

# A Coupled Network of Growth Transform Neurons for Spike-Encoded Auditory Feature Extraction

Ahana Gangopadhyay<sup>1</sup>, Kenji Aono<sup>2</sup>, Darshit Mehta<sup>3</sup>, and Shantanu Chakrabartty<sup>1</sup>

<sup>1</sup>Department of Electrical and Systems Engineering, Washington University in St. Louis, St. Louis, Missouri, USA

<sup>2</sup>Department of Computer Science and Engineering, Washington University in St. Louis, St. Louis, Missouri, USA

<sup>3</sup>Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, Missouri, USA

## Abstract

This paper builds upon our previously reported growth transform based optimization framework to present a novel spiking neuron model and demonstrate its application for spike-based auditory signal processing. Unlike conventional neuromorphic approaches, the proposed Growth Transform (GT) neuron model is tightly coupled to a system objective function, which results in network dynamics that are always stable and interpretable; and the process of spike generation and population dynamics is the result of minimizing an energy functional. We then extend the model to include axonal propagation delays in a manner that the optimized solution of the system or network objective function remains unaffected. The paper characterizes the model for different types of stimuli, and explores how changing different aspects of the cost function can reproduce known single neuron dynamics. We then investigate the properties of a coupled GT neural network that can generate spike-encoded auditory features corresponding to the output of a gammatone filterbank. We show that the discriminatory information is not only encoded in the traditional spike-rates and interspike-interval statistics, but is also encoded in the subthreshold response of GT neurons for inputs that are not strong enough to elicit spikes. We also demonstrate that while different forms of coupling between the neurons could produce compact and energy-efficient representation of the auditory features, the classification performance for a speaker recognition task remains invariant across different types of coupling. Thus, we believe that the proposed GT neuron model provides a flexible neuromorphic framework to systematically design large-scale spiking neural networks with stable and interpretable dynamics.

## I. INTRODUCTION

One of the main paradigms in the design of neuromorphic systems is to be able to achieve robust and efficient recognition performance by mimicking the architecture and the dynamics of biological neural networks. For example, neuromorphic implementations of the cochlea ([1]–[5]) closely mimic the neural dynamics observed in the auditory pathway where sound waves set up mechanical vibrations that are converted to electrical impulses by hair cells present along the cochlear basilar membrane in the inner ear ([6]). Specifically, continuous-time filterbanks have been used to emulate the tonotopic organization observed in the human auditory pathways, where groups of neurons in close proximity respond preferentially to a specific range in the frequency spectrum ([7]). This is illustrated in Figure 1A by an array of spiking neurons that processes the outputs of the filterbanks to produce neural impulses that encode the auditory features. In biological systems, the neural impulses are carried by auditory nerves through several auditory centers before reaching the higher processing stages in the cerebral cortex, undergoing several stages of complex transformations that are important for perception, identification, localization and segregation of sound. In synthetic auditory systems, as shown in Figure 1A, specific higher processing functions like speech or speaker recognition are emulated using standard machine learning modules like support vector machines (SVM) or convolutional neural networks (CNN). Irrespective of the type of the back-end classifier, the system recognition performance is ultimately determined by the discriminatory information present in the auditory features that are encoded by the spiking neuron models. In literature, neuronal models of varying degrees of complexity, ranging from the classic Hodgkin-Huxley model ([8]), FitzHugh-Nagumo model ([9]) and Izhikevich models ([10]) to the simple integrate-and-fire model ([11]) have been used. However, practical implementations of the cochlea usually resort to a simpler form of these models ([12]), where only the limit-cycle statistics like average rates or

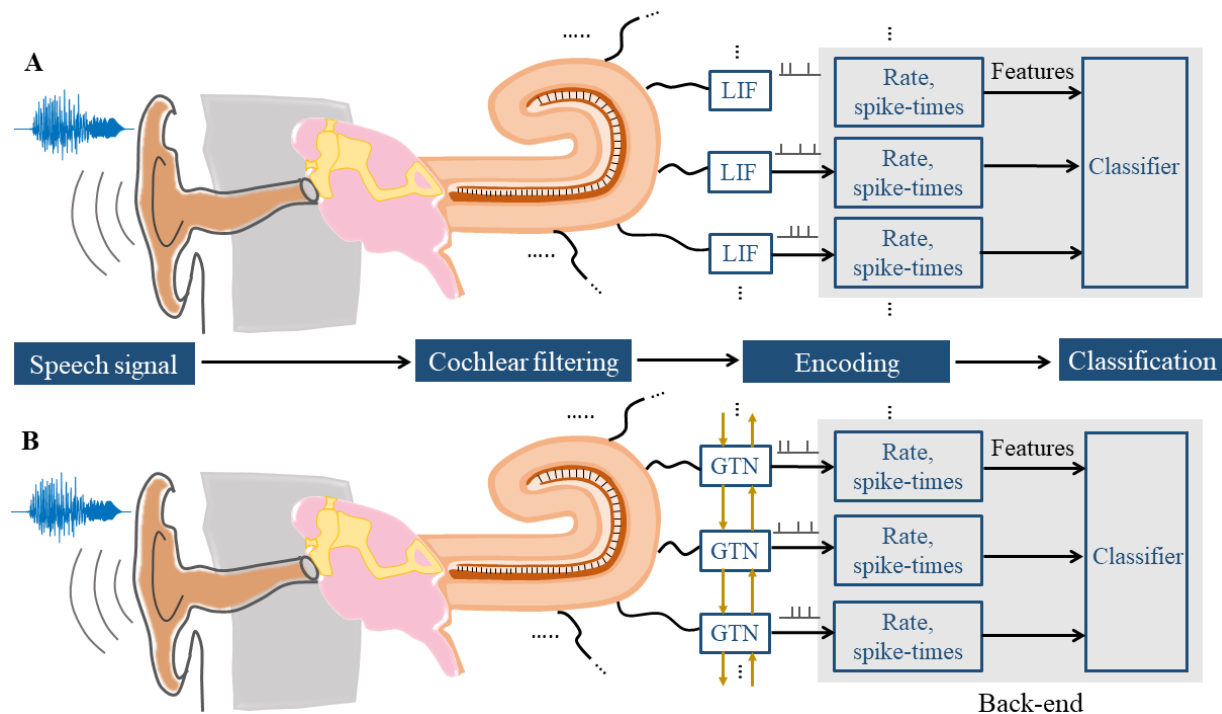


Figure 1. (A) Standard neuromorphic auditory feature extraction using leaky integrate-and-fire (LIF) neuron models followed by a DCT-based decorrelation and signal classification. (B) Proposed architecture using a coupled network of Growth Transform (GT) neurons to achieve a more energy-efficient auditory encoding.

inter-spike intervals of the spike-trains produced are then used for generating auditory features for recognition or classification ([13], [14]). At this level of abstraction, it is not evident how the shape, the nature and the dynamics of each individual spike is related to the overall system objective, and how a population of neurons when coupled together can self-optimize itself to produce an emergent spiking or population response ([15]). This is important for audition since it has been reported that each stage of the auditory pathway also exhibits limit cycles and subthreshold oscillations, both of which are believed to be important for precise encoding of acoustic information ([16], [17]). How these synaptic connections and the rich repertoire of dynamics encode speech signals for discrimination is a problem that remains unsolved, as is how to interpret and control such neural responses in a manner that can achieve a desired system objective.

In this paper, we approach the process of spike generation and encoding in a neuron model from a perspective that is fundamentally different compared to traditional dynamical systems formulations - by designing an energy functional for a network of neurons which produces an emergent spiking dynamics when minimized under realistic physical constraints. This is motivated by the fact that physical processes occurring in nature have an universal tendency to move towards a minimum-energy state - defined by the values taken by the process variables - that is a function of the inputs and the state of the process. Energy-based frameworks are the backbone of most machine learning models, which perform inference by minimizing an energy functional that capture dependencies between the variables ([18]). Energy consideration has either been largely overlooked in conventional representations of spiking neural networks, or been modeled as an external resource to constrain and control neural activity ([19]). This is despite the fact that neural signaling is metabolically expensive, with the neuron expending a significant fraction of its net energy utilization in the activation/deactivation of ion channels for spike generation and transmission ([20]). It therefore makes sense to study the process of spike generation and neural representation of excitation in the context of a network-level energy-minimization problem. ([21]) takes a similar approach by considering an energy function of spike-based network states, where individual neuron states are assumed to be binary in nature (spiking or not spiking). However, there is a lack of an unified model in literature that treats the spike generation process as the direct derivative of an energy functional of continuous-valued neural variables (e.g. membrane potential), that can moreover be used to mimic known neural dynamics from literature.

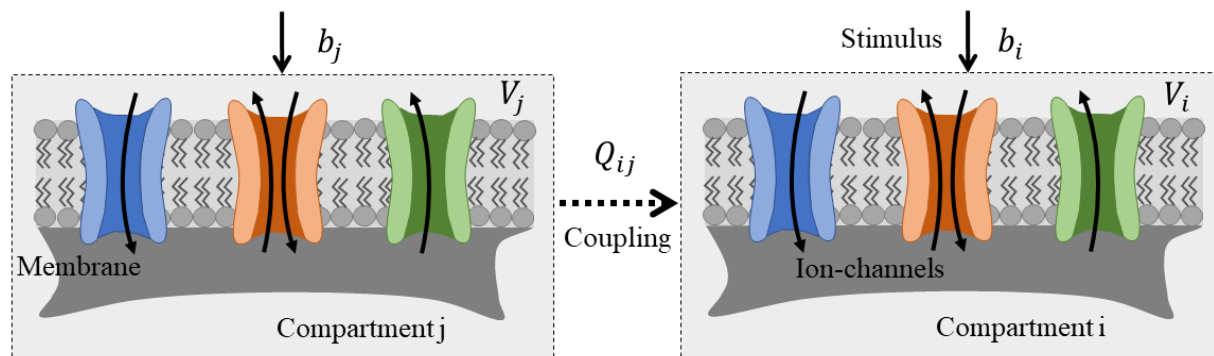


Figure 2. A subsection of the compartmental network is shown. Each neuronal compartment receives its own external stimulus  $b$ . The  $j$ -th compartment is coupled to the  $i$ -th by a coupling strength  $Q_{ij}$ .

In ([22]), we proposed a model of a growth transform (GT) neuron whose dynamical and spiking responses were shown to be strongly coupled with the network objective function. It was shown that while each individual neuron traverses a trajectory in the dual optimization space, the overall network traverses a trajectory in an equivalent primal optimization space. As a result, the network of growth transform neurons was shown to solve classification tasks, while producing stable, unique but interpretable neural dynamics like noise-shaping, spiking and bursting. In this paper, we first introduce a generic energy (cost) function for spiking neurons by applying a current balance equation at each neural compartment in the network, that stochastically maintains a charge-equilibrium in the cells. We then extend the model to incorporate axonal delays without affecting the network stability by exploiting the properties of growth transforms. We show that the GT neural network can exhibit known response characteristics similar to biological neurons by altering different aspects of the energy functional. Finally, as illustrated in Figure 1B, we explore a coupled network of GT neurons for generation and encoding of auditory features that can be used for discriminating speakers. We show that as the input stimulus varies over time, the spiking pattern of the growth transform neuron adapts, effectively tracking and transforming the input, while always being constrained by a system-level objective. Apart from an uncoupled network where the neurons track individual filterbank outputs, we show results for the speaker recognition task with coupled networks having sparse random excitatory and/or inhibitory connections between neurons. This formulation enables us to explore different patterns of connectivity between neurons, which may lead to a better or more compact stimulus representation. For example, inhibitory connections reduce the overall firing rate of the network, but produces similar classification performance as the uncoupled network.

Another novel aspect of this work compared to other neuromorphic approaches is the use of subthreshold response generated by GT neurons for auditory discrimination in addition to spike-based statistics. When the stimulus is not strong enough to elicit spiking from a neuron, its subthreshold activity still encodes information about the stimulus that cannot be obtained from its spiking activity only. Electrophysiological recordings across different sensory modalities and brain regions have shown that subthreshold response of a neuron is stimulus-specific and highly robust [23]–[25]. Rate encoding is a special case of this feature where the membrane potential trace is thresholded to select only the spikes. In this paper, we show that while information carried by the rate or timing of spikes can indeed encode useful information from speech signals, inclusion of subthreshold response results in a finer discrimination between speech signals from different speakers when compared to rate encoding.

## II. METHODS

In this section, we introduce an energy functional for the GT neural network and then derive the basic model of a GT neuron as a result of its minimization. We then extend this model to include axonal propagation delays by exploiting properties of growth transforms.

### A. Energy function for Growth Transform spiking neuron model

The theory underlying growth transforms and its connections with different forms of system objective functions, network dynamics (spiking, bursting and noise-shaping) have been discussed in detail in [22]. Here we summarize

the key concepts underlying the derivation of the GT neuron model based on the energy minimization framework. We will abstract the physics of neuronal spike generation and spike propagation using a multi-compartmental model as shown in Figure 2. Here each compartment could model a neuron which reacts to external stimuli and propagate signals to other compartments (neurons) by producing a rapid change in their transmembrane electrical potential difference (known as an action potential or a spike). Each compartment in the model comprises of ion channels, as shown in Fig. 2, whose roles are to equilibrate the charge across the membrane to its resting state. Also, in this simplified model we assume that the coupling between the compartments  $i$  and  $j$  is given by a transimpedance variable  $Q_{ij}$ .

Assuming that the charge across the membrane remains stable, each compartment maintains a current-balance where the sum of the time-average of the net transmembrane current due to ion-channel activity and the time-average of the time-varying input stimulus received by a compartment (neuron) is zero. Defining the time-averaging operator  $\mathcal{E}(\cdot)$  as an expectation of a signal  $X(t)$  computed over a time range  $[0, T]$  as

$$\mathcal{E}_T(X(t)) = \frac{1}{T} \int_0^T X(t) dt, \quad (1)$$

the current-balance equation for the  $i$ -th neuron can be expressed in terms of its instantaneous membrane potential  $V_i(t)$  as

$$\mathcal{E}_T \left( \sum_{j=1}^N Q_{ij}(V_j(t) - V_{rest}) + b_i(t) + C\Theta(V_i(t)) \right) = 0, \quad i = 1, \dots, N. \quad (2)$$

Here,  $N$  is the total number of neurons in the network,  $Q_{ij} \in \mathcal{R}$  denotes the contribution of the pre-synaptic potential at neuron  $j$  to the net excitation received by the post-synaptic neuron  $i$ ,  $b_i(t) \in \mathcal{R}$  is the instantaneous external current stimulus, and  $\Theta(\cdot)$  refers to the net transmembrane ionic current flow that maps the instantaneous membrane potential  $V_i(t)$  into a spike response.  $V_{rest}$  denotes the resting membrane potential. Note that we are abstracting the effect of multiple ion-channels (for example sodium, potassium or calcium) into a single ion-channel current function. An illustration of the compartmental model is given in Figure 2.  $C$  is a tunable parameter that, as we will show later, models the excitability of a compartment (neuron) to its input. We additionally consider the membrane potential to be always bounded as follows

$$|V_i(t)| \leq V_m \quad \forall t, \quad (3)$$

for some constant  $V_m > 0$ . The principle is similar to the neuron membrane potential being bounded by the ionic potentials on either side of the cell membrane that mediate the changes in membrane potential. For the sake of simplicity, we consider the membrane potential to be normalized, such that

$$|V_i(t)| \leq 1, \quad \forall t. \quad (4)$$

We can decompose the response  $V_i(t)$  of the  $i$ -th neuron and the bias term  $b_i(t)$  into two differential components as follows

$$V_i(t) = V_i^+(t) - V_i^-(t), \quad \text{and} \quad (5)$$

$$b_i(t) = b_i^+(t) - b_i^-(t), \quad i = 1, \dots, N, \quad (6)$$

where  $V_i^+(t)$  and  $V_i^-(t)$  satisfy the following constraints:

$$V_i^+(t) + V_i^-(t) = 1 \quad (7)$$

$$V_i^+(t), V_i^-(t) \geq 0. \quad (8)$$

Note that the Eqs 7 and 8 ensure that the decomposition of  $V_i(t)$  is unique, whereas the decomposition of  $b_i(t)$  is non-unique and does not affect the solution. Similarly, the resting membrane potential (also normalized w.r.t.  $V_m$ ) can be uniquely decomposed as  $V_{rest} = (V_{rest}^+ - V_{rest}^-)$ . Eq. 2 can thus be rewritten as

$$\mathcal{E}_T \left[ \sum_{j=1}^N Q_{ij} \left( (V_j^+(t) - V_{rest}^+) - (V_j^-(t) - V_{rest}^-) \right) + (b_i^+(t) - b_i^-(t)) + C\Theta(V_i^+(t) - V_i^-(t)) \right] = 0 \quad (9)$$

For simplicity, we choose a piece-wise linear form of  $\Theta(\cdot)$  such that it can be decomposed in the following way

$$\Theta(V_i^+(t) - V_i^-(t)) = \Psi(V_i^+(t)) - \Psi(V_i^-(t)) \quad (10)$$

for some function  $\Psi(\cdot)$ . Thus, solving Eq. 9 is equivalent to satisfying the following (stronger) set of conditions:

$$\begin{aligned} \mathcal{E}_T \left[ \sum_{j=1}^N Q_{ij}(V_j^+(t) - V_{rest}^+) + b_i^+(t) + \Psi(V_i^+(t)) \right] &= 0 \\ \mathcal{E}_T \left[ \sum_{j=1}^N Q_{ij}(V_j^-(t) - V_{rest}^-) + b_i^-(t) + \Psi(V_i^-(t)) \right] &= 0, \end{aligned} \quad (11)$$

where  $c$  is a constant. With a transformation of variables, Eq. 11 can be written as

$$\begin{aligned} \mathcal{E}_T \left[ \sum_{j=1}^N Q_{ij}(V_j^+(t) - y_j^+(t)) + \Psi(V_i^+(t)) \right] &= 0 \\ \mathcal{E}_T \left[ \sum_{j=1}^N Q_{ij}(V_j^-(t) - y_j^-(t)) + \Psi(V_i^-(t)) \right] &= 0 \end{aligned} \quad (12)$$

where

$$b_i^\gamma(t) = - \sum_{j=1}^N Q_{ij} u_j^\gamma(t), \quad \text{and} \quad (13)$$

$$y_j^\gamma(t) = V_{rest}^\gamma + u_j^\gamma(t), \quad \gamma = \pm. \quad (14)$$

We assume  $Q$  to be positive definite, which ensures that  $y_j^\gamma$  exists. Thus essentially, the external current stimulus is replaced as the weighted sum of equivalent contributions from pre-synaptic neurons. The equations in 12 can be viewed as a stochastic variant of the first-order conditions of an energy minimization problem subject to the constraints given by Eqs 7 and 8. A similar stochastic first-order framework was used in ([26]) to derive a dynamical system corresponding to  $\Sigma\Delta$  modulation. The equivalent energy minimization problem has the following form

$$\begin{aligned} \arg \min_{V_i^+, V_i^- \in \mathcal{H}} \mathcal{D}\{V_i^+, V_i^-\} &= \arg \min_{V_i^+, V_i^- \in \mathcal{H}} \sum_{i=1}^N \sum_{j=1}^N Q_{ij} (V_j^+ - y_j^+) (V_i^+ - y_i^+) + \sum_{i=1}^N \int \Psi(V_i^+) dV_i^+ \\ &+ \sum_{i=1}^N \sum_{j=1}^N Q_{ij} (V_j^- - y_j^-) (V_i^- - y_i^-) + \sum_{i=1}^N \int \Psi(V_i^-) dV_i^-, \end{aligned} \quad (15)$$

over a domain

$$\mathcal{H} = \{V_i^\gamma : V_i^\gamma \geq 0 \quad \text{and} \quad V_i^+ + V_i^- = 1, \quad \forall i = 1, \dots, N\}. \quad (16)$$

A similar system objective function was used in ([22]) to design a spiking support vector machine (SVM) for a binary classification task. Here, we explore how minimizing an energy function of this form can produce single neuron and population dynamics similar to that observed in biological neuronal networks.

The cost function  $\mathcal{D}$  in Eq. 15 comprises two components: (a) a quadratic term that tries to minimize the error between the variables  $V_i$  and  $y_i$ ; and (b) a regularization term  $\Psi(\cdot)$  that produces spiking. Given any network configuration (set of synaptic connections  $Q_{ij}$  and external inputs  $b_i$ ), the approach provides a way to couple the network-level neural activity to an optimization problem that strives to reach a minimum-energy state corresponding to the optimal solution when disturbed from the equilibrium.  $\Psi(\cdot)$  is designed to be non-differentiable, with a sharp transition near the desired threshold potential (tunable) that produces large excursions in the membrane potential (i.e., spikes) from the resting state when it reaches a threshold. This is because the optimization variables  $V_i^\gamma$  corresponding to the neurons near threshold potential cannot reach the optimal point due to the non-differentiable nature of  $\Psi(\cdot)$ , and will continue to oscillate around the threshold potential, generating spikes as long as the



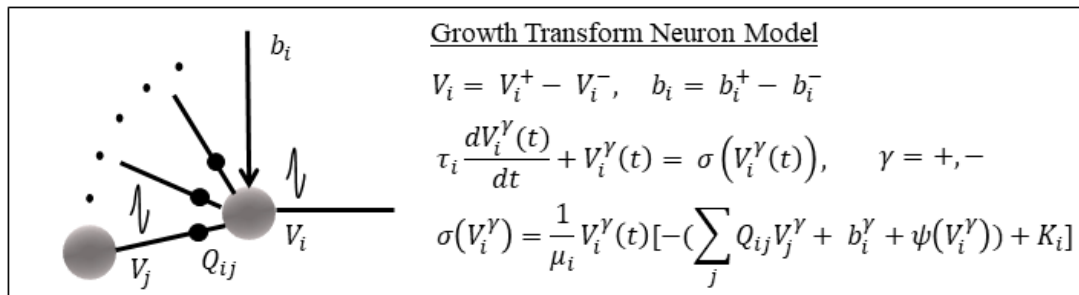


Figure 3. Summary of the Growth Transform Neuron model.

input stimulus is present.  $\Psi(\cdot)$  is designed such that these oscillations take the shape of an action potential in the membrane potential  $V_i$  when  $V_i$  reaches the threshold.

The generation and propagation of action potentials as a way of neural signaling expends energy. The energy efficiency of biological neural networks points to the spiking dynamics being coupled with and optimized with respect to a network-level cost function. Hence at the network level, neuronal activity is essentially guided by two principles: (a) to faithfully represent the external information received in the form of spike-trains, and (b) to do so in a way that minimizes some energy cost in the dynamical system. Specifically, from Eq. 11, we see that the ion-channel activity  $\Psi(\cdot)$  strives to continually minimize the error with which the neural responses track the time-varying input signal, so that Eq 2 is satisfied on an average. The cost function outlined in Eq. 15 also provides the framework to explore different connectivity patterns in the neural network - all within the context of optimizing a network-level cost - that has the potential to learn a better or more compact representation of the stimuli.

In ([22]), we proposed a neuron model that sequentially optimizes the cost function  $\mathcal{D}$  in Eq. 15 subject to constraints  $\mathcal{H}$  using a multiplicative update called growth transforms, a fixed-point algorithm for optimizing continuous objective functions over a manifold like  $\mathcal{H}$  with linear and non-negativity constraints ([27]–[29]). In the following subsection, we extend this discrete-time model to a continuous-time model that incorporates axonal propagation delays.

### B. Growth Transform Neuron Model with Axonal Delay

In biology, different propagation delays along different axons produce considerable variability in the network response. These delays are intrinsically generated in other neuron models like Hodgkin-Huxley, or are synthetically introduced using a time-based kernel in leaky-integrate and fire models ([30]). In most of these cases, it is not clear how coupling the neurons together, combined with the effect of mismatch in the axonal propagation delays, would affect the stability of the network or whether the dynamics could still be interpretable. Here we extend the GT neuron model described previously to incorporate different axonal propagation delays for individual neurons in a way that keeps the neural responses tied to the system objective, and produces a stable output. The approach exploits an interpolation property of growth transforms to implement a dynamic model of a neuron by updating the variables  $V_i^\gamma$ ,  $\gamma = \pm$  such that the polynomial cost function  $\mathcal{D}$  is optimized over the domain  $\mathcal{H}$ . The growth transform updates ([27]) for the minimization of the energy function  $\mathcal{D}(V_i^+, V_i^-)$  over the constrained manifold are given below

$$V_i^\gamma \leftarrow \alpha_i \sigma(V_i^\gamma) + (1 - \alpha_i) V_i^\gamma, \quad [V_i^\gamma \in \mathcal{H}, \quad 0 \leq \alpha_i \leq 1 \quad \forall i], \quad (17)$$

where

$$\sigma(V_i^\gamma) = \frac{1}{\mu_i} V_i^\gamma \left( -\frac{\partial \mathcal{D}(V_i^+, V_i^-)}{\partial V_i^\gamma} + K_i \right) \quad (18)$$

$$= \frac{1}{\mu_i} V_i^\gamma \left( -\sum_{j=1}^N Q_{ij} (V_j^\gamma - y_j^\gamma) - \Psi(V_i^\gamma) + K_i \right) \quad (19)$$

$$= \frac{1}{\mu_i} V_i^\gamma \left( -\sum_{j=1}^N Q_{ij} V_j^\gamma - b_i^\gamma - \Psi(V_i^\gamma) + K_i \right). \quad (20)$$

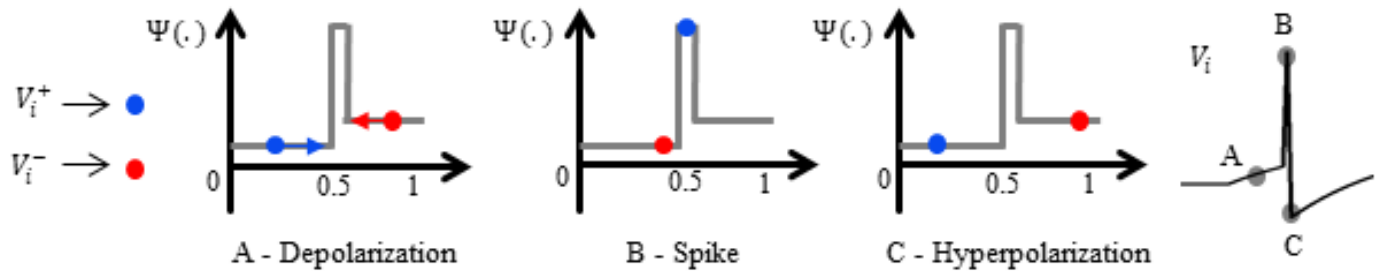


Figure 4. Ion-channel function  $\Psi(\cdot)$  used in this paper, with the paths traced by optimization variables  $V_i^\gamma$  in producing a spike.

The additive factor  $K_i \in \mathcal{R}$  ensures the positivity constraint in Eq. 8 holds, and the normalization factor

$$\mu_i = V_i^+ \left[ -\frac{\partial \mathcal{D}(V_i^+, V_i^-)}{\partial V_i^+} + K_i \right] + V_i^- \left[ -\frac{\partial \mathcal{D}(V_i^+, V_i^-)}{\partial V_i^-} + K_i \right] \quad (21)$$

enforces the conservation constraint in Eq 7. If  $\mathcal{D}(V_i^+, V_i^-)$  was continuously differentiable, growth transform updates would ensure that

$$\mathcal{D}(V_i^\gamma) \leq \mathcal{D}(\alpha_i \sigma(V_i^\gamma) + (1 - \alpha_i) V_i^\gamma), \quad (22)$$

with equality condition being satisfied only when  $V_i^\gamma$  was a critical point of  $\mathcal{D}$ . However since  $\mathcal{D}$  is not continuously differentiable over a sub-domain  $\mathcal{X} \in \mathcal{H}$ , some of the variables  $V_i^\gamma$  whose critical points are not directly reachable by growth transform updates will exhibit limit cycles about the sub-domain  $\mathcal{X}$ . The dynamic properties of the limit cycle will be determined by the shape of  $\Psi(\cdot)$ .

Eq. 17 can be written in the following discrete-time form

$$V_i^\gamma(t) = \alpha_i \sigma(V_i^\gamma(t - \Delta t)) + (1 - \alpha_i) V_i^\gamma(t - \Delta t), \quad [V_i^\gamma(t) \in \mathcal{H}, \quad 0 \leq \alpha_i \leq 1 \quad \forall i], \quad (23)$$

where  $V_i^\gamma(t)$  is the instantaneous value of the membrane potential at time  $t$ , and  $\Delta t$  is a small time-interval. Choosing  $\alpha_i = \frac{1}{\tau_i/\Delta t + 1}$ , where  $\tau = [\tau_i]$ ,  $i = 1, \dots, N$ , can be thought of as the vector of time-constants for the neuronal updates of the evolving dynamical system, we get

$$V_i^\gamma(t) = \frac{1}{\tau_i/\Delta t + 1} \sigma(V_i^\gamma(t - \Delta t)) + \frac{\tau_i/\Delta t}{\tau_i/\Delta t + 1} V_i^\gamma(t - \Delta t) \quad (24)$$

Rearranging the terms, we have

$$\tau_i \left[ \frac{(V_i^\gamma(t) - V_i^\gamma(t - \Delta t))}{\Delta t} \right] + V_i^\gamma(t) = \sigma(V_i^\gamma(t - \Delta t)) \quad (25)$$

In the limiting case, as  $\Delta t \rightarrow 0$ , this reduces to the following continuous-time dynamical system

$$\tau_i \frac{dV_i^\gamma(t)}{dt} + V_i^\gamma(t) = \sigma(V_i^\gamma(t)), \quad i = 1, \dots, N, \quad \gamma = \pm \quad (26)$$

where the growth transform update  $\sigma(V_i^\gamma(t))$  is given by Eq 20. Eq. 26 gives the equivalent growth transform neuron update for solving Eq. 2, and ensures that the network activity always converges irrespective of axonal propagation delays  $\tau_i \in [0, \infty]$ . The neural responses are finally given by the individual membrane potentials  $V_i = (V_i^+ - V_i^-)$ ,  $i = 1, \dots, N$ . It is to be noted that axonal delays are only applied when a spike is not generated, i.e., when  $V_i$  is at resting state potential or during depolarization. The update equations for the GT neuron model are summarized in Figure 3.

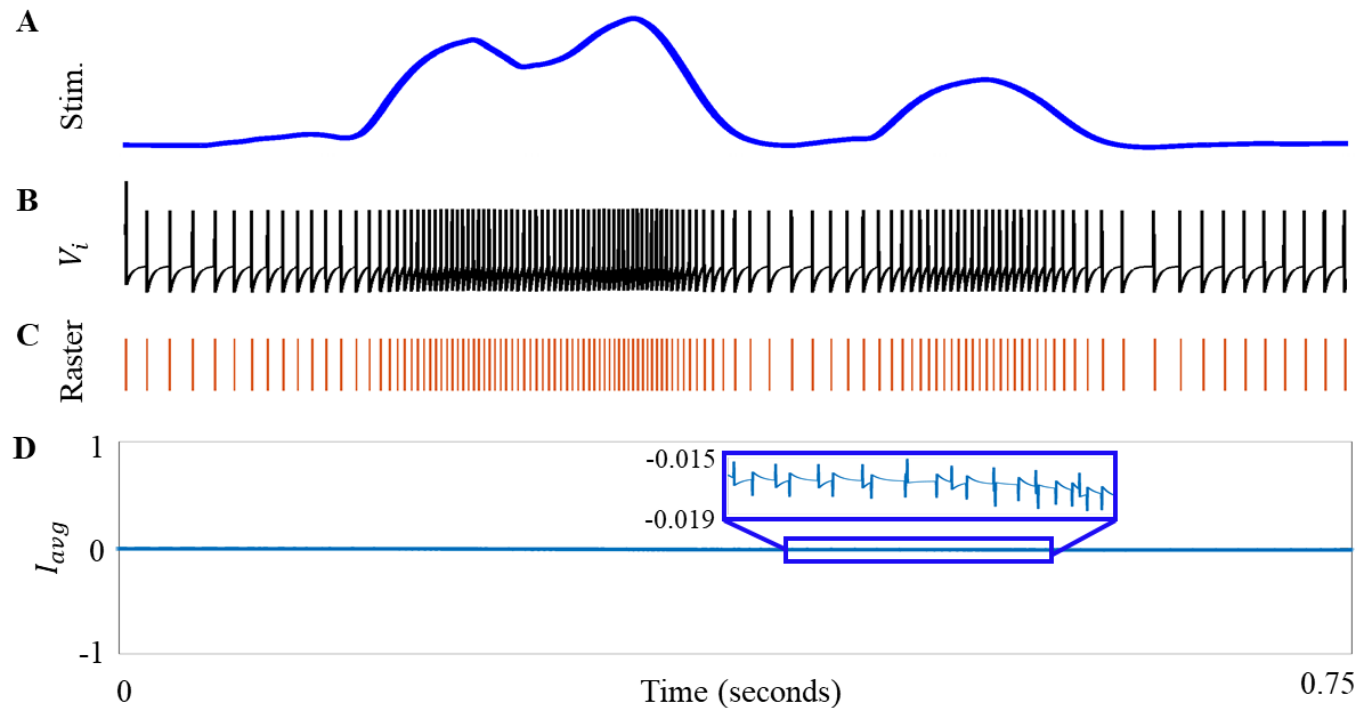


Figure 5. (A) Time-varying stimulus. (B) Membrane potential trace and spike raster. (C) Time-average of the net current at the neuron, given by Eq. 2.

### III. RESULTS

In this section, we first present a detailed characterization of a single GT neuron model, and show how it can replicate many of the neuronal dynamics observed in literature [31] by altering various aspects of the energy functional. We then demonstrate how the GT neuron encodes a time-varying stimulus in its spike pattern using different encoding techniques. Finally, we use a network of GT neurons to perform auditory feature extraction for a speaker recognition task. We provide extensive simulation results with an uncoupled network, and coupled networks with only excitatory, only inhibitory and with either types of connections between neurons, both for no axonal delays and with random axonal propagation delays.

#### A. Single GT Neuron Characterization

The ion-channel function  $\Psi(\cdot)$  used in this paper is shown in Figure 4. The gradient discontinuity in this case was set to  $V^+ = V^- = 0.5$ . In resting state,  $V_i^+ < (0.5 - \epsilon)$  and  $V_i^- > (0.5 + \epsilon)$ , where  $\epsilon$  is the width of the very narrow transition zone that sets the minimum input current needed for a spike as well as its height. When a depolarizing current pulse comes, both  $V_i^+$  and  $V_i^-$  try to move towards 0.5. If the input current is strong enough, the variables cross over the threshold, producing a spike. This adds a large penalty to the gradient term for  $V_i^+$ , and the updates force the variables to come back to a hyperpolarized state (a more negative  $V_i$ ) than where it started from. The resulting membrane potential trace  $V_i$  is also given in Figure 4.

We begin by characterizing the proposed model using a time-varying control input as shown in Figure 5A. The input is processed by a GT neuron, whose response  $V_i$  and the corresponding spike raster are given in Figure 5B. Figure 5C shows the expectation of the net instantaneous current at the neuron given by Eq. 2, taken over a time-window. Although the instantaneous values of the total current at the neuron can be fluctuating considerably depending on the stimulus inputs and ionic currents, growth-transform updates ensure the time-average is always close to zero.

We next investigate the effect of axonal propagation delays at a particular neuron. Figure 6A shows the modulation in the spiking frequency of a neuron due to increasing axonal delay (i.e., increasing  $\tau$ ) in response to the same input.



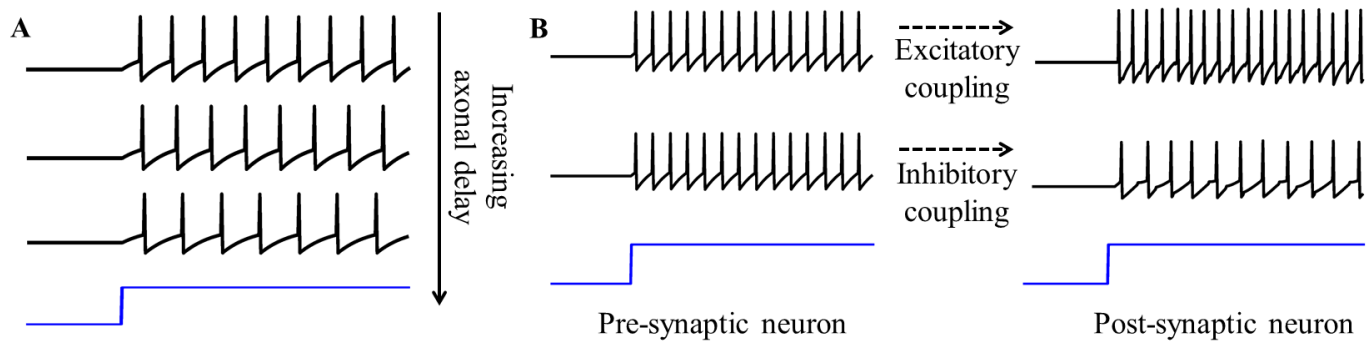


Figure 6. (A) Effect of axonal delay on spike propagation. (B) Effect of excitatory and inhibitory couplings on the post-synaptic neuron.

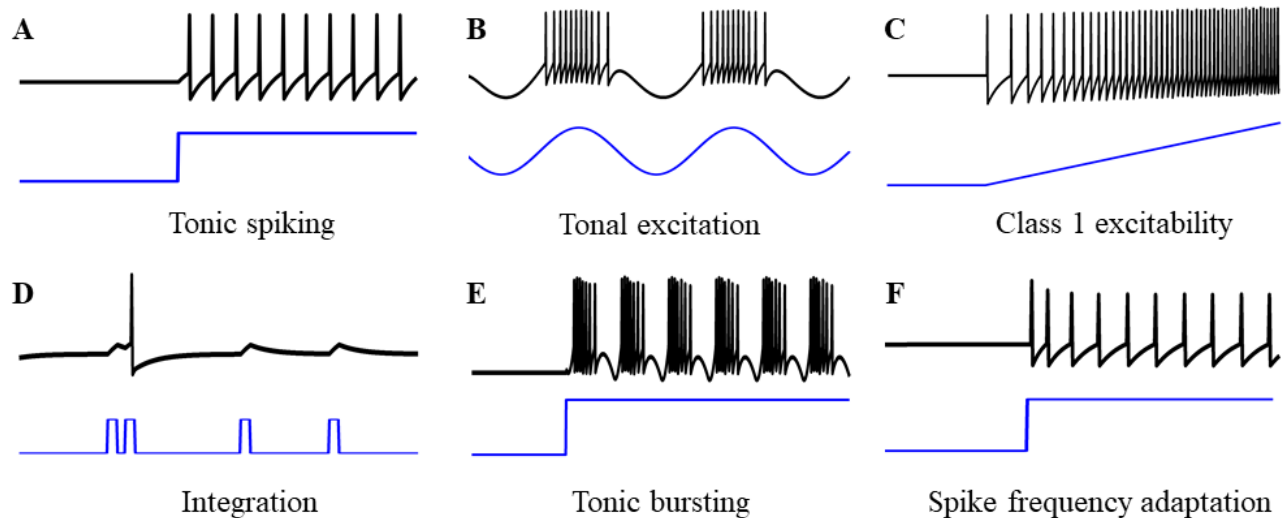


Figure 7. **Single-neuron responses.** Simulations showing different dynamical properties of the GT neuron.

The GT neural network provides a way to couple a network of neurons together that enables the pre-synaptic neuron to excite/inhibit the post-synaptic neuron, depending on the sign of the synaptic weight connecting them. This is illustrated in Figure 6B. An excitatory (positive) coupling increases the firing rate of the post-synaptic neuron, whereas an inhibitory (negative) coupling reduces its firing rate.

### B. Single GT neuron Dynamics

Single neurons show a vast repertoire of response characteristics and dynamical properties that lend richness to their computational properties at the network level. Izhikevich in ([31]) provided an extensive review of different spiking neuron models and their ability to produce the different dynamics observed in biology. We now show how we can reproduce a number of such dynamics by changing different aspects of the cost function in Eq. 15.

- **Tonic spiking.** When stimulated with a constant DC current pulse, a vast majority of neurons fire single, repetitive action potentials for the duration of the stimulus, with or without adaptation. Cortical neurons like regular spiking (RS) excitatory neurons, as well as low-threshold spiking (LTS) and fast spiking (FS) inhibitory interneurons show tonic spiking behavior ([32]–[34]). A simulation of such a firing pattern obtained using our model is shown in Figure 7A.
- **Tonal excitation.** Response of a GT neuron to a sinusoidally varying stimulus is shown in Figure 7B.
- **Class 1 excitability.** Hodgkin identified three classes of neurons on the basis of their patterns of repetitive spiking by studying the response of squid axons to current pulses of various amplitudes ([35], [36]). Class 1 neurons are able to generate action potentials with arbitrarily low frequencies in response to weak but superthreshold stimulus. The frequency of firing varies smoothly with the strength of injected current giving

rise to a continuous  $f - I$  curve. GT neuron inherently shows Class 1 excitability, which is demonstrated by its response to a ramp input in Figure 7C.

- **Integration.** Fig. 7D shows how a single GT neuron works as a leaky integrator, preferentially spiking to high-frequency or closely-spaced input pulses.
- **Bursting.** Bursting neurons fire discrete groups of spikes interspersed with periods of silence in response to a constant stimulus ([32], [33], [37], [38]). Bursting arises from an interplay of fast ionic currents responsible for spiking and slower intrinsic membrane currents that modulate spiking activity, causing the neuron to alternate between activity and quiescence. We model the intrinsic ionic mechanism of the faster process of action potential generation, and the slower process governing burst occurrence and duration by slowly modulating the shape of the underlying ion-channel function  $\Psi(\cdot)$  in a periodic manner in presence of excitation, so that the cell alternates between spiking and silence, thus encoding the input in bursts instead of individual spikes. Simulation of a bursting neuron in response to a constant step input obtained using the growth transform network is shown in Fig. 7E.
- **Spike frequency adaptation.** When presented with a prolonged stimulus of constant amplitude, many cortical cells including regular spiking (RS) neurons and low-threshold spiking (LTS) inhibitory neurons initially respond with a high-frequency spiking that decays to a lower steady-state frequency ([39]). This adaptation in the firing rate is caused by a negative feedback to the cell excitability due to gradual inactivation of depolarizing currents or activation of slow hyperpolarizing currents upon depolarization, and occur at a time-scale slower than the rate of action potential generation. This can be replicated in our model by tuning a single parameter,  $C$ . In the final cost function, this factor  $C$  that models the single-cell excitability, appears as the regularization hyperparameter, setting the relative weightage of the two terms, the net excitation received by a neuron and the net ionic activity in response. In order to model spike frequency adaptation, the excitability factor itself is made to undergo a first-order dynamics as follows

$$\tau_{adap} \frac{dC(t)}{dt} + C(t) = C_f \quad (27)$$

where  $\tau_{adap} \gg \tau$ . The corresponding response is given in Figure 7F.

Software (MATLAB) implementation of single and coupled GT neuron models are available through the web-portal [http://aimlab.seas.wustl.edu/document/gtn\\_gui\\_matlab.zip](http://aimlab.seas.wustl.edu/document/gtn_gui_matlab.zip). The tool enables users to visually see the effect of different hyper-parameters on different GT neuron dynamics and on different population dynamics produced by a network of coupled GT neurons.

### C. Spike-encoding using Growth Transform Neurons

Before we use a network of GT neurons to encode auditory features, we briefly discuss different spike-encoding patterns produced by a single GT neuron in response to the control stimulus given previously. For a time-varying input stimulus, GT neurons show a number of encoding properties similar to those observed in biological neural networks.

- **Rate encoding:** Mean firing rate over time is a popular rate coding scheme that claims that the spiking frequency or rate is modulated with the stimulus intensity ([40]). Rate-based features were obtained from this spike-train by computing the average spiking frequency of the neuron over a moving window.
- **Inter-spike interval-based encoding:** Temporal codes like inter-spike intervals claim that timings between successive spikes carry valuable information about the stimuli ([41]). We explored whether actual spike-timings produced by GT neurons can garner useful information about the input signal. For this, we computed the average inter-spike interval (ISI) for a neuron over a moving window over the entire spike-train.
- **Mean membrane potential:** For the next set of features, we explored whether inclusion of subthreshold activity improves stimulus-encoding. For this, a moving average was computed over the thresholded membrane potential  $V$  itself, that included both spikes and subthreshold activity. The threshold value was chosen to be slightly lower than the threshold membrane potential.

### D. Auditory Feature Extraction

The feature extraction algorithm proposed in this paper is summarized in Figure 8. Gammatone filterbanks ([42]) simulating cochlear filtering in the auditory system produce outputs that track the energy in different frequency bands

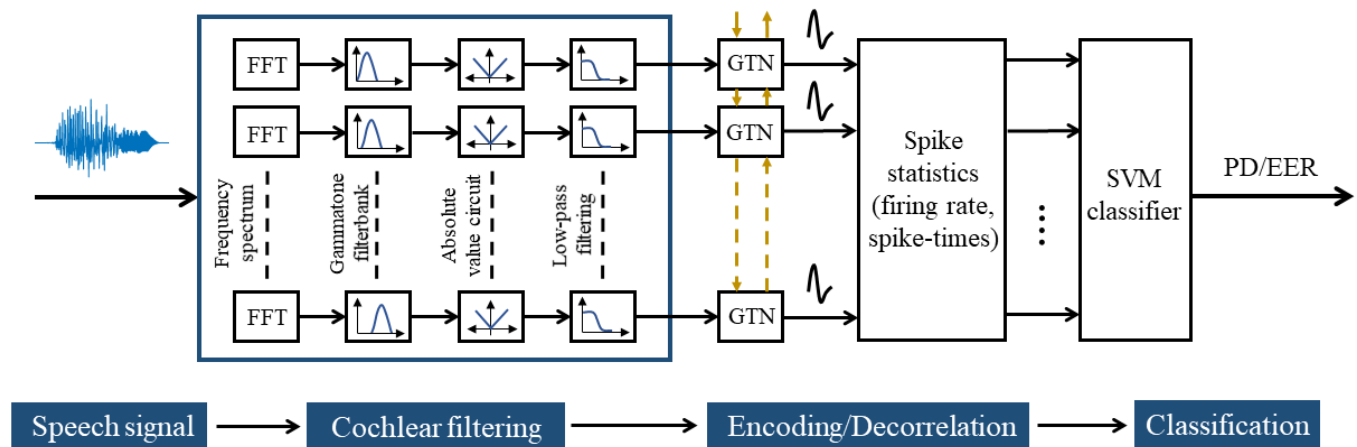


Figure 8. Signal-flow showing the feature extraction procedure for a speaker recognition task. The coupled growth transform neural network processes rectified and low-pass filtered gammatone filterbank outputs and produces spike trains from which different spike statistics are extracted to encode discriminatory features.

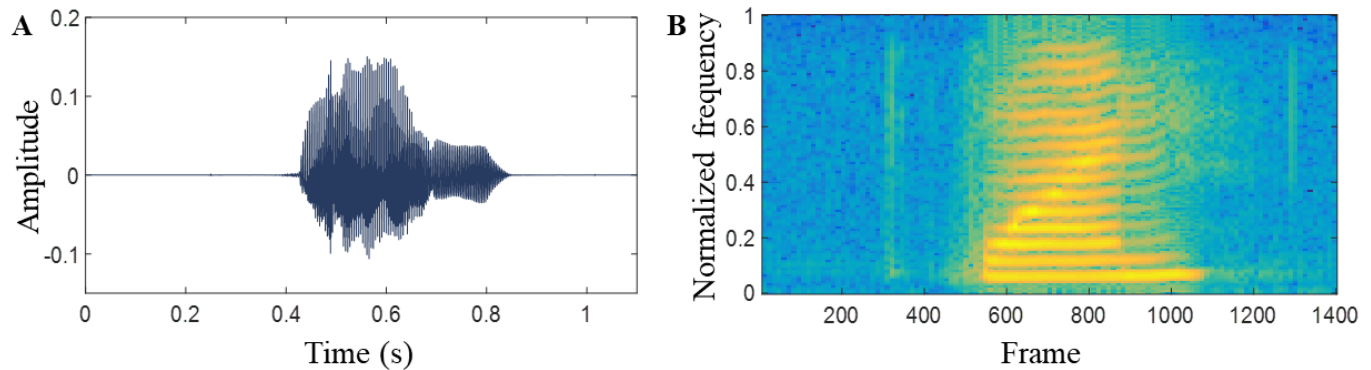


Figure 9. (A) A typical example of a speech signal corresponding to the utterance “one”. (B) The corresponding spectrogram with frame size = 150.

over the entire bandwidth. Each filterbank output  $x_i$  is rectified, low-pass filtered, scaled to  $[0, 1]$  and processed by a GT neuron in a differential manner as  $b_i$ ,  $i = 1, \dots, N$ ,  $N$  being the number of filterbanks used. The growth transform neurons track the energy in individual filterbank outputs, thus performing a basic decoding of speech signals in terms of frequency, intensity and duration. The proposed network outlines an approach where the early stages of acoustic information processing, which is analogous to the functionality of cochlear nuclei, can be connected to a network-level objective function, and the spiking dynamics produced is the direct derivative of a continuous optimization of the network cost as it processes a continually varying input signal in real-time.

In the next two subsections, we explore how spike-encoded features can be extracted from speech signals with uncoupled and coupled GT networks. For this, we use a sample speech signal corresponding to the utterance “one”. Figures 9A and B show the speech signal and its corresponding spectrogram.

1) *Uncoupled GT Neural Network Features*: We first consider an uncoupled network, where the synaptic weight matrix  $Q$  is diagonal. Each neuron therefore independently processes each filterbank output without contributions from the other neurons. For the speech signal in Figure 9A, the spike raster produced by a GT neural network with 20 neurons, each processing a single output channel from a filterbank of size 20, is presented in Figure 10A. Figures 10B, C and D show colormaps for the three feature sets.

2) *Coupled GT Neural Network Features*: Next, we explore how different network connectivity patterns affect the network representation of input stimuli. For this, we considered three coupled networks with excitatory connections, inhibitory connections, and with both excitatory and inhibitory connections. For the last network, neurons were randomly chosen to be either excitatory or inhibitory in nature. The corresponding spike rasters and feature

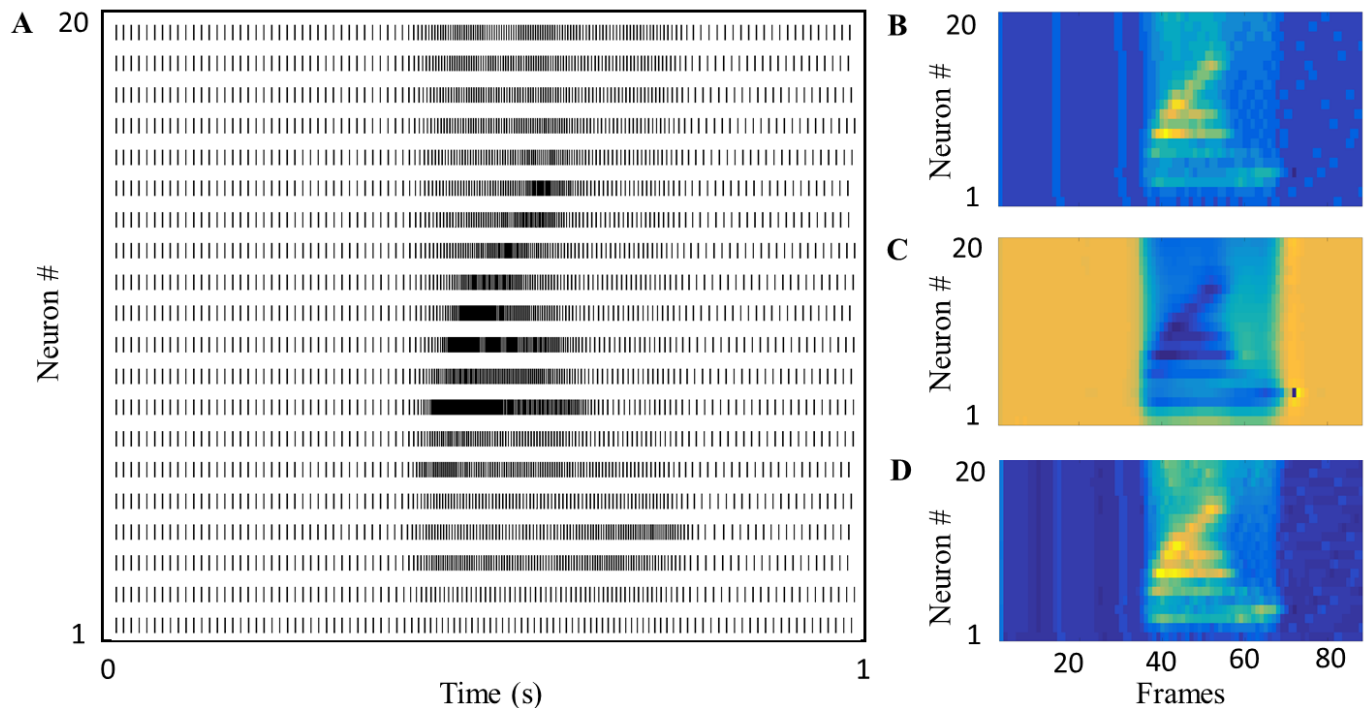


Figure 10. (A) Spike raster produced by the uncoupled growth transform neural network. (B), (C) and (D) Colormaps depicting feature sets with rate, spike-timing and mean membrane potential-based encoding of the speech signal in 9(A) respectively.

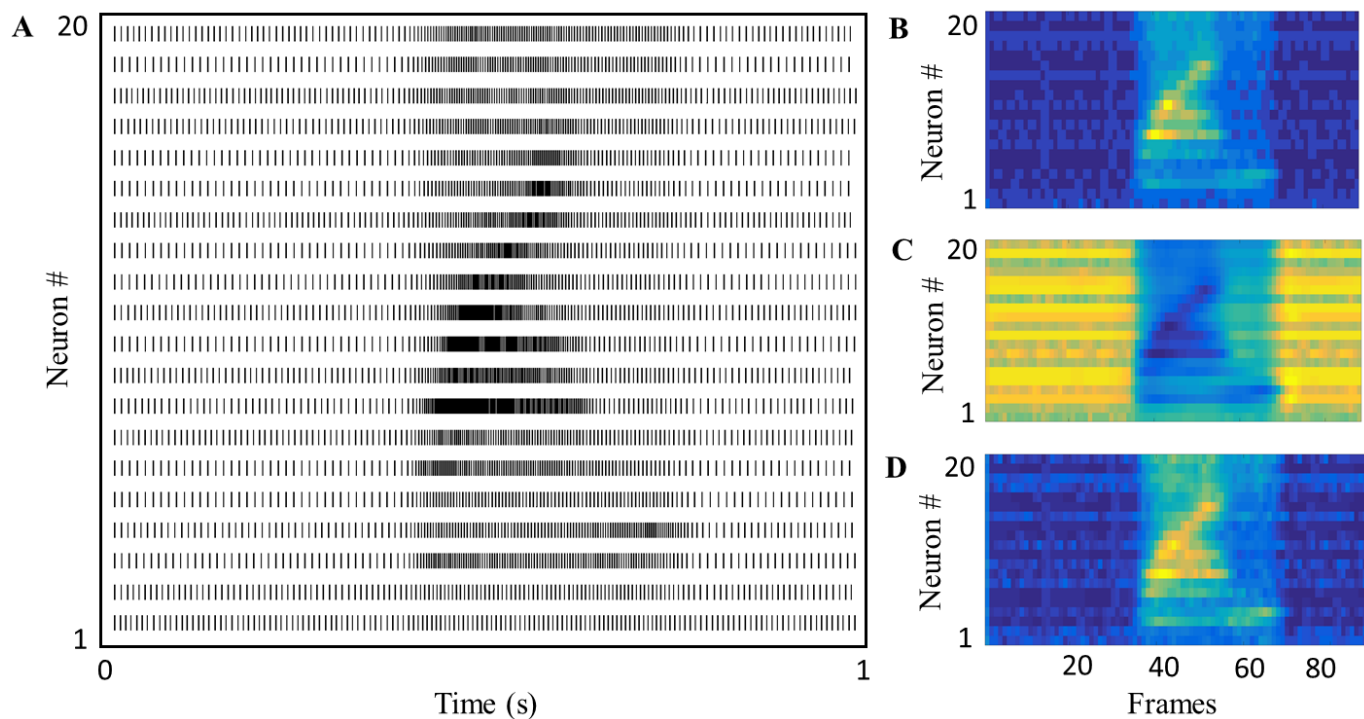


Figure 11. (A) Spike raster produced by the coupled growth transform neural network with sparse random excitatory connections. (B), (C) and (D) Colormaps depicting feature sets with rate, spike-timing and mean membrane potential-based encoding of the speech signal in 9(A) respectively.

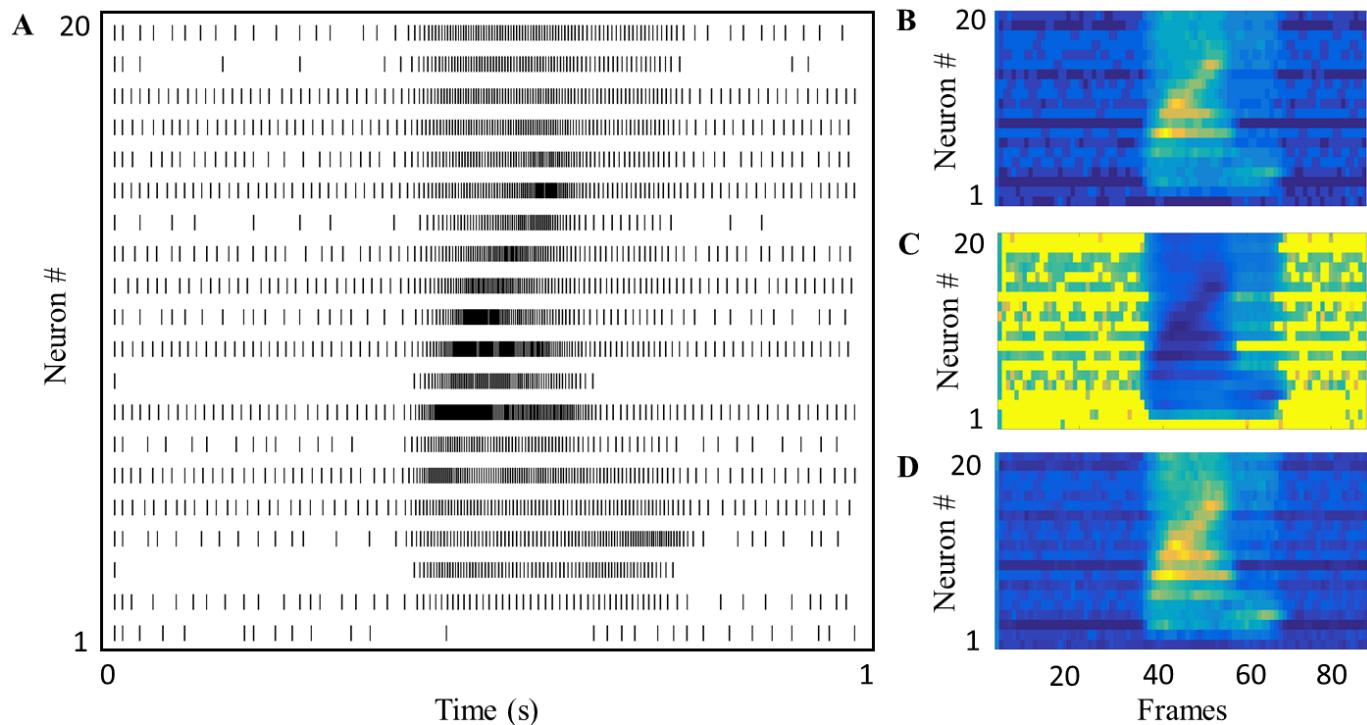


Figure 12. (A) Spike raster produced by the coupled growth transform neural network with sparse random inhibitory connections. (B), (C) and (D) Colormaps depicting feature sets with rate, spike-timing and mean membrane potential-based encoding of the speech signal in 9(A) respectively.

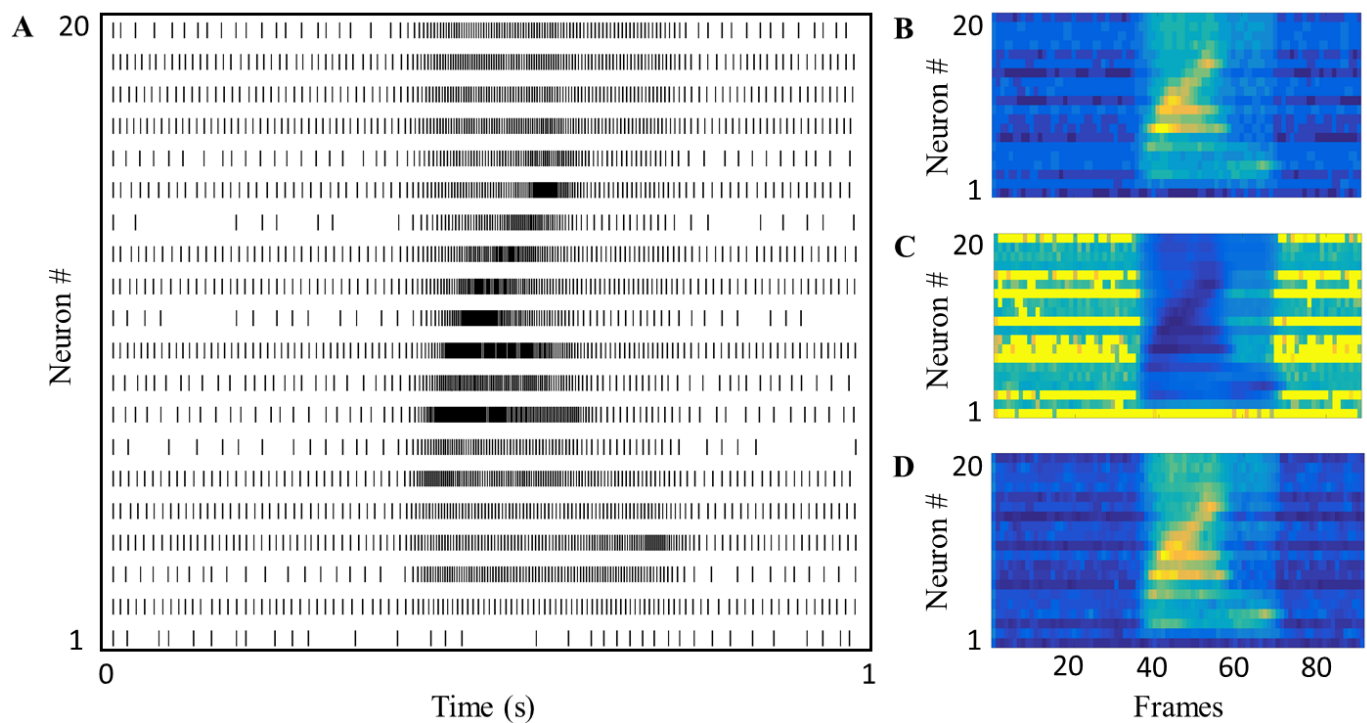


Figure 13. (A) Spike raster produced by the coupled growth transform neural network with sparse random excitatory as well as inhibitory connections. (B), (C) and (D) Colormaps depicting feature sets with rate, spike-timing and mean membrane potential-based encoding of the speech signal in 9(A) respectively.



colormaps are shown in Figures 11, 12 and 13 respectively. We see that different types of coupling affect the spike-patterns and encoding in markedly different ways. As expected, the average firing rate for a coupled network with inhibitory connections is much lower than the uncoupled network.

### E. Speaker Recognition Experiments

A speaker identification task utilizing the growth transform feature extraction algorithm is evaluated against the YOHO database, available through the Linguistic Data Consortium ([43]). This corpus is a collection of combination lock phrases (e.g. 26 – 81 – 56, pronounced twenty-six eighty-one fifty-six) for 138 subjects collected across a three month period. The data of interest are limited to the four enrollment sessions, with 24 utterances per session, collected in an office environment at an 8 kHz sampling rate with 3.8 kHz analog bandwidth. The first session is used for training, the second session for cross-validation, and the remaining third and fourth sessions were applied as a test set to evaluate the performance of the speaker identification system. There is a nominal time interval of three days between the sessions in YOHO. The data of each utterance in a session is initially normalized to the maximum response in the utterance before being passed through a second order gammatone filterbank with a quality factor of seven. Center frequencies of the filters in the filterbank are equally spaced according to the mel-scale up to 4 kHz, where the filterbank size is 20 for all the experiments reported in this paper. The envelopes of the rectified filterbank outputs are fed into the growth transform neural network as external stimuli. The spike-trains produced by the network are smoothed out using a moving window average and used as inputs to a *Gini* support vector machine (Gini-SVM) ([44]) which does the task of identification. The performance metrics reported in this work are the average across all considered speakers' probability of detection (PD) and the equal error rate (EER), which is the point at which the false rejection and false acceptance rates are equivalent. In general, a higher PD or lower EER would be indicative of better performance.

First, the feature set was evaluated using a baseline recognition system without any spiking neurons. The purpose of this evaluation was to determine the upper-limit of recognition that can be achieved by the gammatone filterbank auditory front-end. Then, each feature set was tested for four cases: (a) an uncoupled network, (b) a coupled network with sparse random inhibitory connections, (c) a coupled network with sparse random excitatory connections, and (d) a coupled network where neurons were randomly chosen to be either excitatory or inhibitory with random synaptic weights. For both uncoupled and coupled networks, DCT was applied as a post-processing step to decorrelate the features, and the DC component was dropped to get 19 usable filter outputs.  $\Delta$  features (i.e. the first derivative or velocity) were concatenated to the limit-cycle statistics extracted from the spike-trains. To reduce the effort required by the Gini-SVM, only a subsample of the feature vector (taken at uniform intervals) was fed into the system, for the YOHO utterances that last a couple of seconds, the resulting feature vector length can be several tens of thousands, while the subsampled version might be only a couple hundred.

Results for all the experiments over the entire database are given in Table I. We see that while the performance of the uncoupled network is similar to the baseline, introducing random couplings between neurons do not degrade the classification performance, especially when inter-spike intervals are used as features. This enables us to explore network connectivities that might lead to a more energy-efficient representation of external stimuli. For example, the coupled network with only inhibitory connections, is seen to reduce the overall firing rate of the network by an average of 5% per utterance, while still being able to extract relevant features for encoding of auditory data. Moreover, inclusion of subthreshold activity is seen to result in a finer discrimination in comparison to rate-encoding in almost all the cases.

In order to test how the GT neural network performs with random axonal propagation delays, we conducted the same set of experiments, but with different  $\tau_i$ ,  $i = 1, \dots, N$  chosen randomly for each neuron. The rate of sampling of the input signal was the same as the no-delay case. We see that even with random axonal delays, the classification accuracy using the spike-encoded features does not degrade, even though there is a mismatch between the rates at which different neurons encode the auditory information.

## IV. DISCUSSION

This paper attempts to answer the following fundamental question - Can the spiking dynamics in a synthetic neuron model be derived from a network-level energy-minimization problem? In other words, can we design an appropriate energy functional, which when optimized under realistic physical constraints, produces known neural



Table I  
SPEAKER RECOGNITION RESULTS ON THE YOHO DATABASE USING GT NEURAL NETWORK.

Description	Network	Features	EER (%)	PD (%)
Baseline	No-spikes	Filterbank outputs	2.91	97.69
GT Neural Network, No axonal delays	Uncoupled	Spike-rates	4.32	95.64
		Inter-spike interval	2.15	99.02
		Avg. mem. potential	2.81	98.35
	Excitatory connections	Spike-rates	5.08	93.58
		Inter-spike interval	3.66	96.65
		Avg. mem. potential	3.89	96.33
	Inhibitory connections	Spike-rates	4.67	94.49
		Inter-spike interval	3.56	96.97
		Avg. mem. potential	3.82	96.75
	Excitatory and inhibitory connections	Spike-rates	4.32	94.90
		Inter-spike interval	3.83	96.30
		Avg. mem. potential	2.22	97.5
GT Neural Network, With axonal delays	Uncoupled	Spike-rates	4.10	95.96
		Inter-spike interval	2.18	99.08
		Avg. mem. potential	3.02	98.04
	Excitatory connections	Spike-rates	6.09	91.50
		Inter-spike interval	3.84	96.63
		Avg. mem. potential	4.15	95.79
	Inhibitory connections	Spike-rates	4.98	94.08
		Inter-spike interval	3.71	96.63
		Avg. mem. potential	5.27	93.33
	Excitatory and inhibitory connections	Spike-rates	4.95	94.25
		Inter-spike interval	3.54	97.12
		Avg. mem. potential	3.96	95.00

dynamics? For this, we propose a new spiking neuron model based on growth transform updates where the process of spike generation and encoding is derived from a network cost function that strives to achieve a balance between the net excitation received by a neuron and the net ionic processes that determine its response, so that the time-average of accumulated charge in the cell membrane is zero. We derive a dynamical system formulation of the growth transform updates which takes into account axonal propagation delays in individual neurons. We analyze the model for different inputs and parameter variations, and demonstrate that it is well able to fulfill the current-balance equation from which we derive the neuronal updates.

Electrophysiological recordings show a rich repertoire of response dynamics at the cellular level. We showed how the GT neuron can reproduce a number of such dynamics by altering various parameters of the cost function, and tried to derive mechanistic insights about the significance of each. It was also demonstrated that the GT neuron can encode external stimuli using firing rates and inter-spike intervals, similar to biological neurons. Moreover, the sub-threshold activity of GT neurons was also seen to encode stimulus information in addition to spike-based statistics. These encoding techniques were applied for the task of encoding auditory features in a speaker recognition problem. The neural activity of GT neurons is always coupled to a network-level cost function, which enables us to explore different synaptic connectivities coupling the neurons, without sacrificing network stability. For example, introducing inhibitory couplings between neurons was seen to reduce the overall firing rate of the network, without degrading classification accuracy.

The paper illustrates how spike generation and spike-based encoding in an artificial neuron model could take place within a network-level optimization framework, thereby providing a tool to develop scalable neuromorphic machine learning algorithms. We have only considered fixed synaptic weights between neurons in this paper. An interesting direction of future research would be to learn these weights for an ‘optimal’ representation of the external stimuli, where the optimality condition could be based on spike counts, number of connections, etc. Finally, an end-to-end synthetic auditory system could be implemented within the same framework by replacing the classification module by a network of GT neurons capable of performing higher processing functions like speech or speaker recognition through spike-based encoding and learning.

## CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## AUTHOR CONTRIBUTIONS

AG and SC contributed to the conception and design of the study; KA, AG and DM conducted the simulations; DM designed the MATLAB interface, AG and SC wrote the first draft of the manuscript; AG, KA and SC wrote sections of the manuscript. All authors contributed to manuscript revision, read and approved the submitted version.

## FUNDING

This material is based upon work supported by the National Science Foundation under Grant Nos. CSR-1405273, DGE-0802267, DGE-1143954. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

## REFERENCES

- [1] R. F. Lyon and C. Mead, "An analog electronic cochlea," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 7, pp. 1119–1134, 1988.
- [2] R. Sarpeshkar, C. Salthouse, J.-J. Sit, M. W. Baker, S. M. Zhak, T.-T. Lu, L. Turicchia, and S. Balster, "An ultra-low-power programmable analog bionic ear processor," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 4, pp. 711–727, 2005.
- [3] V. Chan, S.-C. Liu, and A. van Schaik, "Aer ear: A matched silicon cochlea pair with address event representation interface," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 54, no. 1, pp. 48–59, 2007.
- [4] B. Wen and K. Boahen, "A silicon cochlea with active coupling," *IEEE transactions on biomedical circuits and systems*, vol. 3, no. 6, pp. 444–455, 2009.
- [5] S.-C. Liu, A. van Schaik, B. A. Minch, and T. Delbruck, "Asynchronous binaural spatial audition sensor with  $2 \times 64 \times 4$  channel output," *IEEE transactions on biomedical circuits and systems*, vol. 8, no. 4, pp. 453–464, 2014.
- [6] S.-C. Liu and A. van Schaik, "Neuromorphic sensors, cochlea," *Encyclopedia of Computational Neuroscience*, pp. 1982–1986, 2015.
- [7] C. Humphries, E. Liebenthal, and J. R. Binder, "Tonotopic organization of human auditory cortex," *Neuroimage*, vol. 50, no. 3, pp. 1202–1211, 2010.
- [8] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *The Journal of physiology*, vol. 117, no. 4, pp. 500–544, 1952.
- [9] R. FitzHugh, "Impulses and physiological states in theoretical models of nerve membrane," *Biophysical journal*, vol. 1, no. 6, pp. 445–466, 1961.
- [10] E. M. Izhikevich, "Simple model of spiking neurons," *IEEE Transactions on neural networks*, vol. 14, no. 6, pp. 1569–1572, 2003.
- [11] L. F. Abbott, "Lapicque's introduction of the integrate-and-fire model neuron (1907)," *Brain research bulletin*, vol. 50, no. 5, pp. 303–304, 1999.
- [12] J. Lazzaro, J. Wawrzyniec, M. Mahowald, M. Sivilotti, and D. Gillespie, "Silicon auditory processors as computer peripherals," in *Advances in Neural Information Processing Systems*, pp. 820–827, 1993.
- [13] S. Chakrabarty and S.-C. Liu, "Exploiting spike-based dynamics in a silicon cochlea for speaker identification," in *Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on*, pp. 513–516, IEEE, 2010.
- [14] C.-H. Li, T. Delbruck, and S.-C. Liu, "Real-time speaker identification using the aerear2 event-based silicon cochlea," in *Circuits and Systems (ISCAS), 2012 IEEE International Symposium on*, pp. 1159–1162, IEEE, 2012.
- [15] W. Gerstner and W. Kistler, *Spiking Neuron Models Cambridge University Press*. Cambridge, 2002.
- [16] P. O. Kanold and P. B. Manis, "Transient potassium currents regulate the discharge patterns of dorsal cochlear nucleus pyramidal cells," *Journal of Neuroscience*, vol. 19, no. 6, pp. 2195–2208, 1999.
- [17] P. B. Manis, S. C. Molitor, and H. Wu, "Subthreshold oscillations generated by ttx-sensitive sodium currents in dorsal cochlear nucleus pyramidal cells," *Experimental brain research*, vol. 153, no. 4, pp. 443–451, 2003.
- [18] Y. LeCun, S. Chopra, R. Hadsell, M. Ranzato, and F. Huang, "A tutorial on energy-based learning," *Predicting structured data*, vol. 1, no. 0, 2006.
- [19] J. Burrone, P. Taylor, C. Corey, T. Vachnadze, and H. T. Siegelmann, "Energetic constraints produce self-sustained oscillatory dynamics in neuronal networks," *Frontiers in neuroscience*, vol. 11, p. 80, 2017.
- [20] D. Attwell and S. B. Laughlin, "An energy budget for signaling in the grey matter of the brain," *Journal of Cerebral Blood Flow & Metabolism*, vol. 21, no. 10, pp. 1133–1145, 2001.
- [21] Z. Jonke, S. Habenschuss, and W. Maass, "Solving constraint satisfaction problems with networks of spiking neurons," *Frontiers in neuroscience*, vol. 10, p. 118, 2016.
- [22] A. Gangopadhyay and S. Chakrabarty, "Spiking, bursting, and population dynamics in a network of growth transform neurons," *IEEE Transactions on Neural Networks and Learning Systems*, 2017.
- [23] N. J. Priebe, "The relationship between subthreshold and suprathreshold ocular dominance in cat primary visual cortex," *Journal of Neuroscience*, vol. 28, no. 34, pp. 8553–8559, 2008.
- [24] K. E. Perks and T. Q. Gentner, "Subthreshold membrane responses underlying sparse spiking to natural vocal signals in auditory cortex," *European Journal of Neuroscience*, vol. 41, no. 5, pp. 725–733, 2015.

- [25] C. K. Machens, M. Wehr, and A. M. Zador, "Spectro-temporal receptive fields of subthreshold responses in auditory cortex," in *Advances in neural information processing systems*, pp. 149–156, 2003.
- [26] A. Gore and S. Chakrabartty, "A min–max optimization framework for designing  $\sigma\delta$  learners: Theory and hardware," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 3, pp. 604–617, 2010.
- [27] L. E. Baum and G. Sell, "Growth transformations for functions on manifolds," *Pacific Journal of Mathematics*, vol. 27, no. 2, pp. 211–227, 1968.
- [28] P. S. Gopalakrishnan, D. Kanevsky, A. Nádas, and D. Nahamoo, "An inequality for rational functions with applications to some statistical estimation problems," *IEEE Transactions on Information Theory*, vol. 37, no. 1, pp. 107–113, 1991.
- [29] A. Gangopadhyay, O. Chatterjee, and S. Chakrabartty, "Extended polynomial growth transforms for design and training of generalized support vector machines," *IEEE Transactions on Neural Networks and Learning Systems*, 2017.
- [30] T. P. Vogels and L. F. Abbott, "Signal propagation and logic gating in networks of integrate-and-fire neurons," *Journal of neuroscience*, vol. 25, no. 46, pp. 10786–10795, 2005.
- [31] E. M. Izhikevich, "Which model to use for cortical spiking neurons?," *IEEE transactions on neural networks*, vol. 15, no. 5, pp. 1063–1070, 2004.
- [32] A. Agmon and B. Connors, "Repetitive burst-firing neurons in the deep layers of mouse somatosensory cortex," *Neuroscience letters*, vol. 99, no. 1-2, pp. 137–141, 1989.
- [33] D. A. McCormick, B. W. Connors, J. W. Lighthall, and D. A. Prince, "Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex," *Journal of neurophysiology*, vol. 54, no. 4, pp. 782–806, 1985.
- [34] J. R. Gibson, M. Beierlein, and B. W. Connors, "Two networks of electrically coupled inhibitory neurons in neocortex," *Nature*, vol. 402, no. 6757, p. 75, 1999.
- [35] A. L. Hodgkin, "The local electric changes associated with repetitive action in a non-medullated axon," *The Journal of physiology*, vol. 107, no. 2, pp. 165–181, 1948.
- [36] S. A. Prescott, Y. De Koninck, and T. J. Sejnowski, "Biophysical basis for three distinct dynamical mechanisms of action potential initiation," *PLoS computational biology*, vol. 4, no. 10, p. e1000198, 2008.
- [37] C. M. Gray and D. A. McCormick, "Chattering cells: superficial pyramidal neurons contributing to the generation of synchronous oscillations in the visual cortex," *Science*, vol. 274, no. 5284, pp. 109–113, 1996.
- [38] J. C. Brumberg, L. G. Nowak, and D. A. McCormick, "Ionic mechanisms underlying repetitive high-frequency burst firing in supragranular cortical neurons," *Journal of Neuroscience*, vol. 20, no. 13, pp. 4829–4843, 2000.
- [39] B. W. Connors and M. J. Gutnick, "Intrinsic firing patterns of diverse neocortical neurons," *Trends in neurosciences*, vol. 13, no. 3, pp. 99–104, 1990.
- [40] E. D. Adrian, "The impulses produced by sensory nerve endings," *The Journal of physiology*, vol. 61, no. 1, pp. 49–72, 1926.
- [41] D. S. Reich, F. Mechler, K. P. Purpura, and J. D. Victor, "Interspike intervals, receptive fields, and information encoding in primary visual cortex," *Journal of Neuroscience*, vol. 20, no. 5, pp. 1964–1974, 2000.
- [42] R. F. Lyon, A. G. Katsiamis, and E. M. Drakakis, "History and future of auditory filter models," in *Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on*, pp. 3809–3812, IEEE, 2010.
- [43] J. Campbell and A. Higgins, "Yoho speaker verification ldc94s16," 1994.
- [44] S. Chakrabartty and G. Cauwenberghs, "Gini support vector machine: Quadratic entropy based robust multi-class probability regression," *Journal of Machine Learning Research*, vol. 8, no. Apr, pp. 813–839, 2007.
- [45] S. S. Stevens, J. Volkman, and E. B. Newman, "A scale for the measurement of the psychological magnitude pitch," *The Journal of the Acoustical Society of America*, vol. 8, no. 3, pp. 185–190, 1937.
- [46] M. Stopfer, S. Bhagavan, B. H. Smith, and G. Laurent, "Impaired odour discrimination on desynchronization of odour-encoding neural assemblies," *Nature*, vol. 390, no. 6655, pp. 70–74, 1997.
- [47] D. Desmaisons, J.-D. Vincent, and P.-M. Lledo, "Control of action potential timing by intrinsic subthreshold oscillations in olfactory bulb output neurons," *Journal of Neuroscience*, vol. 19, no. 24, pp. 10727–10737, 1999.
- [48] A. Fazel and S. Chakrabartty, "Sparse auditory reproducing kernel (spark) features for noise-robust speech recognition," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 4, pp. 1362–1371, 2012.
- [49] A. Fazel and S. Chakrabartty, "An overview of statistical pattern recognition techniques for speaker verification," *IEEE Circuits and Systems Magazine*, vol. 11, no. 2, pp. 62–81, 2011.
- [50] P. Dayan and L. F. Abbott, *Theoretical neuroscience*, vol. 806. Cambridge, MA: MIT Press, 2001.
- [51] M. A. Escabí, R. Nassiri, L. M. Miller, C. E. Schreiner, and H. L. Read, "The contribution of spike threshold to acoustic feature selectivity, spike information content, and information throughput," *Journal of Neuroscience*, vol. 25, no. 41, pp. 9524–9534, 2005.
- [52] R. Azouz and C. M. Gray, "Adaptive coincidence detection and dynamic gain control in visual cortical neurons in vivo," *Neuron*, vol. 37, no. 3, pp. 513–523, 2003.
- [53] M. H. Higgs and W. J. Spain, "Kv1 channels control spike threshold dynamics and spike timing in cortical pyramidal neurones," *The Journal of physiology*, vol. 589, no. 21, pp. 5125–5142, 2011.
- [54] M. Hausser, "The hodgkin-huxley theory of the action potential," *Nature Neuroscience*, vol. 3, pp. 1165–1165, 2000.
- [55] E. M. Izhikevich, *Dynamical systems in neuroscience*. MIT press, 2007.
- [56] H. R. Wilson and J. D. Cowan, "Excitatory and inhibitory interactions in localized populations of model neurons," *Biophysical journal*, vol. 12, no. 1, pp. 1–24, 1972.
- [57] B. H. Jansen and V. G. Rit, "Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns," *Biological cybernetics*, vol. 73, no. 4, pp. 357–366, 1995.
- [58] R. Krahe and F. Gabbiani, "Burst firing in sensory systems," *Nature Reviews Neuroscience*, vol. 5, no. 1, p. 13, 2004.