Population dynamics of excitatory and inhibitory spiking neurons: quantifying the contribution of spike timing to coding

V Del Prete and ACC Coolen

Department of Mathematics, King's College London
The Strand, London WC2R 2LS, UK

Abstract. In a recent work we have introduced an effective neuron model aiming at bridging the simplified description of integrate and fire neurons and the complex dynamics of conductance based models. Our effective formulation allows us to define a measure of multiple neurons temporal patterns characterizing the dynamics. Here we extend the previous study to analyze the dynamics of two interconnected populations, the first of excitatory, directionally selective, and the second of inhibitory cells, in presence of directional input stimuli. In a mean field approximation we can estimate qualitatively the contribution of spike timing to the information about the directional stimuli.

1. Introduction

One of the main questions addressed both by experimentalists and by theoreticians in modern neuroscience research is how time is used by the brain to process and transmit information. The evidences from the experiments are controversial: information estimates from real data seem to suggest a dominance of rate vs temporal coding in some cases [1, 2]. In other cases spike timing is found to significantly contribute to the total information [3].

From the theoretical point of view, estimating the contribution of spike timing to the information carried about external stimuli is difficult even in the simple case of an isolated cell, and an intimidating computational challenge in the case one deals with the dynamics of a population of neurons. Whether the single neuron dynamics is modeled in detail, via the Hodgkin-Huxley formalism, or via the simpler integrate and fire appoximation, the mathematical analysis at a network level becomes prohibitive in one case, and restricted to an effective mean rate description in the other. In a recent work [4] we have proposed a new effective model which bridges the too simplified description of the integrate and fire neuron and the highly non trivial dynamics of conductance based models. Here we revise and extend the previous study to analyze the dynamics of two interconnected populations of excitatory and inhibitory neurons responding to an external stimulus.

We show that our effective formulation allows us to detect and quantify the presence of pair temporal patterns characterizing the dynamics and to get a qualitative estimate of the contribution of spike timing to coding by means of methods borrowed from non-equilibrium statistical mechanics [5].

2. The single neuron model

The single neuron model has been already briefly introduced in [4]. For each neuron in each population we introduce two *effective* variables: an instantaneous rate ν_i^A and a phase ϕ_i^A ($A = E,I$ respectively for the excitatory and inhibitory population), whose dynamics is characterized as follows:

$$\frac{d}{dt}\phi_i^A = \nu_i^A \qquad \tau_\nu^A \frac{d}{dt}\nu_i^A = f[I_i^A] - \nu_i^A + \tau_\nu^A \xi_i^A \tag{1}$$

where τ_ν^A is a time constant and ξ_i^A denotes Gaussian white noise, with $\langle \xi_i^A(t) \rangle = 0$ and $\langle \xi_i^A(t)\xi_j^B(t') \rangle = (2T^A/\tau_\nu^A)\delta_{ij}\delta_{AB}\delta(t - t')$. $f[I_i^A]$ establishes the functional relationship between the input current I_i^A and ν_i^A, very much like the current-to-rate curve in standard rate models. Recent experiments [6] have shown that real neurons injected with a noisy current show a typical linear rise in their current-to-rate curve, and that two among the various mechanisms responsible for such behaviour are an early facilitation and a later adaptation in the firing of the cell. Moreover implementing these mechanisms in integrate and fire model neurons linearizes the steep initial rise in their current-to-rate curves (see also [7]). In order to take into account these effects we consider a threshold linear function for $f[I_i^A]$, with a threshold at $I = U_{\text{th}}/R$ [8, 6]:

$$f[I_i^A] = \frac{1}{\tau_m^A} \frac{\left[RI_i^A - U_{\text{th}}\right]^+}{U_{\text{th}} - U_{\text{hyp}}}; \tag{2}$$

where $[x]^+ = x\Theta(x)$ and $\Theta(x)$ is the Heaviside step function and one has:

$$I_i^A = I_i^{recA}(t) + S_i^A(t) \tag{3}$$

$I_i^{recA}(t)$ is the contribution to the synaptic current due to recurrent connectivities and $S_i^A(t)$ represents the input from previous stages responding to an external stimulus. We will come back later to the explicit expression of the two contributions $I_i^{recA}(t)$ and $S_i^A(t)$ respectively for the excitatory and inhibitory populations.

From its definition, it is clear that the integer values of ϕ_i^A represent the spike count: a spike is emitted whenever the phase assumes an integer value. Consistent with this picture, the full membrane potential, U_i^{mA} can be described as a function of ϕ_i^A, with spike peaks at integer values of the phase:

$$U_i^{mA} \qquad = U_i^{sA} + U_i^{pA} \tag{4}$$

$$U_i^{sA}(\phi_i^A, \nu_i^A) = \tau_s \Delta_u \, G_{\tau_s}\left[\frac{\sin(\pi\phi_i^A)}{\pi\nu_i^A}\right] \tag{5}$$

$$\tau_m^A \frac{d}{dt}U_i^{pA} \qquad = RI_i^A - U_i^{pA} - 2\tau_m^A(U_i^{pA} + U_{\text{hyp}}) \, G_{\tau_s}\left[\frac{\sin(\pi\phi_i^A)}{\pi\nu_i^A}\right] \tag{6}$$

The function $G_\epsilon(x)$ is a Gaussian distribution with zero mean and standard deviation $\frac{1}{2}\epsilon$.

Eq.(6) for U_i^{pA} characterizes the (passive) sub-threshold activity of the neuron in the form of a leaky integrator, with an additional third term on the right hand side acting both as a "clock" for the spike and as a built-in **analytical** mechanism for the reset. Eq.(5) provides a simplified expression for the supra-threshold activity U_i^{sA}, that is the spike peak. Therefore whenever the phase reaches an integer value the full potential U_i^{mA} will show a peak of the order of Δ_u in voltage and of a duration τ_s in time, and then it will fall down to the negative values of the hyperpolarization depending on the value of U_{hyp}. We assume that the parameters τ_s, Δ_u and U_{hyp} are the same for inhibitory and excitatory cells.

3. Network structure and population dynamics

We assume that out of the total number N of neurons, a fraction εN are excitatory and a fraction $(1 - \varepsilon)N$ are inhibitory. Each of the εN excitatory cells receives input from the other excitatory cells via connectivities $J_{ij} = J^{EE}/N > 0$ and from the $(1 - \varepsilon)N$ inhibitory cells via connectivities $J_{ij} = J^{EI}/N > 0$. Each of the inhibitory cell receives input from the excitatory cells via connectivities $J_{ij} = J^{IE}/N > 0$ and from the inhibitory cells via the connectivities $J_{ij} = J^{II}/N < 0$. Separating the degrees of freedom between sub and supra threshold activity as in eqs.(5),(6) is particularly convenient in the case where the dependence of the synaptic current on the potential of the cell itself can be neglected. In this case, with our choices for the synaptic strengths, the recurrent contributions to the synaptic currents $I_i^{recI}(t)$ and $I_i^{recE}(t)$ in eq.(3), respectively for an inhibitory and excitatory cell, will be:

$$I_i^{recE}(t) = \frac{1}{N}\left[\frac{J^{EE}}{\varepsilon}\sum_j U_j^{sE}(\phi_j^E, \nu_j^E) + \frac{J^{EI}}{1 - \varepsilon}\sum_j U_j^{sI}(\phi_j^I, \nu_j^I)\right]; \qquad (7)$$

$$I_i^{recI}(t) = \frac{1}{N}\left[\frac{J^{IE}}{\varepsilon}\sum_j U_j^{sE}(\phi_j^E, \nu_j^E) + \frac{J^{II}}{1 - \varepsilon}\sum_j U_j^{sI}(\phi_j^I, \nu_j^I)\right]; \qquad (8)$$

Both excitatory and inhibitory cells receive additional inputs $S_i^A(t)$ from from previous stages responding to an external stimulus. For the sake of simplicity we will consider the case of directional coding; the extension to an arbitrary stimulus is straightforward once the tuning properties of the single cells are known. We assume that only excitatory cells have a clear tuning for the direction, while inhibitory cells are broadly tuned. This is somewhat an approximation, since there are some types of inhibitory neurons have been shown to be tuned [9]. Therefore the external input to an excitatory cell i will depend on the direction ϑ of the stimulus and on the neuron preferred direction ϑ_0^i, while the input to an inhibitory cell will only depend on time:

$$S_i^E(t, \vartheta, \vartheta_i^E) = S_0^E(t)\alpha\cos^2(\frac{\vartheta - \vartheta_i^E}{2}) + S_1^E(t)(1 - \alpha); \qquad (9)$$

$$S_i^I(t) = S_1^I(t); \tag{10}$$

$$S_0^E(t) = S_0^E \; e^{-t/\tau_{st}^E}; \quad S_1^E(t) = S_1^E \; e^{-t/\tau_{st}^E}; \quad S_1^I(t) = S_1^I \; e^{-t/\tau_{st}^I}; \tag{11}$$

We assume that the orientation selectivities vary on a much longer time scale with respect to phases and rates, so that they do not depend on time; they are, in the statistical mechanics jargon, a source of *quenched disorder*.

For each of the two populations let us consider the time dependent distribution conditional to the stimulus and to a fixed configuration of the preferred directions: $p_t^A(\boldsymbol{\phi}^A, \boldsymbol{\nu}^A|\vartheta, \{\vartheta_i^E\})$. From standard theoretical arguments (see for example [5]) it is known that its dynamics is determined by a Fokker Planck equation (see for example [10]). It is clear that if we were able to solve it, we would implicitely have access to the detection and temporal evolution of cooperative phenomena (synchronies, time correlations, etc etc). Unfortunately solving a Fokker Planck equation for such a high dimensional system is an extremely hard computational problem. Let us consider the following average distributions:

$$p_t(\phi^E, \nu^E|\vartheta, \vartheta_0^E) = \langle \frac{1}{N\varepsilon} \sum_j \delta[\phi^E - \phi_j^E]\delta[\nu^E - \nu_j^E]\delta[\vartheta_0^E - \vartheta_j^E]\rangle_{p_t} \tag{12}$$

$$p_t(\phi^I, \nu^I|\vartheta, \vartheta_0^E) = \langle \frac{1}{N(1-\varepsilon)} \sum_j \delta[\phi^I - \phi_j^I]\delta[\nu^I - \nu_j^I]\rangle_{p_t} \tag{13}$$

where the averages in brackets are across the full population distributions $p_t^A(\boldsymbol{\phi}^A, \boldsymbol{\nu}^A|\vartheta, \{\vartheta_i^E\})$. The distributions (12),(13) quantify how many neurons *on average* across the population have a given rate, phase and preferred orientation. Notice that even the distribution for the inhibitory neurons depends on the preferred orientations, since inhibitory neurons also receive from the excitatory cells.

It can be shown [5, 4] that the dynamical equation for the distributions (12),(13) in the limit of large numbers of neurons is given by a Ornstein Uhlenbeck process [10, 5] and therefore it admits a gaussian solution. The details of the analytical evaluation will be reported elsewhere. The final solution is given by two systems of equations, which read as follow (we use the usual convention of the letter A for the formulas which are formally the same for excitatory and inhibitory neurons):

$$p_t^A(\phi^A, \nu^A|\vartheta, \vartheta_0^E) = p_t^A(\nu^A|\vartheta, \vartheta_0^E)p_t^A(\phi^A|\nu^A, \vartheta, \vartheta_0^E);$$

$$p_t^A(\nu^A|\vartheta, \vartheta_0^E) = \frac{e^{-\frac{1}{2}[\nu^A-\overline{\nu}^A(t)]^2/[\Sigma^A(t)]^2}}{\Sigma^A(t)\sqrt{2\pi}}; \quad p_t^A(\phi^A|\nu^A, \vartheta, \vartheta_0^E) = \frac{e^{-\frac{1}{2}[\phi^A-a^A(t)-b^A(t)\nu^A]^2/[\sigma^A(t)]^2}}{\sigma^A(t)\sqrt{2\pi}};$$

$$\Sigma^A(t) = \sqrt{e^{-2t/\tau_\nu^A}[\Sigma^A(0)]^2 + T^A(1 - e^{-2t/\tau_\nu^A})}$$

$$[\sigma^A(t)]^2 = [\sigma^A(0)]^2 + \frac{2T^A}{\tau_\nu^A}\int_0^t ds \; [b^A(s)]^2;$$

$$U^A(t) = \int \frac{d\phi^A d\nu^A}{2\pi\sigma^A(t)\Sigma^A(t)} \; e^{-\frac{1}{2}[\nu^A-\overline{\nu}^A(t)]^2/[\Sigma^A(t)]^2 - \frac{1}{2}[\phi^A-a^A(t)-b^A(t)\nu^A]^2/[\sigma^A(t)]^2} \; U^s(\phi^A, \nu^A)$$

$$= \tau_s\Delta_u \int \frac{d\phi^A d\nu^A}{2\pi\sigma^A(t)\Sigma^A(t)} \; e^{-\frac{1}{2}[\nu^A-\overline{\nu}^A(t)]^2/[\Sigma^A(t)]^2 - \frac{1}{2}[\phi^A-a^A(t)-b^A(t)\nu^A]^2/[\sigma^A(t)]^2} \; G_{\tau_s}\left[\frac{\sin(\pi\phi^A)}{\pi\nu^A}\right];$$

$$\frac{d}{dt}b(t) = 1 + \frac{b^A(t)}{\tau_\nu^A}\left\{1 - \frac{2T^A}{[\Sigma^A(t)]^2}\right\};$$

(14)

where $A = E$ for excitatory cells and $A = I$ for inhibitory cells. Moreover one has:

$$\overline{\nu}^E(t) = \overline{\nu}^E(0)e^{-t/\tau_\nu^E} + \int_0^t ds\ e^{-(t-s)/\tau_\nu^E}\int d\nu^E d\phi^E d\vartheta_0^E p_s^E(\phi^E, \nu^E|\vartheta, \vartheta_0^E)$$

$$\mathrm{x}f[\frac{J^{EE}}{\varepsilon}U^E(s) + \frac{J^{EI}}{1-\varepsilon}U^I(s) + S_0^E(s)\alpha\cos^2(\frac{\vartheta - \vartheta_0^E}{2}) + S_1^E(s)(1-\alpha)];$$

$$\frac{d}{dt}a^E(t) = \frac{b^E(t)}{\tau_\nu^E}\left\{\frac{2T^E\overline{\nu}^E(t)}{[\Sigma^E(t)]^2} - \int d\nu^E d\phi^E d\vartheta_0^E p_t^E(\phi^E, \nu^E|\vartheta, \vartheta_0^E)\right.$$

$$\left.\mathrm{x}f[\frac{J^{EE}}{\varepsilon}U^E(t) + \frac{J^{EI}}{1-\varepsilon}U^I(t) + S_0^E(t)\alpha\cos^2(\frac{\vartheta - \vartheta_0^E}{2}) + S_1^E(t)(1-\alpha)]\right\}$$

$$\overline{\nu}^I(t) = \overline{\nu}^I(0)e^{-t/\tau_\nu^I} + \int_0^t ds\ e^{-(t-s)/\tau_\nu^I}f[\frac{J^{IE}}{\varepsilon}U^E(s) + \frac{J^{II}}{1-\varepsilon}U^I(s) + S_1^I(s)];$$

$$\frac{d}{dt}a^I(t) = \frac{b^I(t)}{\tau_\nu^I}\left\{\frac{2T^I\overline{\nu}^I(t)}{[\Sigma^I(t)]^2} - f[\frac{J^{IE}}{\varepsilon}U^E(t) + \frac{J^{II}}{1-\varepsilon}U^I(t) + S_1^I(t)]\right\};$$

(15)

Once we know the distributions $p_t^E(\phi^E, \nu^E|\vartheta, \vartheta_0^E)$, $p_t^I(\phi^I, \nu^I|\vartheta, \vartheta_0^E)$ we can derive the temporal evolution for other quantities providing information about the dynamics. In particular we consider the distribution of two-spike time sequences:

$$\Gamma(\lambda, t, t + \Delta t) = \langle\frac{1}{2N(N-1)}\sum_{i\neq j}[\cos\left(\pi(\phi_i^A(t) - \phi_j^B(t+\Delta t) + \lambda)\right)]^2\rangle$$

$$\simeq \frac{1}{2} + \frac{1}{4}\left\{e^{-2\pi^2([\sigma^A(t)]^2 + b^2(t)[\Sigma^A(t)]^2)}\cos[2\pi(\langle\phi^A(t)\rangle]\cos[2\pi(\langle\phi^A(t)\rangle - \lambda)]\right.$$

$$\left. + e^{-2\pi^2([\sigma^B(t+\Delta t)]^2 + b^2(t+\Delta t)[\Sigma^B(t+\Delta t)]^2)}\sin[2\pi(\langle\phi^B(t+\Delta t)\rangle]\sin[2\pi(\langle\phi^B(t+\Delta t)\rangle - \lambda)]\right\}$$ (16)

where the brackets indicate an average across $p_t^A(\phi^A, \nu^A|\vartheta, \vartheta_0^E)$ and $\langle\phi^A(t)\rangle = a^A(t) + b^A(t)\overline{r}^A(t)$. The integer values of λ represent the difference in the spike count between the two neurons, which might grow with t due to the different average firing rates. Therefore, in order to detect all the events where the two neurons are in phase with a delay equal to Δt one has to consider any integer value of λ at each value of t. Averaging across a given time window one gets an estimate of how frequently two neurons tend to fire with a delay equal to Δt in a given time interval.

It must be said that eq.(16) actually counts how many times two neurons **are in phase**, and not simply the times they **fire together**, which corresponds to the subset where the two phases both assume an integer value. We are currently trying to define new theoretical measures in order to get closer to the true estimate.

We can now calculate the mutual information between the distributions and the input stimuli as a function of time; the separation of degrees of freedom between phases and rates turns to be particularly useful here as it will be clear from the following equations; for the sake of brevity we only report the analytical expression for the population of the excitatory cells:

$$\langle I^E(t)\rangle_{\vartheta_0^E} = \langle\int d\vartheta P(\vartheta)\int d\nu^E d\phi^E p_t(\phi^E, \nu^E|\vartheta, \vartheta_0^E)\log_2\left[\frac{p_t(\phi^E, \nu^E|\vartheta, \vartheta_0^E)}{p_t(\phi^E, \nu^E|\vartheta_0^E)}\right]\rangle_{\vartheta_0^E};$$ (17)

where the average in brackets is across the distribution of orientation selectivities $p(\vartheta_0^E)$, that we assume uniform across the population. It is clear that, since we are including the variability in the firing due to both the rate and the phase, this is the maximum information we can obtain in our mean field approximation for the real population distributions. Let us now integrate the variability due to the phase out of the full distribution:

$$p_t(\nu^E|\vartheta, \vartheta_0^E) = \int d\phi^E p_t(\phi^E, \nu^E|\vartheta, \vartheta_0^E); \tag{18}$$

Now we can consider the information in the approximation of a rate coding:

$$\langle I_{rate}^E(t)\rangle_{\vartheta_0^E} = \langle \int d\vartheta P(\vartheta) \int d\nu^E p_t(\nu^E|\vartheta, \vartheta_0^E) \log_2 \left[\frac{p_t(\nu^E|\vartheta, \vartheta_0^E)}{p_t(\nu^E|\vartheta_0^E)}\right]\rangle_{\vartheta_0^E}; \tag{19}$$

If now we compare the two value of the information we can get a qualitative estimate of the contribution of spike timing to coding.

We are currently solving the numerical equations and calculating the mutual information for different choices of the parameters in order to analyze the emergence of different dynamical regimes.

References

[1] S. Panzeri, S.R. Schultz, A. Treves, and E.T.Rolls. Correlations and the encoding of information in the nervous system. *Proceedings of the Royal Society of London Series B: Biological Sciences*, 266:1001–1012, 1999.

[2] V. Del Prete, A.E.P. Villa, and L. Martignon. Detection of syntonies between multiple spike trains using coarse grain binarization of spike counts distributions. *Network*, 15:13–28, 2004.

[3] R. Petersen, S. Panzeri, and M. Diamond. Population coding of stimulus location in rat somatosensory cortex. *Neuron*, 32:503–514, 2001.

[4] V. Del Prete and ACC Coolen. Non-equilibrium statistical mechanics of recurrent networks with realistic neurons. *Neurocomputing*, 2003. in press.

[5] A.C.C. Coolen. Statistical mechanics of neural networks ii: dynamics. In F. Moss and S. Gielen, editors, *Neuro-informatics and Neural Modelling*, volume 4 of *Handbook of Biological Physics*, pages 597–662, Amsterdam, 2000. Elsevier.

[6] A Rauch, G La Camera, H Luscher, W Senn, and S Fusi. Neocortical pyramidal cells respond as integrate and fire neurons to in vivo like input currents. *J. Neurophysiol.*, 90:1598–1612, 2003.

[7] G.B. Ermentrout. Linearization of f-i curves by adaptation. *Neural Comp.*, 10:1721–1729, 1998.

[8] S Fusi and M Mattia. Collective behaviour of networks with linear (vlsi) integrate and fire neurons. *Neural Comp.*, 11:633–652, 1999.

[9] J A Hirsch, L M Martinez, C Pillai, J M Alonso, Q Wang, and F T Sommer. Functionally distinct inhibitory neurons at the first stage of visual cortical processing. *Nature Neuroscience*, pages 1300–1308, 2003.

[10] CW Gardiner. *Handbook of stochastic methods*. Springer Verlag, Berlin and Heidelberg, 2002.