

Article

A Time-Varying Information Measure for Tracking Dynamics of Neural Codes in a Neural Ensemble

Mohammad R. Rezaei ^{1,2,3}, Milos R. Popovic ^{2,3} and Milad Lankarany ^{1,2,3,*}

¹ Division of Clinical and Computational Neuroscience, Krembil Research Institute, University Health Network, Toronto, ON M5T 0S8, Canada; m.reza.rezaei72@gmail.com

² KITE Research Institute, Toronto Rehabilitation Institute-University Health Network, Toronto, ON M5G 2A2, Canada; milos.popovic@uhn.ca

³ Institute of Biomedical Engineering, University of Toronto, Toronto, ON M5S 3G9, Canada

* Correspondence: milad.lankarany@uhnresearch.ca

Received: 24 May 2020; Accepted: 6 August 2020; Published: 11 August 2020



Abstract: The amount of information that differentially correlated spikes in a neural ensemble carry is not the same; the information of different types of spikes is associated with different features of the stimulus. By calculating a neural ensemble's information in response to a mixed stimulus comprising slow and fast signals, we show that the entropy of synchronous and asynchronous spikes are different, and their probability distributions are distinctively separable. We further show that these spikes carry a different amount of information. We propose a time-varying entropy (TVE) measure to track the dynamics of a neural code in an ensemble of neurons at each time bin. By applying the TVE to a multiplexed code, we show that synchronous and asynchronous spikes carry information in different time scales. Finally, a decoder based on the Kalman filtering approach is developed to reconstruct the stimulus from the spikes. We demonstrate that slow and fast features of the stimulus can be entirely reconstructed when this decoder is applied to asynchronous and synchronous spikes, respectively. The significance of this work is that the TVE can identify different types of information (for example, corresponding to synchronous and asynchronous spikes) that might simultaneously exist in a neural code.

Keywords: information-theoretic measure; rate code; temporal code; multiplexed code; asynchronous and synchronous spikes

1. Introduction

The collective responses of primary sensory neurons constitute fully or partly mixed inputs to cortical neurons; thus, multiple features of the stimulus are to be reliably coded by cortical neurons. The brain uses different coding strategies to represent information underlying those features. Information can be encoded by either the rate of spikes in a relatively long time window—rate code—or by their precise timing—temporal code [1–9]. In temporal coding, information is mostly carried by groups of neurons that fire nearly simultaneously [9,10] (see [11] for other forms of temporal coding) whereas, in the rate coding, the precise timing of spikes is compromised and information across neurons is mostly carried by the rate of asynchronous spikes [12–15].

It has been suggested that the presence of both coding strategies, whose feasibility was demonstrated in different neural systems (e.g., [16,17]), can be considered as a unique way to convey multiple features of the stimulus, i.e., multiplexed coding. In fact, in addition to the rate code, which is largely observed in different neural systems, inter-neuronal correlations within many areas of the brain have a significant functional role in the neural code [18–20]. Temporal correlations between neurons can contribute to coding additional information, which are not represented by the isolated

spike trains. However, it remained unknown to what extent these coding strategies cooperatively contribute to the representation of a mixed stimulus. It is challenging to uncover the distinct roles of differentially correlated spikes—i.e., asynchronous spikes (rate code) and synchronous spikes (temporal code)—in a multiplexed code. To address this challenge, it is crucial to measure the information underlying different types of spikes [16]. Various information-theoretic techniques have been exploited to measure information carried by differentially correlated spikes [15,21]. These methods can be classified into two categories, namely, direct and indirect approaches [19]. In the indirect approach, information is calculated based on the relationship between the stimulus and neural responses. In contrast, in the direct approach, information is obtained based on the statistics of the neural responses, regardless of any assumptions about the stimulus [19]. For example, in the indirect approach, mutual information (MI) between the stimulus and spikes [22–24] is calculated based on their joint probability distribution [23,25,26]. Thus, the computational complexity is expensive. Although some methods, like non-parametric kernel estimation [27] and Gaussian approximation [28], are used to reduce the computational complexity of calculating the joint distribution, the indirect approaches are not sufficiently accurate when applied to multi-dimensional neural activity, i.e., spikes in a neural ensemble [29]. In contrast, information can be measured with less complexity in the direct approaches.

It is worth mentioning that information, in almost all the existing methods, is calculated by the length of the stimulus interval. Nevertheless, the information of a mixed stimulus might be represented by spikes at different time scales. The rate-modulated asynchronous spikes and precisely correlated synchronous spikes occur in different timescales. Thereby, to calculate information in a multiplexed code, it is important to calculate information over time (at each time bin) and in different time scales. In this paper, we propose a time-varying entropy (TVE) measure that calculates the entropy of a neural ensemble in response to a mixed stimulus consisting of slow and fast signals. The simultaneous representation of these signals through synchronous and asynchronous spikes was recently demonstrated [16]. Inspired by [19], we consider spikes as code words with different lengths and time resolutions, and calculate the entropy across homogeneous neurons in a neural ensemble. In this way, we estimate the entropy of spikes at each time bin and show how it varies in different time resolutions, which correspond to different features of the stimulus. Furthermore, by computing the probability distributions and entropies of asynchronous and synchronous spikes in a neural ensemble, we show that these spikes carry different information. The TVEs underlying synchronous and asynchronous spikes have their maximum values when the code words are selected with specific time resolutions. In addition, we demonstrate that the TVEs of synchronous and asynchronous spikes are highly correlated with the fast and slow signals, respectively. Finally, we use a Kalman decoder to reconstruct these features of the stimulus using asynchronous and synchronous spikes. Our results indicate that information underlying synchronous and asynchronous spikes is different and associated with distinct features of the stimulus.

2. Computational Framework

2.1. Responses of a Homogeneous Neural Ensemble to a Mixed Stimulus

According to the feasibility of neural systems to multiplexed coding [16,30], we simulated the activity of a homogeneous neural ensemble in response to a mixed stimulus to explore how much information can be encoded by different patterns of spikes. Each neuron received a mixed signal (I_{mixed}), which consists of a fast signal (I_{fast}) and a slow signal (I_{slow}) [16]. I_{fast} stands for the timing of fast events or abrupt changes in the stimulus and was generated by convolving a randomly (Poisson) distributed Dirac-delta function with a synaptic waveform (normalized to the peak amplitude), $\tau_{rise} = 0.5$ ms, and $\tau_{fall} = 3$ ms. Fast events occurred at a rate of ~ 1 Hz and were scaled by $a_{fast} = 85$ pA.

I_{slow} was generated by an OU process as follows

$$\frac{dI_{slow}}{dt} = -\frac{I_{slow}(t) - \mu}{\tau} + \sigma \frac{\sqrt{2}}{\tau} \xi(t), \quad \xi \sim N(0, 1) \quad (1)$$

where ξ is a random number drawn from a Gaussian distribution, $\tau = 100$ ms is the time constant of the slow signal that produces a slow-varying random walk with an average of $\mu = 15$ pA and a standard deviation of $\sigma = 60$ pA. The mixed signal (I_{mixed}) was obtained by adding I_{fast} and I_{slow} , which are generated independently.

An independent noise (equivalent to the background synaptic activity) was added to each neuron, thus each neuron receives a mixed signal plus noise. Similar to [31], the noise (I_{noise}) was generated by an OU process of $\tau = 5$ ms, $\mu = 0$ pA, and $\sigma = 10$ pA.

The neural ensemble consists of 100 neurons, each of them was modeled by Morris–Lecar equations [32,33]. The equations of a single model neuron receiving a mixed-signal plus noise can be written as follows

$$C \frac{dV}{dt} = I_{mixed}(t) + I_{noise}(t) - \bar{g}_{Na} m_{\infty}(V)(V - E_{Na}) - \bar{g}_K w(V - E_K) - g_L(V - E_L) - \bar{g}_{AHP} z(V - E_K) - g_{exc}(V - E_{exc}) - g_{inh}(V - E_{inh}) \quad (2)$$

where,

$$\frac{dw}{dt} = \phi \frac{w(V) - w}{\tau_{W(V)}} \quad (3)$$

$$\frac{dz}{dt} = \frac{1}{\tau_z} \frac{1 - z}{1 + e^{(\beta_z - V)/\gamma}} \quad (4)$$

$$m_{\infty}(V) = 0.5 \left[1 + \tanh\left(\frac{V - \beta_m}{\gamma_m}\right) \right] \quad (5)$$

$$w_{\infty}(V) = 0.5 \left[1 + \tanh\left(\frac{V - \beta_w}{\gamma_w}\right) \right] \quad (6)$$

$$\tau_w(V) = \frac{1}{\cosh\left(\frac{V - \beta_w}{2\beta_w}\right)} \quad (7)$$

where $\{\bar{g}_{Na} = 20, \bar{g}_k = 20, \bar{g}_L = 20, \bar{g}_{AHP} = 25, g_{exc} = 1.2, g_{inh} = 1.9\} \frac{\text{mS}}{\text{cm}^2}$, $\{E_{Na} = 50, E_K = -100, E_L = -70, E_{exc} = 0, E_{inh} = -70\}$ mV, $\beta_m = -1.2, \gamma_m = 18, \beta_w = -19, \gamma_w = 10, \beta_z = 0, \gamma_z = 2, \tau_a = 20$ ms, $\phi = 0.15$, and $C = 2 \frac{\mu\text{F}}{\text{cm}^2}$. These parameters were set to ensure a neuron operates in a hybrid mode [34], i.e., an operating mode between integration and coincidence detection [35]. The inclusion of background excitatory and inhibitory synaptic conductance in (2) reproduced a “balanced” high-conductance state [36]. The surface area of the neuron was set to $200 \mu\text{m}^2$, so that I_{mixed} is reported in pA, rather than as a density. Figure 1 shows different steps towards constructing the mixed signal and stimulating a neural ensemble. Figure 1A shows how the mixed signal was made by I_{slow} and I_{fast} signals. The spiking activity of the neural ensemble is shown in Figure 1B. Similar to [16], synchronous spikes, sync-spikes, and asynchronous spikes, async-spikes, were distinguished based on a synchrony threshold. Therefore, the dataset consists of the mixed-stimulus and the spiking activities of the neural ensemble where different elements of the mixed stimulus (I_{fast}, I_{slow}) and their related neural activities were shown in different colors.

2.2. Probability Density Estimation

We used a histogram-based method [37] to calculate the probability distributions of spiking activities or word patterns. We considered 100 bins (as 100 is the total number of neurons used in the simulation study) for the construction of the histograms of different types of spikes. For each word pattern, we considered 2^L bins, where L is the length of the word pattern, to construct histograms that include all possibilities. Finally, the histograms are normalized to reach a probability density function.

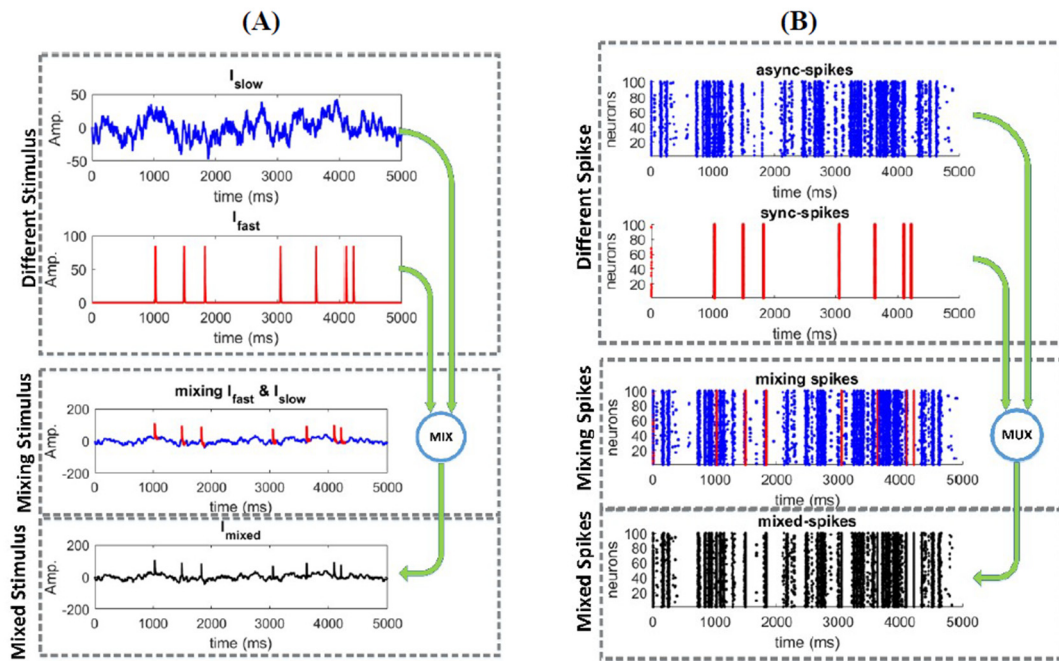


Figure 1. The simulation data consist of the mixed stimulus and spiking activity of the neural ensemble. (A) The mixed stimulus I_{mixed} consists of I_{fast} and I_{slow} (see Section 2) (B) Different patterns of spikes resulted from the neural ensemble (including 100 neurons) as response to a mixed stimulus.

3. Results

3.1. Information Underlying Synchronous and Asynchronous Spikes Are Distinctively Separable

To address whether synchronous and asynchronous spikes convey different information, we test if these spikes are distinctively separable [27]. We use mutual information (MI) to measure the similarities between the probability distributions of these spikes [27,28], i.e., $I(A; S)$, where S and A are random variables drawn from the distributions of synchronous and asynchronous spikes, respectively. The MI can be written as follows [38]

$$I(A; S) = \sum_{a \in A} \sum_{s \in S} p_{(A,S)}(s, a) \log \left(\frac{p_{(A,S)}(s, a)}{P_A(a)P_S(s)} \right) \quad (8)$$

where $P_S(s)$ and $P_A(a)$ are the distributions of synchronous and asynchronous spikes, respectively, and $p_{S,A}(s, a)$ is the joint probability distribution of synchronous and asynchronous spikes. We utilized the histogram-based method suggested in [37] with 100 bins to calculate the probability distributions of each type of spikes. The probability at each bin in the histogram is equal to the number of counts divided by the total number of counts in the histogram. $I(\geq 0)$ is equal to zero if the distributions of sync-spikes ($P_S(s)$) and async-spikes ($P_A(a)$) are independent. To precisely demonstrate the difference between probability distributions of synchronous and asynchronous spikes, we used a non-parametric method to estimate these distributions [39]. This method estimates the probability density function by using normal kernel smoothing function and a bandwidth as follows

$$\hat{f}_h(x) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{x - x_i}{h}\right); -\infty < x < \infty \quad (9)$$

where $\hat{f}_h(x)$ is the approximated histogram, N is the sample size (2×10^5 samples of data in the simulation), $K(\cdot)$ is the kernel function, and h is the bandwidth, which is fixed on 0.4 based on the smoothness of the data.

Figure 2 shows the original and approximated probability distributions of synchronous and asynchronous spikes. The MI between synchronous and asynchronous spikes is nearly zero ($I = 0.003$ and 0.015 for histogram-based and non-parametric methods, respectively), suggesting that the statistical dependencies between their probability distributions are negligible.

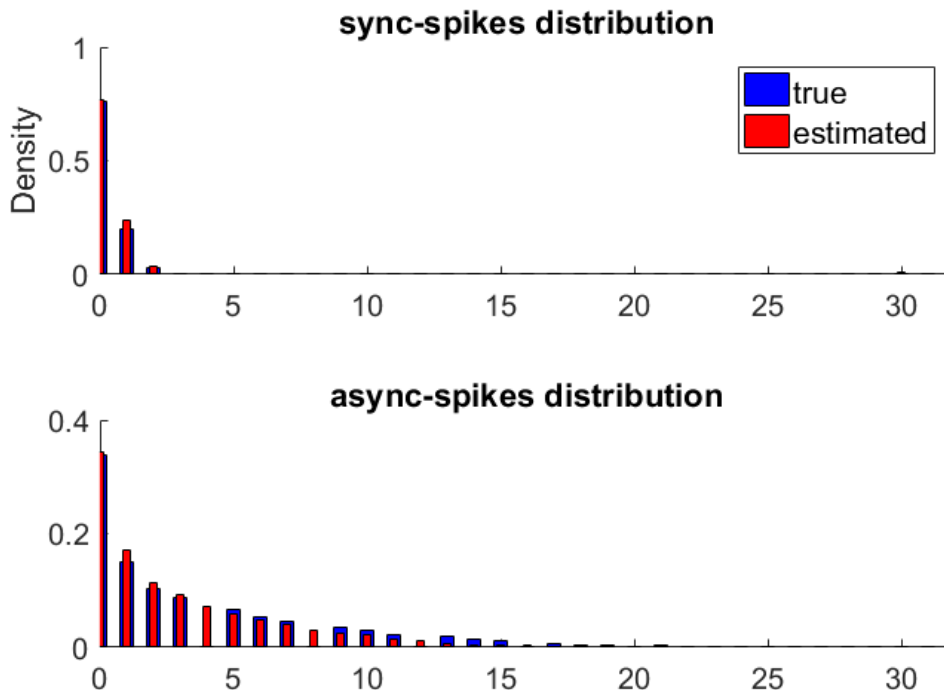


Figure 2. The probability distributions of synchronous (top) and asynchronous (bottom) spikes. For each type of spike, the true distribution was obtained by the histogram method and is shown by blue bars (thick bars). We used a non-parametric method to approximate distributions of synchronous and asynchronous spikes, which are shown by red bars (thin bars).

We also used a statistical hypothesis test to quantify statistical differences between synchronous and asynchronous spikes. A two-sample version of the Kolmogorov test [40,41] was used to detect a wide range of differences between the two distributions. In this way, one can compare the distribution functions of the parent populations of the two samples drawn from the distributions of synchronous and asynchronous spikes. The null hypothesis is that these samples are drawn from an identical distribution function. The statistical test (repeated 1000 times) rejected the null hypothesis at the default significance level of 5% for both histogram-based and non-parametric methods. Our analysis shows that synchronous and asynchronous spikes have different and separable statistical characteristics, which might lead to encoding different types of information.

3.2. Different Types of Spikes in a Multiplexed Code Carry Different Amounts of Information

To quantify the amount of information each type of spike carries, we measure the entropy of synchronous and asynchronous spikes. The entropy determines the variability underlying the probability distributions of spikes [27] and indicates the upper bound of information of spikes. Similar to [42], we considered neural responses as binned spike trains and calculated the entropy of short strings of bins, or words, for each individual neuron. This estimation of entropy depends on two parameters, namely, temporal resolution (δt) and temporal structure or word length (L). The entropy, $H(L, \delta t)$, is defined as follows [19]

$$H(L, \delta t) = -\frac{1}{L\delta t} \sum_{w \in W(L, \delta t)} p(w) \log_2 p(w) \tag{10}$$

where w is a specific word with length L ; $W(L, \delta t)$ is the set of all possible words comprising L bins, and $p(w)$ is the probability of observing a word w in the neural response. The advantage of this method, in comparison to other information measures like the mutual information between stimulus and spikes [27], is that this method is a direct way to estimate information of the spikes, with no need to access the stimulus. After distinguishing synchronous and asynchronous spikes in a neural ensemble (see Section 2), we calculated $H(L, \delta t)$ of each individual neuron for different word lengths (L) and time-bin resolutions (δt) to assess the effect of these variables in extracting information underlying each type of spikes. Figure 3 shows the average of the entropy of individual neurons. The entropy decreases for a low time resolution (i.e., high δt) due to a higher temporal correlation between spikes in a large time bin compared to that in short time bins (see the gradual color contrast vertically in Figure 3A). In addition, the entropy, H , decreases by increasing L due to the integration of dependencies among the time bins, i.e., the longer the word length is, the less uncertain the code word is (see Figure 3B). Given enough data samples for estimating $p(w)$, for an infinite length of L , the entropy can be calculated when $\delta t \rightarrow 0$ and $L \rightarrow \infty$ [19]. However, due to the finite length of data, which leads to a finite length of L , we extrapolated the entropy for an optimum δt (i.e., 0.05 ms) and $L \rightarrow \infty$, and we found a steady-state rate of entropy for different types of spikes (red lines in Figure 3B). As shown in this figure, synchronous and asynchronous spikes convey different rates of entropy (the steady-state rate of entropies for synchronous and asynchronous spikes are about 16.2 and 94.2 bit/s, respectively). In addition, the entropy of all spikes (calculated as the average of the entropy of individual neurons) is about 102 bit/s. The interpretation of an entropy measure of 102 bit/s is that the spiking activity of the neural ensemble can carry as much information as would be required to perfectly discriminate 2^{102} different 1-s-long samples of the stimulus [19]. In the next section, we see if such difference in information measure between synchronous and asynchronous spikes is associated with different features of the stimulus.

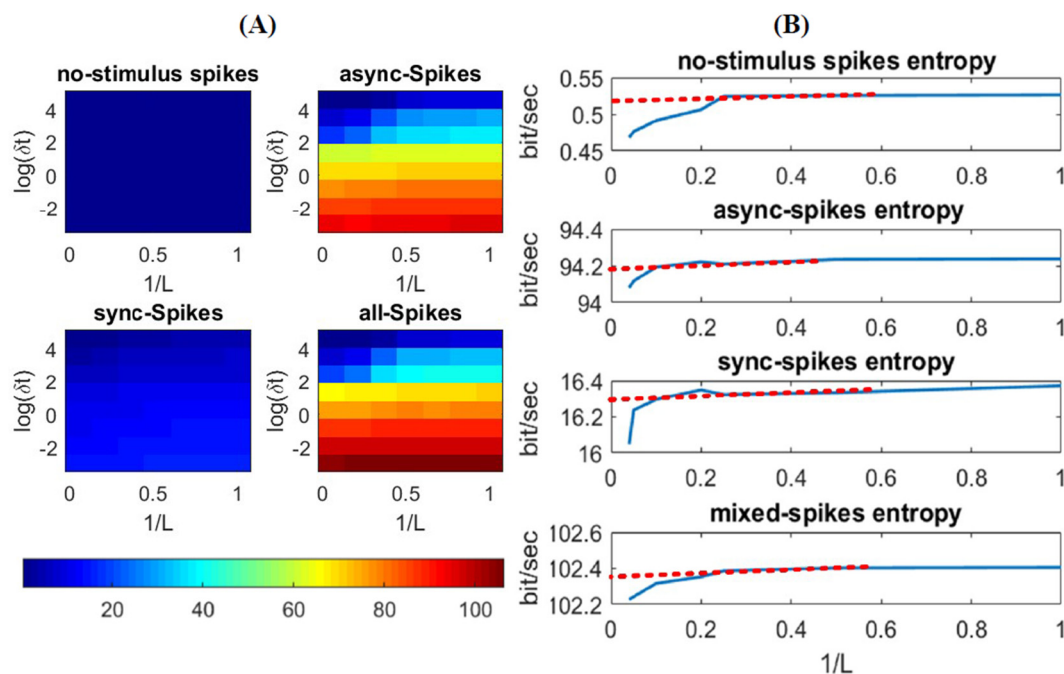


Figure 3. Entropy of different types of spikes. (A) Entropy of different patterns of spikes as a function of time-bin (δt) and word-length (L) (B) Estimated entropy rate of spikes, for no-stimulus, slow, fast, and mixed stimuli, is plotted against the reciprocal of word length, $1/L$. The dashed line and its intersection with y axis represent the value of entropy for $L \rightarrow \infty$, i.e., the minimum value of the entropy.

3.3. Time-Varying Entropy (TVE) Measure

To determine how information of spikes are related to different features of the stimulus, we propose an entropy measure, namely, time-varying entropy (TVE), to calculate the entropy of spikes in a neural ensemble at each time bin. TVE is defined as follows

$$H(L, \delta t, k) = -\frac{1}{L\delta t} \sum_{w \in W_k(L, \delta t)} p(w) \log(w) \quad (11)$$

where k is the index of the time bins and $p(w)$ is the probability of a specific word with length L at time k across the neurons (trials). $W_k(L, \delta t)$ is a set of all possible words with length L and time resolution δt at time-step k across trials. The TVE, in (11), is calculated across neurons and introduces a time-varying entropy measure for an ensemble of neurons. The main difference between the entropy in (10) and that in (11) is that in the former the expected value of (logarithm of) code words is obtained over the length of stimulus, whereas in the latter, it is calculated across neurons, thus providing an entropy measure over time. In words, the entropy in (11) is considered as an information-theoretic measure to calculate information underlying spikes.

To explore the relationship between (all) spikes and the stimulus features, we calculate the correlation between the TVE and stimulus for different combinations of word-lengths (L) and time resolutions (δt). To better visualize how the entropy changes over time, we plotted a few examples of TVE for different L and δt in Figure 4A. Figure 4B shows the relationship between TVE and different features of the stimulus as well as the mixed stimulus as a function of L and δt . As can be seen in this figure, TVE is highly correlated with the I_{fast} for small time bins, implying that the neural code of a neural ensemble utilizes spikes with a very high temporal resolution to represent fast (abrupt) changes in the stimulus (see also Figure 4A for $L = 1$ and $\delta t = 0.05$ ms). For relatively high temporal resolution, TVE increases slightly for shorter L (Figure 4B (top)), confirming that precise timing of spikes is sufficient to represent fast features of the stimulus. Although TVE is calculated for all spikes, one can interpret that spatially correlated spikes in a short time interval in a neural ensemble are considered as synchronous; thus, the code words of $L = 1$ provide better representation for synchronous spikes (i.e., code words are temporally independent).

Figure 4B (middle) shows that the TVE is highly correlated with the I_{slow} for medium time bins, indicating that the neural code of a neural ensemble uses spikes in a relatively low temporal resolution to encode the amplitude of smooth changes (low frequency) in the stimulus (see also Figure 4A for $L = 10$ and $\delta t = 5$ ms). For relatively medium temporal resolution, TVE increases slightly for longer L , suggesting that an appropriate range of temporal correlation within the code words enhances the information representation of the slow features of the stimulus. Figure 4B (bottom) shows the correlation of TVE and the mixed stimulus (I_{mixed}).

Unlike (10), where the entropy is calculated for each individual neuron (over the total stimulation time), the TVE computes the entropy of a neural ensemble over time (at each time bin). One can expect that the average of the TVE over time is equivalent to the average of entropy of individual neurons (as calculated in (10)). Figure 4C shows that the average of the TVE over time is similar to that of individual neurons. As mentioned above, the entropy of all spikes based on (10), and in agreement with [19], is 102 bit/s. As well, the integral of the TVE over time (for the same L and δt) is equal to 92.58 bit/s. It is to be noted that the time-varying entropy in (11) is calculated based on the probability distribution of spikes (for each time-bin) across a limited number of trials (neurons). Therefore, one can expect that the entropy calculated by [19] provides an upper bound of time-varying entropy that an ensemble of neurons with a limited number of neurons can carry.

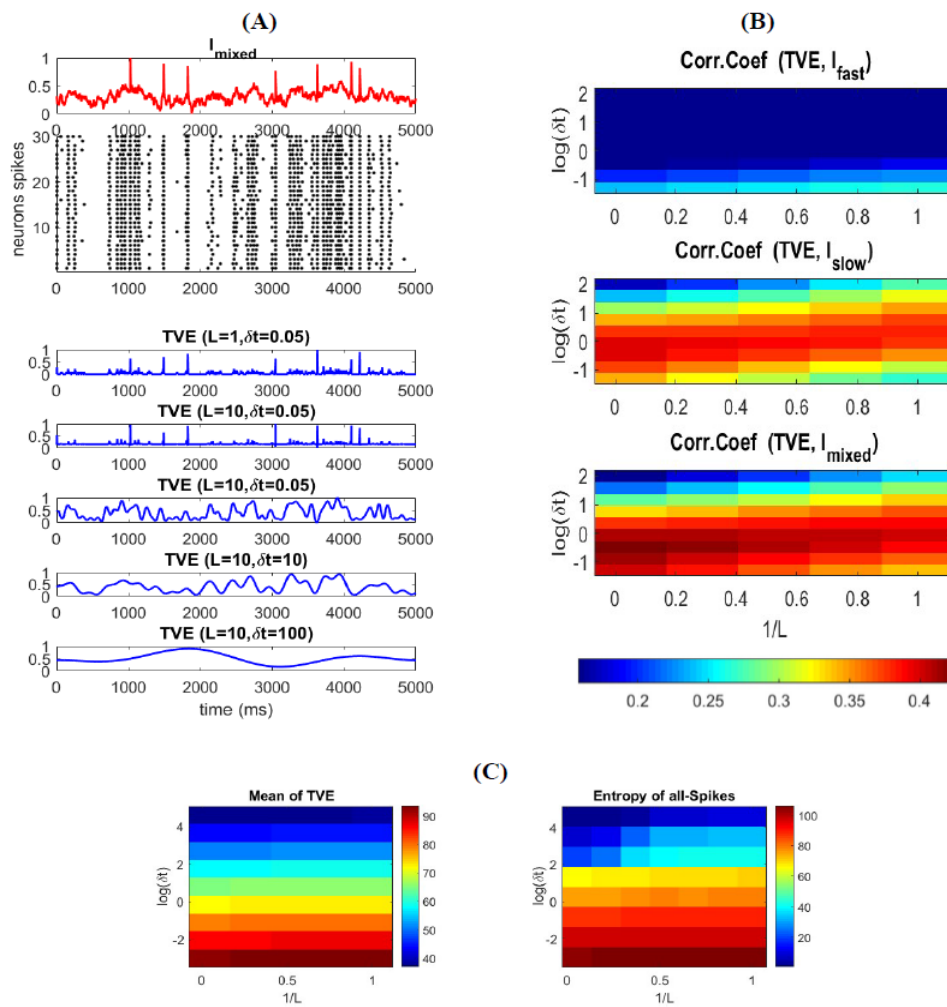


Figure 4. Time-varying entropy (TVE) measure for different type of spikes. (A) TVE for different type of spikes as function of word-length and time-bins resolution. Different parameter sets for TVE enables extracting different types of information. For example, by setting $L = 10$, $\delta t = 0.05$ ms TVE extracts information underlying synchronous spikes. As well, by setting $L = 10$, $\delta t = 10$ ms TVE extracts information related to asynchronous spikes (B) TVE measure correlation coefficient with I_{fast} , I_{slow} , and I_{mixed} . The correlation of TVE with each stimulus is aligned with the figures in panel (A). For $L = 10$, $\delta t = 0.05$ ms TVE is highly correlated with I_{fast} , which drives synchronous spikes. For $L = 10$, $\delta t = 10$ ms TVE is highly correlated with I_{slow} , which provokes asynchronous spikes. Thus, TVE measure can extract information about the stimulus directly from the spikes (C) Mean of Integration of TVE measure over time (left) and entropy of all-spikes calculated in Equation (9) (right).

To better clarify the difference between the entropy in (10) and the TVE in (11), we illustrate in Figure 5 how the entropy is calculated across trials (neurons) for any given time. For specific L and δt , the probability distribution of code words, $p(w)$, can be calculated over the whole simulation time (see (10)). To calculate the TVE, the probability distribution of code words, for any given time, can be calculated across neurons. Figure 5 shows two probability distributions of code words in different time bins, namely, $p(w')$ at t_i and t_j .

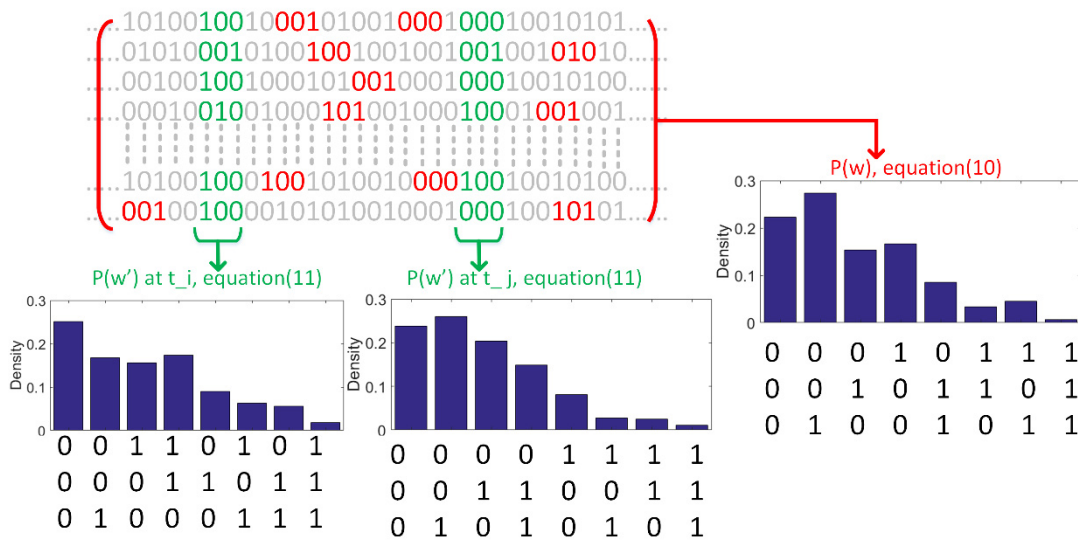


Figure 5. Illustration of calculation of the entropy in (10) and the TVE in (11). The binary sequence of each row indicates the response of each neuron in a neural ensemble. Probability distribution of code words, $p(w)$, over the whole length of data can be calculated based on (10). Two probability distributions underlying two time bins, t_i and t_j , are calculated across neurons (see (11)). The length of code words is equal to 3 and spikes are binned at a resolution (δt), equal to the sampling time of the simulation. Several code words are highlighted by red and green.

Furthermore, to calculate the optimum values of L and δt , which lead to extract maximum information of the mixed stimulus, we build a linear decoder model to reconstruct the stimulus from different combinations of TVE measures. We used a linear regression model with root mean squared error (RMSE) cost function [43] to calculate the linear coefficients and parameter setting of the TVEs. The linear decoder model and the cost functions are written as

$$\hat{y} = w_s * TVE_s + w_a * TVE_a + b; \quad TVE_s = TVE(L_s, \delta t_s), \quad TVE_a = TVE(L_a, \delta t_a) \tag{12}$$

$$\{w_s, w_a\} = \underset{\{w_s, w_a, b, L_s, \delta t_s, L_a, \delta t_a\}}{\operatorname{argmin}} \sqrt{\frac{1}{N} \sum_{k=1}^N (y_k - \hat{y}_k)^2} \tag{13}$$

where TVE_s and TVE_a are the TVE measures for spikes with $\{L_s, \delta t_s\}$ parameters set and those with the $\{L_a, \delta t_a\}$ parameters set, respectively. N is the total number of samples, y_k and \hat{y}_k are the mixed and estimated stimulus at time index k , and $\{w_s, w_a, b\}$ are the regression parameters for TVE_s and TVE_a , respectively. We optimize the linear decoder model for different parameter settings for TVE_s and TVE_a , and select the optimum decoder based on its RMSE performance. Figure 6A shows the true and reconstructed mixed stimulus by (12). The optimum values of L and δt (not shown here), for TVE_s and TVE_a , to reach the best decoding performance are the same as the parameters presented in Figure 4B, for which the highest correlation between TVE (all spikes) and fast and slow signals was obtained. These results justify that TVE with specific (optimum) ranges of δt and L corresponds to different types of information underlying distinct features of the stimulus.

We can represent this relationship clearly by visualizing the TVE measure spectrum for different δt and L through time (note that TVE is less sensitive to the changes in L compared to that in δt ; see (Figure 4A)). The TVE can identify which information is carried by spikes and reconstructs its associated stimulus features. Figure 6B shows the TVE calculated for synchronous, asynchronous spikes, and all spikes for different δt and a fixed $L (=10)$. One can clearly observe that TVE calculated by asynchronous spikes represents information of slowly time-varying changes in the stimulus for medium to high δt . In contrast, the TVE obtained by synchronous spikes represents information

underlying abrupt changes in the stimulus for small δt . Therefore, the TVE calculated by all spikes and for different δt creates a heat map of information (i.e., the TVE spectrum) underlying different features of the stimulus. It is worth mentioning that by integrating the TVE over time (similar to Figure 4C) for synchronous and asynchronous spikes, one can measure how much information is carried by each type of spike. We can represent this relationship clearly by visualizing the TVE measure spectrum for different δt and L through time (note that TVE is less sensitive to the changes in L compared to that in δ ; see (Figure 4A)). The TVE can identify which information is carried by spikes and reconstructs its associated stimulus features. Figure 6B shows the TVE calculated for synchronous, asynchronous spikes, and all spikes for different δt and a fixed $L (=10)$. One can clearly observe that TVE calculated by asynchronous spikes represents information of slowly time-varying changes in the stimulus for medium to high δt . In contrast, the TVE obtained by synchronous spikes represents information underlying abrupt changes in the stimulus for small δt . Therefore, the TVE calculated by all spikes and for different δt creates a heat map of information (i.e., the TVE spectrum) underlying different features of the stimulus. It is worth mentioning that by integrating the TVE over time (similar to Figure 4C) for synchronous and asynchronous spikes, one can measure how much information is carried by each type of spike.

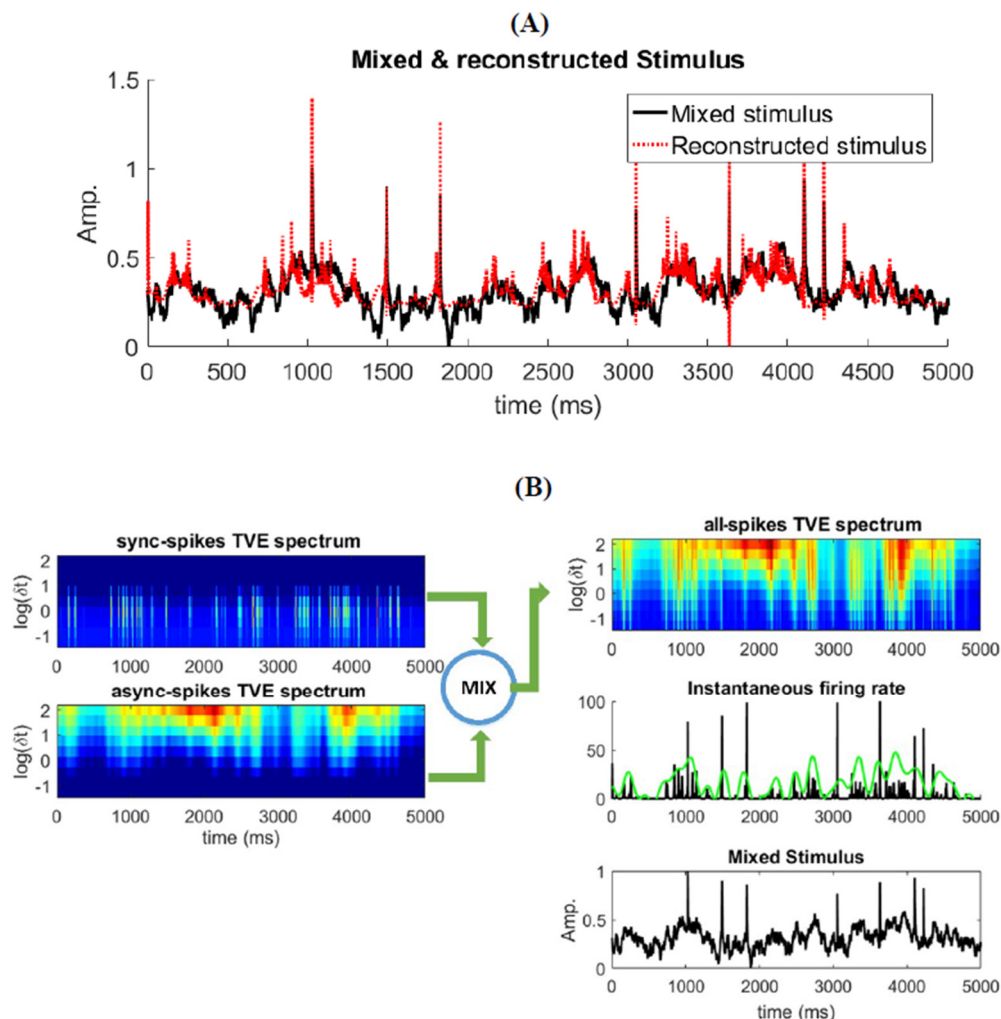


Figure 6. Different elements of the mixed stimulus (I_{fast} , I_{slow}) and their relationship with different type of spikes. (A) Reconstruction of the mixed stimulus by TVE measure. (B) TVE measure spectra for different patterns of spikes (synchronous, asynchronous, and all) given different δt with fixed $L = 10$ through time. Instantaneous firing rate of the neural ensemble is calculated with two different kernel width; green and black graphs are related to kernel width = 100 ms and kernel width = 5 ms, respectively.

Although the TVE spectrum in Figure 6 reveals that synchronous and asynchronous spikes are decodable in different time-resolution scales, the extent to which these spikes can represent the stimulus features relies on multiple factors like the level of background synaptic noise, network size, intrinsic parameters of single neurons, etc. For example, a recent study [44] investigated the necessary conditions underlying reliable representation (and propagation) of time-varying firing rates in feed-forward neural networks with homogeneous neurons. It has been shown that a proper and biologically realistic level of background synaptic noise is substantial to preserve information of a common stimulus. To explore how the level of background synaptic noise alters the co-existence of decodable synchronous and asynchronous spikes (i.e., multiplexing), we consider two extreme cases in which a neural ensemble receives weak and strong synaptic noise. Figure 7 (two top rows) shows the stimulus and the firing rate of a neural ensemble receiving weak ($\sigma = 0.5$ pA), intermediate ($\sigma = 10$ pA), and strong ($\sigma = 50$ pA) synaptic noise. The neural response tends towards synchronous states for a low level of background synaptic noise (Figure 7 (left, second row)). In contrast, the neural response converges to the average firing rate (with some fluctuations) for a high level of synaptic noise (Figure 7 (right, second row)). Therefore, multiplexing—in the sense of decodable synchronous and asynchronous spikes—fails in these extreme cases (see [44] for more details).

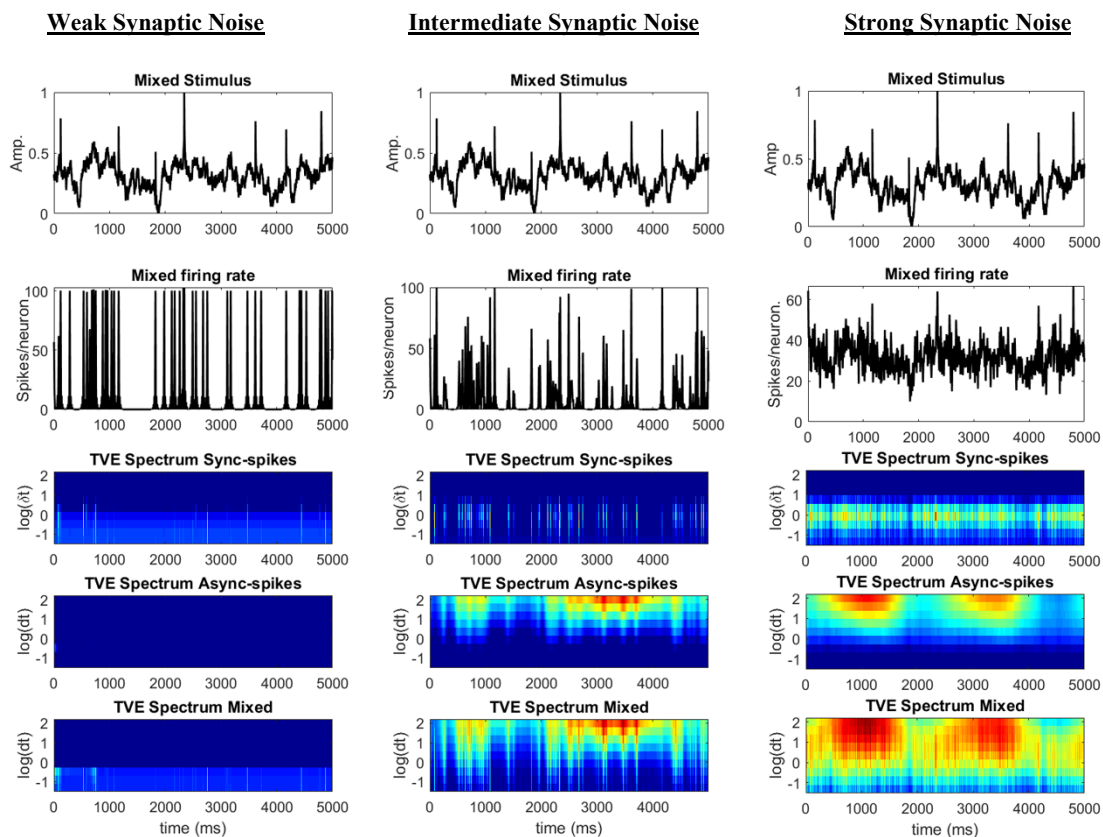


Figure 7. Firing rate and TVE spectrum of spikes for a neural ensemble receiving weak (left, $\sigma = 0.5$ pA), intermediate (middle, $\sigma = 10$ pA), and strong (right, $\sigma = 50$ pA) synaptic noises. The TVE spectrum of different types of spikes is obtained in a similar way, as explained in Figure 6.

For each level of background synaptic noise, the TVE spectrum (similar to Figure 6) is calculated for synchronous, asynchronous, and all spikes. That of synchronous spikes fully represents the TVE spectrum of a neural ensemble receiving weak synaptic noise (see Figure 7 (left, last three rows)). In contrast, the TVE spectrum of a neural ensemble receiving high synaptic noise (Figure 7 (right, last three rows)), is mainly represented by that of asynchronous spikes. As the synchrony threshold is the same for weak, intermediate, and high synaptic noise, this threshold causes several

false-positive synchronous events to be detected when the level of synaptic noise is high. Similar to Figure 6, the TVE spectrum of all spikes for an intermediate level of synaptic noise (see Figure 7 (middle, last three rows)) reveals information underlying slow and fast features of the stimulus. Although the TVE spectrum might not be informative of the decodable information underlying the stimulus when the background synaptic noise level is not biologically realistic (either low or high), the TVE spectrum of all spikes clearly represents that of synchronous and asynchronous spikes for all different levels of synaptic noise.

3.4. Relationship between Mixed Stimulus and Spike Patterns

To identify how synchronous and asynchronous spikes are related to, respectively, the abrupt changes in and the intensity of the stimulus, we develop a decoder model to reconstruct the stimulus from the spikes. A Kalman-filter (KF) decoder model [45,46], a well known state-space approach for neural decoding, is used to reconstruct the stimulus with optimum accuracy underlying linear models [46]. After estimating the parameters of the KF decoder based on mixed stimulus and all spikes, we apply this decoder to synchronous and asynchronous spikes to explore which features of the stimulus are reconstructed. The state-space model of the Kalman-filter decoder can be written as

$$x_{k+1} = Ax_k + w \quad (14)$$

$$z_k = Hx_k + q \quad (15)$$

where x_k and z_k denote the decoded stimulation and the neural firing rate at time index k , respectively. A is the coefficient matrix, $w \approx N(0, W)$ represents the uncertainty underlying x_k , H is a matrix that linearly relates the stimulus to the neural firing, and $q \approx N(0, Q)$ is the measurement noise. We estimate the parameter set $\{A, W, H, Q\}$ from the training data using the following equation.

$$A = \operatorname{argmin}_A \sum_{k=1}^{N-1} \|x_{k+1} - Ax_k\|^2, H = \operatorname{argmin}_H \sum_{k=1}^N \|z_k - Hx_k\|^2 \quad (16)$$

By using Equations (14)–(16), we can reconstruct the stimulation from the spiking activity of the ensemble recursively [46]. Figure 8 shows the decoded stimulus from all spikes, synchronous, and asynchronous spikes using the KF-decoder model. Figure 8A shows that we can reconstruct the fast and slow features of the stimulus by applying the decoder model to synchronous and asynchronous spikes, respectively. Furthermore, to enhance the neural decoding, we filtered (using a Gaussian kernel) synchronous and asynchronous spikes with optimum time resolutions ($\delta t = 0.05$ ms and $\delta t = 10$ ms, respectively) before applying them to the above decoder. Figure 8B shows the reconstructed signals by filtered synchronous and asynchronous spikes. The reconstructed signals are better fitted to the slow and fast features of the stimulus. One can conclude that the information underlying different features of the stimulus is best decoded by different types of spikes, which are integrated with specific time scales.

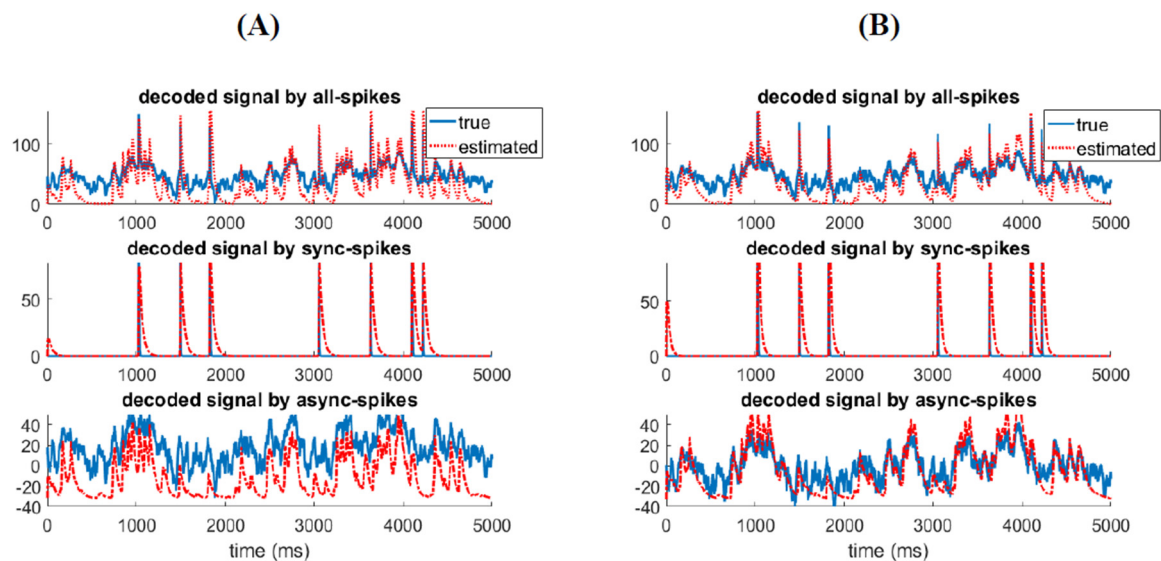


Figure 8. Information decoded by synchronous and asynchronous spikes are associated with different features of the stimulus. **(A)** Kalman-filter decoder model was developed to reconstruct the (mixed) stimulation from all spikes (top). The decoder was then applied to synchronous (middle) and asynchronous spikes (bottom). **(B)** Similar to **(A)** but synchronous (middle) and asynchronous (bottom) spikes were first filtered by a Gaussian kernel with optimum time resolutions ($\delta t = 0.05$ ms for synchronous and $\delta t = 10$ ms for asynchronous spikes) before applying them to the Kalman-filter decoder model.

4. Discussion

In this paper, we demonstrated that differentially correlated spikes in a neural ensemble carry different information, which corresponds to different features of the stimulus. By feeding a mixed stimulus consisting of slow and fast features into an ensemble of homogeneous neurons, we created a multiplexed code in which synchronous and asynchronous spikes can be distinguished. It was shown that the probability distribution of these spikes are distinctively separable. Furthermore, we considered spikes as code words and calculated the entropy of these code words for different lengths and time resolutions. A time-varying entropy (TVE) measure was proposed to calculate the entropy of a neural ensemble at each time bin. By applying TVE to the multiplexed code, we showed that information underlying synchronous and asynchronous spikes was maximized for different time resolutions and lengths. Thus, synchronous and asynchronous spikes carried information of different time scales. However, it was observed that the sensitivity of the TVE to the length of codewords was negligible (specifically for high time resolutions). Finally, we developed a Kalman-based decoder to reconstruct the stimulus from the spikes. We showed that slow and fast features of the stimulus could be fully decoded from the asynchronous and synchronous spike, respectively.

As natural stimuli often operate on multiple time scales [47], the TVE calculates the entropy of a homogeneous neural ensemble in different time resolutions, thus providing a time-varying representation of a neural code in different resolution scales. A recent study [47] introduced multiscale relevance (MSR) measure to characterize the temporal structure of the activities of neurons within a heterogeneous population. It was shown that MSR could capture the dynamical variability of the activity of single neurons across different time scales and detect informative neurons as well as neurons that show a high decoding performance [47]. Despite differences in the architecture of the neural ensemble, types of stimuli as well as other factors like the heterogeneity vs. homogeneity of neurons that differentiated between the scope of our study and that of [47], both studies utilized the entropy as an information-theoretic measure and capitalized on the needs of such a measure in multiscale neural analyses.

The advancements of imaging and electrical recording technologies have provided accessibility to neural activities in population levels; thus, the need for multi-dimensional methods is ever increasing in brain-related studies. Time-varying firing rates of a neural ensemble across time and across multiple experimental conditions can be considered as the starting point for the population-level analyses [48]. Kernel smoothing techniques with optimum kernel bandwidth that maximizes the goodness-of-fit of the density estimate to the underlying rate of spikes are tools for estimating the instantaneous firing rate of a neural ensemble [49]. In this regard, one can use the TVE measure as a simple way to identify the most informative time scales underlying the neural code of an ensemble of neurons.

Nevertheless, in population-level analyses, methods that infer the dynamics underlying neural computations are more demanding than those focused on the representation of neuronal activities [48]. Recently, Elsayed and Cunningham [50] proposed a framework to measure the correlation of the neural activity at the population level across times, neurons, and (experimental) conditions. Although this framework was designed for the rate code and thereby cannot be applied to the temporal code, their proposed methodology [50] determines whether neural population activity exhibits a structure above and beyond that of its set of primary features [48]. Unlike [50], the TVE can track the dynamics of a neural code in multiple time scales and one can apply the TVE to both rate and temporal codes simultaneously. Moreover, it is notable that the constraints on the correlation of the neural activity in a neural ensemble across the experimental condition was relaxed in the present study. Taken together, the TVE not only tracks the dynamics of a neural code—in the sense of detecting synchronous and asynchronous states of a neural ensemble—in different time-resolution scales, but also provides decodable information underlying the stimulus features. The TVE can be extended in our future studies to address richer datasets comprising heterogeneous neurons or networks with feed-forward and recurrent connectivity.

Author Contributions: Conceptualization, M.L.; methodology, M.R.R. and M.L.; software, M.R.R. and M.L.; validation, M.R.R., M.L. and M.R.P.; formal analysis, M.R.R. and M.L.; investigation, M.L.; resources, M.L. and M.R.P.; data curation, M.R.R. and M.L.; writing—original draft preparation, M.R.R. and M.L.; writing—review and editing, M.L. and M.R.P.; visualization, M.R.R.; supervision, M.L. and M.R.P.; project administration, M.L. and M.R.P.; funding acquisition, M.L. and M.R.P. All authors have read and agreed to the published version of the manuscript.

Funding: The present study was supported by Dr. Lankarany's start-up Grant and NSERC Discovery Grant (RGPIN-2020-05868).

Acknowledgments: All the codes in our simulation study are provided in MATLAB and available here: <https://github.com/nsbspl/A-Time-varying-Information-Measure-for-Tracking-Dynamics-of-Neural-Codes-in-a-Neural-Ensemble>.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zuo, Y.; Safaai, H.; Notaro, G.; Mazzoni, A.; Panzeri, S.; Diamond, M.E. Complementary contributions of spike timing and spike rate to perceptual decisions in rat S1 and S2 cortex. *Curr. Biol.* **2015**, *25*, 357–363. [[CrossRef](#)] [[PubMed](#)]
2. Panzeri, S.; Harvey, C.D.; Piasini, E.; Latham, P.E.; Fellin, T. Cracking the neural code for sensory perception by combining statistics, intervention, and behavior. *Neuron* **2017**, *93*, 491–507. [[CrossRef](#)] [[PubMed](#)]
3. Runyan, C.A.; Piasini, E.; Panzeri, S.; Harvey, C.D. Distinct timescales of population coding across cortex. *Nature* **2017**, *548*, 92–96. [[CrossRef](#)] [[PubMed](#)]
4. Kremkow, J.; Aertsen, A.; Kumar, A. Gating of signal propagation in spiking neural networks by balanced and correlated excitation and inhibition. *J. Neurosci.* **2010**, *30*, 15760–15768. [[CrossRef](#)] [[PubMed](#)]
5. Montemurro, M.A.; Panzeri, S.; Maravall, M.; Alenda, A.; Bale, M.R.; Brambilla, M.; Petersen, R.S. Role of precise spike timing in coding of dynamic vibrissa stimuli in somatosensory thalamus. *J. Neurophysiol.* **2007**, *98*, 1871–1882. [[CrossRef](#)]
6. 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* **2001**, *29*, 769–777. [[CrossRef](#)]

7. London, M.; Roth, A.; Beeren, L.; Häusser, M.; Latham, P.E. Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex. *Nature* **2010**, *466*, 123–127. [[CrossRef](#)]
8. Rieke, F.; Warland, D.; van Steveninck, R.R.; Bialek, W. *Spikes: Exploring the Neural Code*; MIT Press: Cambridge, MA, USA, 1999; Volume 7.
9. Abeles, M.; Purpura, Y.; Bergman, H.; Vaadia, E. Synchronization in neuronal transmission and its importance for information processing. *Prog. Brain Res.* **1994**, *102*, 395–404.
10. Diesmann, M.; Gewaltig, M.-O.; Aertsen, A.J.N. Stable propagation of synchronous spiking in cortical neural networks. *Nature* **1999**, *402*, 529–533. [[CrossRef](#)]
11. Panzeri, S.; Brunel, N.; Logothetis, N.K.; Kayser, C. Sensory neural codes using multiplexed temporal scales. *Trend Neurosci.* **2010**, *33*, 111–120. [[CrossRef](#)]
12. Wilson, H.R.; Cowan, J.D. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* **1972**, *12*, 1–24. [[CrossRef](#)]
13. Gerstner, W.; Kistler, W.M. *Spiking Neuron Models: Single Neurons, Populations, Plasticity*; Cambridge University Press: Cambridge, MA, USA, 2002.
14. Kumar, A.; Rotter, S.; Aertsen, A. Conditions for propagating synchronous spiking and asynchronous firing rates in a cortical network model. *J. Neurosci.* **2008**, *28*, 5268–5280. [[CrossRef](#)] [[PubMed](#)]
15. Quiroga, R.Q.; Panzeri, S.J.N.R.N. Extracting information from neuronal populations: Information theory and decoding approaches. *Nat. Rev. Neurosci.* **2009**, *10*, 173–185. [[CrossRef](#)] [[PubMed](#)]
16. Lankarany, M.; Al-Basha, D.; Ratté, S.; Prescott, S.A. Differentially synchronized spiking enables multiplexed neural coding. *Nat. Acad. Sci.* **2019**, *116*, 10097–10102. [[CrossRef](#)] [[PubMed](#)]
17. Kumar, A.; Rotter, S.; Aertsen, A. Spiking activity propagation in neuronal networks: Reconciling different perspectives on neural coding. *Nat. Rev. Neurosci.* **2010**, *11*, 615–627. [[CrossRef](#)] [[PubMed](#)]
18. Reid, R.C.; Victor, J.; Shapley, R. The use of m-sequences in the analysis of visual neurons: Linear receptive field properties. *Vis. Neurosci.* **1997**, *14*, 1015–1027. [[CrossRef](#)]
19. Reinagel, P.; Reid, R.C. Temporal coding of visual information in the thalamus. *J. Neurosci.* **2000**, *20*, 5392–5400. [[CrossRef](#)]
20. Pillow, J.W.; Shlens, J.; Paninski, L.; Sher, A.; Litke, A.M.; Chichilnisky, E.J.; Simoncelli, E.P. Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* **2008**, *454*, 995–999. [[CrossRef](#)]
21. Bastos, A.M.; Schoffelen, J.-M. A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. *Front. Syst. Neurosci.* **2016**, *9*, 175. [[CrossRef](#)]
22. Borst, A.; Theunissen, F.E. Information theory and neural coding. *Nat. Neurosci.* **1999**, *2*, 947–957. [[CrossRef](#)]
23. Piasini, E.; Panzeri, S. Information Theory in Neuroscience. *Entropy* **2019**, *21*, 62. [[CrossRef](#)]
24. Stevens, C.F.; Zador, A.M. Information through a spiking neuron. In *Advances in Neural Information Processing Systems*; NIPS: San Diego, CA, USA, 1996; pp. 75–81.
25. Jordan, M.I.J.S. Graphical models. *Statist. Sci.* **2004**, *19*, 140–155. [[CrossRef](#)]
26. Belghazi, M.I.; Baratin, A.; Rajeswar, S.; Ozair, S.; Bengio, Y.; Courville, A.; Hjelm, R.D. Mine: Mutual information neural estimation. In Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018.
27. Timme, N.M.; Lapish, C.J.E. A tutorial for information theory in neuroscience. *eNeuro* **2018**, *5*. [[CrossRef](#)] [[PubMed](#)]
28. Cover, T.M.; Thomas, J.A. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*, 2nd ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2006.
29. Walters-Williams, J.; Li, Y. Estimation of mutual information: A survey. In *International Conference on Rough Sets and Knowledge Technology*; Springer: Berlin/Heidelberg, Germany, 2009.
30. Pirschel, F.; Kretzberg, J. Multiplexed population coding of stimulus properties by leech mechanosensory cells. *J. Neurosci.* **2016**, *36*, 3636–3647. [[CrossRef](#)]
31. Destexhe, A.; Rudolph, M.; Fellous, J.M.; Sejnowski, T.J. Fluctuating synaptic conductances recreate in vivo-like activity in neocortical neurons. *Neuroscience* **2001**, *107*, 13–24. [[CrossRef](#)]
32. Morris, C.; Lecar, H. Voltage oscillations in the barnacle giant muscle fiber. *Biophys. J.* **1981**, *35*, 193–213. [[CrossRef](#)]
33. Khubieh, A.; Rudolph, M.; Fellous, J.M.; Sejnowski, T.J. Regulation of cortical dynamic range by background synaptic noise and feedforward inhibition. *Neuroscience* **2016**, *26*, 3357–3369. [[CrossRef](#)]

34. Ratté, S.; Hong, S.; De Schutter, E.; Prescott, S.A. Impact of neuronal properties on network coding: Roles of spike initiation dynamics and robust synchrony transfer. *Neuron* **2013**, *78*, 758–772. [[CrossRef](#)]
35. Prescott, S.A.; De Koninck, Y.; Sejnowski, T.J. Biophysical basis for three distinct dynamical mechanisms of action potential initiation. *PLoS Comput. Biol.* **2008**, *4*, e1000198. [[CrossRef](#)]
36. Destexhe, A.; Rudolph, M.; Paré, D. The high-conductance state of neocortical neurons in vivo. *Nat. Rev. Neurosci.* **2003**, *4*, 739–751. [[CrossRef](#)]
37. Shalizi, C. *Advanced Data Analysis from an Elementary Point of View*; Cambridge University Press: Cambridge, UK, 2013.
38. Pérez-Cruz, F. Kullback-Leibler divergence estimation of continuous distributions. In Proceedings of the 2008 IEEE international symposium on information theory, Toronto, ON, Canada, 6–11 July 2008.
39. Van Kerm, P. Adaptive kernel density estimation. *Stata J.* **2003**, *3*, 148–156. [[CrossRef](#)]
40. Fasano, G.; Franceschini, A. A multidimensional version of the Kolmogorov–Smirnov test. *Mon. Not. R. Astron. Soc.* **1987**, *225*, 155–170. [[CrossRef](#)]
41. Massey, F.J., Jr. The Kolmogorov–Smirnov test for goodness of fit. *J. Am. Stat. Assoc.* **1951**, *46*, 68–78. [[CrossRef](#)]
42. Strong, S.P.; Koberle, R.; van Steveninck, R.R.R.; Bialek, W. Entropy and information in neural spike trains. *Phys. Rev. Lett.* **1998**, *80*, 197. [[CrossRef](#)]
43. Seber, G.A.; Lee, A.J. *Linear Regression Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 2012; Volume 329.
44. Hasanzadeh, N.; Rezaei, M.; Faraz, S.; Popovic, M.R.; Lankarany, M. Necessary Conditions for Reliable Propagation of Time-Varying Firing Rate. *Front. Comput. Neurosci.* **2020**, *14*, 64. [[CrossRef](#)]
45. Wu, W.; Black, M.J.; Gao, Y.; Bienenstock, E.; Serruya, M.; Shaikhouni, A.; Donoghue, J.P. Inferring hand motion from multi-cell recordings in motor cortex using a Kalman filter. In Proceedings of the SAB'02-Workshop on Motor Control in Humans and Robots: On the Interplay of Real Brains and Artificial Devices, Edinburgh, UK, 10 August 2002; pp. 66–73.
46. Wu, W.; Black, M.J.; Gao, Y.; Bienenstock, E.; Serruya, M.; Shaikhouni, A.; Donoghue, J.P. Neural decoding of cursor motion using a Kalman filter. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 2003; pp. 133–140.
47. Cubero, R.J.; Marsili, M.; Roudi, Y. Multiscale relevance and informative encoding in neuronal spike trains. *J. Comput. Neurosci.* **2020**, *48*, 85–102. [[CrossRef](#)] [[PubMed](#)]
48. Pillow, J.W.; Aoi, M.C. Is population activity more than the sum of its parts? *Nat. Neurosci.* **2017**, *20*, 1196–1198. [[CrossRef](#)]
49. Shimazaki, H.; Shinomoto, S. Kernel bandwidth optimization in spike rate estimation. *J. Comput. Neurosci.* **2010**, *29*, 171–182. [[CrossRef](#)]
50. Elsayed, G.F.; Cunningham, J.P. Structure in neural population recordings: An expected byproduct of simpler phenomena? *Nat. Neurosci.* **2017**, *20*, 1310. [[CrossRef](#)]

