Adaptation with correlated noise predicts negative interval correlations in neuronal spike trains

Robin S. Sidhu[1], Erik C. Johnson[2], Douglas L. Jones[1], Rama Ratnam[3*],

1 Department of Electrical & Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, United States of America
2 The Johns Hopkins University Applied Physics Laboratory, Laurel, MD, United States of America
3 Division of Biological and Life Sciences, School of Arts and Sciences, Ahmedabad University, Ahmedabad, Gujarat, India

* rama.ratnam@ahduni.edu.in (RR)

Abstract

Negative correlations in the sequential evolution of interspike intervals (ISIs) are a signature of memory in neuronal spike-trains. They provide coding benefits including firing-rate stabilization, improved detectability of weak sensory signals, and enhanced transmission of information by improving signal-to-noise ratio. Here we predict observed ISI serial correlations from primary electrosensory afferents of weakly electric fish using an adaptive threshold model with a noisy spike threshold. We derive a general relationship between serial correlation coefficients (SCCs) and the autocorrelation function of added noise. Observed afferent spike-trains fall into two categories based on the pattern of SCCs: non-bursting units have negative SCCs which remain negative but decay to zero with increasing lags (Type I SCCs), and bursting units have oscillatory (alternating sign) SCCs which damp to zero with increasing lags (Type II SCCs). Type I SCCs are generated by low-pass filtering white noise before adding it to the spike threshold, whereas Type II SCCs are generated by high-pass filtering white noise. Thus, a single parameter (the sign of the pole of the filter) generates both types of SCCs. The filter pole (equivalently time-constant) is estimated from the observed SCCs. The predicted SCCs are in geometric progression. The theory predicts that the limiting sum of SCCs is -0.5, and this is confirmed from the expressions for the two types of filters. Observed SCCs from afferents have a limiting sum that is slightly larger at -0.475 ± 0.04 (mean $\pm$ s.d.). The theoretical limit of the sum of SCCs leads to a perfect DC block in the power spectrum of the spike-train, thereby maximizing signal-to-noise ratio during signal encoding. The experimentally observed sum of SCCs is just short of this limit. We conclude by discussing the results from the perspective of optimal coding.

Author summary

Many neurons spontaneously emit spikes (impulses) with a random time interval between successive spikes (interspike interval or ISI). The spike generation mechanism can have memory so that successive ISIs are dependent on one another and exhibit correlations. An ISI which is

longer than the mean ISI tends to be followed by an ISI which is shorter than the mean, and vice versa. Thus, adjacent ISIs are negatively correlated, and further these correlations can extend over multiple ISIs. A simple model describing negative correlations in ISIs is an adaptive threshold with noise added to the spike threshold. A neuron becomes more resistant (refractory) to spiking immediately after a spike is output, with refractoriness increasing as more spikes are spaced closer together. Refractoriness reduces as spikes are spaced further apart. We show that a neuron can generate experimentally observed patterns of correlations by relating it to the noise in the spike threshold. Two different types of filtered noise (low-pass and high-pass) generate the observed patterns of correlations. We show that the theoretical sum of the sequence of correlations has a limiting value which maximizes the information a neuron can transmit. The observed sum of correlations is close to this limit.

1 Introduction

The spiking activity of many neurons exhibit memory which stabilizes the neurons' firing rate and makes it less variable than a Poisson process. A signature of these memory effects can be found in the serial correlations of interspike intervals (ISIs) which display a prominent negative correlation between adjacent ISIs. This is a result of long intervals following short intervals so that fluctuations from the mean ISI are damped over long time-scales, thereby stabilizing the firing rate [1, 2]. Negative correlation between adjacent ISIs, which is the first serial correlation coefficient (ρ_1), can assume a range of values [3] from near-zero (close to a Poisson spike train, e.g., [4, 5]) to values close to -0.9 [1]. While more negative values may suggest a stronger memory effect, the relationship between the extent of memory in the spike train and their ISI correlations is by no means clear, in part due to the difficulty in determining joint interval distributions of arbitrarily high orders [1, 6].

Negative correlations in spike trains have been known for many years. Early reports came from lateral line units of Japanese eel [7], the retina of the cat [8], and subsequently in several neuronal systems [3–5, 9–15] (see [3, 16] for tabulations of negative correlations). In the active sense of some wave-type weakly electric fish [11, 13], primary electrosensory afferents exhibit the strongest known correlations in their adjacent ISIs [1]. These electrosensory neurons are an excellent model system for studying memory effects and regularity in firing due to their high spontaneous spike-rates [1, 6, 17]. It is likely that negative correlations have an adaptive significance, because they can facilitate optimal detection of weak electrosensory signals [1, 2, 11, 17–19] and maximize information transmission [20, 21]. The wider implications of negative correlations for neuronal processing are yet to be fully explored.

At a mechanistic level, negative correlations can result from increased refractoriness following an action potential, leading to an elevation in firing threshold. This was historically called "accommodation" [22], and can be attributed to a decaying or adaptive threshold following a spike which relaxes back to its normative value without reset [17, 23–31]. Thus, if two spikes are output in quick succession with an interval smaller than the mean ISI, the upward threshold shift will be cumulative, making the membrane more refractory, and so a third spike in the sequence will occur with an ISI that is likely to be longer than the mean. In this way, the time-varying adaptive threshold serves as a moving barrier, carrying with it a memory of prior spiking activity.

Several reports have suggested that an adaptive threshold may have a biophysical basis in outward potassium currents, including voltage gated potassium currents and calcium-dependent or AHP currents (e.g., [5, 30, 32–36]). An adaptive threshold model is known to accurately

predict spike-times in neurons ranging from cortical [37–39] to peripheral [39]. Further, we recently showed that a spike-timing code based on an adaptive threshold is an optimal neural code which minimizes coding error (estimation error) for a fixed long-term spike-rate (energy consumption) [39–41]. Thus, adaptation with its signature serial interval correlations, are likely tied to fundamental mechanisms responsible for a timing-based neural code.

Adaptive threshold models are deterministic. To make them stochastic, colored noise or Gaussian noise (or both) are added to the threshold or to the input signal so that negative correlations can be observed [17, 28, 33, 42]. In these models the first SCC ρ_1 (between adjacent ISIs) is close to or equal to -0.5, and all remaining correlation coefficients ρ_i, $i \geq 2$, are identically zero. Experimental spike trains demonstrate broader trends, where ρ_1 can assume values smaller or greater than -0.5, and the remaining SCCs can be non-zero for several lags, sometimes with damped oscillations and sometimes monotonically increasing to zero [1]. Lindner and co-workers [30, 31] developed a multi-parameter adaptive threshold model with additive colored and Gaussian noise to introduce noise fluctuations of different time-scales (slow and fast, respectively). They showed that the time-scale of noise fluctuations determined the various patterns of SCCs, including positive correlations. All of these patterns had a geometric structure (i.e., $\rho_k/\rho_{k-1} = \text{constant}$). Urdapilleta [43] also obtained a geometric structure with monotonically decaying SCC pattern with $\rho_k < 0$. These latter studies show that the role of noise fluctuations, in particular the time-scale of fluctuations, are important in determining patterns of SCCs. Thus, a model that can accurately predict experimentally observed SCCs for all lags has the potential to isolate mechanisms responsible for ISI fluctuations and negative correlations, and provide insights into neural coding. This is the goal of the current work.

Serial interspike interval correlations observed in primary P-type afferents of weakly electric fish *Apteronotus leptorhynchus* are modeled using an adaptive threshold with a noisy firing threshold. The autocorrelation function of the noise determines the SCCs of the ISIs. The form of the autocorrelation function determines the value of ρ_1 and its deviation from -0.5, and determines subsequent ρ_k. Theoretically, the limiting sum of the SCCs is -0.5, and experimental SCCs are close to this sum. Explicit closed-form relations between the correlation function and SCCs are provided. This model is parsimonious, and in addition to predicting spike-times as shown earlier, it reproduces observed ISI SCCs.

2 Results

2.1 Experimental results

Serial correlation coefficients (SCCs) were estimated from the baseline activity of 52 P-type primary electrosensory afferents in the weakly electric fish *Apteronotus leptorhynchus* (see Materials and Methods). SCCs from two example afferents (Fig. 1) demonstrate the patterns of negative SCCs observed in spike trains, and motivate this work. Statistical properties of these and other spike trains were reported earlier [1], with a qualitative description of SCCs and some descriptive statistics. A complete analytical treatment is undertaken here. Two types of serial interspike interval statistics can be identified according to the value taken by ρ_1 (the first SCC, between adjacent ISIs).

1. Type I: $-0.5 < \rho_1 < 0$. Subsequent ρ_k are negative and diminish to 0 with increasing k (Fig. 1A). The ISIs of these afferents are unimodal (shown later) and their spike trains do not exhibit bursting activity.

2. Type II: $\rho_1 < -0.5$. Subsequent ρ_k alternate between positive (ρ_{2k}) and negative (ρ_{2k+1}) values, are progressively damped, and diminish to zero (Fig. 1C). The ISIs of these spike trains are bimodal with a prominent peak at an ISI equal to about one period of the electric organ discharge (EOD) (shown later). These spike trains exhibit strong bursting.

Afferent fibers sampled from individual fish were a mix of Types I and II. Additionally we identify a third type that has not been observed in experimental data (at least by these authors) but is fairly common in some adaptive threshold models (e.g., [17, 28], see also Discussion). We call this Type III, and for this pattern of SCCs $\rho_1 = -0.5$ and subsequent ρ_k are identically zero (Fig. 1). The Type III pattern is a singleton (i.e., there is only one SCC pattern in this class).

SCCs for observed Type I and Type II units ρ_k ($k \geq 1$) sum to approximately -0.5. For Type III, the sum is exactly -0.5.

Figure 1. Representative serial interspike interval correlations (SCCs) from two P-type primary electrosensory afferent spike trains from two different fish. Panels depict patterns of correlation observed in the data. A. Type I ($\rho_1 > -0.5$): Non-bursting unit with first SCC $\rho_1 = -0.36$. Remaining $\rho_k < 0$ diminish to 0. Sum of SCCs (Σ) over 15 lags is -0.48. B. Type II ($\rho_1 < -0.5$): Strongly bursting unit with $\rho_1 = -0.7$ with marked alternating positive and negative correlations. Σ over 15 lags is -0.49. Spike trains sampled at 60 μs. Mean ISI $\pm$ SD (in ms): 2.42 ± 0.72 (A), and 6.04 ± 3.59 (B). Electric organ discharge (EOD) frequency: 948 Hz (A), and 990 Hz (B).

The baseline spike-train statistics of 52 P-type afferents reported earlier [1] were analyzed in detail in this work. All observed units showed a negative correlation between succeeding ISIs ($\rho_1 < 0$) (Fig. 2A). Experimental spike-trains demonstrated an average $\bar{\rho}_1 = -0.52 \pm 0.14$ (mean $\pm$ s.d.) ($N = 52$). Nearly half of these fibers had $\rho_1 > -0.5$ ($N = 24$) while the remaining fibers had $\rho_1 < -0.5$ ($N = 28$). The second SCC (ρ_2) between every other ISI (Fig. 2B) assumed positive ($N = 36$) or negative ($N = 16$) values with an average of $\bar{\rho}_2 = 0.10 \pm 0.18$. The value assumed by ρ_2 is linearly dependent on ρ_1 with a positive ρ_2 more likely if $\rho_1 < -0.5$ (Fig. 2C). The linear relationship is described by the equation $\rho_2 = -1.18\rho_1 - 0.51$, with a standard error (SE) of 0.008. This is close to the line $\rho_2 = -\rho_1 - 0.5$, the significance of which is discussed further below. Finally, the sum of the SCCs for each fiber taken over the first fifteen lags (excluding zeroth lag) is $\sum_k \rho_k = -0.475 \pm 0.04$, a number which, for most fibers, falls just short of -0.5 (Fig. 2D).

Figure 2. Population summaries of interspike interval (ISI) serial correlation coefficients (SCCs) for P-type primary afferent spike-trains ($N = 52$) [1]. A. Histogram of SCC between adjacent ISIs (ρ_1), $\bar{\rho}_1 = -0.52 \pm 0.14$ (mean $\pm$ s.d.), range -0.82 to -0.25 B. Histogram of SCC between every other ISI (ρ_2), $\bar{\rho}_2 = 0.10 \pm 0.18$, range -0.18 to 0.57. C. Anti-diagonal relationship between observed ρ_1 (abscissa) and ρ_2 (ordinate) (filled circles). Straight line describes best fit, $\rho_2 = -1.18\rho_1 - 0.51$ with SE $= 0.008$. D. Mean sum of SCCs for the population over 15 lags, $\sum \rho_k = -0.475 \pm 0.04$. The population histograms in panels A and B were reported earlier with a different bin width (see [1], Figs. 7A and 7B therein, respectively).

2.2 Deterministic adaptive threshold model

In the simplest form of the adaptive threshold model [39], the spike-initiation threshold is a dynamic variable governed by two time-varying functions: the sub-threshold membrane potential $v(t)$ and a time-varying or dynamic threshold $r\,(t)$. In the sub-threshold regime $v(t) < r\,(t)$. Most models (for e.g., [37]) assume that a spike is emitted when $v(t)$ exceeds $r\,(t)$ from below. To induce memory, the dynamic threshold is never reset (Fig. 1A) but suffers a jump in value whenever a spike is generated and then gradually relaxes to its quiescent value. This "jump and decay" is a fixed function which is called the adaptive threshold, and it usually takes the form $h\,(t) = A \exp\,(-t/\tau)$ where A is the instantaneous jump in threshold following a spike at $t = 0$, and τ is the time-constant of relaxation. Between two spikes $t_k < t \leq t_{k+1}$ the dynamic threshold is $r(t_k^-) + h(t - t_k)$ where $r(t_k^-)$ is the value assumed immediately before the spike at $t = t_k$. It captures the sum over the history of spiking activity.

Adaptive threshold models in the literature typically integrate the input using a low-pass filter (i.e., pass it through a leaky integrator) or a perfect integrator and then reset the membrane voltage to a constant v_r following a spike (e.g., [17,31]). Some leaky integrators have been assumed to be non-resetting (e.g., [37]). In the form of the model considered in this work and earlier [39–41] the voltage $v(t)$ is not integrated (i.e., there is no filtering), and it is not reset following a spike. These assumptions remove the effects of filtering in the soma and dendrites, and remove all the nonlinearities except for the spike-generator which we assume generates a sequence of Dirac-delta functions.

For modeling spontaneous activity we consider a steady DC-level bias voltage $v > 0$. In its most general form, a spike is fired whenever $v - r\,(t) \;= \gamma$ where γ is a constant spike-threshold (Fig. 3A). Historically, and in much of the literature $\gamma = 0$ (for e.g., [17,31,37]). We use the more general form with a constant, non-zero γ. The specific value assumed by γ plays a role in optimal estimation of the stimulus [39,41] but does not influence serial correlation coefficients. This is addressed in the discussion. The major advantage of these simplifications is that they allow us to focus on the role of the adaptive threshold element $h\,(t)$ in generating SCCs.

Figure 3. Schematic of neuron with adaptive threshold coding. A. v is a constant bias voltage that generates spontaneous activity, $r\,(t)$ is an adaptive threshold, and γ is a spike threshold such that a spike is emitted when $v - r\,(t) = \gamma$. Following a spike, the adaptive threshold suffers a jump of fixed magnitude A. The membrane voltage is non-resetting. Spike times are t_0, t_1, t_2, ..., etc. Historically, the spike threshold was set to zero, i.e., $\gamma = 0$ and a spike is fired when $v = r\,(t)$. Here, the more general form with non-zero γ is considered for reasons entered in the discussion. B. The spike encoder with adaptive threshold can be viewed as a feedback control loop where the spike-train $\delta\,(t - t_i)$ is filtered by $h\,(t)$ to generate the time-varying adaptive threshold $r\,(t)$ which can be viewed as an estimate of the membrane voltage $v(t)$. The comparator generates the estimation error $e(t) = v - r\,(t)$ (see also Panel A). The estimation or coding error drives the spike generator, and a spike is fired whenever $e(t) \geq \gamma$. The simplest form for the adaptive threshold (the estimation filter) $h\,(t)$ is a low-pass filter (an RC membrane) given by $h\,(t) = A \exp\,(-t/\tau)$. This is the form shown in panel A.

To make explicit the presence of memory, we note that the condition for firing the kth spike at t_k is met when

$$v - r\,(t_k) = v - (v - \gamma + A) \sum_{i=1}^{k-1} \exp\left(-\frac{t_k - t_i}{\tau}\right) = \gamma\,. \tag{1}$$

It is evident that memory builds up because of a summation over history of spiking activity. In the literature, a stochastic extension of this model is usually $v - r(t) + w(t) = 0$, where w is independent Gaussian noise.

The spike encoder with adaptive threshold implicitly incorporates a feedback loop (Fig. 3B) and so a different view of the above model is to think of the dynamic threshold $r(t)$ as an ongoing estimate of the membrane voltage $v(t)$. Here, the adaptive threshold $h(t) = A\exp(-t/\tau)$ acts as a linear low-pass filter. It filters the spike train to form an ongoing estimate $r(t)$ of the voltage $v(t)$. The instantaneous error in the estimate (the coding error) is then $e(t) = v(t) - r(t)$ (Fig. 3B). When the error exceeds γ a spike is output and the threshold increases by A to reduce the error. The time-varying adaptive threshold tracks $v(t)$ much like a home thermostat tracks a temperature set-point (Fig. 3A). Viewed in this way, it is the estimation error, and not the signal $v(t)$, which directly drives the spike generator and determines when the next spike should be generated (Fig. 3B).

From Fig 3A we can approximate the exponential with a piece-wise linear equation when the ISI is small. If t_{i-1} and t_i, $i \geq 1$, are successive spike-times (Fig. 3B), then the time evolution of the adaptive threshold $r(t)$ is given by

$$\begin{aligned} r(t) &= (v - \gamma + A)\exp(-\frac{t - t_{i-1}}{\tau}), \quad t_{i-1} < t \leq t_i\,, \\ &= (v - \gamma + A)(1 - \frac{t - t_{i-1}}{\tau}) + \mathrm{O}((t - t_{i-1})^2)\,. \end{aligned} \tag{2}$$

Note that v, γ, A, τ are constant, and so we can define $m = (v - \gamma + A)/\tau$ so that the slope of the decaying adaptive threshold is $-m$. The ISI can be obtained directly as (see Materials and Methods for details)

$$t_i - t_{i-1} = \frac{A\tau}{v - \gamma + A} = \frac{A}{m} = \text{Constant}\,, \tag{3}$$

which is the deterministic firing rule for a spike generator with a constant, DC-level input voltage.

2.3 Stochastic extension of adaptive threshold

In the schematic shown in Fig. 3B, noise injected in the body of the feedback loop will reverberate around the loop and generate memory. Here we consider a noisy threshold γ. Subsequently we will provide some results for a noisy time-constant τ in the adaptive threshold element. Fig. 4 depicts the stochastic adaptive threshold model where the spike threshold (γ) is a stochastic process. All other aspects of the model are unchanged from the deterministic model (Fig. 3). Let γ be a discrete wide-sense stationary process with mean $\mathbf{E}[\gamma]$, discrete auto-correlation function $\mathbf{R}(k)$ and power $\mathbf{E}\left[\gamma^2\right] = \mathbf{R}(0)$. The spike threshold with additive noise assumes the random value γ_i, $i \geq 1$ immediately after the $(i-1)^{\text{th}}$ spike and remains constant in the time interval $t_{i-1} < t \leq t_i$ (Fig 4A) [21,44]. Thus, the ith spike is emitted when the error satisfies the condition

$$e(t_i) \geq \gamma_i, \quad \text{for}\, t > t_{i-1}. \tag{4}$$

Subsequently the adaptive threshold jumps to a higher value specified by $v - \gamma_i + A$, and the noisy spike threshold assumes a new value γ_{i+1}. From Fig. 4A proceeding as before

$$t_i - t_{i-1} = \frac{1}{m}\left\{\gamma_i - \gamma_{i-1} + A\right\}. \tag{5}$$

The mean ISI is therefore $\mathbf{E}\left[t_i - t_{i-1}\right] = A/m$ as in the deterministic case given by Eq. 3. From the assumption of wide-sense stationarity the auto-correlation function $\mathbf{E}\left[\gamma_i\,\gamma_j\right]$ can be written as $\mathbf{R}\left(t_i - t_j\right)$. This is a discrete auto-correlation function whose discrete nature is made clear in the following way (see also Materials and Methods). Denote the mean of the kth-order interval by $T_k = \mathbf{E}[t_{i+k} - t_i]$, then the mean ISI is T_1 $(= A/m)$, and further $T_k = kT_1$. A realization of the random variable γ is generated once every ISI and thus, $\mathbf{R}$ takes discrete values at successive multiples of the mean ISI, i.e., $\mathbf{R}\left(T_1\right)$, $\mathbf{R}\left(T_2\right)$, etc., and will be denoted by $\mathbf{R}\left(1\right)$, $\mathbf{R}\left(2\right)$, etc., respectively. As noted before, $\mathbf{R}\left(0\right)$ is noise power. That is, we can write

$$\mathbf{R}\left(t_{i+k} - t_i\right) = \mathbf{R}\left(T_k\right) = \mathbf{R}\left(k\right)\ , \tag{6}$$

where k is the number of intervening ISIs. The general formula for serial correlation coefficients at lag k is well-known [45], and it is

$$\rho_k = \frac{\mathrm{Cov}\left[\left(t_i - t_{i-1}\right),\left(t_{i+k} - t_{i+k-1}\right)\right]}{\mathrm{Cov}\left[\left(t_i - t_{i-1}\right)\ \left(t_i - t_{i-1}\right)\right]^{1/2}\mathrm{Cov}\left[\left(t_{i+k} - t_{i+k-1}\right),\left(t_{i+k} - t_{i+k-1}\right)\right]^{1/2}}\,. \tag{7}$$

For a wide-sense stationary process the covariances are constant and so the subscript i can be dropped. From the relations and definitions in Eqs. 5, 6, and 7, and after some routine algebra (see Materials and Methods), we obtain

$$\rho_0 = 1\,, \tag{8}$$

$$\rho_k = -\frac{\mathbf{R}\left(k-1\right) - 2\,\mathbf{R}\left(k\right) + \mathbf{R}\left(k+1\right)}{2\left(\mathbf{R}\left(0\right) - \mathbf{R}\left(1\right)\right)}\,, \qquad k \geq 1\,. \tag{9}$$

Further below we determine the $\mathbf{R}\left(k\right)$ from experimental data. The serial-correlation coefficients given by Eqs. 8 and 9 are independent of the slope m of the decay rate of the adaptive threshold, and its gain A. Thus, for a constant input the observed correlation structure of the spike-train is determined solely by the noise added to the deterministic firing threshold γ.

Figure 4. Adaptive threshold neuron with a stochastic firing threshold. Panel descriptions follow Fig. 3 and only differences are noted. A. The spike threshold is $\gamma_i = v - r\left(t\right)$ where γ_i is a random value generated at t_{i-1}^{+}, and held constant until the next spike at t_i. The discrete noise sequence γ_i is generated by a discrete wide-sense stationary process with mean γ and unknown discrete auto-correlation function $\mathbf{R}\left(k\right)$. The goal is to determine $\mathbf{R}\left(k\right)$ which will generate a prescribed sequence of ISI serial correlation coefficients ρ_k. To reduce clutter, the spike threshold $v - \gamma_i$ is depicted as γ_i. B. Block diagram showing the stochastic modification of the spike threshold. Symbols and additional description as in text and Fig. 3.

2.3.1 Limiting sum of SCCs and power spectra

We make the assumption that the process γ is aperiodic and the auto-correlation function $\mathbf{R}\left(N\right) \to 0$ when $N \to \infty$. That is, the noise process decorrelates over long time-scales. From this assumption and Eq. 9 it follows that (see Materials and Methods)

$$\sum_{k=1}^{\infty} \rho_k = -\frac{1}{2}\,. \tag{10}$$

This is the limiting sum of ISI serial correlation coefficients. Let the mean and variance of ISIs be denoted by T_1 and V_1, respectively, and the coefficient of variation of ISIs as $C = \sqrt{V_1}/T_1$. If $P(\omega)$ is the power spectrum of the spike train, then the DC-component of the power is given by [45]

$$\lim_{\omega \to 0} P(\omega) = \frac{C^2}{2\pi T_1} \left(1 + 2 \sum_{k=1}^{\infty} \rho_i \right). \tag{11}$$

Introducing Eq. 10 into Eq. 11, we obtain

$$\lim_{\omega \to 0} P(\omega) = 0\,, \tag{12}$$

yielding a perfect DC block.

2.3.2 Type I serial correlation coefficients

Type I afferent spiking activity demonstrates serial correlations where $-0.5 < \rho_1 < 0$, and subsequent ρ_k are negative and diminish to 0 with increasing k (Fig. 1A). These spike trains have a unimodal ISI and do not display bursting activity. They can be generated by a process where threshold noise is low-pass filtered white noise (a Gauss-Markov or Ornstein-Uhlenbeck process).

Consider a first-order low-pass filter whose time-constant is $\tau_\gamma = RC$, where R is the resistance and C the capacitance. Denote $a = \exp(-T_1/\tau_\gamma) < 1$, where T_1 is the mean ISI as before. The impulse response of the filter in discrete-time is $h(n) = a^n\, u(n)$ where $u(n)$ is the Heaviside function, and n denotes integer multiples of T_1. Then it can be shown (see Materials and Methods)

$$\mathbf{R}(k) = \frac{a^k}{1 - a^2}\,. \tag{13}$$

Eqs. 9 and 13 yield the geometric series

$$\rho_k = -\frac{a^{k-1}(1-a)}{2}\,, \quad k \geq 1\,. \tag{14}$$

From Eq. 14, and noting that $0 < a < 1$, we conclude that $\rho_1 > -0.5$, $\rho_k < 0$ for all k, $|\rho_{k-1}| > |\rho_k|$, and $\rho_k \to 0$ as $k \to \infty$. This is the observed Type I pattern of SCCs. Further, summing the geometric series Eq. 14 yields $\sum_{k \geq 1} \rho_k = -0.5$ as stated in Eq. 10. The low-pass filter parameter a can be estimated from experimentally determined SCCs. From Eq. 14 this is

$$a = 1 + 2\rho_1\,. \tag{15}$$

In practice, a better estimate is often obtained from the ratio $a = \rho_2/\rho_1$.

We model a Type I P-type primary electrosensory afferent depicted in Fig. 1A using a noisy threshold with (low-pass) correlation function $\mathbf{R}(k)$ specified by Eq. 13. The correlation function (Fig. 5A) and adaptive threshold parameters were determined from the experimental data, and tuned so that they matched the afferent SCCs, ISI, and joint distributions (Fig. 6). The top row (Figs. 6A-C) depicts the ISI distribution, joint ISI distribution, and the serial correlation coefficients, respectively. Type I spike trains do not display bursting activity and their ISI distribution is unimodal. The bottom row (Figs. 6D-F) shows data from a matched model using low-pass filtered noise. SCCs of adjacent ISIs ρ_1 are -0.39 (data) and -0.34 (model). The observed pattern of Type I SCCs is reproduced.

Figure 5. Normalized auto-correlation functions of threshold noise. A. Type I SCC matched to the afferent depicted in Fig. 1A with no bursting activity. B. Type II SCC from an afferent with moderate bursting activity. C. Type II SCCs matched to the afferent depicted in Fig. 1B with strong bursting activity. Autocorrelation function that matches Type III SCCs will be zero for all non-zero lags (not shown).

Figure 6. Type I serial correlation coefficients (SCCs). The top row depicts experimental spike-train from a P-type primary electrosensory afferent. A. The spike train has a unimodal interspike interval distribution (ISI) and does not display bursting activity. Abscissa is ISI in multiples of electric organ discharge (EOD) period (1.05 ms), and ordinate is probability. B. Joint interspike interval distribution showing probability of observing successive intervals $ISI\,(i)$ (abscissa) and $ISI\,(i+1)$ (ordinate). The size of the circle is proportional to the probability of jointly observing the adjacent ISIs, i.e., $P\,(i,\,i+1)$. C. Serial correlation coefficients (ordinate) as a function of lag measured in multiples of mean ISI (abscissa). Spike-train sampled at EOD period. SCCs for this afferent are shown in Fig. 1A at a different sampling rate (see Materials and Methods). The bottom row depicts results for matched model using low-pass filtered white noise. Panel descriptions as in top row, except in F where open circles denote experimental data taken from C. To generate model results, $v = 1.845$ mV; adaptive threshold parameters: $A =$ 0.15 mV and $\tau = 30$ ms; low-pass filter: $\tau_\gamma = 2.74$ ms ($a = 0.4$) and $\mathbf{R}\,(0) = 1.07 \times 10^{-3}$ mV^2 (SNR = 35 dB).

2.3.3 Type II serial correlation coefficients

Type II afferent spiking activity demonstrates serial correlations where $\rho_1 < -0.5$ and successive ρ_k alternate between positive (ρ_{2k}) and negative (ρ_{2k+1}) values, are progressively damped, and diminish to zero (Fig. 1B). These spike trains have bimodal ISIs and display bursting activity. They can be generated by a process where threshold noise is high-pass filtered white noise.

Consider a first-order high-pass filter whose time-constant is $\tau_\gamma = RC$, where R is the resistance and C the capacitance. Denote $a = \exp\left(-T_1/\tau_\gamma\right) < 1$, where T_1 is the mean ISI as before. The impulse response of the filter in discrete-time is $h\,(n) = (-a)^n\,u\,(n)$ where $u\,(n)$ is the Heaviside function, and n denotes integer multiples of T_1. This filter is similar to the low-pass filter used for generating Type I serial correlations with a replaced by $-a$. It can be shown (see Materials and Methods)

$$\mathbf{R}\,(k) = \frac{(-a)^k}{1-a^2}\,. \tag{16}$$

Eqs. 9 and 16 yield the geometric series

$$\rho_k = -\frac{(-a)^{k-1}\,(1+a)}{2}\,, \quad k \geq 1\,. \tag{17}$$

From Eq. 17, and noting that $|a| < 1$, we conclude that $\rho_k < -0.5$, $\rho_{2k} > 0$, $\rho_{2k+1} < 0$, $|\rho_{2k}| > |\rho_{2k+1}|$, and $\rho_k \to 0$ as $k \to \infty$. This is the observed Type II pattern of SCCs. Further, summing the geometric series (Eq. 17) yields $\sum_{k\geq 1} \rho_k = -0.5$ as stated in Eq. 10. The high-pass filter parameter a can be estimated from experimentally determined SCCs. From Eq. 17 this is

$$a = -(1 + 2\rho_1)\,. \tag{18}$$

As noted for Type I SCCs a better estimate is often obtained from the ratio $a = -\rho_2/\rho_1$. Finally, note that Type II serial correlations can be obtained from Type I serial correlations by replacing a in Eqs. 13, 14, and 15 with $-a$, to obtain Eqs. 16, 17, and 18, respectively.

We model Type II P-type primary electrosensory afferents using a noisy threshold with (high-pass) correlation function $\mathbf{R}(k)$ specified by Eq. 16. The noise correlation function and adaptive threshold parameters were determined from the experimental data, and tuned so that they matched the afferent SCCs, ISI, and joint distributions. We demonstrate with two examples: i) moderate bursting activity with a broad bimodal ISI distribution (Fig. 7A-C) along with the matched model (Fig. 7D-F). The high-pass noise correlation function is depicted in (Fig. 5B). SCCs of adjacent ISI ρ_1 are -0.59 (data) and -0.62 (model). ii) strong bursting activity with a well-defined bimodal distribution (Fig. 8A-C) along with the matched model (Fig. 8D-F). The high-pass noise correlation function is depicted in (Fig. 5C). SCCs of adjacent ISI ρ_1 are -0.7 (data) and -0.7 (model). The observed patterns of Type II SCCs are reproduced.

Figure 7. Type II serial correlation coefficients (SCCs) with moderate bursting activity. Description is identical to Fig. 6. Top row depicts experimental spike train from a P-type primary electrosensory afferent. Bottom row depicts results for matched model using high-pass filtered white noise. To generate model results, $v = 1.845$ mV; adaptive threshold parameters: $A = 0.28$ mV and $\tau = 26$ ms; high-pass filter: $\tau_\gamma = 3.18$ ms ($a = 0.29$) and $\mathbf{R}(0) = 9.22 \times 10^{-3}$ mV^2 (SNR = 26 dB).

Figure 8. Type II serial correlation coefficients (SCCs) with strong bursting activity. Description is identical to Fig. 6. Top row depicts experimental spike train from a P-type primary electrosensory afferent. SCCs for this afferent are shown in Fig. 1B at a different sampling rate (see Materials and Methods). Bottom row depicts results for matched model using high-pass filtered white noise. To generate model results, $v = 1.845$ mV; adaptive threshold parameters: $A = 0.19$ mV and $\tau = 60$ ms; high-pass filter: $\tau_\gamma = 16.89$ ms ($a = 0.69$) and $\mathbf{R}(0)$: 6.4×10^{-3} mV^2 (SNR = 27 dB).

2.3.4 Type III serial correlation coefficients

Type III afferent spiking activity demonstrates serial correlations where $\rho_1 = -0.5$ and all $\rho_k = 0$ for $k \geq 2$. This is a degenerate case, resulting in a singleton with a unique set of SCCs. Such spike trains can be generated by a process where spike threshold noise is uncorrelated, and hence white. In this case, $\mathbf{R}(0) = \sigma_\gamma^2$ is noise power, and $\mathbf{R}(k) = 0$ for $k \geq 1$. We see immediately from Eqs. 8 and 9 that the ρ_k have the prescribed form. Further, we note that trivially $\sum_{k\geq 1} \rho_k = -0.5$ as stated in Eq. 10. We mention in passing, and for later discussion, that the use of unfiltered white noise is equivalent to passing white Gaussian noise through an all-pass filter ($\tau_\gamma \to 0$, i.e., $a \to 0$). A spike-train with exactly Type III SCCs has not been observed in the experimental data presented here.

3 Discussion

All experimentally observed P-type spike-trains from *Apteronotus leptorhynchus* demonstrated negative correlations between adjacent ISIs (Figs. 1, 2A). Thus, negative SCCs between adjacent ISIs may be an obligatory feature of neural coding, at least in this species. A broad experimental observation is the roughly equal division of spike-trains into units with $\rho_1 > -0.5$ (non-bursting or Type I units, Fig. 1A, $N = 24$) and units with $\rho_1 < -0.5$ (bursting or Type II units, Fig. 1C, $N = 28$) (using a different method of classification, Xu *et al.* [46] reported 31% of 117 units as bursting). Further, for Type I units $\rho_2 < 0$, whereas for Type II units $\rho_2 > 0$ (Fig. 2C). The former give rise to over-damped SCCs which remain negative and diminish to zero, while the latter give rise to under-damped or oscillatory SCCs which also diminish to zero. The dependence of ρ_2 on ρ_1 is linear for the most part and follows the equation $\rho_2 = -1.18\rho_1 - 0.51$ (Fig. 2C). The limiting sum of SCCs $\sum_{k\geq 1} \rho_k$ is close to -0.5 irrespective of the type of SCC pattern (Fig. 2D) [19]. Although the dominant SCCs are ρ_1 and ρ_2, their sum $\rho_1 + \rho_2$ is not close to -0.5 because the linear relationship between these coefficients deviates from the line $\rho_2 = -\rho_1 - 0.5$ (re: $\rho_2 = -1.18\rho_1 - 0.51$, Fig. 2C). Thus, more terms (lags) are needed to bring the sum of SCCs close to the limiting value, and this results in significant correlations extending over multiple lags (time-scales). As discussed in an earlier work [1], long-short (short-long) ISIs create memory in the spike-train and keep track of the deviations of successive ISIs from the mean ISI (T_1). These deviations are a series of "credits" and "debits" which may not balance over adjacent ISIs, but will eventually balance so that a given observation of k successive ISIs returns to the mean with $t_k - t_0 \approx kT_1$. Such a process will exhibit long-range dependencies that may not be captured by SCCs.

That the dependencies may extend over multiple ISIs is confirmed from an analysis of joint dependencies of intervals extending to high orders [1, 6]. All 52 units in *Apteronotus leptorhynchus* were at least second-order Markov processes, with about half ($N = 24$) being fifth-order or higher [1]. Further, SCCs were not correctly predicted when only the adjacent ISI dependencies were preserved, i.e., were considered to be first-order Markov [1, 47, 48]. Indeed, an examination of the sequence of SCCs provides no indication of the extent of memory. For instance, short-duration correlations do not necessarily imply that ISI dependencies are limited to fewer adjacent intervals. Long-duration dependencies may be present even when the correlation time is short [6]. Conversely, a first-order Markov process produces a ringing in the serial correlogram ($\rho_k = \rho_1^k$) that can continue for ISIs much longer than two adjacent ISIs [45, 49]. In fact, for some P-type electrosensory afferent spike trains, the observed ISIs exhibited SCCs whose magnitudes were smaller than the SCCs for the matched first-order Markov model even though the experimental data was at least second-order or higher (see Fig. 8, [1]).

These experimental observations and dependency analyses motivated us to ask whether we could develop a model to reproduce a prescribed sequence of SCCs $\rho_1, \rho_2, \ldots, \rho_k$. A model with a time-varying adaptive threshold is a physiologically plausible choice and is used here in its simplest form. Adaptive threshold models typically have three components: 1) dynamics of membrane voltage $v(t)$ in response to an input signal, 2) a spike or impulse generator, and 3) a time-varying adaptive threshold $r(t)$ which is elevated following a spike and subsequently relaxes monotonically. A spike is fired when the membrane voltage meets the (relaxing) threshold, thus forming a feedback loop (Fig. 3B). In many models, the feedback loop is implicitly defined (as in conductance-based models), but it may also be explicitly defined as is done here so that $v(t) - r(t)$ reaches a fixed spike-initiation threshold. This threshold is usually

taken to be zero, i.e., $v(t) = r(t)$ [37], but here we assume it to be a non-zero value ($\gamma = A/2$) following [39]. This choice does not alter model behavior or the analyses presented here (see Materials and Methods) but has relevance to optimal coding and is discussed further below.

Models that do not utilize an adaptive threshold are not discussed here because they are outside the scope of this work. The reader may consult [3, 43, 50, 51] for more details and references to such models. Most adaptive threshold models that address negative correlations assume some form of perfect or leaky integrate-and-fire dynamics for the first two components listed above [17, 20, 28–30, 52], and a first-order low-pass filter without reset for the adaptive threshold (i.e., an exponential kernel). Noise is added to the adaptive threshold or to the input current and this results in negative serial correlations. Model complexity has been a major drawback in determining the precise role of noise in shaping ISI correlations (see for instance observations made in [20, 53]). Resets, hard refractory periods, sub-threshold dynamics due to synaptic filtering, and sometimes multiple sources of noise obscure the effects of signal propagation through the system and obscure signal dependencies. Thus with few exceptions (see below), adaptive threshold models have been qualitative. They demonstrate some features of experimentally observed ISI distributions, and at best correlations between adjacent ISIs (i.e., ρ_1) [26, 29, 42]. On the other hand, reduced model complexity can result in a lack of biophysical plausibility. Thus a judicious choice of models should expose desired mechanisms while retaining enough important features of the phenomena.

In recent years deterministic adaptive threshold models with an exponential kernel have been used to predict spike-times from cortical and peripheral neurons [37, 39, 54] (see [38] for an early review) and predict peri-stimulus time histograms [39, 41] with good accuracy. Capturing spike-times accurately is perhaps the first requirement in our analysis, and this gives confidence that the model may tell us something about ISI correlations. We eliminated sub-threshold dynamics and resets so that there is only one nonlinear element, the spike generator [39, 41]. These are not serious restrictions, and they make the analysis tractable. We follow the usual practice of representing the adaptive threshold element with a first-order low-pass filter with time-constant τ. The absence of reset implies that the time-varying adaptive threshold which carries memory is a simple convolution of the spike train with an exponential kernel (Fig. 3B). We inject noise precisely in one of two places, either to perturb the spike threshold γ (Fig. 4) or perturb the time-constant τ (Fig. 9). The two forms of perturbation are formally equivalent (see Materials and Methods). We linearize the exponential so that we can obtain analytical solutions of SCCs. This is applicable at asymptotically high spike-rates [39, 40] and applies well to P-type afferent spike trains because of their high baseline firing rates (about 250-300 spikes/s, [1, 46, 55]). To fit ISI and joint-ISI distributions of individual P-type afferents, the parameters of the adaptive threshold element $h(t)$ (A and τ) are obtained from the afferent spike-train [39]. The noise filter coefficient is directly obtained from ρ_1 through Eqs. 15 and 18 or from the ratio ρ_2/ρ_1, and thence the auto-correlation functions through Eqs. 13 or 16. These model elements and procedures allow us to determine the shaping of ISI correlations. The major results are

1. ISI correlations are determined by the auto-correlation function, $\mathbf{R}$, of the noise process (Eqs. 8-9).

2. Non-bursting units and bursting units are described by the same functional relationship between ISI SCCs and $\mathbf{R}$.

3. Non-bursting spike trains (with unimodal ISI distribution) are generated by low-pass filtered white noise (e.g., Fig. 6).

4. Bursting spike trains (with bi-modal ISI distribution) are generated by high-pass filtered white noise (e.g., Figs. 7 and 8).

5. The two filter types (low- and high-pass) are described by the sign of a single parameter, the filter pole a. The filter pole is directly related to the time-constant of the filter. This parameter can be uniquely determined from the correlation, ρ_1, between adjacent ISIs (Eqs. 15 and 18). More robust estimates are obtained from the ratio ρ_2/ρ_1. SCCs at subsequent lags are related to a as terms in a geometric progression (Eqs. 14 and 17).

6. While more complex patterns of SCCs can be produced by other types of filters, only Type I and Type II SCC patterns are observed in P-type afferent spike-trains. Type III SCC patterns are mentioned here because they are commonly reported in modeling studies. They are restricted to adjacent ISIs (discussed further below).

7. The expression for ISI SCCs is independent of the adaptive threshold parameters (A and τ), the signal (v), and the firing threshold (γ). It is dependent only on the noise correlation function, including noise power $\mathbf{R}(0)$.

8. The model fits ISI and joint-ISI distributions.

9. For both non-bursting and bursting units (low- and high-pass filtered noise, respectively) the theoretical prediction of the sum of ISI SCCs is exactly -0.5. The sum of SCCs over all afferent spike trains is close to this limit -0.475 ± 0.04. ($N = 52$).

10. SNR ($10 \log_{10}(v^2/\mathbf{R}(0))$ is generally larger than 20 dB, i.e., fluctuations in threshold (noise) are small compared to the input signal. This is in keeping with the hypothesis that spike-time jitter is small in comparison with the mean ISI.

There are two components to ISI serial correlations as is apparent from Eq. 5 where the ith ISI is given by $t_i - t_{i-1} = (\gamma_i - \gamma_{i-1} + A)/m$. The first component is due to the difference $\gamma_i - \gamma_{i-1}$ which is coupled to the next (adjacent) interval $t_{i+1} - t_i$ by the common term γ_i. This term appears with opposing signs in adjacent ISIs and hence results in a negative correlation which does not extend beyond these ISIs. If the γ_i are uncorrelated then it can be shown that the adjacent ISI correlation $\rho_1 = -0.5$ and all other $\rho_k = 0$, for $k \geq 2$. Thus, for independent random variables, this result follows from the property of a differentiator and it is not indicative of memory beyond adjacent ISIs. The second component of ISI correlations is due to long-range correlations $\mathbf{R}$ in the random process γ which extend beyond adjacent ISIs. These correlations are endogenous, possibly biophysical in origin, and could be shaped by coding requirements. The two components to ISI correlations are made clear by restating Eq. 9 for ρ_1 as

$$\rho_1 = -\frac{1}{2} + \frac{\mathbf{R}(1) - \mathbf{R}(2)}{2(\mathbf{R}(0) - \mathbf{R}(1))}. \tag{19}$$

The two components are thus separable. For a wide-sense stationary process, $\mathbf{R}(0) > \mathbf{R}(k)$ for all $k \geq 1$, and so the denominator of the second term is always positive. Thus the deviation of ρ_1 from -0.5 is determined by the sign of $\mathbf{R}(1) - \mathbf{R}(2)$. This term is positive for non-bursting units (Type I SCCs), and it is negative for bursting units (Type II SCCs). The singleton case (Type III SCCs) results because noise is uncorrelated and so the term vanishes. In this case only the first component is present. For exponentially correlated noise (from an Ornstein-Uhlenbeck or Gauss-Markov process) or low-pass filtered noise, $\mathbf{R}(1) - \mathbf{R}(2) > 0$

(Fig. 5A), and this produces a non-bursting Type I pattern. For high-pass filtered noise, $\mathbf{R}(1) - \mathbf{R}(2) < 0$ (Fig. 5B, C), and this produces a bursting Type II pattern with a bimodal ISI distribution. Thus, a single parameter (the sign of the pole of the filter) can create all the observed patterns of negatively correlated SCCs.

These results can be directly compared with results from an earlier study that explored patterns of SCCs generated by a perfect integrate-and-fire neuron under two conditions [30]: i) a deterministic adaptation current with fast, white gaussian noise input, and ii) a stochastic adaptation current with slow, exponentially correlated (channel) noise. A deterministic adaptation current with fast noise produced patterns of SCCs that we report here as Types I and II. These patterns were characterized by a parameter ϑ which is analogous to the filter-pole a used here, with Type I pattern ($a > 0$) corresponding to $\vartheta > 0$, and Type II pattern ($a < 0$) corresponding to $\vartheta < 0$. A pattern similar to the Type III SCC pattern reported here resulted when $\vartheta = 0$. This is the same as $a = 0$, i.e., an all-pass noise filter which allows white noise to pass through unchanged with infinite bandwidth ($\tau_\gamma \to 0$ when $a \to 0$). While we define Type III SCCs to have only one pattern ($\rho_1 = -0.5$, and $\rho_k = 0$ for $k \geq 2$) Lindner and co-workers [30] report that $\rho_k = 0$ when $k \geq 2$, but the value of ρ_1 is governed by an additional parameter and could take a range of values, with $\rho_1 = -0.5$ appearing as a limiting case. The case of stochastic adaptation currents produced only positive ISI correlations which we do not observe or model here. In a subsequent report Lindner and co-workers [31] extended the model to more general integrate-and-fire models and obtained a relationship between ISI SCCs and the phase-response curve (PRC). The patterns of SCCs reported here most likely correspond to their Type I PRC.

The SCCs reported here follow a geometric progression with the filter-pole a being the ratio parameter (see Eqs. 14 and 17). Lindner and co-workers [30, 31] reported that the patterns of negative correlations follow a geometric progression with the ratio parameter being ϑ. Similarly, a geometric progression was also reported by Urdapilleta [43]. Note that a first-order Markov process also follows a geometric progression with ratio parameter ρ_1 [45, 49]. The limiting sum of SCCs reported here (Eq. 10) is exactly -0.5, whereas [31] report a sum that asymptotically approaches a value that is slightly larger than -0.5. The average sum of SCCs in our experimental data (Fig. 2D) is also slightly larger than -0.5 (-0.475 ± 0.04.), and this merits further investigation.

In considering the results presented here and those from the more general model of Lindner and co-workers, it appears that the presence of the feedback loop (the coupling between the membrane voltage and the adaptation variable) may account for almost all the properties of SCCs if the noise fluctuation is shaped appropriately. Furthermore, this study shows that noise injected in the spiking threshold γ or the time-constant τ has the effect of reverberating around the feedback loop and introducing negative correlations extending over multiple ISIs. This suggests an endogenous source of noise, resulting for example from probabilistic transitions between conformational states in voltage-gated ion channels [54, 56–59]. This is plausible because the amount of noise to be added is small, generally weaker than 20 dB SNR (see Figs. 6, 7, and 8). That only weak noise may be necessary was also reported in [31], and in a study on threshold shifts [54]. An ion channel that may be a substantial contributor to adaptation currents is the non-inactivating M-current (KCNQ/Kv7 family) [60] which is a voltage-gated potassium channel with slow dynamics (relative to the mean ISI). This channel may be responsible for spike-timing precision, and hence a timing-based neural code (see [39] for a discussion). The possible role of the M-current and the specific sources of noise have been explored earlier (see [30] above in relation to types of SCCs, and [5]). Both studies [5, 30]

included a modified conductance-based Traub-Miles model [61, 62] frequently used in modeling M-currents. They concluded that negative correlations were a consequence of a deterministic adaptation current with fast white Gaussian noise input rather than a stochastic adaptation current with slow (low-pass filtered) correlated noise. These reports should be contrasted with the current study which shows that negative correlations can be generated by a deterministic adaptation current with a noisy spike firing threshold, or a non-deterministic adaptation current with a noisy time-constant. Further, either slow noise (low-pass filtered) or fast noise (high-pass filtered) can produce negative correlations. Therefore the results obtained here show that the time-scale of noise fluctuations dictate only the type of SCC pattern that is produced (i.e., bursting or non-bursting).

Negatively correlated spike trains have implications for information transmission [20]. These spike trains have a power spectra $P(\omega)$ that rolls off towards DC [21], effectively reducing low-frequency noise and improving SNR in the baseband (also referred to as noise-shaping). The connection between DC power $P(\omega = 0)$ and the ISI SCCs is given by Eq. 12 [45], where it can be seen that negative correlations have a tendency to reduce noise at zero-frequency. From the theoretical limit of the sum of SCCs (-0.5) reported here in Eq. 10, DC-power vanishes thus yielding a perfect DC block allowing low-frequency signals to be transmitted with very high SNR. The limiting sum of observed SCCs is only just a little larger than the optimum sum (-0.5) and this allows some noise to bleed through at zero-frequency.

The model presented here sets the spike threshold $\gamma = A/2$. We showed previously that the time-varying adaptive threshold $r(t)$ is an estimator of the input signal $v(t)$, and the mean-squared estimation error $(v(t) - r(t))_{\mathrm{MSE}}$ is minimized when the firing threshold is $\gamma = A/2$ [39, 40]. The error increases for any other value of γ, in particular $\gamma = 0$ which is the value usually adopted in the literature. Thus, for optimal timing of spikes the ongoing estimation error must be bounded below by $\gamma = A/2$. When this bound is reached, the neuron fires a spike and raises the adaptive threshold variable by A. Viewed in this manner, the spike generator directly encodes the error and not the signal (Figs. 3B, 4B). This is much more efficient than directly encoding the signal (or some filtered version of the signal). The coding error is equivalent to quantization error in digital coding. This coding mechanism is analogous to one-bit delta modulation [63], but it is asymmetric because the threshold is one-sided. Thus, coding with an adaptive threshold is analogous to lossy digital coding or source coding. That is, a neuron performs data compression.

The adaptive threshold model parameters A and τ are fitted from the baseline (spontaneous) spike-train through the relationship $A\tau = v/r_s$, where r_s is the long-term baseline spike rate, and v is the bias input. This allows us to reduce the degrees of freedom to one, to either A or τ which is then determined by optimizing the model spike train to match experimental spike-times (see [39] for details of the procedure). Thus, the adaptive threshold parameters are not estimated from the observed SCCs. This fits with our understanding of the adaptive threshold as an optimal estimator with its parameters being determined by coding quality and an energy constraint [39, 40]. Therefore, from the point of view of optimal coding, the deterministic adaptive threshold parameters should not be estimated from the SCCs. When the parameters are fixed, a noisy threshold can be generated by estimating the filter parameter a from ρ_1 or ρ_2/ρ_1 to match the observed SCCs. This decoupling of adaptive threshold parameters from SCCs implies that a correlation function $\mathbf{R}$ can be generated from a specified sequence of SCCs (determined, say experimentally) without reference to the adaptive threshold. Thus, a family of spike trains with arbitrary mean ISI, ISI and joint-ISI distributions can be generated, all of which have the same sequence or pattern of SCCs. Some iteration will be

needed to fit the ISI and joint-ISI distributions from the adaptive threshold parameters A and τ as is done here, but this is not too difficult.

The simplified adaptive threshold model used here demonstrates the influence of noise on ISI correlations. A single-parameter (the sign of the noise filter coefficient or pole) captures the pattern of observed SCCs, and in this respect it agrees with a detailed dynamical model [31]. Further, this model can fit observed spike-times with good accuracy [37, 39], is energy-efficient while maximizing coding fidelity [39, 40], and is computationally inexpensive. Thus, it can be useful for understanding timing-based neural codes.

4 Materials and Methods

4.1 Experimental procedures

Spike data from P-type primary electrosensory afferents were collected in a previous *in vivo* study in *Apteronotus leptorhynchus*, a wave-type weakly electric fish [1]. All animal procedures including animal handling, anesthesia, surgery, and euthanasia received institutional ethics (IACUC) approval from the University of Illinois at Urbana-Champaign, and were carried out as previously reported [1]. No new animal procedures were performed during the course of this work. Of relevance here are some details on the electrophysiological recordings. Briefly, action potentials (spikes) were isolated quasi-intracellularly from P-type afferent nerve fibers in quiescent conditions under ongoing electric organ discharge (EOD) activity (this is the so-called baseline spiking activity of P-type units). An artificial threshold was applied to determine spike onset times, and reported at the ADC sampling rate of 16.67 kHz (sampling period of 60 μs). The fish's ongoing quasi-sinusoidal EOD was captured whenever possible to determine the EOD frequency from the power spectrum, or the EOD frequency was estimated from the power spectrum of the baseline spike train. Both methods reported almost identical EOD frequencies. EOD frequencies typically ranged from $750 - 1000$ Hz (see [1] for more details).

4.2 Data analysis

P-type units fire only once every EOD cycle and this forms a convenient time-base to resample the spike train [1]. Spike times resampled at the EOD rate are reported as increasing integers n_1, n_2, $n_3, \ldots$, etc with the first spike in the data record assumed to start on the first EOD cycle ($n_1 = 1$). Resampling removes the phase jitter in spike-timing but retains long-term correlations due to memory effects in the spike train. Serial correlation coefficients or SCCs were estimated from time-stamps at the original ADC rate of 16.67 kHz (as shown in Fig. 1) and at the EOD frequency (as shown in Figs 6C, 7C, and 8C) [1]. In some afferents there is a small difference between the two estimates, particularly estimates of ρ_1, but this has negligible effect on the results. SCCs were estimated from the resampled spike trains as follows:

$$\rho_k = \frac{\sum_{i=1}^{M-k} (n_i - T_1)(n_{i+k} - T_1)}{(\sum_{i=1}^{M-k} (n_i - T_1)^2 \sum_{i=1}^{M-k} (n_{i+k} - T_1)^2)^{1/2}}, \quad k \geq 0, \tag{20}$$

where T_1 is the mean ISI, and M is the number of spikes in a block, typically ranging from $500 - 1000$ spikes. The SCCs were estimated in non-overlapping blocks and then averaged.

4.3 Deterministic and stochastic adaptive threshold model

In the deterministic adaptive threshold model (Fig. 3A), if t_{i-1} and t_i, $i \geq 1$, are successive spike-times, then the time evolution of the adaptive threshold $r(t)$ is given by

$$\begin{aligned} r(t) &= (v-\gamma+A)\exp(-\frac{t-t_{i-1}}{\tau}), \quad t_{i-1} < t \leq t_i, \\ &= (v-\gamma+A)(1-\frac{t-t_{i-1}}{\tau}) + \mathrm{O}((t-t_{i-1})^2). \end{aligned} \tag{21}$$

At $t = t_{i-1}^+$, $r\left(t_{i-1}^+\right) = v-\gamma+A$, and at $t = t_i$, $r(t_i) = v-\gamma$. So,

$$r(t) - r(t_{i-1}) = -\frac{v-\gamma+A}{\tau}(t-t_{i-1}), \quad t_{i-1} < t \leq t_i. \tag{22}$$

Let $m = (v-\gamma+A)/\tau > 0$, so that the slope of the adapting threshold is $-m$, then

$$r(t) - r(t_{i-1}) = -m(t-t_{i-1}), \quad t_{i-1} < t \leq t_i. \tag{23}$$

Noting that $r(t_i) = v-\gamma$,

$$\frac{A}{m} = (t_i - t_{i-1}), \tag{24}$$

which is the deterministic firing threshold for a constant, DC-level signal.

4.3.1 Random firing threshold

We will now provide a stochastic extension to the model by making the spike threshold (γ) noisy while leaving all else unchanged (Fig. 4). Let γ be a discrete wide-sense stationary process with mean $\mathbf{E}[\gamma]$, discrete auto-correlation function $\mathbf{R}(k)$ and power $\mathbf{E}\left[\gamma^2\right] = \mathbf{R}(0)$. The spike threshold with additive noise assumes the value γ_i, $i \geq 1$ immediately after the $(i-1)^{\text{th}}$ spike and remains constant in the time interval $t_{i-1} < t \leq t_i$ (Fig 4A) [21,44]. Thus, the ith spike is emitted when the error satisfies the condition

$$e(t_i) \geq \gamma_i, \quad \text{for}\, t > t_{i-1}. \tag{25}$$

Subsequently the adaptive threshold jumps to a higher value specified by $v - \gamma_i + A$, and the noisy spike threshold assumes a new value γ_{i+1}. From Fig. 4A, proceeding as before

$$t_i - t_{i-1} = \frac{1}{m}\left\{\gamma_i - \gamma_{i-1} + A\right\}. \tag{26}$$

The mean ISI is therefore $\mathbf{E}[t_i - t_{i-1}] = A/m$ as in the deterministic case given by Eq. 3. Thus the correlation function for the ISI sequence is

$$\begin{aligned} \mathbf{E}[(t_i - t_{i-1})^2] &= \frac{1}{m^2}\mathbf{E}\left[(\gamma_i - \gamma_{i-1} + A)(\gamma_i - \gamma_{i-1} + A)\right], \\ &= \frac{A^2}{m^2} - \frac{2}{m^2}(\mathbf{E}[\gamma_i\,\gamma_{i-1}] - \mathbf{E}[\gamma^2]), \end{aligned} \tag{27}$$

with covariance

$$\mathrm{Cov}\left[(t_i - t_{i-1}), (t_i - t_{i-1})\right] = -\frac{2}{m^2}(\mathbf{E}\left[\gamma_i\,\gamma_{i-1}\right] - \mathbf{E}\left[\gamma^2\right]). \tag{28}$$

More generally, for the sum of k ISIs

$$\begin{aligned}\mathbf{E}[(t_{i+k}-t_i)^2] &= \frac{1}{m^2}\,\mathbf{E}[(\gamma_{i+k}-\gamma_i+A)^2],\\ &= \frac{kA^2}{m^2}-\frac{2}{m^2}(\mathbf{E}[\gamma_{i+k}\,\gamma_i]-\mathbf{E}[\gamma^2])\,,\end{aligned} \tag{29}$$

with covariance

$$\mathrm{Cov}\left[(t_{i+k}-t_i)\,,(t_{i+k}-t_i)\right]=-\frac{2}{m^2}(\mathbf{E}\left[\gamma_{i+k}\,\gamma_i\right]-\mathbf{E}\left[\gamma^2\right]). \tag{30}$$

From Eqs. 27 and 29 it can be seen that there are two terms in each of these ISI correlation functions. The first is due to the input signal (which gives a mean ISI of A/m) and the second is due to the discrete-event noise process γ. From the assumption of wide-sense stationarity the auto-correlation function $\mathbf{E}\left[\gamma_i\,\gamma_j\right]$ can be written as $\mathbf{R}\left(t_i-t_j\right)$. The discrete nature of the correlation function $\mathbf{R}$ is made clear in the following way. Denote the mean of the kth-order interval by $T_k=\mathbf{E}[t_{i+k}-t_i]$, then the mean ISI is T_1 $(=A/m)$, and further $T_k=kT_1$. The random variable γ is generated once every ISI and thus, the discrete auto-correlation function $\mathbf{R}$ takes values at successive multiples of the mean ISI, i.e., $\mathbf{R}\left(T_1\right)$, $\mathbf{R}\left(T_2\right)$, etc., and will be denoted by $\mathbf{R}\left(1\right)$, $\mathbf{R}\left(2\right)$, etc., respectively. As noted before, $\mathbf{R}\left(0\right)$ is noise power. That is, we can write

$$\mathbf{R}\left(t_{i+k}-t_i\right)=\mathbf{R}\left(T_k\right)=\mathbf{R}\left(k\right)\,, \tag{31}$$

and therefore

$$\mathrm{Cov}\left[(t_i-t_{i-1})\,,(t_i-t_{i-1})\right] = \frac{2}{m^2}\left\{\mathbf{R}\left(0\right)-\mathbf{R}\left(1\right)\right\}\,, \tag{32}$$

$$\mathrm{Cov}\left[(t_{i+k}-t_i)\,,(t_{i+k}-t_i)\right] = \frac{2}{m^2}\left\{\mathbf{R}\left(0\right)-\mathbf{R}\left(k\right)\right\}\,. \tag{33}$$

The cross-covariance of two ISIs (t_i-t_{i-1}) and $(t_{i+k}-t_{i+k-1})$ separated by lag $k\geq 1$ is

$$\begin{aligned}\mathrm{Cov}\left[(t_i-t_{i-1})\,,(t_{i+k}-t_{i+k-1})\right] &= \frac{1}{m^2}\,\mathbf{E}[((\gamma_i-\gamma_{i-1})+A)\,((\gamma_{i+k}-\gamma_{i+k-1})+A)]-\frac{A^2}{m^2}\,,\\ &\quad \frac{1}{m^2}\,\mathbf{E}[\gamma_i\,\gamma_{i+k}-\gamma_i\,\gamma_{i+k-1}+\gamma_i\,A-\gamma_{i-1}\,\gamma_{i+k}\\ &\quad +\gamma_{i-1}\,\gamma_{i+k-1}-\gamma_{i-1}\,A+\gamma_{i+k}\,A-\gamma_{i+k-1}\,A\\ &\quad +A^2]-\frac{A^2}{m^2}\,.\end{aligned} \tag{34}$$

This yields,

$$\begin{aligned}\mathrm{Cov}\left[(t_i-t_{i-1})\,,(t_{i+k}-t_{i+k-1})\right] &= \frac{1}{m^2}\,\mathbf{E}[\gamma_i\,\gamma_{i+k}-\gamma_i\,\gamma_{i+k-1}-\gamma_{i-1}\,\gamma_{i+k}+\gamma_{i-1}\,\gamma_{i+k-1}]\,,\\ &= -\frac{1}{m^2}\left\{\mathbf{R}\left(k-1\right)-2\,\mathbf{R}\left(k\right)+\mathbf{R}\left(k+1\right)\right\}\,.\end{aligned} \tag{35}$$

The general formula for serial correlation coefficients at lag k is [45],

$$\rho_k=\frac{\mathrm{Cov}\left[(t_i-t_{i-1})\,,\,(t_{i+k}-t_{i+k-1})\right]}{\mathrm{Cov}\left[(t_i-t_{i-1})\;(t_i-t_{i-1})\right]^{1/2}\,\mathrm{Cov}\left[(t_{i+k}-t_{i+k-1})\,,\,(t_{i+k}-t_{i+k-1})\right]^{1/2}}\,. \tag{36}$$

For a wide-sense stationary process the covariances are constant and so the subscript i can be dropped. Introducing Eqs. 32 and 35 into Eq. 36 yields

$$\rho_0 = 1, \tag{37}$$

$$\rho_k = -\frac{\mathbf{R}(k-1) - 2\,\mathbf{R}(k) + \mathbf{R}(k+1)}{2\,(\mathbf{R}(0) - \mathbf{R}(1))}\,, \quad k \geq 1\,. \tag{38}$$

The serial-correlation coefficients (SCCs) given by Eqs. 37-38 are independent of the slope m of the reconstruction filter (the decay rate of the adaptive threshold) and the reconstruction filter gain A. Thus, the observed correlation structure of the spike-train are determined solely by the noise statistics of the firing threshold γ.

4.3.2 Random time-constant

We can transform the adaptive threshold model with a random spike threshold from the previous section and Fig. 4 so that the time-constant of the adaptive threshold filter $h(t)$ is a random variable with mean τ (Fig. 9). From Eq. 26 the random variate $\gamma_i - \gamma_{i-1}$ which appears in the time-base can be transformed into the random variate m_i which is the slope of the linearized adaptive threshold (green line, Fig. 9). From

$$m_i = \frac{\tau_i}{v - \gamma + A}\,, \tag{39}$$

we obtain

$$A\tau_i = \tau\,\{\gamma_i - \gamma_{i-1} + A\}\,. \tag{40}$$

From which,

$$\mathbf{E}[\tau_i] = \tau\,. \tag{41}$$

As noted, this is the mean value of the adaptive threshold filter time-constant. It is immediately apparent from Eqs. 26 and 40 that the covariance of the sequence τ_i (sampled at the ISIs) is the same as the covariance of the ISI sequence up to a scale-factor, and therefore the serial correlations of τ ($\rho_{\tau,\,k}$, $k \geq 0$) are the same as the serial correlations of the ISIs. We can establish this as follows. From Eq. 40,

$$\mathrm{Cov}(\tau_i\,, \tau_i) = \frac{2\tau^2}{A^2}\,\{\mathbf{R}(0) - \mathbf{R}(1)\}\,, \tag{42}$$

and for $k \geq 1$,

$$\mathrm{Cov}(\tau_{i+k}\,, \tau_i) = -\frac{\tau^2}{A^2}\,\{\mathbf{R}(k-1) - 2\,\mathbf{R}(k) + \mathbf{R}(k+1)\}\,. \quad k \geq 1. \tag{43}$$

From Eqs. 42 and 43 we can determine the serial correlation coefficients of the random filter time-constants to be

$$\rho_{\tau,\,0} = 1, \tag{44}$$

$$\rho_{\tau,\,k} = -\frac{\mathbf{R}(k-1) - 2\,\mathbf{R}(k) + \mathbf{R}(k+1)}{2\,(\mathbf{R}(0) - \mathbf{R}(1))}, \quad k \geq 1. \tag{45}$$

The right side of the above equations, Eqs. 44 and 45, are the same as Eqs. 37 and 38, the expressions for the SCCs of a spike train generated with a noisy adaptive threshold. Thus, the random filter time-constants have the same serial correlation coefficients as the interspike

intervals, $\rho_{\tau,k} = \rho_k$. Thus, a noisy threshold and a noisy filter time-constant will generate spike trains with the same ISI SCCs, and either method can be used. Here we generate spike trains using a noisy threshold.

Figure 9. Adaptive threshold with a stochastic filter time-constant. B. Geometry of the spiking mechanism. The input is a deterministic piece-wise constant signal (v, here shown to be DC-valued). Two equivalent spike-generation processes are possible: Case (i) The spike threshold γ is noisy while the adaptive threshold filter $h(t)$ has fixed time-constant τ. This is the same as Fig. 4A. The spike threshold is a constant term $v - \gamma$ plus a random perturbation x_i generated at t_{i-1}^{+}, and held constant until the next spike at t_i (black). The time-varying adaptive threshold is shown in red. Case (ii) The adaptive filter $h(t)$ has a random time-constant with mean τ while the spike threshold is $v - \gamma$, and is constant. Threshold noise x_i from Case (i) can be transformed into a noisy filter time-constant τ_i which is fixed between spikes, and takes correlated values over successive ISIs. The estimate from this process is shown in green. Note that the spike-times for both cases are identical. B. Block diagram illustrating the feedback from the estimator that generates optimal spike times. The time-constant τ is a random variable that is constant between spikes, and is correlated over interspike intervals (ISIs). It is of interest to ask what is the relationship between the correlation functions in the two cases. Symbols and additional description as in text.

4.4 Sum of serial correlation coefficients

We now determine the limiting sum of the SCCs over all lags. Let us first transform ρ_k as follows by defining

$$z_k = -2\rho_k \left(\mathbf{R}\left(0\right) - \mathbf{R}\left(1\right)\right) , \quad k \geq 1 . \tag{46}$$

Then from Eq. 38, z_k is a moving-average process such that

$$z_k = \mathbf{R}\left(k-1\right) - 2\,\mathbf{R}\left(k\right) + \mathbf{R}\left(k+1\right) , \tag{47}$$

from which, it is easy to show that

$$\sum_{k=1}^{N} z_k = \mathbf{R}\left(N+1\right) - \mathbf{R}\left(N\right) + \mathbf{R}\left(0\right) - \mathbf{R}\left(1\right) , \tag{48}$$

where N is the Nth-order interval. We make the assumption that the process γ is aperiodic and the auto-correlation function $\mathbf{R}\left(N\right) \to 0$ when $N \to \infty$, and so $\left(\mathbf{R}\left(N+1\right) - \mathbf{R}\left(N\right)\right) \to 0$. That is, the process decorrelates over long time-scales. Thus,

$$\lim_{N\to\infty} \sum_{k=1}^{N} z_k = \mathbf{R}\left(0\right) - \mathbf{R}\left(1\right) . \tag{49}$$

By summing Eq. 46 over all k, and inserting the result from Eq. 49, we have the limiting sum of the ISI serial correlation coefficients for the spike-train

$$\sum_{k=1}^{\infty} \rho_k = -\frac{1}{2} . \tag{50}$$

4.5 Noise correlation function for Type I SCCs

Consider a first-order low-pass filter whose time-constant is $\tau_\gamma = RC$, where R is the resistance and C the capacitance. Denote $a = \exp\left(-T_1/\tau_\gamma\right) < 1$, where T_1 is the mean ISI as before. The impulse response of the filter in discrete-time is $h\left(n\right) = a^n\, u\left(n\right)$ where $u\left(n\right)$ is the Heaviside function, and n denotes integer multiples of T_1. For white noise input, the output is wide sense stationary, and its discrete auto-correlation function $\mathbf{R}\left(k\right)$ is

$$\begin{aligned}
\mathbf{R}\left(k\right) &= \sum_{n=-\infty}^{\infty} \left(a^n\, u\left(n\right)\right)\left(a^{n-k}\, u\left(n-k\right)\right), \\
&= \sum_{n=k}^{\infty} a^n\, a^{n-k}, \\
&= a^k \sum_{m=0}^{\infty} a^{2m}, \\
&= \frac{a^k}{1-a^2}. \qquad (51)
\end{aligned}$$

4.6 Noise correlation function for Type II SCCs

Consider a first-order high-pass filter whose time-constant is $\tau_\gamma = RC$, where R is the resistance and C the capacitance. Denote $a = \exp\left(-T_1/\tau_\gamma\right) < 1$, where T_1 is the mean ISI as before. The impulse response of the filter in discrete-time is $h\left(n\right) = \left(-a\right)^n\, u\left(n\right)$ where $u\left(n\right)$ is the Heaviside function, and n denotes integer multiples of T_1. For white noise input, the output is wide sense stationary, and its discrete auto-correlation function $\mathbf{R}\left(k\right)$ is

$$\begin{aligned}
\mathbf{R}\left(k\right) &= \sum_{n=-\infty}^{\infty} \left(\left(-a\right)^n\, u\left(n\right)\right)\left(\left(-a\right)^{n-k}\, u\left(n-k\right)\right), \\
&= \sum_{n=k}^{\infty} \left(-a\right)^n\, \left(-a\right)^{n-k}, \\
&= \left(-a\right)^k \sum_{m=0}^{\infty} a^{2m}, \\
&= \frac{\left(-a\right)^k}{1-a^2}. \qquad (52)
\end{aligned}$$

This filter form (Eq. 52) is the same as the low-pass filter used for generating Type I serial correlations (Eq. 51) with a replaced by $-a$.

References

1. Ratnam R, Nelson ME. Nonrenewal statistics of electrosensory afferent spike trains: implications for the detection of weak sensory signals. J Neurosci. 2000;20(17):6672–6683.

2. Goense JB, Ratnam R. Continuous detection of weak sensory signals in afferent spike trains: the role of anti-correlated interspike intervals in detection performance. J Comp Physiol A. 2003;189(10):741–759.

3. Farkhooi F, Strube-Bloss MF, Nawrot MP. Serial correlation in neural spike trains: Experimental evidence, stochastic modeling, and single neuron variability. Phys Rev E. 2009;79(2):021905.

4. Lowen SB, Teich MC. Auditory-nerve action potentials form a nonrenewal point process over short as well as long time scales. J Acoust Soc Am. 1992;92:803–806.

5. Fisch K, Schwalger T, Lindner B, Herz AV, Benda J. Channel noise from both slow adaptation currents and fast currents is required to explain spike-response variability in a sensory neuron. J Neurosci. 2012;32(48):17332–17344.

6. van der Heyden MJ, Diks CGC, Hoekstra BPT, J D. Testing the order of discrete Markov chains using surrogate data. Physica D. 1998;117:299—313.

7. Katsuki Y, Yoshino S, Chen J. Action currents of the single lateral-line nerve fiber of fish. I. On the spontaneous discharge. Jpn J Physiol. 1950;1:87–99.

8. Kuffler SW, Fitzhugh R, B BH. Maintained activity in the cat's retina in light and darkness. J Gen Physiol. 1957;40:683—702.

9. Neiman A, Russell DF. Stochastic biperiodic oscillations in the electroreceptors of paddlefish. Phys Rev Lett. 2001;86:3443–3446.

10. Johnson DH, Tsuchitani C, Linebarger DA, Johnson MJ. Application of a point process model to responses of cat lateral superior olive units to ipsilateral tones. Hear Res. 1986;21:135–159.

11. Hagiwara S, Morita H. Coding mechanisms of electroreceptor fibers in some electric fish. J Neurophysiol. 1963;26:551–567.

12. Amassian VE, Macy J, Waller HJ, Leader HS, Swift M. Transformations of afferent activity at the cuneate nucleus. In: Gerard RW, Duyff J, editors. Information Processing in the Nervous System. Excerpta Medica Foundation, Amsterdam; 1964. p. 235–254.

13. Bullock TH, Chichibu S. Further analysis of sensory coding in electroreceptors of electric fish. Proc Natl Acad Sci U S A. 1965;54:422—429.

14. Calvin WH, Stevens CF. Synaptic noise and other sources of randomness in motoneuron interspike intervals. J Neurophysiol. 1968;31:574–587.

15. Nawrot MP, Boucsein C, Rodriguez-Molina V, Aertsen A, Grün S, Rotter S. Serial interval statistics of spontaneous activity in cortical neurons in vivo and in vitro. Neurocomputing. 2007;70(10-12):1717–1722.

16. Avila-Akerberg O, Chacron MJ. Nonrenewal spike train statistics: causes and functional consequences on neural coding. Exp Brain Res. 2011;210(3-4):353–371.

17. Chacron MJ, Longtin A, Maler L. Negative interspike interval correlations increase the neuronal capacity for encoding time-dependent stimuli. J Neurosci. 2001;21(14):5328–5343.

18. Goense JB, Ratnam R, Nelson ME. Burst firing improves the detection of weak signals in spike trains. Neurocomputing. 2003;52:103–108.

19. Ratnam R, Goense JBM. Variance stabilization of spike-trains via non-renewal mechanisms: The impact on the speed and reliability of signal detection. Computational Neuroscience Meeting (CNS*2004) (Baltimore, MD); 2004.

20. Chacron MJ, Lindner B, Longtin A. Noise shaping by interval correlations increases information transfer. Phys Rev Lett. 2004;92(8):080601.

21. Chacron MJ, Lindner B, Longtin A. ISI correlations and information transfer. Fluctuation Noise Lett. 2004;4(01):L195–L205.

22. Hill AV. Excitation and accommodation in nerve. Proc R Soc Lond B Biol Sci. 1936;119:305–355.

23. Buller AJ. A model illustrating some aspects of muscle spindle physiology. J Physiol. 1965;179:402–416.

24. Hagiwara S. Analysis of interval fluctuation of the sensory nerve impulse. Jpn J Physiol. 1954;4(3):234–240.

25. Goldberg JM, Adrian HO, Smith FD. Response of neurons of the superior olivary complex of the cat to acoustic stimuli of long duration. J Neurophysiol. 1964;27:706–749.

26. Geisler CD, Goldberg JM. A stochastic model of the repetitive activity of neurons. Biophys J. 1966;6(1):53–69.

27. ten Hoopen M. On an impulse generating mechanism. Jpn J Physiol. 1964;14:607–614.

28. Brandman R, Nelson ME. A simple model of long-term spike train regularization. Neural Comput. 2002;14(7):1575–1597.

29. Chacron MJ, Longtin A, St-Hilaire M, Maler L. Suprathreshold Stochastic Firing Dynamics with Memory in P-Type Electroreceptors. Phys Rev Lett. 2000;85:1576–1579. doi:10.1103/PhysRevLett.85.1576.

30. Schwalger T, Fisch K, Benda J, Lindner B. How noisy adaptation of neurons shapes interspike interval histograms and correlations. PLoS Comput Biol. 2010;6. doi:10.1371/journal.pcbi.1001026.

31. Schwalger T, Lindner B. Patterns of interval correlations in neural oscillators with adaptation. Front Comput Neurosci. 2013;7:164. doi:10.3389/fncom.2013.00164.

32. Benda J, Herz AVM. A universal model for spike-frequency adaptation. Neural Comput. 2003;15:2523–2564.

33. Prescott SA, Sejnowski TJ. Spike-rate coding and spike-time coding are affected oppositely by different adaptation mechanisms. J Neurosci. 2008;28(50):13649–13661.

34. Liu YH, Wang XJ. Spike-frequency adaptation of a generalized leaky integrate-and-fire model neuron. J Comput Neurosci. 2001;10(1):25–45.

35. Benda J, Maler L, Longtin A. Linear versus nonlinear signal transmission in neuron models with adaptation currents or dynamic thresholds. J Neurophysiol. 2010;104:2806—2820.

36. Jolivet R, Lewis TJ, Gerstner W. Generalized integrate-and-fire models of neuronal activity approximate spike trains of a detailed model to a high degree of accuracy. J Neurophysiol. 2004;92:959—976.

37. Kobayashi R, Tsubo Y, Shinomoto S. Made-to-order spiking neuron model equipped with a multi-timescale adaptive threshold. Front Comput Neurosci. 2009;3.

38. Gerstner W, Naud R. How good are neuron models? Science. 2009;326:379–380.

39. Jones DL, Johnson EC, Ratnam R. A stimulus-dependent spike threshold is an optimal neural coder. Front Comput Neurosci. 2015;9:61.

40. Johnson EC, Jones DL, Ratnam R. Minimum Squared-Error, Energy-Constrained Encoding by Adaptive Threshold Models of Neurons. In: Proc IEEE Int Symp Info Theory. IEEE; 2015. p. 1337–1341.

41. Johnson EC, Jones DL, Ratnam R. A minimum-error, energy-constrained neural code is an instantaneous-rate code. J Comput Neurosci. 2016;40(2):193–206.

42. Chacron MJ, Pakdaman K, Longtin A. Interspike interval correlations, memory, adaptation, and refractoriness in a leaky integrate-and-fire model with threshold fatigue. Neural Comput. 2003;15(2):253–278.

43. Urdapilleta E. Onset of negative interspike interval correlations in adapting neurons. Phys Rev E. 2011;84(4):041904.

44. Gestri G, Mastebroek H, Zaagman W. Stochastic constancy, variability and adaptation of spike generation: performance of a giant neuron in the visual system of the fly. Biol Cybern. 1980;38(1):31–40.

45. Cox DR, Lewis PAW. The Statistical Analysis of Series of Events. John Wiley and Sons, New York, NY; 1966.

46. Xu Z, Payne JR, Nelson ME. Logarithmic time course of sensory adaptation in electrosensory afferent nerve fibers in a weakly electric fish. J Neurophysiol. 1996;76(3):2020–2032.

47. Poggio GF, Viernstein LJ. Time series analysis of impulse sequences of thalamic somatic sensory neurons. J Neurophysiol. 1964;27:517–545.

48. Rodieck RW. Maintained activity of cat retinal ganglion cells. J Neurophysiol. 1967;30(5):1043–1071.

49. Nakahama H, Ishii N, Yamamoto M. Markov process of maintained impulse activity in central single neurons. Kybernetik. 1972;11(2):61–72.

50. Longtin A, Laing C, Chacron MJ. Correlations and memory in neurodynamical systems. In: Processes with Long-range Correlations. Springer; 2003. p. 286–308.

51. Lindner B, Schwalger T. Correlations in the sequence of residence times. Phys Rev Lett. 2007;98(21):210603.

52. Lindner B. Interspike interval statistics of neurons driven by colored noise. Phys Rev E. 2004;69(2):022901.

53. Lindner B, Chacron MJ, Longtin A. Integrate-and-fire neurons with threshold noise: A tractable model of how interspike interval correlations affect neuronal signal transmission. Phys Rev E. 2005;72(2):021911.

54. Fontaine B, Peña JL, Brette R. Spike-threshold adaptation predicted by membrane potential dynamics in vivo. PLoS Comput Biol. 2014;10(4):e1003560.

55. Bastian J. Electrolocation. J Comp Physiol. 1981;144(4):465–479.

56. White JA, Klink R, Alonso A, Kay AR. Noise from voltage-gated ion channels may influence neuronal dynamics in the entorhinal cortex. J Neurophysiol. 1998;80(1):262–269.

57. White JA, Rubinstein JT, Kay AR. Channel noise in neurons. Trends Neurosci. 2000;23(3):131–137.

58. Schneidman E, Freedman B, Segev I. Ion channel stochasticity may be critical in determining the reliability and precision of spike timing. Neural Comput. 1998;10(7):1679–1703.

59. Van Rossum M, O'Brien BJ, Smith RG. Effects of noise on the spike timing precision of retinal ganglion cells. J Neurophysiol. 2003;89(5):2406–2419.

60. Brown D, Adams P. Muscarinic suppression of a novel voltage-sensitive K+ current in a vertebrate neurone. Nature. 1980;283(5748):673–676.

61. Traub RD, Miles R. Neuronal Networks of the Hippocampus. vol. 777. Cambridge University Press; 1991.

62. Ermentrout B. Linearization of FI curves by adaptation. Neural Comput. 1998;10(7):1721–1729.

63. Jayant NS, Noll P. Digital Coding of Waveforms: Principles and Applications to Speech and Video. Englewood Cliffs, NJ. 1984; p. 115–251.

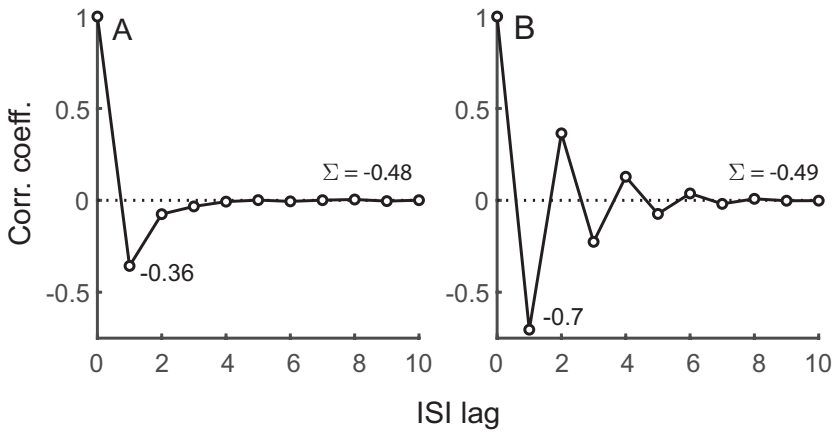

Fig 1. Representative serial interspike interval correlations (SCCs) from two P-type primary electrosensory afferent spike trains from two different fish. Panels depict patterns of correlation observed in the data. A. Type I ($\rho_1 > -0.5$): Non-bursting unit with first SCC $\rho_1 = -0.36$. Remaining $\rho_k < 0$ diminish to 0. Sum of SCCs (Σ) over 15 lags is -0.48. B. Type II ($\rho_1 < -0.5$): Strongly bursting unit with $\rho_1 = -0.7$ with marked alternating positive and negative correlations. Sum of SCCs (Σ) over 15 lags is -0.49. Spike trains sampled at 60 μs. Mean ISI $\pm$ SD (in ms): 2.42 ± 0.72 (A), and 6.04 ± 3.59 (B). Electric organ discharge (EOD) frequency: 948 Hz (A), and 990 Hz (B).

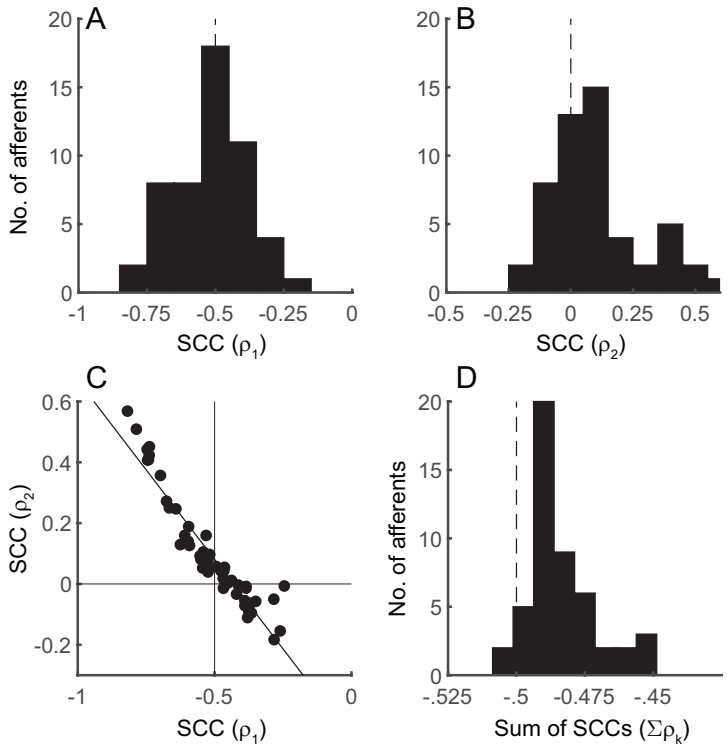

Fig 2. Population summaries of interspike interval (ISI) serial correlation coefficients (SCCs) for P-type primary afferent spike-trains($N = 52$) [1]. A. Histogram of SCC between adjacent ISIs (ρ_1), $\bar{\rho}_1 = -0.52 \pm 0.14$ (mean $\pm$ s.d.), range -0.82 to -0.25 B. Histogram of SCC between every other ISI (ρ_2), $\bar{\rho}_2 = 0.10 \pm 0.18$, range -0.18 to 0.57. C. Anti-diagonal relationship between observed ρ_1 (abscissa) and ρ_2 (ordinate) (filled circles). Line describes best fit, $\rho_2 = -1.18\rho_1 - 0.51$ with SE $= 0.008$. D. Mean sum of SCCs for the population over 15 lags, $\sum \rho_k = -0.475 \pm 0.04$. The population histograms in panels A and B were reported earlier with a different bin width (see [1], Figs. 7A and 7B therein, respectively).

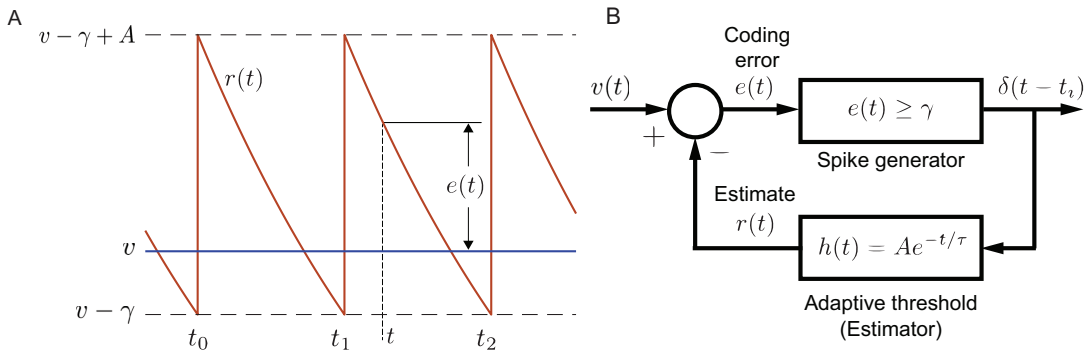

Fig 3. Schematic of neuron with adaptive threshold coding. A. v is a constant bias voltage that generates spontaneous activity, $r(t)$ is an adaptive threshold, γ is a spike threshold such that a spike is emitted when $v - r(t) = \gamma$. Following a spike, the adaptive threshold suffers a jump of fixed magnitude A. The membrane voltage is non-resetting. Spike times are t_0, t_1, t_2, ..., etc. Historically, the spike threshold was set to zero, i.e., $\gamma = 0$ and a spike is fired when $v = r(t)$. Although it makes no difference to the results presented here, the more general form with γ is considered for reasons entered in the discussion. B. The spike encoder with adaptive threshold can be viewed as a feedback control loop where the spike-train $\delta(t - t_i)$ is filtered by $h(t)$ to generate the time-varying adaptive threshold $r(t)$ which can be viewed as an estimate of the membrane voltage $v(t)$. The comparator generates the estimation error $e(t) = v - r(t)$ (see also Panel A). The estimation or coding error drives the spike generator, and a spike is fired whenever $e(t) \geq \gamma$. The simplest form for the adaptive threshold (the estimation filter) $h(t)$ is a low-pass filter (an RC membrane) given by $h(t) = A\exp(-t/\tau)$. This is the form shown in panel A.

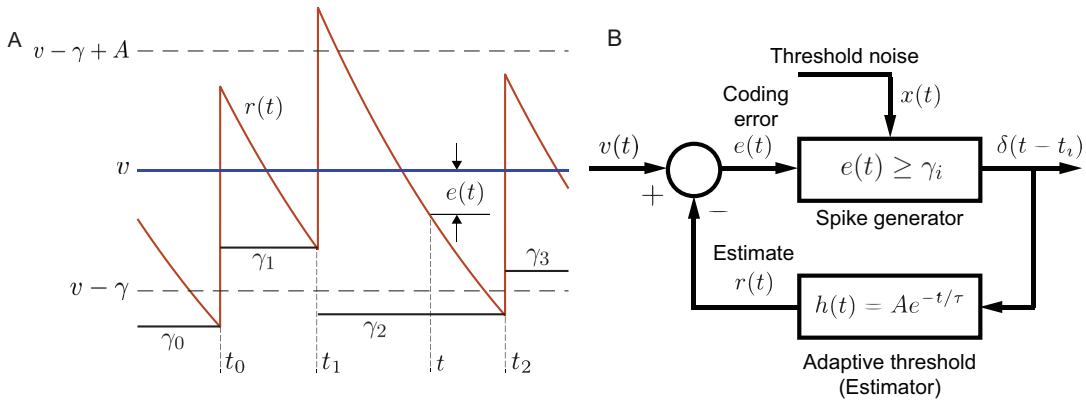

Fig 4. Adaptive threshold neuron with a stochastic firing threshold. Panel descriptions follow Fig. 3 and only differences are noted. A. The spike threshold is $\gamma_i = v - r(t)$ where γ_i is a random value generated at t_{i-1}^{+}, and held constant until the next spike at t_i. The discrete noise sequence γ_i is generated by a discrete wide-sense stationary process with mean γ and unknown discrete auto-correlation function $\mathbf{R}(k)$. The goal is to determine $\mathbf{R}(k)$ which will generate a prescribed sequence of ISI serial correlation coefficients ρ_k. To reduce clutter, the spike threshold $v - \gamma_i$ is depicted as γ_i. B. Block diagram showing the stochastic modification of the spike threshold. Symbols and additional description as in text and Fig. 3.

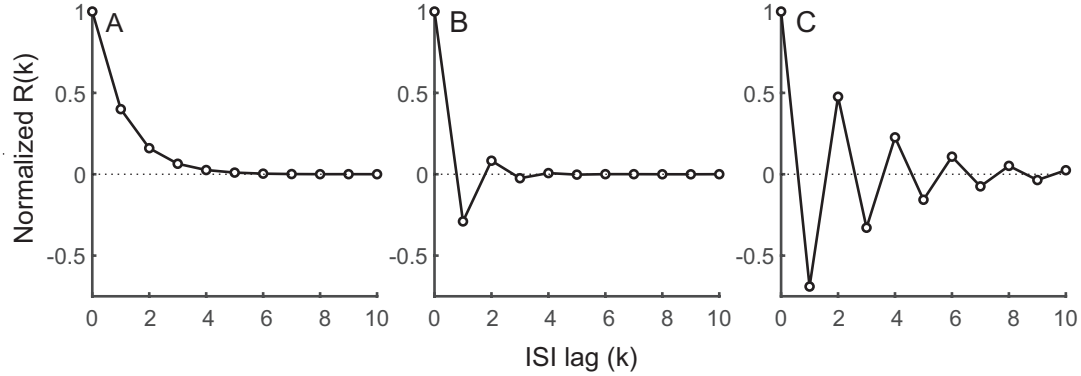

Fig 5. Normalized auto-correlation functions of threshold noise. A. Type I SCC matched to the afferent depicted in Fig. 1A with no bursting activity. B. Type II SCC from an afferent with moderate bursting activity. C. Type II SCCs matched to the afferent depicted in Fig. 1B with strong bursting activity. Autocorrelation function that matches Type III SCCs will be zero for all non-zero lags (not shown).

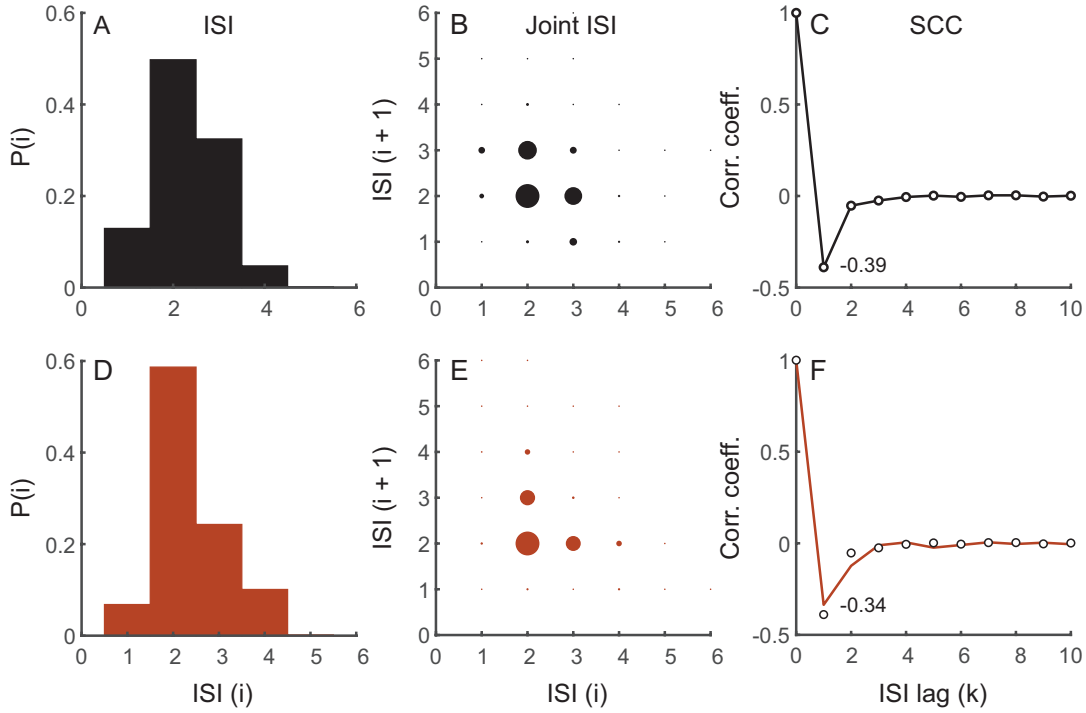

Fig 6. Type I serial correlation coefficients (SCCs). Top row depicts experimental spike-train from a P-type primary electrosensory afferent. A. The spike train has a unimodal interspike interval distribution (ISI) and does not display bursting activity. Abscissa is ISI in multiples of electric organ discharge (EOD) period (1.05 ms) and ordinate is probability. B. Joint interspike interval distribution showing probability of observing successive intervals ISI (i) (abscissa) and ISI $(i+1)$ (ordinate). The size of the circle is proportional to the probability of jointly observing the adjacent ISIs, i.e., $P\,(i,\, i+1)$. C. Serial correlation coefficients (ordinate) as a function of lag measured in multiples of mean ISI (abscissa). Spike-train sampled at EOD period. SCCs for this afferent are shown in Fig. 1A at a different sampling rate (see Materials and Methods). Bottom row depicts results for matched model using low-pass filtered white noise. Panel descriptions as in top row, except in F where open circles denote experimental data taken from C. To generate model results, v: 1.845 V, adaptive threshold parameters: A: 0.15 and time-constant: 30 ms, low-pass filter time constant τ_γ: 2.74 ms ($a = 0.4$), and $\mathbf{R}\,(0)$: 1.07×10^{-3} V^2 (SNR = 35 dB).

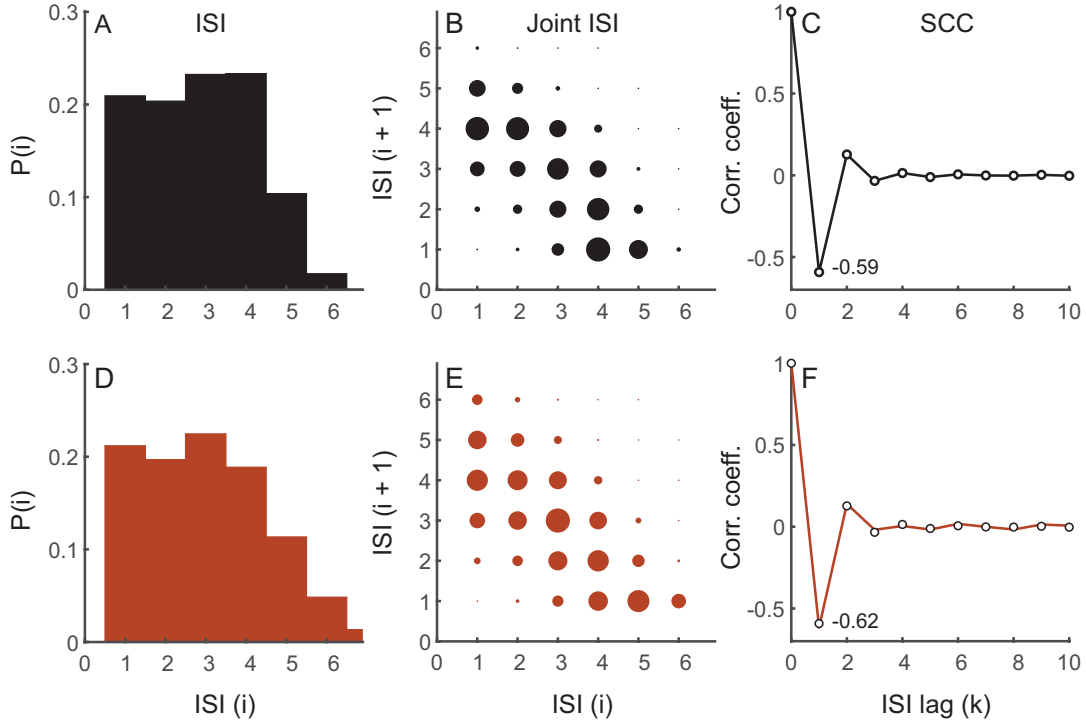

Fig 7. Type II serial correlation coefficients (SCCs) with moderate bursting activity. Description is identical to Fig. 6. Top row depicts experimental spike train from a P-type primary electrosensory afferent. Bottom row depicts results for matched model using high-pass filtered white noise. To generate model results, v: 1.845 V, adaptive threshold parameters: A: 0.28 and time-constant: 26 ms, high-pass filter time constant τ_γ: 3.18 ms ($a = 0.29$), and $\mathbf{R}(0)$: 9.22×10^{-3} V^2 (SNR = 26 dB).

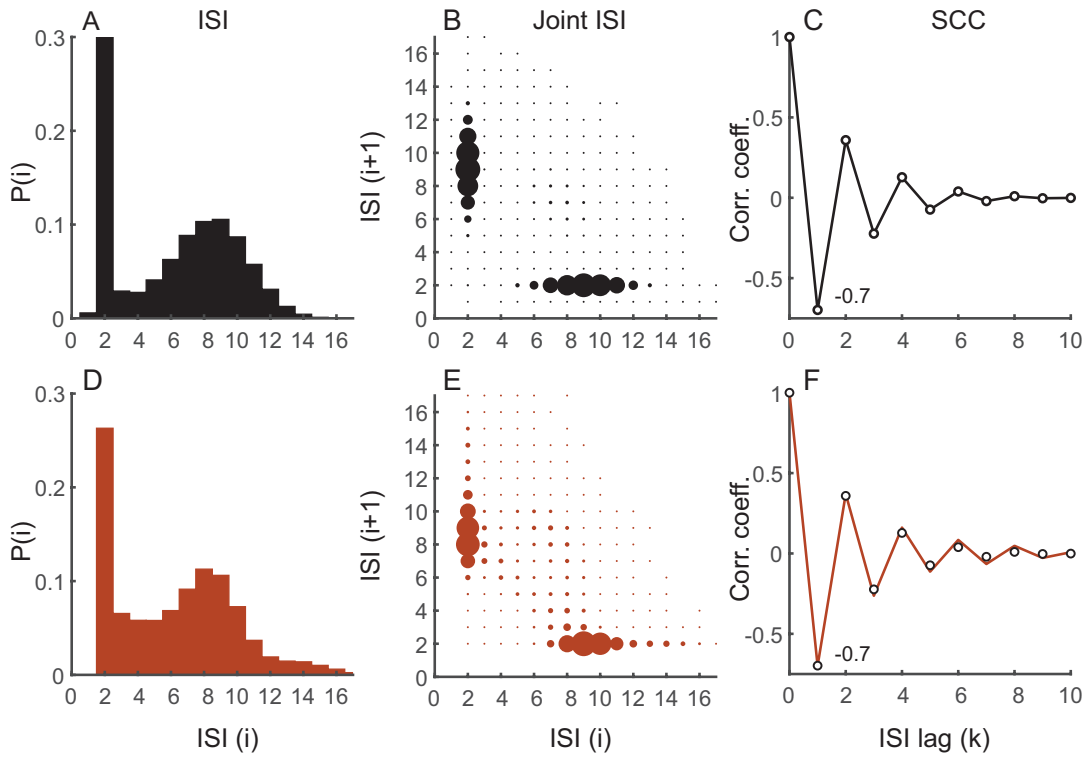

Fig 8. Type II serial correlation coefficients (SCCs) with strong bursting activity. Description is identical to Fig. 6. Top row depicts experimental spike train from a P-type primary electrosensory afferent. SCCs for this afferent are shown in Fig. 1B at a different sampling rate (see Materials and Methods). Bottom row depicts results for matched model using high-pass filtered white noise. To generate model results, v: 1.845 V, adaptive threshold parameters: A: 0.19 and time-constant: 60 ms, high-pass filter time constant τ_γ: 16.89 ms ($a = 0.69$), and $\mathbf{R}(0)$: 6.4×10^{-3} V^2 (SNR = 27 dB).

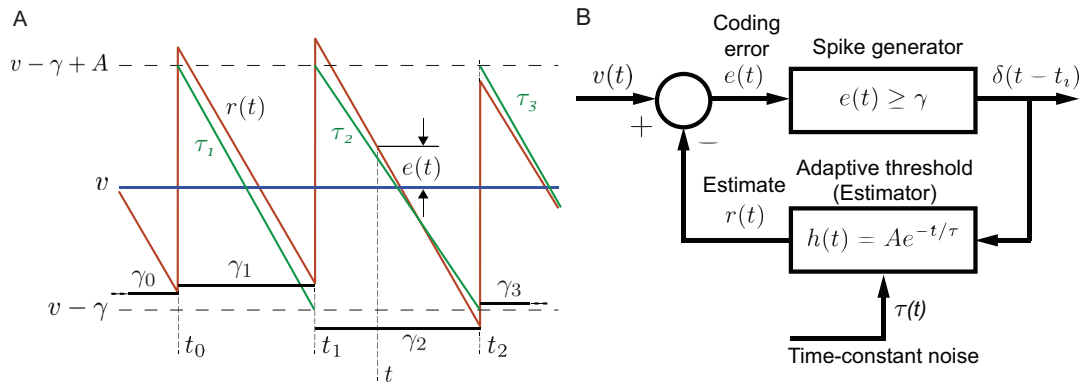

Fig 9. Adaptive threshold with a stochastic filter time-constant. A. Geometry of the spiking mechanism. The input is a deterministic piece-wise constant signal (v, here shown to be DC-valued) and the spike threshold is $\gamma = v - r(t)$. Two equivalent spike generation processes are possible: Case (i) The spike threshold γ is noisy while the adaptive threshold filter $h(t)$ has fixed time-constant τ. This is the same as Fig. 4A where γ_i is a discrete random variable that is generated at t_{i-1}^+ and held constant until the next spike at t_i (black). The time-varying adaptive threshold is shown in red. Case (ii) The adaptive filter $h(t)$ has a random time-constant with mean τ while the spike threshold is γ, and is constant. Threshold noise γ_i from Case (i) can be transformed into a noisy filter time-constant τ_i which is fixed between spikes, and takes correlated values over successive ISIs. The estimate from this process is shown in green. Note that the spike-times for both cases are identical. B. Block diagram illustrating the feedback from the estimator that generates optimal spike times. The time-constant τ is a random variable that is constant between spikes, and is correlated over interspike intervals (ISIs). It is of interest to ask what is the relationship between the correlation functions in the two cases. Symbols and additional description as in text.