LETTER Communicated by Carl van Vreeswijk

Collective Behavior of Networks with Linear (VLSI) Integrate-and-Fire Neurons

Stefano Fusi
Maurizio Mattia
INFN, Sezione di Roma, Dipartimento di Fisica, Università di Roma "La Sapienza," Rome, Italy

We analyze in detail the statistical properties of the spike emission process of a canonical integrate-and-fire neuron, with a linear integrator and a lower bound for the depolarization, as often used in VLSI implementations (Mead, 1989). The spike statistics of such neurons appear to be qualitatively similar to conventional (exponential) integrate-and-fire neurons, which exhibit a wide variety of characteristics observed in cortical recordings. We also show that, contrary to current opinion, the dynamics of a network composed of such neurons has two stable fixed points, even in the purely excitatory network, corresponding to two different states of reverberating activity. The analytical results are compared with numerical simulations and are found to be in good agreement.

1 Introduction

The integrate-and-fire (IF) neuron has become popular as a simplified neural element in modeling the dynamics of large-scale networks of spiking neurons. A simple version of an IF neuron integrates the input current as an RC circuit (with a leakage current proportional to the depolarization) and emits a spike when the depolarization crosses a threshold. We will refer to it as the *RC neuron*. Networks of neurons schematized in this way exhibit a wide variety of characteristics observed in single and multiple neuron recordings in cortex in vivo. With biologically plausible time constants and synaptic efficacies, they can maintain spontaneous activity, and when the network is subjected to Hebbian learning (subsets of cells are repeatedly activated by the external stimuli), it shows many stable states of activation, each corresponding to a different attractor of the network dynamics, in coexistence with spontaneous activity (Amit & Brunel, 1997a). These stable activity distributions are selective to the stimuli that had been learned. When the network is presented a familiar stimulus (similar to one that was previously learned), the network is attracted toward the learned activity distribution most similar to the stimulus. At the end of this relaxation process (in the attractor) a subset of neurons cooperates to maintain elevated firing rates. This selective activity is sustained throughout long delay intervals,

as observed in cortical recordings in monkeys performing delay-response tasks (Miyashita & Chang, 1988; Wilson, Scalaidhe, & Goldman-Rakic, 1993; Amit, Fusi, & Yakovlev, 1997). Moreover, extensive simulations revealed that for these networks, spike time statistics and cross-correlations are quite like in cortical recordings in vivo (Amit & Brunel, 1997b). If such collective behavior, including dynamic learning is a relevant computational module, an electronic implementation would be called for.

Except for simple testing of small-scale pilot systems, electronic implementation implies VLSI technology. In VLSI the building block is the transistor, and the corresponding guiding principles are the economy in the number of transistors, closely connected to the area of the chip, and the reduction of power consumption. In an analog VLSI (aVLSI) implementation, the natural minimalist version of an IF neuron, as canonized by Mead (1989), operates in current mode and therefore integrates linearly the input current. This aVLSI implementation of the neuron has many desirable features. It operates with current generators and hence very low power consumption, an essential feature for integrating a large number of neurons on a single chip. It is also a natural candidate for working with transistors in the weak-inversion regime, which brings another significant reduction in consumption (Mead, 1989). One can also implement an aVLSI dynamic synapse with similar attractive electronic characteristics (Annunziato, 1995; Diorio, Hasler, Minch, & Mead, 1996; Elias, Northmore, & Westerman, 1997; Annunziato, Badoni, Fusi, & Salamon, 1998).

Here we will concentrate on the statistical properties of the spikes generated by an aVLSI neuron, as a function of the statistics of the input current, and on the dynamics of networks composed of aVLSI neurons, keeping the distributions of synaptic efficacies fixed. We ask the following question: Given that the depolarization dynamics of the aVLSI neuron is significantly different from that of the RC neuron, can the collective dynamics found in a network of RC neurons be reproduced in networks of neurons of the aVLSI type?

2 RC Neuron Versus aVLSI Neuron

The RC neuron below threshold is an RC circuit integrating the input current with a decay proportional to the depolarization of the neuron's membrane $V(t)$:

$$\frac{dV(t)}{dt} = -\frac{V(t)}{\tau} + I(t), \tag{2.1}$$

where $I(t)$ is the net charging current, expressed in units of potential per unit time, produced by afferent spikes, and τ $(= RC)$ is the integration time constant of the membrane depolarization. When $V(t)$ reaches the threshold θ, the neuron emits a spike, and its potential is reset to H, following an absolute refractory period τ_{arp}.

On the other hand, the aVLSI neuron below threshold can be schematically considered as a linear integrator of the input current,

$$\frac{dV(t)}{dt} = -\beta + I(t), \tag{2.2}$$

with the constraint that if $V(t)$ is driven below the resting potential $V = 0$, it remains 0, as in the presence of a reflecting barrier. β is a constant decay ($\beta > 0$) that, in the absence of afferent currents, drives the depolarization to the resting potential. The spiking condition remains unmodified, and the reset potential H must be positive or zero. In the rest of the article, it will be set to 0. As for the RC neuron, the absolute refractory period sets the maximum emission frequency ($\nu_{\max} = 1/\tau_{arp}$).

2.1 Afferent Current. We assume that at any time t, the source of depolarization $I(t)$ (afferent current) is drawn randomly from a gaussian distribution with mean $\mu_I(t)$ and variance $\sigma_I^2(t)$ per unit time, so from equation 2.2, the depolarization is a stochastic process obeying:

$$dV = \mu(t)dt + \sigma(t)z(t)\sqrt{dt}, \tag{2.3}$$

where $\mu(t) = -\beta + \mu_I(t)$ is the total mean drift at time t, $\sigma(t) = \sigma_I(t)$ is the standard deviation, and $z(t)$ is a random gaussian process with zero-mean and unit variance.

For instance, if a neuron receives Poissonian trains of spikes from a large number of independent input channels, the dynamics is well approximated by equation (2.3) (Tuckwell, 1988; Amit & Tsodyks, 1991; Amit & Brunel, 1997a).

Here we assume that the input current $I(t)$ is uncorrelated in time. This is a good approximation for VLSI applications, but for biological neurons, we should take into account the time correlations introduced by the synaptic dynamics. We did not investigate the effect of these correlations on the behavior of a network of aVLSI neurons. It can be studied as for RC neurons (Brunel & Sergi, 1999).

2.2 SD and ND Regime: Some Key Features. If the reflecting barrier is absent or the distance H between the reset potential and the barrier is much greater than $\theta - H$ (Gerstein & Mandelbrot, 1964), then linear integrator dynamics can operate only in a positive drift regime ($\mu > 0$); otherwise the probability density function (p.d.f.) of the first passage time—the first time the depolarization $V(t)$ crosses θ starting from the reset potential—has no finite moments, that is, the mean emission rate of the neuron vanishes (Cox & Miller, 1965). For such a neuron, the current-to-rate transduction function depends on only the mean drift and is linear for a wide range of positive drifts. If the absolute refractory period is not zero, then the transduction

function is convex, showing some nonlinearity when the neuron works in a saturation regime (near the maximum frequency $1/\tau_{arp}$). Otherwise, the transduction function is a threshold-linear function. This regime is signal dominated (SD). The question about the collective behavior of SD neurons is underlined by the following consideration: for a threshold-linear transduction function, the coexistence of spontaneous activity with structured delay activity is not possible. Each of the two types of behavior is implementable in a network with "linear" neurons (van Vreeswijk & Hasselmo, 1995) but not both.

If, on the contrary, there is a reflecting barrier not too far from θ, then the statistics of the input current can be such that the neuron can also operate in a different regime. If spikes are emitted mostly because of large, positive fluctuations, the neuron is working in the noise-dominated (ND) regime. This happens when the mean drift is small or negative and the variability is large enough (see also Bulsara, Elston, Doering, Lowen, & Lindenberg, 1996).

When the neuron can operate in both SD and ND regimes—and we will show that this is the case for the aVLSI neuron—then the current-to-rate transduction function is nonlinear (convex for large drifts, concave for small and negative drifts) and mean-field theory exhibits the coexistence of two collective stable states. In particular, the nonlinearity due to the ND regime is a necessary element for obtaining spontaneous and selective activity in more complex networks of excitatory and inhibitory neurons (Amit & Brunel, 1997a).

3 Statistical Properties of aVLSI Neurons

3.1 Current-to-Rate Transduction Function and Depolarization Distribution. In order to obtain the current-to-rate transduction function (the mean emission rate as a function of μ and σ in stationary conditions), we define $p(v, t)$ as the probability density that at time t the neuron has a depolarization v. For the diffusion process of equation 2.3, $p(v, t)$ obeys the Fokker-Planck equation (see, e.g., Cox & Miller, 1965):

$$\frac{1}{2}\sigma^2(t)\frac{\partial^2 p}{\partial v^2} - \mu(t)\frac{\partial p}{\partial v} = \frac{\partial p}{\partial t}. \tag{3.1}$$

This equation must be complemented by boundary conditions restricting the process to the interval $[0, \theta]$:

- At $v = 0$: A reflecting barrier, since no process can pass below 0.
- At $v = \theta$: An absorbing barrier. All processes crossing the threshold are absorbed and reset to H.

Formally, this is equivalent to the conditions that $p(v, t) = 0$ at $v = \theta$ (see, e.g., Cox & Miller, 1965), and that no process is lost when absorbed at θ or

reflected at $v = 0$ (for simplicity, we start by assuming that $\tau_{arp} = 0$), that is,

$$\int_0^\theta p(v, t)dv = 1. \tag{3.2}$$

This implies that the rate at which processes are crossing the threshold is the same as the rate at which they reenter at 0. Moreover, no process can cross the reflecting barrier from above, so the net flux of processes going through the reflecting barrier is due only to processes coming from the threshold. Integrating over v on both sides of equation (3.1), imposing the boundary condition at $v = \theta$ and the normalization condition in equation 3.2, one gets:

$$\frac{1}{2}\sigma^2 \frac{\partial p}{\partial v}\bigg|_{v=\theta} = \left[\frac{1}{2}\sigma^2 \frac{\partial p}{\partial v} - \mu p\right]_{v=0}. \tag{3.3}$$

The probability per unit time of crossing the threshold and reentering from the reflecting barrier is given by the flux of processes at $v = \theta$:

$$\nu(t) = -\frac{1}{2}\sigma^2 \frac{\partial p}{\partial v}\bigg|_{v=\theta}. \tag{3.4}$$

If one considers a large ensemble of neurons, the diffusion equation has a natural interpretation in terms of replicas of identical neurons. Each neuron can be considered as representative for one particular realization of the stochastic process (i.e., a single instance of the process corresponding to a particular choice of $I(t)$). So $p(v, t)$ can be seen as the depolarization distribution across all the neurons of the network at time t. $\nu(t)$ is the mean fraction of neurons crossing the threshold per unit time at time t, and if the network state is stationary, it is also the mean emission frequency for any generic neuron.

If $\tau_{arp} > 0$, the realizations of a stochatic process in which the neuron crosses the threshold θ must be delayed before coming back to the reset value. The flux of the processes reentering from $v = 0$ at time t must be equal to the flux of processes that were crossing the threshold at time $t - \tau_{arp}$:

$$\frac{1}{2}\sigma^2 \frac{\partial p}{\partial v}\bigg|_{v=\theta} = -\nu(t) \text{ and } \left[\frac{1}{2}\sigma^2 \frac{\partial p}{\partial v} - \mu p\right]_{v=0} = -\nu(t - \tau_{arp}). \tag{3.5}$$

Since some processes are spending time inside the absolute refractory period, a new normalization condition must be imposed,

$$\int_0^\theta p(v, t)dv + \int_{t-\tau_{arp}}^{t} \nu(t')dt' = 1, \tag{3.6}$$

where the integrand is composed of two parts: the probability density of being in the linear integration interval $[0, \theta]$ and the probability of being in the absolute refractory period.

For steady statistics of the input current and in a stationary regime ($\partial p/\partial t = 0$), we have that $\nu(t)$ is constant ($= \nu$), and the density function is given by solving equation 3.1 with the boundary conditions of equations 3.5,

$$p(v) = \frac{\nu}{\mu}\left[1 - \exp\left(-2\frac{\mu}{\sigma^2}(\theta - v)\right)\right] \qquad \text{for } v \in [0, \theta], \tag{3.7}$$

where ν is determined by imposing equation (3.6),

$$\nu \equiv \Phi(\mu, \sigma) = \left[\tau_{arp} + \frac{\sigma^2}{2\mu^2}\left(\frac{2\mu\theta}{\sigma^2} - 1 + e^{\frac{-2\mu\theta}{\sigma^2}}\right)\right]^{-1}, \tag{3.8}$$

which gives the mean emission rate of the aVLSI neuron as a function of mean and variance of the input current.

3.2 Interspike Interval Distribution and Variability. The probability density of the interspike intervals (ISI) in stationary conditions with $\tau_{arp} = 0$, is computed following Cox and Miller (1965). The first passage time T is a random variable with a p.d.f. $g(H, T)$ that depends on the initial value of the depolarization, that is, the reset potential H. If H is considered as a variable, not as in the previous section where it was kept fixed ($H = 0$), it can be shown that $g(H, t)$ satisfies a backward Kolmogorov diffusion equation,

$$\frac{1}{2}\sigma^2\frac{\partial^2 g}{\partial H^2} + \mu\frac{\partial g}{\partial H} = \frac{\partial g}{\partial t}, \tag{3.9}$$

that can be solved by using the Laplace transform $\gamma(H, s)$ of $g(H, T)$:

$$\gamma(H, s) \equiv \int_0^\infty e^{-st'} g(H, t')dt'.$$

The nth order derivative of $\gamma(H, s)$ with respect to $-s$ calculated at $s = 0$ is the nth moment of the first passage time T.

The equation for the Laplace transform is:

$$\frac{1}{2}\sigma^2\frac{\partial^2 \gamma}{\partial H^2} + \mu\frac{\partial \gamma}{\partial H} = s\gamma.$$

The boundary conditions restricting the process between the reflecting barrier at 0 and the absorbing barrier in θ are translated in terms of the Laplace transform in the following equations:

$$\gamma(\theta, s) = 1; \qquad \left.\frac{\partial \gamma(H, s)}{\partial H}\right|_{H=0} = 0.$$

The solution is:

$$\gamma(0, s) = \frac{ze^{\theta C}}{z\cosh(\theta z) + C\sinh(\theta z)} ,$$

where $C \equiv \mu/\sigma^2$, $z \equiv \sqrt{\mu^2 + 2s\sigma^2}/\sigma^2$, and $H = 0$.

$g(0, T)$ will be evaluated numerically by antitransforming $\gamma(0, s)$. If $\tau_{arp} \neq 0$, the mean emission frequency calculated from the expected value μ_T of the ISI is given by

$$\nu = \frac{1}{\tau_{arp} + \mu_T} . \tag{3.10}$$

This expression reproduces equation 3.8.

The coefficient of variability (CV), defined as the ratio between the square root of the variance σ_T and the mean μ_T of the ISI, can be calculated as

$$CV(\mu, \sigma) \equiv \frac{\sigma_T}{\mu_T} = \frac{\sqrt{e^{-2m} + 4e^{-m}(m+1) + 2m - 5}}{e^{-m} + (1 + \mu\tau_{arp}/\theta)m - 1} , \tag{3.11}$$

where $m = 2\mu\theta/\sigma^2$ and:

$$\mu_T = \tau_{arp} - \left.\frac{\partial\gamma}{\partial s}\right|_{s=0}$$

$$\sigma_T^2 = \left.\left(\frac{\partial^2\gamma}{\partial s^2} - \left(\frac{\partial\gamma}{\partial s}\right)^2\right)\right|_{s=0} .$$

4 SD Versus ND Regime: Results

If $\mu\theta/\sigma^2 \gg 1$, the depolarization dynamics is dominated by the deterministic part of the current and the neuron is operating in the SD regime. In Figure 1 (top) we show an example of a simulated neuron operating in this regime: the depolarization grows, fluctuating around the linear ramp determined by the constant drift, until it emits a spike. Since positive and negative fluctuations tend to cancel, the neuron fires quite regularly, and the average ISI is θ/μ.

In contrast, in the ND regime (see Figure 1B), the neuron spends most of the time fluctuating near the reflecting barrier and emits a spike only when a large fluctuation in the input current drives the depolarization above the threshold. Since the fluctuations are random and uncorrelated, the neuron fires irregularly and the ISI distribution is wide (see below). In this regime the process is essentially dominated by the variance of the afferent current.

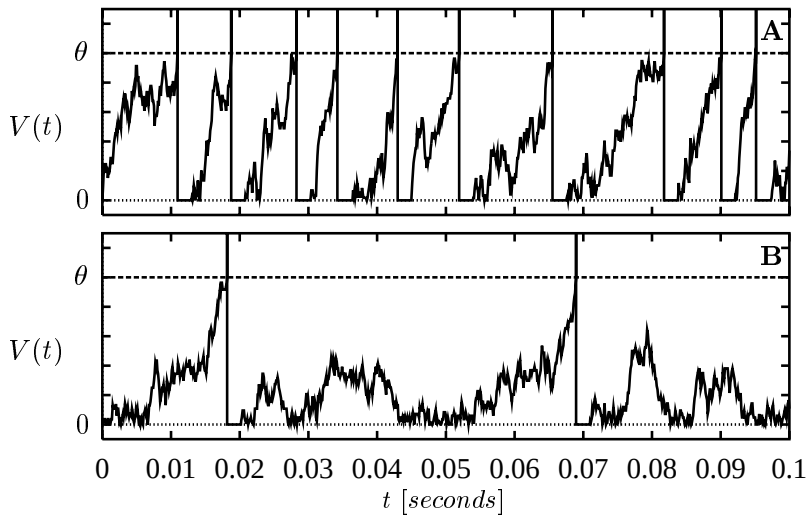

Figure 1: Realizations of stochastic processes representing depolarization dynamics simulated in (A) SD and (B) ND regimes. Time is expressed in seconds. Parameters: (A) $\mu = 102\theta$ Hz, $\sigma^2 = 28.1\theta^2$ Hz, producing a mean firing rate $\nu = 94$ Hz; (B) $\mu = -10.1\theta$ Hz, $\sigma^2 = 14.4\theta^2$ Hz, mean rate $\nu = 8.1$ Hz. $\tau_{arp} = 2$ ms in both cases. In the SD regime, the process is dominated by the deterministic part of the input current. The noisy linear ramp is clearly visible. In the ND regime, the depolarization fluctuates under threshold, waiting for the large, positive fluctuation of the input current to drive $V(t)$ above threshold.

4.1 Depolarization Distribution. In Figure 2 we plot the probability density $p(v)$, given by equation 3.7, for different regimes. In the signal-dominated regime, $p(v)$ is almost uniform because the neuron tends to go from 0 to θ at constant speed. As one moves toward the ND regime, the probability density changes concavity and tends to concentrate at the reflecting barrier, $v = 0$.

4.2 ISI Distribution and Coefficient of Variability. In the SD regime, the ISI depends essentially on the mean drift μ. As one moves toward higher frequencies (i.e., large drift), the neuron tends to fire more regularly, and the ISI distribution tends to be peaked around $T = \theta/\mu$ (see Figure 3). As σ increases and μ decreases, moving toward the ND regime, the curve spreads and the distribution extends to a wide range of ISI. The qualitative behavior of the ISI distribution is quite similar to the one described for the RC neurons in Amit & Brunel (1997b), which, in turn, resembles the ISI distribution of cortical recordings (Tuckwell, 1988; Usher, Stemmler, Koch, & Olami, 1994).

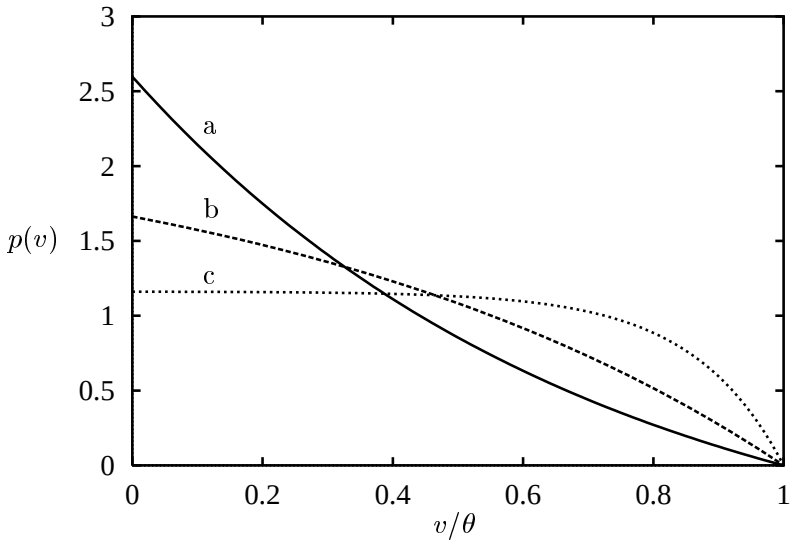

Figure 2: Probability density function $p(v)$ in three regimes: (a) ND, (b) intermediate and (c) SD. Parameters: (a) $\mu = -10.1\theta$ Hz, $\sigma^2 = 14.4\theta^2$ Hz; (b) $\mu = 10.0\theta$ Hz, $\sigma^2 = 16.0\theta^2$ Hz; (c) $\mu = 102\theta$ Hz, $\sigma^2 = 28.1\theta^2$ Hz. In the ND regime $p(v)$ is concentrated well below the threshold, near the reset potential ($v = 0$). As μ increases, the curve changes concavity and tends to a uniform distribution, which is the density function for a deterministic process ($\sigma = 0$).

The behavior of the coefficient of variability CV (see equation 3.11) shows more clearly the relation between the spread of the ISI distribution and the average of the first passage time. In Figure 4 we plot CV versus μ and σ. When $\sigma = 0$ (left side), the depolarization walk is deterministic and the variability is 0. For negative drifts, when both the mean and the standard deviation of the ISI are zero, we assumed conventionally CV=0. For $\sigma > 0$ and $\sigma^2 \ll 2|\mu|\theta$, two regions can be distinguished. At negative drift ($\mu < 0$), CV is almost 1 and the spike emission process is Poissonian. At positive drifts, the deterministic part dominates, and CV is small. As σ increases, the mean frequency saturates to $1/\tau_{arp}$, and the coefficient of variability tends to 0. This is due to the fact that the large fluctations in the afferent current drive the depolarization above the threshold immediately after the absolute refractory period, even in the absence of drift.

4.3 Current-to-Rate Transduction Function: Sources of Nonlinearity. In Figure 5 we plot the current-to-rate transduction function given by equation 3.8 as a function of μ for three different values of σ. For any value of

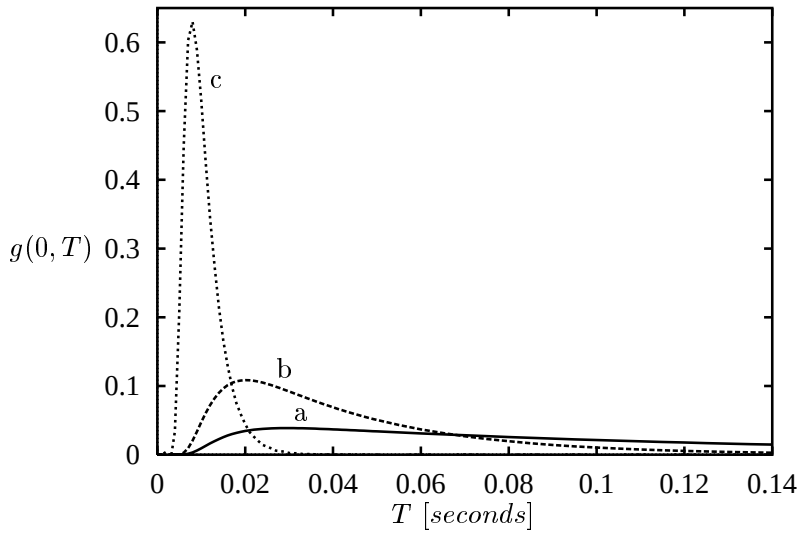

Figure 3: ISI distribution $g(0, T)$ at (a) negative, (b) intermediate, and (c) positive drift. Parameters: (a) $\mu = -10.1\theta$ Hz, $\sigma^2 = 14.4\theta^2$ Hz; (b) $\mu = 10.0\theta$ Hz, $\sigma^2 = 16.0\theta^2$ Hz; (c) $\mu = 102\theta$ Hz, $\sigma^2 = 28.1\theta^2$ Hz. The variability coefficients CV and mean first passage times $E[T]$ in each regime are: (a) $CV = 0.79$, $E[T] = 0.12$ s; (b) $CV = 0.56$, $E[T] = 0.045$ s, and (c) $CV = 0.24$, $E[T] = 0.010$ s. The ISI distribution is widespread for negative drift and tends to a peaked distribution as μ goes to positive values.

σ, the absolute refractory period introduces a first source of nonlinearity in the region of high frequencies since Φ saturates at $\nu = 1/\tau_{arp}$.

If $\sigma^2 \ll |\mu|\theta$, the random walk is dominated by the drift, which is the deterministic part of the current (SD regime), and we have

$$\Phi \simeq \begin{cases} \dfrac{\mu}{\theta + \tau_{arp}\mu} & \text{if } \mu > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{4.1}$$

For a wide range of drifts, this function is well approximated by a threshold-linear function Φ_{tl} defined as

$$\Phi_{tl}(\mu) = \begin{cases} \dfrac{\mu}{\theta} & \text{if } \mu > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{4.2}$$

In the SD regime, the nonlinearity due to the absolute refractory period makes the curve convex for any $\mu > 0$, as in Gerstein and Mandelbrot (1964).

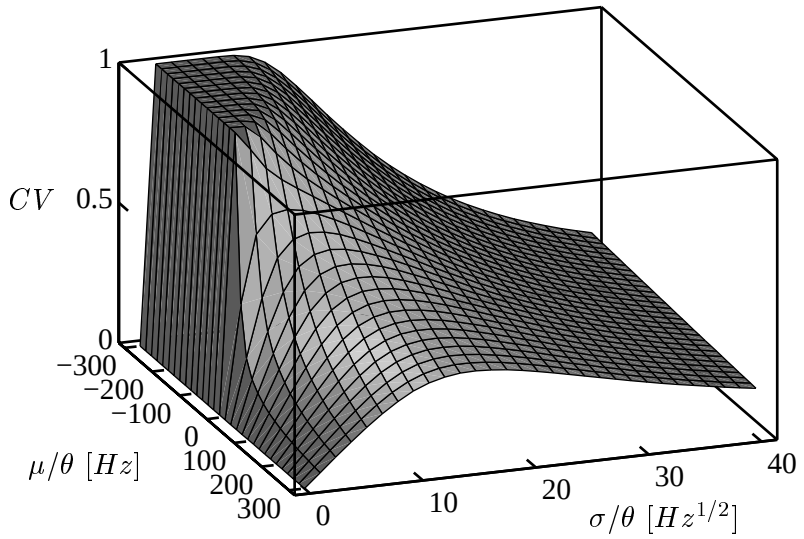

Figure 4: Coefficient of variability CV in the space (μ, σ). The variability is much higher in the negative drift regime. When $\mu < 0$ and $\sigma^2 \ll 2|\mu|\theta$, then $CV \to 1$, and the spike emission process becomes Poissonian.

As σ increases, Φ departs from the threshold-linear behavior, and nonzero frequencies are produced also for negative drifts. The curve is convex for large, positive drift and concave for small, negative drifts. The variance in the input current produces the second source of nonlinearity.

To expose the differences between the response function Φ and the threshold-linear function Φ_{tl}, we plot the surface $\Phi(\mu, \sigma) - \Phi_{tl}(\mu)$ in Figure 6. At $\sigma = 0$, the only source of nonlinearity is due to τ_{arp}: a dark shadow, corresponding to negative differences, appears in the region of high frequencies. As σ increases, a region of nonzero frequencies pops up around $\mu = 0$. The region in which Φ differs from threshold linear grows as one moves toward large variances, eventually covering a wide range of negative and positive values of μ (bright region). This is the second source of nonlinearity and corresponds to the ND regime.

5 Network Dynamics: Double Fixed Point

The extended mean-field theory (Amit & Brunel, 1997a) allows us to study the dynamics of any population of neurons randomly interconnected, provided that one knows the current-to-rate transduction function. In the most general case, the afferent current to any neuron is composed of two parts: one from spikes emitted by other neurons in the same population and the

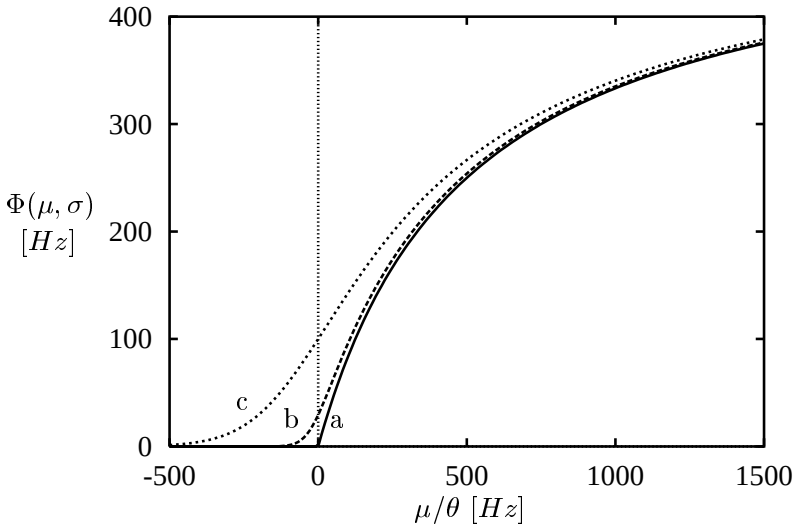

Figure 5: Current-to-rate transduction function $\Phi(\mu,\sigma)$ for different variances of afferent current: (a) $\sigma^2 = 0$, (b) $\sigma^2 = 31.4\theta^2$ Hz, and (c) $\sigma^2 = 121\theta^2$ Hz. The firing rate in the region around $\mu = 0$ is rather sensitive to changes in the variance. $\phi(\mu,\sigma)$ passes from a threshold-linear function at $\sigma = 0$ to a nonlinear function when $\sigma > 0$. If $\mu\theta \gg \sigma^2$, the transduction function is almost independent of σ. The nonlinearity that appears for large μ is due to τ_{arp}: Φ tends to the asymptotic frequency $1/\tau_{arp}$.

other from outside. If (1) the mean number of afferent connections is large, (2) the mean charging current produced by the arrival of a single spike (the mean synaptic efficacy) is small relative to threshold, and (3) the emission times of different neurons can be assumed uncorrelated (these conditions are satisfied in many known cases; see, e.g., Amit & Brunel, 1997b, and van Vreeswijk & Sompolinsky, 1996), then the current $I(t)$ is gaussian and μ and σ^2 are linear functions of the instantaneous probability of emission $\nu(t)$:

$$\mu(t) = a_\mu \nu(t) + b_\mu(t)$$

$$\sigma^2(t) = a_\sigma \nu(t) + b_\sigma(t).$$

The part depending on $\nu(t)$ is due to the recurrent connections inside the population, while the offset is generated by the spikes coming from outside and by the constant decay β. The a's and b's are variables depending on the statistics of the connectivity, the synaptic efficacy, the decay β, and the external afferents.

The conditions enumerated are approximately satisfied for RC neurons in many cases, and the mean-field description is a good approximation.

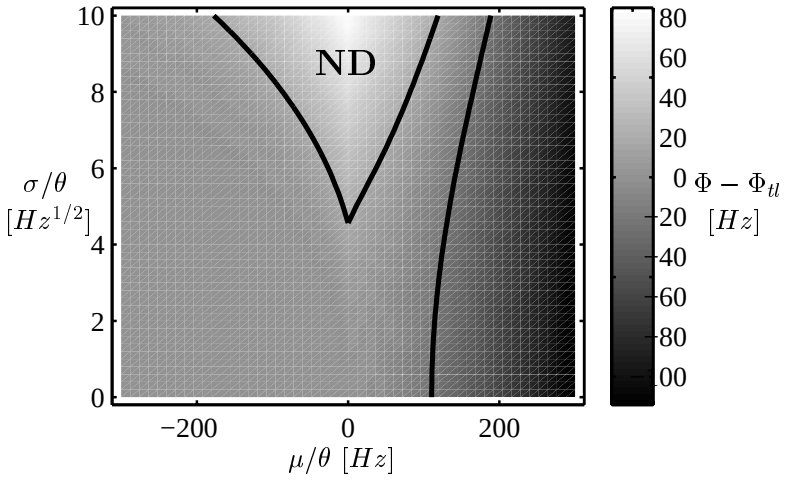

Figure 6: Difference between the aVLSI neuron current-to-rate transduction function $\Phi(\mu, \sigma)$ and the threshold-linear function $\Phi_{tl}(\mu)$ in the space (μ, σ). Dark and bright regions correspond, respectively, to negative and positive differences (see the scale on the right). Contour lines are at differences of $+20$ Hz and -20 Hz. The two sources of nonlinearity are due to the absolute refractory period (for large μ) and the ND regime (the brightest region).

In order to have a fixed point of the population dynamics, the rate that determines the statistics of the afferent current must be equal to the mean emission rate. In formal terms, the following self-consistency mean-field equation must be satisfied:

$$\nu = \Phi(\mu(\nu), \sigma(\nu)). \tag{5.1}$$

If the function Φ is linear in ν, as in the case of a neuron operating in the SD regime, then only one stable fixed point at nonzero emission rate is possible (van Vreeswijk & Hasselmo, 1995). Having two stable fixed points in a single population of excitatory neurons requires a change in the convexity. In the case of Φ of equation 3.8, the two nonlinearities described in the previous section are sufficient to allow for three fixed points (see Figure 7). Two of them, corresponding to the lowest and the highest frequencies, are stable, and the one in the middle is unstable and is the border of the two basins of attraction. In the stable state of low frequency, the neurons are working in the ND regime ($\mu\theta/\sigma^2 = -0.17$), while in the state of high frequency, the signal is dominating ($\mu\theta/\sigma^2 = 7.75$) and the behavior of the network is almost unaffacted by σ.

The example in Figure 7 with a single population of excitatory neurons shows that the mathematical properties of the current-to-rate transduction

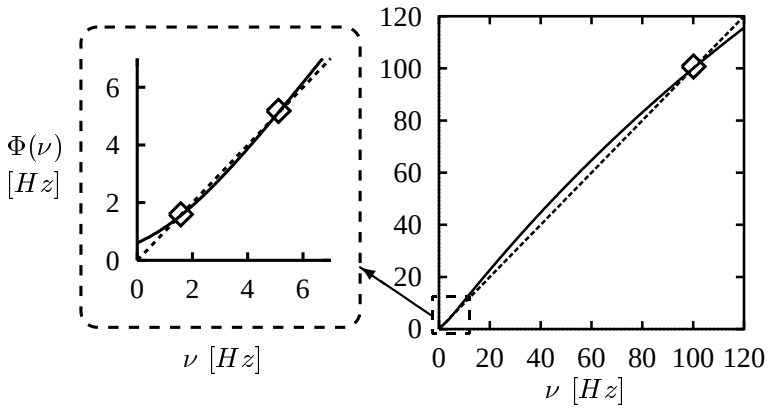

Figure 7: Fixed points ($\diamond$) of network dynamics: graphical solution of the self-consistency equation. Solid line = mean firing rate $\Phi(\nu)$; dashed line = ν. The rectangle on the left is an enlargement of the low-frequency region. Drift and variance: $\mu(\nu) = (-2.52 + 1.25\nu)\theta$/s and $\sigma(\nu)^2 = (1.88 + 0.021\nu)\theta^2$/s. There are three intersections between $\Phi(\nu)$ and ν: two correspond to stable fixed points ($\nu = 1.52$ Hz at negative drift and 99.1 Hz at positive drift) and one to an unstable fixed point (5.0 Hz).

function make it possible to have a double fixed point for the dynamics of the network.

In fact, the nonlinearity near zero is a sufficient condition for having in more complex networks the coexistence of spontaneous activity and many selective delay activity states. Without it, the low-rate fixed point corresponds to a state in which all the neurons are quiescent and the existence of a low-rate, highly variable spontaneous activity is not possible (see section 6 and van Vreeswijk & Hasselmo, 1995).

5.1 Simulations: A Toy Network. In order to check the assumptions of the extended mean-field theory, we present a simulation of a single population of randomly interconnected excitatory neurons. Simulations and mean-field analysis have been carried out for a complete system composed of a structured network of excitatory and inhibitory neurons (Amit & Mattia, 1998). Here we concentrate on the properties of a network of excitatory aVLSI neurons.

We study a network composed of $N = 1000$ excitatory neurons and choose the parameters in such a way that $\Phi(\nu)$ is the same as in Figure 7. Each neuron has a probability $c = 0.075$ (connectivity level) of a direct synaptic contact with any other neuron in the network. Such contacts are chosen randomly to have an amorphous recurrent synaptic structure. Hence each neuron receives on average cN recurrent synapses plus an external

stochastic afferent current. The simulation dynamics is summarized by the following two equations:

1. The depolarization dynamics of any generic neuron i, given by equation 2.2,

$$\frac{dV_i(t)}{dt} = -\beta + I_i(t).$$

In our case $\beta = 115.2\theta/\text{s}$ which means that in absence of afferent spikes the depolarization decays from θ to 0 in 8.7 ms.

2. The expression of the current $I_i(t)$,

$$I_i(t) = \sum_{j=1}^{N} J_{ij} \sum_k \delta(t - t_j^{(k)} - d) + I_{ext}(t),$$

which is composed of two parts. The first term is due to the spikes coming from the recurrent afferents, and the second represents the external current. J_{ij} is 0 if there is no direct contact from neuron j to neuron i. Otherwise it represents the amplitude of the postsynaptic potential (PSP) provoked in neuron i by a spike emitted by neuron j. In the present case these PSPs are equal for all the synaptic connections: $J = 0.0167\theta$, implying that 60 simultaneous afferent spikes would provoke a postsynaptic spike. The sum over k is extended to all the spikes emitted by the neurons: $t_j^{(k)}$ is the time at which neuron j has emitted the kth spike, which is received by neuron i after a delay $d = 2$ ms. The external current $I_{ext}(t)$ is gaussian white noise with $\mu_{ext} = 112.7\theta/\text{s}$ and $\sigma_{ext}^2 = 1.88\theta^2/\text{s}$.

For such a network the mean drift and the variance, introduced in the mean-field theory in the previous section, are:

$$\mu = cNJ\nu + \mu_{ext} - \beta = (1.25\nu - 2.52)\theta/s$$
$$\sigma^2 = cNJ^2\nu + \sigma_{ext}^2 = (0.02\nu + 1.88)\theta/s.$$

With this form of μ and σ^2, we carry out the mean-field analysis as described in Amit and Brunel (1997a), and we find the two stable stationary states of Figure 7. In the low-rate state ($\nu = 1.52$ Hz) the recurrent contribution $\mu_{rec} \equiv cNJ\nu$ to the mean drift is much smaller than the external contribution ($|\mu_{rec}| = 0.62\theta$ Hz $\ll \mu_{\text{ext}}$), whereas in the high-frequency stable state ($\nu =$ 99.1 Hz) they are almost of the same magnitude ($\mu_{rec} = 124.9\theta$ Hz $\simeq \mu_{ext}$).

The range of variability of J that allows for the three fixed points is limited ($J \in [0.015, 0.018]$) when all the other parameters are kept fixed. This is not an intrinsic limitation of the aVLSI neuron, but it is due to the fact that we are showing a toy example with a single excitatory population. In the presence of a dynamical inhibitory population, the range of variability of J is much wider (see, e.g., Amit & Brunel, 1997a and section 6).

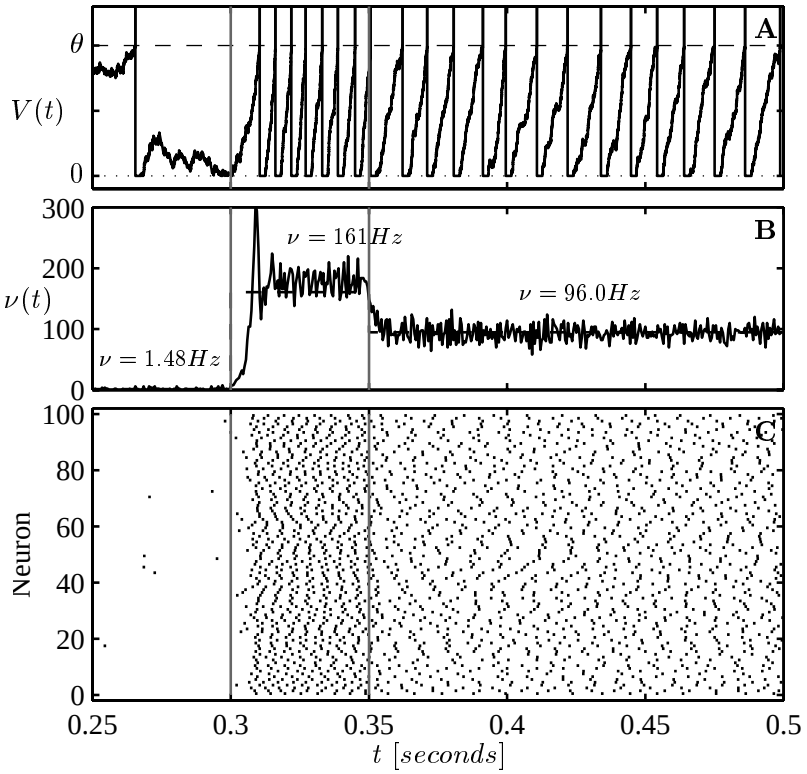

Figure 8: Simulation of the network dynamics described in Figure 7, to show the existence of two states of activation. (A) Depolarization as a function of time of a sample neuron. (B) $\nu(t)$ as a function of time. The dashed lines correspond to the mean rates in the three intervals (also the numbers are reported). (C) Raster of the spikes emitted by 100 different neurons. The simulation starts with the network that already relaxed into the low-rate fixed point. After 50 ms the network is stimulated by increasing the external current. After 50 ms the stimulation is removed and the network relaxes into the high rate stable state. See the text for discussion.

In Figure 8 we show the results of a simulation. We start with all the neurons in a quiescent state and $V = 0$. After a short time interval ($\sim$ 100 ms) the network relaxes into a low-rate stable state, and we start "recording" from the neurons; then we stimulate the network for 50 ms by increasing the mean and the variance of the external current I_{ext} by a factor 1.5. This stimulation drives the dynamics of the network to the basins of attraction of the second fixed stable point. Finally we restore the original I_{ext}: the network relaxes to the stable state at high frequency. The external currents in the first

interval (prestimulation) and in the last interval (poststimulation) are the same.

Figure 8A shows the depolarization dynamics of a sample neuron as a function of time: in the prestimulation interval it is working in the ND regime, while in the poststimulation interval the neurons are in the SD regime. Figure 8B shows the probability of firing $\nu(t)$ per unit time (expressed in sp / s), sampled every 0.5 ms. And in Figure 8C, we see the rasters of 100 neurons: different lines correspond to different neurons in the same run. In both regimes there is no evidence for waves of synchronization.

The mean emission rate predicted by the theory (ν_{th}) in both stable states is in good agreement with the mean frequency obtained by averaging $\nu_s(t)$ (the fraction of neurons emitting spikes at time t in the simulation) over a time interval of 1 s on 10 different runs. For the low-rate theoretical fixed point, we have $\nu_{th} = 1.52$ Hz and in the simulation $\nu_{sim} = 1.45 \pm 0.14$ Hz, while for the high-rate fixed point $\nu_{th} = 99.1$ Hz and $\nu_{sim} = 94.5 \pm 1.7$ Hz. A similar quantitative agreement between theory and simulation was obtained for networks of RC neurons (Amit & Brunel, 1997b).

Also, the coefficient of variability of the ISI is quite close to the theoretical prediction: $CV_{th} = 0.87$ and $CV_{sim} = 0.88$ for a neuron emitting at mean rate 1.52 Hz in the low-rate stable state, and $CV_{th} = 0.14$ and $CV_{sim} = 0.11$ for a cell with mean frequency 99.5 in the high-rate fixed point. Even if the degree of variability is quite different for the two stable rates, the mean-field theory predictions capture both. In particular it is clear from the simulations that the fact that the CV is low (high frequency) does not affect the reliability of the theoretical predictions.

In Figure 9 we compare the distribution of the depolarization predicted by equation 3.7 with the results of the simulations for the two stable frequencies. Again the agreement is remarkably good for both the ND and SD regimes.

6 Discussion

Linear IF neurons are natural candidates for an aVLSI implementation of networks of spiking neurons. The study we performed shows that there is an operating regime in which the statistical properties of the firing process are qualitatively similar to those characterizing RC neurons. In a purely excitatory network, the collective behavior of linear neurons can sustain two stable states of activation corresponding to two different dynamical regimes (SD and ND). The reflecting barrier (which allows for the ND regime) is fundamental in order to have the change in the convexity required for two stable fixed points. Without this source of nonlinearity, it is not possible to have the coexistence of two states of activation in which both rates are different from zero. The role of the reflecting barrier is twofold. First, it is necessary in order to have the ND regime since, without it, the response function would not depend on σ. Second, it plays a role analogous to the

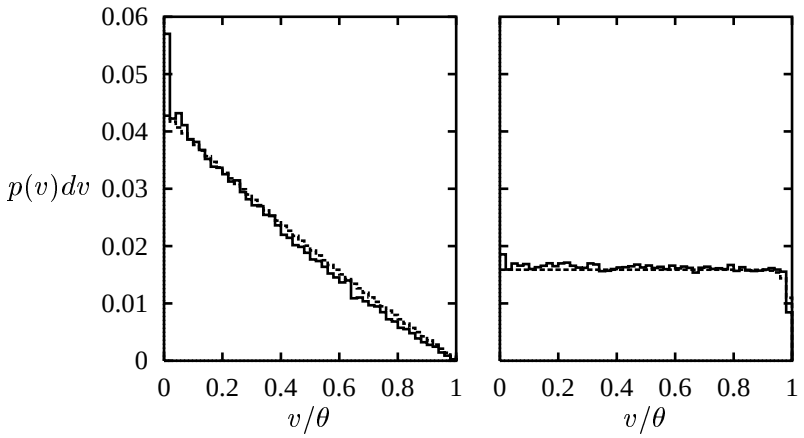

Figure 9: Comparison of the depolarization distributions $p(v)$ predicted by the theory (dashed line) and calculated from the simulations (solid line) (left: low-rate stable state; right: high-rate stable state). The bin is $dv = \theta/50$ in both cases. The agreement between theory and simulation is good, except in proximity of the reflecting barrier, where the quality of the diffusion approximation is degraded by the discreteness of the PSPs. This discrepancy is more evident when the neurons spend more time near the reflecting barrier (low rates).

exponential decay of the RC neuron in decorrelating the depolarization value from events (input spikes) in the past. For the RC neuron, the events in the past are forgotten in a time of order τ (see equation 2.1) because of the exponential decay, while for the linear neurons, they are forgotten whenever the depolarization would tend to be driven below the reflecting barrier. The time needed to forget any past event, in the absence of the input current ($I(t) = 0$), is θ/β, which can be considered as the time constant of the linear neuron.

In more complex networks, composed of RC or threshold-linear neurons, the second nonlinearity, which in the simple example presented in section 5 is due to the absolute refractory period, is generated dynamically by adding a population of strong inhibitory neurons (Amit & Brunel, 1997a; van Vreeswijk & Hasselmo, 1995). The introduction of an inhibitory population allows having both low spontaneous and a variety of low selective delay activity states. In a single population, spontaneous activity cannot be sustained by a feedback current that is of the same magnitude as the external current, which is more biologically plausible. With a single population, as in the case of the example shown in this article, the external current must be dominant for the low-rate stable solution. The coexistence of two fixed points with selective subpopulation is an outcome of the dynamical balance between excitation and inhibition, and it does not require a fine-tuning of

the parameters. Moreover, with a population of inhibitory neurons, a peak appears in the cross-correlations of spike emission times, as in experimental data for cortical recordings (Amit & Brunel, 1997b; Brunel & Hakim, 1998; Amit & Mattia, 1998).

The formalism used to derive the current-to-rate transduction function allows us also to study the dynamics of the transients (Knight, Manin, & Sirovich, 1996). Stability conditions and global oscillations are currently being investigated for RC neurons (Brunel & Hakim, 1999) and will be studied for aVLSI neurons in a future work.

Acknowledgments

We are grateful to D. J. Amit and P. Del Giudice for many useful suggestions that greatly improved a previous version of the manuscript and to N. Brunel for many helpful discussions.

References

Amit, D. J., & Brunel, N. (1997a). Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cerebral Cortex*, *7*, 237–252.

Amit, D. J., & Brunel, N. (1997b). Dynamics of a recurrent network of spiking neurons before and following learning. *Network*, *8*, 373–404.

Amit, D. J., Fusi, S., & Yakovlev, V. (1997). Paradigmatic working memory (attractor) cell in IT. *Neural Computation*, *9*, 1071–1092.

Amit, D. J., & Mattia, M. (1998). *Simulations and mean field analysis of a structured recurrent network of linear (VLSI) spiking neurons before and following learning*. Unpublished manuscript. Rome: Instituto di Fisica, University of Rome "La Sapienza."

Amit, D. J., & Tsodyks, M. V. (1991). Quantitative study of attractor neural network retrieving at low spike rates: I. Substrate—spikes, rates and neuronal gain. *Network*, *2*, 259.

Annunziato, M. (1995). *Hardware implementation of an attractor neural network with IF neurons and stochastic learning*. Thesis, Università degli Studi di Roma "La Sapienza."

Annunziato, M., Badoni, B., Fusi, S. & Salamon, A. (1998). Analog VLSI implementation of a spike driven stochastic dynamical synapse. In L. Niklasson, M. Boden, T. Ziemke (Eds.), *Proceedings of the 8th International Conference on Artificial Neural Networks, Skovde, Sweden* (Vol 1, pp. 475–480). Berlin: Springer-Verlag.

Brunel, N. & Sergi, S. (1999). Firing frequency of leaky integrate-and-fire neurons with synaptic dynamics, *Journal of Theoretical Biology*, in press.

Brunel, N., & Hakim, V. (1999). Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. *Neural Computation*, in press.

Bulsara, A. R., Elston, T. C., Doering, C. R., Lowen, S. B., & Lindenberg, K. (1996). Cooperative behaviour in periodically driven noisy integrate-fire models of neuronal dynamics. *Physical Review E, 53*, 3958–3969.

Cox, D. R., & Miller, H. D. (1965). *The theory of stochastic processes*. London: Metheuen.

Diorio, C., Hasler, P., Minch, B. A., & Mead, C. (1996). A single transistor silicon synapse. *IEEE Trans. Electronic Devices, 43*, 1972–1980.

Elias, J., Northmore, D. P. M., & Westerman, W. (1997). An analog memory circuit for spiking silicon neurons. *Neural Computation, 9*, 419–440.

Gerstein, G. L., & Mandelbrot, B. (1964). Random walk models for the spike activity of a single neuron. *Biophysical Journal, 4*, 41–68.

Knight, B., Manin, D., & Sirovich, L. (1996). Dynamical models of interacting neuron populations in visual cortex. In E. C. Gerf (Ed.), *Symposium on Robotics and Cybernetics; Computational Engineering in System Applications*. Liue, France: Cite Scientifique.

Mead, C. (1989). *Analog VLSI and neural system*. Reading, MA: Addison-Wesley.

Miyashita, Y., & Chang, H. S. (1988). Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature, 331*, 68.

Tuckwell, H. C. (1988). *Introduction to theoretical neurobiology* (Vol. 2). Cambridge: Cambridge University Press.

Usher, M., Stemmler, M., Koch, C., & Olami, Z. (1994). Network amplification of local fluctuations causes high spike rate variability, fractal firing patterns and oscillatory local field potentials. *Neural Computation, 6*, 795.

van Vreeswijk, C. A., & Hasselmo, M. E. (1999). Self-sustained memory states in a simple model with excitatory and inhibitory neurons. *Biol. Cybern*., in press.

van Vreeswijk, C. A., & Sompolinsky, H. (1996). Chaos in neural networks with balanced excitatory and inhibitory activity. *Science, 274*, 1724–1726.

Wilson, F. A. W., Scalaidhe, S. P. O., & Goldman-Rakic, P. S. (1993). Dissociation of object and spatial processing domains in primate pre-frontal cortex. *Science, 260*, 1955.

Received August 6, 1997; accepted June 10, 1998.