The neuron’s response at extended timescales

Daniel Soudry and Ron Meir

Department of Electrical Engineering, the Laboratory for Network Biology Research, Technion 32000, Haifa, Israel

Many systems are modulated by unknown slow processes. This hinders analysis in highly nonlinear systems, such as excitable systems. We show that for such systems, if the input matches the sparse “spiky” nature of the output, the spiking input-output relation can be derived. We use this relation to reproduce and interpret the irregular and complex 1/f response observed in isolated neurons stimulated over days. We decompose the neuronal response into contributions from its long history of internal noise and its “short” (few minutes) history of inputs, quantifying memory, noise and stability.

Many models, especially in biology, are accurate only below a certain timescale - due to the existence of additional slow processes. If these slow processes are not well characterized, it may be hard to predict how they will affect the dynamics at longer timescales. This is especially true if the dynamics are far from equilibrium, highly nonlinear and contain feedback, a regime where excitability is a typical dynamical phenomenon [1]. Among many types of excitable systems (e.g., [1] and references therein), a neuron is a prototypical example - where Action Potentials (AP - a stereotypical voltage “spike”) are generated in response to stimulation [2]. AP generation is indeed affected by many slow processes [3] - with new processes being discovered at an explosive rate [4,5]. This may entail a complex stochastic and history-dependent Input-Output (I/O) relation, on multiple timescales [6,7]. In general, it is hard to identify, simulate or analyze such an I/O due to the large number of processes which are unknown or lacking known parameters.

We find the situation simplifies considerably if we use (experimentally relevant [10,13]) sparse spike inputs, similar to the typical output of the neuron (Fig. 1A&B). We derive, for a general biophysical stochastic neuron model (Eqs. 1-3) with a few assumptions, a concise description for the I/O (Eqs. 17) based on biophysically meaningful parameters. This I/O is well described by an ‘engineering-style’ block diagram with feedback (Fig. 1C), which can be used to decompose the effects of noise and input on the response. Beyond the conceptual lucidity, such a linear I/O allows the utilization of well known statistical tools to derive all second order statistics, construct linear optimal estimators and perform parameter identification. These results hold numerically, even sometimes when our assumptions break down.

We demonstrate the utility of our results on recent experiments [13] where synthaptically isolated individual neurons, from rat cortical culture, were stimulated with extra-cellular sparse current pulses for an unprecedented duration of days. The neurons exhibited 1/fα statistics [14], responding in a complex and irregular manner from seconds to days. Using our results, we are able to reproduce and analyze the origins of this behavior in a biophysical model, showing that slow processes span a wide range of timescales - with slower processes being “noisier”, due to low ion channel population numbers. The model suggests the 1/fα statistics of the response originates from the long history of internal noise, while input fluctuations only affect the response on a (relatively) short timescale of a few minutes.

The voltage dynamics of an isopotential neuron are determined by ion channels, protein pores which change conformations stochastically with voltage-dependent rates [16]. On the population level, such dynamics are generally very well described by models of the form [17,19]

\[ \dot{V} = f(V, r, s, I(t)) \]  
\[ \dot{r} = A_r(V) r + B_r(V, r) \xi_r \]  
\[ \dot{s} = A_s(V) s + B_s(V, s) \xi_s \]
with voltage $V$, stimulation current $I(t)$, rapid variables $r$ (e.g., $m, n, h$ in the Hodgkin-Huxley (HH) model [2]), slow variables $s$ (e.g., slow sodium inactivation [20]), rate matrices $A_{r/s}$, white noise processes $\xi_{r/s}$ (with zero mean and unit variance), and matrices $B_{r/s}$ which can be written explicitly using the rates and ion channel numbers [21] ($D = BB^\dagger$ is the diffusion matrix [21][22]). For simplicity, we assumed $r$ and $s$ are not coupled directly, but this is non-essential [23][24]. The parameter space can be constrained [19], since we consider here only excitable, non-oscillatory neurons which do not fire spontaneously and which have a single resting state - as common for cortical cells, e.g., [13]. Such biophysical neuronal models (Eqs. [1][3]) are generally complex non-linear models, containing many variables and unknown parameters (sometimes ranging in the hundreds [25][26]), not all of which can be identified [27]. Therefore, such models are notoriously difficult to tune, highly susceptible to overfitting and computationally expensive [28][30]. Also, the high non-linearity usually prevents exact mathematical analysis of such models at their full level of complexity [31].

**Model reduction** However, much of the complexity in such models can be overcome under a well defined and experimentally relevant settings [10][14], if we use sparse inputs, similar in nature to the spikes commonly produced by the neuron. This is done by “averaging out” Eqs. [1][3] using similar methods to those in [19]. Specifically, suppose $I(t)$ is a pulse train arriving at times $\{t_m\}$ (Fig. 1A, top), so $T_m = t_{m+1} - t_m \gg \tau_{AP}$ with $\tau_{AP}$ being the timescale of an AP (Fig. 1B). Our aim is to describe the AP occurrences $Y_m$, where $Y_m = 1$ if an AP occurred immediately after the $m$-th stimulation, and 0 otherwise (Fig. 1A, bottom). To do so, we need to integrate Eqs. [1][3] between $t_m$ and $t_{m+1}$. Since $T_m \gg \tau_{AP}$ the rapid AP generation dynamics of $(V, r)$ relax to a steady state before $t_{m+1}$. Therefore, the neuron AP “remembers” any history before $t_m$ only through $s_m = s(t_m)$. Given $s_m$, the response of the fast variables $(V, r)$ to the $m$-th stimulation spike will determine the probability to generate an AP. This probability, $p_{AP}(s)$, collapses all the relevant information from Eqs. [1][2] and can be found numerically from the pulse response of Eqs. [1][3] with $s$ held fixed [21]. In order to integrate the remaining Eq. [3] we define the averaged rate matrix

$$A(Y_m, T_m) = \tau_{AP} T_m^{-1} (Y_m A_+ + (1 - Y_m) A_-) + (1 - \tau_{AP} T_m^{-1}) A_0,$$

where $A_+$, $A_-$ and $A_0$ are the averages of $A_s$ during an AP response, a failed AP response and rest, respectively. Assuming $T_m \ll \tau_s$, we obtain, to first order

$$s_{m+1} = s_m + T_m A(Y_m, T_m) s_m + n_m,$$

where $n_m$ is a white noise process with zero mean and variance $T_m D(Y_m, T_m, s_m)$ (defined similarly to $T_m A(Y_m, T_m)$). Note that this simplified linear discrete time map has far fewer parameters than the full model, since it is written explicitly only using the averaged microscopic rates of $s$ (through $A$ and $D$), population sizes (through $D$), the probability to generate an AP given $s$, $p_{AP}(s)$, and the relevant timescales. This effective model exposes the large degeneracy in the parameters of the full model and leads to significantly reduced simulation times and mathematical tractability.

**Linearization** Intrinsic ion channel noise can be exploited to linearize the neuronal dynamics, rendering it more tractable than the (less realistic) noiseless case [19]. Suppose that $\{T_m\}$ has stationary statistics with mean $T_s$ so that $\tau_{AP} \ll T_m \ll \tau_s$ with high probability. Since $s$ is slow and AP generation is rather noisy in this regime [19] (so $p_{AP}(s_m)$ is slowly varying [24]), we assume a stable excitability fixed point $s_*$ exists, so perturbations $\delta s_m = s_m - s_*$ are small and we can linearize $p_{AP}(s_m) \approx p_s + w^\top \delta s_m$. The mean AP firing rate can be found self-consistently (and rather accurately, Fig. 2A) from the location of the fixed point $s_*$

$$\langle Y_m \rangle = p_s = p_{AP}(s_*(p_s, T_*)),$$

while the perturbations around the fixed point are described by the linear system

$$\delta s_{m+1} = F \delta s_m + d I_T + a \tilde{Y}_m + n_m, \quad (6)$$

$$\tilde{Y}_m = w^\top \delta s_m + e_m \quad (7)$$

with $F = I + T_s A(p_s, T_*)$, $\langle n_m n_m^\top \rangle = T_s D(p_s, T_*, s_*)$, $e_m$ is a white noise process (other parameters in [37]). This linear I/O, which contains feedback from the ‘output’ $\tilde{Y}_m$ to the state variable $\delta s_m$ (Fig. 1C), can be very helpful mathematically and all its parameters are directly related to well motivated biophysical variables. Moreover, this formulation makes it now possible to construct optimal linear estimators for $Y_m$ and $s_m$ [22] (Fig. 2B), perform parameter identification, and use standard tools [22] to find all second order statistics in the system, such as correlations or Power Spectral Densities (PSD). For example, the PSD for $Y_m$ is

$$S_Y(f) = w^\top H_c(f) \left(D(p_s, T_*, s_*) + d d^\top S_T(f) H_c^\top (-f) w + T_s p_s (1 - p_s) (1 + w^\top H_c(f) a_0)^2 \right) \quad (8)$$

where $H_c(f) = (2 \pi f I - A(p_s, T_*) - T_s^{-1} a w^\top)^{-1}$, for $f \ll T_s^{-1}$. Again, note the large degeneracy here - many different parameters will generate the same PSD. Other immediately derivable statistics are $S_{\delta s}$, the cross-PSDs $S_{\delta Y}(f)$, $S_{\delta s Y}(f)$ and $S_{\delta Y}^2(f)$ as well as the also respective correlations. Our exact results agree very well with the numerical solution of Eqs. [1][3] (Fig 2C-D), even in some cases where the underlying assumptions do not hold (specifically if $T_m \sim \tau_s$ and $s_*$ is unstable [24]).
Modeling the Experiment  Next, we use the expression for the PSD (Eq. 5) to find conditions under which the experimental results of [13] can be reproduced, given our assumptions. We will show that kinetic processes must possess a large range of slow timescales, where the number of relevant ion channels with each timescale scales with an exponent $\alpha$. Fitting a specific model we reproduce the experimental results in Fig. 3.

Previous work [19] used a stochastic HH model with slow sodium inactivation to reproduce the experimental results of [13] up to a timescale of minutes. However, this model cannot explain dynamics on longer timescales (Fig. 7C). Specifically, we require $S_Y(f) \sim f^{-\alpha}$ for $f < 10^{-2}$Hz, with $\alpha \approx 1.4$ when $T_m = T_s$, as in [13]. If the total number of ion channel states of all channel types is finite, then we can decompose Eq. 8

$$S_Y(f) = c_0 + \sum_{i=1}^{M} \frac{c_i \lambda_i}{(2\pi f)^2 + \lambda_i^2},$$

where $c_i$ are some (derivable) constants and the poles $\lambda_i$ are solutions of the characteristic equation

$$|A - A(p_s, T_s) - T_s^{-1}aw^\top| = 0.$$

In order to approximate a PSD of the form $f^{-\alpha}$, we require the poles to cover a large range [15]. Though the $T_s^{-1}aw^\top$ feedback term can tune the location of the poles (through the variable $a$, see Fig. 4C), comparison with experiment implies that the observed $f^{-\alpha}$ dependence was not generated in this way, since it exists even near the critical stimulation frequency, where $p_s \rightarrow 1, w \rightarrow 0$. Therefore, the eigenvalues of $A(p_s, T_s)$, the average rate matrix, must span a large range of (inverse) timescales. However, the existence of this range is not sufficient - we also require [13] $c_i \propto \lambda_i^{-\alpha}$ so that

$$S_Y(f) \sim \int \frac{d\lambda \lambda^{1-\alpha}}{(2\pi f)^2 + \lambda^2} \propto f^{-\alpha}.$$

Analyzing Eq. 8 we find that we can generate this in a robust way (independent of $T_s$), that is consistent with other experimental observations, by having a scaling relation in $N_i$, the number of the corresponding ion channels (affecting $D(p_s, T_s, s_m)$, the variance of $n_m$ in Eq. 9). Specifically, by setting $N_i \propto \lambda_i^{-\alpha}$, implying that slower processes are noisier. To fit a specific model consider the stochastic HH model which augments the basic HH model with a slow sodium inactivation variable [19, 24].

We extend this model by replacing the sodium inactivation variable $s$ with $\sum_{i=1}^{M} s_i/M$, where $s_1$ is identical to $s$, and for $\{s_i\}_{i=2}^{M}$ the rates scale as $c^i$ and the channel numbers scale as $N_i \propto c^{\alpha i}$. In order to obtain $\alpha = 1.4$ for the duration of the experiment we require $c = 0.2$ and $M = 5$. This reproduces well the observed scaling relations (Fig. 3).

Predictions  First, we consider the duration of the neuronal memory. To quantify this more precisely, we note that in the frequency domain we can use spectral factorization [22] and write the response of the linear system (Eqs. 6-7) as

$$\tilde{Y}(f) = H_{\text{signal}}(f) \tilde{T}(f) + H_{\text{noise}}(f) \nu(f)$$

with $H_{\text{signal}}(f) = w^\top H_{\text{c}}(f) d$ and $\nu(f)$ is the Fourier transform of a zero mean white noise representing internal neuronal noise. In our model, the $f^{-\alpha}$ behavior is generated by fluctuations in internal neuronal noise, so $H_{\text{noise}}(f) \sim f^{-\alpha}$. However, $H_{\text{signal}}(f) \sim c$ for $f \rightarrow 0$ since it is not affected by $N_i$. This entails that the neuron possesses a long memory for its intrinsic fluctuations but a “finite memory” of its input - i.e. perturbations in $Y_m$ due to perturbations in $T_m$ will decay exponentially with a finite timescale. Specifically, in the fitted model, $H_{\text{signal}}(f)$ has the shape of a simple low pass filter with a timescale of $\sim 100$sec. This could also be probed experimentally directly by applying a sinusoidal $T_m$ input

$$T_m = T_s + \sum_{i} T_i \sin (2\pi f_i T_s m),$$

since the linear response of the model will generate a direct probe of $H_{\text{signal}}(f)$ (Fig. 4A). Note that this prediction of input memory timescale of $\sim 100$sec is valid only for spiking input, in the context of our model.
However, \( \text{Var} (\alpha \rightarrow f) \) is generally hard if internal noise is high [34], since continuous input signals may very well generate long memory effects [33].

Second, we quantify the “noisiness” of the neuronal response. In the fitted model the neuronal response under periodic stimulation is very noisy - in the sense that linear optimal estimation of \( Y_m \) performs similarly to the trivial predictor \((Y_m = \langle Y_m \rangle - 0.5)\), even if \( s_m \) is fully known (since \( e_m \) has a large variance). In order to improve predictability of \( Y_m \), we can increase the variability of the input \( T_m \). To test this we examined data from a similar experiment where the variability of \( T_m \) was higher than the internal noise at certain frequencies, so it was possible to estimate \( S_Y(f) = H_{signal}(f) S_f(f) \) in these regions (which is generally hard if internal noise is high [34]), which seems to correspond well with our fitted model (Fig. 3B). Therefore, input variability in inter-spike intervals \( T_m \) can increase signal to noise ratio, as was previously observed for current amplitude variability over shorter time scales of milliseconds [33].

Finally, consider the stability of the neuronal response for even longer timescales. Generally, \( S_Y(f) \sim f^{-\alpha} \) as \( f \rightarrow 0 \) with \( \alpha > 1 \) implies \( \text{Var} (Y_m) \sim m^{\alpha-1} \) as \( m \rightarrow \infty \). However, \( \text{Var} (Y_m) \leq 0.25 \) since \( Y_m \) is binary, and therefore the scaling behavior must have some cutoff. Extrapolating the experimental results suggests that for \( \alpha = 1.4 \) this cutoff could be reached approximately only after four years of ongoing experiments - which is comparable with the lifespan of a rat (in comparison, \( \alpha = 2 \) gives a cutoff of a few days). In general, \( M = 5 \) poles should allow coverage of all this timescale range [15]. However, in our model, the scaling is limited by channel numbers. Since typically \( 1 < N_i < 10^6 [36] \), scaling in \( N_i \) can generate \( f^{-\alpha} \) in the PSD over \( 6/\alpha \) Orders of Magnitude (OM). For \( \alpha = 1.4 \), we get 4.2 OM, while [13] covered about 3 OM.

To summarize this section, we used the linearity of our derived I/O to decompose the contributions of inputs and internal noise to the response of the neuron. This decomposition shows that even though the neuron can “remember” its intrinsic fluctuations over timescales of days, its memory of past inputs can be relatively “short”. For example, the input memory of our fitted model decay exponentially with a timescale of \( \sim 100 \) sec. We suggest an experiment to test this directly. Additionally, this linear decomposition allows us to quantify the signal-to-noise ratio of the I/O and show that it increases with input spike time variability. Finally, we set upper limits on the timescale range of the observed \( 1/f^\alpha \) behavior.

Acknowledgments The authors are grateful to O. Barak, N. Brenner, Y. Elhanati, A. Gal, T. Knaf\, Y. Kafri and S. Marom for insightful discussions and reviewing parts of this manuscript. The research was partially funded by the Technion V.P.R. fund and by the Intel Collaborative Research Institute for Computational Intelligence (ICRI-CI).

[1] B. Lindner, Phys. Rep. 392, 321 (2004).
[2] A. L. Hodgkin and A. F. Huxley, J. Physiol. 117, 500 (1952).
[3] S. Marom, Prog. Neurobiol. 90, 16 (2010).
[4] B. P. Bean, Nat. Rev. Neurosci. 8, 451 (2007).
[5] P. J. Sjostrom, E. A. Rancz, A. Roth, and M. Häusser, Physiol. Rev. 88, 769 (2008).
[6] D. Debanne, E. Campanac, A. Bialowas, and E. Carlier, Physiol. Rev. 91, 555 (2011).
[7] S. B. Lowen, L. S. Liebovitch, and J. A. White, Phys. Rev. E (1999).
[8] G. Gilboa, R. Chen, and N. Brenner, J. Neurosci. 25, 6479 (2005).
[9] B. P. Bean, Nat. Rev. Neurosci. 8, 451 (2007).
[10] P. J. Sjostrom, E. A. Rancz, A. Roth, and M. Häusser, Physiol. Rev. 88, 769 (2008).
[11] D. Debanne, E. Campanac, A. Bialowas, and E. Carlier, Physiol. Rev. 91, 555 (2011).
[12] R. De Col, K. Messlinger, and R. W. Carr, J. Physiol. 586, 1089 (2008).
[13] A. Gal, D. Eytan, A. Wallach, M. Sandler, J. Schiller, and S. Marom, J. Neurosci. 30, 16332 (2010).
[14] J. H. Goldwyn, J. T. Rubinstein, and E. Shea-Brown, J. Neurophysiol. 108, 1430 (2012).
[15] M. S. Keshner, Proc. IEEE 70, 212 (1982).
[16] B. Hille, Ion Channels of Excitable Membranes (Sinauer Associates, Sunderland, MA 01375, 2001), 3rd ed.
[17] R. Fox and Y. Lu, Phys. Rev. E 49, 3421 (1994).
[18] J. H. Golwyn, N. S. Iemnovo, M. Fambare, and E. Shea-Brown, Phys. Rev. E 83, 041908 (2011).
[19] D. Soudry and R. Meir, Front. Comput. Neurosci. 6 (2012).
[20] W. K. Chandler and H. Meves, J. Physiol. 211, 707 (1970).
[21] P. Orio and D. Soudry, PLoS One 7, e36670 (2012).
[22] C. Gardiner, Handbook of stochastic methods (Springer, Verlag Berlin Heidelberg, 2004), 3rd ed.
[23] G. Wainrib, M. Thieullen, and K. Pakdaman, J. Comput. Neurosci. 32, 327 (2011).
[24] See Supplemental Material.
[25] C. Koch and I. Segev, Methods in Neuronal Modeling: from Ions to Networks, vol. 484 (MIT press, Cambridge, 1989), 2nd ed., ISBN 0-262-11231-0.
[26] A. Roth and M. Häusser, J. Physiol. 535, 445 (2001).
[27] Q. J. M. Huys, M. B. Ahrens, and L. Paninski, J. Neurophysiol. 96, 872 (2006).
[28] W. Gerstner and R. Naud, Science 326, 379 (2009).
[29] M. Migliore, C. Cannia, W. W. Lytton, H. Markram, and M. L. Hines, J. Comput. Neurosci. 21, 119 (2006).
[30] S. Druckmann, T. K. Berger, F. Schürmann, S. Hill, H. Markram, and I. Segev, PLoS Comput. Biol. 7, e1002133 (2011).
[31] B. Ermentrout and D. Terman, Mathematical Foundations of Neuroscience, vol. 35 (Springer Verlag, New York, 2010).
[32] B. D. O. Anderson and J. B. Moore, Optimal Filtering, vol. 11 (Prentice hall, Englewood Cliffs, NJ, 1979).
[33] D. Soudry and R. Meir, Front. Comput. Neurosci. 4 (2010).
[34] J. S. Bendat and A. G. Piersol, Random Data Analysis and Measurement Procedures, vol. 11 (Wiley, New York, NY, 2000), 3rd ed.
[35] Z. F. Mainen and T. J. Sejnowski, Science 268, 1503 (1995).
[36] P. Rowat, Neural Comput. 19, 1215 (2007).
[37] \( \langle e_m \rangle = \langle e_m n_m \rangle = 0, \langle e_m^2 \rangle = p_* (1 - p_*), n_m d = A_0 s_* \) and \( a = \tau_{AP} (A_+ - A_-) s_* \).