Statistical structure of the trial-to-trial timing variability in synfire chains

Dina Obeid,[1, *] Jacob A. Zavatone-Veth,[2, †] and Cengiz Pehlevan[1, 3, ‡]

[1] *John A. Paulson School of Engineering and Applied Sciences,*
Harvard University, Cambridge, Massachusetts 02138, USA
[2] *Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*
[3] *Center for Brain Science, Harvard University, Cambridge, Massachusetts 02138, USA*
(Dated: March 21, 2020)

Timing and its variability are critical for behavior. Consequently, neural circuits that take part in the control of timing and in the measurement of temporal intervals have been the subject of much research. Here, we provide an analytical and computational account of the temporal variability in what is perhaps the simplest model of a timing circuit, the synfire chain. We elucidate the statistical structure of its trial-to-trial timing variability and show that it is well-described by a statistical model which consists of local, global, and jitter components. We relate each of these components to distinct biological mechanisms. A synfire chain architecture is thought to underlie the circuit that takes part in the control and timing of zebra finch song; we find that the structure of trial-to-trial timing variability in the synfire chain is consistent with experimental observations of the song's temporal variability. Our study therefore provides a neuronal account of behavioral variability in zebra finches.

I.   INTRODUCTION

Neural circuits are charged with behaviorally crucial timing tasks. The neural mechanisms underlying behavioral timing have been extensively studied experimentally [1–5], establishing links between temporal variations of neural activity and that of behavior [6–13]. Understanding timing variability in neural circuits is therefore necessary for understanding timing variability in behavior. In this paper, we study temporal variability in one of the most basic neural network models of timing, the synfire chain [14–16]. While the synfire chain is theoretically well-studied [15, 17–19], and experiments support its existence in biological systems [20], a theoretical account of its temporal variability is still lacking.

The synfire chain is a feedforward network composed of pools of neurons that produce traveling waves of synchronized spiking activity, as illustrated in Figure 1a. The synchronization of spikes within a pool, and the sequential propagation of activity across pools (Figure 1a) allow the synfire chain to serve multiple timing functions in a very natural way. First, it can be used to keep time by simply counting the pool which the spiking activity has reached. Second, it can be used to produce precisely timed intervals defined by the time elapsed between when activity arrives at a given pool and when it arrives at a subsequent pool (Figure 1a).

The synfire chain can sustain activity indefinitely given sufficiently many pools [15, 17], or by arranging it in a circular topology such that the final pool connects back to the first one [17]. Experiments support the existence of a synfire chain architecture in the songbird premotor

* dinaobeid@seas.harvard.edu
† jzavatoneveth@g.harvard.edu
‡ cpehlevan@seas.harvard.edu

cortex [20] for millisecond-scale precise time-keeping of the birdsong, with total song durations of few hundreds of milliseconds [6, 21]. While our results are applicable to arbitrary time scales, our numerical examples will respect the experimental time scale.

We are interested in the so-called trial-to-trial variability in the timing of neural activity of a synfire chain. Such variability arises from the millisecond-scale tempo differences across multiple propagations of the spiking activity in the chain. Trial-to-trial variability may be actively regulated [22] or caused by the inherent noise and fatigue in the neural system.

While it might seem small, millisecond-scale neural variability has been experimentally shown to correlate with behavioral variability at the same timescale [9, 11, 13]. Therefore, our findings may have direct implications for behavioral timing. Indeed, we will show that the statistical structure of temporal variability in a synfire chain can explain some of the salient features of temporal variability in birdsong [21, 23, 24].

We address these questions first in a simplified and analytically solvable model of trial-to-trial variability in a chain of individual neurons (Section II). We derive analytical expressions that describe the magnitude and statistical structure of temporal variability, and verify our results with simulations. Next, we address temporal variability in a biologically plausible model of a synfire chain, which includes multiple neurons per pool, by numerical simulations (Section III). We show that our results from the analytically tractable model qualitatively carry over to this more complex model. We further study the dependence of the various components of variability on the number of neurons per pool of the chain. We show that the statistical structure and magnitude of variability in a synfire chain is consistent with that observed in the zebra finch song (Section IV). In zebra finches, experimental studies support the existence of a synfire chain structure in the premotor nucleus HVC [20], which takes
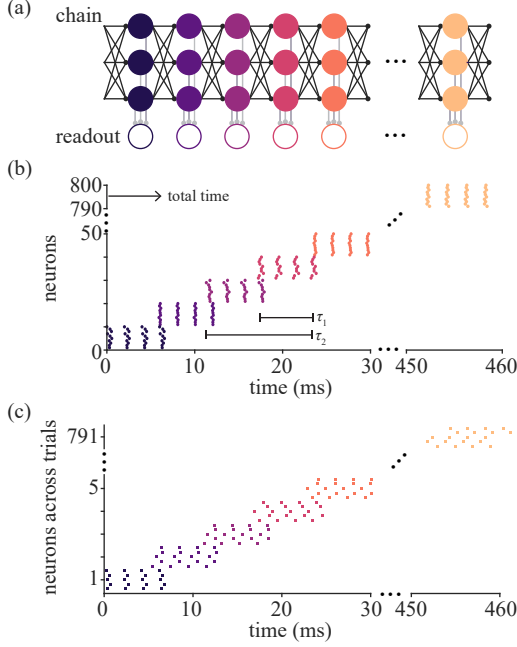
FIG. 1. Timekeeping and trial-to-trial variability in synfire chains. (a) The synfire chain is a feedforward network of pools of neurons. In the figure, circles represent neurons and arrows represent synapses. A simple scheme to measure time intervals composed of K successive pools, is to mark the beginning and end of the interval by the first spike of readout neurons corresponding to first and last pools. (b) Spike trains produced by the synfire chain show a synchronized activity that progresses pool-by-pool. Each point corresponds to the time of a spike produced by a neuron in the chain. Color denotes membership in a layer, each of which consists of a pool of 10 neurons. Various time intervals are shown. (c) Trial-to-trial variability in the synfire chain. Spike times of six neurons from different pools are plotted in different colors. Five different trials for each neuron are shown.

part in the production and timing of the birdsong.

II. TRIAL-TO-TRIAL TIMING VARIABILITY IN A CHAIN OF SINGLE NEURONS

In this section, we describe the statistical structure of trial-to-trial timing variability in an analytically tractable model: a chain of single leaky integrate-and-fire (IF) neurons. In this simple model, each neuron is driven by excitatory synaptic input from the previous neuron in the chain at time t_{ps}, which we model by $I_s\Theta(t - t_{ps})$, where $\Theta(t)$ is the Heaviside step function. Additionally, we model the drive to each neuron from outside the chain by the sum of a constant current I_0, and the noise due to synaptic transmission and other cellular processes by a zero-mean Gaussian process $\sqrt{\tau}\eta(t)$ with autocorrelation $\sigma^2\tau\delta(t - t')$, where τ is the membrane time constant and σ controls the standard deviation of the noise [25, 26]. The sub-threshold dynamics of the membrane potential

V of a given neuron in the chain is then governed by the Langevin equation [25, 26]

$$\tau\dot{V}(t) = -V(t) + I_0 + I_s\Theta(t - t_{ps}) + \sqrt{\tau}\eta(t). \quad (1)$$

When the neuron's membrane potential reaches a firing threshold V_{th}, the neuron produces a spike and resets its membrane potential to V_r.

A. Local variability in a chain of single neurons

We want to study the variability in the first-spike times of successive neurons in the chain. This problem differs from the standard treatment of noisy and leaky IF neurons [25, 26] in that we are interested in trial-to-trial variability of intervals between different neurons' spikes rather than long-time statistics of the intervals between spikes generated by a single neuron. However, we can map this problem to previous results in the literature [25, 26] by dividing it into two threshold-crossing problems. First, we can determine the probability distribution of a given neuron's membrane potential at time t_{ps}, using the fact that it receives no synaptic input before the previous neuron's first spike. Then, given that its membrane potential at time t_{ps} is V_0 with probability $P(V_0)$, we can think of the trial-to-trial variability in that neuron's time to first-spike after t_{ps} as the variability in the inter-spike intervals of a single leaky IF neuron with $V_r = V_0$.

If we assume that t_{ps} is long enough such that $V(t_{ps})$ has equilibrated, the solution to the first problem is given by the stationary distribution of the membrane potential, which was found in [25] to be

$$P(V) = \Theta(V_{th} - V)\frac{2\nu\tau}{\sigma}\exp\left(-\frac{(V - I_0)^2}{\sigma^2}\right)$$
$$\times \int_{(V-I_0)/\sigma}^{(V_{th}-I_0)/\sigma} du\,\Theta\left(u - \frac{V_r - I_0}{\sigma}\right)e^{u^2}, \quad (2)$$

where the firing rate ν is given by

$$\nu = \frac{1}{\tau\sqrt{\pi}}\left[\int_{(V_r-I_0)/\sigma}^{(V_{th}-I_0)/\sigma} du\,e^{u^2}(1 + \mathrm{erf}(u))\right]^{-1}. \quad (3)$$

We will be interested in the limit of very low firing rates $V_{th} - I_0 \gg \sigma$, which leads to the standard approximations (see Appendix A 1 and [25]) of the firing rate as

$$\nu \approx \frac{V_{th} - I_0}{\sigma\tau\sqrt{\pi}}\exp\left(-\frac{(V_{th} - I_0)^2}{\sigma^2}\right), \quad (4)$$

and the membrane potential distribution as

$$P(V) \approx \frac{1}{\sqrt{\pi}\sigma}\exp\left(-\frac{(V - I_0)^2}{\sigma^2}\right), \quad (5)$$

which is the stationary limit of an Ornstein-Uhlenbeck process with boundaries set at infinity.

We can calculate the mean and variance of the first-spike-interval, T_{fs}, defined as the time elapsed from t_{ps} to the arrival of the first spike, using the mapping between our problem and that of determining the statistics of the inter-spike intervals of a single leaky IF neuron. Conditioned on $V(t_{ps}) = V_0$, these statistics are given by the standard expressions [25, 26]

$$\langle T_{fs} \rangle_{V_0} = \tau\sqrt{\pi} \int_{(V_0-I_0-I_s)/\sigma}^{(V_{th}-I_0-I_s)/\sigma} du\, e^{u^2}(1 + \mathrm{erf}(u)) \quad (6)$$

and

$$\langle \delta T_{fs}^2 \rangle_{V_0} = 2\pi\tau^2 \int_{(V_0-I_0-I_s)/\sigma}^{(V_{th}-I_0-I_s)/\sigma} dx\, e^{x^2}$$
$$\times \int_{-\infty}^{x} du\, e^{u^2}(1 + \mathrm{erf}(u))^2. \quad (7)$$

We can then combine these results to compute the mean and variance of the first-spike-interval across trials. If we approximate the distribution of initial membrane potentials by the stationary Ornstein-Uhlenbeck process limit (5) in the low firing rate regime $V_{th} - I_0 \gg \sigma$, we obtain the lowest-order asymptotic expansions

$$\frac{\langle T_{fs} \rangle}{\tau} \sim \log\left(\frac{I_s}{I_0 + I_s - V_{th}}\right) - \frac{\sigma^2}{4(I_0 + I_s - V_{th})^2} \quad (8)$$

and

$$\frac{\langle \delta T_{fs}^2 \rangle}{\tau^2} \sim \frac{\sigma^2}{2(I_0 + I_s - V_{th})^2} \quad (9)$$

in the limit of large synaptic input $I_s + I_0 - V_{th} \gg \sigma$ (a detailed derivation of these expressions is given in Appendix A 2). The scaling of this variability with I_s and σ for fixed I_0 and V_{th} is illustrated in Figure 2 (see Section II C). In Appendix A 3, we also derive asymptotics for $\langle T_{fs} \rangle$ and $\langle \delta T_{fs}^2 \rangle$ using the alternative approximation $P(V_0) \approx \delta(V_0 - I_0)$.

The above calculation gives the variability in the timing of one neuron with respect to the previous neuron. If one defines an interval, T, by a collection of K subsequent neurons, i.e. $T \equiv \sum_{a=j}^{j+K} T_{fs}^a$ for some j, variance simply adds since noise is assumed to be independent for each neuron.

B. Correlated variability in a chain of single neurons

Thus far, we have only considered sources of variability that are independent across neurons. However, in biological neural networks, there are many possible mechanisms that could introduce correlated variability, such as synaptic depression or correlated external input [26–29]. Here, we consider a simple model for correlated variability due to neural fatigue. In this model, the spiking
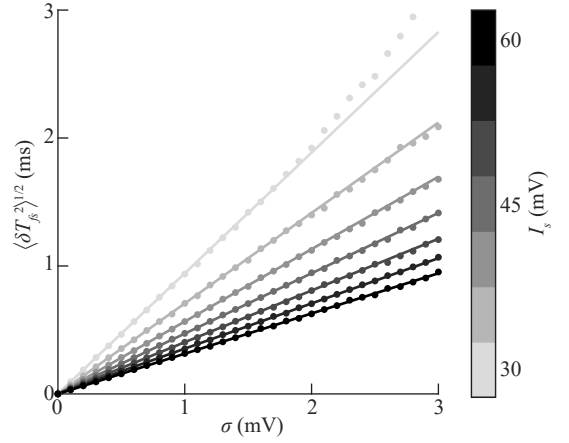

FIG. 2. Scaling of local trial-to-trial timing variability in a chain of single neurons. The results of numerical experiment (Section II C) are shown as dots, and the solid lines show the asymptotic approximations obtained in Section II A. The ordinate shows the standard deviation $\langle \delta T_{fs}^2 \rangle^{1/2}$ of the first-spike-interval, and the abscissa shows the standard deviation σ of the noise. Increasing values of I_s are indicated by darker shades of gray.

threshold in a given trial is increased by some multiple $m \in \{0, 1, \ldots, m_{\max}\}$ of a small increment δV_{th} relative to the baseline threshold V_{th}. We assume that, across trials, m is drawn from some distribution with mean $\langle m \rangle$ and variance $\langle \delta m^2 \rangle$. Working in the regime in which $m_{\max}\delta V_{th} \ll I_0 + I_s - V_{th}$, we can use our previously-obtained asymptotic expansions (8, 9) of the mean and variance of the first-spike-interval conditioned on V_{th} to obtain

$$\frac{\langle T_{fs} \rangle}{\tau} \sim \log\left(\frac{I_s}{I_0 + I_s - V_{th}}\right) + \frac{\delta V_{th}}{I_0 + I_s - V_{th}}\langle m \rangle, \quad (10)$$

and

$$\frac{\langle \delta T_{fs}^2 \rangle}{\tau^2} \sim \frac{\sigma^2}{2(I_0 + I_s - V_{th})^2}\left[1 + 2\frac{\delta V_{th}}{I_0 + I_s - V_{th}}\langle m \rangle\right]$$
$$+ \frac{\delta V_{th}^2}{(I_0 + I_s - V_{th})^2}\langle \delta m^2 \rangle, \quad (11)$$

to lowest order in both $\delta V_{th}/(I_0+I_s-V_{th})$ and $\sigma/(I_0+I_s-V_{th})$ (see Appendix A 2 for details). If we now consider two different neurons a and b, the trial-to-trial covariance of their first-spike-intervals T_{fs}^a and T_{fs}^b will be

$$\frac{\langle \delta T_{fs}^a \delta T_{fs}^b \rangle}{\tau^a \tau^b} \sim \frac{\delta V_{th}^a \delta V_{th}^b}{(I_0^a + I_s^a - V_{th}^a)(I_0^b + I_s^b - V_{th}^b)}\langle \delta m^2 \rangle, \quad (12)$$

to lowest order in $\delta V_{th}^a/(I_0^a + I_s^a - V_{th}^a)$ and $\delta V_{th}^b/(I_0^b + I_s^b - V_{th}^b)$.

Therefore, we obtain a model in which the trial-to-trial covariance matrix of the first-spike-intervals of neurons in the chain is the sum of an diagonal, local-to-a-neuron component and a rank-one global component. Concretely, if we assume for simplicity that
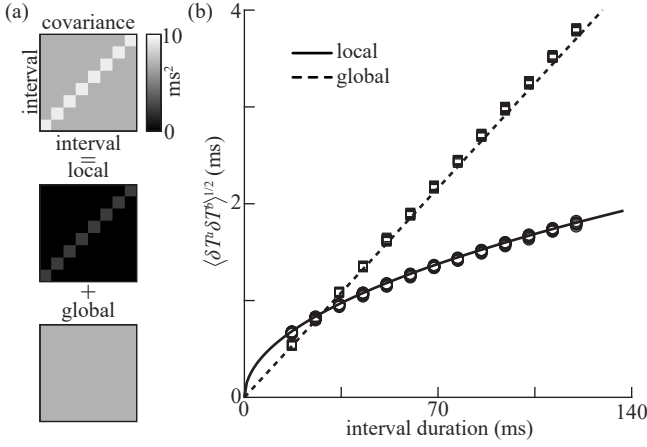
FIG. 3. Correlated trial-to-trial timing variability in a chain of single neurons. (a) Schematic representation of the decomposition of the interval-interval covariance matrix into local and global components as described in Section II B. (b) Scaling of the local and global variability with interval duration. Numerical experiment and asymptotic theory are shown by circles and a solid line for local variability, and squares and a dashed line for global variability (see Section II C). Realizations of the random sampling used to generate intervals are plotted as individual markers.

all neurons in the chain are identical, the covariance matrix of the first-spike-intervals of the neurons in the chain is given as $\boldsymbol{\Sigma}_{fs} = \sigma_I^2 \mathbf{I} + \sigma_G^2 \mathbf{1}\mathbf{1}^\top$, where $\sigma_I^2 = \tau^2 \frac{\sigma^2}{2(I_0 + I_s - V_{th})^2} \left[1 + 2\frac{\delta V_{th}}{I_0 + I_s - V_{th}}\langle m \rangle\right]$ is the local component of the first-spike-interval variance, $\sigma_G^2 = \tau^2 \frac{\delta V_{th}^2}{(I_0 + I_s - V_{th})^2}\langle \delta m^2 \rangle$ is the global component of the first-spike-interval variance, $\mathbf{I}$ is the identity matrix, and $\mathbf{1}$ is the ones vector.

If we define an interval by combining K subsequent neurons in the chain, i.e. $T_i \equiv \sum_{a=(i-1)K+1}^{iK} T_{fs}^a$, we can see that the resulting interval-interval covariance matrix $\boldsymbol{\Sigma}_T$ will have the same structure: $\boldsymbol{\Sigma}_T = K\sigma_I^2\mathbf{I} + K^2\sigma_G^2\mathbf{1}\mathbf{1}^\top$. We note that this scaling of local and global components of variability is independent of the details of the single-neuron model, and follows from the structure of the first-spike-interval covariance. This variance decomposition and scaling with interval length are illustrated in Figure 3.

C. Numerical simulation

We compare the theoretical asymptotics we obtained in Section II A and Section II B to the results of numerical experiment. To study the scaling of local variability with noise variance and synaptic strength as shown in Figure 2, we perform 10^4 realizations of a single-neuron simulation. In these experiments, we fix $\tau = 20$ ms, $I_0 = -70$ mV, and $V_{th} = -45$ mV while varying σ and I_s. We integrate the Langevin equation (1) using the Euler-

Maruyama method [30] augmented by the reset rule with a timestep of 10^{-3} ms; we find empirically that increasing or decreasing the timestep by factors of ten does not influence the qualitative results. For all parameter values tested, we observe good agreement between our asymptotic approximation and the experimental results for the mean first-spike-interval. As shown in Figure 2, for the lowest synaptic strengths and largest noise variances, we observe some discrepancy between asymptotic theory and experiment. This is unsurprising, since in that regime $I_s + I_0 - V_{th}$ is only around five times greater than the standard deviation of the noise, hence higher-order terms in the expansion are non-negligible.

To study the influence of introducing neural fatigue on the covariance between intervals defined by collections of neurons as described in Section II B, we simulate a chain of 80 identical neurons using the same methodology as described above. In these experiments, we fix $\sigma = 1$ mV and $I_s = 45$ mV. Over the 10^4 realizations performed, we draw the parameter m from the discrete uniform distribution on $\{0, \ldots, 249\}$, with the spiking threshold increment set in terms of the baseline threshold V_{th} as $\delta V_{th} = 10^{-3} V_{th}$. We then define intervals of varying lengths by grouping together uniformly randomly sampled sequences of neurons as described in Section II B.

To decompose the resulting interval-interval covariance matrices into local and global components, we use the factor analysis method described in [24]. In short, this method assumes a Gaussian generative model for interval durations, and decomposes the duration of an interval a (out of P) in a trial μ, T_μ^a as

$$T_\mu^a = T^a + l^a \epsilon_\mu^a + g^a \chi_\mu \\ + (1 - \delta_{aP}) j^a \xi_\mu^a - (1 - \delta_{a1}) j^{a-1}\xi_\mu^{a-1}, \quad (13)$$

where T^a is the average duration of the interval, $\epsilon_\mu \sim \mathcal{N}(0,1)$ and $\xi_\mu \sim \mathcal{N}(0,1)$ are random variables drawn independently for each interval and trial, and $\chi_\mu \sim \mathcal{N}(0,1)$ is drawn once per trial and shared across all intervals. l^a, g^μ and j^a are nonnegative parameters. This decomposition leads to a covariance matrix of the form

$$\Sigma_{ab} = \delta_{ab}\left(l^{a2} + (1 - \delta_{aP})j^{a2} + (1 - \delta_{a1})j^{a-1^2}\right) \\ + g^a g^b - (\delta_{a,b+1} + \delta_{a+1,b})j^{a2}. \quad (14)$$

From this covariance matrix, it is apparent that constants l^a can be associated with local variability and g^a with global variability. j^a describes another type of variability called jitter that we haven't encountered yet [24]. Jitter describes noise in determining interval boundaries and causes negative correlations between nearby intervals [24]. We fit this model to the intervals generated by our network using the expectation-maximization algorithm described in [24].

In Figure 3, we plot the square root of the local and global components of variability as a function of interval duration to more clearly illustrate their scaling with

the interval duration. We observe good qualitative agreement between the asymptotic theory presented in Section II B and the results of these numerical experiments.

III. TRIAL-TO-TRIAL TIMING VARIABILITY IN A SYNFIRE CHAIN

In Section II, we considered a chain of single neurons with simplified dynamics and coupling for the sake of analytical tractability. In this section, we study variability in a more realistic neural network, a synfire chain [14, 15, 31–33]. A synfire chain is a feed-forward network of multiple pools of neurons, also termed nodes or layers, which are linked by excitatory synaptic connections. We model the neurons in the synfire chain as bursting neurons, and add a set of readout neurons. This will later allow us to compare the statistical structure of trial-to-trial timing variability in this model with the structure of variability in a biological system, the zebra finch, in which the underlying neural circuit generating the behavior is believed to operate as a synfire chain (Section IV) [6, 20]. We note that the structure of variability does not change if we model neurons that emit a single spike rather than bursts.

We construct a chain of N pools, each composed of M identical neurons. Every neuron in a given pool i is connected to all neurons in the next pool $i + 1$ with equal weights. As before, we approximate the external input to each neuron $I(t)$ by the sum of an average part $g(t)$ and a zero-mean Gaussian process $\sqrt{\tau_m}\eta(t)$ with autocorrelation $\langle \eta(t)\eta(t') \rangle = \sigma_{\mathrm{neuron}}\delta(t - t')$. Neurons in the first pool receive an extra input $J(t)$, which is modeled as a rectangular pulse of height J_0 and width T_p. In addition to the per-neuron noise, we include another noise term, which we refer to as "shared" noise. This noise $\xi(t)$ is generated by a white Gaussian process but is shared across all neurons in a given pool, with mean $\langle \xi(t) \rangle = 0$ and autocorrelation $\langle \xi(t)\xi(t') \rangle = \sigma_{\mathrm{pool}}\delta(t-t')$. Our motivation for including this additional noise term will become clear later.

Then, the sub-threshold dynamics of the membrane potential of neuron j in pool i are given by

$$\tau_m \dot{V}_j^{(i)} = E_l - V_j^{(i)} + J_j^{(i)}(t) + \sqrt{\tau_m}\eta_j^{(i)}(t) + g_j^{(i)}(t) + \sqrt{\tau_m}\xi^{(i)}(t), \quad (15)$$

where $J_j^{(i)}$ is zero for all $i > 1$. The synaptic input $g_j^{(i)}(t)$ is modeled by the low-pass-filtered spike train [26]

$$\tau_s \dot{g}_j^{(i)} = -g_j^{(i)} + \frac{\tau_s I_s}{M} \sum_{k=1}^{M}\sum_{l=0}^{S-1} \delta(t - t_{ps,k}^{(i-1)} - \tau_b l), \quad (16)$$

where $t_{ps,k}^{(i-1)}$ denotes the first-spike time of the kth neuron of the $(i-1)^{\mathrm{th}}$ pool and I_s sets the strength of synaptic coupling. The neurons are modeled to emit a burst of S spikes separated by a fixed interval τ_b, rather than a single spike. When a neuron's membrane potential reaches the firing threshold V_{th}, it is then fixed at that threshold until the specified burst duration has elapsed, at which point it is reset to V_r, and once again evolves according to the sub-threshold dynamics (15). As in Section II B, we model neural fatigue as a small increase in the membrane potential threshold of all the neurons in the chain after each trial. We also considered a different neural mechanism for fatigue, a simplified model of synaptic depression. However, the structure of the resulting trial-to-trial timing variability was independent of which neural mechanism of fatigue we implement.

For each pool in the chain, we have a single readout neuron, which receives synaptic input from all neurons of that pool along with a white Gaussian noise input with mean zero and autocorrelation $\sigma_{\mathrm{readout}}\delta(t - t')$ (Figure 1a). The dynamics of membrane potential and synaptic inputs for the readout neurons are similar to that in (15) and (16) with the appropriate inputs.

We will study the statistical structure of timing variability in this synfire chain model using numerical simulations. As for the simple model (II C), we integrate the Langevin dynamics (15, 16) using the Euler-Maruyama method, with a timestep of 10^{-2} ms. Unless otherwise noted, we simulate a chain of $N = 81$ pools of $M = 32$ neurons each. We set the reset and resting potentials to $-70\,\mathrm{mV}$, the baseline spiking threshold to $-45\,\mathrm{mV}$, and the synaptic strength to $I_s = 45\,\mathrm{mV}$. We fix the membrane constant τ_m to $20\,\mathrm{ms}$, the synaptic time constant τ_s to $5\,\mathrm{ms}$, the number of bursts S to 4, and the spacing of bursts τ_b to $2\,\mathrm{ms}$. These values were chosen to respect the time scales of bursting in zebra finch HVC neurons [6, 20, 34]. Unless otherwise noted, we let the strengths of the per-neuron, per-pool, and readout noise be $0.5\,\mathrm{mV}$, $1\,\mathrm{mV}$, and $3\,\mathrm{mV}$, respectively. As in Section II C, we fix the spiking threshold increment δV_{th} to $10^{-3}V_{th}$, and draw the multiplicative increment factor from the discrete uniform distribution on $\{0, \ldots, 249\}$ for each trial. As propagation in a synfire chain is not always successful [15, 33]. We consider a trial to be successful if the total number of spikes in the chain is between $4N$ and $4N + 0.1\,(4N)$, and if all readout neurons fire once. In the experiments on which Figures 4, 5, and 6 were based, two trials out of 1000 were excluded from our simulations with all sources of noise. In Figure 7 no trials were excluded from the inset.

A. Relating neuronal mechanisms to different components of trial-to-trial timing variability

In Section II B, we showed that introducing correlated variability to a chain of IF neurons through a simple model of neural fatigue yields a spike interval covariance matrix that is the sum of a local component and a rank-one global component. To test whether this structure is present in the trial-to-trial timing variability of the
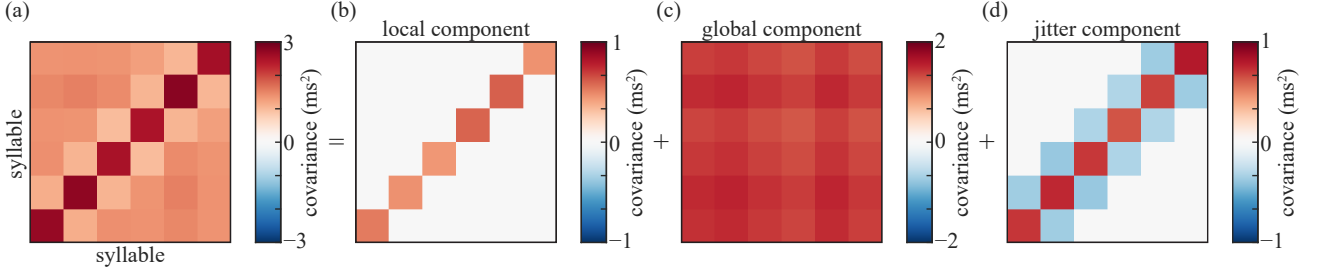
6

FIG. 4. Factor decomposition of the interval duration covariance matrix. (a) Full model covariance matrix. (b-d) The covariance matrices of the latent factors resulting from applying the analysis method of [24] to the full model shown in (a).
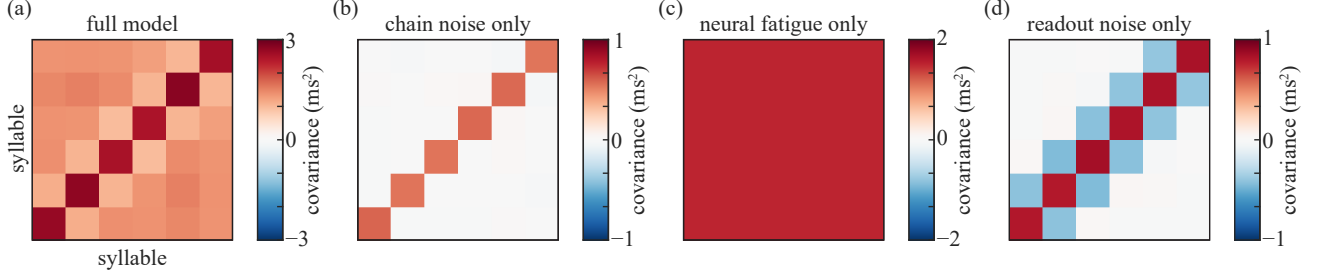

FIG. 5. Interval duration covariance matrices in the synfire chain model. (a). Interval duration covariance matrix for the full model. (b-d). Interval duration covariance matrices due to chain noise, neural fatigue, and readout noise alone, respectively.

synfire chain model, we define intervals by grouping together sequences of ten neurons, yielding intervals with a mean duration of 59.5 ms and the covariance structure shown in Figure 4a. Then, we take the full model covariance matrix and use the generative model proposed in [24] to decompose it into three components, a local component, a global component, and a jitter component, Figure 4. We found that the resulting decomposition explained the covariance matrix well, with a standardized root mean squared residual of 0.0067 [24]. Thus, the statistical structure of trial-to-trial timing variability in this synfire chain model is consistent with that of the simple model, with the addition of the component corresponding to readout noise.

Next, we delineate the neural mechanisms behind the components of variability by selectively including different sources of noise. First, we include only the chain noise, which comprises of the shared and the neuron-specific noise terms in the input to a neuron shown in Eq. (15). In this case, we recover a diagonal covariance matrix, which we identify as local variability Figure 5b. If we include only neural fatigue, we recover the global component of variability producing a nearly rank-1 covariance matrix, Figure 5c. If we only include noise in the readout neurons, the duration of neighboring intervals are anti-correlated (Figure 5d), corresponding to jitter. This exercise allows us to identify distinct cellular and synaptic mechanisms that explain distinct components of temporal variability: chain noise contributes to the local component, fatigue to the global component,

and readout noise to the jitter component.

B. Scaling of the components of variability with interval duration

In Section II B, we observed that, if one groups multiple pools together to form intervals, the local component of variability should scale with the square root of the interval length, while the global component should scale linearly with interval length. This scaling is independent of the details of the model, and simply follows from the assumption that the total trial-to-trial variability of the spike interval can be decomposed into a local component and a global component, both of which are uniform in magnitude across pools. Applying the factor analysis method introduced in [24] to the spike times produced by the detailed synfire chain model (Figure 4), we find that the scaling of local and global components of variability with interval length is consistent with this prediction (Figure 6).

C. Scaling of local variability with pool size

If we vary the number of neurons per pool M, we find that the interval duration variance due to per-neuron component of chain noise falls as $\frac{1}{M}$ (Figure 7). Thus, despite the fact that the system is nonlinear, such noise adds quasi-linearly. Therefore, to have a non-negligible
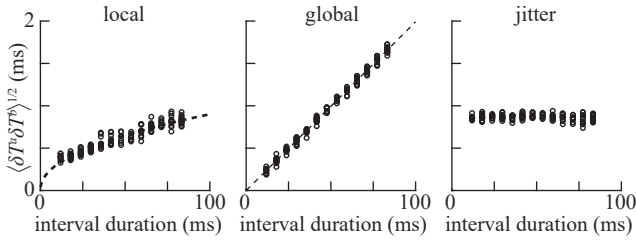
FIG. 6. Scaling of local, global, and jitter components of variability with interval duration in the detailed synfire chain model with 32 neurons per pool. Circular markers indicate the results of numerical simulations, and dashed lines show power-law fits to the data, with exponents of 0.46 and 1.00 for the local and global components, respectively. No fit is shown for the jitter component, as there does not exist a statistically significant Spearman correlation between it and the interval duration ($\rho = -0.09$, $p = 0.25$).

local component of variability in this model, we must assume that neurons belonging to the same pool receive shared noise. In Figure 7, we see that the interval variability due to this noise mechanism is roughly independent of the number of neurons per pool. Varying the number of neurons per pool did not have an effect on the readout noise or fatigue. This is shown in the inset of Figure 7.

IV. THE STRUCTURE OF VARIABILITY IN A SYNFIRE CHAIN IS CONSISTENT WITH THAT OBSERVED IN ZEBRA FINCH SONG

Glaze and Troyer [21, 23, 24] studied trial-to-trial variability in the highly stereotyped zebra finch song. Zebra finch songs consist of several introductory notes, followed by a few renditions of a motif, sung in a very repetitive manner. Motifs contain 3 to 8 syllables. Syllables range from 50-100 ms and are separated by gaps. The timing of the song is controlled by clock-like bursting in the premotor nucleus HVC, in particular in HVC neurons projecting to Robust Nucleus of the Arcopallium (RA). Many studies suggest that the underlying neural circuit behavior is consistent with a synfire chain model [6, 7, 20, 34, 35]. Further, experimental evidence supports millisecond scale correlations between HVC activity and the song [9, 11, 13]. Thus, we want to test if the trail-to-trail variability in a synfire chain is also consistent with the trial-to-trial variability observed in the song duration.

Our first observation is that the full model covariance matrix of interval duration in the synfire chain model (Fig. 5a and Section III) has the same structure and magnitude as song timing covariance reported by Glaze and Troyer [24]. These authors [24] showed that the generative model given in Eq. (13) is a good descriptor of the statistical structure of the birdsong temporal variability.
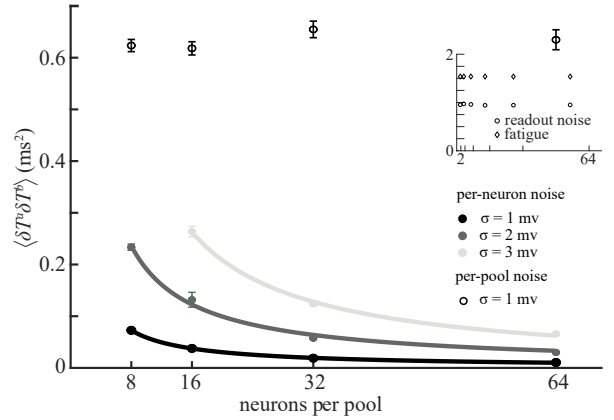

FIG. 7. Scaling of interval variability with the number of neurons per pool due to different noise sources. A per-neuron noise (filled circles) and shared noise among neurons in the same pool (open circles). The filled black, gray, and light gray circles show the interval variability due to per-neuron noise with $\sigma = 1\,\mathrm{mV}$, $2\,\mathrm{mV}$, and $3\,\mathrm{mV}$, respectively. Solid lines are power-law fits to the data, with exponents of -0.95 (black line), -0.94 (gray line) and -1.04 (light gray line). Open circles show the same thing but for noise that is shared among neurons in the same pool with $\sigma = 1\,\mathrm{mV}$. Error bars show the standard error of the mean. The data point for per neuron noise of $\sigma = 3\,\mathrm{mV}$ and $M = 8$ was excluded because the chain propagation failed in more than 10% of trials. Inset: Variability due to readout noise and fatigue as a function of neurons per pool.

We found the same thing for our synfire chain model in Section III A. The magnitudes of variability we report for a 100 ms interval is about 1 ms (for all components), again consistent with data, Fig. 8; data from [24]. Further, our results in Section III A allows us to connect behavioral variability to neural mechanisms. Our model suggests that the the chain noise contributes to the local component of the song variability, fatigue contributes to the global component, and readout noise contributes to the jitter component.

To further validate our model, we examine how the different components of variability in syllable duration scale with syllable duration in birdsong (Fig. 8; data from [24]). The scaling of the local and global components of variability of syllable duration is similar to that of the synfire chain model. In the experimental data, it was observed that the jitter component of syllable duration variability is not significantly correlated with interval duration (Figure 8); a similar lack of scaling is also observed in our model (Figure 6).

Finally, we present a prediction of the synfire chain model for the songbird. Previous studies estimated the number of RA-projecting HVC neurons in adult male zebra finches to be in the tens of thousands [36]. In turn, we can estimate the number of neurons per pool in the HVC synfire chain to be about on the order of hundreds of neu-
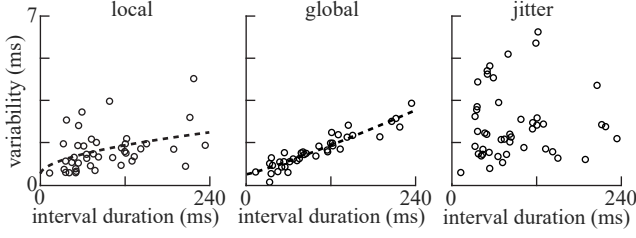
FIG. 8. Scaling of the local, global, and jitter components of syllable timing variability with interval duration in zebra finch song (data from Glaze & Troyer [24]). As in Figure 6, we fit the relationships between interval duration and the local and global components of variability with power laws, yielding exponents of 0.53 and 1.14, respectively. No fit is shown for the jitter component, as it is not significantly correlated with interval duration.

rons per pool, assuming a ∼5ms delay between activity reaching successive pools in the chain [6] and a song that is ∼500ms long. Our simulations (Section III C and Figure 7) show that in order to have a realistic level of local song variability with a pool size in the hundreds, neurons within the same pool should receive a shared noise input. A possible source of such noise can be the area LMAN, which projects to HVC indirectly [37]. LMAN is known to act as a source of variability in the song [8, 22], and lesioning it reduces temporal variability [8].

V. CONCLUSION

In this paper, we presented analytical and computational analyses of the trial-to-trial timing variability in synfire chains. We first show how variability scales with input strength and noise level in a simple, analytically tractable chain of single neurons. We also show how trial-to-trial variability scales with interval duration in this simple model. Then, we show that our main results carry to a more realistic synfire chain. We found that the statistical structure of timing variability in a synfire chain is well-described by a generative model which consists of local, global and jitter components. Remarkably, we were able to relate each of these components to distinct biological mechanisms.

Our findings have important implications for the relationship between neural and behavioral variability. Even the most stereotyped of animal behaviors are significantly variable [23], and it has been argued that some of this variability is rooted in neural activity [28, 29, 38]. If so, temporal variability of behavior should reflect the statistics and structure of temporal variability of neural circuits that represent or govern the behavior's timing [2, 5]. Indeed, interrelationships between the timings of neural dynamics and behavior have been observed in various experimental studies [6–13]. In this paper, we argued that the temporal variability observed in the synfire

chain is consistent with the magnitude and structure of the timing variability in the zebra finch song. Thus, our findings provide a concrete example of a detailed match between neural and behavioral variability. Furthermore, we predict that neurons within the same pool of the HVC synfire chain should receive a shared noise input.

A notion of tempo variation that we did not consider arises from structural changes to the chain, such as homeostatic and synaptic plasticity [8, 34, 35], or experimental perturbations [7]. In birdsong, these mechanisms can lead to tempo changes on the order of tens of milliseconds [7, 8, 34, 35], and, when naturally occurring, require thousands of song repetitions to take effect [8].

Appendix A: Asymptotics for the chain of integrate-and-fire neurons

1. Stationary membrane potential distribution in the low-rate limit

Here, we review the approximate firing rate and stationary membrane potential distribution in the low-rate limit $V_{th} - I_0 \gg \sigma$ [25]. Inspecting the equation

$$\nu = \frac{1}{\tau\sqrt{\pi}} \left[\int_{\frac{V_r - I_0}{\sigma}}^{\frac{V_{th} - I_0}{\sigma}} du\, e^{u^2}(1 + \mathrm{erf}(u)) \right]^{-1}, \quad \text{(A1)}$$

we can see that the integral is dominated by the upper limit due to the exponential. Making the change of variables $u' \equiv u/h$, $h \equiv (V_{th} - I_0)/\sigma$, $v \equiv (V_0 - I_0)/(V_{th} - I_0)$, we have

$$\begin{aligned}
\nu &= \frac{\sigma}{\tau\sqrt{\pi}(V_{th} - I_0)} \left[\int_v^1 du\, e^{h^2 u^2}(1 + \mathrm{erf}(hu)) \right]^{-1} \\
&\approx \frac{\sigma}{2\tau\sqrt{\pi}(V_{th} - I_0)} \left[\int_v^1 du\, e^{h^2 u^2} \right]^{-1} \\
&\approx \frac{\sigma}{2\tau\sqrt{\pi}(V_{th} - I_0)} \left[\left. \frac{e^{h^2 u^2}}{2uh^2} \right|_v^1 \right]^{-1} \\
&\approx \frac{V_{th} - I_0}{\sigma\tau\sqrt{\pi}} \exp\left(-\frac{(V_{th} - I_0)^2}{\sigma^2} \right). \quad \text{(A2)}
\end{aligned}$$

where we integrated by parts in the third line. By a similar argument, we can approximate the stationary distri-

bution of the membrane potential using the approximation

$$\int_{\frac{V-I_0}{\sigma}}^{\frac{V_{th}-I_0}{\sigma}} du\, \Theta\left(u - \frac{V_0 - I_0}{\sigma}\right) e^{u^2} \approx \frac{1}{2\nu\tau\sqrt{\pi}}, \quad (A3)$$

in this limit.

2. Moments of the first-spike-interval in the low-rate stationary limit

To derive the moments of the first-spike-interval in the low-rate stationary approximation, we start with the standard results (given as 6 and 7 in the main text) for the mean and variance conditioned on V_0 [25, 26]:

$$\langle T_{fs} \rangle_{V_0} = \sqrt{\pi}\tau \int_{(V_0 - I_0 - I_s)/\sigma}^{(V_{th} - I_0 - I_s)/\sigma} dy\, e^{y^2} [1 + \mathrm{erf}(y)] \quad (A4)$$

and

$$\langle \delta T_{fs}^2 \rangle_{V_0} = 2\pi\tau^2 \int_{(V_0 - I_0 - I_s)/\sigma}^{(V_{th} - I_0 - I_s)/\sigma} dx\, e^{x^2}$$
$$\times \int_{-\infty}^0 dy\, e^{(x+y)^2} [1 + \mathrm{erf}(x+y)]^2. \quad (A5)$$

Considering the mean first-spike-interval, we use the integral representation of the error function as

$$1 + \mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{-\infty}^0 du\, e^{-(u+x)^2} \quad (A6)$$

to write

$$\frac{\langle T_{fs} \rangle_{V_0}}{\tau} = 2 \int_{(V_0 - I_0 - I_s)/\sigma}^{(V_{th} - I_0 - I_s)/\sigma} dy\, e^{y^2} \int_{-\infty}^0 du\, e^{-(u+y)^2}$$
$$= \int_{-\infty}^0 \frac{du}{u} e^{-u^2}$$
$$\times \left[e^{2u(I_0 + I_s - V_0)/\sigma} - e^{2u(I_0 + I_s - V_{th})/\sigma} \right], \quad (A7)$$

where we note the cancellation in the bracketed integrand that ensures that it does not diverge as $u \to 0^-$. We can then easily compute the expectation over the approximate distribution (5) of V_0 to obtain

$$\frac{\langle T_{fs} \rangle}{\tau} = \int_{-\infty}^0 \frac{du}{u} e^{-u^2} \left[e^{u^2 + 2I_s u/\sigma} - e^{2u(I_0 + I_s - V_{th})/\sigma} \right], \quad (A8)$$

which, though it does not have a simple closed-form solution, is the integral of a bounded entire function that decays exponentially fast at infinity (provided that $I_s > 0$), and is therefore well-behaved.

To obtain a similar integral expression for the variance of the first-spike-interval, we recall the law of total variance

$$\langle \delta T_{fs}^2 \rangle = \langle \langle \delta T_{fs}^2 \rangle_{V_0} \rangle + \langle \delta \langle T_{fs} \rangle_{V_0}^2 \rangle, \quad (A9)$$

where the outer angle brackets denote averaging over the distribution of V_0, and follow the same procedure that we used to derive $\langle T_{fs} \rangle$ to obtain

$$\frac{\langle \langle \delta T_{fs}^2 \rangle_{V_0} \rangle}{\tau^2} = 4 \int_{-\infty}^0 du \int_{-\infty}^0 dv \int_{-\infty}^0 dw\, \frac{e^{-(u+v+w)^2 + 2uv}}{u+v+w}$$
$$\times \left[e^{(u+v+w)^2 + 2(u+v+w)I_s/\sigma} - e^{2(u+v+w)(I_0 + I_s - V_{th})/\sigma} \right], \quad (A10)$$

and

$$\frac{\langle \delta \langle T_{fs} \rangle_{V_0}^2 \rangle}{\tau^2} = \int_{-\infty}^0 \frac{du}{u} \int_{-\infty}^0 \frac{dv}{v} e^{-u^2 - v^2}$$
$$\times \left[e^{(u+v)^2 + 2I_s(u+v)/\sigma} - e^{u^2 + v^2 + 2I_s(u+v)/\sigma} \right]. \quad (A11)$$

With these integral expressions in hand, we can now derive asymptotic expansions for the moments. For brevity, we define the dimensionless scalars $\alpha \equiv (I_0 + I_s - V_{th})/\sigma$ and $\beta \equiv (V_{th} - I_0)/\sigma$; we will work in the limit of low baseline firing rates $\beta \gg 1$ and large synaptic inputs $\alpha \gg 1$. Rescaling u by 2α in (A8), we can write the mean first-spike-interval as

$$\frac{\langle T_{fs} \rangle}{\tau} = \int_{-\infty}^0 \frac{du}{u} e^u \left[e^{(\beta/\alpha)u} - e^{-u^2/4\alpha^2} \right]$$
$$= \log\left(\frac{\alpha + \beta}{\alpha}\right) - \int_{-\infty}^0 \frac{du}{u} e^u \left[e^{-u^2/4\alpha^2} - 1 \right], \quad (A12)$$

where we have split the integral into two pieces by adding and subtracting one from the integrand and evaluated the first of the remaining integrals. Expanding the remaining integrand other than the overall exponential weight e^u as a power series and integrating term-by-term using the relationship of the integrand to the gamma function [39], we obtain the divergent asymptotic series

$$\frac{\langle T_{fs} \rangle}{\tau} \sim \log\left(\frac{\alpha + \beta}{\alpha}\right) + \frac{(-1)^k (2k-1)!}{4^k k! \alpha^{2k}}$$
$$\sim \log\left(\frac{\alpha + \beta}{\alpha}\right) - \frac{1}{4\alpha^2} + \mathcal{O}(\alpha^{-4}), \quad (A13)$$

which yields the lowest-order approximation given in the main text.

We now consider $\langle \delta T_{fs}^2 \rangle$. Converting the integral over the negative octant in (A10) to an integral over the positive octant, making the change of variables $x \equiv u$, $y \equiv v + w$, $z \equiv w$, and parameterizing the domain of

integration such that we integrate first over $z \in [0, y]$, we have

$$\frac{\langle\langle\delta T_{fs}^2\rangle_{V_0}\rangle}{\tau^2} = 2 \int_0^\infty dx \int_0^\infty dy \, \frac{e^{-2\alpha(x+y)}}{x(x+y)} \left[e^{2xy} - 1\right]$$
$$\times \left[e^{-(x+y)^2} - e^{-2\beta(x+y)}\right]. \tag{A14}$$

Then, adding the expression for $\langle\delta\langle T_{fs}\rangle_{V_0}\rangle$ given in (A11) to the above expression for $\langle\langle\delta T_{fs}^2\rangle_{V_0}\rangle$ as prescribed by the law of total variance (A9), we have

$$\frac{\langle\delta T_{fs}^2\rangle}{\tau^2} = \int_0^\infty dx \int_0^\infty dy \, e^{-2\alpha(x+y)} \left[e^{2xy} - 1\right]$$
$$\times \left[\frac{2e^{-(x+y)^2}}{x(x+y)} + \frac{1}{xy}\left(\frac{x-y}{x+y}\right) e^{-2\beta(x+y)}\right]. \tag{A15}$$

As it is anti-symmetric about the line $y = x$, the second term in the bracketed integrand will vanish under integration over the positive quadrant, leaving

$$\frac{\langle\delta T_{fs}^2\rangle}{\tau^2} = 2 \int_0^\infty dx \int_0^\infty dy \, \frac{e^{-(x+y)^2-2\alpha(x+y)}}{x(x+y)} \left[e^{2xy} - 1\right]. \tag{A16}$$

Rescaling x and y by 2α and making the change of variables $u \equiv x + y$, $v \equiv x$, we have

$$\frac{\langle\delta T_{fs}^2\rangle}{\tau^2} = 2 \int_0^\infty \frac{du}{u} \int_0^u \frac{dv}{v} \, e^{-u^2/4\alpha^2 - u} \left[e^{v(u-v)/2\alpha^2} - 1\right]. \tag{A17}$$

Expanding the bracketed portion of the integrand as a power series and observing that

$$\int_0^u dv \, v^k (u-v)^{k+1} = \frac{k!(k+1)!}{(2k+2)!} u^{2k+2}, \tag{A18}$$

we have, integrating over u term-by-term,

$$\frac{\langle\delta T_{fs}^2\rangle}{\tau^2} = \sum_{k=0}^\infty \frac{k!}{2^k(2k+2)!\alpha^{2k+2}} \int_0^\infty du \, e^{-u^2/4\alpha^2 - u} u^{2k+1}. \tag{A19}$$

To allow us to apply standard asymptotic results to the remaining integral, we note that it is related to Tricomi's confluent hypergeometric function $U(a, b, z)$ as $\alpha^{2k+2}(2k+1)! \, U(k+1, 1/2, \alpha^2)$ [39], hence, shifting indices for convenience, we can write

$$\frac{\langle\delta T_{fs}^2\rangle}{\tau^2} \sim \sum_{k=1}^\infty \frac{(k-1)!}{2^k k} U\left(k, \frac{1}{2}, \alpha^2\right). \tag{A20}$$

Using the standard result that

$$U(a, b, z) \sim z^{-a} \Bigg[1 - \frac{a(a-b+1)}{z}$$
$$+ \frac{a(a+1)(a-b+2)(a-b+1)}{2z^2}$$
$$+ \mathcal{O}(z^{-3})\Bigg], \tag{A21}$$

for $|z| \gg 1$ [39], we have

$$\frac{\langle\delta T_{fs}^2\rangle}{\tau^2} \sim \frac{1}{2\alpha^2} - \frac{1}{8\alpha^4} + \mathcal{O}(\alpha^{-6}), \tag{A22}$$

which yields the lowest-order approximation given in the main text.

To obtain the asymptotic approximations for the timing variability in the simple model for neural fatigue given in the main text (10, 11, 12), we start from the asymptotic expansions without fatigue (8, 9), and apply the laws of total expectation and total variance given the assumed distribution of the parameter m. We then expand the resulting expressions about the baseline spiking threshold V_{th} to lowest order in $\delta V_{th}/(I_s + I_0 - V_{th})$, assuming that $m_{max} \delta V_{th} \ll I_s + I_0 - V_{th}$, yielding the asymptotic approximations (10) and (11).

3. Moments of the first-spike-interval in a delta-function approximation

In the previous appendix and in the main text, we considered the approximation of the distribution of initial membrane potentials by the stationary Gaussian limit (5). In this appendix, we consider a delta-function approximation $P(V_0) \approx \delta(V_0 - \langle V_0\rangle)$. This approximation maps directly to the standard treatment of leaky IF neurons with the appropriate replacement of V_r by I_0. Here, we review the derivation of the corresponding asymptotic results [26, 27]. In the limit $V_{th} - I_0 \gg \sigma$ of low firing rates, we have $\langle V_0\rangle = I_0$, hence we fix $V_0 = I_0$ in this approximation. Considering the mean first-spike-interval, we again start from the standard expression (6) with V_0 set to I_0, and rescale $\sigma y \mapsto y$, yielding

$$\frac{\langle T_{fs}\rangle}{\tau} = \frac{\sqrt{\pi}}{\sigma} \int_{-I_s}^{V_{th}-I_0-I_s} dy \, e^{y^2/\sigma^2} \left(1 + \text{erf}\left(\frac{y}{\sigma}\right)\right). \tag{A23}$$

In the limit $I_s + I_0 - V_{th} \gg \sigma$ of large synaptic inputs, the quantity y in the above integrand is always negative, and we have $y/\sigma \ll -1$. Using the asymptotic expansion of the error function for $x \ll -1$ [39],

$$\text{erf}(x) \sim -1 + \frac{e^{-x^2}}{\sqrt{\pi}|x|} \left(1 - \frac{1}{2x^2} + \dots\right), \tag{A24}$$

we obtain

$$\frac{\langle T_{fs}\rangle}{\tau} \sim \log\left(\frac{I_s}{I_s + I_0 - V_{th}}\right)$$
$$- \frac{\sigma^2}{4}\left(\frac{1}{(I_s + I_0 - V_{th})^2} - \frac{1}{I_s^2}\right) \tag{A25}$$

to lowest order. Similarly, for the variance of the first-spike-interval, we start with the standard expression (7) with $V_0 = I_0$. Again rescaling the variables of integration

by σ and using the asymptotic form of the error function, we obtain the lowest-order asymptotic approximation

$$\frac{\langle \delta T_{fs}^2 \rangle}{\tau^2} \sim \frac{\sigma^2}{2} \left(\frac{1}{(V_{th} - I_s - I_0)^2} - \frac{1}{I_s^2} \right). \tag{A26}$$

Comparing these expressions to the corresponding results (8, 9) in the approximation of the initial membrane potential distribution by the stationary Gaussian distribution (5), we observe that they are identical up to the presence of the $-I_s^{-2}$ terms in the lowest-order approximations. The presence of these terms in the delta-function approximation means that the variability decreases more rapidly with increasing synaptic strength and increases less rapidly with increasing noise variance σ^2 than in the Gaussian approximation.

[1] R. B. Ivry and R. M. Spencer, The neural representation of time, Current opinion in neurobiology **14**, 225 (2004).

[2] M. D. Mauk and D. V. Buonomano, The neural basis of temporal processing, Annu. Rev. Neurosci. **27**, 307 (2004).

[3] J. T. Coull, R.-K. Cheng, and W. H. Meck, Neuroanatomical and neurochemical substrates of timing, Neuropsychopharmacology **36**, 3 (2011).

[4] H. Merchant, D. L. Harrington, and W. H. Meck, Neural basis of the perception and estimation of time, Annual review of neuroscience **36**, 313 (2013).

[5] J. J. Paton and D. V. Buonomano, The neural basis of timing: distributed mechanisms for diverse functions, Neuron **98**, 687 (2018).

[6] R. H. Hahnloser, A. A. Kozhevnikov, and M. S. Fee, An ultra-sparse code underlies the generation of neural sequences in a songbird, Nature **419**, 65 (2002).

[7] M. A. Long and M. S. Fee, Using temperature to analyse temporal dynamics in the songbird motor pathway, Nature **456**, 189 (2008).

[8] F. Ali, T. M. Otchy, C. Pehlevan, A. L. Fantana, Y. Burak, and B. P. Ölveczky, The basal ganglia is necessary for learning spectral, but not temporal, features of birdsong, Neuron **80**, 494 (2013).

[9] C. Tang, D. Chehayeb, K. Srivastava, I. Nemenman, and S. J. Sober, Millisecond-scale motor encoding in a cortical vocal area, PLoS biology **12**, e1002018 (2014).

[10] G. B. Mello, S. Soares, and J. J. Paton, A scalable population code for time in the striatum, Current Biology **25**, 1113 (2015).

[11] K. H. Srivastava, C. M. Holmes, M. Vellema, A. R. Pack, C. P. Elemans, I. Nemenman, and S. J. Sober, Motor control by precisely timed spike patterns, Proceedings of the National Academy of Sciences **114**, 1171 (2017).

[12] E. D. Remington, D. Narain, E. A. Hosseini, and M. Jazayeri, Flexible sensorimotor computations through rapid reconfiguration of cortical dynamics, Neuron **98**, 1005 (2018).

[13] S. J. Sober, S. Sponberg, I. Nemenman, and L. H. Ting, Millisecond spike timing codes for motor control, Trends in neurosciences **41**, 644 (2018).

[14] M. Abeles, Local circuits: an electrophysiological study (1982).

[15] M. Diesmann, M.-O. Gewaltig, and A. Aertsen, Stable propagation of synchronous spiking in cortical neural networks, Nature **402**, 529 (1999).

[16] N. F. Hardy and D. V. Buonomano, Neurocomputational models of interval and pattern timing, Current Opinion in Behavioral Sciences **8**, 250 (2016).

[17] M. Li and H. Greenside, Stable propagation of a burst through a one-dimensional homogeneous excitatory chain model of songbird nucleus hvc, Physical Review E **74**, 011918 (2006).

[18] D. Z. Jin, F. M. Ramazanoğlu, and H. S. Seung, Intrinsic bursting enhances the robustness of a neural network model of sequence generation by avian brain area hvc, Journal of computational neuroscience **23**, 283 (2007).

[19] D. Z. Jin, Generating variable birdsong syllable sequences with branching chain networks in avian premotor nucleus hvc, Physical Review E **80**, 051902 (2009).

[20] M. Long, D. Jin, and M. Fee, Support for a synaptic chain model of neuronal sequence generation, Nature **468**, 394 (2010).

[21] C. M. Glaze and T. W. Troyer, Temporal structure in zebra finch song: implications for motor coding, Journal of Neuroscience **26**, 991 (2006).

[22] B. P. Olveczky, A. S. Andalman, and M. S. Fee, Vocal experimentation in the juvenile songbird requires a basal ganglia circuit., PLoS biology **3**, e153 (2005).

[23] C. M. Glaze and T. W. Troyer, Behavioral measurements of a temporally precise motor code for birdsong, Journal of Neuroscience **27**, 7631 (2007).

[24] C. M. Glaze and T. W. Troyer, A generative model for measuring latent timing structure in motor sequences, PLoS One **7**, e37616 (2012).

[25] N. Brunel, Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons, Journal of computational neuroscience **8**, 183 (2000).

[26] H. Tuckwell, *Introduction to theoretical neurobiology: nonlinear and stochastic theories*, Vol. 8 (Cambridge Univ Pr, 2005).

[27] F. Wan and H. Tuckwell, Neuronal firing and input variability, J. theor. Neurobiol **1**, 197 (1982).

[28] A. A. Faisal, L. P. Selen, and D. M. Wolpert, Noise in the nervous system, Nature reviews neuroscience **9**, 292 (2008).

[29] A. Renart and C. K. Machens, Variability in neural activity and behavior, Current opinion in neurobiology **25**, 211 (2014).

[30] N. A. Gershenfeld, *The nature of mathematical modeling* (Cambridge university press, 1999).

[31] M. Abeles, *Corticonics: Neural circuits of the cerebral cortex* (Cambridge University Press, 1991).

[32] J. Hertz, Theoretical aspects of neural computation. in (p. 135-144). wong ky m. king i. and yeung dy (1997).

[33] S. Goedeke and M. Diesmann, The mechanism of synchronization in feed-forward neuronal networks, New Journal of Physics **10**, 015007 (2008).

[34] C. Pehlevan, F. Ali, and B. P. Ölveczky, Flexibility in motor timing constrains the topology and dynamics of pattern generator circuits, Nature communications **9**, 977 (2018).

[35] T. M. Otchy, S. B. Wolff, J. Y. Rhee, C. Pehlevan, R. Kawai, A. Kempf, S. M. Gobes, and B. P. Ölveczky, Acute off-target effects of neural circuit manipulations, Nature **528**, 358 (2015).

[36] C. Walton, E. Pariser, and F. Nottebohm, The zebra finch paradox: song is little changed, but number of neurons doubles, Journal of Neuroscience **32**, 761 (2012).

[37] K. Hamaguchi and R. Mooney, Recurrent interactions between the input and output of a songbird cortico-basal ganglia pathway are implicated in vocal sequence variability, Journal of Neuroscience **32**, 11671 (2012).

[38] A. K. Dhawale, M. A. Smith, and B. P. Ölveczky, The role of variability in motor learning, Annual review of neuroscience **40**, 479 (2017).

[39] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, Vol. 55 (US Government printing office, 1948).