Fluctuation-Dissipation Relations for Spiking Neurons

Benjamin Lindner

Bernstein Center for Computational Neuroscience Berlin, Philippstraße 13, Haus 2, 10115 Berlin, Germany and Physics Department of Humboldt University Berlin, Newtonstraße 15, 12489 Berlin, Germany

(Received 1 April 2022; revised 27 September 2022; accepted 17 October 2022; published 31 October 2022)

Spontaneous fluctuations and stimulus response are essential features of neural functioning, but how they are connected is poorly understood. I derive fluctuation-dissipation relations (FDR) between the spontaneous spike and voltage correlations and the firing rate susceptibility for (i) the leaky integrate-and-fire (IF) model with white noise and (ii) an IF model with arbitrary voltage dependence, an adaptation current, and correlated noise. The FDRs can be used to derive thus far unknown statistics analytically [model (i)] or the otherwise inaccessible intrinsic noise statistics [model (ii)].

DOI: 10.1103/PhysRevLett.129.198101

Small physical systems often display considerable fluctuations that can be characterized by correlation functions or power spectra. Fluctuation-dissipation relations connect the statistics of these spontaneous fluctuations of certain observables to their mean response to a time-dependent perturbation. Originally proposed for equilibrium thermodynamic systems [1,2], they have been extended to nonequilibrium setups with a steady state [3–6]. Fluctuation-dissipation theorems can be used to infer the response properties from observations of purely spontaneous activity, to prove in a model-free way that a system operates outside thermodynamic equilibrium [7,8], or to test whether a system obeys a Markovian description [9,10]; for general reviews on applications of fluctuation-dissipation relations (FDRs), see the comprehensive reviews [11,12].

Fluctuations are especially prominent in neural systems, specifically in the spike generation of neurons (nerve cells) in the brain, which is reflected in a long history of stochastic modeling in neuroscience [13,14]. Neurons are notoriously noisy due to intrinsic sources of fluctuations (e.g., channel noise and unreliable synaptic transmission); in the recurrent networks of the cortex, the nonlinear interactions among many pulse-generating units lead to a strong chaotic variability (a network noise) even if single units follow a completely deterministic dynamics (i.e., the above mentioned channel noise, for instance, is neglected) and even if external (noisy) stimulation is absent. Most importantly, the response to external signals is of overarching importance for nerve cells, as it characterizes the transmission and processing of information, which is the main task of these cells. So, it is of vital importance to understand potential connections between the statistics of spontaneous activity and the response to a time-dependent perturbation in the case of spiking neurons.

Let us consider a paradigmatic stochastic model of computational neuroscience, the leaky integrate-and-fire (IF) model with white noise \(\xi(t)\) and a time-dependent current signal \(s(t)\):

\[
\frac{dv}{dt} = -v + \mu + s(t) + \sqrt{2D}\xi(t). \tag{1}
\]

The voltage across the nerve membrane \(v(t)\) upon reaching a threshold \(v_T\) is reset to \(v_R < v_T\) and, simultaneously, the time instant is registered as a spike time \(t_s\). The most important output of this model is the spike train, \(x(t) = \sum \delta(t - t_s)\) (this is what is communicated to other cells). In Eq. (1) time and voltage are measured in multiples of the membrane time constant \(\tau_m\) and the threshold-reset distance, respectively. The mean constant input \(\mu\) and the intensity of the white noise \(D\) are important parameters that determine the stochastic regime of the model [15].

For the spontaneous activity \(s(t) \equiv 0\) the power spectrum of the spike train can be analytically calculated and expressed in terms of parabolic cylinder functions \(D_\alpha(x)\) [16]:

\[
S_{xx}(\omega) = r_0 \left|\frac{D_{10}(z_T)^2 - \frac{\omega^2}{4z_T^2}D_{10}(z_R)^2}{D_{10}(z_T) - e^{\frac{\omega^2}{4z_T}}D_{10}(z_R)}\right|^2. \tag{2}
\]

Here \(z_T/R = (\mu - v_{T/R})\sqrt{D}\) and \(r_0 = \langle x(t) \rangle\) is the stationary firing rate [17] with the angular brackets indicating an ensemble average.

The response to a weak signal \(s(t)\) is quantified by the time-dependent rate modulation \(r(t) \approx r_0 + K_s \ast s(t)\), given in terms of a convolution with the linear-response function \(K_s(t)\) or in terms of the susceptibility \(\chi_s(\omega)\) (the Fourier transform of \(K_s\)), which can be expressed by confluent hypergeometric functions [18] or, equivalently, again in terms of parabolic cylinder functions [19]:

\[
\chi_s(\omega) = \frac{i r_0 \omega / \sqrt{D} D_{10^{-1}}(z_T) - e^{\frac{\omega^2}{4z_T}}D_{10^{-1}}(z_R)}{D_{10}(z_T) - e^{\frac{\omega^2}{4z_T}}D_{10}(z_R)}. \tag{3}
\]
There is some structural similarity in the expressions for power spectrum and susceptibility—both are given in terms of ratios of differences of parabolic cylinder functions, but, apparently, it is not possible to express one in a simple way by the other. So, even in this case, where we know the explicit solutions for the two characteristics of spontaneous fluctuations and of the response to a stimulus, it does not help us to connect them in a fluctuation-dissipation relation. The situation is similar (analytical expressions are known but cannot be related) for IF models with shot noise [20,21], with dichotomous background noise [22], or escape noise [23].

Here I connect the statistics of spontaneous spiking and the firing rate response to a weak signal by means of a simple calculation, which is markedly different from the typical derivation of the standard FDR [24] and also to recent calculations for IF models in discrete time and embedded in networks [25]. The approach here builds on two ideas: (i) the reset can be incorporated into the Langevin dynamics by means of the spike train (see, e.g., Refs. [26] or [27]), which permits one to average this and related equations, leading by the Rice method to equations for spectral measures, and (ii) by means of the Furutsu-Novikov theorem [28,29], I can relate the noise-spike-train correlator to the exact linear-response function. I first outline this calculation for the simple model in Eq. (1) and then treat the biophysiologically more realistic and dynamically richer exponential IF model endowed with correlated (colored) current noise.

LIF model with white noise.—Without signal current \( s(t) \equiv 0 \), one can rewrite Eq. (1) as follows:

\[
d\xi = -\mu + \sqrt{2D_s(t)} - (v_T - v_R)\xi(t),
\]

where the last term formally imposes the fire-and-reset rule: the delta functions in \( \xi(t) \) will push the voltage back from the threshold at \( v_T \) to the reset point \( v_R \). Having incorporated the reset rule into the equation, one can now take averages over a stationary ensemble. A direct average of Eq. (4), for instance, yields \( \langle d\xi/dt \rangle \langle v \rangle = 0 = \langle -\mu - (v_T - v_R)v_0 \rangle \), which leads to \( \langle v \rangle = \mu - (v_T - v_R)v_0 \), a nontrivial relation between mean membrane voltage and firing rate.

More importantly, one obtains relations for the second-order statistics, specifically cross spectra and power spectra, as follows. Taking the complex-conjugated Fourier transform of Eq. (4) at nonvanishing frequency, I get \( \langle \tilde{v} \rangle = \sqrt{2D_s} - (v_T - v_R)\tilde{x} \). Multiplication with \( \tilde{x} \) and averaging over the noise ensemble yields

\[
\langle \tilde{x}^2 \rangle = 2D\langle \tilde{x}^2 \rangle - (v_T - v_R)\langle \tilde{x} \rangle.
\]

To calculate \( \langle \tilde{x}^2 \rangle \) one may invoke the Furutsu-Novikov theorem [28,29]. To see this works for the specific problem, I split the Gaussian noise \( \tilde{\xi} \) into \( N \) independent statistically identical sub processes \( \tilde{\xi}_n \) with power spectra \( S_{nn} = S_{\xi\xi}/N \). The average is now \( \langle \tilde{x}^2 \rangle = \sum \langle \tilde{x}_n^2 \rangle = N\langle \tilde{x}_n^2 \rangle \) (with an arbitrary index \( n \)) and can be split into averages over the single noise realization \( \tilde{\xi}_n \) and over realizations of the remaining noise processes \( \tilde{\xi}_{n'} \):

\[
\langle \tilde{x}^2 \rangle = \langle \langle \tilde{x}_n^2 \rangle \rangle_{\tilde{\xi}_{n' \neq \tilde{\xi}_n}} = \langle \langle \tilde{x}_n \rangle \rangle_{\tilde{\xi}_{n' \neq \tilde{\xi}_n}}^2.
\]

For \( N \to \infty \) the single subprocess \( \tilde{\xi}_n \) becomes infinitely weak, satisfying perfectly the requirement of linearity of the Fourier transform (for a frozen realization of \( \tilde{\xi}_n \)) satisfies \( \langle \tilde{x}_n \rangle_{\tilde{\xi}_{n' \neq \tilde{\xi}_n}} = \chi(w; D)\sqrt{2D_{\tilde{\xi}_n}} \). Here I have highlighted the explicit dependence of the susceptibility on the noise intensity (that I later will omit again)—for \( N \to \infty \) practically all of the noise (i.e., the remaining processes \( \tilde{\xi}_{n'} \)) act as a background noise. Combining all the results above, one obtains \( \langle \tilde{x}^2 \rangle = \sqrt{2D_s}（w）S_{\xi\xi} = \sqrt{2D_s}（w）（because S_{\xi\xi} = 1 for the white noise），resulting in

\[
\chi_s（w） = \frac{v_T - v_R}2S_{xx}（w） + (1+iw)S_{xx}（w）/2D.
\]

On the left hand side, we find the susceptibility of the firing rate with respect to a weak time-dependent signal, as can be, for instance, determined by a periodic stimulation \( s(t) = e \cos（o_0t） \) from the rate modulation \( r(t) = \langle x(t) \rangle = r_0 + |x_s（o_0）| \cos（o_0t - \arg x_s（o_0）） \) [here argument is the complex argument] or, equivalently, by a weak broadband stimulus as was used in Fig. 1(a). On the right side of Eq. (7) are statistics of the spontaneous activity \( s(t) \equiv 0 \): Besides the spike-train power spectrum, the cross spectrum between the subthreshold membrane voltage and the generated spikes emerges—this is the missing link between the spontaneous fluctuation statistics and the response statistics. For a selected parameter set, the relation is tested and confirmed in Fig. 1(a).

Because in the special case of the LIF model [Eq. (1)] one knows most of the statistics by explicit expressions, one can use the relation above to determine the cross spectrum between \( v(t) \) and \( x(t) \) analytically:

\[
S_{xx}（w） = \frac{2D\chi_s（w） - (v_T - v_R)S_{xx}（w）}{1+iw},
\]

which by virtue of Eq. (2) and Eq. (3) can be expressed by parabolic cylinder functions and is confirmed in Fig. 1(b) by numerical simulations.

Exponential IF model with adaptation current and colored noise.—We now turn to a more general and biophysically more realistic model, which, in its essential ingredients, has been justified on theoretical grounds [30] but also extracted from data [31]. As suggested by Brette and Gerstner [32], I include an adaptation current [33] and replace the linear leak term with the function \( f(v) = \mu - v + \Delta_s \exp[(v - v_a)/\Delta_a] \) (see also Ref. [30]):
Im\[Svx(\omega)\] 

\[ Re[\chi(\omega)] \] 

\[ Im[\chi(\omega)] \] 

\[ \text{fluctuation statistics} \] 

\[ \text{theory} \] 

FIG. 1. White-noise-driven leaky IF model. (a) Confirmation of the FDR for real (top) and imaginary part (bottom) as functions of frequency for the left side (response properties) and the right side (spontaneous activity) of Eq. (7) for \( \mu = 0.8, D = 0.1 \) and a broadband stimulus (uniform power for \( |\omega| < 2\pi \cdot 100 \)) and a small variance of \( \langle s^2(t) \rangle = 0.1 \). For both sets of simulations, a time step of \( \Delta t = 10^{-4} \) and a time window of \( T \approx 100 \) were used (for real neurons with \( \tau_m = 10 \ms \), this would translate into a time window of 1s). (b) Cross spectrum between spike train and subthreshold voltage according to the analytical solution, Eq. (8) (solid lines), and from simulations (symbols).

\[
\frac{dv}{dt} = f(v) - a + \eta(t) - (v_T - v_R) x(t) + s(t),
\]

\[
\frac{da}{dt} = \Delta_a x(t).
\]

Additionally, the white noise \( \xi(t) \) has been replaced by a temporally correlated (colored) Gaussian noise \( \eta(t) \) with a nonflat power spectrum \( S_{\eta\eta}(\omega) \). We note that the parameter \( v_i < v_T \) sets a kind of soft threshold, but we still keep a hard threshold at \( v_T \) and a corresponding reset rule (this has been already incorporated above). The variable \( a(t) \) acts as an inhibitory current that pushes the voltage away from threshold. It evolves according to the slow dynamics given in the second equation (the ratio of its time constant to the membrane time constant is typically \( \tau_a \gg 1 \)) but every spike generated by the model kicks the adaptation variable up by the amount \( \Delta_a \), which implements the negative feedback that results in the spike-frequency adaptation seen in so many brain cells [33].

One can again use the same methods to derive equations for correlation functions and cross spectra and power spectra, namely, Fourier transformation, multiplication with \( \tilde{x} \), averaging, and once more invoking the Furutsu-Novikov theorem yield a fluctuation-dissipation relation for the adapting exponential integrate-and-fire model with colored noise:

\[
\chi_s = \frac{(v_T - v_R + \frac{\Delta_\eta}{1 + i\omega \tau}) S_{xx} + i \omega S_{xe} - S_{sf}(v)}{S_{\eta\eta}}. \tag{10}
\]

Again, on the left hand side we have exclusively the response to a weak stimulus, which can be determined by means of a periodic or a broadband stimulus \( s(t) \). On the right hand side, we find exclusively statistics of the spontaneous activity for \( s(t) \equiv 0 \). Several observations can be made: (i) the adaptation dynamics enters only by modifying the prefactor of the spike-train power spectrum in a frequency-specific manner, and (ii) instead of only the cross spectrum of the subthreshold membrane voltage and the spike train (which still appears on the rhs), we now also get the cross spectrum of the spike train \( x(t) \) and the subthreshold nonlinearity \( f[v(t)] \), which in some situations can be extracted from experiments [31].

As examples I pick two cases in Fig. 2, which both confirm the FDR for the adapting neuron with colored noise. For both cases it is not known how to calculate analytically any of the statistics shown or used here.

In Fig. 2(a) I use a slow adaptation current and a low-pass filtered noise (its correlation time is 10 times the membrane time constant). The adaptation current leads to a high-pass shape of the susceptibility [34], and the cut-off frequency of the noise results in additional shoulders in real and imaginary parts of \( \chi(\omega) \); both features are in marked contrast to white-noise-driven IF models without adaptation [15]. Remarkably, the numerical fluctuations of the susceptibility determined by stimulation or from the spontaneous statistics via Eq. (10) behave very differently: at small and up to intermediate frequencies (\( \omega < 1 \)) the rhs of Eq. (10) provides the more reliable estimate of the susceptibility while in the high-frequency limit it is the other way around.

In Fig. 2(b) the intrinsic noise process is more complex: two independent low-pass-filtered noise processes with distinct cut-off frequencies (noise from two populations of ion channels with different kinetics; see, e.g., Ref. [35]) and a narrow band noise, a harmonic noise [36] (stochastic network oscillations; see, e.g., Ref. [37]). In the susceptibility, one still sees adaptation-mediated high-pass filtering, shoulders at the cut-off frequencies of the low-pass filtered noise, and a dip in the real part at the frequency of the narrow band noise—complex effects that deserve closer investigation in future studies. In any case, Eq. (10) is confirmed again, and the fluctuation statistics reflects the rich statistical features of the susceptibility.
In the test of Eq. (10), I had to assume knowledge of the noise spectrum $S_{m}(\omega)$. In the more typical case, in which $S_{m}(\omega)$ is not known, but one can determine the response function and the spontaneous spectra independently, Eq. (10) can be used to infer the otherwise inaccessible noise spectrum $S_{m}(\omega)$. The result gives a satisfying estimate of the true noise spectrum both for the case of a simple Lorentzian, i.e., a low-pass filtered noise [cf. lowest panel in Fig. 2(a)] and for the more complex case of three distinct contributions [two Lorentzians and a narrow band spectrum, cf. lowest panel in Fig. 2(b)].

**Conclusions.**—In this Letter the relation between spontaneous fluctuations and the response to external perturbations have been worked out for an important class of spiking neuron models of the IF type. Two applications of the FDRs have been demonstrated: The analytical derivation of the cross spectrum between spike train and subthreshold voltage for the LIF model with white noise and the determination of the (in experimental situations often unknown) colored-noise spectrum from the response and the spontaneous activity.

However, there is still work to be done: With the same approach based on Eq. (4), it is easily possible to derive relations for the susceptibility of the subthreshold membrane voltage and its spontaneous power spectrum. The simple method introduced here can be applied to all types of neuron models with spike-associated reset such as the two-dimensional Izhikevich model [38] or the generalized IF model [39] if the voltage dynamics is driven by Gaussian noise. Moreover, it can even be extended to recurrent networks of IF models delay coupled by current synapses, e.g., to the celebrated Brunel model [40]; here one obtains a system of equations for susceptibilities, power and cross spectra of spike trains, and subthreshold voltage variables, which is currently under investigation.

I note that an entirely independent set of fluctuation-dissipation relations can be derived by the more common approach to nonequilibrium thermodynamic systems with a steady state going back to Agarwal [3,4] and discussed also more recently in the literature [5,9,10]. This approach will lead to relations in terms of the correlation function of the conjugated variable that is a highly nonlinear function of the membrane voltage steady state distribution. Finally, there is also work on fluctuation-dissipation relations for IF models in discretized time by Cessac et al. [25], the relation of which to the results derived here has to be clarified.

The division into subthreshold voltage and spike train used in the derived relations may appear as an artifact of the integrate-and-fire framework. However, it is meaningful because the spike train represents the important signal that is communicated to other neurons. In this context, it will also be useful to generalize the analysis to conductance-based neuron models of the Hodgkin-Huxley type. As the above mentioned multidimensional IF models approximate conductance-based models in many situations surprisingly well, I expect that the FDRs found here may also hold true (at least approximately) for these more detailed models of neural firing.

Finally, I mention constraints due to the FDR that are worth further investigation. First of all, in the context of neural information transmission, I note that both susceptibility and spike-train power spectra appear in a frequency-resolved measure of signal transfer, the coherence function [41], via $C(\omega) \approx |\chi(\omega)|^2 S_{ss}(\omega)/S_{xx}(\omega)$ [where $S_{ss}(\omega)$ is...
the spectrum of the input signal); relations like Eqs. (7) and (10) between \( \chi(\omega) \) and the spontaneous spike-train spectrum \( S_{xx}(\omega) \) impose (yet to be explored) constraints on the neural encoding capabilities. Secondly, for neurons in recurrent networks and in the asynchronous state, the FDR can be simplified because in this situation due to the consistency of input and output statistics, the noise spectrum \( S_{nn}(\omega) \) and spike-train spectrum \( S_{xx}(\omega) \) become proportional to each other (see, e.g., Refs. [42–45]). Such constraints on the single-neuron susceptibility within a recurrent network may be useful to simplify stochastic mean-field theories that contain single-neuron susceptibilities and spontaneous power spectra as their essential building blocks (see, e.g., Refs. [40,46]).

I would like to thank Davide Bernardi, Jonas Ranft and Magnus Richardson for comments on an earlier version of this paper.

[17] The firing rate of the LIF model with white noise is given by

\[
\tau_0^{-1} = \sqrt{\pi} \int_{z_0}^{\infty} e^{z^2} \text{erfc}(z) dz,
\]

see e.g. L. M. Ricciardi, Diffusion Processes and Related Topics on Biology (Springer-Verlag, Berlin, 1977).
[24] In the standard approach, the response to a step perturbation is expressed by the linear response on the one hand and by a Taylor expansion of the steady state under constant perturbation on the other hand.