Supporting Information

Nesse et al. 10.1073/pnas.1008587107

SI Text

SI Methods. Simulated spike trains were generated from a modified Morris–Lecar (ML) model of an excitable neural membrane (1, 2) driven by a fluctuating input current. The voltage V is governed by the current balance relation between the capacitive current (left side) and the negative membrane ionic currents (right side):

$$C\frac{dV}{dt} = g_c m_\infty(V)(V_c - V) + g_k W(V_k - V) + g_l(V_l - V) + g_H H(V_H - V) + I_m + g_x x(t). \quad \text{[S1]}$$

The first three terms are the familiar spike-initiating sodium membrane current, the spike-terminating potassium current, and the passive leak current, respectively, where W is the proportion of conducting potassium channels. The fourth current is a slower-to-activate and slower-to-decay voltage-gated adaptation current (AC). In keeping with its name, the voltage depolarization of each spike activates this current and produces a postspike hyperpolarization. The dynamics of H and W are similar in form and are determined by the equation

$$\tau_X(V)\frac{dX}{dt} = X_\infty(V) - X, \quad \text{[S2]}$$

in which X represents W or H. The function $X_\infty(V)$ is a positive-sloped sigmoidal shaped function: A spike (V high) thus implies that X_∞ goes into an active phase ($X_\infty \sim 1$). The activation functions are of the form $W_\infty(V) = 1/2(1 + \tanh[(V - \gamma_W)/\zeta_W])$ [and analogously for $m_\infty(V)$], and $H_\infty(V) = 1/(1 + \exp[-\alpha_H(V - \beta_H)])$. For W, the activation leads to a quick termination of spiking on a fast time scale determined by $\tau_W(V) = 5\cosh[(V - \gamma_Z)/(2\zeta_W)]^{-1}$, where the factor 5 ms sets the repolarization time scale at the peak of spiking. The activation and decay of H is slower than W. The H time constant is $\tau_H(V) = \tau + (\tau^+ - \tau)/(1 + \exp[-\alpha_H(V - \beta_H)])$, where $\tau^+ = 50/3$ ms is the time scale of the activation phase when $V \gg \beta_H$ and $\tau = 400$ ms is the decay time scale in the nonspiking phase $V \ll \beta_H$, with $\beta_H = 12$ mV and $\alpha_H = 20$ mV^{-1}. Constant current input is given by $I_m = 38$ nA/cm^2, and noisy fluctuating current, representing the input stimulus, is given by $x(t)$ with dynamics

$$\tau_x\frac{dx}{dt} = -x + \sigma_x\xi(t), \quad \text{[S3]}$$

where $\xi(t)$ is uncorrelated Gaussian white noise: $\langle\xi(t)\rangle = 0$ and $\langle\xi(t)\xi(t')\rangle = \delta_{t,t'}$. A time scale of $\tau_x = 5$ ms was chosen to induce input correlations on the order of fast glutamate AMPA receptors, and $\sigma_x = 0.3$ sets the noise strength. The model and parameter choices are similar to those in Rinzel and Ermentrout (2), not counting the slow process $H(t)$ that we have introduced here. The remaining parameters are $C = 20$, $g_c = 4.4$, $\gamma_m = -1.2$, $\zeta_m = 18$, $V_c = 120$, $g_k = 8$, $\gamma_W = 12$, $\zeta_W = 17.4$, $V_k = -84$, $g_l = 2$, $V_l = -60$, and $g_x = 45$.

The parameters of the ML model were selected to generate standard neuronal dynamics, in which spike generation is determined through a saddle node bifurcation [type 1 excitability (2)]; however, our results are applicable to type 2 excitability as well. The constant injected current $I_m = 38$ was selected to set the membrane potential just slightly below the deterministic bifurcation threshold, so that fluctuations in $x(t)$ trigger spiking. The fluctuating amplitude current strength, set by $\sigma_x = 0.3$ and $g_x = 45$, generates robust spiking, but the induced voltage perturbations are much smaller than the spike amplitudes and do not obscure the spike kinetics appreciably, and thus the spike shape from spike to spike is very uniform. In the next two sections we explore the modulation of the injected current I_m.

Approximation of $H(t)$ and Derivation of $Q(h|h')$. We approximate $H(t)$ with a nonsmooth but continuous curve, cleaved into two smooth phases that form the interspike interval (ISI) cycle Δt, starting with the decaying phase of $H(t)$, followed by the activation phase occurring during the action potential: $\Delta t = \Delta_{\text{decay}} + \Delta_{\text{spike}}$. Suppose that Δ_{decay} begins at $t = 0$, in which the peak adaptation value is $H(0) \equiv h_i$. The decay phase dynamics are $H(t) \approx h_i e^{-t/\tau}$. At the beginning of the activating phase ($t = \Delta_{\text{decay}}$) the AC is $H(\Delta_{\text{decay}}) \approx h_i e^{-\Delta_{\text{decay}}/\tau}$. Then, the AC exponentially approaches unity with time scale τ^+: $H(t) \approx e^{-(t-\Delta_{\text{decay}})/\tau^+}[H(\Delta_{\text{decay}}) - 1] + 1$ for $\Delta_{\text{decay}} \le t \le \Delta t_i$. If Δ_{spike} were very long, the activation would asymptote to unity and the current would be maximally activated. However, it is reasonable to assume Δ_{spike} is a very small portion of the ISI ($\Delta t_i \approx \Delta_{\text{decay}}$), and so we consider the activation phase to be effectively instantaneous. Note further that in *Fitting the ML Model* (below) we detail how to account for both the spike duration and absolute refractory period when the firing rate is so high that they can no longer be considered small relative to the ISI. Taken together, the two phases define a linear map from h_i to h_{i+1}:

$$h_{i+1} = \Lambda + (1 - \Lambda)h_i \exp(-\Delta t_i/\tau) \equiv f(\Delta t_i, h_i), \quad \text{[S4]}$$

which is Eq. **1** in the main text.

To assess the fit of [**S4**] to the ML model, we simulated for many spikes over a wide range of injected current levels ($I_m = 34$–80). Fig. S1 plots the preactivation of $H(t)$ (abscissa) and the resulting change in activation Δh postspike (ordinate). By using the approximating model [**S4**], the preactivation is predicted to be $H_{\text{pre}} = h_i e^{-\Delta t/\tau}$, and the difference between pre- and postactivation is $\Delta h = h_{i+1} - H_{\text{pre}} = \Lambda - \Lambda H_{\text{pre}}$. Hence, Eq. **S4** predicts a linear relationship with y intercept Λ and slope $-\Lambda$, which is confirmed in Fig. S1.

Here we derive the Markov transition probability $Q(h|h')$ and the associated limiting density $q_\infty(h)$. For a generic emission rate function $\lambda[H(t)]$, with $H(t) = h'e^{-t/\tau}$, the ISI density is derived from a hazard function formalism. Let the hazard function $P(\Delta t|h')$ be the proportion of an ensemble, conditioned on h', that has not fired a spike in time $\Delta t > 0$ from the time 0 where the systems in the ensemble originally spiked. If $\lambda(H)$ is the firing probability per unit time, then the hazard function is governed by the differential equation

$$\frac{dP}{d\Delta t}(\Delta t|h') = -\lambda[H(\Delta t)]P(\Delta t|h'), \quad \text{[S5]}$$

expressing that the reduction in P per unit time is the emission rate (λ) times the proportion (P). Eq. **S5** has the solution $P(\Delta t|h') = 1 - \exp[-\int_0^{\Delta t}\lambda(h'e^{-u/\tau})du]$. The ISI probability density then is the negative time derivative of P:

$$-\frac{dP}{d\Delta t} \equiv p(\Delta t|h') = \lambda(h'e^{-\Delta t/\tau})\exp\left(-\int_0^{\Delta t}\lambda(h'e^{-u/\tau})du\right). \quad \text{[S6]}$$

Using the inverse of the h map ([**S4**]), $\Delta t = f^{-1} = \tau\ln\{[h'(1 - \Lambda)]/(h - \Lambda)\}$, we perform a change of variable on [**S6**] from the time domain to achieve an expression solely in the h domain:

$$Q(h|h') \equiv -p(f^{-1}|h')\frac{\partial f^{-1}}{\partial h}(h)\Theta(f^{-1}), \qquad \textbf{[S7]}$$

where the partial derivative is the Jacobian of the change of variables. For the specific emission rate model

$$\lambda(H) = \alpha \exp(-\beta H) \qquad \textbf{[S8]}$$

used in the main text and this supplement, **[S7]** reduces to

$$Q(h|h') = \lambda\left(\frac{h-\Lambda}{1-\Lambda}\right)\frac{\tau}{h-\Lambda}e^{[-\tau\int_0^{f^{-1}(h,h')}\lambda(h'e^{-z/\tau})dz]}\Theta(f^{-1}). \qquad \textbf{[S9]}$$

We make the substitution $u = (1-\Lambda)e^{-z/\tau}$ in the integral of Eq. **S9** to get

$$Q(h|h') = \lambda\left(\frac{h-\Lambda}{1-\Lambda}\right)\frac{\tau}{h-\Lambda}e^{[-\tau\int_{h-\Lambda}^{h'(1-\Lambda)}\lambda(\frac{u}{1-\Lambda})\frac{1}{u}du]}\Theta(f^{-1}). \qquad \textbf{[S10]}$$

This substitution allows for the convenient description of $Q(h|h')$ in terms of exponential integral-type functions. We define the function L as

$$L(h,h') \equiv \tau\int_{h-\Lambda}^{h'(1-\Lambda)}\lambda\left(\frac{u}{1-\Lambda}\right)\frac{1}{u}du \qquad \textbf{[S11]}$$

so that

$$Q(h|h') = L'(h,h')e^{-L(h,h')}\Theta(f^{-1}), \qquad \textbf{[S12]}$$

where the L prime (L') indicates differentiation in h. The function L (**[S11]**) is in the exponential integral class of functions, diverging to infinity for $h \to \Lambda$ and limiting to zero as $h \to h'(1-\Lambda)+\Lambda$. Alternatively, in the adaptation independence regime the upper integral limit in **[S11]** can be replaced with unity and the factor $\Theta(f^{-1})$ can be dropped. In this case, the Markov transition reduces to

$$Q(h|h') = q_\infty(h) = L'(h)e^{-L(h)}. \qquad \textbf{[S13]}$$

Hence, in the adaptation independence regime, $q_\infty(h)$ is analytically computable from L. Outside of the adaptation independence regime, $q_\infty(h)$ can only be computed numerically. We defined a sufficiently refined partition of the h space and represented the Markov transition $Q(h|h')$ as a Markov matrix. The limiting density $q_\infty(h)$ then is the leading eigenvector associated with the unique unit eigenvalue of the Markov matrix, which we solved for using ARPACK routines.

Fitting the ML Model. We stated in the main text that the spike emission rate model $\lambda[H(t)] = \alpha\exp[-\beta H(t)]$ coupled with the h map (**[S4]**) provides a good fit to the behavior of the ML model [and other similar models (3, 4)] over a range of baseline excitability levels. Here we establish this fact. Fig. S2 plots results from Monte Carlo simulations of the ML model over a range of baseline injected current levels I_m, with parameters as detailed in *SI Methods* and shown in Figs. 1 and 2 of the main text. The main panel shows the h distributions from the ML model, along with the respective analytic approximations $q_\infty(h)$. The distribution of $I_m = 38$ of Fig. 2 of the main text is shown second from the left (green). As input current is increased, the $q_\infty(h)$ broadens and translates to higher activation states. This translation is concomitant with an increase in the mean firing rate (*Upper Left Inset*). Consistent with the results in Fig. 3 of the main text, the coefficient of variation (CV) and CV_Z (*Lower Left Inset*) exhibit local minima for distinct current levels where the single ISI and multiple ISI sequences, respectively, show points of maximal regularity and are fit acceptably by the analytical approximation (see below for a discussion of the quantitative discrepancies between the two models).

The autocorrelation coefficients ρ_k ($k = 1,2,3,4$) of the ISI sequence (Fig. S2, *Upper Right Inset*) exhibit the same qualitative behavior shown in Fig. 3 of the main text. The emergence of nonzero correlations in the h sequence indicates a loss of independence (*Lower Right Inset*), which occurs concomitantly with the point of minimum ISI correlation (ρ_1), the emergence of a significant nonzero secondary eigenvalue η_2 of the Markov transition $\mathbb{Q}$, and a minimum in the CV_Z. The confluence of these behaviors is precisely the same combination of phenomena exhibited by the loss of adaptation independence in the analytical model shown in Fig. 3 of the main text.

The divergence between the CV of the ML model and the analytical approximation is because of variability (noise) in the h map (**[S4]**) shown in Fig. S1 and concomitant variability in the repolarization current W. Additionally, for very high inputs with firing rates approaching 20 Hz, the ML model is near the maximal firing rate. In this regime a variety of conditions that allow the analytical model to approximate the ML system become weakened. The first condition is that the spike duration Δ_{spike} of the ML model is very small relative to the ISI. For the parameters that we have used in this article, Δ_{spike} is approximately 5 ms, and the absolute refractory period is approximately 15 ms. This nonvarying time interval of 20 ms comprises a large proportion of the overall 50 ms average ISI when firing at 20 Hz. Hence, it must be considered in the analytical model. Of course, by choosing faster repolarization kinetics closer to cortical-level spike dynamics, the spike duration and refractory period would be shorter and the assumption of an effectively instantaneous spike and refractory period would not be violated. However, here we demonstrate that the analytical rate model (**[S5]**–**[S10]**) can be modified slightly to achieve a good fit to the spike statistics even with the violated assumption and still be consistent with the results in the main text. To do so, we added in a constant $\Delta_{spike} = 5$ ms spike duration to each ISI and also included an absolute refractory period $\Delta_{refract} = 15$ ms to the $Q(h_{i+1}|h_i)$ transition for the results in Fig. S2, which is achieved by including $\Delta_{refract}$ in the Heaviside factor of Eq. **7** of the main text: $\Theta[f^{-1}(h_{i+1},h_i)] \to \Theta[f^{-1}(h_{i+1},h_i) - \Delta_{refract}]$. For low firing rates (0–10 Hz), the inclusion of these small time intervals insignificantly affect our main results but help the model fit better for higher firing rates.

Second, the analytical rate model assumes the input fluctuations that trigger spiking have a fast autocorrelation time scale relative to the mean ISI. In the ML model we have used input noise with a correlation time scale of 5 ms. In this regime, 5 ms is only marginally different from the ISI time scale. Consequently, input correlations could induce positive ISI correlations at these high firing rates, resulting in greater irregularity in the spike train (higher CV). Of course, such high firing rates, approaching the absolute refractory period bound, are toxic to neurons and are therefore considered nonphysiological and will not be studied further.

To get a quantitative fit over the range of input currents in Fig. S2, we defined a map from input current I_m to the input conductance level s used in the analytical model approximation: $\lambda(H-s)$. The input conductance is linearly related to the current by

$$s = \frac{1}{60}(I_m - 38), \qquad \textbf{[S14]}$$

in which the input level $I_m = 38$ used in the main text (Figs. 1 and 2) sets the zero point for the s input: $s = 0$. The divisive factor 60 was empirically determined to fit the firing rate and h distributions of the ML model over a range of input currents.

Adaptation independence can be understood by visualizing the Markov transition $Q(h_{i+1}|h_i)$ in both the independence and nonindependence regimes. Fig. S3 shows $Q(h_{i+1}|h_i)$ transition functions represented as 90×90 matrices for two of the input current levels used in the previous Fig. S2. Dark blue coloring represents values near zero; warmer tones represent positive values. Fig. S3*A* shows Q and its associated second eigenvalue η_2 (numerically computed) in the independence regime for $I_m = 38$, the same input level used in Figs. 1 and 2 of the main text. Each column of the matrix is effectively identical and proportional to $q_\infty(h_{i+1})$ (see Fig. S2 above at $I_m = 38$), which is the eigenvector associated with the leading unit eigenvalue $\eta_1 = 1$ of the matrix. The numerically computed second eigenvalue η_2 is effectively zero. All eigenvalues of Markov matrices are bounded by a unit modulus by Perron–Frobenius theory; hence, the magnitude of η_2 is meaningful because it measures the degree of independence by determining the rate of convergence to the limiting distribution q_∞—lower η_2 values imply faster convergence. The value η_2 defines the rate of contraction of the subspace orthogonal to q_∞ (in the standard L^2 inner product space). Because $\eta_2 \sim 0$ in Fig. S3*A*, the convergence to q_∞ occurs in effectively one spike.

In contrast, Fig. S3*B* shows Q for the input level $I = 72$. Clearly, the columns of the matrix are not all identical as they were in Fig. S3*A*, implying nonindependence. Consistent with this, the secondary spectrum is nonzero: $\eta_2 \approx 0.052$, which follows from the choice of a large baseline level (high firing rate). Hence, the contraction proceeds as $\eta_2^{(i)}$. For even higher input levels I_m, η_2 becomes even larger (see Fig. S2), and so the rate of convergence is slower. In Fig. S3*B* we have drawn the bounding line $h_{i+1} = \Lambda + (1-\Lambda)h_i$ (dashed black line), which represents the maximal h_{i+1} that is achievable for each h_i. In Fig. S3*A*, which has a lower input level, the bounding line would be drawn off the axes and so is not a real constraint on the dynamics. The bounding line illustrates how the loss of independence is realized. Low h_i values are restricted from mapping to high h_{i+1} values that are achievable for slightly larger h_i values, even with very short ISIs. Note also that the positive matrix values in Fig. S3*B* do not exactly border the bounding line. Instead, there is a small strip of zero values just below the bounding line. This strip exists because we are using the model parameters used in Fig. S2, which includes an explicit finite spike duration and an absolute refractory period, which disallows very short, near-zero ISIs that would allow h_{i+1} values arbitrarily close to the bounding line. These extra features add realism to the model but do not change the qualitative results appreciably.

Independence can be modulated by changes in noise intensity σ_x. Lower noise levels induce lower-variance q_∞ distributions. Zero noise is the trivial case where the neuron can only encode information in the firing rate and the h_i are constant and so are trivially independent (or a lack of any firing for low input currents I_m). Conversely, very high noise levels effectively overwhelm the AC, thereby blocking adaptation independence. To understand this point in more detail we refer to Nesse et al. (3), where it was noted that to properly fit the parameter β in the emission rate model $[\lambda(H) = \alpha \exp(-\beta H)]$ to the ML model, it must be inversely dependent on the noise strength (i.e., $\beta \propto 1/\sigma_x^2$). Conversely, it was also noted that α is proportional to σ_x^2. Hence, adjusting the noise level has multiple effects on spiking behavior. If we define $\alpha \equiv \sigma_x^2 \alpha_0$ and $\beta \equiv \beta_0/\sigma_x^2$ and rearrange $\lambda(H)$ to get

$$\lambda(H) = \exp\left(-\frac{\beta_0}{\sigma_x^2}[H - \sigma_x^2 \ln(\alpha_0 \sigma_x^2)/\beta_0]\right). \qquad \textbf{[S15]}$$

adaptation independence then holds when $\frac{\beta_0}{\sigma_x^2}\Lambda - \ln(\alpha_0 \tau \sigma_x^2) \gg 1$ (see Eq. **8** of the main text). Certainly, large enough values of σ_x will violate this condition. Conversely, if σ_x is very small, the condition will hold even for $H(t) \sim 0 \ll \Lambda$. That is, the adaptation activation h_i is trivially independent because the h_i are effectively constant (i.e., the cell fires regularly). For intermediate noise levels, the condition holds only for near-peak values of $H(t)$, that is, $H(t) \geq \Lambda$. In this zone, adaptation independence is nontrivial because the h_i values are variable. This zone is the main focus of this article because the h_i can encode information only if they vary.

ISI Correlations. As shown in Fig. S2, and Fig. 3 in the main text, the ISI correlation structure depends critically on adaptation independence and input level. The serial ISI covariance can be computed by using $q_\infty(h_i)$. The serial covariance between two ISIs that are k ISIs apart is

$$\mathrm{Cov}(\Delta t_i, \Delta t_{i+k}) = \langle \Delta t_i \Delta t_{k+1} \rangle - \langle \Delta t_i \rangle^2. \qquad \textbf{[S16]}$$

By inserting $f^{-1}(h_{i+1}, h_i) = \Delta t_i$ into this ISI covariance formula we get

$$\begin{aligned}\mathrm{Cov}(\Delta t_i, \Delta t_{i+k}) = \tau^2 \langle \{\ln[h_i(1-\Lambda)] - \ln(h_{i+1} - \Lambda)\} \\ \times \{\ln[h_{i+k}(1-\Lambda)] - \ln(h_{i+k+1} - \Lambda)\}\rangle \\ - \tau^2 \langle \{\ln[h_i(1-\Lambda)] - \ln(h_{i+1} - \Lambda)\}\rangle^2. \qquad \textbf{[S17]}\end{aligned}$$

To simplify the equation, we define the functions $R_k \equiv \ln[h_{i+k}(1-\Lambda)]$ and $U_k \equiv \ln(h_{i+k} - \Lambda)$. Expanding **[S17]** results in seven distinct terms:

$$\begin{aligned}\mathrm{Cov}(\Delta t_i, \Delta t_{i+k})/\tau^2 &= \langle R_0 R_k \rangle - \langle R_0 U_{k+1} \rangle - \langle R_k U_1 \rangle + \langle U_1 U_{k+1} \rangle \\ &\quad - \langle R_0 \rangle^2 + 2\langle R_0 \rangle \langle U_1 \rangle - \langle U_1 \rangle^2 \\ &= \mathrm{Cov}(R_0, R_k) + \mathrm{Cov}(U_1, U_{k+1}) \\ &\quad - \mathrm{Cov}(R_k, U_1) - \mathrm{Cov}(R_0, U_{k+1}). \qquad \textbf{[S18]}\end{aligned}$$

If the h_i are nonindependent, then there is no simplification of **[S18]**. However, recall that if two random variables a and b are independent, then $\langle ab \rangle = \langle a \rangle \langle b \rangle$. Therefore, if the h_i sequence is independent, then there are simplifications. If so, then for $k > 1$, all terms cancel, such that $\mathrm{Cov}(\Delta t_i, \Delta t_{i+k}) = 0$. If $k = 1$, then **[S18]** reduces to two terms:

$$\mathrm{Cov}(\Delta t_i, \Delta t_{i+1})/\tau^2 = -\langle R_1 U_1 \rangle + \langle R_0 \rangle \langle U_1 \rangle. \qquad \textbf{[S19]}$$

The equation above (**[S19]**) is the equality in Eq. **9** in the main text. Second, note that R_0 can be replaced with R_1 in **[S19]** with no consequence, and so we drop the subscripts for simplicity. We can then write

$$\begin{aligned}\mathrm{Cov}(\Delta t_i, \Delta t_{i+1})/\tau^2 &= -\mathrm{Cov}(R, U) \qquad \textbf{[S20]} \\ &= -\int q_\infty (R - \langle R \rangle)(U - \langle U \rangle) dh. \qquad \textbf{[S21]}\end{aligned}$$

Now note that R_k and U_k are both strictly monotonic increasing functions of h_{i+k}. Hence, by Chebyshev's algebraic inequality, also known as the covariance inequality, we have $\mathrm{Cov}(R, U) \geq 0$, implying that the covariance between successive ISIs is nonpositive in the adaptation independence regime, as stated in Eq. **9** of the main text.

Mutual Information. To compute I_{AC}, a precise definition of the input stimulus is required. For our model $\lambda(H) = \alpha e^{-\beta H}$ we define the input variable $y \in [0,1]$, uniformly distributed, as a probability of spiking in a small time interval δt (5, 6): If $y \leq \lambda(H)\delta t$, it triggers a spike; otherwise, not. This input approximates noisy input currents $x(t)$ with short autocorrelation time scales relative to the mean ISI and δt. In the case of the ML Monte Carlo model,

the input stimulus was a fast fluctuating current $x(t)$, defined as Ornstein–Uhlenbeck (OU) noise ([**S3**]). Without knowledge of the initial condition, $x(t)$ is Gaussian distributed with standard deviation σ_x. We start by partitioning the time domain into bins of δt width. To connect y to the input current $x(t)$, we must assume there exists a transformation $\mathcal{T}$: $[x(t),\delta t] \to y$ for each time bin. We assume that the time scale of OU correlation is short relative to time discretization: $\tau_x \ll \delta t$. Hence, for a given bin interval δt, the values of $x(t)$ will be effectively independent of those in a subsequent time bin. If we take a particular $x(t)$ (e.g., the midpoint value of the bin or the average over the bin) to be representative of the interval, then

$$y = \mathcal{T}[x(t)] \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x(t)}{\sigma_x}} e^{-u^2} du, \qquad \textbf{[S22]}$$

where σ_x is the standard deviation of x. Each y is uniformly distributed on the unit interval and independent of subsequent y values in other bin intervals, and, thus, the stochastic process that generates the y values is independent and identically distributed. The variable y in this definition then represents the effect of $x(t)$ on spiking probability over a time interval δt.

We partition the y space $[0,1]$ into n intervals of width $1/n$. The input entropy for the time interval δt is $\mathcal{H}(y) = \ln(n)$, measured in "nats" (7). If the cell does not spike, then the uncertainty about y is diminished because $y > \lambda(H)\delta t$, and so the conditioned entropy is reduced and is approximately $\ln\{n[1 - \lambda(H)\delta t]\}$. Conversely, if it does spike, then the entropy is also reduced to be approximately $\ln[n\lambda(H)\delta t]$. Consider a single ISI that is determined by two h variables h_1 and h_2. The ISI $\Delta t = f^{-1}(h_2,h_1)$ then is divided into M intervals of width δt, so that $M\delta t \approx \Delta t$. The reduction in entropy of x that occurs with knowledge of h_1 and h_2 is

$$I[x;(h_1,h_2)] \cong \qquad \textbf{[S23]}$$

$$I[y;(h_1,h_2)] = \mathcal{H}(y) - \mathcal{H}[y|(h_1,h_2)] \qquad \textbf{[S24]}$$

$$= M \ln(n) \qquad \textbf{[S25]}$$

$$- \sum_{i=0}^{M-1} \ln\{n[1 - \lambda(h_1 e^{-i\delta t/\tau})\delta t]\} \qquad \textbf{[S26]}$$

$$- \ln[n\lambda(h_1 e^{-\delta t M/\tau})\delta t]. \qquad \textbf{[S27]}$$

The first term ([**S25**]) is the input stimulus entropy $[\mathcal{H}(y)]$ over the ISI and subsequent spike, and the remaining terms ([**S26**] and [**S27**]) form the conditioned entropy ($\mathcal{H}[y|(h_1,h_2)]$) over the ISI. The term in [**S27**] represents the uncertainty of y at the last (Mth) time interval at the end of the ISI when the cell spikes. We specify n large enough, so that $\lambda(h_1 e^{-\Delta t/\tau})\delta t \gg 1/n$ for all likely h_1 values. The term in [**S25**] cancels with equal terms in [**S26**] and [**S27**], so n disappears from the equation. Furthermore, note that $\ln(1+\epsilon) \approx \epsilon$ for small ϵ so that the sum in [**S26**] can be approximated by an integral: $-\sum_{i=0}^{M-1} \ln[1 - \lambda(h_1 e^{-i\delta t/\tau})\delta t] \cong \sum_{i=0}^{M-1} \lambda(h_1 e^{-i\delta t/\tau})\delta t \cong \int_0^{\Delta t} \lambda(h_1 e^{-z/\tau}) dz$. With simplification, the terms in [**S25**]–[**S27**] reduce to

$$I[y;(h_1,h_2)] \cong -\ln[\lambda(h_1 e^{-\Delta t/\tau}) e^{-\int_0^{\Delta t} \lambda(h_1 e^{-z/\tau}) dz} \delta t] \qquad \textbf{[S28]}$$

$$= -\ln[p(\Delta t|h_1)\delta t], \qquad \textbf{[S29]}$$

where the approximation ($\cong$) in [**S28**] becomes more precise as δt becomes smaller.

The mutual information per spike I_{AC} between the stimulus and the AC states can be computed in two distinct ways: a temporal-based formulation in which the temporal domain is uniformly partitioned into fixed intervals δt and an AC-based formulation with a uniform δh partition. We start with the uniform-δt resolution. After substituting $\Delta t = f^{-1}(h_2,h_1)$ and then multiplying and dividing the log argument in [**S28**] by the Jacobian $\tau/(h_2 - \Lambda)$, we get

$$I[y;(h_1,h_2)] \cong -\ln[Q(h_2|h_1)] - \ln[(h_2 - \Lambda)\delta t/\tau]. \qquad \textbf{[S30]}$$

The argument of the logarithm of the second term in [**S30**], $\delta h(h_2) \equiv (h_2 - \Lambda)\delta t/\tau$, is required to achieve a uniform temporal resolution δt. Note that this uniform temporal resolution defines a variable h resolution. The mutual information $I_{\mathrm{AC}}(\delta t)$ is calculated by averaging [**S30**] over h_2 and h_1:

$$I_{\mathrm{AC}}(\delta t) \cong \mathcal{H}(h_2|h_1) - \int_{\Lambda}^{1} q_\infty(h_2) \ln[(h_2 - \Lambda)\delta t/\tau] dh_2, \qquad \textbf{[S31]}$$

where the first term is the conditional entropy $\mathcal{H}(h_2|h_1)$. If there is adaptation independence, then I_{AC} is not conditioned on h_1, so that $\mathcal{H}(h_2) = \mathcal{H}(h_2|h_1)$, in which case the independent h sequence possesses all of the stimulus information available in the temporally correlated spike train. This powerful result enables a simple coding of input fluctuations by the independent h sequence that does not require conditional probabilities to decode.

To get the δh-uniform information $I_{\mathrm{AC}}(\delta h)$, note that the second term in [**S31**] is the information associated with the change of variables from the uniform time resolution δt to the nonuniform h resolution. The nonuniform h resolution requires a very fine partition [small $\delta h(h_2)$] for low h_2 values in order to discriminate long ISIs uniformly well. It is perhaps unrealistic to expect a biological detector of the AC states to have a highly nonuniform resolution. Therefore, we consider the second term in [**S31**] to be unobtainable information in practice. To obtain a uniform h-resolution information, we replace the second term in [**S31**] with $-\ln(\delta h)$:

$$I_{\mathrm{AC}} \equiv I_{\mathrm{AC}}(\delta h) \cong \mathcal{H}(h_2|h_1) - \ln(\delta h), \qquad \textbf{[S32]}$$

which is equivalently the h-sequence entropy rate $\mathcal{H}(h_2|h_1)$ plus a constant. Eq. **S32** above is the mutual information that we refer to in Eq. **14** of the main text. Furthermore, this approximation ($\cong$) in [**S31**] and [**S32**] becomes more precise as δt and δh, respectively, become smaller (provided $\tau_x \ll \delta t$). Hence, we state [**S32**] as an equality in the main text. Note also, to compare [**S31**] and [**S32**] with other information (see main text), we have ignored the additive constant resolution terms $\ln(\delta t)$ and $\ln(\delta h)$ in favor of a nonambiguous universal reference point, as is customary (7).

Information Gain. The Kullback–Leibler (KL) divergence is defined as follows:

$$D_{\mathrm{KL}}(s,\Delta s) = \int_{\Lambda}^{1} q_\infty^{s+\Delta s}(h') \int_{\Lambda}^{1} Q^{s+\Delta s}(h|h') \ln\left(\frac{Q^{s+\Delta s}(h|h')}{Q^{s}(h|h')}\right) dh dh'. \qquad \textbf{[S33]}$$

By using the representation of $Q(h|h')$ defined in Eq. **S12** above, the divergence is calculated to be

$$D_{\mathrm{KL}}(s,\Delta s) = \beta\Delta s - (1 - e^{-\beta\Delta s}) \times \cdots \qquad \textbf{[S34]}$$

$$\int q_\infty^{s+\Delta s}(h') \int Q^{s+\Delta s}(h|h') L(h,h') dh dh'. \qquad \textbf{[S35]}$$

We find the integral factor in [**S35**] is equivalent to unity, shown as follows. For simplicity we drop the $s + \Delta s$ superscripts and note that $\Theta[f^{-1}(h,h')] > 0$ when $h \le \Lambda + (1 - \Lambda)h'$, thereby defining the inner integral upper integration limit:

$$\int_{\Lambda}^{1} q_{\infty}(h')\left(\int_{\Lambda}^{\Lambda+(1-\Lambda)h'} L'e^{-L}Ldh\right)dh' \quad \text{[S36]}$$

$$= \int_{\Lambda}^{1} q_{\infty}(h')\left([-e^{-L}L]_{\Lambda}^{\Lambda+(1-\Lambda)h'} + \int_{\Lambda}^{\Lambda+(1-\Lambda)h'} Qdh\right)dh' \quad \text{[S37]}$$

$$= \int_{\Lambda}^{1} q_{\infty}(h')dh' = 1, \quad \text{[S38]}$$

because $L(\Lambda,h') = \infty$ and $L[\Lambda + (1-\Lambda)h',h'] = 0$. Note that the divergence is independent of s, depending only on Δs:

$$D_{\mathrm{KL}}(s,\Delta s) = \beta\Delta s - (1 - e^{-\beta\Delta s}). \quad \text{[S39]}$$

Poisson Information Gain. In this section we detail how to compute the information gain $D_{\mathrm{KL}}(s,\Delta s)$ for a Poisson process. To make the Poisson process comparable to the AC model, we specify that the change in the Poisson firing rate is the same as the firing-rate change of the AC system for a baseline change Δs at a given s. Note that D_{KL} for a Poisson process of a given rate $\gamma(s)$ is

$$D_{\mathrm{KL}}^{\mathrm{Poisson}}(s,\Delta s) \equiv -\ln\left(\frac{\gamma(s)}{\gamma(s+\Delta s)}\right) - 1 + \frac{\gamma(s)}{\gamma(s+\Delta s)}. \quad \text{[S40]}$$

The proportional change in firing rate for the AC system is $\Delta f_A/f_A$. A Poisson process that undergoes the equivalent proportional firing-rate change can be described by the rate function $\gamma(s)$, where $\gamma(s+\Delta s) \equiv f_A$ and the difference in firing rate is $\Delta f_A \equiv \gamma(s+\Delta s) - \gamma(s)$. Hence, the proportional rate change is $\Delta f_A/f_A = 1 - \gamma(s)/\gamma(s+\Delta s)$. Plugging this into [**S40**], we get

$$D_{\mathrm{KL}}^{\mathrm{Poisson}}(s,\Delta s) = -\ln\left(1 - \frac{\Delta f_A}{f_A}\right) - \frac{\Delta f_A}{f_A}. \quad \text{[S41]}$$

A sufficiently small baseline change Δs will produce a small $\Delta f_A/f_A$, in which case we can Taylor expand [**S41**] to the first nonzero order so that

$$D_{\mathrm{KL}}^{\mathrm{Poisson}}(s,\Delta s) \approx \frac{1}{2}\left(\frac{\Delta f_A}{f_A}\right)^2. \quad \text{[S42]}$$

Dividing [**S43**] by $\Delta f_A/f_A$, we get the information gain per proportional change in firing rate:

$$D_{\mathrm{KL}}^{\mathrm{Poisson}}(s,\Delta s) f_A/\Delta f_A \approx \frac{1}{2}\left(\frac{\Delta f_A}{f_A}\right). \quad \text{[S43]}$$

Recall from *Information Gain* that AC systems have a constant information gain D_{KL} over s, so the factor $f_A/\Delta f_A$ accounts for all the variation as a function of s. Hence, the Poisson and AC-system information gain per proportional change in firing rate are inversely related to first order, as stated in the main text. Hence, as the AC information gain per proportional change in firing rate increases to the optimal input point shown in Fig. 3 of the main text, the Poisson process decreases over the same input range.

Notes to Numerical Computations. In this section we summarize briefly the numerical methods used to compute the results in all the figures of the main text.

Fig. 1. (*A* and *B*) Numerical simulations of the ML model with parameters are listed in *SI Methods*. The autocorrelation functions (*C* and *D*) were computed from 10,000-spike, free-running simulations.

Fig. 2. (*A*) From the same simulations as Fig. 1, we measured the $H(t)$ at the local minimum point just prior to spike [$H(t)$ preactivation] and the subsequent local maximum point of activated level h_i just after spiking. For visual clarity, we plotted only every tenth spike of the 10,000 spikes simulated (the same method was employed in Fig. S1). (*B*) From the 10,000 Monte Carlo-generated spikes, we created empirical probability distributions. The analytical approximation $q_{\infty}(h)$ was computed from Eq. **S14** above in *Approximation of H(t) and Derivation of Q(h|h′)* or, equivalently, Eq. **7** of the main text with the integral upper limit set to unity. (*C*) Conditional ISI distributions were empirically generated by simulating the ML model for 3,000 spikes for each value of h. Each spike was simulated with randomized input initial conditions for $x(t)$, but a fixed initial condition with $h = 0.25$, 0.30, and 0.35, respectively, for each example curve. The variable silent periods listed are approximate. The analytical conditioned ISI densities were computed as $p(\Delta t|h) = \lambda(he^{-\Delta t/\tau})\exp[-\int_0^{\Delta t}\lambda(he^{-u/\tau})du]$, for each h value. (*D*) Just as in *A* and *B*, the Monte Carlo-generated distribution was constructed from the 10,000 spikes. The unconditioned analytic density was computed from $p(\Delta t|h)$ and $q_{\infty}(h)$ shown in *B* and *C* and by using Eq. **5** of the main text.

Fig. 3. (*A*) We computed $q_{\infty}^{s}(h)$ by constructing the full $Q^{s}(h,h')$ functions and found the leading eigenvector associated with the unit eigenvalue (by using ARPACK routines), which is guaranteed by Frobenius–Perron theory to be the limiting density $q_{\infty}^{s}(h)$. Note that the analytical derivation of $q_{\infty}^{s}(h)$ in Eq. **S14** of *Approximation of H(t) and Derivation of Q(h|h′)* is identical to the leading eigenvector only in the adaptation independence regime. (*B*) The CV, CV_Z, ρ_1, ρ_2, and the firing rate were computed from Monte Carlo simulations of the analytic model. The second eigenvalue η_2 was computed from the matrices $Q^{s}(h,h')$ by using ARPACK routines. Note that the CV_Z was computed from moments of Z (Eq. **10** of the main text) and not from the analytic reduction of Eq. **13** of the main text, which is valid only in the independent regime. (*C*) For $s \geq -0.06$, both the mutual information of h and the renewal ISI process were computed from $q_{\infty}^{s}(h)$ and $p(\Delta t|h)$, as described in *A*. For input values $s < -0.06$ the computation of $q_{\infty}^{s}(h)$ becomes singular and is thus inaccurate. For these inputs we have instead constructed empirical h distributions from Monte Carlo simulations of the analytic model. Note that because the ISI distribution (Eq. **5** of the main text) approaches Poisson statistics for these low inputs we have omitted them because they would be redundant. The mutual information of the equivalent-rate Poisson process was computed from the empirically obtained firing rate in *B*. (*D*) The D_{KL} for the h process is computed to be a constant for fixed Δs and β (Eq. **16** of the main text and Eq. **S39** of *Information Gain*) and is divided by the proportional change in firing rate obtained from the firing rate in *B*. For the ISI renewal process, as in *C*, the lower input levels ($s < -0.06$) required computation of D_{KL} from distributions constructed from Monte Carlo simulations; however, for larger input levels $s \geq -0.06$, we used the analytically derived ISI distributions as in Eq. **5** of the main text.

1. Morris C, Lecar H (1981) Voltage oscillations in the barnacle giant muscle fiber. *Biophys J* 35:193–213.

2. Rinzel J, Ermentrout B (1998) *Methods of Neural Modelling*, eds C. Koch, I. Segev (Massachusetts Institute of Technology, Cambridge, MA), 2nd Ed.

3. Nesse W-H, Del Negro C-A, Bressloff P-C (2008) Oscillation regularity in noise-driven excitable systems with multi-time-scale adaptation. *Phys Rev Lett* 101:088101.
4. Muller E, Buesing L, Schemmel J, Meier K (2007) Spike-frequency adapting neural ensembles: Beyond mean adaptation and renewal theories. *Neural Comput* 19:2958–3010.
5. Strong S-P, Koberle R, de Ruyter van Steveninck R-R, Bialek W (1998) Entropy and information in neural spike trains. *Phys Rev Lett* 80:197–200.
6. Butts D-A, et al. (2007) Temporal precision in the neural code and the timescales of natural vision. *Nature* 449:92–95.
7. Cover T, Thomas J (1991) *Elements of Information Theory* (Wiley-Interscience, New York).

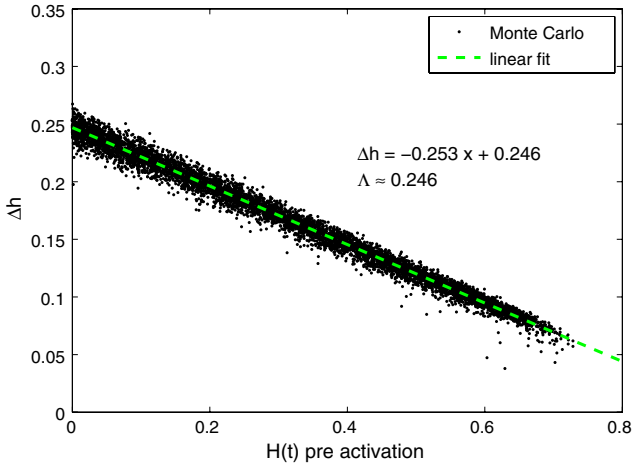

Fig. S1. Monte Carlo-generated data of the ML model ([**S1**]) plotting the preactivation $H(t)$ level (the level just prior to spiking) and the postspike activation change Δh over a large range of input currents $I_m = 34$–80 (see Fig. S2).

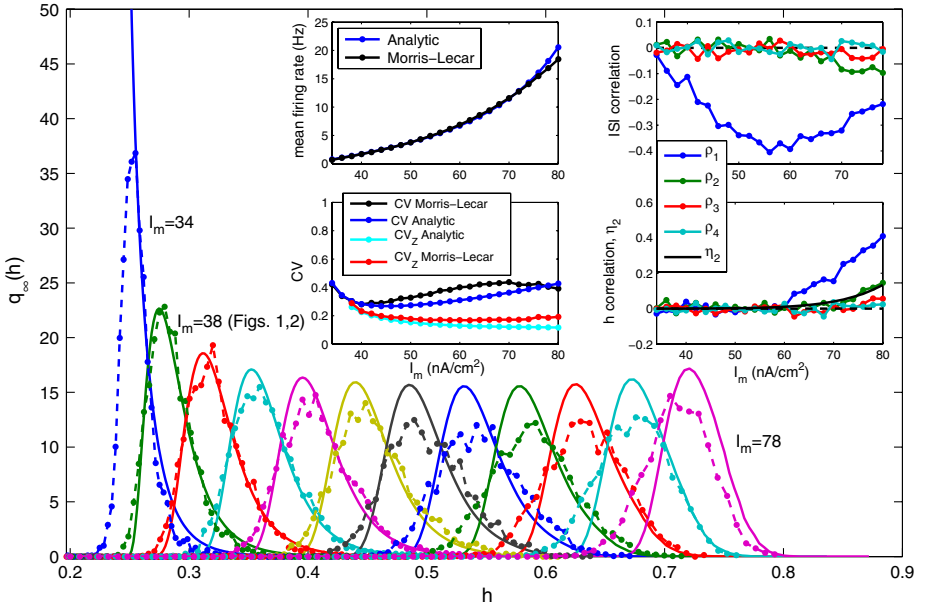

Fig. S2. Monte Carlo-generated h distributions (dot-dashed lines) from the ML model ([**S1**]) plotted with respective analytical approximations (q_∞, solid lines) over a range of injected currents $I_m = 34$–80 $(\mathrm{nA/cm^2})$. The mean firing rate (*Upper Left Inset*) and the CV and CV_Z (*Lower Left Inset*) are well fit by the analytical approximation. The autocorrelation coefficients ρ_k ($k = 1,2,3,4$) of the ISI sequence (*Upper Right Inset*) and autocorrelation of h sequence (*Lower Right Inset*) exhibit the same qualitative behavior shown in Fig. 3 of the main text. The minimum point of the first-order ISI correlation (ρ_1) coincides with the emergence of a significant nonzero secondary eigenvalue η_2 of the Markov operator Q (*Lower Right Inset*) and the minimum point of the CV_Z, which is also consistent with Fig. 3 of the main text.

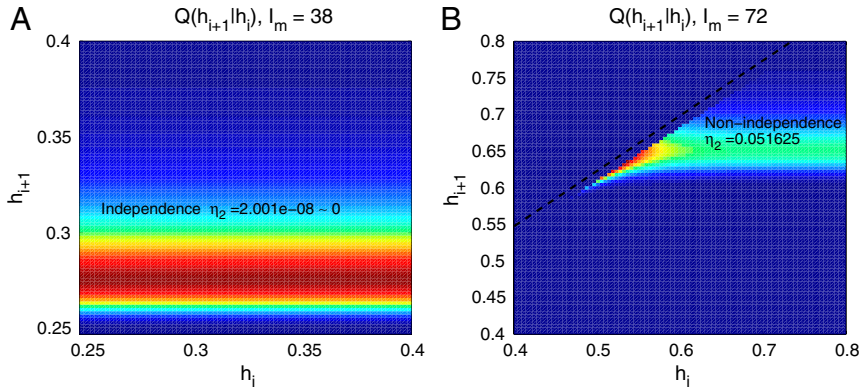

Fig. S3. Color renderings of the transition functions $Q(h_{i+1}|h_i)$, represented as 90 by 90 matrices, for an input level in the independent regime (*A*) and an input level in the nonindependent regime (*B*). Also listed is the secondary spectrum η_2 in both panels. The blue color indicates matrix values at or very near zero. Warmer colors indicate positive values. In the nonindependent regime in *B*, we have drawn the bounding line $h_{i+1} = \Lambda + (1 - \Lambda)h_i$ (dashed black line), which represents the maximal h_{i+1} that is achievable for each h_i. In *A*, which has a lower input level, the bounding line would be drawn off the axes.