



OPEN The impact of spike timing precision and spike emission reliability on decoding accuracy

Wilten Nicola^{1,2,3,5}✉, Thomas Robert Newton^{1,5} & Claudia Clopath⁴

Precisely timed and reliably emitted spikes are hypothesized to serve multiple functions, including improving the accuracy and reproducibility of encoding stimuli, memories, or behaviours across trials. When these spikes occur as a repeating sequence, they can be used to encode and decode a potential time series. Here, we show both analytically and in simulations that the error incurred in approximating a time series with precisely timed and reliably emitted spikes decreases linearly with the number of neurons or spikes used in the decoding. This was verified numerically with synthetically generated patterns of spikes. Further, we found that if spikes were imprecise in their timing, or unreliable in their emission, the error incurred in decoding with these spikes would be sub-linear. However, if the spike precision or spike reliability increased with network size, the error incurred in decoding a time-series with sequences of spikes would maintain a linear decrease with network size. The spike precision had to increase linearly with network size, while the probability of spike failure had to decrease with the square-root of the network size. Finally, we identified a candidate circuit to test this scaling relationship: the repeating sequences of spikes with sub-millisecond precision in area HVC (proper name) of the zebra finch. This scaling relationship can be tested using both neural data and song-spectrogram-based recordings while taking advantage of the natural fluctuation in HVC network size due to neurogenesis.

The firing of spikes is metabolically costly, with the brain consuming approximately 20% of the body's energy at any time¹. Indeed, studies show that the firing and transmission of action potentials via synapses consumes the majority of the predicted energy used by cells in the cerebral cortex². Thus, to maximize the use of any particular spike in encoding/decoding stimuli across trials, a neural circuit may devote resources to increasing the reliability of emitting existing spikes and the timing precision of these emitted spikes, rather than adding extra spikes. This may also be a useful strategy in improving the performance of neuromorphic circuits, which use hardware instantiated spiking neurons for computations^{3,4}.

Indeed, there are existing examples in neuroscience where neural circuits prioritize the reliability of spike emission and the precision of spike times, rather than encoding with a large volume of spikes. For example, the neural circuits controlling the singing behaviour of the zebra finch prioritize spike timing precision and emission reliability. The zebra finch singing behaviour is dependent on multiple nuclei post-learning. Here, we focus on the HVC (proper name), the Robust Nucleus of the Archopallium (RA), and the hypoglossal nucleus^{5–11}. During the approximately half second bout of singing in a single song, HVC neurons which project to area RA (HVC_{RA} neurons) fire a precise chain of spikes where each HVC_{RA} neuron fires a burst at a well-defined moment in time¹⁰. This chain of spikes covers the entire time interval of singing, even during the silent intervals between the different segments (or syllables) of a single song¹². This chain of spikes is also highly reproducible across singing bouts, with individual spikes displaying sub-millisecond precision^{8,10}. For example, the first spike fired in a burst for HVC_{RA} neurons was estimated to have a precision of 0.73 ± 0.3 ms. At RA, the individual neurons also display highly reproducible spike sequences with sub-millisecond (0.28 ± 0.3 ms) precision⁸. RA serves as one of the final command signals to produce singing through the vocal organ with the hypoglossal nucleus acting as the final relay. The neural architecture at the heart of this reproducible and precise behaviour is a reproducible, precise, and sparse sequence of spikes.

If populations of neurons use precisely timed and reliably emitted spikes to encode stimuli or produce behaviours, how would that impact the ability to decode out these stimuli or behaviours? Indeed, there is substantial computational modelling and analytical results that investigate the accuracy of decoding a stimulus. For

¹University of Calgary, Calgary, Canada. ²Department of Cell Biology and Anatomy, Calgary, Canada. ³Hotchkiss Brain Institute, Calgary, Canada. ⁴Department of Bioengineering, Imperial College London, London, UK. ⁵These authors contributed equally: Wilten Nicola and Thomas Robert Newton. ✉email: wilten.nicola@ucalgary.ca

example, if the spikes are only emitted as part of a rate code, then prior work shows that we should expect that decoding accuracy only scales with the square-root of the number of spikes fired^{13,14}. However, there are also proposed neural circuits that rely on balanced excitatory/inhibitory networks that do not depend on precise and repeatable sequences of spikes yet have decoding errors that decrease linearly with network size^{13,15–17}. These networks fire spikes sparsely^{15–18} through a coordinated network structure where a neuron's time varying inputs are precisely balanced between excitation and inhibition. The precise balance of inputs allows a neuron to compute the decoding error in representing a stimulus, or dynamical system. A spike is only fired when a threshold decision boundary is crossed^{15–18}, at which point the error in reproducing a stimulus or dynamical system is reset. These networks are robust to neural heterogeneity^{19,20}, neuronal model implementation¹⁷, and can be learned locally with biologically plausible spike timing-dependent plasticity rules^{21–23}. Owing to the error reset mechanism, the spike trains produced have Poisson-like statistics, without explicitly being drawn from a Poisson point process^{15–17}.

In this study, we investigate the decoding accuracy of time-series data from precisely timed and regular sequences of spikes. By using synthetically constructed spike-trains and mathematical analysis, we derived sufficient conditions for sequences of spikes to have a qualitative difference over rate codes in the accuracy of decoding time series data. This time series data can represent behaviours, stimuli, or other signals but must be suitably smooth in a mathematical sense. In networks containing perfectly precise and reliable spikes, the root mean squared error (RMSE) of a time-series decoded with a linear decoder decreases linearly with the number of spikes or neurons used in the decoding. This is similar to the aforementioned efficient codes that rely on E/I balance^{13,15–17}. We tested these findings in numerical simulations with synthetically generated spike-trains and synthetically generated signals to be decoded. This linear decrease in the RMSE with network size also holds when the spike times are no longer perfectly precise and reliable. However, the spike timing precision and spike emission reliability of a neuron must increase along with the size of a network. The error incurred by approximating the system with spikes can be regarded as the bias in a bias variance decomposition, while the variability in the replay of a signal caused by imprecise or unreliable spikes can be interpreted as the variance. Finally, we discuss how to test the scaling-relationship experimentally in the zebra finch circuit. This circuit exhibits a natural increase in the HVC network size via neurogenesis. We predict that if the decoding accuracy of the stereotypical zebra finch song increases linearly with HVC network size, then so must the precision of HVC spikes.

Results

Before considering how precise spike times impact the neural code, neural encoding and neural decoding will be defined. As an organism performs a behaviour or perceives a stimulus which is represented with the function of time, $x(t)$, spikes are emitted as a result of the stimulus or in pursuit of enacting the behaviour. We will assume that the signal $x(t)$ occupies an interval $t \in [0, T]$ with all spikes generated in the encoding process confined to this interval. As such, the stimulus becomes encoded by some encoding transformation E into a series of spikes which can be represented as a matrix.

$$\boldsymbol{\tau} = E(x(t)) \quad (1)$$

Here the vector $\boldsymbol{\tau}$ is an $N \times n_t$ matrix where N corresponds to the number of neurons, and n_t is a discretization of time. Thus, we define the matrix $\boldsymbol{\tau}$ as

$$\tau_{ij} = \begin{cases} 1 & \text{if Neuron } i \text{ spiked in } I_j \\ 0 & \text{otherwise} \end{cases}$$

where the I_j th time interval corresponds to $I_j = [\frac{jT}{n_t}, \frac{(j+1)T}{n_t}]$. We will not explicitly specify how the function E generates the encoding matrix $\boldsymbol{\tau}$, but rather consider the conditions on a mapping that allow for accurate decoding. For example, the mapping E may generate the same realization of spikes for $x(t)$ every time yielding reliable and precisely timed spikes. Alternatively, E may be stochastic in nature and the spike times may have some noise or “jitter” present, or subsets of the spikes may even fail to be emitted. Thus, the function E in (1) defines a basic neural encoding scheme.

Decoding can be thought of as a kind of inverse to Eq. (1). In neural decoding, the goal is to reconstruct $x(t)$ from $\boldsymbol{\tau}$. In general, inverting Eq. (1) is not possible. However, $x(t)$ can be approximated with $\hat{x}(t)$ and some decoding operator D :

$$\hat{x}(t) = D(\boldsymbol{\tau}) = D(E(x(t))) \quad (2)$$

One approach to decoding $x(t)$ is to convolve the spikes with some sort of kernel, $K(t)$ and then consider the decoder D to be a linear transformation, $D = \boldsymbol{\phi}^x$

$$\hat{x}(t) = \sum_{j=1}^N \phi_j^x r_j(t) = \sum_{j=1}^N \sum_{t_{jk} < t} \phi_j^x K(t - t_{jk}) \quad (3)$$

where t_{jk} is the j th spike fired by the k th neuron. The Kernel function $K(t)$ the spikes. Here, we consider the simple exponential filter ($K(t) = \exp(-t/\tau)$) with a filtering time constant of 10 ms. The function $r_j(t)$ is the filtered spike history for neuron j . We can interpret \boldsymbol{r} as a temporal basis set. The linear decoder $\boldsymbol{\phi}^x$ transforms the temporal basis set $\boldsymbol{r}(t)$ into the approximation to the decoded signal ($\hat{x}(t)$).

Typically, the decoder is constructed by using one or a subset of repetitions of the behaviour or stimulus (and spikes) to train the decoder. The trained decoder is later tested with repetitions that were not used to construct the initial decoder.

Neural encoding and decoding raise the question of how accurate can we decode out the function $x(t)$ with $\hat{x}(t)$. The Root-Mean-Squared Error (RMSE) is commonly, although by no means exclusively used as the metric to measure this accuracy^{14–17}. For example, if we consider the audio-time series of a zebra finch singing as $x(t)$, then the RMSE measures the difference between the decoded song ($\hat{x}(t)$) and the recording with a RMSE of 0 indicating perfect decoding accuracy. The RMSE is defined as:

$$RMSE = \sqrt{\int_0^T (\hat{x}(t) - x(t))^2 dt} = \sqrt{MSE} \quad (4)$$

where MSE denotes the mean-squared-error. With encoding, decoding, and the RMSE of decoding now defined, the question of how accurate can the stimulus $x(t)$ be reconstructed naturally emerges. More formally, for progressively larger networks ($N \rightarrow \infty$), what is the limiting behaviour of the RMSE?

$$RMSE = \sqrt{\int_0^T (\hat{x}(t) - x(t))^2 dt} = O(N^\lambda), \quad N \rightarrow \infty \quad (5)$$

Efficient coding networks were proposed in^{15–17} to 1) sparsely fire spikes and 2) accurately and robustly represent signals or internal dynamics all with one-neural architecture. Briefly, these networks use precisely balanced slower connections and faster inhibitory connections which can be learned or pre-configured as a result of an error reducing optimization criterion^{15–18,21–23}. These networks are robust to losing neurons, and can accurately produce a desired output. These networks precisely track a desired signal as a decoded representation $\hat{x}(t)$. When the error in this representation reaches a critical threshold (or bounding box¹⁸), a spike is emitted to reset the local error in the representation to 0. The neuron that ultimately fires this spike inhibits the other neurons in the network to maximize the efficiency of the code. Thus, for these networks, the λ value in Eq. 5 is -1 and adding more neurons decreases the RMSE linearly with the network size.

As we will demonstrate numerically and analytically, for precisely timed sequences of spikes, such as those fired by HVC_{RA} neurons or pyramidal neurons in RA in the zebra finch, the RMSE also has a $\lambda = -1$ and decreases linearly with the number of neurons:

$$RMSE = \sqrt{\int_0^T (\hat{x}(t) - x(t))^2 dt} \propto \frac{1}{N} \quad (\text{Precise spike times and reliable spikes}). \quad (6)$$

This is qualitatively different from and in stark contrast to alternate codes. For example, if the timing of spikes is unimportant for the neural code, and only the number of spikes per unit of time (the rate) encodes information, the RMSE decreases with the square root ($\lambda = -1/2$) of the number of neurons^{15–17}:

$$RMSE = \sqrt{\int_0^T (\hat{x}(t) - x(t))^2 dt} \propto \frac{1}{\sqrt{N}} \quad (\text{Rate Code}). \quad (7)$$

The bias-variance decomposition

To test how the RMSE changes when spikes become imprecisely timed or unreliably emitted, a perturbed decoded approximation $\tilde{x}(t)$ is defined as the approximation to $x(t)$ when the decoder ϕ^x is applied to the second set of spikes which suffer from spike failure, where spikes fail to be emitted at their designated time, or spike jitter, where the spike time of a spike is randomly perturbed. Then the MSE can be decomposed as

$$\begin{aligned} MSE &= \int_0^T (\tilde{x}(t) - x(t))^2 dt = \|\tilde{x} - x\|^2 \\ &= \|\tilde{x} - \hat{x} + \hat{x} - x\|^2 \end{aligned} \quad (8)$$

$$\begin{aligned} &\leq (\|\tilde{x} - \hat{x}\| + \|\hat{x} - x\|)^2 = \|\hat{x} - x\|^2 + \|\tilde{x} - \hat{x}\|^2 + 2\|\hat{x} - x\|\|\tilde{x} - \hat{x}\| \\ &= \text{Bias}^2 + \text{Variance} + 2\text{Bias} \cdot \text{Std} \end{aligned} \quad (9)$$

where the square of the RMSE is the mean squared error (MSE) and $\|\cdot\|$ is the standard L_2 function norm defined over $[0, T]$. The first term in Eq. (8) can be interpreted as the bias in the bias-variance trade-off, while the second term is the variance (square of the standard deviation). The first term quantifies how well we can approximate $x(t)$ with a spike train, while the second term is the error incurred in assuming that the spike train is perfectly repeatable.

Precise and reliable spike times lead to a linear decrease in the root mean squared error (RMSE) with increasing network sizes

To investigate how spike timing precision and spike emission reliability impact decoding accuracy, we constructed synthetic spike trains where each neuron fires a variable number of spikes. These spikes were generated randomly from a Poisson distribution, with the same realization of spikes drawn on each trial. Further, the spikes are reliable and precisely timed from trial to trial as they are drawn from the same realization, however they can be easily and independently perturbed in subsequent trials (Fig. 1A–B). We constructed optimal linear

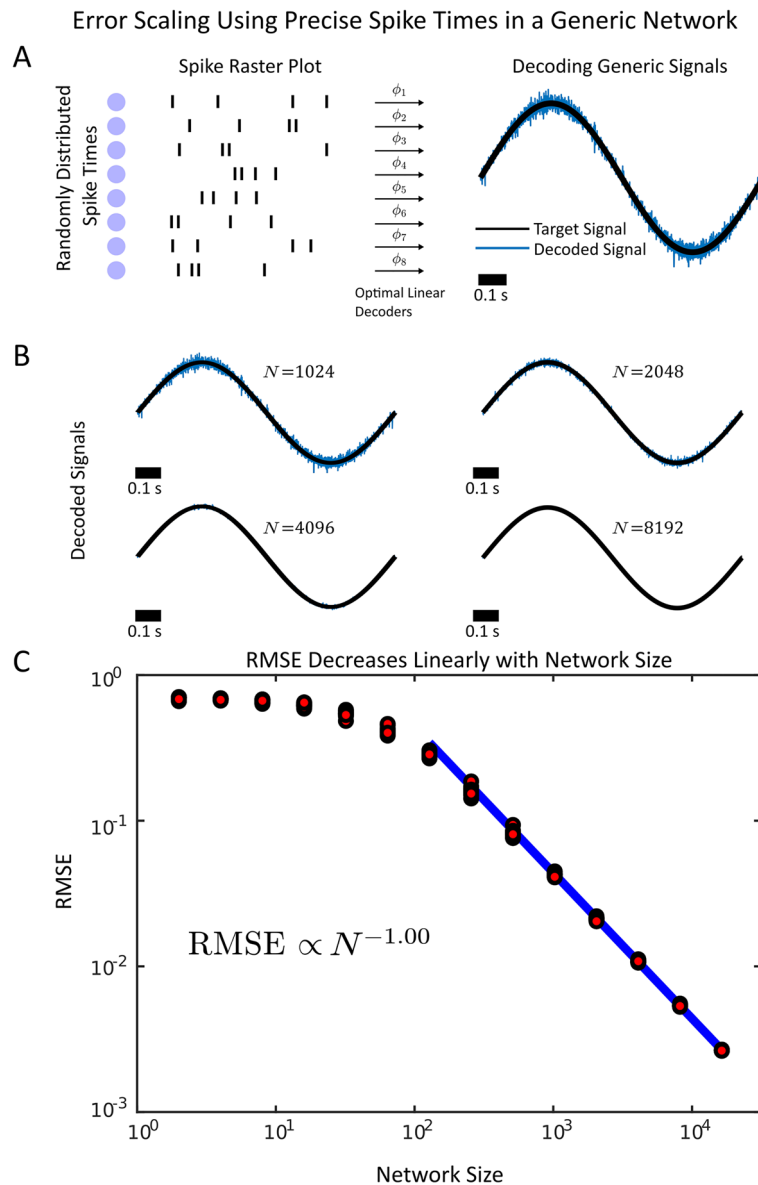


Figure 1. (A) (Left) A diagram of a randomly generated spike sequence used to decode a time series with optimal linear decoders. A decoded signal (blue) along with the target supervisor (black) is shown in the right. The signal is decoded using $N = 1024$ neurons. (B) Decoding the time series with larger networks. The spikes generated are drawn with the same statistics as in (A) and filtered identically. The signal is decoded for progressively larger networks ($N = 2^{10}$ to $N = 2^{14}$). (C) The loglog plot of the RMSE vs network size. The RMSE decreases linearly with the network size. The linear fit was performed to the mean of the 10 realizations from $N = 2^{10}$ to $N = 2^{14}$. The largest five.

decoders that could decode out a target signal, in this case, a simple sinusoidal oscillator with a period of 1 Hz (Fig. 1A). We found that the RMSE decreased linearly ($RMSE \propto N^{-1.00}$) with the network size for sufficiently large networks. Thus, if spike times are perfectly precise, they can be used to decode out a continuous signal with N^{-1} RMSE scaling. The spikes do not need to explicitly encode features of the signal, but merely tile time with sufficiently large numbers. This is an identical RMSE scaling to efficient codes^{15–17} and superior to the error scaling achieved by using firing rates ($RMSE \propto N^{-1/2}$).

Next, we investigated for what general conditions this result would hold (Supplementary Material S1 and Fig. S1). We found that the linear decrease in the RMSE with network size is a general, mathematical result that occurs so long as the time series that the spikes must approximate is sufficiently smooth (differentiable). This criterion can be relaxed to the function being Lipschitz continuous (Supplementary Fig. S1). The spikes need not be ordered into a chain or generated randomly as above but must be precise from trial-to-trial and sufficiently dense in time (Supplementary Material S1). Under these conditions, the efficient coding error scaling $RMSE \propto N^{-1}$ is likely to be achieved for sufficiently large network sizes. Further, the spikes can overlap and interfere with each other extensively (Supplementary Fig. S2) and need not be sparsely distributed or orthogonal.

As long as the spikes are precisely timed, the RMSE will still decrease linearly with the network size (Supplementary Fig. S2) under the smoothness conditions considered in the Supplementary Material. Alternatively stated, in the case of perfectly timed and perfectly reliable spikes and a suitably smooth supervisor, the bias term in Eq. (8) decreases linearly with the network size, while the variance is 0 as the spikes contain no degree of imprecision or unreliability.

The precision of spike times and reliability of spikes must increase with the network size to preserve $\text{RMSE} \propto N^{-1}$ scaling

While spike times in the zebra finch HVC_{RA} neurons and RA neurons are precise, they are not perfectly precise. From trial-to-trial, the spike times often display sub-millisecond levels of jitter. This is true for neurons in both the HVC and RA^{8,10} where the jitter is $O(0.1)$ ms (depending on the neuronal subtype). Further, neurons can also fail in firing spikes owing to unreliable synapses²⁴. Thus, we considered just how precise and reliable spikes should be when increasing the network size to maintain a linear decrease in the RMSE.

First, we considered spike-jitter from trial-to-trial by constructing a decoder with initially precise spikes. Then, we perturbed the spike times randomly and reapplied the decoder on subsequent trials, and measured the RMSE (Fig. 2A). We found that if the standard deviation of the jitter (σ) was fixed, the N^{-1} error scaling in the RMSE was no longer present ($\text{RMSE} \propto N^{-0.085}$). This was numerically confirmed for trained optimal linear decoders (Fig. 2). The sub-linear decrease in the RMSE was also confirmed for other fixed values of the standard deviation of the spike jitter (Supplementary Fig. S3).

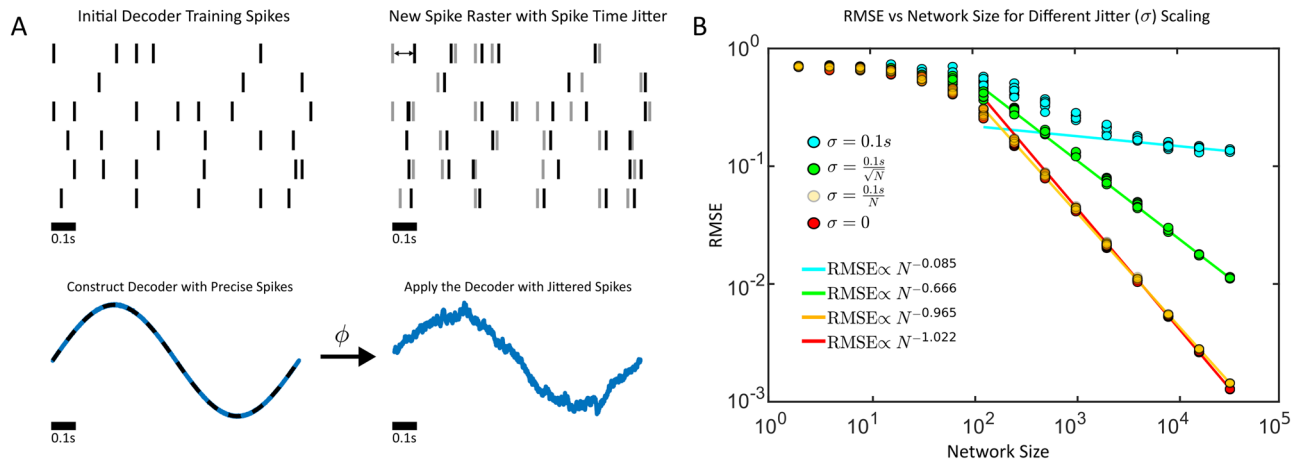
If a fixed amount of timing imprecision would destroy the linear decrease in the RMSE, could populations of neurons increase their spike timing precision with network size to restore it? To investigate this hypothesis, we considered the possibility that spikes might become more precisely timed with increasing network sizes by having the jitter decrease with larger networks (either $\sigma \propto \sqrt{N^{-1}}$ or $\sigma \propto N^{-1}$, Fig. 2B). We found that if the spike times became more precise linearly ($\sigma \propto N^{-1}$) with the network size, this was sufficient to restore the linear decrease in the RMSE with network size ($\text{RMSE} \propto N^{-0.965}$).

Next, we considered the case where spikes might abruptly fail (Fig. 2C,D). We again found that if the probability of spike failure (p_F) was fixed and independent of the network size, the linear decrease in the RMSE with increasing network size would no longer be present ($\text{RMSE} \propto N^{-0.130}$). However, once again, if the spikes are less likely to fail with larger networks ($p_F \propto \sqrt{N^{-1}}$), the linear decrease in the RMSE would be restored ($\text{RMSE} \propto N^{-1.022}$). These results were qualitatively and quantitatively similar to the case where spikes were randomly emitted (rather than randomly failed, Supplementary Fig. S4). Thus, spike timing imprecision or spike-failure are themselves not sufficient to eliminate the linear decrease in the RMSE, but the spikes can become more reliable and more precisely timed with progressively larger network sizes to maintain $\text{RMSE} \propto N^{-1}$ scaling. The probability of a spike failing can decrease with the square-root of the network size while the standard deviation of spike jitter should decrease linearly with the network size as a sufficient condition to preserve a linear decrease in RMSE scaling.

The impacts of spike timing imprecision, and spike emission reliability were also explored analytically, albeit with simplifying assumptions for tractability (Supplementary Material S2). Indeed, in the bias-variance decomposition, sufficient (but not necessary) conditions were determined to restrict the variance to scale like N^{-2} , which leads to the RMSE scaling like $\propto N^{-1}$. First, we considered the case of a simple neuron-specific jitter to a filtered, but differentiable, spike train. This would be the case for example if the spikes were filtered with Gaussian filters where every spike for a single neuron is jittered by the same amount, which differs from the jitter applied to the other neurons' spikes. If the sum of the absolute decoders, given by $\sum_{j=1}^N |\phi_j^{x(t)}|$, is bounded for $N \rightarrow \infty$, then $\sigma \propto N^{-1}$ is a sufficient condition for $\text{RMSE} \propto N^{-1}$ as $N \rightarrow \infty$. For the case of failing spikes, we considered evenly distributed spikes with box filters. In this case, each spike does not have any redundancy in the form of overlapping spikes from other neurons to aid in the coding of a signal, as all the spike trains are orthogonal. In this scenario, it is sufficient for the spike failure rate to scale with $p_F \propto N^{-2}$, which differs radically from the numerical results of $p_F \propto \sqrt{N^{-1}}$ to maintain $\text{RMSE} \propto N^{-1}$. This discrepancy where the numerical simulations result in looser restrictions on the probability of spike failure may be due to the simplifying assumption of non-redundant spiking in the analytical derivation for tractability. To investigate this more thoroughly, the simulations in Fig. 2 were conducted with larger initial p_F ($p_F = 0.05$ and $p_F = 0.5$ Supplementary Fig. S3). In the latter case, we did find that $\text{RMSE} \propto -0.703$ when the $p_F \propto \frac{0.5}{\sqrt{N}}$. This may imply that the decrease in RMSE with p_F is dependent on the initial p_F with larger initial failure probabilities overwhelming any redundancy that additional spikes provides as the network becomes larger, or alternatively, the sub-linear decrease in RMSE with network size for larger p_F may be a numerical transient.

In Figs. 1 and 2, the decoders were trained with a single repetition of the supervisor and spike train. We tested the hypothesis that more repetitions would improve RMSE scaling with N by training a network with 10 trials of imprecise (Fig. 3A,C) or unreliable spikes (Fig. 3B). The linear decrease in RMSE with increasing network size was only maintained under identical scaling conditions as in Fig. 2 for the spike jitter ($\text{RMSE} \propto N^{-0.958}$) where the standard deviation decreased linearly with network size, and spike failure ($\text{RMSE} \propto N^{-0.953}$) where the probability of spike failure decreased with the square root of the network size. However, there was a notable improvement in RMSE scaling when using multiple trials when the probability of spike failure was constant ($p_F = 0.02$). When multiple trials were used, $\text{RMSE} \propto N^{-0.764}$ while a single training trial lead to $\text{RMSE} \propto N^{-0.243}$. It is possible that with more training trials, the exponent would shift further closer to -1 . Thus, to preserve the linear decrease in RMSE with network size, even when multiple trials are used to train a decoder, the spikes should still become more reliably emitted and more precisely timed with network size to maintain a linear decrease in the RMSE with network size.

Optimal Spike Timing Codes Require Increasing Spike Time Precision with Network Size



Optimal Spike Timing Codes Require Increasing Spike Reliability with Network Size

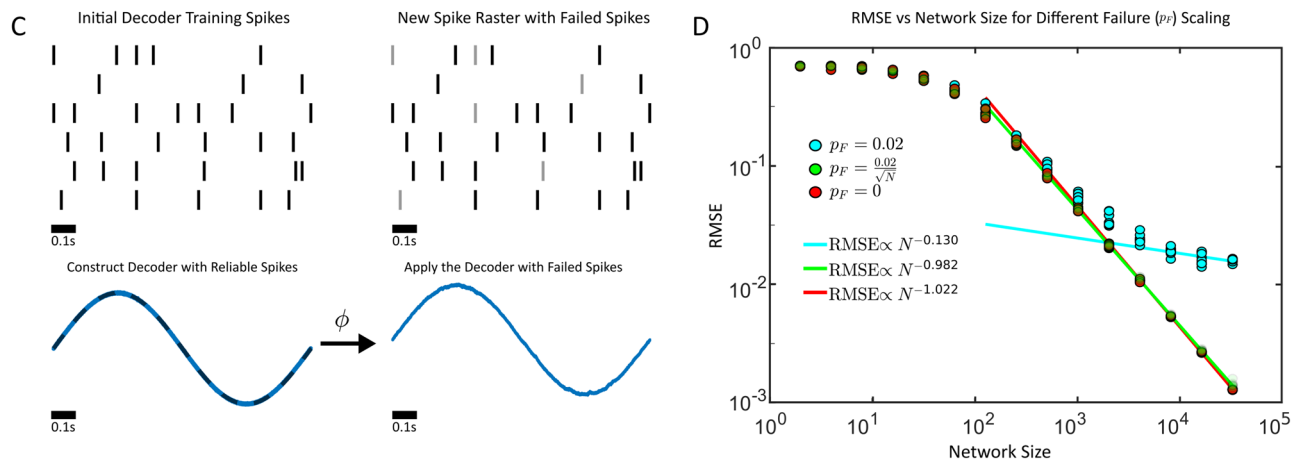


Figure 2. (A) Synthetic spike trains are generated and used to train an optimal decoder on a simple sinusoidal oscillator with a period of 1 second. The spike trains are then jittered and the decoder is reapplied, with the resulting root mean squared error (RMSE) measured. (Top) A cartoon of the spike generation protocol: the spikes are first generated to train a decoder. A new spike raster is then generated by jittering the spikes with the decoder reapplied. A sample curve for $N = 2^{14}$ shown where the spikes are jittered randomly with mean 0 and standard deviation $\sigma = 0.1s$. The supervisor is overlaid (black dashed line) during training. (B) The RMSE as a function of the network sized (N) for differing jitter amounts σ . The jitter is either fixed at 0.1 seconds (blue), decreases like the square-root of the network size (green), linearly with the network size (orange) or remains fixed with network size. (C) Synthetic spike trains are generated and decoders are trained as in (A), only now a random subset of spikes fail with p_F denoting the spike failure probability. A $p_F = 1$ implies that all spikes fail, while a $p_F = 0$ implies that the spikes are perfectly reliable. (Top) A cartoon of the spike generation protocol: the spikes are first generated to train a decoder. A new spike raster is then generated by dropping spikes randomly with probability p_F with the decoder reapplied after. A sample curve for $N = 2^{14}$ shown where the spikes are dropped randomly with $p_F = 0.02$. The supervisor is overlaid (black-dashed) line during training. (D) The RMSE as a function of the network sized (N) for differing spike failure amounts p_F .

The impacts of supervisor complexity and smoothness on linear RMSE scaling

The decoded signals thus far were simple smooth functions (e.g., sinusoidal inputs). We sought to determine how increasing the potential complexity of a supervisor would impact the scaling of the RMSE. First, we considered a multi-dimensional supervisor consisting of a sequence of pulses of variable duration (Fig. 4). The pulses represent different notes for the song “Ode to Joy” (see¹⁴ for further details). We found that precisely timed and reliably emitted spikes still lead to a linear decrease in the RMSE with network size ($\text{RMSE} \propto N^{-1.019}$, Fig. 4A). However, if the spikes are imprecise, the standard deviation of the spike jitter must decrease linearly with the network size to preserve the linear error scaling of the RMSE ($\text{RMSE} \propto N^{-0.981}$). If the spikes are unreliable, the probability of spike failure must also decrease with the square root of the network size to preserve the linear error scaling of the RMSE ($\text{RMSE} \propto N^{-1.008}$). Collectively, these results imply that the linear decrease of the RMSE with network size is robust to different supervisors as the RMSE scaling was identical to the results in Figs. 1, 2 and 3.

Using Multiple Trials with Jittered Spikes During Training

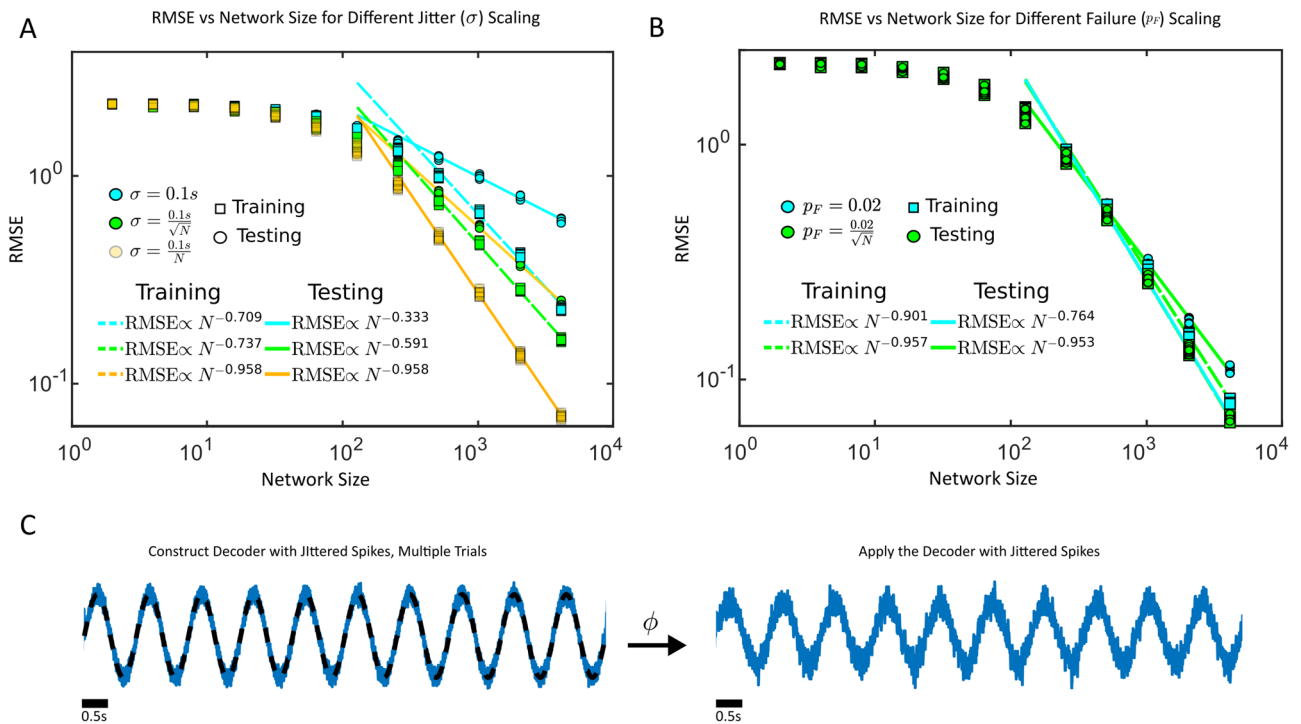


Figure 3. (A) The impact of multiple trials on RMSE scaling. Randomly generated spike sequences with imprecise and reliable spike times are used to linearly decode a sinusoidal oscillator with a period of 1 second. The decoder is trained with 10 repetitions of a spike train with jittered spikes and tested with another 10 repetitions to measure the RMSE. A total of 10 realizations of this procedure are used for each N . The spike jitter amount decreases linearly with N (yellow), with the square-root of N (green), or stays constant (teal). (B) Randomly generated spike sequences with precise spike times but unreliable spikes are used to linearly decode a generic sinusoidal oscillator with a period of 1 second. The decoder is trained with 10 repetitions of a spike train with failed spikes and tested with another 10 repetitions to measure the RMSE. A total of 10 realizations of this procedure are used for each N . (C) The target time series (black dashed lines) and the decoded estimate (blue). The decoder is reapplied in the left on a test trial of jittered spikes. The decoded signal is trained and tested in (C) with $N = 2^{12}$ neurons and a jitter of $\sigma = 0.1$ s. Note that for (A) and (B), the maximum network size considered was $N = 2^{12}$.

Next, we considered if the continuity of the decoded signal was critical for the linear decrease in RMSE with larger network sizes (Fig. 5). To that end, we utilized a simple but non-smooth signal (the sign function) that exhibits a discontinuity (Fig. 5A–C). Surprisingly, even when the spikes are both reliable and precise, the RMSE did not decrease linearly with the network size, but at a slower rate ($\text{RMSE} \propto N^{-0.772}$, Fig. 5A–B). This signal is also more poorly decoded when the spikes suffer from spike jitter or failure, however the slower RMSE scaling is restored when the standard deviation of the jitter decreases linearly with the network size ($\text{RMSE} \propto N^{-0.721}$, or when the spike failure rate decreases with the square root of the network size ($\text{RMSE} \propto N^{-0.771}$). The source of the discontinuity is where the magnitude of the error is the largest (Supplementary Fig. 5). We note that the point of discontinuity also explicitly breaks the smoothness conditions (and more generally the Lipschitz condition) required for $\text{RMSE} \propto N^{-1}$ scaling in the derivation (Supplementary Material S1). Further, we show that approximating a nonsmooth step function will change the bias scaling in the bias-variance decomposition to bias $\propto N^{-1/2}$ (Supplementary Material S2). This was derived analytically under simplifying assumptions about the nature of the spike train (evenly distributed single spikes) and filtering (box-filtering) which differs from the more realistic numerical examples considered. Nevertheless, the smoothness of the supervisor impacts the RMSE scaling via the bias term in the bias-variance decomposition.

Testing the RMSE versus network size in the zebra finch

We investigated what the impact of precise spike times would have on the RMSE of a decoded behaviour: the song produced by the zebra finch. We first constructed a simplified model of the zebra finch circuit with synthetically constructed spike trains that mimic the statistics of HVC_{RA} neurons (Fig. 6A). In particular, each HVC_{RA} neuron fires a single burst of spikes, and the bursts are sampled from a uniform distribution over the entire duration of the song. A decoder (ϕ^x) is then constructed using a zebra finch song recording (see Materials and Methods, ¹⁴, Fig. 6B). We considered three scenarios: a fixed jitter at $\sigma = 0.3$ ms a jitterless network with $\sigma = 0$ ms, and a network for which the jitter would decrease with N with $\sigma = \frac{1.5}{N}$. Note that we took a larger jitter value for small N so that the overall jitter in the latter case would be comparable to 0.3 ms for large N . We found that larger networks

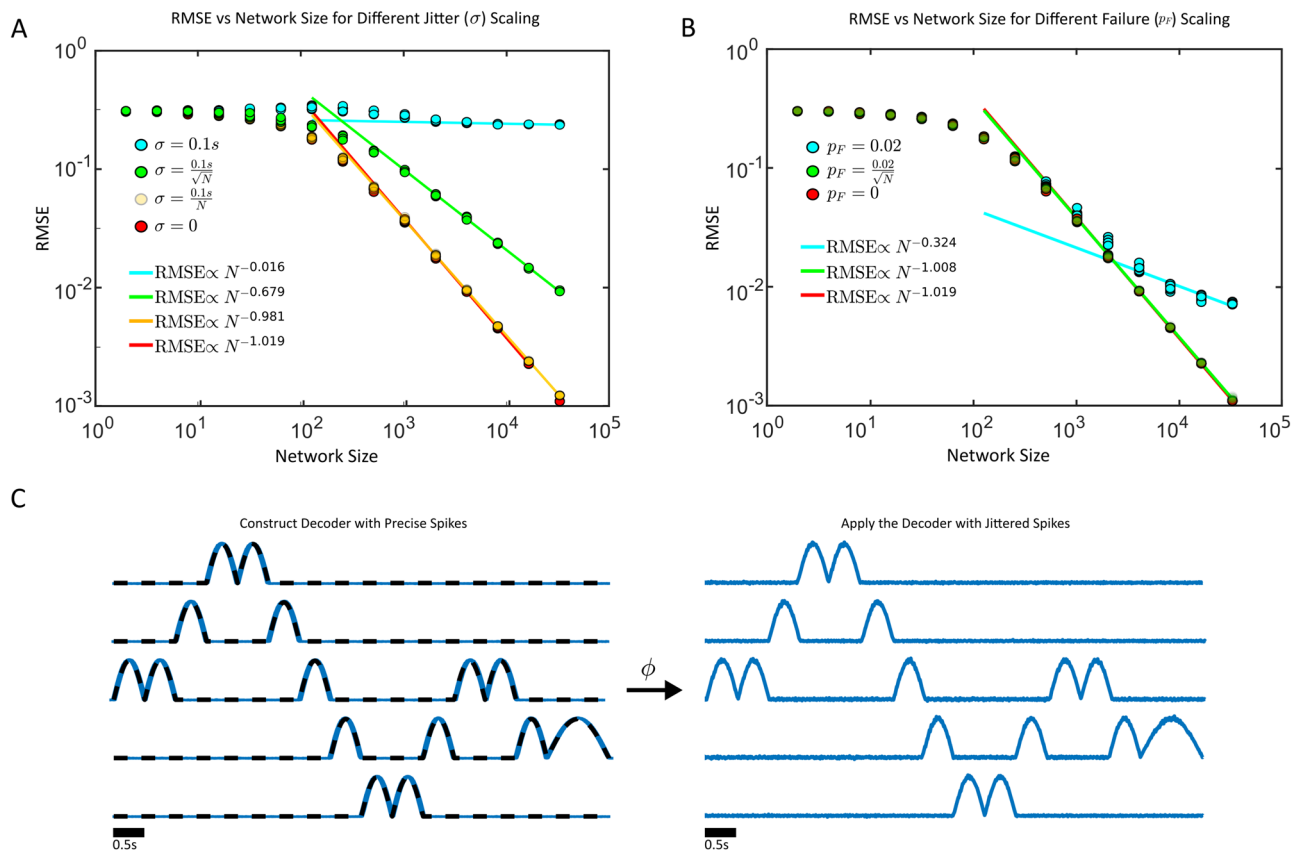


Figure 4. (A) Imprecise but reliable spikes are used to decode a 5-dimensional sequence of pulses that replicate the song “Ode to Joy”. The root mean squared error is plotted versus the network size. The spike jitter amount decreases linearly with N (yellow), with the square-root of N (green), or stays constant at $\sigma = 0.1$ s (teal), or stays constant at 0 (red). (B) Precisely timed but unreliable spikes with a probability of spike failure p_F are used to decode Ode to Joy. The spike failure probability either decreases with the square root of the network size (green), or is fixed at $p_F = 0.02$ (teal), or is fixed to 0 (red). (C) The decoded Ode to Joy signal at $N = 2^{15}$ with $\sigma \propto \sqrt{N}^{-1}$ (right). The jitterless signal (blue) and the supervisor (black) are shown in the left. The pulse width denotes either a quarter note or a half note, while the pulse dimension denotes the specific note.

(increasing N) of HVC_{RA} neurons resulted in more accurate decoding (Fig. 6B). For sufficiently large networks, the RMSE decreased approximately linearly when the jitter was either not present ($RMSE \propto N^{-0.9}$), or decreased linearly with the network size ($RMSE \propto N^{-0.89}$). We remark that the RMSE scaling exponent is slightly larger than -1, which may be due to discontinuities or rapid signal changes in the spectrogram (Supplementary Material S2). If the jitter was fixed at $\sigma = 0.3$ ms, the RMSE decreased sublinearly with network size ($RMSE \propto N^{-0.43}$).

While the RMSE is important for computational studies, it is difficult for an experimental neuroscientist to actually utilize and measure for a simple reason: they do not have access to the signal a circuit intends to represent, only the output of the circuit itself. Indeed, the RMSE represents the accuracy of a behaviour once that behaviour is known. Thus, we tested whether the precision of a behaviour also increased with network size by measuring the standard deviation of the decoded spectrograms (Fig. 6C–D). We found that the behavioural variability (see Materials and Methods for definition) of the spectrograms displayed similar scaling relationships (slopes of -0.43, -0.93, and -0.89) as the RMSE (slopes of -0.43, -0.9, and -0.89) for all conditions ($\sigma = 0.3$ ms, $\sigma = 0$ ms, $\sigma \propto N^{-1}$). This result allows an experimentalist to record a small number of repetitions (10 in this case) of animal singing, and use the stereotypy of a song as a proxy for the RMSE in measuring these scaling relationships.

Discussion

Precisely timed spikes have been hypothesized to somehow encode information^{25–34}. Here, we explored the hypothesis that spike timing precision and spike emission reliability are specifically used to improve the accuracy of decoding a time series, as measured by the root mean squared error (RMSE). The RMSE in decoding any behaviour or sensory stimulus decreases linearly with the network size when using an optimal linear decoder with precisely timed and reliable spikes. This is similar to efficient coding circuits^{15–18,21–23}, or the transition between rate and timing codes considered by others³⁴. We found that this linear decrease in RMSE was robust to different supervisors, and different distributions of spikes, so long as the spikes were precise and reliable, and sufficiently smooth. If the standard deviation in the spike times remained fixed with network size, the linear decrease in the RMSE with network size would be destroyed. If, however, spike emission reliability and spike timing precision

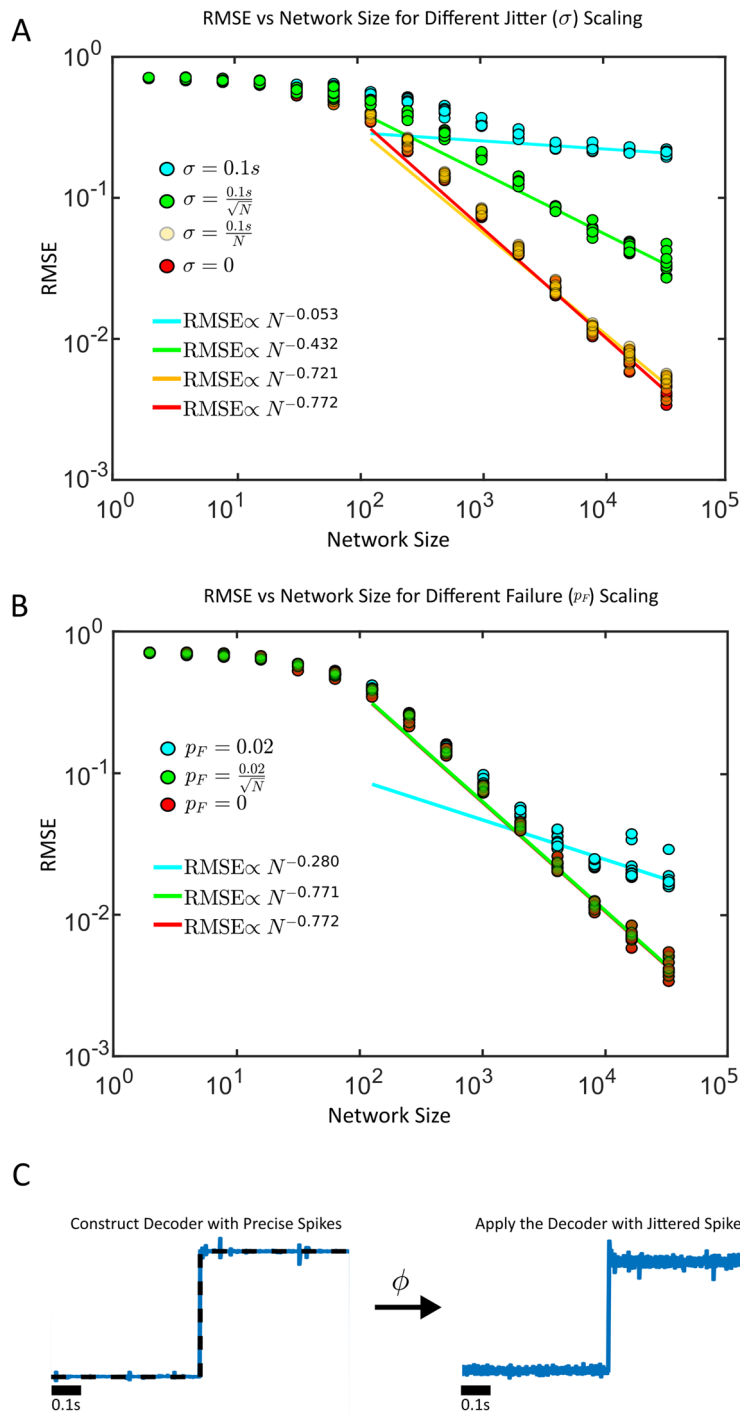


Figure 5. (A) Imprecise but reliable spikes are used to decode a nonsmooth function $f(t) = \text{sign}(t - 0.5)$. The root mean squared error is plotted versus the network size. The spike jitter amount decreases linearly with N (yellow), with the square-root of N (green), or stays constant at $\sigma = 0.1$ s (teal), or stays constant at $\sigma = 0$ (red). (B) Precisely timed but unreliable spikes with a probability of spike failure p_F are used to decode the sign function. The spike failure probability either decreases with the square root of the network size (green), or is fixed at $p_F = 0.02$ (teal), or is fixed at $p_F = 0$ (red). (C) The decoded sign function signal at $N = 2^{15}$ with $\sigma \propto \sqrt{N}^{-1}$ (right). The jitterless signal (blue) and the supervisor (black) are shown in the left.

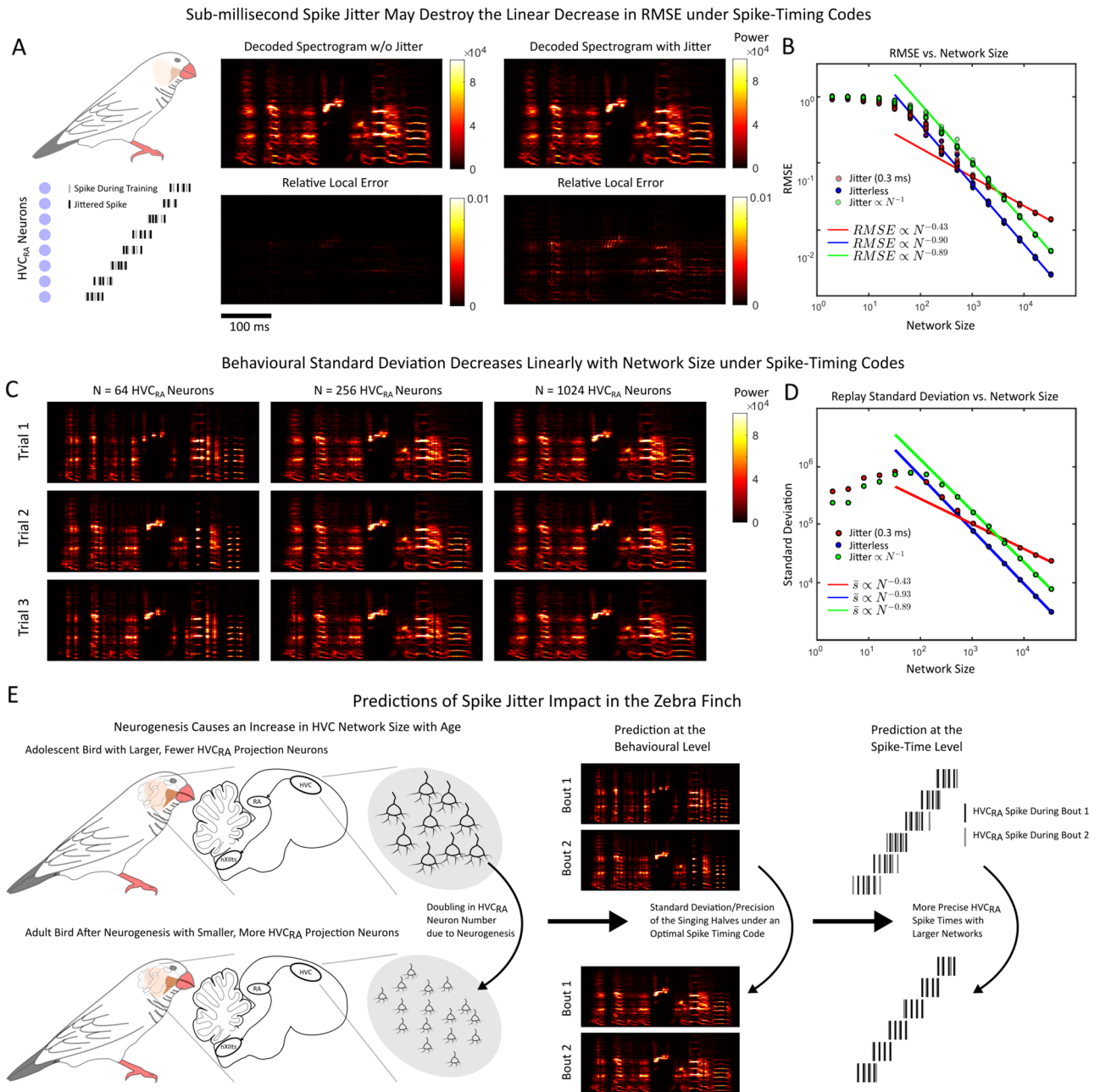


Figure 6. (A) Synthetic HVC_{RA} synfire chain spike sequences are created and used to train a decoder to decode out a sample spectrogram. These spikes are generated as bursts of 4 spikes per cell. The time the initial spike in a burst is randomly sampled from a uniform distribution on [0, 0.88]s. These spikes are then jittered with varying amounts of jitter with the decoder reapplied and the RMSE measured. (B) Root Mean Squared Error (RMSE) between the song spectrogram and the decoded output as a function of the network size (N) with varying amounts of jitter. Sub-millisecond jitter can destroy the $\approx N^{-1}$ scaling in the RMSE. The standard deviation of the jitter was taken to be $\sigma = 0$ (blue), $\sigma = 0.3$ ms (red), and $\sigma = \frac{15}{N}$ (green). (C) Sample decoded spectrograms for different values of the network size for the fixed jitter $\sigma = 0.3$ ms case. (D) The spectrogram variance $\bar{\sigma}^2$ (see Materials and Methods) displays identical scaling relationships as the RMSE. The standard deviation of the jitter was taken to be $\sigma = 0$ (blue), $\sigma = 0.3$ ms (red), and $\sigma = \frac{15}{N}$ (green). (E) The prediction of linear RMSE scaling in the zebra finch circuit. An increase in network size (in this case a doubling) due to neurogenesis leads to a proportional increase in behavioural stereotypy/precision (or decrease in the behavioural standard deviation), and an increase in the precision of spike times in HVC_{RA} neurons.

increased with the network size, the linear decrease in RMSE with network size would be retained. The standard deviation of the spike jitter had to decrease linearly with network size while the probability of spike failure had to decrease with the square-root of the network size for a linear decrease in the RMSE with network size to be maintained.

The sub-linear increase in the RMSE with network size can in many of the cases considered here, be attributed to the variance in the bias-variance decomposition of the RMSE. In both analytical results, and numerical simulations, we found that for smooth supervisors, the bias term in the RMSE would decrease linearly with the network size, while the variance would decrease sub-linearly unless the variance could be controlled to also decrease with square of the network size by firing spikes more reliably and more precisely. We remark that as the bias and the variance can scale at different rates, accurate empirical estimates of the asymptotic scaling relationship in the RMSE may require larger network sizes than those considered here.

The numerical and analytical exploration of the coding accuracy of a stimulus considered here can inform reservoir computing approaches with spikes^{14,35–37}, where the learning is primarily in the readout weights, and more generally, neuromorphic computing, where spiking neurons are used in hardware implementations of neural circuits^{3,4}. Our results show that there are diminishing returns in adding either more spikes or more neurons to improve decoding, unless the spike timing reliability and spike failure rates can be suitably controlled with larger network sizes. These results show that hardware resources should be dedicated to improving the precision of spikes or their reliability, instead of increasing the network size for networks that are already large. This is in line with recent results that examine network size versus error rates in non-spiking networks³⁹.

The zebra finch is well positioned to test hypotheses about how the precision in reproducing behaviours is determined by the number of neurons controlling said behaviour. First, the circuit controls a stereotyped behaviour that is readily elicited and easily recorded with a microphone in head-fixed animals. Second, new HVC_{RA} neurons naturally form through the process of neurogenesis^{40–45} in HVC, and roughly double in number from an average of around 40,000 HVC_{RA} neurons in the first year of life, to around 80,000 HVC_{RA} neurons by year 11 (see Fig. 2 in⁴⁰, Fig. 6). Our work predicts that singing bouts will become more regular or equivalently, decrease in their variance linearly with the size of the network if the spikes also increase in precision and reliability. The precise trend comparing the behavioural precision with the network size of HVC_{RA} neurons has not been currently ascertained. However, there are a pair of studies that separately imply that the total number of HVC_{RA} neurons increases with age⁴⁰ and that the singing precision also increases with age⁴⁴. However, the authors in⁴⁴ postulated that the increase in behavioural precision was inversely related to the rate of neurogenesis, rather than the overall number of HVC_{RA} neurons.

As an alternative to naturally allowing the HVC_{RA} neurons to increase in number with neurogenesis, one can selectively inhibit HVC_{RA} neurons with optogenetics or ablate a fixed proportion of the HVC nucleus entirely¹¹. This would deactivate a random subset of HVC_{RA} neurons. After the animal is given a suitable amount of time to recover and relearn from the deactivation of said neurons, the resulting circuit should produce the identical song, but with fewer HVC_{RA} neurons. This recovery period should also be smaller than the (very slow) timescale of HVC_{RA} neurogenesis. Here, we would expect the behavioural precision to decrease linearly with the proportion of neurons deactivated. We note that selective perturbations in efficient codes were also studied in¹⁸ (see Figs. 4, 5, 6), and the impacts on the bounding box of an efficient code. We also note that it is possible that a trained decoder would be able to predict the resulting changes to the song after perturbation, or potentially predict the natural fluctuations of the song with a sufficiently large training sample.

Further, we remark that precise and reliably emitted spikes are sufficient to generate a linear decrease of the RMSE with network size, however, this is not the only possibility. In fact, linear error scaling was previously predicted under efficient spike time coding schemes^{15–23}. Here, neurons explicitly code the error in the representation of a stimulus, behaviour, or internal dynamic state with their voltages. When the error reaches a critical threshold, a neuron fires a spike to explicitly reset the error to 0. The end result is a network that also produces a linear scaling of the RMSE with network size. However, owing to the explicit error correction mechanism in these circuits, the spike times for individual neurons are not precisely timed across multiple trials, but are precisely timed to reset the error in a representation^{15–17,21}. Here, we suggest an alternate mechanism for N^{-1} error scaling in a circuit: stabilize the spike times and spike emission reliability for larger networks.

This theoretical work demonstrates the sufficient conditions for a neural circuit to qualitatively improve the decoding of a stimulus or behaviour with precisely timed and reliably emitted spikes. Under sufficiently precise and reliable spiking, a network can double its size to double the accuracy of encoding/decoding a stimulus. However, said network must also double the precision of all spikes in the network, and increase the reliability with which spikes are emitted, but by less than double the network size.

Materials and methods

Generating synthetic spike trains

Each network considered here constitutes N synthetically generated spike trains, with each spike train representing a single neuron. Thus, the synthetic spike times, t_{ji} which corresponds to the i th spike fired by the j th neuron that could have their spike timing precision and spike emission reliability controlled. These spikes were then filtered with a synaptic filter:

$$r_j(t) = \sum_{t_{ji} < t} K(t - t_{ji}). \quad (10)$$

The filtering or Kernel function, $K(t)$ was taken to be a single-exponential filter for all numerical simulations:

$$K(t) = \exp\left(-\frac{t}{\tau_s}\right) \quad (11)$$

where $\tau_s = 10$ ms was used as the synaptic filter, approximating the time constant of AMPA synapses. The generation of spike times t_{ji} is described in greater detail below.

HVC_{RA} spike train

Each HVC_{RA} neuron consisted of a synthetically generated spike train consisting of 4 spikes in an isolated burst. The spikes were separated by 3 ms inter-spike-intervals, with the initial spike time drawn from a uniform distribution on the interval $[0, T]$, where $T = 0.88$ seconds is the duration of the song-recording. To vary the network size N , the networks were successively doubled in size from $N = 2^1$ to $N = 2^{15} = 32768$. The networks in Fig. 6 were simulated 10 times for a fixed N , with different seeds in each network, thereby generating different starting times for each HVC_{RA} burst.

To jitter the spikes, each spike time t_{ji} was randomly perturbed by ϵ_{ji} , where ϵ_{ji} was a normally distributed random variable with mean 0, and standard deviation, σ . The values of σ vary within figures with perfectly precise spikes ($\sigma = 0$ milliseconds), spikes with sub-millisecond precision ($\sigma = 0.3$ milliseconds, Fig. 6), and spikes that become increasingly precise with larger network sizes ($\sigma \propto \frac{1}{N}$).

Poisson generated spike trains

The randomly generated spikes in Figs. 1, 2, 3, 4 and 5 were generated from a Poisson process with firing rate $\nu = 2$ Hz, for successively larger networks which were doubled in size from $N = 2^1$ to $N = 2^{14}$. As in the song-bird example, the spikes were perturbed in time with a normally distributed random variable with mean 0 and standard deviation σ . To implement spike-failure, decoders were first constructed with all spikes generated. These decoders remained fixed, even after the spike times or spike reliability is altered. Then, the fixed decoders were applied to the same spike trains but with each spike having a probability of p_F of failure. To implement spike interference, spikes were added randomly to neurons in time with probability p_I . The total number of spikes added was $p_I n_{spikes}$ where n_{spikes} was the number of spikes generated in the nominal, reliable spiking case. The added spikes were randomly distributed from a uniform distribution in time, and across neurons.

Constructing linear decoders

Linear decoders were constructed by first generating filtered spike-trains that contained no jitter or failure. Then the solution to the optimal linear decoder for a given signal $x(t)$ is:

$$\phi^x = \left(\int_0^T \mathbf{r}(t)\mathbf{r}(t)^T dt \right)^{-1} \int_0^T \mathbf{r}(t)x(t) dt \quad (12)$$

The decoder can then be applied to new spike trains to decode $\hat{x}(t)$ when these new spikes display either imprecision in their spike times or the failure of spike emission relative to the initial spike train used to construct ϕ^x . For the simple sinusoidal example, $x(t) = \sin(2\pi t)$, with the approximation $\hat{x}(t)$ given by:

$$\hat{x}(t) = \sum_{j=1}^N \phi_j r_j(t), \quad (13)$$

while for the spectrogram, each frequency component of the song has its own decoder:

$$\phi^{p(f)} = \left(\int_0^T \mathbf{r}(t)\mathbf{r}(t)^T dt \right)^{-1} \int_0^T \mathbf{r}(t)p(f, t) dt \quad (14)$$

The term $p(f, t)$ is the power of the frequency component f at time t in the spectrogram, as defined in¹⁴, while the decoder component $\phi^{p(f)}$ is the optimal decoder for the power frequency component $p(f, t)$, yielding the approximation:

$$\hat{p}(f, t) = \sum_{j=1}^N \phi_j^{p(f)} r_j(t)$$

The frequency range considered varies from a low of $f = 172.27$ Hz to a high of $f = 10$ kHz, discretized with 229 evenly distributed points, thereby making $\phi^{p(f)}$ a $N \times 229$ dimensional decoder.

Measuring the root mean-squared error

The RMSE for the decoded oscillators in Figs. 2F–H and 3 is given by:

$$RMSE = \sqrt{\int_0^1 (\hat{x}(t) - x(t))^2 dt} \quad (15)$$

where $x(t) = \sin(2\pi t)$, and $\hat{x}(t)$ is the decoded approximation to $x(t)$.

The RMSE for the decoded spectrogram is computed with:

$$RMSE = \sqrt{\int_0^F \int_0^T (\hat{p}(f, t) - p(f, t))^2 dt df} \quad (16)$$

where $p(f, t)$ is the value of the song spectrogram at frequency f and time t over the range $f = 0$ to $F = 10$ kHz and from time $t = 0$ to $T = 0.88$ s (song duration). For each value of N , 10 trials (corresponding to different decoders) were used to obtain the average value of the RMSE for that fixed value of N . The linear fits to the RMSE were constructed on the loglog scale with the MATLAB *polyfit* function, for sufficiently large networks ($N \geq 2^{10}$).

Measuring the stereotypy in the zebra finch networks

To estimate the variability from song-bout to song-bout, we computed the integrated variance, $\bar{\sigma}_{SB}$:

$$\bar{s}^2 = \int_F \int_0^T \sigma_{SB}^2(f, t) dt df \quad (17)$$

where

$$\sigma_{SB}^2(f, t) = \sum_{k=1}^m (p_k(f, t) - \langle p_k(f, t) \rangle)^2 \quad (18)$$

$$\langle p_k(f, t) \rangle = \frac{1}{m} \sum_{k=1}^m p_k(f, t) \quad (19)$$

where $\hat{p}_k(f, t)$ denotes the k th trial's decoded spectrogram for $k = 1, 2, \dots, 10$.

Time discretization

To simulate the filtered spike-trains and decoded signal over a time interval $[0, T]$ we discretize time into M evenly distributed points (T_1, \dots, T_M) with step size Δt , which we took as $\Delta t = 10^{-5}$ ms, and $T_k = (k\Delta)$ and $M = T/\Delta t$. We then have a corresponding discretized filtered spike-train vectors (r^1, \dots, r^M) which are defined by:

$$r_j^k = \sum_{t_{ji} < T_k} K(T_k - t_{ji}) \quad (20)$$

where K is the filtering function and r_j^k is the j^{th} neuron's filtered spike-train at time step k and t_{ji} is the i th spike fired by the j th neuron. We then generate the decoded output for each time step T_k by the vector multiplication $\hat{x} = (\phi^x)^T r^k$. We can improve the efficiency of this calculation for the exponential filter

$$K(t) = \exp\left(-\frac{t}{\tau_s}\right) \quad (21)$$

using the exponentiation identity $e^{x+y} = e^x e^y$. If for neuron j , no spikes occur in the time interval $[T_{k-1}, T_k]$ then we have the update

$$r_j^k = r_j^{k-1} \exp\left(-\frac{\Delta t}{\tau_s}\right) \quad (22)$$

and if a spike does occur in the interval $[T_{k-1}, T_k]$ at time t_{ji} then we have the update

$$r_j^k = \exp\left(-\frac{T_k - t_{ji}}{\tau_s}\right) + r_j^{k-1} \exp\left(-\frac{\Delta t}{\tau_s}\right) \quad (23)$$

Thus at each time step we do not need to compute the sum over all previous spikes and can instead perform a single update. For Fig. 6, the spike interpolation scheme in between time steps in Eq. 23 was not used, and a basic time step of $\Delta t = 10^{-5}$ was utilized.

Training/Testing trials

For Figs. 1, 2, 4 and 5, Supplementary Figs. S3–S4, and (6 songbird) a single trial of spikes was used to train the optimal linear decoders ϕx , while for Fig. 3, a total of 10 trials were used to train the optimal linear decoders. The test RMSE for Fig. 3 used 10 trials for 10 different random network seeds (spike trains). The test RMSE for Figs. 1, 2, 4 and 5, and Supplementary Figs. S3–S4, 10 different network seeds were used with 1 trial to estimate the test RMSE. For all simulated networks with the exception of the songbird networks (6), the trials used for training/testing were appended consecutively in a large matrix for $r(t)$. To avoid the initial transients for $r(t) = 0$, an initial buffer trial of spikes was added. This buffer trial would be used to initialize the filtered spike trains $r(t)$ but would not be used to train the optimal linear decoders.

Data availability

A hyperlink to the songbird spectrogram can be found on the modelDB⁴⁶ link above, with the raw recording from⁴⁷. The raw recording was processed as in¹⁴.

Code availability

The relevant code to generate the results of this paper can be found on modelDB⁴⁶, under Accession no. 2015409 (direct hyperlink <https://modeldb.science/2015409>). The read-only access code for referees is ro1331. The accompanying code on modelDB will be released to the public after publication of this manuscript.

Received: 24 March 2023; Accepted: 1 April 2024

Published online: 08 May 2024

References

- Brady, S., Siegel, G., Wayne Albers, R. & Price, D. L. *Basic Neurochemistry: Molecular, Cellular and Medical Aspects* (Elsevier, 2005).
- Howarth, C., Gleeson, P. & Attwell, D. Updated energy budgets for neural computation in the neocortex and cerebellum. *J. Cerebral Blood Flow Metab.* **32**(7), 1222–1232 (2012).
- Paugam-Moisy, H. & Bohte, S. M. Computing with spiking neuron networks. *Handbook of Natural Computing* **1**, 1–47 (2012).
- Davies, M. *et al.* Advancing neuromorphic computing with loihi: A survey of results and outlook. *Proc. IEEE* **109**(5), 911–934 (2021).
- Nottebohm, F., Stokes, T. M. & Leonard, C. M. Central control of song in the canary, *Serinus canarius*. *J. Compar. Neurol.* **165**(4), 457–486 (1976).
- Mooney, R. Different subthreshold mechanisms underlie song selectivity in identified HVC neurons of the zebra finch. *J. Neurosci.* **20**(14), 5420–5436 (2000).
- Long, M. A., Jin, D. Z. & Fee, M. S. Support for a synaptic chain model of neuronal sequence generation. *Nature* **468**(7322), 394–399 (2010).
- Leonardo, A. & Fee, M. S. Ensemble coding of vocal control in birdsong. *J. Neurosci.* **25**(3), 652–661 (2005).
- Kozhevnikov, A. A. & Fee, M. S. Singing-related activity of identified HVC neurons in the zebra finch. *J. Neurophysiol.* **97**(6), 4271–4283 (2007).
- Hahnloser, R. H. R., Kozhevnikov, A. A. & Fee, M. S. An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature* **419**(6902), 65 (2002).
- Roberts, T. F., Gobes, S. M. H., Murugan, M., P Ölveczky, B. & Mooney, R. Motor circuits are required to encode a sensory model for imitative learning. *Nat. Neurosci.* **15**(10), 1454–1459 (2012).
- Picardo, M. A. *et al.* Population-level representation of a temporal sequence underlying song production in the zebra finch. *Neuron* **90**(4), 866–876 (2016).
- Barlow, H. *et al.* Possible principles underlying the transformation of sensory messages. *Sensory Commun.* **1**(01), 217–233 (1961).
- Nicola, W. & Clopath, C. Supervised learning in spiking neural networks with force training. *Nat. Commun.* **8**(1), 1–15 (2017).
- Denève, S. & Machens, C. K. Efficient codes and balanced networks. *Nat. Neurosci.* **19**(3), 375 (2016).
- Boerlin, M., Machens, C. K. & Denève, S. Predictive coding of dynamical variables in balanced spiking networks. *PLoS Comput. Biol.* **9**(11), e1003258 (2013).
- Schwemmer, M. A., Fairhall, A. L., Denève, S. & Shea-Brown, E. T. Constructing precisely computing networks with biophysical spiking neurons. *J. Neurosci.* **35**(28), 10112–10134 (2015).
- Calaim, N., Dehmelt, F. A., Gonçalves, P. J. & Machens, C. K. The geometry of robustness in spiking neural networks. *Elife* **11**, e73276 (2022).
- Gutierrez, G. J. & Denève, S. Population adaptation in efficient balanced networks. *Elife* **8**, e46926 (2019).
- Zeldenrust, F., Gutkin, B. & Denève, S. Efficient and robust coding in heterogeneous recurrent networks. *PLoS Comput. Biol.* **17**(4), e1008673 (2021).
- Brendel, W., Bourdoukan, R., Verterchi, P., Machens, C. K. & Denève, S. Learning to represent signals spike by spike. *PLoS Comput. Biol.* **16**(3), e1007692 (2020).
- Bourdoukan, R., & Deneve, S. Enforcing balance allows local supervised learning in spiking recurrent networks. *Adv. Neural Inf. Process. Syst.* **28** (2015)
- Ralph Bourdoukan, David Barrett, Sophie Deneve, and Christian K Machens. Learning optimal spike-based representations. *Advances in neural information processing systems*, 25, 2012.
- Allen, C. & Stevens, C. F. An evaluation of causes for unreliability of synaptic transmission. *Proc. Natl. Acad. Sci.* **91**(22), 10380–10383 (1994).
- Bohte, S. M. The evidence for neural information processing with precise spike-times: A survey. *Nat. Comput.* **3**(2), 195–206 (2004).
- VanRullen, R., Guyonnet, R. & Thorpe, S. J. Spike times make sense. *Trends Neurosci.* **28**(1), 1–4 (2005).
- Panzeri, S., Petersen, R. S., Schultz, S. R., Lebedev, M. & Diamond, M. E. The role of spike timing in the coding of stimulus location in rat somatosensory cortex. *Neuron* **29**(3), 769–777 (2001).
- Gutkin, B., Ermentrout, G. B. & Rudolph, M. Spike generating dynamics and the conditions for spike-time precision in cortical neurons. *J. Comput. Neurosci.* **15**(1), 91–103 (2003).
- Borst, A. & Theunissen, F. E. Information theory and neural coding. *Nat. Neurosci.* **2**(11), 947–957 (1999).
- Gütig, R. To spike, or when to spike?. *Curr. Opin. Neurobiol.* **25**, 134–139 (2014).
- Tully, P. J., Lindén, H., Hennig, M. H. & Lansner, A. Spike-based Bayesian-Hebbian learning of temporal sequences. *PLoS Comput. Biol.* **12**(5), e1004954 (2016).
- Quiñero, R. & Panzeri, S. Extracting information from neuronal populations: Information theory and decoding approaches. *Nat. Rev. Neurosci.* **10**(3), 173–185 (2009).
- Victor, J. D. & Purpura, K. P. Nature and precision of temporal coding in visual cortex: A metric-space analysis. *J. Neurophysiol.* **76**(2), 1310–1326 (1996).
- Theunissen, F. & Miller, J. P. Temporal encoding in nervous systems: A rigorous definition. *J. Comput. Neurosci.* **2**, 149–162 (1995).
- Nicola, W. & Clopath, C. A diversity of interneurons and Hebbian plasticity facilitate rapid compressible learning in the hippocampus. *Nat. Neurosci.* **22**(7), 1168–1181 (2019).
- DePasquale, B., Cueva, C. J., Rajan, K., Escola, G. S. & Abbott, L. F. full-force: A target-based method for training recurrent networks. *PLoS ONE* **13**(2), e0191527 (2018).
- Paugam-Moisy, H., Martinez, R. & Bengio, S. Delay learning and polychronization for reservoir computing. *Neurocomputing* **71**(7–9), 1143–1158 (2008).
- Paugam-Moisy, H. & Bohte, S. M. Computing with spiking neuron networks. *Handbook of Natural Computing* **1**, 1–47 (2012).
- Raman, D. V., Rotondo, A. P. & O’Leary, T. Fundamental bounds on learning performance in neural circuits. *Proc. Natl. Acad. Sci.* **116**(21), 10537–10546 (2019).
- Walton, C., Pariser, E. & Nottebohm, F. The zebra finch paradox: Song is little changed, but number of neurons doubles. *J. Neurosci.* **32**(3), 761–774 (2012).
- Kirn, J. R. & Nottebohm, F. Direct evidence for loss and replacement of projection neurons in adult canary brain. *J. Neurosci.* **13**(4), 1654–1663 (1993).

42. Pytte, C. L. *et al.* Adult neurogenesis is associated with the maintenance of a stereotyped, learned motor behavior. *J. Neurosci.* **32**(20), 7052–7057 (2012).
43. Brenowitz, E. A. & Larson, T. A. Neurogenesis in the adult avian song-control system. *Cold Spring Harb. Perspect. Biol.* **7**(6), a019000 (2015).
44. Pytte, C. L., Gerson, M., Miller, J. & Kirn, J. R. Increasing stereotypy in adult zebra finch song correlates with a declining rate of adult neurogenesis. *Dev. Neurobiol.* **67**(13), 1699–1720 (2007).
45. Nordeen, K. W. & Nordeen, E. J. Projection neurons within a vocal motor pathway are born during song learning in zebra finches. *Nature* **334**(6178), 149–151 (1988).
46. McDougal, R. A. *et al.* Twenty years of modeldb and beyond: Building essential modeling tools for the future of neuroscience. *J. Comput. Neurosci.* **42**, 1–10 (2017).
47. Crandall, S. R. & Nick, T. A. Neural population spiking activity during singing: adult and longitudinal developmental recordings in the zebra finch. *CRCNS. org* <https://doi.org/10.6080/KONP22C8> (2014).

Author contributions

W.N. and C.C. and T.R.N. wrote the main manuscript. W.N. and T.R.N. performed the simulations. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-58524-7>.

Correspondence and requests for materials should be addressed to W.N.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024