2007

Decimal Spike Code Maximizes Neural Memory Against Retrieval Time

Bo Deng
University of Nebraska-Lincoln, bdeng@math.unl.edu

[Issued 2007.]

Decimal Spike Code Maximizes Neural Memory Against Retrieval Time

Bo Deng[1]

Abstract: For a new class of circuit models for single neurons we demonstrate here that if the circuits' plastic spike-bursts are used for memory alphabet, then using the first 10 bases tend to maximize the amount of information that can be retrieved from and stored into the circuits in a unit time.

1. Introduction. Any computing device has to have the following basic components: an alphabet to code information, a processor to manipulate the information, and a memory to store the input/output of the processor. The Chinese abacus, *suanpan*, uses a number system in beads by which numbers are manipulated essentially in 4 bases but stored in the increments of 5s and 10s. Its upper and lower beads on rods serve both as the processer and the storage. Modern computers use the binary alphabet for information coding, central processing units (CPU) for data manipulation, random access memories (RAM) for temporary storage and some memory devices for long-term storage. Regardless of a device's computing logics or its data's semantics, its processer and memory are characterized by the amount of information that can be shuttled into and out from each component in a unit of time, which we will refer to as the *information retrieval rate* or just the *retrieval rate* for short. As fundamental as it must be, the retrieval rate question rarely caught the attention of theoreticians because all manmade devices, electronic devices in particular, come with preset performance parameters which either include or determine their retrieval rates. However, the situation is completely different when a neuron is assumed to function as a processor or a memory storage for neural computing. Nature must have installed a capacity retrieval rate to each neuron and we know practically nothing about it.

We will study the neural retrieval rate problem using a new class of circuit models introduced in [3]. More specifically, we will use the transmembrane spike-bursts of the models for the memory alphabet, and demonstrate that using the first 10 spike-burst bases tends to maximize the retrieval rate of the circuits.

2. Neural Circuit Models. In this paper we will view a neuron as an information processing unit either as a logic gate or as a memory device. Whatever the neural information maybe, it must be coded by a physical state of of the neuron. Following the view of [5], the alphabetical state need be a transient and plastic state for flexibility, but a steady-state-like state for robustness for high tolerance against errors. Spike-bursting states of the following circuit models of neurons from

[1]Department of Mathematics, University of Nebraska-Lincoln, Lincoln, NE 68588. Email: bdeng1@math.unl.edu

Table 1: Circuit Variable and Parameter Legend

		Variables
$V_{\rm C}$	—	Membrane voltage with the default current direction being outward.
$A_{\rm Na}$	—	Outward current through the Na^+ ion pump with $A_{\rm Na} > 0$.
$-A_{\rm K}$	—	Inward current through the K^+ ion pump with $A_{\rm K} > 0$.
$I_{\rm Na,p}$	—	Current through Na^+'s serial passive channels.
$I_{\rm A} = A_{\rm Na} - A_{\rm K}$	—	Net current through the parallel Na^+-K^+ ion pump.
$I_{\rm S} = A_{\rm Na} + A_{\rm K}$	—	Absolute current through the parallel Na^+-K^+ ion pump, $I_{\rm S} > 0$.
$I_{\rm ext}(t)$	—	Inward external current.
		Parameters
$\bar{E}_{\rm Na} > 0$	—	The resting potential of Na^+'s passive channels.
$\bar{E}_{\rm K} < 0$	—	The resting potential of K^+'s passive channels.
C	—	Membrane capacitance.
$\lambda_{\rm J}$	—	Ion J's pump coefficient.
$\gamma_{\rm J}$	—	Ion J's pump resistance.
		IV-Characteristics and Parameters
$I = f_{\rm K}(V)$	—	For K^+'s electro and diffusive channels in parallel, with maximal conductance $g_{\rm K} > 0$, minimal diffusion $d_{\rm K} + g_{\rm K} < 0$, and the diffusion domination voltage range $[v_1, v_2]$ with $v_1 > 0$. Specifically, $f_{\rm K}(V) = g_{\rm K}V + d_{\rm K}(V - v_1)(v_1 < V < v_2) + d_{\rm K}(v_2 - v_1)(v_2 < V)$.
$V = h_{\rm Na}(I)$	—	For Na^+'s electro and diffusive channels in series, with maximal conductance $g_{\rm Na} > 0$, minimal diffusion $g_{\rm Na}d_{\rm Na}/(d_{\rm Na} + g_{\rm Na}) < 0$, and the diffusion domination current range $[i_1, i_2]$ with $i_1 > 0$. Specifically, $h_{\rm Na}(I) = \frac{1}{g_{\rm Na}}I + \frac{1}{d_{\rm Na}}(I - i_1)(i_1 < I < i_2) + \frac{1}{d_{\rm Na}}(i_2 - i_1)(i_2 < I)$.

[3, 4, 5] meet these criteria,

$$\begin{cases} CV_{\rm C}{}' = -[I_{\rm Na,p} + f_{\rm K}(V_{\rm C} - \bar{E}_{\rm K}) + A_{\rm Na} - A_{\rm K} - I_{\rm ext}] \\ A_{\rm Na}{}' = \lambda_{\rm Na}A_{\rm Na}[V_{\rm C} - \gamma_{\rm Na}(A_{\rm Na} - \delta A_{\rm K})] \\ A_{\rm K}{}' = \lambda_{\rm K}A_{\rm K}[-V_{\rm C} + \gamma_{\rm K}(\delta A_{\rm Na} - A_{\rm K})] \\ \epsilon I_{\rm Na,p}{}' = V_{\rm C} - \bar{E}_{\rm Na} - h_{\rm Na}(I_{\rm Na,p}). \end{cases} \tag{1}$$

The definitions for the variables and parameters of the equations are given in Table 1, where a `Matlab` notation is used for unit step functions: $(a < x) = 1$ if and only if $a < x$ and $(a < x) = 0$ otherwise. Also, $(a < x < b) = (a < x) \cdot (x < b)$. The equations above represent only two prototypical types of many from [3]: the $pK^+_-sNa^+_+$ model when $\delta = 1, \gamma_{\rm Na} = \gamma_{\rm K} = \gamma$, and the $pK^+_{-d}sNa^+_{+d}$ model for $\delta = 0$. For the former, the Na^+ and K^+ ion pumps are coupled as one unit and for the latter, the pumps are decoupled, each operating independently from the other.

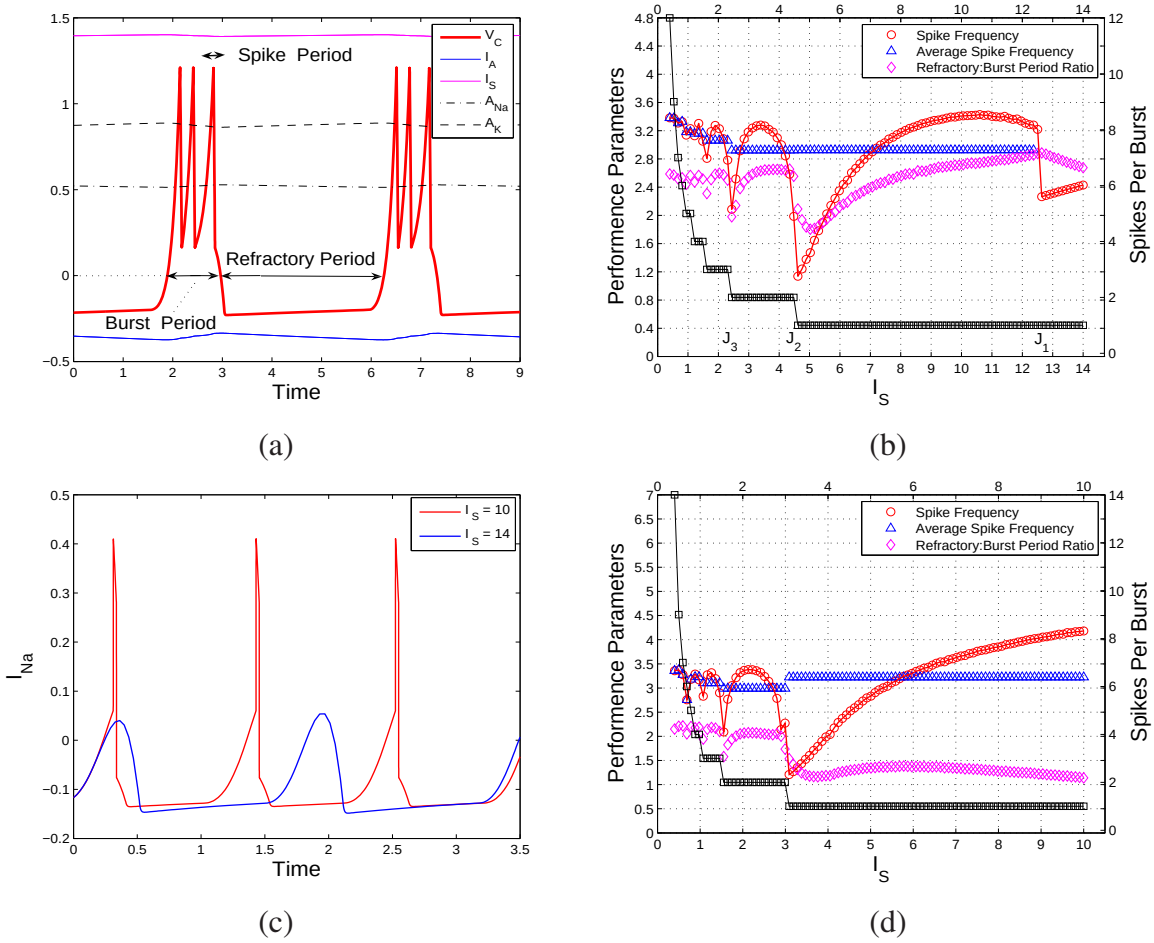


Figure 1: (a) Terminology Legend. (b) The $\text{pK}^+_-\text{sNa}^+_+$ model ($\delta = 1$ for Eq.(1)) with parameter values: $g_{\text{K}} = 1, d_{\text{K}} = -1.25, v_1 = 0.5, v_2 = 2, g_{\text{Na}} = 0.16, d_{\text{Na}} = -0.1, i_1 = .06, i_2 = 0.28, g_{\text{Cl}} = 0.01, d_{\text{Cl}} = 0, \bar{E}_{\text{Na}} = 0.6, \bar{E}_{\text{K}} = -0.7, \bar{E}_{\text{Cl}} = -0.6, \lambda_{\text{Na}} = \lambda_{\text{K}} = 0.05, \gamma_{\text{Na}} = \gamma_{\text{K}} = 0.1, \delta = 1, C = 0.01, I_{\text{ext}} = 0, \epsilon = 0.0005$. (c) 1-spike bursts for an I_{S} value from (J_2, J_1) and a pulse-only burst for another I_{S} value $> J_1$. (d) The $\text{pK}^+_{-d}\text{sNa}^+_{+d}$ model ($\delta = 0$ for Eq.(1)) with the same IV-curves. The same parameter values as (b) except $\lambda_{\text{Na}} = \lambda_{\text{K}} = 0.1, \gamma_{\text{Na}} = 0.1, \gamma_{\text{K}} = 0.05, \delta = 0$.

The key variable of our focus is the absolute ion pump current $I_{\text{S}} = A_{\text{Na}} + A_{\text{K}} > 0$ through the two-way Na^+-K^+ ion pump across the cell membrane. It was demonstrated in [4, 5] that for each fixed set of some values in the voltage variable V_{C}, the net pump current I_{A}, and the passive sodium current $I_{\text{Na,p}}$, the domain of the absolute pump current I_{S} can be partitioned into subintervals $0 < \cdots < J_n < J_{n-1} < \cdots < J_2 < J_1$ so that for $I_{\text{S}} \in (J_{k+1}, J_k)$, the circuit dynamics gives rise to a k-spike burst. Fig.1(a) shows a typical 3-spike burst. It looks like an asymptotic state but in fact only a transient state — if the time lasts long enough the 3-spike burst will repeat a few times before turning into a 2-spike or 4-spike burst. Such a transient but steady-state-like state is referred to as a *metastable* state and the property of a metatstable state changing into another metastable state of a different spike number is referred to as *plastic*, see [4, 5].

Fig.1(b) captures the metastability and plasticity in a more systematic way. Here, each spike-

burst starts at a set of fixed values in $V_{\rm C}, I_{\rm A}, I_{\rm Na,p}$ but a varying value in $I_{\rm S}$ along the horizontal axis. The fixed values correspond to an end boundary point of the refractory region with $V_{\rm C}(0) = 0, V'_{\rm C}(0) > 0$ and the spike-bursts are defined to end at an initial boundary point of the refractory region with $V_{\rm C}(\tau) = 0, V'_{\rm C}(\tau) < 0$, where $\tau = \tau(I_{\rm S})$ is the time duration, or the *burst period* of a k-spike-burst if $I_{\rm S} \in (J_{k+1}, J_k)$ for some natural number $k \geq 1$. For each $I_{\rm S} > 0$, the plot shows these quantities: the number of spikes of the immediate spike-burst, the spike frequency $k/\tau(I_{\rm S})$ if $I_{\rm S} \in (J_{k+1}, J_k)$, the average spike frequency $k/\bar{\tau}_k$ with $\bar{\tau}_k$ being the average k-spike burst period:

$$\bar{\tau}_k = \frac{1}{J_k - J_{k+1}} \int_{J_{k+1}}^{J_k} \tau(I_{\rm S}) dI_{\rm S}, \tag{2}$$

and the refractory:burst period ratio. The k-spike-burst interval (J_{k+1}, J_k) lies directly underneath the step function plot for the per-burst spike number for $k \geq 2$. For the 1-spike-burst interval, the end point J_1 is defined when the unit spike is changed qualitatively as shown in Fig.1(c), showing all bursts losing their spikes for $I_{\rm S} > J_1$. For the remainder discussion, we will measure the spike-burst ending point J_1 by its proportionality, β, against the 1-spike-burst ending point J_2:

$$J_1 = \beta J_2 \quad \text{or} \quad \beta = \frac{J_1}{J_2}. \tag{3}$$

To be used later, the β-value is estimated to be $\beta = 12.5/4.6052 = 2.7143$ and $\beta > 3.3$ for the circuit of Fig.1(b) and Fig.1(d), respectively.

3. Memory Entropy. For the remainder of the paper, whenever a k-spike burst appears as the dynamical state of the circuit, we will consider it to present the kth base (or letter), b_k, of an alphabet, $\mathcal{A}$, which is referred to as the memory alphabet of the circuit model, see Table 2 for its definition and other definitions.

For physical representations of memory bases, we have used punch cards, magnetized tapes, semiconductors, etc. for manmade devices. In the case of a DVD disc, for example, it is a piece of the physical space of the disc that is altered in one of two possible ways to store a binary bit 0 or 1. For our neuron circuit models, we will assume instead that it is the state of the plastic variable $I_{\rm S}$ in the range $(0, J_1)$ that codes a particular spike-burst base. That is, the physical medium for the memory alphabet is the absolute ion pump current $I_{\rm S}$. In fact, as shown in [5], the power through the Na^+-K^+ pump for most spike-bursts is approximately proportional to the current $I_{\rm S}$. In other words, the alphabet's physical medium can be considered to be the power spectrum of the spike-bursts, which in turn is related to the biochemical energy conversion of neurons by their intercellular ARPass process. In this sense, we refer to (J_{k+1}, J_k) as the *base interval*, and $w_k = J_k - J_{k+1}$ as the *base width* for the k-spike burst base b_k.

Since the absolute ion pump current $I_{\rm S}$ for a working neuron should never go down to zero, we expect that there is an effective lower limit of the variable, which without loss of generality is taken to be J_{n+1} for some $n \geq 2$. That is, the power spectrum interval $[J_{n+1}, J_1]$ is the expected effective and practical range for the spike-bursts of a given circuit model. The range in turn is partitioned

Table 2: Definitions for Memory Parameters

$\mathcal{A}_n = \{b_1, b_2, \ldots, b_n\}$	—	Spike-burst alphabet.
(J_{k+1}, J_k)	—	Base b_k's coding interval in variable $I_{\rm S}$.
$w_k = \lvert J_k - J_{k+1}\rvert$	—	Base width (or memory width) for base b_k.
$q = \{q_1, q_2, \ldots, q_n\}$	—	Partitioning distribution for the n-base alphabet $\mathcal{A}_n$ with $q_k = w_k / \sum_{i=1}^n w_i$.
$H_n(q) = \sum_{k=1}^n q_k \log_2 \frac{1}{q_k}$	—	Memory entropy in bit per base.
$\tau(I_{\rm S})$ with $I_{\rm S} \in (J_{k+1}, J_k)$	—	Burst period or base processing time for base b_k at $I_{\rm S}$.
$\tau_{{\rm min},k} = \min_{\{I_{\rm S}\in(J_{k+1},J_k)\}} \tau(I_{\rm S})$	—	The minimal processing time for base b_k.
$\bar{\tau}_k = \frac{1}{w_k}\int_{J_{k+1}}^{J_k} \tau(I_{\rm S}) dI_{\rm S}$	—	The average processing time for base b_k.
$T_n(q, \bar{\tau}) = \sum_{k=1}^n \bar{\tau}_k q_k$	—	The averaged per-base processing time with $T_n(q, \bar{\tau}) = \frac{1}{\lvert J_1 - J_{n+1}\rvert}\int_{J_{n+1}}^{J_1} \tau(I_{\rm S}) dI_{\rm S}$.
$T_n(q, \tau_{\rm min}) = \sum_{k=1}^n \tau_{{\rm min},k} q_k$	—	The averaged per-base minimal processing time with $T_n(q, \tau_{\rm min}) = \frac{1}{\lvert J_1 - J_{n+1}\rvert}\sum_{k=1}^n \min_{\{\tau\}}\int_{J_{k+1}}^{J_k} \tau(I_{\rm S}) dI_{\rm S}$.
$R_n(q, \bar{\tau}) = H_n(q)/T_n(q, \bar{\tau})$	—	The average or mean retrieval rate in bit per time.
$C_n(q, \tau_{\rm min}) = H_n(q)/T_n(q, \tau_{\rm min})$	—	The capacity retrieval rate in bit per time.

into a collection of n subintervals, $\{(J_{k+1}, J_k) : 1 \le k \le n\}$ with each subinterval corresponding to a spike-burst base of the memory alphabet $\mathcal{A}_n$.

The questions of why does a spike-burst appear and what does it mean are not the concern of this paper. The first type question may have to do with if it is for a computational manipulation or a memory storage and the second type question may have to do with if it codes, e.g., a piece of sensory information or internal memory. Instead, we will consider the likely probabilistic way in which a spike-burst appears.

From the perspective of the circuit, it does not exist for a particular type of computing logic nor for a particular type of information. Thus, its memory medium $I_{\rm S}$ is expected to be used equally in the effective range $[J_{n+1}, J_1]$. From the perspective of the computing logics and memory data, especially from the view of biological evolution, the effective coding range of the circuit is a black box in the sense that any piece is as good as any other piece, and any piece will be used by a particular logic or particular memory, and on average any piece will be used equally likely as any other piece. In other words, the 0th order approximation to the intrinsic amount of information that a spike-burst base carries is the amount when each encoding value $I_{\rm S}$ is used equally in the effective range $[J_{n+1}, J_1]$. Probabilistically speaking, $I_{\rm S}$ as a random variable assumes the uniform distribution over $[J_{n+1}, J_1]$.

Given this uniform distribution of $I_{\rm S}$ over its effective range for all possible logics and memory data, the amount of information that the k-spike burst base b_k carries depends on the (conditional) probability that it occurs, and the probability is its base width to the total width of the effective

range: $q_k = w_k / \sum_{i=1}^n w_i$. According to Shannon ([6]), base b_k contains $\log_2 \frac{1}{q_k}$ bit of information, and on average, a spike-burst base contains

$$H_n(q) = \sum_{k=1}^{n} q_k \log_2 \frac{1}{q_k}$$

bit of information per base. For example, if the effective range codes the first 12 spike-bursts for the circuit Fig.1(b) as shown, then on average a base contains about $H_{12} = 1.6912$ bits of information. Similarly, $H_n = 1.5706,\ 1.6122,\ 1.6436,\ 1.6679$ for $n = 8, 9, 10, 11$, suggesting the more memory bases the greater information coded per base on average.

4. Average and Minimum Retrieval Times. The processing time of a particular alphabetical state in this paper is defined to be the time the state takes to express itself. Depending on the particular nature of the process which the state is in, the process may be construed for computational manipulation or for data storage. We will refer to the processing time collectively without distinction as the *retrieval time*.

For a k-spike burst with $I_{\rm S} \in (J_{k+1}, J_k)$, the particular retrieval time is the burst period $\tau(I_{\rm S})$, c.f. Table 2. Because of the assumption of the uniform distribution of the encoding variable $I_{\rm S}$ over the base interval (J_{k+1}, J_k) for b_k, the average retrieval time for b_k is $\bar\tau_k$ as defined by (2), see also Table 2. As a result, a typical spike-burst base takes on average $T_n(q, \bar\tau) = \sum_{k=1}^b q_k \bar\tau_k$ in time per base.

On the other hand, each base has a minimal retrieval time $\tau_{{\rm min},k} = \min_{I_{\rm S}\in(J_{k+1},J_k)} \tau(I_{\rm S})$ which is equal to $\tau(I_{\rm S,min})$ for some $I_{\rm S,min} \in (J_{k+1}, J_k)$, and the averaged minimal retrieval time is $T_n(q, \tau_{\rm min}) = \sum_{k=1}^b q_k \tau_{{\rm min},k}$ in time per base. Note that the relation $T_n(q, \tau_{\rm min}) < T_n(q, \bar\tau)$ automatically holds.

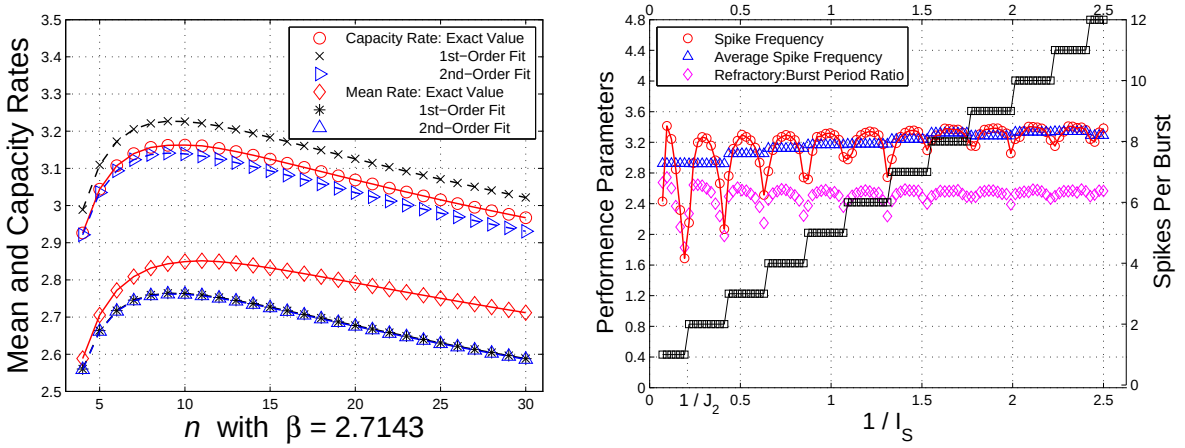
Fig.1(b,d) show some plots related to the retrieval times. More specifically, they show the spike frequency $k/\tau(I_{\rm S})$ as a function of $I_{\rm S} \in (J_{k+1}, J_k)$ for all k as well as the average spike frequency $k/\bar\tau_k$. The maximal spike frequency is related to the minimal burst period by $k/\tau_{{\rm min},k}$ for each k. With the individual bases' periods derivable from the spike frequency plots, the two types of averaged retrieval times can be computed by definition. For example, the numerical values for the average retrieval time and the averaged minimal retrieval time for the circuit of Fig.1(b) are $T_{12}(q, \bar\tau) = 0.5936,\ T_{12}(q, \tau_{\rm min}) = 0.5362$, respectively.

5. The Result. Thus, for a neural circuit with an effective memory region $I_{\rm S} \in [J_{n+1}, J_1]$ for some $n \geq 2$, the average amount of information that can be processed or retrieved or stored in a unit amount of time is by definition:

$$R_n(q, \bar\tau) = \frac{H_n(q)}{T_n(q, \bar\tau)} = \frac{\sum_{k=1}^n q_k \log_2(1/q_k)}{\sum_{k=1}^n q_k \bar\tau_k},$$

in bit per time. We will call it the *mean retrieval rate*. On the other hand, the maximal possible retrieval rate is by definition

$$C_n(q, \tau_{\rm min}) = \frac{H_n(q)}{T_n(q, \tau_{\rm min})} = \frac{\sum_{k=1}^n q_k \log_2(1/q_k)}{\sum_{k=1}^n q_k \tau_{{\rm min},k}},$$



(a) (b)

Figure 2: (a) The mean rate R_n and the capacity rate C_n and their 1st-order and 2nd-order approximations. (b) The same plot as Fig.1(b) except against the reciprocal $1/I_{\rm S}$.

which is referred to as the *capacity retrieval rate*. It holds automatically that $C_n(q, \tau_{\min}) > R_n(q, \bar{\tau})$.

The central question is what is the size of the memory alphabet n so that the retrieval rates R_n, C_n are maximal? For the circuit of Fig.1(b), the numerical rates of R_n, C_n for $n = 2, 3, 4, \ldots, 30$ are shown in Fig.2(a). It shows that $C_{10} \geq C_n$ and $R_{11} \geq R_n$. In other words, the optimal choice in the number of memory alphabet for the circuit is around 10.

For the remainder of this section, we will derive some simpler approximating formulas for the mean and capacity rates based on an empirical general property of the circuit models so that we do not have to reply exclusively on the brute-force of numerical calculations for all components of the rates. The empirical property is gleaned from Fig.2(b) which clearly shows a linear regression between $1/I_{\rm S}$ and the bursting spike number n. Thus, for some parameter $\mu > 0$ we will assume $1/J_k - 1/J_2 = \mu(k-2)$ which in turn is

$$J_k = \frac{J_2}{1 + J_2\mu(k-2)}, \quad \text{for } \; k \geq 2. \tag{4}$$

Together with the relationship (3), the memory distribution q can be calculated, and hence the memory entropy $H_n(q)$. For the circuit model of Fig.2(b), the numerical value of J_2 is 4.6053 and the least-square best fit of μ is 0.2246.

Next, we approximate the average and the averaged minimal spike-burst periods by what we refer to as the mth-order approximations. We use the mean rate R_n for the illustration because the case for the capacity rate C_n is completely analogous. Again, Fig.2(b) gives an empirical basis for the approximations. More specifically, the spike frequencies $k/\bar{\tau}_k$ appear to cluster near a constant, and the approximating scheme is based on how that constant is chosen for a particular approximation. We will denote by $\sigma_k = \bar{\tau}_k/k, \; k \geq 1$, which represents the averaged spike period for the k-spike bursts. By definition, the 1st-order approximation of the average k-spike burst

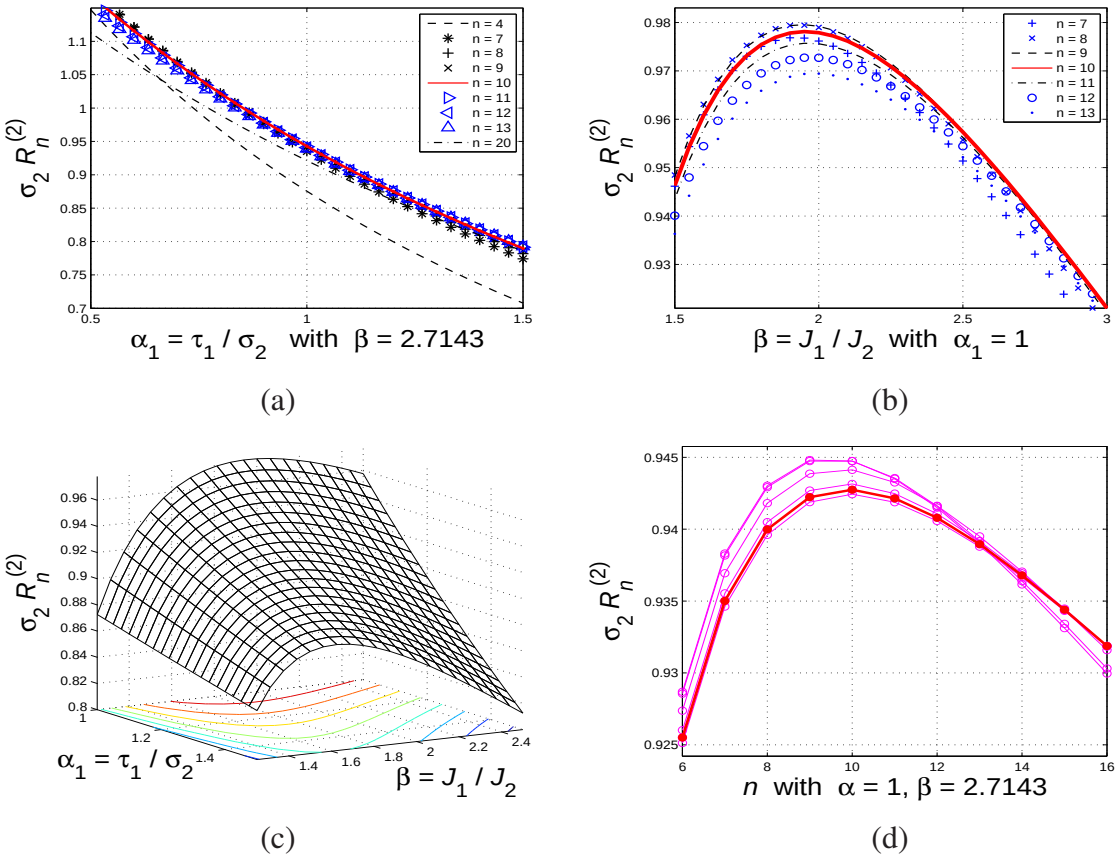


Figure 3: (a) A dimensionless 2nd-order mean rate comparison for the circuit model of Fig.1(b). (b) A similar plot but against parameter β. $R_{10}^{(2)}$ dominates for $\beta > 2.5$. (c) The dimensionless mean rate of $R_{10}^{(2)}$ for the same model as a function of α_1 and β. (d) $\sigma_2 R_n^{(2)}$ v.s. n for a fixed set of α_1, β. Six models are shown for which the μ and J_2 values are $\mu = 0.30, 0.43, 0.01, 0.10, 0.19, 0.10$, respectively, $J_2 = 4.50, 2.86, 13.33, 14.29, 5.71, 11.11$. The first, the second, and the third pairs of the μ values are for the piecewise linear, the arctangent, and the cubic nonlinearities, respectively. The first and the second μ-value of each pairs is for the $\text{pK}^+_-\text{sNa}^+_+$ and $\text{pK}^+_{-d}\text{sNa}^+_{+d}$ model, respectively. $R_{10}^{(2)}$ becomes more dominant for larger β.

period $\bar{\tau}_k$ is

$$\frac{k}{\bar{\tau}_k} = \frac{1}{\sigma_1} \quad \text{for all} \quad k \geq 1, \quad \text{equivalently} \quad \bar{\tau}_k = k\sigma_1. \tag{5}$$

That is, only the first burst period $\bar{\tau}_1$ remains exact but the remaining ones are approximated. This gives rise to the 1st-order approximation of the average burst period, $T_n(q, \bar{\tau})$ with $\bar{\tau}$ replaced by (5), and the corresponding 1st-order approximation, denoted by $R_n^{(1)}$, of R_n. Note that $R_n^{(1)}$ is parameterized by these parameters $J_2, \mu, \beta, \sigma_1$ which can be estimated from plots such as Fig.2(b).

By definition as well, the 2nd-order approximation of the average burst period $\bar{\tau}_k$ is obtained by assuming the same spike frequency σ_2 for spike frequencies $k/\bar{\tau}_k$ of all $k \geq 2$ and allowing the first two burst periods to remain exact:

$$\bar{\tau}_k = k\sigma_2 \quad \text{for all} \quad k \geq 2, \quad \text{and} \quad \bar{\tau}_1 = \alpha_1\sigma_2,$$

where $\alpha_1 = \bar{\tau}_1/\sigma_2$ is the proportionality of the exact 1-spike-burst period $\bar{\tau}_1$ measured against the projected 1-spike period σ_2. These two relations together give the 2nd-order approximation of the average burst period and the corresponding 2nd-order approximation, $R_n^{(2)}$, of R_n. Similarly, $R_n^{(2)}$ is parameterized by $J_2, \mu, \beta, \sigma_2$ as well as α_1. This approximation process can go on to higher orders as well. For example, the 3rd-order approximation for the burst periods are $\bar{\tau}_1 = \alpha_1\sigma_3, \bar{\tau}_2 = 2\alpha_2\sigma_3, \bar{\tau}_k = k\sigma_3$ for all $k \geq 3$. That is, the first three burst periods are exact for some corresponding proportionalities α_1, α_2 and all the higher ones are approximated. We note that $R_n^{(n)} = R_n$ for a choice of $\{\alpha_k : 1 \leq k \leq n-1\}$. Similarly, the mth-order approximation of the averaged minimal burst period $\tau_{\min}$ is defined analogously and so is the mth-order approximation of C_n by $C_n^{(m)}$.

Fig.2(a) show the 1st-order and the 2nd-order approximations to both the mean rate R_n and the capacity rate C_n for the circuit model of Fig.1(b). It shows that the 2nd-order approximation is quantitatively good. Because of this conclusion, Fig.3 shows various comparisons using only the dimensionless 2nd-order approximations, $\sigma_2 R_n^{(2)}$. Taken all together, it can be concluded that having 10 memory bases for our neuronal circuit models tends to maximize the memory entropy against its retrieval time.

Following the same arguments from [5], factoring the refractory period (Fig.1) into the retrieval rates will not quantitatively change the optimal solution around $n = 10$ for the maximal retrieval rates in both the mean and the absolute maximum.

6. Discussions. The derivation of this conclusion above also suggests a possible evolutionary scenario as to how the optimal mode was reached. One can imagine that the neurons evolved to first acquire the 1-spike burst interval (J_2, J_1), then the next interval (J_3, J_2), and so on. With each new acquisition of base intervals, the memory retrieval rate improved a little (Fig.3(d)) until the maximal rate was reached. Nature may have tried but abandoned the next base interval because the new acquisition did not do better. The conclusion of this paper is that the number n should be around 10 for the effective region (J_{n+1}, J_1) which we have used throughout the paper.

This paper follows the same entropical optimization idea ([2, 4]) to analyze informational trade-off against time. The same methodology led to the entropical explanations to the DNA code in 4 bases ([1]), the origin of 2 sexes ([2]), the best mean transmission rate in 4 bases for neural information ([5]), the fastest transmission rate in the Golden Ratio distribution of binary neural information ([5]), and here the neural computing/storage preference in 10 bases. The entropical optimization theory hypothesizes that biological systems prefer the most informational diversity expressed in the shortest temporal constraints as possible. Undoubtedly, it will take some carefully-designed experiments to test if any of the entropical explanations is wrong.

One can certainly make a strong case that the emerging of the decimal number system was the most critical event in the history of human's technological and scientific advancement. It eventually displaced all other systems throughout the world. A leading folklore theory attributes its adoption to the number of fingers of our hands. Such a theory, however, cannot rule out other alternative theories on the same logical ground. After all, there are 20 digits from our four limbs, 8 fingers and 2 thumbs from two hands, 4 fingers or 5 digits from each hand, but none of the

corresponding number systems is used today to satisfy our daily computational needs. If our neuron models is a good approximation of the real kind, the result of this paper may have given a more plausible explanation to the decimal system's origin and prevalence than the hand theory. That is, of all number systems, the decimal system would just make it feel more visceral for our ancestors because it was conformed to the preferred computing mode of their neurons as well ours. The hand theory seems reasonable only because it correctly reflects the deeper explanation in the neuron theory. Unexpectedly, the neuron theory also seems to suggest that our two hands might have co-evolved with our neuron's decimal memory system to fill in an evolutionary niche. In this last regard, suanpan could be considered perhaps the first artificial and morphological interface to our brain. Its number system seems to match the best mode for the most data transmission in 4 bases ([5]), and the best mode for the most memory retrieval in 10 bases demonstrated here.

References

[1] Deng, B., *Why is the number of DNA bases 4?,* Bulletin of Math. Biol., **68**(2006), pp.727–733.

[2] Deng, B., *The origin of 2 sexes through optimization of recombination entropy against time and energy,* Bulletin of Math. Biol., **69**(2007), pp.2105–2114.

[3] Deng, B., *General circuit models of neurons with ion pump,* preprint, (2007).

(http://www.math.unl.edu/∼bdeng1/Papers/DengNeuroCircuit.pdf)

[4] Deng, B., *Neural metastability and plasticity mediated by two ion pumps,* preprint, (2007).

(http://www.math.unl.edu/∼bdeng1/Papers/DengNeuralMetaStability.pdf)

[5] Deng, B., *Quadrary spike code for optimal neural communication and informational preference to Golden Ratio distribution,* preprint, (2007).

(http://www.math.unl.edu/∼bdeng1/Papers/DengNeuralCodePreferences.pdf)

[6] Shannon, C.E., *A methematical theory of communication,* Bell System Technical Journal, **27**(1948), pp.379–423, and pp.623–656.

Supplements. The figure below shows the simulations for two more different circuit models whose complete equations can be found in [3, 4]. Fig.4(a,b) show that the absolute minimal uni-spike period for all k-spike bursts can be the same for all $k \geq 2$ as discussed in the main text. Fig.4(c,d) show typical plots for spike-burst frequencies. In such cases, the maximal spike frequency for each class of k-spike bursts increases in k, and approaches a limit. The circuit choice for SEED system has such parameter values so that the maximal uni-spike frequency sequence stays constant for all $k \geq 2$ as shown in (a,b) and Fig.1. Such optimal parameter values are fairly easy to find.

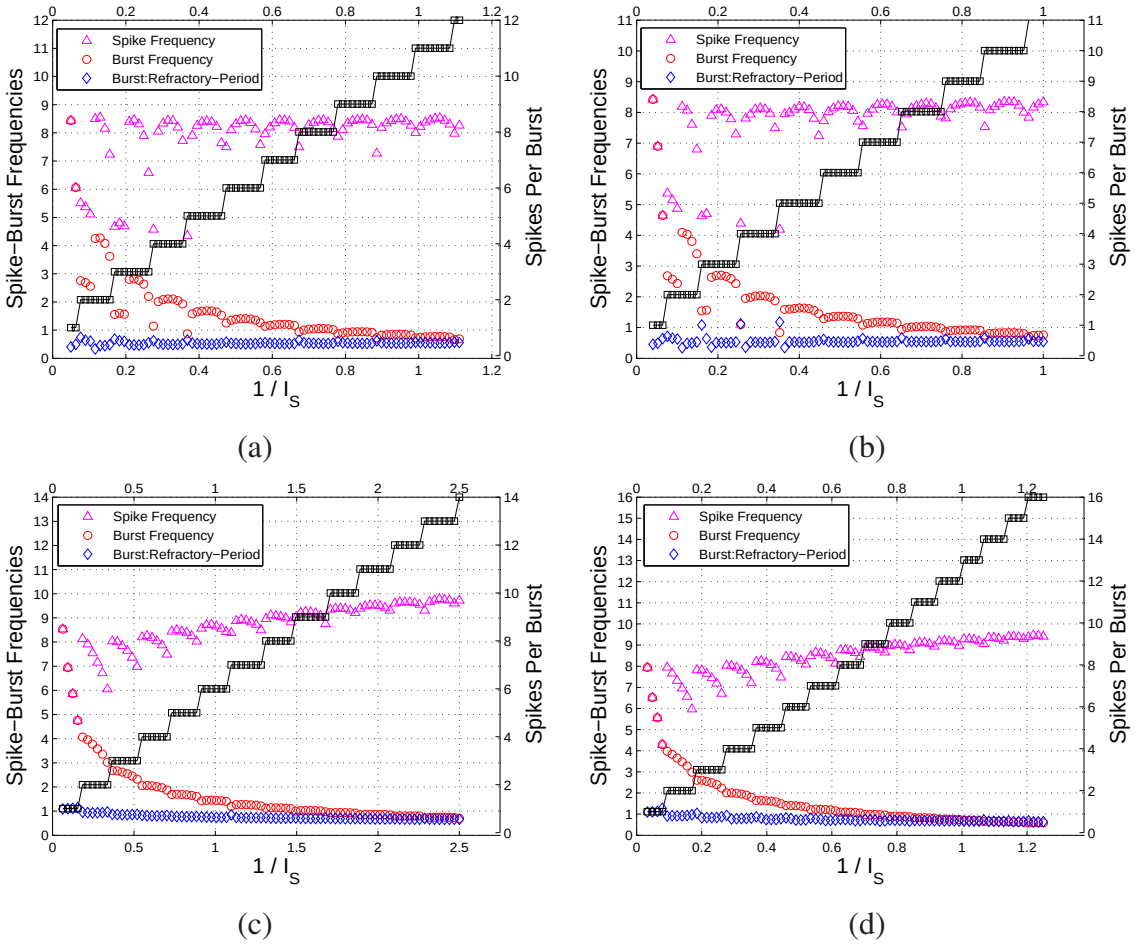


Figure 4: (a) $\text{pK}^+_-\text{sNa}^+_+$ model with the arctangent IV-curves. The same parameter values as Fig.1 except for $g_{\text{K}} = 5, d_{\text{K}} = -5.5, v_1 = 0.45, v_2 = 1.55, i_1 = .1, i_2 = 0.33, \lambda_{\text{Na}} = \lambda_{\text{K}} = 0.1$. (b) $\text{pK}^+_{-d}\text{sNa}^+_{+d}$ model with the arctangent IV-curves. The same parameter values as (a). (c) $\text{pK}^+_-\text{sNa}^+_+$ model with the cubic IV-curves. The same parameter values as Fig.1 except for $a = 1, v_1 = 0.5, v_2 = 2.5, b = 30, i_1 = 0.25, i_2 = 0.8, \lambda_{\text{Na}} = \lambda_{\text{K}} = 0.5$. (d) $\text{pK}^+_{-d}\text{sNa}^+_{+d}$ model with the cubic IV-curves. The same parameter values as (c) except for $\lambda_{\text{Na}} = 0.5, \lambda_{\text{K}} = 0.1, \gamma_{\text{Na}} = 0.1, \gamma_{\text{K}} = 0.5$.