ROBUST ONLINE RECONSTRUCTION OF CONTINUOUS-TIME SIGNALS FROM A LEAN SPIKE TRAIN ENSEMBLE CODE

Anik Chattopadhyay, Arunava Banerjee, *Member, IEEE,*

***Abstract*—Sensory stimuli in animals are encoded into spike trains by neurons, offering advantages such as sparsity, energy efficiency, and high temporal resolution. This paper presents a signal processing framework that deterministically encodes continuous-time signals into biologically feasible spike trains, and addresses the questions about representable signal classes and reconstruction bounds. The framework considers encoding of a signal through spike trains generated by an ensemble of neurons using a convolve-then-threshold mechanism with various convolution kernels. A closed-form solution to the inverse problem, from spike trains to signal reconstruction, is derived in the Hilbert space of shifted kernel functions, ensuring sparse representation of a generalized Finite Rate of Innovation (FRI) class of signals. Additionally, inspired by real-time processing in biological systems, an efficient iterative version of the optimal reconstruction is formulated that considers only a finite window of past spikes, ensuring robustness of the technique to ill-conditioned encoding; convergence guarantees of the windowed reconstruction to the optimal solution are then provided. Experiments on a large audio dataset demonstrate excellent reconstruction accuracy at spike rates as low as one-third of the Nyquist rate, while showing clear competitive advantage in comparison to state-of-the-art sparse coding techniques in the low spike rate regime.**

***Index Terms*—coding, integrate-and-fire, spike, convolution, reconstruction.**

I. INTRODUCTION

In most animals, sensory stimuli are communicated to the brain via ensembles of discrete, spatio-temporally compact electrical events generated by neurons, known as action potentials or spikes[1]. The conversion of continuous-time stimuli to spike trains occurs at an early stage of sensory processing, such as in the retinal ganglion cells in the visual pathway and spiral ganglion cells in the auditory pathway. Nature likely resorts to spike-based encoding due to several advantages: sparsity of representation[2], energy efficiency[3], noise robustness[4], high temporal precision[5], and facilitation of downstream computation[6,7]. Olshausen and Field[2] demonstrated how efficient codes could arise from learning sparse representations of natural stimuli, resulting in striking similarities to observed biological receptive fields. Similarly, Smith and Lewicki[8,9] showed that auditory filters could be estimated by training a population spike code model with natural sounds. While these studies highlight important aspects of spike-based encoding, they rely on existing dictionary learning techniques (e.g., matching pursuit or basis pursuit) to obtain the sparse codes, raising the question of the biological feasibility of computing such codes. The formal study of encoding continuous-time sensory stimuli via biologically plausible spiking neurons falls under the field of neural coding. Based on how neural spike responses are represented, studies can be broadly categorized[10] into (i) rate coding, where spike train responses are converted into an average rate, and (ii) temporal coding, where the precise timing of spikes convey information about the stimuli. In the rate coding literature, spike responses to stimuli are converted to an average instantaneous rate $r(t)$, and stimulus reconstruction is typically formulated probabilistically by choosing the stimulus s that maximizes the likelihood $P(s|r)$. Rate coding is criticized for losing temporal precision, especially since studies have shown that neurons can exhibit sub-millisecond precision[11]. Temporal coding, in contrast, although emphasizes the timing of individual spikes, stimulus reconstruction in this literature is often formulated probabilistically (e.g., Bayesian Inference[12]) or through a linear transformation of spike-responses (e.g., reverse correlation[13]). Probabilistic approaches to reconstruction complicate the development of a deterministic signal processing framework from spike trains, where simple linear transformations may be too restrictive for representing a generalized class of signals. Recent advances in temporal coding schemes, such as in Brendel et al.[14], leverage recurrent networks and have shown near perfect reconstruction for certain signals. However, these techniques involve complex training procedures and do not provide reconstruction guarantees for a generalized class of signals. This article is therefore motivated to build a signal processing framework that deterministically encodes continuous-time signals into biologically plausible spike trains, addressing questions of representable signal classes and reconstruction error bounds.

Related Work and Contribution: From a signal processing standpoint, the very coarse $\Sigma\Delta$ quantization of bandlimited signals, as investigated in[15], in effect represents a spike train encoding. This scheme encodes an oversampled signal into a stream of 1/0 bits by thresholding the cumulative quantization error. Similarly, the Time encoding Machine[16] encodes a signal $X(t)$ into a sequence of times $\{t_k\}$, corresponding to instants at which the signal crosses certain thresholds using an integrate-and-fire mechanism. In both cases, the signal class is bandlimited, and reconstruction guarantee is provided in the oversampled regime above the Nyquist rate. However, evidence suggests that spike codes can achieve high reconstruction accuracy while being much sparser than warranted by a rate code (e.g., the H1 neuron in the fly[17]). Our framework

The supplementary material is available at *http://ieeexplore.ieee.org*.

considers a generalized class—Finite Rate of Innovation[18] (FRI) signals—not limited to bandlimited signals. It involves an encoding model for a spiking neuron that convolves the input signal with a generalized kernel function and produces spikes via thresholding on that convolved signal. Using a set of generalized convolution kernels for spiking neurons is biologically justified. In sensory pathways, an input signal undergoes multiple graded transformations before reaching a spiking neuron (e.g., in the visual pathway, signals pass through amacrine, horizontal, bipolar cells, etc., before being converted into spike trains at the retinal ganglion cells). We model this transformation using a linear convolution kernel, while the overall transformation into spike trains remains nonlinear due to the thresholding operation. The use of such generalized kernels in encoding can also be seen in[19,20]. However, there the inverse problem from spike trains to the reconstructed signal is solved in the Fourier domain (i.e., using *sinc* interpolation), restricting it to the bandlimited class to provide reconstruction guarantees in the oversampled regime. In our framework, we derive a closed-form solution to the inverse problem in the Hilbert space of shifted kernel functions themselves, enabling our framework to represent a generalized FRI class of shifted kernel functions and allowing us to provide reconstruction bounds in the sparse regime. By representing reconstructed signals sparsely using shifted kernel functions, our framework resembles the Convolutional Sparse Coding (CSC) technique[21]. However, while CSC encodes signals through a complex optimization procedure, our framework leverages the efficiency of fast, real-time biological signal processing, showing clear computational advantages, especially in the low spike rate regime (section VII).

Our main contributions are: (1) We present a comprehensive signal processing framework for encoding using biological spike trains and decoding. While reconstruction is not known to be a biological phenomena, it serves as a direct method for evaluating the efficacy of the coding framework rather than using other approaches such as information theory. (2) In Sections III, IV, V, we derive a closed-form solution to the inverse problem of reconstructing the signal from a spike train ensemble, and then identify the feasibility conditions for perfect and approximate reconstruction for the class of FRI signals. (3) In Section VI, we establish the robustness of our framework. We first observe that the optimal reconstruction, formulated in Section III, is poorly conditioned, especially when the input signal is long, therefore producing many spikes. Inspired by biological signal processing, where real-time response to spike-encoded inputs is necessary, we develop an efficient linear-time iterative version of the optimal reconstruction of Section III that considers only a window of past spikes. We prove that this window-based reconstruction converges to the optimal solution under realistic assumptions. (4) In Section VII, we validate our technique through experiments on a large corpus of audio signals. The results show remarkable reconstruction accuracy at a low spike rate, achieving a median of about 20dB SNR at one-third the Nyquist rate. Additionally, we compare our results against a state-of-the-art convolutional sparse coding technique, demonstrating superior accuracy and runtime in the low spike rate regime.

Notations followed in the paper:

- t_i: Time of occurrence of the i^{th} spike of the ensemble.
- T_i: The threshold value of the corresponding neuron.
- $X(t)$ or X: The input signal. For ease of notation we drop the time t as function argument and simply indicate the input as X instead of $X(t)$.
- X^*: The reconstructed signal.
- m: The number of kernels that comprise our framework.
- N: The total number of spikes produced by the system.
- (t_i, ϕ^{j_i}): Tuple denotes the i^{th} spike of the ensemble, produced by kernel ϕ^{j_i} at time t_i.
- $\Phi^{j_i}(t_i - t)$ or ϕ_i: The kernel function producing the i^{th} spike, inverted and shifted to the time of the spike's occurrence t_i. For brevity, this *function* is alternatively termed as the i^{th} *spike* instead of the tuple notation (t_i, ϕ^{j_i}) and is denoted via the shorthand ϕ_i whenever appropriate. Also, the mathematical definition of the term "spike" must not be confused with the real physical object representing the elicitation of a neuron's action potential. These distinctions should be clear from the context.
- $\mathcal{S}(V)$: Subspace spanned by V in a Hilbert space $\mathcal{H}, V \subseteq \mathcal{H}$.
- $\mathcal{P}_v(u)$: In a Hilbert space $\mathcal{H}$, the projection of u on a vector v for $u, v \in \mathcal{H}$.
- $\mathcal{P}_{\mathcal{S}(V)}(u)$: In a Hilbert space $\mathcal{H}$, the projection of u on the subspace $\mathcal{S}(V)$ for $u \in \mathcal{H}, V \subseteq \mathcal{H}$. Note that this notation is similar to the above $\mathcal{P}_v(u)$ except here the projection is taken with respect to a subspace $\mathcal{S}(V)$ instead of a single vector v. This should be clear from the context.

II. Coding

For encoding, we make the following assumptions:

1) We consider the set of input signals $\mathcal{F}$ to be the class of all *finite-support, bounded functions* (formally, $\mathcal{F} = \{X(t) | t \in [0, \tau], |X(t)| \leq b\}$, for some arbitrary but fixed $\tau, b \in \mathbb{R}^+$) that satisfy a finite rate of innovation bound[22]. Naturally, $X(t) \in L^2$, i.e., square integrable.
2) We assume an ensemble of m spiking neurons $\Phi = \{\Phi^j | j \in Z^+, 1 \leq j \leq m\}$, each characterized by a continuous kernel function $\Phi^j(t)$, where $\forall j, \Phi^j(t) \in C[0, \tau]$, $\tau \in \mathbb{R}^+$. Also we assume that each kernel Φ^j is normalized, i.e., $||\Phi^j||_2 = 1, \forall j$.
3) Lastly, we assume that Φ^j has a time varying threshold $T^j(t)$. The ensemble of kernels encodes an input signal $X(t)$ into a sequence of spikes $\{(t_i, \Phi^{j_i})\}$, where the i^{th} spike of the ensemble is produced by the j_i^{th} kernel Φ^{j_i} at time t_i if and only if: $\int X(t)\Phi^{j_i}(t_i - t)dt = T^{j_i}(t_i)$.

Due to assumptions (1) & (2), our framework operates within a Hilbert space $\mathcal{H}$ of bounded-support, square-integrable functions $L^2[0, \tau]$ for some $\tau \in \mathbb{R}$, endowed with the standard inner product: $\langle f, g \rangle = \int fg, \forall f, g \in \mathcal{H}$. Since these functions reside in a Hilbert space, the algebraic operations on the inputs or the kernels (e.g., the use of the orthogonal projection operator) performed later in this paper are well-defined.

In our implementation, a threshold function is assumed in which the time varying threshold $T^j(t)$ of the j^{th} kernel remains constant at C until that kernel produces a spike, at which

time an *after-hyperpolarization potential (ahp)* increments the threshold by a value M. This increment then returns to zero linearly within a refractory period δ. Formally,

$$T^j(t) = C + \sum_{\substack{p \text{ satisfying} \\ t_p^j \in [t-\delta, t]}} M(1 - \frac{t - t_p^j}{\delta}) \quad (C, M, \delta \in \mathbb{R}^+) \quad (1)$$

where the sum is taken over all spike times t_p^j in the interval $[t-\delta, t]$ at which the kernel Φ^j generated a spike. This threshold function allows a neuron to remain quiescent as long as the signal is uncorrelated with its kernel Φ^j; it fires when the correlation reaches a certain threshold and continues to fire at higher threshold levels communicating increasing correlation levels, only inhibited by its previous spikes. This phenomenon of probing signals via a sequence of spikes is depicted in Fig.1 for one kernel. In Eq.1, the *ahp* model is assumed to approximate the behavior of a biological neuron which undergoes a brief refractory period, typically lasting 1 msec immediately after producing a spike. In our model this is achieved by setting the value of M to a value much higher than the supremum norm of the input signal. As will be evident from subsequent sections, such a model of *ahp* not only bounds the interspike intervals thus ensuring the stability of the spiking framework, it also allows our model to spike at dynamically changing threshold levels starting from a baseline threshold at C for varying correlation levels between the input signal and the kernels, leading to better representational capacity of our model. In what follows, we present a set of corollaries based on the assumptions of our encoding model.

Corollary 0.1. *Let Φ^j be a function in $C[0,\tau]$, where $\tau \in \mathbb{R}^+$ and $||\Phi^j||_2 = 1$. Let $X(t) \in \mathcal{F} = \{f(t) \mid t \in [0,\tau'], |f(t)| \leq b\}$, where $b, \tau' \in \mathbb{R}^+$, be the input to our model. Then: (a) The convolution $C^j(t)$ between $X(t)$ and $\Phi^j(t)$, defined by $C^j(t) = \int X(t')\Phi^j(t-t')dt'$ for $t \in [0, \tau+\tau']$, is a bounded and continuous function. Specifically, one can show that $|C^j(t)| < b\sqrt{\tau}$ for all $t \in [0, \tau+\tau']$. (b) Suppose the parameter M in Eq. 1 is chosen such that $M > 2b\sqrt{\tau}$. Then the interspike interval between any two spikes produced by a given neuron Φ^j is greater than $\frac{\delta}{2}$.*

Proof: The proof has been detailed in section I of[23] since the corollary is intuitively evident. □

Corollary 0.2. *Suppose the assumptions of the Corollary 0.1 hold true. Then, the spike rate generated by our framework for any input signal $X(t)$ is bounded. Consequently, the maximum number of preceding spikes that can overlap with any given spike is bounded above by a constant value.*

Proof. Based on Corollary 0.1, the total spike rate is bounded above by $\frac{2m}{\delta}$, where m is the number of kernels employed by our model. Since each kernel $\Phi^j \in C[0,\tau]$ is of compact support, the maximum number of preceding spikes any given spike can overlap with is bounded above by $\tau\frac{2m}{\delta}$, where τ is the maximum length of support for any kernel Φ^j. □

Corollary 0.3. *Let $X(t)$ be an input signal and Φ^j be a kernel. Suppose the convolution between $X(t)$ and $\Phi^j(t)$ at time t_p is denoted by $C^j(t_p) = \langle X(t), \Phi^j(t_p - t)\rangle > 0$, and let the absolute refractory period be δ as modeled in Eq. 1. If the baseline threshold C is set such that $0 < C \leq C^j(t_p)$, then the kernel Φ^j must produce a spike in the interval $[t_p - \delta, t_p]$, according to the threshold model defined in Eq. 1.*

Proof: We prove by contradiction, utilizing the continuity of the convolution function, $C^j(t)$. **Case 1:** No spike is produced by the kernel Φ^j before or at t_p. By definition $C^j(0) = 0$. However, since $0 < C \leq C^j(t_p)$, the continuous function $C^j(t)$ must intersect the baseline threshold C between $t = 0$ and $t = t_p$ by the intermediate value theorem. This contradiction implies that the kernel Φ^j must produce a spike prior to or at time t_p. **Case 2:** Assuming spikes were generated before or at time t_p by the kernel Φ^j, let t_l be the time of the last spike produced by Φ^j before or at t_p. Suppose $t_l < t_p - \delta$ to ensure no spike in the interval $[t_p - \delta, t_p]$. Then, by Eq. 1, the threshold of kernel Φ^j at t_p is $T^j(t_p) = C$. However, $C^j(t_p) = \langle X(t), \Phi^j(t_p - t)\rangle \geq C$. Since $C^j(t_l) = T^j(t_l)$ and $C^j(t)$ is continuous, and considering Eq. 1, where the *ahp* rises by a high value M at t_l and then linearly decreases to C before t_p, the intermediate value theorem implies that $C^j(t)$ must cross the threshold between t_l and t_p. However, since t_l was the last spike of Φ^j before t_p, this contradicts our assumption. Hence, there must be a spike in $[t_p - \delta, t_p]$. □

Assumption 1. $||\mathcal{P}_{\mathcal{S}(\{\phi_1,\ldots,\phi_{n-1}\})}(\phi_n)|| \leq \beta < 1$, *for some $\beta \in \mathbb{R}$, $\forall n \in \{1, ..., N\}$, where N is the total number of spikes produced by the system. In words, the norm of the projection of every spike onto the span of all previous spikes is bounded from above by some constant strictly less than 1.*

Justification: Corollary 0.1 suggests that for appropriately chosen parameters to the threshold Eq. 1, the spikes produced by the same kernel are sufficiently disjoint in their support. Therefore, each new spike ϕ_n comes with a component that is disjoint in support with respect to the spikes produced by the same kernel. Since the biological kernels of our framework are causal in nature, due to the disjoint component in time a new spike maintains an orthogonal component with respect to all the previous spikes by the same kernel. For spikes produced by different kernels, we observe that different biological kernels correspond to different frequency responses (e.g., it has been observed that the responses of the auditory nerve can be well approximated by a bank of linear gammatone filters[8,24]). Since there are only finitely many kernels in our framework, this leads to the fact that a new spike is poorly represented by the previous spikes produced by other kernels. Overall, a new spike ϕ_n, for appropriately chosen *ahp* parameters in Eq. 1 will not be fully represented by previous spikes either due to disjoincy in support or frequency. Hence, the overall set of spikes grows as a *linearly independent* set. The technical need for this assumption will become clear in later sections.

III. Decoding

The objective of the decoding module is to reconstruct the original signal from the encoded spike train ensemble. We consider the prospect of inverting the coding scheme. Ideally, we seek a signal that satisfies the same set of constraints as the original signal at spike times $\{t_i\}$ apropos the set

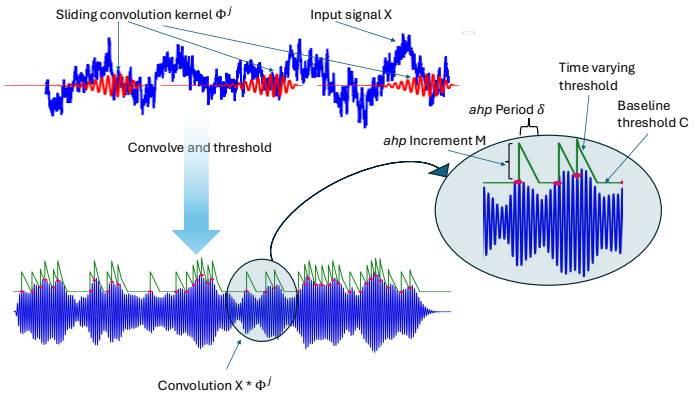

Fig. 1. The convolve and threshold mechanism described in the coding model for a single kernel. Top: a sample signal (in blue) is shown overlayed with a convolution kernel (in red). Below: the result of convolution in blue and the threshold function for the kernel in green. Spikes times are marked at the threshold crossing level with red dots.The zoomed-in portion (enclosed in an oval) on the right highlights the parameters of the threshold Eq. 1.

of kernels in ensemble Φ, and furthermore, that does not generate spikes in the interspike intervals. Since directly imposing non-spiking is intractable, we instead look for the signal with the smallest magnitude that generates the spikes $\{t_i\}$. Since for any fixed Φ^j, convolution is a linear operator on the input, this substitute constraint is well-founded. Separately, noting that the problem is ill-posed when only spike time constraints are considered and recognizing the possibility of independent (Gaussian) measurement noise that would be reflected in spike time jitter, we choose to regularize the reconstructed signal for stability using an L^2 penalty[25]. Formally, the reconstruction X^* of the input signal X is formulated to be the solution to:

$$\begin{aligned} X^* &= \underset{\tilde{X}}{\operatorname{argmin}} ||\tilde{X}||_2^2 \\ \text{s.t.} &\int \tilde{X}(\tau)\Phi^{j_i}(t_i-\tau)d\tau = T^{j_i}(t_i); 1 \le i \le N \end{aligned} \tag{2}$$

where $\{(t_i, \Phi^{j_i}) | i \in \{1, ..., N\}\}$ is the set of all spikes generated by the encoder. Providentially, an L^2 minimization in the objective of (2) reduces the convex optimization problem to a solvable linear system of equations as described below.

IV. Signal Class for Perfect Reconstruction

We observe that in general the encoding of $L^2[0,T]$ signals into spike trains is not an injective map; the same set of spikes can be generated by different signals so as to result in the same convolved values at the spike times. Naturally, with a finite and fixed ensemble of kernels Φ, one cannot achieve perfect reconstruction for all $L^2[0,T]$ signals. Assuming, additionally, a finite rate of innovation, as $\mathcal{F}$ was previously defined changes the story. We now restrict ourselves to a subset $\mathcal{G}$ of $\mathcal{F}$ defined as $\mathcal{G} = \{X | X \in \mathcal{F}, X = \sum_{p=1}^{N} \alpha_p \Phi^{j_p}(t_p - t), j_p \in \{1, ..., m\}, \alpha_p \in \mathbb{R}, t_p \in \mathbb{R}^+, N \in Z^+\}$ and address the question of reconstruction accuracy. Essentially $\mathcal{G}$ consists of all linear combinations of arbitrarily shifted inverted kernel functions. N is bounded above by the total number of spikes that the ensemble Φ can generate over $[0,T]$. For the class $\mathcal{G}$ the *perfect reconstruction theorem* is presented below. The theorem is proved with the help of two lemmas.

Theorem 1. *(Perfect Reconstruction Theorem) Let $X \in \mathcal{G}$ be an input signal. Then for appropriately chosen time-varying thresholds of the kernels, the reconstruction X^*, resulting from the proposed coding-decoding framework is accurate with respect to the L^2 metric, i.e., $||X^* - X||_2 = 0$.*

Lemma 1. *The solution X^* to the reconstruction problem Eq. 2 can be written as: $X^* = \sum_{i=1}^{N} \alpha_i \Phi^{j_i}(t_i - t)$ where the coefficients $\alpha_i \in \mathbb{R}$ can be uniquely solved from a system of linear equations if the set of spikes $\{\phi_i = \Phi^{j_i}(t_i - t)\}_{i=1}^{N}$ produced is* linearly independent.

Proof: An argument similar to that of the Representer Theorem[26] on Eq. 2 directly results in: $X^* = \sum_{i=1}^{N} \alpha_i \Phi^{j_i}(t_i - t)$ where the α_i's are real valued coefficients. This holds true because any component of X^* orthogonal to the span of the $\Phi^{j_i}(t_i - t)$'s does not contribute to the convolution (inner product) constraints. In essence, X^* is an orthogonal projection of X on the span of the spikes $\{\phi_i = \Phi^{j_i}(t_i - t) | i \in \{1, 2, ..., N\}\}$. Hence, the coefficients can be derived by solving the linear system: $P\alpha = T$ where P is the $N \times N$ Gram matrix of the spikes, i.e., $[P]_{ik} = \langle \Phi^{j_i}(t_i - t), \Phi^{j_k}(t_k - t) \rangle$, and $T = \langle T^{j_1}(t_1), \ldots, T^{j_N}(t_N) \rangle^T$. Furthermore, the system has a unique solution if P is invertible, which is the case if the set of spikes $\{\phi_i\}_{i=1}^{N}$ is linearly independent. This in turn follows from Assumption 1. Notably, when the P-matrix is non-invertible we still obtain a unique reconstruction X^* (see section II of[23]) by calculating the α's via the pseudo-inverse of P, but in such a case the system becomes ill-conditioned, an issue that is analyzed separately in section VI. □

Lemma 2. *Let X^* be the reconstruction of an input signal X with $\{\phi_i\}_{i=1}^{N}$ being the set of generated spikes. Then, for any arbitrary signal $\tilde{X}$ within the span of the spikes given by $\tilde{X} = \sum_{i=1}^{N} a_i \phi_i$, $a_i \in \mathbb{R}$, the following holds: $||X - X^*|| \le ||X - \tilde{X}||$.*

Proof: Follows straightforwardly from the fact that X^* is an orthogonal projection on the span (see section II of[23]). □

Exploring further, for a given input signal X, if S_1 and S_2 are two sets of spike trains where $S_1 \subset S_2$ produced by two different kernel ensembles, the second a superset of the first, then Lemma 2 further implies that the reconstruction due to S_2 is at least as good as the reconstruction due to S_1 because the reconstruction due to S_1 is in the span of the shifted kernel functions of S_2 as $S_1 \subset S_2$. This immediately leads to the conclusion that for a given input signal the more kernels we add to the ensemble the better the reconstruction, provided the kernels maintain linear independence.

Proof of Theorem 1: The proof of the theorem follows directly from Lemma 2. Since the input signal $X \in \mathcal{G}$, let X be given by: $X = \sum_{p=1}^{N} \alpha_p \Phi^{j_p}(t_p - t) \quad (\alpha_p \in \mathbb{R}, t_p \in \mathbb{R}^+, N \in Z^+)$. Assume that the time varying thresholds of the kernels in our kernel ensemble Φ are set in such a manner that the following conditions are satisfied: $\langle X, \Phi^{j_p}(t_p - t) \rangle = T^{j_p}(t_p) \quad \forall p \in \{1, ..., N\}$ i.e., each of the kernels Φ^{j_p} at the very least produces a spike at time t_p against X (regardless of other spikes at other times). Clearly then X lies in the span

of the set of spikes generated by the framework. Applying Lemma 2 it follows that: $||X - X^*||_2 \leq ||X - X||_2 = 0$. $\square$

V. Approximate Reconstruction

Theorem 1 stipulates the conditions under which perfect reconstruction is feasible in the purview of our framework. Specifically the theorem shows the ideal conditions—when the input signal lies in the span of shifted kernel functions and the spikes are generated at certain desired locations—where perfect reconstruction is attainable. However, under realistic scenarios such conditions may not be feasible and hence the need for quantification of reconstruction error as the system deviates from the ideal conditions. For example, even though corollary 0.3 shows that a spike can be produced arbitrarily close to the desired location by setting the *ahp* parameters C and the δ of Eq. 1 at reasonably low values, it begs the question to what extent the reconstruction suffers due to small deviations in spike times. Likewise, the input signal may not perfectly fit in the signal class $\mathcal{G}$, i.e. the input may not be exactly representable by the kernel functions due to the presence of internal or external noise. Under such non-ideal scenarios how much the reconstruction suffers is addressed in the following theorem.

Theorem 2. *(Approximate Reconstruction Theorem). Let the input signal X be represented as $X = \sum_{i=1}^{N} \alpha_i f^{p_i}(t_i - t)$, where $\alpha_i \in \mathbb{R}$ and $f^{p_i}(t)$ are bounded functions on finite support. Assume that there is at least one kernel function Φ^{j_i} in the ensemble for which $||f^{p_i}(t) - \Phi^{j_i}(t)||_2 < \delta$ for all $i \in \{1, ..., N\}$. Additionally, assume each kernel Φ^{j_i} produces a spike within a γ interval of t_i, for some δ,$\gamma \in \mathbb{R}^+$, for all i. Further, assume the functions f^{p_i} satisfy a frame bound type of condition: $\sum_{k \neq i} \langle f_{p_i}(t - t_i), f_{p_k}(t - t_k) \rangle \leq \eta \; \forall \, i \in \{1, ..., N\}$, and that the kernel functions are Lipschitz continuous. Under such conditions, the L^2 error in the reconstruction X^* of the input X has bounded SNR.*

Proof: The proof follows from continuity arguments and the use of bounds on the eigen values of the Gram matrix P (see section III in[23]).

VI. Stability of the solution and Windowed Iterative Reconstruction:

Theorem 2 shows that even under non-ideal conditions, our technique can keep the reconstruction error in check through suitable parameter choices. However, this may increase spike rates, as implied by Corollaries 0.1 and 0.3. Higher spike rates exacerbate the condition number of the P matrix referred to in Lemma 1. Instabilities in P render the solutions practically unusable in applications with finite precision floating point representations, due to quantization error. The *ahp* partially mitigates this issue for finite-sized P matrices by ensuring linear independence among the spikes. The inhibitory effect of the *ahp*, as alluded to in assumption 1, results in spikes that are sufficiently disjoint in time, leading to production of a linearly independent set of spikes. But the condition number can get progressively worse as we process longer signals and the size of P grows arbitrarily large. The following Theorem establishes a relation between the condition number of the P matrix and the spike count, revealing that, in the worst case, despite realistic assumptions about spike non-overlap, the condition number can deteriorate exponentially.

Theorem 3 (Condition Number Theorem)**.** *Let $\{P_k\}$ denote the set of all Gram matrices corresponding to any set of k successive spikes $\{\phi_1, \ldots, \phi_k\}$, i.e., $P_k[i,j] = \langle \phi_i, \phi_j \rangle$, where each spike satisfies the Assumption 1: $||\mathcal{P}_{S(\{\phi_1,\ldots,\phi_{i-1}\})}(\phi_i)|| \leq \beta < 1$, $\forall i \in \{2, \ldots, k\}$. If C_k denotes the least upper bound on the condition number of the class of matrices $\{P_k\}$, then $(1-\beta^2)^{-k+1} \leq C_k \leq (1 + (k-1)\beta)(\frac{1-\beta^2}{2})^{-k+1}$*

· **Proof:** The condition number of P_k is defined as $\frac{\Lambda_{max}(P_k)}{\Lambda_{min}(P_k)}$, where $\Lambda_{max}(P_k)$ and $\Lambda_{min}(P_k)$ denote the maximum and minimum eigen values of P_k. First we find the infimum L_k on Λ_{min} over all $\{P_k\}$, the class of all gram-matrices of k successive spikes. We find the infimum inductively on k. By definition,

$$\Lambda_{min}(P_k) = \min_{e,||e||=1} ||\Sigma_{i=1}^{k} e_i \phi_i||^2$$

(where $e = [e_1, ..., e_k]^T$, a k-vector)

$$= \min_{e,||e||=1} [e_k^2 + 2e_k \langle \phi_k, \Sigma_{i=1}^{k-1} e_i \phi_i \rangle + ||\Sigma_{i=1}^{k-1} e_i \phi_i||^2]$$

$$\geq \min_{e,||e||=1} [e_k^2 - 2|e_k|\beta\sqrt{1-e_k^2}||\Sigma_{i=1}^{k-1} \frac{e_i}{\sqrt{1-e_k^2}}\phi_i|| + ||\Sigma_{i=1}^{k-1} \frac{e_i}{\sqrt{1-e_k^2}}\phi_i||^2] \quad \text{(since, } ||\mathcal{P}_{S\{\phi_1,\ldots,\phi_{i-1}\}}(\phi_i)|| \leq \beta)$$

$$\geq \min_{e,||e||=1} [e_k^2 - 2|e_k|\beta z\sqrt{1-e_k^2} + (1-e_k^2)z^2] \tag{3}$$

$$\text{(denoting, } z = ||\Sigma_{i=1}^{k-1} \frac{e_i}{\sqrt{1-e_k^2}}\phi_i||)$$

Now we set $|e_k| = \cos\theta, |e_k| \leq 1$ to obtain:

$$\Lambda_{min}(P_k) \geq \min_{e,||e||=1} [\cos^2\theta - 2\beta z \cos\theta \sin\theta + z^2 \sin^2\theta]$$

$$\geq \min_{e,||e||=1} [\frac{1+z^2}{2} + \frac{1-z^2}{2}\cos 2\theta - \beta z \sin 2\theta]$$

$$\geq \min_{e,||e||=1} [\underbrace{\frac{1+z^2}{2} - \sqrt{(\frac{1-z^2}{2})^2 + \beta^2 z^2}}_{g(z^2)}] \tag{4}$$

But, $z^2 = ||\Sigma_{i=1}^{k-1} \frac{e_i}{\sqrt{1-e_k^2}}\phi_i||^2 \geq L_{k-1}$

(Since, $\Sigma_{i=1}^{k-1} \frac{e_i^2}{1-e_k^2} = 1$ we get this inductively)

Since for $|\beta| < 1$ the expression $g(z^2)$ in 4 is a monotonic in z^2, and L_k is the infimum of $\Lambda_{min}(P_k)$, we may write,

$$L_k \geq [\frac{1+L_{k-1}}{2} - \sqrt{(\frac{1-L_{k-1}}{2})^2 + \beta^2 L_{k-1}}] \tag{5}$$

But L_k is a lower bound of Λ_{min}, and all the inequalities above are tight. Specifically, Eq. 3 & 4 show how given P_{k-1}, a Gram matrix of $k-1$ successive spikes with $\Lambda_{min} = L_{k-1}$, one can choose ϕ_k and e to result in a matrix P_k, so that $\Lambda_{min}(P_k)$ achieves the lower bound of Eq. 5. Therefore, the inequality of Eq. 5 can be turned into an equality.

$$L_k = [\frac{1+L_{k-1}}{2} - \sqrt{(\frac{1-L_{k-1}}{2})^2 + \beta^2 L_{k-1}}]$$

$$L_k = [\frac{(1-\beta^2)L_{k-1}}{\frac{1+L_{k-1}}{2} + \sqrt{(\frac{1-L_{k-1}}{2})^2 + \beta^2 L_{k-1}}}] \quad (6)$$

Since, $|\beta| \geq 0$, setting $|\beta| = 0$ in denominator of Eq. 6,

$$L_k \leq (1-\beta^2)L_{k-1} \quad (7)$$

Again, $L_1 = 1$ and using induction we can get $L_k \leq 1$.

Therefore, setting $L_{k-1} = 1, \beta = 1$ in denominator of Eq. 6,

$$L_k \geq \frac{(1-\beta^2)L_{k-1}}{2} \quad (8)$$

$$\Rightarrow \frac{(1-\beta^2)L_{k-1}}{2} \leq L_k \leq (1-\beta^2)L_{k-1} \qquad \text{(Using 8 \& 7)}$$

$$\Rightarrow (\frac{1-\beta^2}{2})^{k-1} \leq L_k \leq (1-\beta^2)^{k-1} \quad (9)$$

Eq. 9 establishes a bound on the infimum of Λ_{min}. To complete the proof and establish a bound on C_k we need to show a bound on the supremum of $\Lambda_{max}(P_k)$, call it U_k. A bound on U_k can be shown as follows:

$$\Lambda_{max}(P_k) \leq \sup_i (P_k[i,i] + \sum_{i \neq j} |P_k[i,j]|)$$

(using Gershgorin Circle Theorem)

$$= \sup_i \{\langle \phi_i, \phi_i \rangle + \sum_{i \neq j} |\langle \phi_i, \phi_j \rangle|\} \leq (1 + (k-1)\beta)$$

(Since $||\mathcal{P}_{S\{\phi_1,\ldots,\phi_{i-1}\}}(\phi_i)|| \leq \beta \Rightarrow |\langle \phi_i, \phi_j \rangle| \leq \beta$)

$$\Rightarrow 1 \leq U_k \leq (1 + (k-1)\beta) \quad (10)$$

($1 \leq U_k$ is trivial because $\Lambda_{max} = 1$ for $P_k = I$)

Combining Eq. 9 & 10 we get:

$$(1-\beta^2)^{-k+1} \leq C_k \leq (1 + (k-1)\beta)(\frac{1-\beta^2}{2})^{-k+1} \qquad \square$$

The above theorem provides a tight upper bound on the condition number of the P-matrix and clearly shows how the condition number can degrade even if the spikes are sufficiently disjoint in time, i.e. $\beta \approx 0$. This is where the combined effect of causality of the kernels and the *ahp* comes to our defense. We observe that the addition of the $n+1^{th}$ spike on an existing set of n spikes can only affect the solution substantially within a finite time window which in turn is facilitated by the effect of the *ahp*, which ensures that new spikes maintain reasonable separation with the previous spikes and therefore have a fading effect on the reconstruction back into the past. This observation is consistent with biological systems where an animal has to respond in real-time and therefore if an additional spike changes the conceived reconstruction too far into the past it is of little use. This simple observation enables us to encode and reconstruct signals in an online mode within a finite window of past spikes, leading to remarkable efficiency of our proposed coding scheme. In this section we first provide a mathematical formulation of our window-based reconstruction scheme and then we derive conditions under which the window-based reconstruction converges to the optimal reconstruction formulated in Lemma 1. As it turns out, such conditions are easily met in the context of naturally occurring biological kernels, enabling our biologically motivated coding scheme to result in a very efficient and robust solution, as evidenced by our experimental results (section VII).

Windowed iterative reconstruction: Lemma 1 establishes the fact that the reconstruction X^* by our framework is essentially the projection onto the span of all spikes, i.e., $X^* = \mathcal{P}_{\mathcal{S}(\{\phi_1,\ldots\phi_N\})}(X)$. This observation enables us to formulate the reconstruction iteratively by updating an existing reconstruction on a set of n spikes $\{\phi_i\}_1^n$ with each new incoming spike ϕ_{n+1}, instead of solving the $P\alpha = T$ equation for the full set of spikes as shown in lemma 1. The iterative update of the reconstruction then follows from the formula:

$$\mathcal{P}_{\mathcal{S}(\bigcup_{i=1}^{n+1}\{\phi_i\})}(X) = \mathcal{P}_{\mathcal{S}(\bigcup_{i=1}^{n}\{\phi_i\})}(X) + \langle X, \phi_{n+1}^{\perp} \rangle \frac{\phi_{n+1}^{\perp}}{||\phi_{n+1}^{\perp}||^2} \quad (11)$$

where $\phi_{n+1}^{\perp}$ is the orthogonal complement of the additional $n+1$-th spike with respect to the span of all previous spikes, i.e. $\phi_{n+1}^{\perp} = \phi_{n+1} - \mathcal{P}_{\mathcal{S}\{\phi_1,\ldots,\phi_n\}}(\phi_{n+1})$. The above iterative scheme Eq. 11, motivates us to formulate a windowed iterative reconstruction. Here, when a new spike ϕ_{n+1} appears, instead of calculating its orthogonal complement with respect to the span of all previous spike, the orthogonal projection of ϕ_{n+1} is computed with respect to the span of w previous spikes, where w is chosen as a fixed window size. Mathematically, for the $(n+1)$-th incoming spike, we define the *windowed iterative reconstruction*, $X_{n+1,w}^*$, for an input signal X with window size w iteratively as:

$$X_{n+1,w}^* = X_{n,w}^* + \langle X, \phi_{n+1,w}^{\perp} \rangle \frac{\phi_{n+1,w}^{\perp}}{||\phi_{n+1,w}^{\perp}||^2} \quad (12)$$

where $\phi_{n+1,w}^{\perp}$ is defined as:

$$\phi_{n+1,w}^{\perp} = \phi_{n+1} - \mathcal{P}_{\mathcal{S}(\bigcup_{n-w+1}^{n}\{\phi_i\})}(\phi_{n+1})$$

The idea is that $\phi_{n+1,w}^{\perp}$ closely approximates $\phi_{n+1}^{\perp}$ for reasonably large window size w, allowing us to formulate an iterative reconstruction based only on a finite window of w spikes rather than inverting a large P-matrix of size $N \times N$ as formulated in lemma 1. Eq. 12 involves computing $\phi_{n+1,w}^{\perp}$ for each new spike, derived by inverting the $w \times w$ gram matrix corresponding to the previous w spikes. Since w is chosen as a finite constant independent of N, it speeds up the decoding process and holds the condition number of the solution in check as per Theorem 3. The following Theorem 4 establishes how the window-based solution formulated in Eq. 12 converges to the optimal solution of Eq. 1 under an assumption feasible in the context of spikes produced by our framework via biological kernels. The assumption is:

Assumption 2. $||\mathcal{P}_{\mathcal{S}(\{\phi_1,\ldots,\phi_N\}\setminus\{\phi_n\})}(\phi_n)|| \leq \beta < 1$, *for some* $\beta \in \mathbb{R}, \forall n \in \{1,\ldots,N\}$. *In words, the norm of the projection of each spike onto the span of the remaining spikes is bounded above by a constant strictly less than 1.*

Justification: Assumption 2 extends Assumption 1 to the future. For large N, it is possible to construct a spike sequence where each spike satisfies Assumption 1, yet the norm of the projection of an individual spike onto the set

of remaining spikes (including both past and future spikes), i.e. $||\mathcal{P}_{\mathcal{S}(\bigcup_1^N\{\phi_i\}\setminus\{\phi_n\})}(\phi_n)|| \rightarrow 1$. An example illustrating this scenario is shown in Fig. 5, where each spike corresponds to one full cycle of a sine wave, and for any given spike ϕ_n the half wave of its tail (head) precisely overlaps with the half wave of the head (tail) of its previous (next) spike ϕ_{n-1} (ϕ_{n+1}). One can show that in such a scenario $||\mathcal{P}_{\mathcal{S}(\bigcup_1^N\{\phi_i\}\setminus\{\phi_n\})}(\phi_n)||$ converges to 1 for large values of n and N (see section IV of[23] for details). This convergence occurs because within the compact support of a spike, the spike is fully represented by the overlapping components of the neighboring spikes. Conversely, if a spike is poorly represented within its compact support by the overlapping components of its neighboring spikes, it satisfies the condition of Assumption 2. Specifically, if within its compact support, a spike ϕ_n can produce a component $\hat{\phi_n}$, orthogonal to all overlapping parts of the neighboring spikes with $||\hat{\phi_n}|| > 0$, then ϕ_n satisfies the condition of Assumption 2, because $\hat{\phi_n}$ is orthogonal to the overlapping parts of the neighboring spikes and hence $\hat{\phi_n}$ is orthogonal to every spike other than ϕ_n. This observation enables us to assert Assumption 2 for spikes produced by our framework via the biological kernels (e.g., the gammatone kernels corresponding to the auditory processing[24]) which inherently exhibit causal and fading memory properties, so that the overlapping parts of the neighboring spikes do not represent the given spike within its support. Fig. 2 visually illustrates this property using the dot product of overlapping portions of gammatone kernels. We show that due to the *ahp*, overlapping spikes corresponding to the same gammatone kernel poorly represent each other as they temporally separate. Additionally, spikes corresponding to gammatones of different frequencies poorly represent each other, irrespective of temporal shifts. Thus, with appropriately chosen *ahp* parameters, the spikes produced by our framework's biological kernels are poorly represented within their supports by overlapping components of neighboring spikes, either due to temporal disjoincy or frequency mismatch. Since every spike overlaps with only finitely many other spikes (corollary 0.2), Assumption 2 holds. An important consequence of Assumption 2 is that the norm of the projection of every unit vector in a subspace of a finite partition of spikes onto the remaining subspace of spikes is strictly less than 1. This condition forms the basis for establishing the convergence of *windowed iterative reconstruction* to optimal reconstruction in our subsequent windowing theorem 4. We formally state the condition in Lemma 3 before presenting Windowing theorem 4.

Lemma 3. *Let $S = \{\phi_i\}_{i=1}^N$ denote the set of spikes generated by our framework, satisfying Assumption 2, i.e., $\forall n \in \{1,...,N\}$, $||\mathcal{P}_{\mathcal{S}(\bigcup_{i=1}^N\{\phi_i\}\setminus\{\phi_n\})}(\phi_n)|| \leq \beta$, where $\beta \in \mathbb{R}$ is a constant strictly less than 1. Consider a subset $V \subseteq S$ of a finite size d, $d < N$. Then, for every $v \in \mathcal{S}(V)$ with $||v|| = 1$, $\exists \beta_d < 1$, such that $||\mathcal{P}_{\mathcal{S}(S\setminus V)}(v)|| \leq \beta_d$ where β_d is a real constant that depends on β and d. Specifically, we can show that $\beta_d^2 \leq (1 + \frac{1-\beta^2}{d^2\beta^2})^{-1} < 1$.*

Proof: Proof in Section V of[23]. □

Theorem 4 (Windowing Theorem)**.** *For an input signal X with bounded L^2 norm, suppose our framework produces a set of $n+1$ successive spikes $S = \{\phi_1, ..., \phi_{n+1}\}$, sorted by their time of occurrence and satisfying Assumption 2. The error in the iterative reconstruction of X with respect to the last spike ϕ_{n+1} due to windowing, as formulated in Eq. 12, is bounded. Specifically,*

$$\forall \epsilon > 0, \exists w_0 > 0 \text{ s.t. } ||\mathcal{P}_{\phi_{n+1,w}^{\perp}}(X) - \mathcal{P}_{\phi_{n+1}^{\perp}}(X)|| < \epsilon, \quad \forall w \geq w_0 \text{ and } w \leq n \tag{13}$$

where w_0 is independent of n for arbitrarily large $n \in \mathbb{N}$.

Import and Proof Idea: The Theorem implies that the error from windowing can be made arbitrarily small by choosing a sufficiently large window size w_0, independent of n, when n is arbitrarily large. At first glance, one might think that the condition in Eq. 13 is trivially satisfied by choosing $w_0 = n$, i.e., a window inclusive of all spikes. However, the key aspect of the theorem is that w_0 should be independent of n when n is arbitrarily large. This allows us to use the same window size regardless of the number of previous spikes, even for large signals producing many spikes, thereby maintaining the condition number of the overall solution as per Theorem 3. Our proof demonstrates this by showing that the reconstruction error converges geometrically as a function of the window size, depending only on the spike rate which in turn depends on the *ahp* parameters in Eq. 1, but not on n. The proof hinges on a central lemma showing that the L_2 norm of the difference between $\phi_{n+1,w}^{\perp}$ and $\phi_{n+1}^{\perp}$ decreases steadily as w increases, based on the assumptions stated in the theorem. This ensures that $\phi_{n+1,w}^{\perp}$ becomes a good approximation of $\phi_{n+1}^{\perp}$. The proof of the theorem then follows by establishing a bound on $\|\mathcal{P}_{\phi_{n+1,w}^{\perp}}(X) - \mathcal{P}_{\phi_{n+1}^{\perp}}(X)\|$ for a given choice of w for any bounded input X. The lemma is provided below:

Lemma 4. *Under the conditions of Theorem 4, for any $\delta > 0$, there exists $w_0 \in \mathbb{N}^+$ such that $\|\phi_{n+1,w}^{\perp} - \phi_{n+1}^{\perp}\| < \delta \ \forall w \geq w_0, w \leq n$, where the choice of w_0 is independent of n for arbitrarily large $n \in \mathbb{N}^+$.*

Proof Idea: The proof leverages Corollary 0.2, which states that the maximum number of spikes overlapping in time is bounded by a constant $d \in \mathbb{N}^+$, dependent on the *ahp* parameters. Using this corollary, we partition the set of all spikes in time into a chain of subsets, where each subset overlaps only with its neighboring subsets and is disjoint from all others. Each subset contains at most d spikes. The proof then demonstrates that the error in approximating $\phi_{n+1}^{\perp}$ due to windowing, i.e., $\|\phi_{n+1,w}^{\perp} - \phi_{n+1}^{\perp}\|$, decreases faster than a geometric sequence as more of these partitions are included within the window. This convergence is illustrated schematically in Fig. 3.

Proof of Lemma 4: Let the set of spikes $S = \{\phi_1, ..., \phi_n\}$ be partitioned into a chain of subsets of spikes $v_1, ..., v_m$ in descending order of time, defined recursively. The first subset v_1 consists of all previous spikes overlapping with the support of ϕ_{n+1}. Recursively, v_{i+1} is the set of spikes overlapping with the support of any spike in v_i for all $i \geq 1$. This process continues until the first spike ϕ_1 is included in the final subset v_m, where $m \leq n$. An example of this partitioning is illustrated in Fig. 3. The individual spikes in each partition are

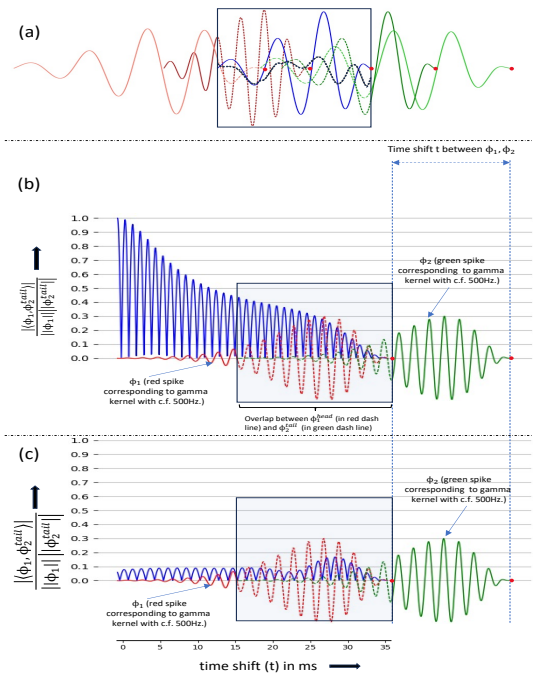

Fig. 2. Illustration of why a spike produced by our framework using gammatone kernels is poorly represented by other overlapping spikes within its temporal support. **(a)** Shows a diagram of five spikes; the central spike (blue dashed line within the rectangle) overlaps with two past spikes (red) and two future spikes (green). Despite the overlaps, the distinct shapes of the gammatone kernels result in a poor representation of the central spike, leaving a component orthogonal to its neighbors (shown in black). **(b)** Details the interaction between two spikes, ϕ_1 and ϕ_2 (red and green curves), both using a gammatone kernel with a center frequency of 500 Hz. It examines how well ϕ_1 is represented by the overlapping tail of ϕ_2 (denoted ϕ_2^{tail}, green dashed line), focusing on representation within this overlapping support only. The graph in blue measures the dot product $\frac{|\langle \phi_1, \phi_2^{tail} \rangle|}{||\phi_1|| ||\phi_2^{tail}||}$ as a function of the time lag t between the two spikes, illustrating how the representation deteriorates rapidly as the lag increases, thus demonstrating the poor representational quality induced by the *ahp-effect*. **(c)** A similar plot of interaction between two spikes, ϕ_1 and ϕ_2 as in **(b)**, except here the center frequencies of the corresponding gammatone kernels are different (500 Hz for ϕ_1 and 400 Hz for ϕ_2). The blue graph represents the dot product $\frac{|\langle \phi_1, \phi_2^{tail} \rangle|}{||\phi_1|| ||\phi_2^{tail}||}$ as a function of time shift t, highlighting systematic poor representation due to frequency differences. Each spike's time of occurrence is marked by a red dot.

indexed as follows:

$$v_i = \{\phi_{p_i}, ..., \phi_{p_{i-1}-1}\}, \forall i \leq m$$
$$\text{where } 1 = p_m < ... < p_1 < p_0 = n + 1.$$

That is, the i^{th} partition v_i consists of spikes indexed from p_i to $p_{i-1} - 1$. Since the spikes $\phi_1, ..., \phi_n$ are sorted in order of their occurrence time, if both ϕ_{p_i} and $\phi_{p_{i-1}-1}$ are in partition v_i, then by construction, $\phi_j \in v_i$ for all $p_i \leq j \leq p_{i-1} - 1$.

Claim 4.1. *The number of spikes in every partition, v_i $\forall i \in \{1, ..., m\}$, is bounded by some constant $d \in \mathbb{N}^+$.*

Proof: This corollary follows from the observation that all spikes in a given partition v_{i+1} overlap in time with at least one spike from the preceding partition v_i, specifically the spike whose support extends furthest into the past, for all $i \in \{2, .., m\}$ (see Fig. 3). In the case of the partition v_1, each spike overlaps with ϕ_{n+1}. Then the proof follows from the corollary 0.2. □

Now, let $V_1, ..., V_m$ be subspaces spanned by the subsets of spikes $v_1, ..., v_m$ respectively. Before proceeding with the rest of the proof, we introduce the following notations.

$$\tilde{\phi}_i = \phi_i - \mathcal{P}_{\mathcal{S}(\bigcup_{j=i+1}^{n} \{\phi_j\})}(\phi_i) \quad \forall i \in \{1, ..., n\}$$

where $\tilde{\phi}_i$ denotes the orthogonal complement of the spike ϕ_i with respect to all the future spikes up to ϕ_n. We also denote,

$$\tilde{V}_k = \mathcal{S}(\bigcup_{j=p_k}^{p_{k-1}-1} \{\tilde{\phi}_j\}) \quad \forall k \in \{1, ..., m\}$$

where $\tilde{V}_k$ is the subspace spanned by spikes in the partition v_k, with each spike orthogonalized with respect to all future spikes.

Claim 4.2. *For the subspaces defined on the partitions as above, the following holds:*

$$\tilde{V}_k = \{x - \mathcal{P}_{\sum_{j=1}^{k-1} V_j}(x) | x \in V_k\} = \mathcal{S}\{\tilde{\phi}_{p_k}, ..., \tilde{\phi}_{p_{k-1}-1}\}.$$

Proof: The proof of Claim 4.2 follows from the properties of orthogonal projection in a Hilbert space (see[23]). □

Following the Claim, we define the subspace U_k as below:

$$U_k = \sum_{i=k}^{m} \tilde{V}_i = \mathcal{S}(\bigcup_{j=1}^{p_{k-1}-1} \{\tilde{\phi}_j\}) \tag{14}$$

$$= \{x - \mathcal{P}_{\sum_{j=1}^{k-1} V_j}(x) | x \in \sum_{i=k}^{m} V_i\} \tag{15}$$

where the equality between 14 and 15 follows from Claim 4.2. Lastly, define $\phi^{\perp}_{n+1,v_k}$ as:

$$\phi^{\perp}_{n+1,v_k} = \phi_{n+1} - \mathcal{P}_{\sum_{i=1}^{k} V_i}(\phi_{n+1})$$

i.e. $\phi^{\perp}_{n+1,v_k}$ is the orthogonal complement of ϕ_{n+1} with respect to window of spikes up to partition v_k. Now we proceed to quantify the norm of the difference of $\phi^{\perp}_{n+1}$ and $\phi^{\perp}_{n+1,v_k}$. For that denote e_k as follows:

$$\begin{aligned}
e_k &= ||\phi^{\perp}_{n+1,v_k} - \phi^{\perp}_{n+1}|| \\
&= ||\mathcal{P}_{\sum_{i=1}^{k} V_i}(\phi_{n+1}) - \mathcal{P}_{\sum_{i=1}^{m} V_i}(\phi_{n+1})|| \\
&= ||\mathcal{P}_{\sum_{i=1}^{k} V_i}(\phi_{n+1}) - \mathcal{P}_{\sum_{i=1}^{k} V_i + U_{k+1}}(\phi_{n+1})|| \text{(by def. of } U_k) \\
&= ||\mathcal{P}_{\sum_{i=1}^{k} V_i}(\phi_{n+1}) - (\mathcal{P}_{\sum_{i=1}^{k} V_i}(\phi_{n+1}) + \mathcal{P}_{U_{k+1}}(\phi_{n+1}))|| \\
&\qquad \text{(Since by construction } U_{k+1} \perp \Sigma_{i=1}^{k} V_i) \\
&\Rightarrow e_k = ||\mathcal{P}_{U_{k+1}}(\phi_{n+1}))|| \qquad (16)
\end{aligned}$$

Note that by definition $e_m = 0$ and by Assumption 2 we get,

$$e_k = ||\mathcal{P}_{U_{k+1}}(\phi_{n+1}))|| \leq |\beta| < 1, \ \forall k \in \{1, ..., m\} \tag{17}$$

Now Assume that:

$$\mathcal{P}_{U_{k+1}}(\phi_{n+1})) = \alpha_k \tilde{\psi}_k, \text{where } \alpha_k \in \mathcal{R}, \tilde{\psi}_k \in U_{k+1}$$

It follows from definition of U_{k+1} that $\tilde{\psi}_k$ is of the form:

$\tilde{\psi}_k = \psi_k - \mathcal{P}_{\sum_{j=1}^{k} V_i}(\psi_k)$ for some $\psi_k \in \sum_{i=k+1}^{m} V_i$

w.l.o.g. assume $||\psi_k|| = 1$ by appropriately choosing α_k.

Since $\alpha_k\tilde{\psi}_k$ is projection of $\phi_{n+1}, \alpha_k = \frac{\langle \phi_{n+1}, \tilde{\psi}_k \rangle}{||\tilde{\psi}_k||^2}$ (18)

For $k > 0$ combining 18 and 16 we obtain,

$$e_k = ||\alpha_k\tilde{\psi}_k|| = |\alpha_k|||\tilde{\psi}_k|| = \frac{|\langle \phi_{n+1}, \tilde{\psi}_k \rangle|}{||\tilde{\psi}_k||}$$
$$= \frac{|\langle \phi_{n+1}, \psi_k - \mathcal{P}_{\sum_{j=1}^{k} V_j}(\psi_k) \rangle|}{||\tilde{\psi}_k||}$$
$$= \frac{|\langle \phi_{n+1}, \mathcal{P}_{\sum_{j=1}^{k} V_j}(\psi_k) \rangle|}{||\tilde{\psi}_k||} = \frac{|\langle \phi_{n+1}, \mathcal{P}_{\sum_{j=1}^{k-1} V_j + \tilde{V}_k}(\psi_k) \rangle|}{||\tilde{\psi}_k||}$$

(for $k > 0, \psi_k \in \sum_{i=k+1}^{m} V_i \perp \phi_{n+1}$ due to disjoint support)

$$= \frac{|\langle \phi_{n+1}, \mathcal{P}_{\sum_{j=1}^{k-1} V_j}(\psi_k) + \mathcal{P}_{\tilde{V}_k}(\psi_k) \rangle|}{||\tilde{\psi}_k||} \text{ (by def. } \tilde{V}_k \perp \sum_{j=1}^{k-1} V_j)$$

$$\Rightarrow e_k = \frac{|\langle \phi_{n+1}, \mathcal{P}_{\tilde{V}_k}(\psi_k) \rangle|}{||\tilde{\psi}_k||} \; (k > 0, \psi_k \in \sum_{i=k+1}^{m} V_i \perp \sum_{j=1}^{k-1} V_j) \quad (19)$$

Likewise, $e_{k-1} = ||\mathcal{P}_{U_k}(\phi_{n+1}))|| = ||\mathcal{P}_{U_{k+1} + \tilde{V}_k}(\phi_{n+1}))||$

$$\Rightarrow e_{k-1}^2 = ||\mathcal{P}_{U_{k+1}}(\phi_{n+1}))||^2 + ||\mathcal{P}_{\tilde{V}_k}(\phi_{n+1}))||^2$$

(Since by construction $U_{k+1} \perp \tilde{V}_k$)

$$\Rightarrow e_{k-1}^2 = e_k^2 + ||\mathcal{P}_{\tilde{V}_k}(\phi_{n+1}))||^2 \quad (20)$$

Assume that $\mathcal{P}_{\tilde{V}_k}(\phi_{n+1}) = \beta_k\tilde{\theta}_k$, where $\tilde{\theta}_k \in \tilde{V}_k, \beta_k \in \mathcal{R}$

$$\Rightarrow ||\mathcal{P}_{\tilde{V}_k}(\phi_{n+1}))|| = ||\beta_k\tilde{\theta}_k|| = \frac{|\langle \phi_{n+1}, \tilde{\theta}_k \rangle|}{||\tilde{\theta}_k||} \quad (21)$$

From 20 & 21 we get, $e_{k-1}^2 = e_k^2 + \frac{|\langle \phi_{n+1}, \tilde{\theta}_k \rangle|^2}{||\tilde{\theta}_k||^2}$ (22)

Again, for $k > 0$ we further analyze e_k from 19 to obtain:

$$e_k = \frac{|\langle \phi_{n+1}, \mathcal{P}_{\tilde{V}_k}(\psi_k) \rangle|}{||\tilde{\psi}_k||}$$
$$= \frac{|\langle \phi_{n+1}, \mathcal{P}_{\mathcal{S}(\{\tilde{\theta}_k\})}(\psi_k) + \mathcal{P}_{\tilde{V}_k \setminus \mathcal{S}(\{\tilde{\theta}_k\})}(\psi_k) \rangle|}{||\tilde{\psi}_k||} \quad (23)$$

The expression 23 is essentially written by breaking ψ_k into two mutually orthogonal subspaces: $\mathcal{S}(\{\tilde{\theta}_k\})$, subspace of $\tilde{V_k}$ spanned by $\tilde{\theta}_k$, and $\tilde{V_k} \setminus \mathcal{S}(\{\tilde{\theta}_k\}$, the subspace of $\tilde{V_k}$ orthogonal to the subspace $\mathcal{S}(\{\tilde{\theta}_k\})$. Also, observe that $\phi_{n+1} \perp \tilde{V_k} \setminus \mathcal{S}(\{\tilde{\theta}_k\})$ since $\mathcal{P}_{\tilde{V}_k}(\phi_{n+1})) = \beta_k\tilde{\theta}_k$. Therefore,

$$e_k = \frac{|\langle \phi_{n+1}, \mathcal{P}_{\mathcal{S}(\{\tilde{\theta}_k\})}(\psi_k) \rangle|}{||\tilde{\psi}_k||} = \frac{|\langle \phi_{n+1}, \frac{\langle \psi_k, \tilde{\theta}_k \rangle \tilde{\theta}_k}{||\tilde{\theta}_k||^2} \rangle|}{||\tilde{\psi}_k||}$$

$$\Rightarrow e_k = \frac{|\langle \phi_{n+1}, \tilde{\theta}_k \rangle|}{||\tilde{\theta}_k||} \frac{|\langle \psi_k, \frac{\tilde{\theta}_k}{||\tilde{\theta}_k||} \rangle|}{||\tilde{\psi}_k||} \quad (24)$$

Combining 22 and 24 we obtain:

$$e_{k-1}^2 = e_k^2 + e_k^2 \frac{||\tilde{\psi}_k||^2}{|\langle \psi_k, \frac{\tilde{\theta}_k}{||\tilde{\theta}_k||} \rangle|^2} \quad (25)$$

Since both ψ_k and $\frac{\tilde{\theta}_k}{||\tilde{\theta}_k||}$ are unit norm $|\langle \psi_k, \frac{\tilde{\theta}_k}{||\tilde{\theta}_k||} \rangle| \leq 1$.

$$\Rightarrow e_{k-1}^2 \geq e_k^2(1 + ||\tilde{\psi}_k||^2) \quad (26)$$

Eq. 26 demonstrates that the sequence $\{e_k\}$ converges geometrically to 0 when $(1+||\tilde{\psi}_k||^2) > 1$, i.e. $||\tilde{\psi}_k|| > 0$. To complete the proof of Lemma 4, we need to establish a positive lower bound on $||\tilde{\psi}_k||$, ensuring the convergence of the sequence e_k. The following Corollary provides the necessary lower bound on $||\tilde{\psi}_k||$.

Claim 4.3. *Following Claim 4.1, if the number of spikes in each partition v_i is bounded by d, then $||\tilde{\psi}_k||^2 \geq 1 - \beta_d^2$, where $\beta_d > 0$ is as defined in Lemma 3.*

Proof: Since ψ_k is a unit vector in $\sum_{i=k+1}^{m} V_i$ and has disjoint support with respect to the future partitions $v_{k-1}, ..., v_1$, except for the partition the v_k, the projection of ψ_k on V_k and the application of Claim 4.1 and Lemma 3 lead to the result. See[23] for details. □

Finally, combining 26 with Claim 4.3, we obtain:

$$e_{k-1}^2 \geq e_k^2(1 + (1 - \beta_d^2)) \Rightarrow e_k^2 \leq \frac{e_{k-1}^2}{\gamma^2} \quad (27)$$

where $\gamma^2 = (1 + (1 - \beta_d^2))$ is a constant strictly greater than 1. Since $e_m = 0$, Eq. 27 shows that the sequence $\{e_k\}_{k=1}^{m}$ converges to 0 faster than geometrically. Thus, $\forall \delta > 0, \exists k_0 \in \mathbb{N}^+$such that $e_k < \delta$, for all $k \geq k_0$ and $m \geq k$. Given the geometric drop in Eq. 27 and the bound $e_1 < 1$ (Eq. 17), for arbitrarily large m (hence n) the choice of k_0 is independent of m and depends only on β_d, which is determined by the *ahp* parameters in Eq. 1). Since the number of spikes in each partition is bounded by d (Claim 4.1), choosing a window size $w_0 = k_0 * d$ ensures $\forall \delta > 0, \exists w_0 \in \mathbb{N}^+$ such that $||\phi_{n+1,w}^{\perp} - \phi_{n+1}^{\perp}|| < \delta$ for all $w \geq w_0$ and $w \leq n$. For arbitrarily large n, the choice of w_0 is independent of n. □

Proof of Theorem 4: Having established a bound on the norm of the difference between $\phi_{n+1}^{\perp}$ and $\phi_{n+1,w}^{\perp}$, we need to bound the norm of the difference between the projections of the input signal X with respect to these vectors. Specifically, we seek to bound $||\mathcal{P}_{\phi_{n+1,w}^{\perp}}(X) - \mathcal{P}_{\phi_{n+1}^{\perp}}(X)||$ based on the window size. We use the following notations: $\tilde{X} = \mathcal{P}_{\mathcal{S}(\{\phi_{n+1,w}^{\perp}, \phi_{n+1}^{\perp}\})}(X), X_u = \mathcal{P}_{\phi_{n+1}^{\perp}}(X), X_v = \mathcal{P}_{\phi_{n+1,w}^{\perp}}(X)$, so that $\tilde{X}, X_u$ and X_v lie in a plane (Fig. 4). Assume the angle between $\tilde{X}$ and $\phi_{n+1}^{\perp}$ is a, and the angle between $\tilde{X}$ and $\phi_{n+1,w}^{\perp}$ is b. Hence, the angle between $\phi_{n+1,w}^{\perp}$ and $\phi_{n+1}^{\perp}$ is $a - b$ (Fig. 4). Note that X may not lie in the same plane as X_u and X_v; thus, we consider the projection $\tilde{X}$ of X on the plane. We denote $p_w = \phi_{n+1,w}^{\perp} - \phi_{n+1}^{\perp}$. Further analysis shows $p_w = \mathcal{P}_{\mathcal{S}(\{\tilde{\phi}_1, ..., \tilde{\phi}_w\})}(\phi_{n+1,w}^{\perp})$, and thus $p_w \perp \phi_{n+1}^{\perp}$ (see[23] for details). Therefore, from Fig. 4 we observe:

$$\sin^2(a - b) = \frac{||p_w||^2}{||\phi_{n+1,w}^{\perp}||^2} \quad (28)$$

Now we can quantify the norm of the difference in two the projections of X w.r.t. $\phi_{n+1}^{\perp}$ and $\phi_{n+1,w}^{\perp}$ as follows:

$$||\mathcal{P}_{\phi^{\perp}_{n+1,w}}(X) - \mathcal{P}_{\phi^{\perp}_{n+1}}(X)|| = ||X_u - X_v||^2$$
$$= ||X_u||^2 + ||X_v||^2 - 2||X_u||||X_v||cos(a-b)$$
$$= ||\tilde{X}||^2[\cos^2 a + \cos^2 b - 2\cos a \cos b \cos(a-b)] \quad \text{(see Fig. 4)}$$
$$= ||\tilde{X}||^2 \sin^2(a-b) = ||\tilde{X}||^2 \frac{||p_w||^2}{||\phi^{\perp}_{n+1,w}||^2}$$

(Using trigonometric identities & Eq. 28, see[23] for details)

$$\leq \frac{||X||^2}{1-\beta^2}||\phi^{\perp}_{n+1,w} - \phi^{\perp}_{n+1}||^2 \quad \text{(Since } ||\tilde{X}|| \leq ||X|| \text{ and}$$
$$||\phi^{\perp}_{n+1,w}||^2 \geq ||\phi^{\perp}_{n+1}||^2 \geq (1-\beta^2) \text{ by assumption)}$$

Finally, since the input signal has a bounded L_2 norm (i.e. $||X||$ is bounded), setting $\delta = \epsilon\frac{\sqrt{1-\beta^2}}{||X||}$ in Lemma 4 gives a window size w_0 such that $||\mathcal{P}_{\phi^{\perp}_{n+1,w}}(X) - \mathcal{P}_{\phi^{\perp}_{n+1}}(X)|| \leq \frac{||X||}{\sqrt{1-\beta^2}}||\phi^{\perp}_{n+1,w} - \phi^{\perp}_{n+1}|| < \epsilon, \forall\epsilon > 0, \forall w \geq w_0$ and $w \leq n$. The choice of w_0 is independent of n for arbitrarily large $n \in \mathbb{N}$. This completes the proof of the theorem. □

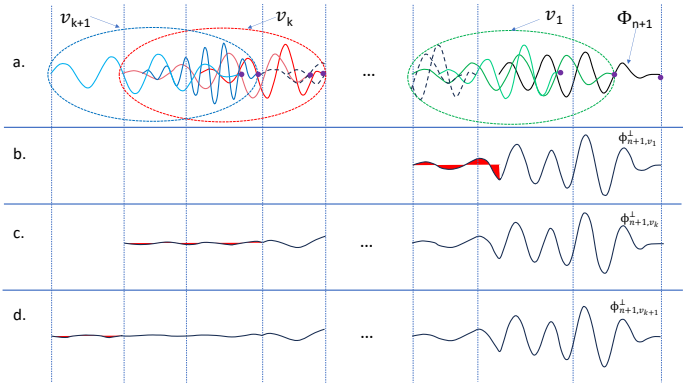

Fig. 3. Illustration of Lemma 4. The diagram shows the convergence of the windowed orthogonal complement $\phi^{\perp}_{n+1,w}$ of spike ϕ_{n+1} to $\phi^{\perp}_{n+1}$ by orthogonalizing ϕ_{n+1} across the partitions of spikes. (a) Displays all spikes up to ϕ_{n+1} (black), with partitions circled: v_1 (green), v_k (red), and v_{k+1} (blue). Spikes contained in each partition are shaded accordingly, with the time of each spike marked by a purple dot. (b), (c), and (d) show orthogonal complements $\phi^{\perp}_{n+1,v_1}$, $\phi^{\perp}_{n+1,v_k}$, and $\phi^{\perp}_{n+1,v_{k+1}}$ respectively. The support of $\phi^{\perp}_{n+1,w}$ extends as more partitions are included, with the extending tail for each additional partition highlighted in red. This tail's diminishing energy as more partitions are added illustrates Lemma 4.

VII. Experiments on Real Signals

The proposed framework was tested on audio signals. **Dataset:** We chose the Freesound Dataset Kaggle 2018, an audio dataset of natural sounds referred in[27], containing 18,873 audio files. All audio samples in this dataset are provided as uncompressed PCM 16bit, 44.1kHz, mono audio files. **Set of Kernels:** We chose gammatone filters $(at^{n-1}e^{-2\pi bt}\cos(2\pi ft + \phi))$ in our experiments since they are widely used as a reasonable model of cochlear filters in auditory systems[24]. Going by Theorem 2, increasing the number of kernels clearly enhances the representation capacity, but it comes at the cost of decreased runtime efficiency of reconstruction and higher spike rate of the ensemble, which in turn can adversely affect the stability of the system. Keeping this trade-off in mind, we ran an experiment with a small set of audio snippets, as detailed in Section VII of the

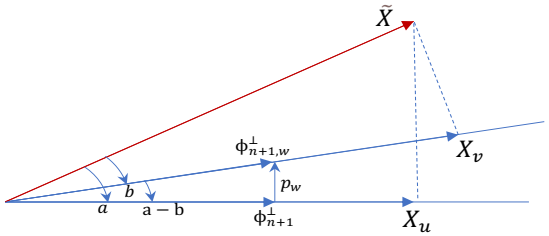

Fig. 4. Figure illustrating the vector projections of the input signal X onto vectors $\phi^{\perp}_{n+1}$ and $\phi^{\perp}_{n+1,w}$. The red vector represents $\tilde{X}$, the projection of X within the plane formed by $\phi^{\perp}_{n+1}$ and $\phi^{\perp}_{n+1,w}$. The vectors $\phi^{\perp}_{n+1}$ and $\phi^{\perp}_{n+1,w}$, as well as the projections X_u and X_v of $\tilde{X}$ onto them, are indicated in blue. The vector p_w, representing the difference between $\phi^{\perp}_{n+1,w}$ and $\phi^{\perp}_{n+1}$, is also shown in blue. The angles a between $\tilde{X}$ and X_u, b between $\tilde{X}$ and X_v, and $a-b$ between $\phi^{\perp}_{n+1,w}$ and $\phi^{\perp}_{n+1}$ are marked.

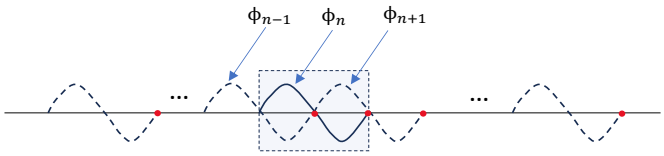

Fig. 5. The scenario illustrating the need for Assumption 2. See text for details. For derivation of $||\mathcal{P}_{\mathcal{S}(\bigcup_1^N\{\phi_i\}\setminus\{\phi_n\})}(\phi_n)|| \to 1$ see[23].

supplementary[23] and based on the results chose a set of 50 gammatone kernels for our large scale experiment. **Results:** The proposed framework was tested extensively against the mentioned dataset. Comprehensive results with 600 randomly selected audio snippets using 50 kernels are shown in Fig. 6. The complete source code used for the experiment is available in GitHub[28]. In the experiment, kernels were normalized, and parameters C and M for the time-varying threshold function Eq. 1 were selected through a systematic grid search on a smaller dataset of 20 randomly chosen snippets. The details of the grid search process and range of values tested for the hyperparameters are detailed in Section VII of[23]. δ was then varied across trials to generate codes at different spike rates (see[23] for details). In each trial, an audio snippet of length $\approx 2.5s$ was processed with fixed parameter values, except for the refractory period δ which was stepped down across trials leading to improvement in reconstructions at higher spike rates. δ varied from approximately $250ms$ to $5ms$. After converting an audio snippet into a spike train ensemble, reconstruction was performed iteratively using a fixed-sized window of spikes as described in Section VI. The refractory period δ was systematically varied, but each trial on an audio snippet was run with a single value of δ. On each trial a reconstruction was performed with a suitable choice of the window size which was decided based on the refractory period δ. The choice of the window size, going by Theorem 4, is critical and subject to the trade-off between reconstruction accuracy and efficiency of reconstruction. Therefore, on a small scale experiment, we systematically tested different window sizes for reconstruction at varying spike rates, as detailed in Section VII of the supplementary. Based on the results, for reconstructions in our large scale experiments we chose a variable window size from 5k to 15k that was inversely proportional to the refractory period.

The results in Fig. 6 show that increasing the spike rate by

tuning the refractory period allows near-perfect reconstruction, aligning with our theoretical analysis. Some variability in reconstruction accuracy across signals can be attributed to dataset idiosyncrasies, e.g., certain audio samples could be noisy or ill-represented in the kernels. However, the overall trend shows promise, achieving an average of $\approx 20dB$ SNR at $1/3$th Nyquist Rate. This in conjunction with the fact that signals are represented in this scheme only via set of spike times and kernel indexes (thresholds can be inferred) shows potential for an extremely efficient coding mechanism. Since the generation of spikes requires scanning through convolutions in a single pass, encoding is highly efficient. However, decoding is slightly more time-consuming because it involves solving the linear system $P\alpha = T$ to derive the coefficients. But then reconstruction is performed iteratively on a finite window, as described in Section VI. Considering $O(w^3)$ as the time complexity for inverting a $w \times w$ matrix, the overall time complexity of decoding is $O(Nw^3)$, where N is the length of the signal and w is the chosen window size. Thus the overall process still remains linear, making it a suitable choice for lengthy continuous-time signals.

Comparison With Convolutional Sparse Coding: Our proposed framework is comparable to Convolutional Sparse Coding (CSC) techniques[21]. In CSC, the objective is to efficiently represent signals using a small number of basis functions convolved with sparse coefficients. Mathematically, for a given signal Y, the CSC model can be expressed as: $Y \approx \Sigma_{k=1}^{K} d_k \star x_k$ where d_k are the convolutional filters, x_k are the sparse feature maps, K is the number of filters and $\star$ denotes convolution. The similarity with our proposed framework becomes clear when we consider our reconstruction formulation: $X^* = \sum_{i=1}^{N} \alpha_i \Phi^{j_i}(t_i - t)$. Here, the kernels Φ^k serve as the convolution filters d_k, and the sparse feature map x_k in our framework can be likened to a vector of coefficients α_i, where each α_i corresponds to the spikes of the kernel Φ^k, and is time-shifted to the appropriate occurrences of spikes. In other words, our framework finds a sparse representation of a signal through its own biological spiking mechanism. Finding sparse code for signals, in general is an NP-Hard problem[29], and several heuristic-based approaches are used to address this challenge. In[30], our proposed framework was compared against one such heuristic-based implementation of CSC, specifically Convolutional Orthogonal Matching Pursuit (COMP), using a small dataset of about 20 audio snippets. COMP employs a greedy technique to iteratively find dictionary atoms and is relatively slow due to the orthogonalization it performs to the atoms in each step. Most of the current leading CSC algorithms[31,32,33] use L_1 regularization as a relaxation of L_0 regularization in their sparse reconstruction objectives, making them amenable to efficient optimization methods such as the Alternating Direction Method of Multipliers (ADMM). In this paper, we extensively compare our technique with the efficient CBPDN algorithm implemented within the state-of-art SPORCO python library[34]. In this experiment, we used 200 randomly chosen audio snippets from the same dataset, utilizing 10 gammatone kernels. The snippet lengths varied from 0.5s to 3.5s to compare processing times as a function of snippet length. Fig. 7 shows the reconstruction accuracy comparison between our framework and CBPDN. CBPDN, which is based on L_1 optimization, achieved reconstructions at different sparsity levels by varying the regularization parameter λ, while our framework achieved different spike rates by varying the refractory period (see[28] for exact parameter values). Fig. 7 shows that our framework outperforms CBPDN in the sub-Nyquist spike rate regime (0.1-0.6 Nyquist rate), achieving higher average SNR values at consistently lower spike rates. As shown in Fig. 7(b), our framework generally yields better SNR across spike rate bins, except in the initial very low spike rate and low SNR region. This exception can be attributed to dataset-specific idiosyncrasies, as CBPDN produced significantly fewer reconstructions in this range compared to our framework, as confirmed by the bar plot in Fig.7(c). Overall, our framework produces a larger fraction of reconstructions in the high SNR region as compared to CBPDN as is evident from the bar plot in 7(a). Fig. 8 compares the runtimes of our framework and CBPDN on a 10-core Intel(R) Xeon(R) CPU E5-2650 v3 server. The average processing time, shown as a function of snippet length, demonstrates that our framework also outperforms CBPDN in terms of runtimes, particularly in the aforementioned low spike rate regime. Although the processing time is platform-dependent, our framework offers better asymptotic complexity. Leading CSC algorithms[31,32,33] transform long temporal signals to the Fourier domain, resulting in $O(N \log N)$ complexity due to FFT, whereas our method achieves $O(Nw^3)$. While the overall processing times (considering both encoding and decoding) of both frameworks are comparable, our framework encodes signals into spike times very efficiently using a convolve-and-threshold mechanism in minimal time. This is particularly useful in online settings, such as when a signal needs to be streamed. In such cases, our framework allows signals to be efficiently conveyed via only the spike times and indices (with thresholds or coefficients inferred). In contrast, CBPDN and other CSC algorithms, being inherently based on optimization, cannot achieve such efficient encoding and require both the coefficients and the timing of the atoms for signal representation. Overall, our framework shows promise as a superior alternative to CSC techniques in specific scenarios, particularly for low spike rate representation of natural audio signals with biological kernels, as demonstrated in our experiments.

VIII. Conclusion

The experimental results establish the efficiency and robustness of the proposed spike-based encoding framework, which clearly outperforms state-of-the-art CSC techniques in the low spike rate regime. Notably, this high-fidelity coding and reconstruction is achieved through a simplified abstraction of a complex biological sensory processing system, which typically involves numerous neurons across multiple layers with diverse goals—such as feature extraction, decision making, classification and more—rather than being limited to reconstruction. For context, the human auditory system's cochlear nerve contains about 50,000 spiral ganglion cells (analogous to 50,000 kernels). The fact that our framework, with a single layer of roughly 100 neurons using a simple convolve-and-threshold model, achieves such high-quality reconstruction

underscores the potential of fundamental biological signal processing principles. Our framework differs fundamentally from the Nyquist-Shannon theory, primarily in its mode of representation and coding. Instead of sampling the value of a function at uniform or non-uniform pre-specified points, our coding scheme identifies the non-uniform points where the function takes specific convolved values. This efficient coding scheme, combined with our proposed window-based fast processing of continuous-time signals, shows great promise for achieving significant compression in real-time signal communication.

REFERENCES

[1] F. Rieke, D. Warland, R. de Ruyter van Steveninck, and W. Bialek, *Spikes: Exploring the Neural Code*. Cambridge, MA, USA: MIT Press, 1999.

[2] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, no. 6583, pp. 607–609, 1996.

[3] S. B. Laughlin and T. J. Sejnowski, "Communication in neuronal networks," *Science*, vol. 301, no. 5641, pp. 1870–1874, 2003.

[4] M. London and M. Häusser, "Dendritic computation," *Annual Review of Neuroscience*, vol. 28, pp. 503–532, 2005.

[5] G. Buzsáki, *Rhythms of the Brain*. Oxford, UK: Oxford University Press, 2006.

[6] P. Földiák, "Forming sparse representations by local anti-hebbian learning," *Biological Cybernetics*, vol. 64, no. 2, pp. 165–170, 1990.

[7] D. Graham and D. Field, "Sparse coding in the neocortex," *Evolution of Nervous Systems*, vol. 3, pp. 181–187, 2007.

[8] M. S. Lewicki, "Efficient coding of natural sounds," *Nature Neuroscience*, vol. 5, pp. 356–363, 2002.

[9] E. C Smith and M. Lewicki, "Efficient auditory coding," *Nature*, vol. 439, pp. 978–82, 2006.

[10] M. E. Auge D., Hille J. and et al., "A survey of encoding techniques for signal processing in spiking neural networks," *Neural Processing Letters*, vol. 47, no. 2, pp. 321–345, 2023.

[11] Z. F. Mainen and T. J. Sejnowski, "Reliability of spike timing in neocortical neurons," *Science*, vol. 268, pp. 1503–1506, 1995.

[12] J. W. Pillow, J. Shlens, L. Paninski, A. Sher, A. M. Litke, E. J. Chichilnisky, and E. P. Simoncelli, "Spatio-temporal correlations and visual signaling in a complete neuronal population," *Nature*, vol. 454, pp. 995–999, 2008.

[13] F. Rieke, D. Warland, R. de Ruyter van Steveninck, and W. Bialek, *Spikes: Exploring the Neural Code*. Cambridge, MA: MIT Press, 1997.

[14] W. Brendel, R. Bourdoukan, P. Vertechi, C. K. Machens, and S. Deneve, "Learning to represent signals spike by spike," *PLoS Computational Biology*, vol. 13, no. 10, p. e1005799, 2017.

[15] I. Daubechies and R. DeVore, "Approximating a bandlimited function using very coarsely quantized data:

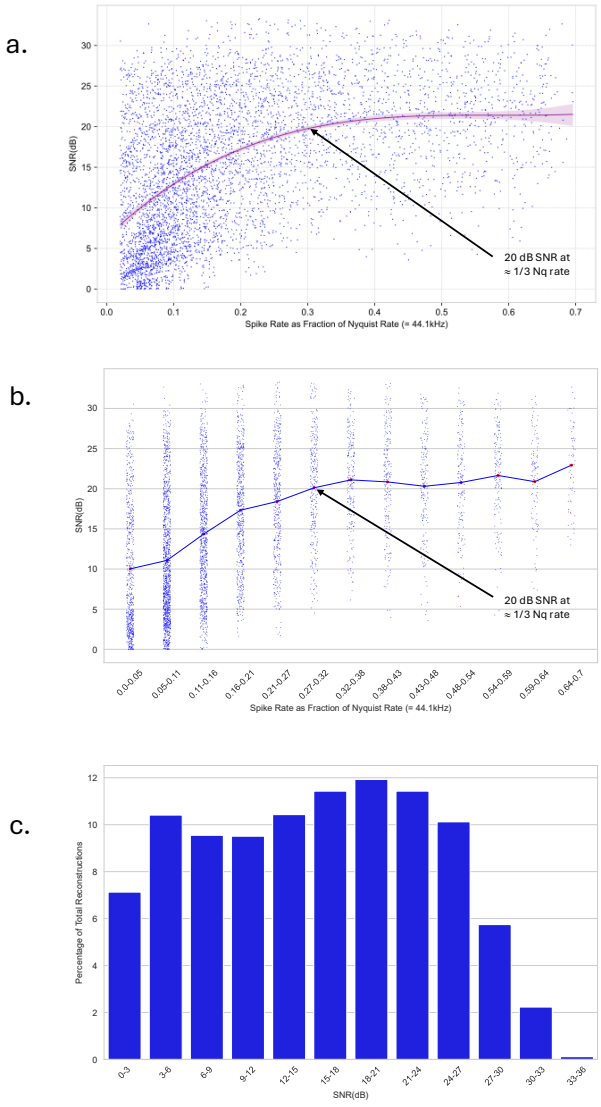

Fig. 6. Comprehensive results of experiments on 600 audio snippets with 50 kernels. **(a)** Scatter plot of reconstructions where each dot represents a single reconstruction performed on one of the 600 sound snippet for a particular setting of the *ahp* parameters as described in Section VII. The plot shows the SNR value of the reconstructions (y-axis) against corresponding spike-rate of the ensemble (x-axis). The trendline in the plot is generated using Seaborn's polynomial regression of order 3, which fits a third-degree polynomial to the data using least squares regression. he trendline passes through an average SNR of 20 dB at an average spike rate of approximately one-third of the Nyquist rate, as indicated by the arrow in the plot. **(b)** The strip plot illustrates the distribution of SNR values across different spike rate bins along the x-axis. The spike rate is binned from 0 to 0.7 of the Nyquist rate at nearly uniform intervals, while the y-axis represents the SNR values of the reconstructions. The blue line passing through the strip plot connects the average SNR values in each bin, with the averages indicated by red markers. As marked in the plot, the average SNR achieved at a spike rate of approximately one-third of the Nyquist rate is around 20 dB, which is consistent with plot (a). **(c)** Bar graph showing the distribution of all reconstructions from (a) at different SNR values. Reconstructions are binned at intervals of 3dB in the range 0 to 36 dB SNR (x-axis) and the bars show the percentage of all reconstructions that fall within the corresponding bin (y-axis). As is evident, the maximum fraction of reconstructions falls in the bin of 18-21 dB.

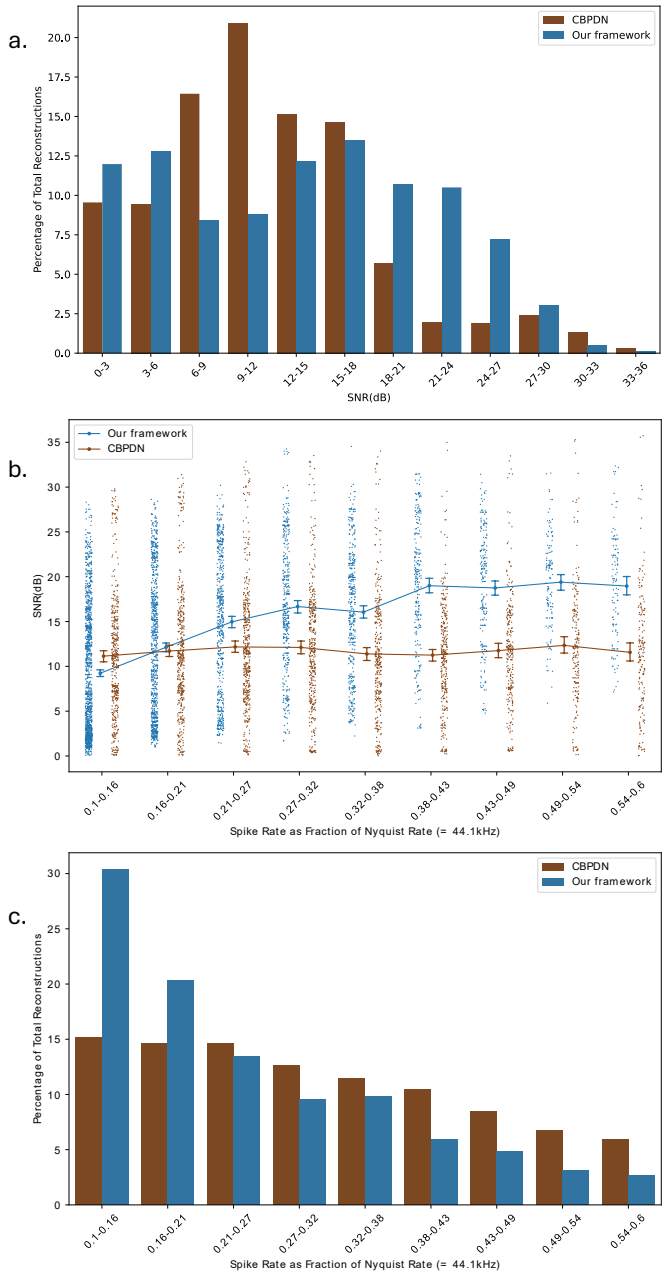

Fig. 7. Experimental comparison of our framework and cbpdn in terms of reconstruction accuracy and spike rate. **(a)** Bar plot comparing reconstruction accuracy. The x-axis represents SNR values binned at 3dB intervals (0-36 dB). Each bar indicates the fraction of reconstructions within the corresponding bin (brown for cbpdn and blue for our framework). The maximum fraction of reconstructions for cbpdn falls in the 9-12dB bin as opposed to the 15-18dB bin for our framework. Also our framework produced significantly more reconstructions in the high SNR region (right side of the graph). **(b)** Strip plot comparing SNR across spike rates. The x-axis bins spike rates as fractions of the Nyquist rate, with brown and blue dots representing individual reconstructions for CBPDN and our framework, respectively. The blue and brown lines indicate average SNR per bin. Within the demonstrated spike rate range (0.1-0.6 Nyquist rate) our framework consistently achieves higher reconstruction accuracy. While CBPDN briefly outperforms in the lowest spike rate bin, the limited available reconstructions for CBPDN (as seen in (c)) make this trend less reliable. **(c)** Bar plot comparing spike rates for the same reconstructions as in (a) and (b). The x-axis represents spike rates (0.1–0.6 Nyquist rate) binned at equal intervals. Most reconstructions from our framework occur at lower spike rates (left side of the graph, where blue bars are taller), while still maintaining better accuracy as demonstrated in (a) and (b).

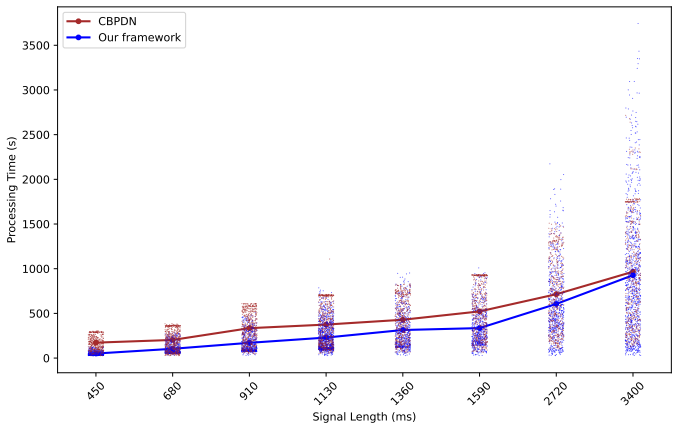

Fig. 8. Comparison of runtimes between our framework and cbpdn. The x-axis shows the different lengths of the audio snippets in msec and the strips of dots (brown dots for cbpdn and blue dots for our framework) represent the distribution of reconstruction runtimes corresponding to each length. The two lines connect the average processing times at each length for both our framework (blue) and cbpdn (brown). In terms of overall runtime our framework outperforms cbpdn. Although there are a few points produced by our framework in the high spike regime with higher processing times due to large matrix inversions for higher spike rates, this is greatly dependent on the hardware used.

A family of stable sigma-delta modulators of arbitrary order," *Annals of Mathematics*, vol. 158, no. 2, pp. 679–710, 2003.

[16] A. A. Lazar and L. T. Toth, "Time encoding and perfect recovery of bandlimited signals," in *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03).*, vol. 6, April 2003, pp. VI–709.

[17] I. Nemenman, G. D. Lewen, W. Bialek, and R. R. d. R. van Steveninck, "Neural coding of natural stimuli: information at sub-millisecond resolution," *PLoS computational biology*, vol. 4, no. 3, p. e1000025, 2008.

[18] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, 2002.

[19] A. A. Lazar and E. A. Pnevmatikakis, "Faithful representation of stimuli with a population of integrate-and-fire neurons," *Neural Computation*, vol. 17, no. 6, pp. 1345–1373, 2005.

[20] ——, "Multichannel time encoding with integrate-and-fire neurons," *Neurocomputing*, vol. 65, pp. 401–414, 2005.

[21] C. Garcia-Cardona and B. Wohlberg, "Convolutional dictionary learning: A comparative review and new algorithms," *IEEE Transactions on Comp. Imaging*, vol. 4, no. 3, pp. 366–381, 2018.

[22] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, 2002.

[23] A. Chattopadhyay and A. Banerjee, "Supplementary material," 2024, supplementary Material.

[24] R. Patterson, I. Nimmo-Smith, J. Holdsworth, and P. Rice, "An efficient auditory filterbank based on the gammatone function," 01 1988.

[25] J. N. Franklin, "On tikhonov's method for ill-posed

problems," *Mathematics of Computation*, vol. 28, no. 128, pp. 889–907, 1974.

[26] B. Schölkopf, R. Herbrich, and A. J. Smola, "A generalized representer theorem," in *International Conference on Computational Learning Theory*. Springer, 2001, pp. 416–426.

[27] E. Fonseca, M. Plakal, F. Font, D. P. W. Ellis, X. Favory, J. Pons, and X. Serra, "General-purpose tagging of freesound audio with audioset labels: Task description, dataset, and baseline," 2018.

[28] A. Chattopadhyay and A. Banerjee, "Locallyoptimal-spiking," 2024, gitHub Repository. [Online]. Available: https://github.com/crystalonix/LocallyOptimalSpiking.git

[29] G. Davis, S. Mallat, and M. Avellaneda, "Adaptive greedy approximations," *Constructive Approximation*, vol. 13, no. 1, pp. 57–98, 1997.

[30] A. Chattopadhyay and A. Banerjee, "Beyond rate coding: Signal coding and reconstruction using lean spike trains," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp. 1–5.

[31] B. Wohlberg, "Efficient convolutional sparse coding," in *2014 IEEE Int. Conf. on Acoustics, Speech and Signal Proc. (ICASSP)*, 2014, pp. 7173–7177.

[32] F. Heide, W. Heidrich, and G. Wetzstein, "Fast and flexible convolutional sparse coding," in *2015 IEEE Conf. on CVPR*, 2015, pp. 5135–5143.

[33] B. Wohlberg, "Boundary handling for convolutional sparse representations," in *2016 IEEE ICIP*, 2016, pp. 1833–1837.

[34] ——, "Sporco: A python package for standard and convolutional sparse representations," 01 2017, pp. 1–8.