologies using our



information-theoretical formalism will enable us to develop a graphical technique similar to the morphoelectrotonic transform (Zador, Agmon-Snir, & Segev, 1995), which will allow us to visualize the information transmission ability of the entire dendritic tree. Such a procedure requires the numerical computation of the Green's function between different locations along the dendritic tree and the soma. The expressions we have derived will allow us to quantify the information loss (in the detection/estimation paradigms) between the two locations. We believe that this procedure will provide an important graphical abstraction of the dendritic tree from an information-theoretical standpoint and is the subject of our ongoing efforts.

### Acknowledgments

---

This research was supported by NSF, NIMH, and the Sloan Center for Theoretical Neuroscience. We are grateful to the reviewers in helping us improve the quality of this article. We thank our collaborators, Peter Steinmetz and Miki London, for their invaluable suggestions and Idan Segev, Elad Schneidman, Yosef Yarom, Fabrizio Gabbiani, Andreas Andreou, and Pamela Abshire for illuminating discussions. We also acknowledge initial discussions with Bill Bialek and Tony Zador on the use of information theory to understand single-neuron biophysics.

### References

---

- Allman, J. (1990). Evolution of neocortex. In E. G. Jones & A. Peters (Eds.), *Cerebral cortex* (vol. 8A, pp. 269–283). New York: Plenum Press.
- Andreou, A. G., & Furth, P. M. (1998). An information-theoretic framework for comparing the bit-energy of signal representation at the circuit level. In E. S. Sinencio & A. G. Andreou (Eds.), *Low voltage, low power integrated circuits and systems*. New York: IEEE Press.
- Bernander, O., Koch, C., & Douglas, R. J. (1994). Amplification and linearization of distal synaptic input to cortical pyramidal cells. *J. Neurophysiol.*, 72(6), 2743–2753.
- Bialek, W., & Rieke, F. (1992). Reliability and information-transmission in spiking neurons. *Trends Neurosci.*, 15(11), 428–434.
- Bialek, W., Rieke, F., van Steveninck, R. R. D., & Warland, D. (1991). Reading a neural code. *Science*, 252(5014), 1854–1857.
- Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *J. Neurosci.*, 12, 4745–4765.
- Colbert, C. M., & Johnston, D. (1996). Axonal action-potential initiation and Na<sup>+</sup> channel densities in the soma and axon initial segment of subicular pyramidal neurons. *J. Neurosci.*, 16(21), 6676–6686.
- Courant, R., & Hilbert, D. (1989). *Methods of mathematical physics* (Vol. 1). New York: Wiley.
- Cover, T. M., & Thomas, J. A. (1991). *Elements of information theory*. New York: Wiley.



- DeFelice, L. J. (1981). *Introduction to membrane noise*. New York: Plenum Press.
- Deweese, M., & Bialek, W. (1995). Information-flow in sensory neurons. *Nuovo Cimento D*, 17, 733–741.
- Gabbiani, F. (1996). Coding of time-varying signals in spike trains of linear and half-wave rectifying neurons. *Network: Computation in Neural Systems*, 7(1), 61–85.
- Gabbiani, F., & Koch, C. (1998). Principles of spike train analysis. In C. Koch & I. Segev (Eds.), *Methods in neuronal modeling: From ions to networks* (2nd ed.). Cambridge, MA: MIT Press.
- Gabbiani, F., Metzner, W., Wessel, R., & Koch, C. (1996). From stimulus encoding to feature extraction in weakly electric fish. *Nature*, 384(6609), 564–567.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Helstrom, C. (1968). *Statistical theory of signal detection*. Oxford: Pergamon Press.
- Jack, J. J. B., Noble, D., & Tsien, R. (1975). *Electric current flow in excitable cells*. Oxford: Oxford University Press.
- Johnston, D., Magee, J. C., Colbert, C. M., & Cristie, B. R. (1996). Active properties of neuronal dendrites. *Ann. Rev. Neurosci.*, 19, 165–186.
- Johnston, D., & Wu, S. M. (1995). *Foundations of cellular neurophysiology*. Cambridge, MA: MIT Press.
- Koch, C. (1999). *Biophysics of computation: Information processing in single neurons*. New York: Oxford University Press.
- Magee, J., Hoffman, D., Colbert, C., & Johnston, D. (1998). Electrical and calcium signaling in dendrites of hippocampal pyramidal neurons. *Annu. Rev. Physiol.*, 60, 327–346.
- Mainen, Z. F., & Sejnowski, T. J. (1998). Modeling active dendritic processes in pyramidal neurons. In C. Koch & I. Segev (Eds.), *Methods in neuronal modeling* (2nd ed. pp. 171–210). Cambridge, MA: MIT Press.
- Manwani, A., & Koch, C. (1998). Synaptic transmission: An information-theoretic perspective. In M. Jordan, M. S. Kearns, & S. A. Solla (Eds.), *Advances in neural information processing systems*, 10 (pp. 201–207). Cambridge, MA: MIT Press.
- Manwani, A., Segev, I., Yarom, Y., & Koch, C. (1998). Neuronal noise sources in membrane patches and linear cables: An analytical and experimental study. *Soc. Neurosci. Abstr.*, p. 1813.
- Newsome, W. T., Britten, K. H., & Movshon, J. A. (1989). Neuronal correlates of a perceptual decision. *Nature*, 341, 52–54.
- Papoulis, A. (1991). *Probability, random variables, and stochastic processes*. New York: McGraw-Hill.
- Parker, A. J., & Newsome, W. T. (1998). Sense and the single neuron: Probing the physiology of perception. *Ann. Rev. Neurosci.*, 21, 227–277.
- Perkel, D. H., & Bullock, T. H. (1968). Neural coding. *Neurosci. Res. Prog. Sum.*, 3, 405–527.
- Poor, H. V. (1994). *An introduction to signal detection and estimation*. New York: Springer-Verlag.
- Rall, W. (1959). Branching dendritic trees and motoneuron membrane resistivity. *Exp. Neurol.*, 1, 491–527.

- Rall, W. (1960). Membrane potential transients and membrane time constant of motoneurons. *Exp. Neurol.*, 2, 503–532.
- Rall, W. (1967). Distinguishing theoretical synaptic potentials computed for different soma-dendritic distributions of synaptic input. *J. Neurophysiol.*, 30(5), 1138–1168.
- Rall, W. (1969a). Distributions of potential in cylindrical coordinates and time constants for a membrane cylinder. *Biophys. J.*, 9(12), 1509–1541.
- Rall, W. (1969b). Time constants and electrotonic length of membrane cylinders and neurons. *Biophys. J.*, 9(12), 1483–1508.
- Rall, W. (1989). Cable theory for dendritic neurons. In C. Koch & I. Segev (Eds.), *Methods in neuronal modeling: From synapses to networks* (pp. 9–62). Cambridge, MA: MIT Press.
- Rieke, F., Bodnar, D. A., & Bialek, W. (1995). Naturalistic stimuli increase the rate and efficiency of information-transmission by primary auditory afferents. *Proceedings of the Royal Society of London Series B Biological Sciences*, 262(1365), 259–265.
- Rieke, F., Warland, D., & Bialek, W. (1993). Coding efficiency and information rates in sensory neurons. *Europhysics Letters*, 22(2), 151–156.
- Rieke, F., Warland, D., van Steveninck, R. R. D., & Bialek, W. (1997). *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press.
- Segev, I., & Burke, R. E. (1998). Compartmental models of complex neurons. In C. Koch & I. Segev (Eds.), *Methods in neuronal modeling: From ions to networks* (2nd ed.). Cambridge, MA: MIT Press.
- Shadlen, M. N., & Newsome, W. T. (1998). The variable discharge of cortical neurons: Implications for connectivity, computation, and information coding. *J. Neurosci.*, 18(10), 3870–3896.
- Shannon, C. E. (1949). *A mathematical theory of communication*. Urbana, IL: University of Illinois Press.
- Stevens, C. F., & Zador, A. (1996). Information through a spiking neuron. In D. S. Touretzky, M. C. Mozer, & M. E. Hasselmo (Eds.), *Advances in neural information processing systems*, 8, Cambridge, MA: MIT Press.
- Strong, S. P., Koberle, R., van Steveninck, R. D. R., & Bialek, W. (1998). Entropy and information in neural spike trains. *Phys. Rev. Lett.*, 80(1), 197–200.
- Theunissen, F. E., & Miller, J. P. (1991). Representation of sensory information in the cricket cercal sensory system II: Information theoretic calculation of system accuracy and optimal tuning-curve widths of four primary interneurons. *J. Neurophysiol.*, 66(5), 1690–1703.
- Theunissen, F., & Miller, J. P. (1995). Temporal encoding in nervous systems: A rigorous definition. *J. Comput. Neurosci.*, 2(2), 149–162.
- Tuckwell, H. C. (1988a). *Introduction to theoretical neurobiology I: Linear cable theory and dendritic structure*. New York: Cambridge University Press.
- Tuckwell, H. C. (1988b). *Introduction to theoretical neurobiology II: Nonlinear and Stochastic Theories*. New York: Cambridge University Press.
- Tuckwell, H. C., & Walsh, J. B. (1983). Random currents through nerve membranes. I. Uniform Poisson or white noise current in one-dimensional cables. *Biol. Cybern.*, 49(2), 99–110.

- Vallbo, A. B. (1995). Single-afferent neurons and somatic sensation in humans. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 237–252). Cambridge, MA: MIT Press.
- Vallbo, A. B., & Johansson, R. S. (1976). Skin mechanoreceptors in the human hand: Neural and psychophysical thresholds. In Y. Zotterman (Ed.), *Sensory functions of the skin in primates*. Oxford: Pergamon.
- van Steveninck, R. D., & Bialek, W. (1988). Real-time performance of a movement-sensitive neuron in the blowfly visual system—Coding and information-transfer in short spike sequences. *Proceedings of the Royal Society of London Series B Biological Sciences*, 234(1277), 379–414.
- van Steveninck, R. D., & Bialek, W. (1995). Reliability and statistical efficiency of a blowfly movement-sensitive neuron. *Philosophical Transactions of the Royal Society of London Series B*, 348(1325), 321–340.
- Wan, F. Y., & Tuckwell, H. C. (1979). The response of a spatially distributed neuron to white noise current injection. *Biol. Cybern.*, 33(1), 39–55.
- Wessel, R., Koch, C., & Gabbiani, F. (1996). Coding of time-varying electric field amplitude modulations in a wave-type electric fish. *J. Neurophysiol.*, 75(6), 2280–2293.
- Wiener, N. (1949). *Extrapolation, interpolation and smoothing of stationary time series*. Cambridge, MA: MIT Press.
- Yuste, R., & Tank, D. W. (1996). Dendritic integration in mammalian neurons, a century after Cajal. *Neuron*, 16(4), 701–716.
- Zador, A. (1998). Impact of synaptic unreliability on the information transmitted by spiking neurons. *J. Neurophysiol.*, 79, 1219–1229.
- Zador, A. M., Agmon-Snir, H., & Segev, I. (1995). The morphoelectrotonic transform: A graphical approach to dendritic function. *J. Neurosci.*, 15(3), 1669–1682.

---

Received August 14, 1998; accepted March 15, 1999.