Neural System Identification With Spike-Triggered Non-Negative Matrix Factorization

Shanshan Jia, Zhaofei Yu, *Member, IEEE*, Arno Onken, Yonghong Tian, *Senior Member, IEEE*, Tiejun Huang, *Senior Member, IEEE*, and Jian K. Liu

***Abstract*—Neuronal circuits formed in the brain are complex with intricate connection patterns. Such complexity is also observed in the retina with a relatively simple neuronal circuit. A retinal ganglion cell (GC) receives excitatory inputs from neurons in previous layers as driving forces to fire spikes. Analytical methods are required to decipher these components in a systematic manner. Recently a method called spike-triggered non-negative matrix factorization (STNMF) has been proposed for this purpose. In this study, we extend the scope of the STNMF method. By using retinal GCs as a model system, we show that STNMF can detect various computational properties of upstream bipolar cells (BCs), including spatial receptive field, temporal filter, and transfer nonlinearity. In addition, we recover synaptic connection strengths from the weight matrix of STNMF. Furthermore, we show that STNMF can separate spikes of a GC into a few subsets of spikes, where each subset is contributed by one presynaptic BC. Taken together, these results corroborate that STNMF is a useful method for deciphering the structure of neuronal circuits.**

***Index Terms*—Neural network, neural spike, nonlinearity, non-negative matrix factorization, receptive field, system identification.**

I. Introduction

NEURONAL circuits in the brain are highly complex. Even for the retina, a relatively simple neuronal circuit, the underlying structure, in particular, functional characteristics, are still not completely understood. However, the retina serves as a typical model for both deciphering the structure of neuronal circuits [1]–[6] and testing novel methods for neuronal coding [7]–[11]. The retina consists of three layers with photoreceptors, bipolar cells (BCs), and ganglion cells (GCs), together with inhibitory horizontal and amacrine cells in between, as illustrated in Fig. 1(a). GCs, as the only output neurons of the retina, send visual information via the optic tracts and thalamus to cortical areas for higher cognition. Each GC receives inputs from a number of excitatory BCs as a driving force to generate spikes [Fig. 1(b)].

Due to the clear input–output relation, the retina is well suited for studying encoding/decoding of the stimulus (visual optical image) with neuronal responses (spikes in retinal GCs). For the purpose of system identification, characterizing its neuronal circuit is not trivial. However, most methods for deciphering neuronal circuits rely on certain experimental techniques. Traditionally, one can detect the connection between neurons with single or multiple electrodes [12], [13]. With the advance of experimental techniques, large-scale multielectrode array and calcium imaging can simultaneously record hundreds or thousands of cells [14], [15]. Therefore, a systematic method for analyzing these cells is highly desirable.

A recent work proposed such a method, called spike-triggered non-negative matrix factorization (STNMF), to analyze the underlying structural components of the retina [16]. Non-negative matrix factorization (NMF) has been proposed to capture the local structure of a given dataset [17]. It is widely used in computer vision [18], [19]; signal processing [20]–[22]; machine learning [23]–[25]; gene expression [26]; and neuroscience [27]–[31]. The ability to learn local parts from the whole dataset was further improved by sparseness constraints [32], [33]. Such a sparse coding is naturally related to the receptive field structure of sensory neurons, which is typically found in the visual system [34], [35].

In the recent study [16], by analyzing the spikes recorded from the retinal GCs, STNMF was shown to identify physical locations of subunit bipolar neurons of the previous layer that are pooling to a target retinal GC. Here, we significantly extend this approach to demonstrate how STNMF can be used for characterizing various functional properties of the retinal circuit. It is difficult to demonstrate the power of STNMF for a biological neuronal circuit, even the retina, as there are many unknowns due to the limitations of current experimental techniques for measuring a complete map of the retina. Therefore, in this study, we first use a clearly defined minimal network model as proof of principle to explain STNMF, and then demonstrate it with the retinal GC data.

Manuscript received August 9, 2018; revised February 16, 2020; accepted November 30, 2020. Date of publication January 5, 2021; date of current version June 16, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 61806011, Grant 61961130392, Grant 61825101, and Grant U1611461; in part by the National Postdoctoral Program for Innovative Talents under Grant BX20180005; in part by the China Postdoctoral Science Foundation under Grant 2018M630036; in part by the Zhejiang Lab under Grant 2019KC0AB03 and Grant 2019KC0AD02; and in part by the Royal Society Newton Advanced Fellowship under Grant NAF-R1-191082. This article was recommended by Associate Editor D. Tao. *(Corresponding authors: Zhaofei Yu; Jian K. Liu.)*

Shanshan Jia, Zhaofei Yu, Yonghong Tian, and Tiejun Huang are with the National Engineering Laboratory for Video Technology, School of Computer Science and Technology, Peking University, Beijing 100871, China (e-mail: jssll00@126.com; yuzf12@pku.edu.cn; yhtian@pku.edu.cn; tjhuang@pku.edu.cn).

Arno Onken is with the Institute for Adaptive and Neural Computation, School of Informatics, University of Edinburgh, Edinburgh EH8 9YL, U.K. (e-mail: aonken@inf.ed.ac.uk).

Jian K. Liu is with the Centre for Systems Neuroscience, Department of Neuroscience, Psychology and Behaviour, University of Leicester, Leicester LE1 7HA, U.K. (e-mail: jian.liu@leicester.ac.uk).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCYB.2020.3042513.

Digital Object Identifier 10.1109/TCYB.2020.3042513

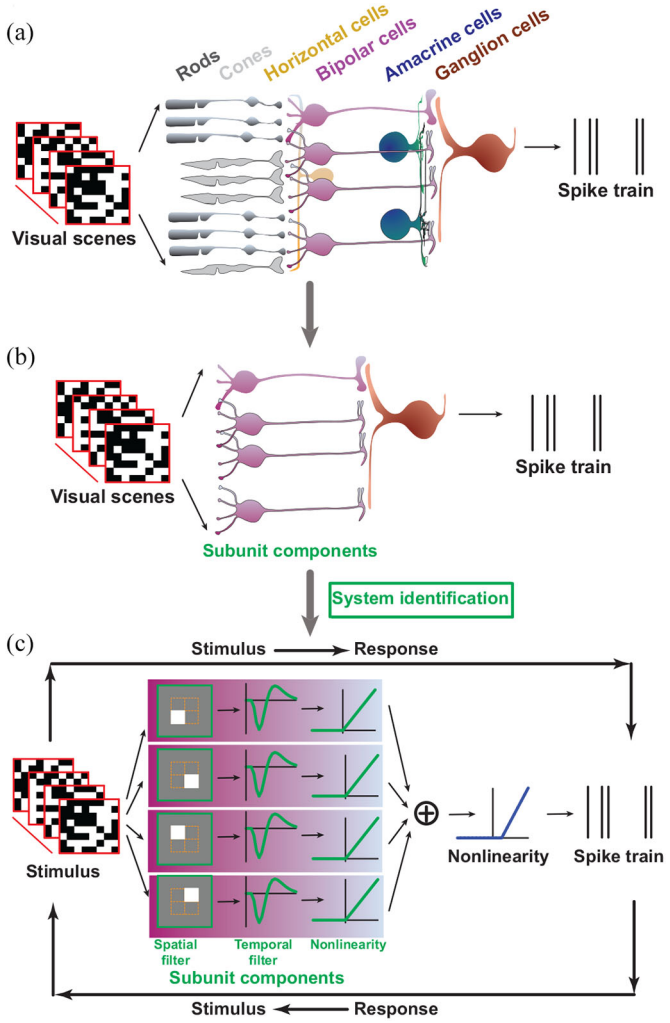

Fig. 1. Illustration of retinal neuronal network and GC model. (a) Illustration of retinal network. Light signals come into the retina from photoreceptors (rods and cones), transfer to BCs, and then send the output as spikes of the GC. In between, there are horizontal cells and amacrine cells as inhibitory modulations. (b) Minimal neural network of one GC consisted of a few BCs as subunit components. System identification is to uncover properties of these subunit components. (c) Model illustration. Stimuli are a sequence of random black–white images in 8×8 pixels. The network has four subunits, where each one has a spatial filter covering a 2×2 region indicated by the white color (1 in white and 0 everywhere else), temporal OFF filter and threshold-linear nonlinearity. Summation of the output of each subunit is pooling to a readout unit, passes the final nonlinearity, and generates the activity as a spiking probability to determine actual spikes according to the Poisson process.

The remainder of this article is structured as follows. First, we explain the workflow of STNMF as a general framework for system identification of the neural network using a modeled retinal GC in Section II. Section II shows a complete picture of neural network components, including synaptic subunit structures, synaptic connections, and their weights between presynaptic BCs and postsynaptic GC. Then, we demonstrate a novel feature of STNMF for classification of all the spikes of a GC into a few subsets of spikes, where each subset of spikes is mainly contributed by one presynaptic BC. When applying STNMF to biological data of retinal GCs, similar results are found as for the artificial data with known ground truth. For each retinal GC, STNMF finds a set of presynaptic BCs together with their contributed spikes. This article concludes with a summary and discussion in Section IV.

II. Methods

For a biological neuronal circuit, even the retina as illustrated in Fig. 1(a), there are many unknowns. Thus, we use a clearly defined minimal network model of the retinal GC to explain the framework STNMF as a method of system identification for the neural network.

A. Ganglion Cell Models

A simulated GC in Fig. 1 is modeled by a typical linear–nonlinear model that has been shown to capture biological retinal neuronal responses [9], [16], [36]. The model cell has four excitatory subunits with a size of 2×2 pixels each. The setup of the model is equivalent to a neural network with two layers: 1) an input layer with four subunits and 2) an output layer with one single readout unit. Inhibitory neurons are not modeled here, since they are barely triggered to show the effect on the receptive field of the GC under the stimulation condition of white noise.

Input stimuli are given by a sequence of random binary black or white checkers. Similar to those filters in neurons, each subunit has a static spatial filter and temporal filter. The different subunits have different spatial locations, such that each subunit can only "read" input stimuli at one specific location, but ignores other parts outside of this location. After input stimuli are convolved by these subunits with spatial and temporal filters, filter outputs then pass a rectification stage in the form of a threshold-linear nonlinearity. The outcomes of all subunits are summed up with a weight for each subunit and pooled to the output unit. Finally, the summation is rectified by another threshold-linear nonlinearity with a higher positive threshold to obtain the final output. Note that since the summation is already positive, a higher threshold is needed to reduce baseline activity and generate spare spiking activity. In the end, a spike train is sampled from the Poisson process.

Note that the current model is implemented for OFF-type retinal GCs with subunits having OFF polarity, that is, the linear filter (as a multiplication of spatial and temporal filter) prefers the negative part of stimulus images. One can simply tell the polarity by fixing the spatial filter to be positive [indicated in white, comparing to black–white stimulus in Fig. 1(c)], and checking the dominant part (the first peak close to spiking time) of the temporal filter to be positive or negative.

Such a model can be considered as a minimal network of GCs with four BCs as subunit components. The recent study [16] used STNMF to identify physical locations of subunit BCs that are pooling to a target retinal GC. However, no functional properties of the bipolar subunits were uncovered there. Here, we use this model to demonstrate how STNMF can be used for characterizing functional properties of bipolar subunits, which includes spatial and temporal filters, nonlinearities, synaptic connections and strengths, and more importantly, subsets of GC spikes contributed by each BC.

B. Spike-Triggered Analysis

The STNMF method is based on a simple and useful method for system identification in visual neuroscience, so-called spike-triggered average (STA) [7], which is similar to the first-order kernel in the Volterra/Wiener kernel series expansion [37]

$$r(t) = \int_{\mathcal{R}} k(\tau)s(t-\tau)d\tau + \int_{\mathcal{R}^2} h(\tau_1,\tau_2)s(t-\tau_1)s(t-\tau_2)d\tau_1\tau_2 + \cdots \quad (1)$$

When stimuli are Gaussian, both kernels $k(\tau)$ and $h(\tau_1, \tau_2)$ can be estimated by reverse correlation. Specifically, for the ith spike r^i occurring at time t_i, one collects a segment of stimuli $s(\tau)^i = s(t_i - \tau)$ that preceded that spike, where the lag τ denotes the timescale of history, into an ensemble of spike-triggered stimuli $\{s(\tau)^i\}$, then averages it over all spikes to obtain the STA filter $k(\tau) = \langle s(\tau)^i \rangle_i$. When the stimuli are spatial–temporal white noise, the 3-D STA filter can be decomposed by a singular value decomposition to obtain the principle temporal filter and spatial receptive field [38]. An illustration of the spatial receptive field of the GC model obtained by STA is shown in Fig. 2.

C. Spike-Triggered Non-Negative Matrix Factorization Analysis

The procedure of STNMF analysis is similar to the one described in [16]. Briefly, to reduce computation costs for the STNMF analysis, we first apply a preprocessing for spike-triggered stimulus ensemble: for the ith spike, the corresponding stimulus segment $s(\tau)^i$ is weighted averaged by the temporal STA filter k_t: $\bar{s}^i = s(\tau)^i \cdot k_t(\tau)$, such that time dimension τ is collapsed. It results in a single frame of stimulus image for the ith spike, called effective stimulus image $\bar{s}^i$. With the ensemble of effective stimulus images $S = \{\bar{s}^i\}_i$ for all spikes as in Fig. 2, one can apply NMF directly in a similar way for face images [17]. Specifically, $S = (s_{ij})$ is an $N \times P$ matrix with indices $i = 1, \ldots, N$ for all N spikes, and $j = 1, \ldots, P$ for all P image pixels. We use a semi-NMF algorithm [39] as

$$S \approx WM \quad (2)$$

where weight matrix W is $N \times K$, module matrix M is $K \times P$, and K is the number of modules. Both stimuli S and weights W can be negative, but modules M are still non-negative.

The idea of semi-NMF can be understood from the perspective of clustering. One could consider the data matrix $S = (s_1, \ldots, s_P)$ as a collection of P vectors as columns. Each vector s_j is a sequence of effective stimulus images at a specific spatial location since the number of pixels is corresponding to the total space of an image. Suppose we have K-means clustering on S with cluster centroids $W = (w_1, \ldots, w_K)$. Thus, each w_k is a sequence of weights, in which each individual weight w_{ik} is the strength between the ith spike-triggered (effective) stimulus image and the module k. Larger w_{ik} means stronger correlation between the ith spike and the module k. Therefore,

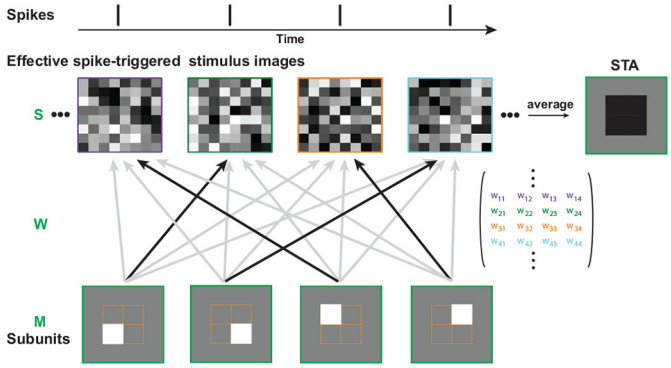

Fig. 2. Illustration of STNMF analysis. For each spike, there is an effective spike-triggered stimulus image. Averaging of this ensemble yields a single STA filter. STNMF reconstructs the image ensemble by approximating with a set of modules and a matrix of weights, such that one of the modules is strongly correlated to one of spikes/images indicated by stronger (black lines) or weaker (gray lines) weights.

the matrix W reflects connection weights between spikes and modules/subunits. Biologically, this is equivalent to the synaptic weight from a presynaptic neuron to a postsynaptic neuron. In this way, STNMF essentially becomes a clustering method by connecting those spikes generated by subunits with a set of modules, such that each module/subunit contributes its corresponding spikes locally at a specific space as illustrated in Fig. 2.

If we let $M = m_{kj}$ denote cluster indicators, that is, $m_{kj} = 1$, if s_j belongs to the cluster k, $m_{kj} = 0$, otherwise. We can write the K-means clustering objective function as

$$F_{K-\text{means}} = \sum_{j=1}^{P}\sum_{k=1}^{K} m_{kj}\|s_j - w_k\|_2^2 = \|S - WM\|_F^2 \quad (3)$$

where $\|v\|$ denotes L_2 norm of a vector v and $\|A\|$ denotes the Frobenius norm of a matrix A. The above objective can be alternatively considered as an objective function for matrix approximation. The difference is that M is not binary but non-negative $M \in R^+$. In addition, a sparseness constraint is added on the columns of M [40], such that

$$F = \|S - WM\|_F^2 + \lambda \sum_{j=1}^{P} \|M_j\|_1^2 \quad (4)$$

where the sparsity parameter $\lambda = 0.1$ throughout the current study, and $\|v\|_1$ is L_1 norm of a vector v. The minimization of F was implemented as an alternating optimization of W and M using the NMF MATLAB toolbox [41].

III. Results

A. Subunit Filters Revealed by STNMF

We set up a minimal model of the retinal GC as in Fig. 1(c) in order to investigate how upstream BCs affect spiking activity of the target GC (see Section II for details). The GC model has four subunits as excitatory BCs that have spatial and temporal filters to compute stimulus. The final spikes of GC are the only output of this model. With the input of stimulus images and the output of GC spikes, the question is how

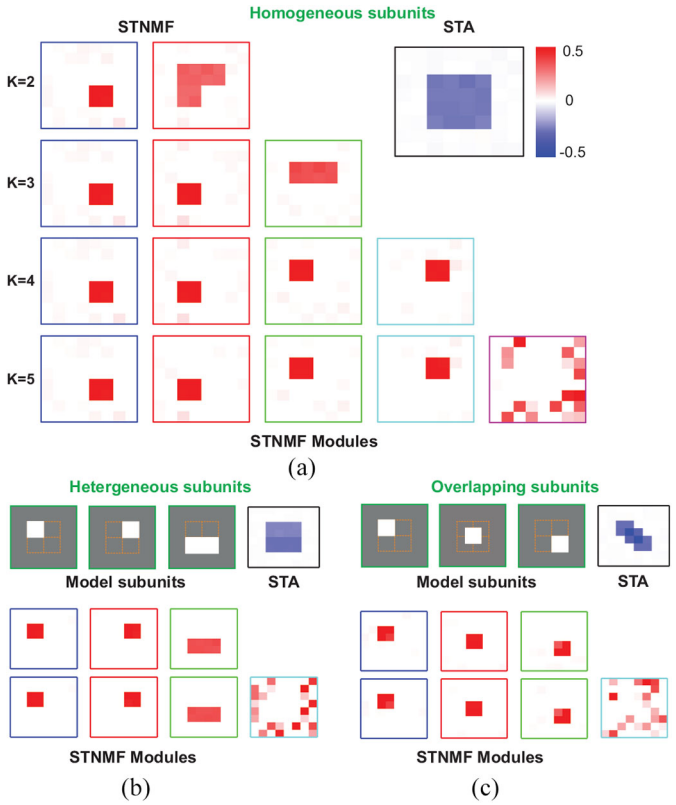

Fig. 3. Exact subunit spatial filters revealed by STNMF. (a) STNMF faithfully recovers the original subunits when $K = 4$. (Left) STNMF modules with the parameter $K = 2, 3, 4, 5$ when subunits are homogeneous. (Inset) Receptive field as spatial STA. Color bar indicates positive (red) and negative (blue). For better visualization, STA is flipped as negative. (b) Heterogeneous subunits by STNMF. (Top) Model subunits with small and big regions, but STA is similar. (Bottom) STNMF modules with $K = 3, 4$. (c) Similar to (b) but with model subunits overlapped in space.

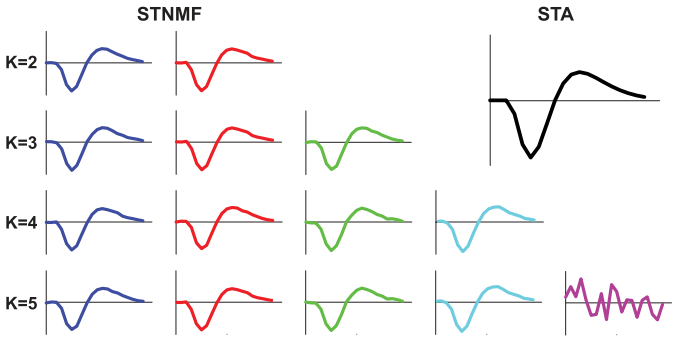

Fig. 4. Recovered subunit temporal filter for different subunits with $K = 2, 3, 4, 5$. (Inset) STA temporal filter. Same color as in Fig. 3(a).

to achieve system identification to find those computational components used by the model.

Similar to typical experimental protocols [7], [16], we use visual stimuli consisting of a sequence of white noise as black and white checkers randomly distributed in space and time domain. Under this stimulation, the receptive field of GC can be computed from spiking response by a method called STA [7] (see Section II). The STA is equivalent to an average characteristic of GCs, so the shape of STA is a combination of all subunits in space as shown in Fig. 2. Note that inhibitory neurons are not included in the model as the surround of the receptive field is barely triggered under the white noise stimulation [16].

However, computations are done by the subunits of the model in the first case. Extracting details of these subunits can be achieved by another recently proposed method, called STNMF [16]. The framework of STNMF (see Section II) is illustrated in Fig. 2. In the previous study [16], STNMF was shown to identify spatial receptive fields of subunits of modeled GCs. Here, we recap this finding and in addition, show that the working principle of STNMF is to reveal underlying nonlinear computations of the network.

STNMF can be seen as a type of method for clustering (see Section II). Similar to other clustering methods, the number of clusters, here modules K, is unknown. As K is a free parameter, one has to choose K before using STNMF. Similar to the previous study [16], the number of meaningful subunit modules obtained by STNMF is not changed when K is large enough, that is, larger than the actual number of subunits used in the model [Fig. 3(a)]. In other words, the result is convergent when using a large K, which is also seen by the convergence of the Akaike information criterion when K is larger [42]. This unique advantage of STNMF, together with the constraint of nonnegativity condition, distinguishes STNMF from other traditional classification methods (see Section II).

With $K = 4$, STNMF finds the exact number and structure of subunits. When $K = 5$, the extra subunit in Fig. 3(a) is just noise with a low degree of coherence or autocorrelation in space. This signature can be used to determine the number of subunits when the actual number is unknown in real biological data [16].

To test the robustness of STNMF, we apply perturbations of the subunit structure to the model. The hypothesis is that the underlying computation is corresponding to the subunit structure. GC spiking responses are induced by subunit computation. If STNMF only reflects the property of stimulus images, such as using NMF for face images [17], without taking into account spiking computation, then the change of model subunits will not change the STNMF subunit output. Instead, when STNMF can capture the underlying computation of the network, one would expect that STNMF captures the change of the local structure of subunits. Therefore, we test the hypothesis that subunits identified by STNMF are changed when the computation of the network is changed.

We manipulate the spatial structure of subunits as in Fig. 3(b) and (c). One perturbation is to have different sizes of subunits. Fig. 3(b) shows the case where the network has three subunits: two are in the same size, and one has a doubled size. By analyzing spikes, a similar STA is obtained. However, STNMF precisely captures three subunits, although they have different sizes. Another perturbation is to have overlapped space between subunits. Fig. 3(c) shows that the network has three overlapping subunits. Similarly, the STA is a combination of all subunits. STNMF, on the other hand, can recover all three subunits separately. In all the cases, stimulus images are the same, but spikes are different due to changes of subunits and computations. Taken together, these results show that

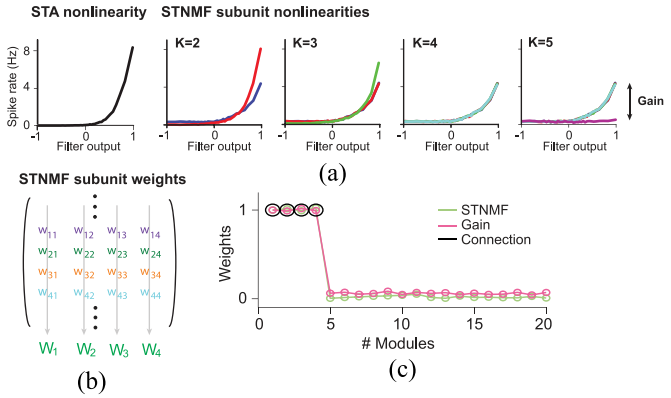

Fig. 5. Subunit connection weights revealed by STNMF. (a) Nonlinearities of each filter. (Left) Nonlinearity of STA filter. (Right) Nonlinearities of each STNMF subunits with $K = 2, 3, 4, 5$. Same color as in Fig. 3(a). (b) STNMF subunit weights. Each column of the weight matrix is corresponding to a subunit. The sum of each column represents STNMF weight for each subunit. (c) STNMF weights (green), nonlinearity gains (purple) and connection weights (black) are identical.

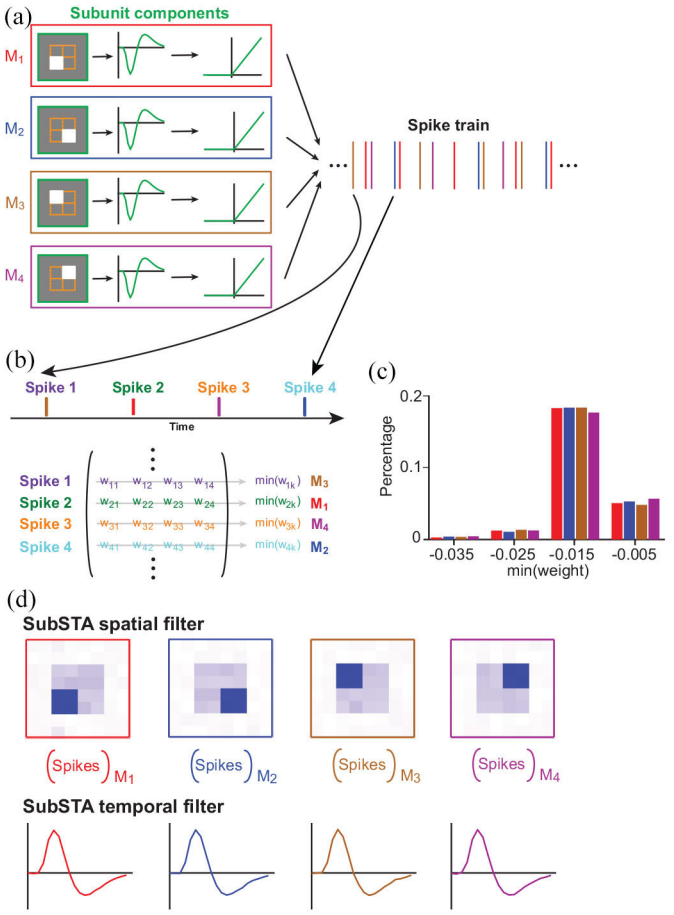

Fig. 6. Classification of spikes by STNMF. (a) Each spike can be seen as a contribution from one subunit M_j. (b) Each spike is labeled with a corresponding subunit according to the minimal value of STNMF weight per row. Each column aligns with a subunit, and each row with a spike. (c) Histogram of weight minimums from four subsets of spikes showing a uniform distribution across four subunits. (d) SubSTA spatial and temporal filters computed from one subset of spikes with STA analysis.

STNMF indeed captures the computation within the network, but not stimulus images themselves.

After recovering spatial subunit filters, one can also obtain the corresponding temporal filter for each subunit (Fig. 4). Temporal filter can be computed with the spatial filter obtained by STNMF. A sequence of stimulus images is convolved by each spatial filter and then summed over all pixels to obtain a 1-D output. The spike-triggered analysis is then applied to the output to find temporal filters.

Note that temporal filters of subunits are not obtained by STNMF directly, since here the effect of time has been removed during preprocessing of STNMF analysis. However, it is possible to obtain both spatial and temporal modules simultaneously (see [42]).

B. Subunit Connection Weight Revealed by STNMF

Besides spatial and temporal filters for each subunit, there is one last component in the model that needs to be identified: connection weight of each subunit. For this purpose, we calculate the nonlinearity of each subunit by using its spatial and temporal filter and then averaged as a histogram mean [43]. To do so, stimuli are first convolved with spatial and temporal filters to obtain a generator signal. It is then binned into 40 bins with variable bin sizes so that each bin contains the same number of data points. Then, the nonlinearity is displayed as a histogram mean by plotting the average generator signal against the average spike rate for each bin.

Fig. 5(a) shows that nonlinearities are changing with the parameter K. When $K = 4$, all nonlinearities of four subunits are overlapped, since subunit connection weights used in the model are the same. Similarly, when $K = 5$, a weak (flat) nonlinearity for the fifth subunit occurs due to noise. The strength of nonlinearity can be characterized by the gain or magnitude of nonlinearity. The gain reflects how much the subunit contributes to spiking responses, so it is closely related to the subunit weight.

As STNMF focuses on all the spikes, one expects that gain can be revealed by STNMF. Indeed, we find the gain can be extracted from the weight matrix of STNMF [Fig. 5(b)]. By averaging each column of the weight matrix, one can obtain a weight W_j for each subunit j. Interestingly, the weight W_j is identical to the connection weight of subunit j. All of these three measures, STNMF weight, nonlinearity gain, and connection weight, are matched very well [Fig. 5(c)]. These results indicate that STNMF subunit weights provide a good estimate for actual subunit connection weights. Such information is difficult to obtain in biological data due to limitations of experimental techniques [16].

C. Classification of Spikes by STNMF

In the GC model, final spikes are contributed by four subunits, thus, each spike could be induced by one subunit. Inspired by the clustering viewpoint of STNMF, one may wonder if STNMF can be used to classify all the GC spikes to four subsets of spikes, where each subset is mainly, if not completely, contributed by one specific subunit.

The STNMF weight matrix is subunit specific for every column, but it is also spike specific for every row. As each row corresponds to one individual spike, every spike can be labeled or classified according to one subunit as illustrated in Fig. 6(a). Note that the model is designed for OFF-type GCs,

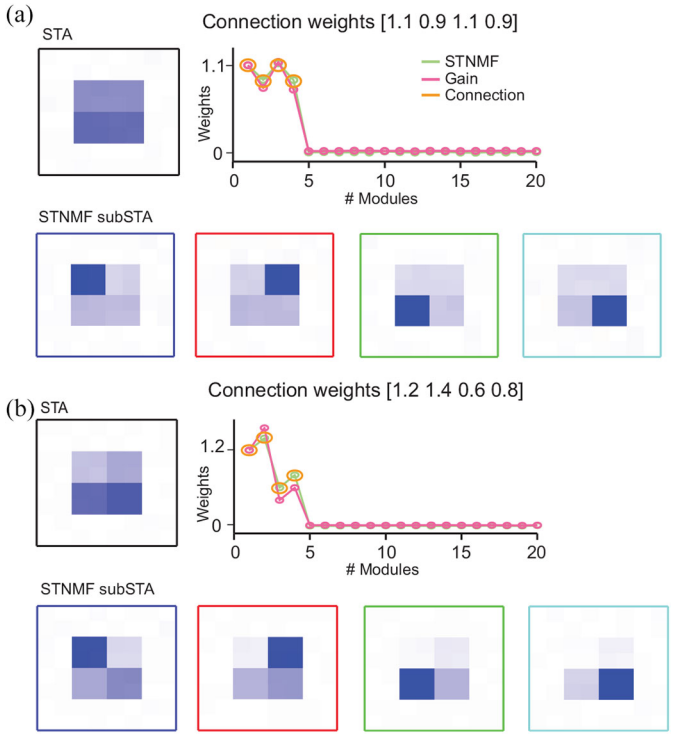

Fig. 7. Effect of heterogeneous connection weights. (a) Connection weights [1.1, 0.9, 1.1, 0.9]. (Top, left) Spatial STA filter. (Top, right) Module weights computed as normalized gain, NMF weight, and connection weight designed in the model. (Bottom) SubSTA computed from subsets of spikes. (b) Similar to (a) but with connection weights as [1.2, 1.4, 0.6, 0.8].

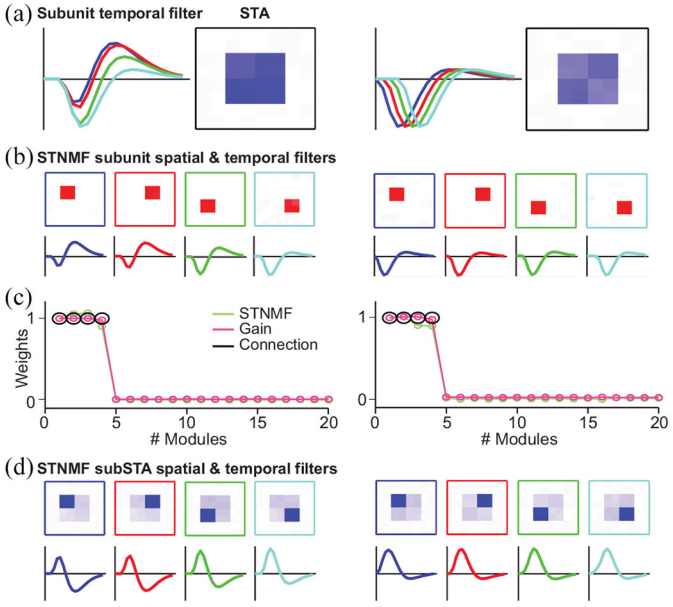

Fig. 8. Effect of different temporal filters on STNMF results. (a) Heterogeneous temporal filters with different amplitudes (left) and delays (right). Spatial STAs are similar. (b) STNMF subunit spatial and temporal filters. Note both amplitude and delay changes are recovered by STNMF. (c) Matching STNMF weight with nonlinearity gain and connection weight. (d) Spatial and temporal filters of STNMF subSTA computed from each subset of spikes. Note here the polarity of STNMF subSTAs is opposite to STNMF subunits that are always positive.

and since subunits are always non-negative, one can take the minimal value per row in the weight matrix w_{ij}, for instance, $\min(w_{1k}) = \min_j(w_{1j})$ for the first row, that is, the first spike as in Fig. 6(b). The minimum index j is this spike's label for the subunit k. The value of the minimum weight can measure the contribution made by this subunit.

Now, we can classify every spike into one specific kth subunit M_k since $k = \min_j\{1, 2, 3, 4\}$. For instance, for the first spike, spike 1 is associated with the first row and the minimum value is at the third column. Therefore, the first spike should be associated with the third subunit M_3 with $k = 3$. After doing this loop for all rows/spikes, we can label each spike with a specific subunit. For this particular model cell, we obtain four subsets of spikes for four subunits, respectively. For every spike, there is a min (weight). The histogram of these min(weight) shows that these weights are indeed uniformly distributed across four subunits as in Fig. 6(c), meaning the connection weight of every subunit is the same as in the model.

Then, for each subset of spikes, we compute STA to obtain the spatial and temporal filters of each subunit as in Fig. 6(d). These spatial filters are similar to STNMF subunits. We name these filters as "subSTAs." As subSTA reflects the contribution of one subunit to GC spikes, one can test the robustness of subSTA by modifying the strength of the subunit connection. We manipulate the connection weights of four subunits as [1.1, 0.9, 1.1, 0.9] and [1.2, 1.4, 0.6, 0.8]. Three measures, STNMF weights, nonlinearity gains and connection weights, are tightly matched (Fig. 7). As a result, the subSTAs computed from subsets of spikes classified by STNMF are also faithfully similar to the subunits used in the model.

With preprocessing of data for STNMF, we obtain the ensemble of effective stimulus images, where temporal correlation can be removed. Then, the question is that whether the results obtained by STNMF could be changed when temporal filters in the model change. In order to test this, we use two different perturbations for the temporal filter: 1) different amplitudes and 2) different delays [Fig. 8(a)]. Similar to the previous study [16], results are very robust. All properties of subunits, such as spatial and temporal filters [Fig. 8(b)], and overlapped STNMF weight, nonlinearity gain, and connection weight [Fig. 8(c)], are robust. In addition, the subSTAs obtained by classified spikes are also faithfully reproduced. Using spatial subSTAs, we obtain temporal components, which are closely matched to subunit temporal filters designed in the model [Fig. 8(d)].

D. Application of STNMF to Real Retinal Data

We apply STNMF to retinal GC data published previously [16], [42]. Briefly, salamander retinal GCs were recorded with a multielectrode array using similar stimulation of spatiotemporal white noise as in the model above. An overview of the application of STNMF to one GC is shown in Fig. 9, which is similar to pervious results [16]. The standard spike-triggered analysis can obtain the spatial receptive field, temporal filter, and nonlinearity [Fig. 9(a)]. For this particular cell, STNMF can find nine subunits, that is, BCs converged to this GC. Computational properties of

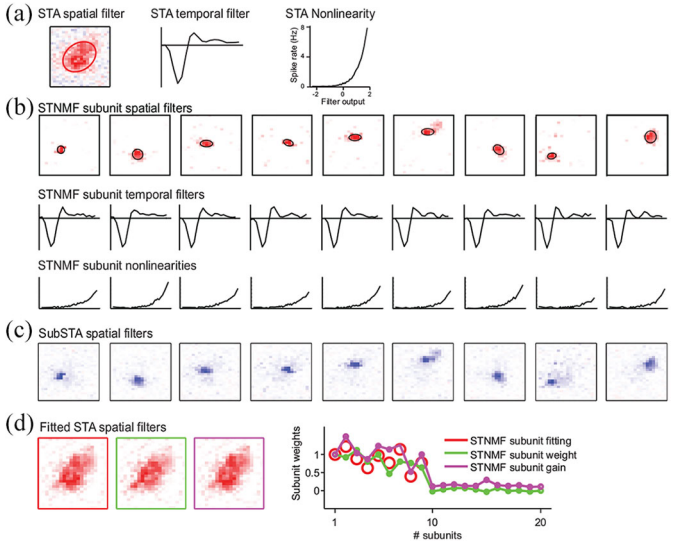

Fig. 9. Overview of biological properties of a GC identified by STNMF. (a) STA spatial filter, temporal filter, and nonlinearity. (b) STNMF subunit spatial filters, temporal filters, and their nonlinearities. Circles are outlines fitted with 2-D Gaussian. (c) SubSTA spatial filters from subsets of classified GC spikes. (d) Reconstructed STA spatial filter by fitting with STNMF subunit (red), STNMF weights from weight matrix (green), and STNMF subunit gain from nonlinearity (orange), respectively. These three measures are highly correlated (right).

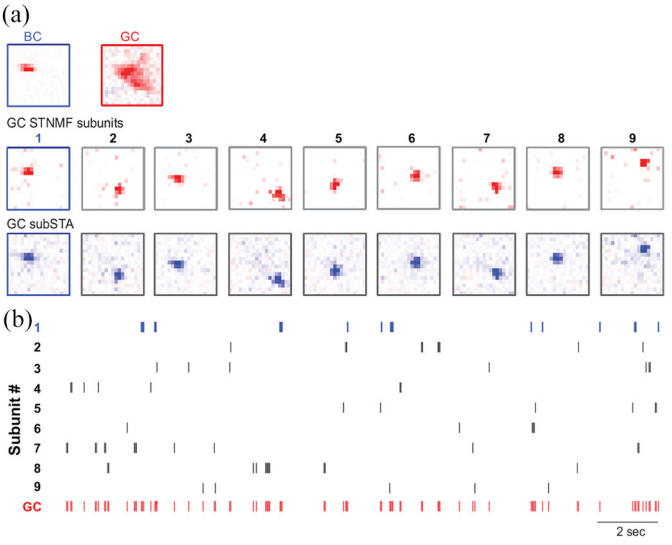

Fig. 10. Subsets of spikes contributed by BC. (a) Receptive fields of BCs and GC (top). STNMF subunit receptive fields (middle). SubSTA computed from subsets of classified GC spikes. (b) Subsets of GC spikes classified by STNMF. The first spike train is contributed by the BC together with the original GC spike train. BC data taken from [16].

these BCs include spatial receptive field, temporal filter, and nonlinearity [Fig. 9(b)]. All the spikes from this GC can be further classified into nine subsets of spikes according to each subunit. Another STA analysis can obtain subSTA spatial filter for each subunit. Similar to the above results, these subSTAs are highly matched to subunit receptive fields identified by STNMF [Fig. 9(c)].

The synaptic strength of each subunit can be computed in three different ways: 1) weights of fitting GC receptive field with all subunit receptive fields; 2) subunit weights calculated from the weight matrix of STNMF as above; and 3) subunit gains calculated from each nonlinearity of the subunit. Synaptic strengths from the last two measures can also be used to fit the GC receptive field, which results in a similar result [Fig. 9(d) left]. Although there is no ground truth about actual synaptic weights between BCs and the GC, all these three measures are highly correlated as in Fig. 9(d) right.

Once all GC spikes are classified into subsets of BC spikes, one can test if these BC spikes are contributed by one actual BC. The dataset in [16] provides such a possibility as there is one simultaneously recorded one BC with a large population of GCs. One example is shown in Fig. 10(a) with the receptive fields of BC and GC. Again, subunits identified by STNMF and subSTAs computed by subsets of subunit spikes are highly overlapped. The first subunit is highly overlapping with the recorded BC, which indicates actual connection between BC and GC.

As GC spikes can be classified into subsets of spikes for each BC in Fig. 10(b), one can test if there is a functional connection between the first subunit and recorded BC, apart from that they are physically located at the same place. To explore this, we calculate correlation coefficient (CC) between the BC membrane potential and subunit spike trains. We find that there is stronger correlation between the first subunit and BC (CC = 0.14), compared to correlations of other subunits (CC = 0.02 ± 0.01, mean $\pm$ std).

CC is a measure of linear correlation, which can miss nonlinear correlation. To investigate whether there is nonlinear coupling between the BC membrane potential and other subunits, additional measures of dependence are needed. The relationship between subunits and the BC membrane potential is complicated by the fact that subunit activity consists of a discrete number of events (i.e., neuronal spikes), whereas the BC membrane potential is a continuous quantity. We analyze this relationship computing STA of the BC membrane potential, binning spike trains into short time windows (33 ms) and modeling the joint distribution of spike counts and membrane potential.

First, we calculate STA of BC membrane potential and find that it gives an amplitude for the 1st subunit as 1.35 and other subunits as 0.14 ± 0.08. Then, we also apply vine copula with various parametric copula families, which is a general statistical model of joint distributions representing couplings between discrete signals, for example, subunit spike counts, and continuous signals, for example, BC membrane potential. When signals are mixed as discrete quantities or continuous quantities, vine copulas as models of the mixed joint distributions is quite useful [44]. These models include various choices for parametric bivariate copulas. Here, using Gaussian, student t, and rotated Clayton copula families, we use canonical vine to extend these bivariate models to multivariate models [44]. We fit these models to mixed data and use copula parameters to quantify coupling strengths. For all copula families, coupling strength of zero corresponds to independence. Here, we find that vine copula gives a coupling strength for the first subunit as 7.12 and other subunits as 0.19 ± 0.18. These results indicate that subset of spikes from the first subunit is indeed contributed by this specific BC.

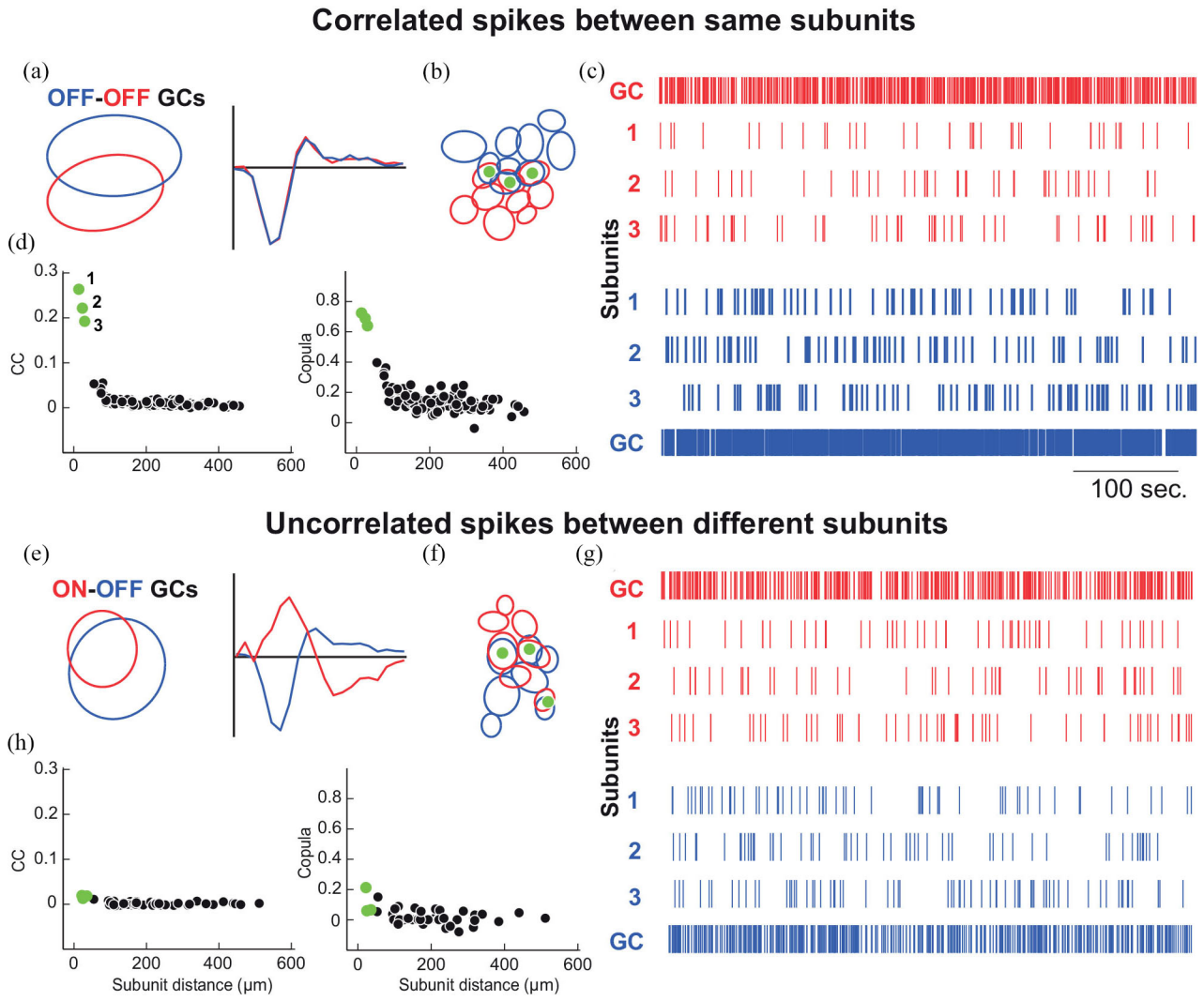

Fig. 11. Subunit spikes induced by BCs but not driven by shared stimuli. (a)–(d) Correlated spikes between same subunits. (a) Pair of OFF-OFF GCs with overlapping receptive fields and identical OFF-type temporal filters. (b) Subunits obtained by STNMF for this pair of GCs. Three green pairs of subunits shared between these two GCs. (c) Three subsets of spikes corresponding to three pairs of BCs colored in green in (b). Original GC spikes are shown at the top and bottom. (d) Correlation coefficient (CC, left) and copula strength (right) for all pairs of BCs in (b). Top three numbered by 1, 2, and 3 are three subsets of spikes in (c). (e)–(h) Similar to (a)–(d) but uncorrelated spikes between different subunits for a pair of BCs with one OFF and one ON type of GC. Three pairs of BCs colored in green are overlapped, but have different identities as ON and OFF BCs as indicated by their temporal filters in (e). Correlations between BCs are close to zero.

Further investigation of coupling between subunits can be done at the population level. For each GC, there are a few subunits found by STNMF. One can look at pairs of two GCs such that there are some overlapped subunits as illustrated by spatial receptive fields in Fig. 11(a), where one pair of GCs with the same type of fast OFF GCs [16] is shown, together with another pair of GCs with different cell types as fast OFF and ON cells. Such information about cell types can be seen by their temporal filters, where fast OFF GCs have identical filter shape and ON GC has positive polarity at its first peak in Fig. 11(b) and (f).

For each GC, a few subunits revealed by STNMF are shown in Fig. 11(c). There are considerable overlapped subunits in a pair of GGs. When subsets of spikes are obtained by STNMF, there is a possibility that correlation between spikes is induced by overlapped spatial location, rather than produced by the same BC. Since spikes driven by the same stimulus input are highly correlated, in particular, when spatiotemporal white noise stimuli are replaced by spatially uniform white noise [43], two trains of subunit spikes could be correlated when these two subunits are located in the same space. As a result, they read the same stimuli at this spacial location. Such a possibility can be examined by a population analysis of GCs.

For the same type of GCs with overlapped subunits, the sample pair shown in Fig. 11(a) has three overlapped subunits in Fig. 11(b). Therefore, there are three shared subunit spike trains from each GC in this pair as in Fig. 11(c). We find that for each pair of overlapped subunits, their spike trains are highly correlated as characterized by CC and vine copula coupling strength as in Fig. 11(d). However, the result obtained from different types of GCs is different as shown in Fig. 11(e)–(h). The sample pair shown in Fig. 11(e) has one OFF cell and one ON cell. This pair also shows highly overlapped subunits

that have three spike trains classified by STNMF for each GC. In contrast to the pair of the same cell type, there is no correlation in subunit spike trains between overlapped subunits. The biological view is that there are ON BCs in ON GC and OFF BCs in OFF GC. Although overlapped ON and OFF BCs are located in close by spatial location due to the 3-D structure of the retina, they are driven by the same stimulus but generate different spikes only when stimuli present different parts: bright images for ON spikes versus dark images for OFF spikes. As a result, these two spike trains from ON and OFF BCs are not correlated. In other words, they are decorrelated due to the uncorrelated stimuli.

Taken together, these results show that correlations between subunit spikes are not driven by stimuli, but by the same BC identified by STNMF. These results confirm and extend the previous observation [16], where the identity of BC was justified by subunit physical location only. Our results here go one step further to show that the identity of BCs can be detected by means of functional properties, that is, subsets of GC spikes contributed by the BC.

IV. Discussion

In this study, we propose STNMF as a useful method for system identification of neuronal circuits. With a simple network model of retinal GCs with clearly defined subunit components, connections, and weights, STNMF is able to reveal all these structural components within the network. Furthermore, STNMF allows us to classify the whole set of spikes of a GC into a few subsets of spikes, such that each subset of spikes is mainly contributed by one specific subunit. When applying STNMF to the retinal data, biological network components can be revealed. In particular, the classification of GC spikes shows that a subset of spikes is mainly contributed by one BC that connects to the target GC.

Besides confirming what has been shown in the previous study [16], where STNMF detected a layout of physical location of BCs, here we significantly extend the power of the STNMF approach by analyzing the STNMF weight matrix. As a result, the weight matrix reflects functional properties of BCs, including their synaptic weights and contributed spikes for the downstream GC. Therefore, STNMF is useful for uncovering relevant functional and structural properties of neuronal circuits.

A. Neuronal Circuit at Single Cell Level

From the viewpoint of postsynaptic neuron, properties of neuronal circuits revealed by STNMF include locations of presynaptic neurons and their synaptic connections and strengths/weights.

Structural components of single postsynaptic neuron are revealed as a layout of presynaptic neurons. The organization of such a layout could be complex or simple. Depending on the type of neurons and animal species, the number of presynaptic neurons could be very large or small. For instance, in the cerebellum, Purkinje cells have a large dendritic tree with thousands of presynaptic connections [45], whereas unipolar cells have only one presynaptic fiber [46], and granular cells have an average of four presynaptic fibers [47]. In salamander retina that was used in the current study, there are a few BCs per GC [16].

Synaptic connections and weights are more difficult to identify. Traditionally, direct measure of these properties is established by pairwise (or triple and more) electrodes recording from presynaptic and postsynaptic neurons [12], [13]. Here, we find that STNMF can identify these properties using simulated cells. Verification of this observation by experiments is possible for large-scale recordings of spiking and/or imaging of the calcium signal activity of a population of neurons, where inferring connections between neurons are feasible, for example, by means of graph theory or complex network analysis [48].

B. Classifying Spikes of Postsynaptic Neurons

A postsynaptic neuron receives signal from a set of presynaptic neurons in multiple channels. Each presynaptic signal is ubiquitous in that the information from input to output is transformed in nonlinear fashion. Such nonlinearity is evidenced by spiking activity of a neuron, where incoming signal with mixed positive and negative signs is eventually transferred to a sequence of digital spikes. Such a feature becomes a fundamental principle of neuronal computation since the spiking mechanism was uncovered 60 years ago [49].

STNMF implements the analysis of every single spike for one postsynaptic neuron. As a result, STNMF naturally labels every spike to one of the presynaptic neurons during the process of factorization. The relationship between spikes and presynaptic neurons is encoded in the STNMF weight matrix. Here, we decoded this information and classified all the spikes of a GC into a few subsets of spikes such that one subset of spikes is corresponding to one presynaptic neuron. In other words, these subsets of spikes are closely correlated to activities of presynaptic neurons.

Although the activity of a BC in the retina is traditionally viewed as graded signal without spikes, it could still generate large deflections of the membrane potential that is similar to a spike event [50], we find that there exists strong correlation between its membrane potential and spikes. Both are generated by the stimulus of white noise checkers with a size of 30 μm. Therefore, it is possible, when stimuli are strong enough, to trigger strong activity in BC membrane potential, which, in turn, can produce spiking activity in the connected GC. Indeed, one recent study found that one BC could trigger GC spikes under white noise bars stimulus by fitting a two-layer linear–nonlinear network similar to the model used in our current study [51]. As for other parts of the brain, traditional view is that one presynaptic neuron may not be enough to drive a postsynaptic neuron to fire a spike. However, a caveat here is that dendritic spikes could be larger than what we expected [52].

Simultaneous recording of upstream BC and downstream GC in the retina is ideal to test the identity of subunits revealed by STNMF [16]. Our current results go one step further to uncover functional identity and potential contribution of BCs for their downstream GCs. The recent advance in experimental techniques make it possible to record simultaneously

signals in soma and multiple dendrites with both imaging and electrophysiology [53], [54]. This protocol could provide an interesting test for the utility of STNMF.

C. System Identification of Neural Networks

Retinal GCs carry out visual computations from stimulus to responses. One of the central problems is to find encoding and decoding principles between stimulus and response [55], for which a number of possible methods of system identification have been proposed in both visual neuroscience and computer vision [10], [56], [57].

Input–output relation of sensory information has been traditionally modeled by dynamic functions, for example, the Laguerre–Volterra model [58], [59], or trainable network models through unsupervised (e.g., spike-timing-dependent plasticity) [60] or supervised learning [61]. In contrast, detailed neuroscience knowledge provides a bottom-up approach with neural network models [57], [62], whereas the underlying network structure needs to be cleverly designed by hand or selected from a massive pool of possible network architectures [63].

It has been observed in neuroscience experiments that specific features are encoded by specific neurons in visual systems, and also in other sensory systems [64]. NMF itself can be viewed as a generative model [65], whereas the convolutional neural network is supervised. However, both types of methods can be used to extract underlying features from data. Their potential usages for modeling input–output relation is evident: local structure features play an important role in computation [16], [66], [67]. Indeed, recent studies show that some NMF variants can go beyond shallow layered networks, like our modeled retina network with only two layers, to use a framework of deep architectures [68]–[72] to learn a hierarchy of attributes of given datasets. The combination of NMF and deep convolutional neural network holds promise to uncover hierarchical structures of neural networks [73]–[77]. Therefore, further extension of our current STNMF is likely to be fruitful for understanding deep architecture of neuronal systems in the brain.

ACKNOWLEDGMENT

The authors would like to thank Y. Zheng, Y. Zhang, L. He, Y. Yue, and K. Du for helpful discussions.

REFERENCES

[1] M. Helmstaedter, K. L. Briggman, S. C. Turaga, V. Jain, H. S. Seung, and W. Denk, "Connectomic reconstruction of the inner plexiform layer in the mouse retina," *Nature*, vol. 500, no. 7461, pp. 168–174, 2013.

[2] H. Zeng and J. R. Sanes, "Neuronal cell-type classification: Challenges, opportunities and the path forward," *Nat. Rev. Neurosci.*, vol. 18, no. 9, pp. 530–546, 2017.

[3] R. E. Marc, B. W. Jones, C. B. Watt, J. R. Anderson, C. Sigulinsky, and S. Lauritzen, "Retinal connectomics: Towards complete, accurate networks," *Progr. Retinal Eye Res.*, vol. 37, pp. 141–162, Nov. 2013.

[4] H. S. Seung and U. Sümbül, "Neuronal cell types and connectivity: Lessons from the retina," *Neuron*, vol. 83, no. 6, pp. 1262–1272, 2014.

[5] J. R. Sanes and R. H. Masland, "The types of retinal ganglion cells: Current status and implications for neuronal classification," *Annu. Rev. Vis. Sci.*, vol. 38, pp. 221–246, Jul. 2015.

[6] J. B. Demb and J. H. Singer, "Functional circuitry of the retina," *Annu. Rev. Vis. Sci.*, vol. 1, pp. 263–289, Nov. 2015.

[7] E. J. Chichilnisky, "A simple white noise analysis of neuronal light responses," *Network*, vol. 12, no. 2, pp. 199–213, 2001.

[8] J. W. Pillow *et al.*, "Spatio-temporal correlations and visual signalling in a complete neuronal population," *Nature*, vol. 454, no. 7207, pp. 995–999, 2008.

[9] J. M. McFarland, Y. Cui, and D. A. Butts, "Inferring nonlinear neuronal computation based on physiologically plausible inputs," *PLoS Comput. Biol.*, vol. 9, no. 7, Jul. 2013, Art. no. e1003143.

[10] Z. Yu *et al.*, "Toward the next generation of retinal neuroprosthesis: Visual computation with spikes," *Engineering*, vol. 6, no. 4, pp. 449–461, Apr. 2020.

[11] Y. Zhang *et al.*, "Reconstruction of natural visual scenes from neural spikes with deep neural networks," *Neural Netw.*, vol. 125, pp. 19–30, May 2020.

[12] S. Song, P. J. Sjöström, M. Reigl, S. Nelson, and D. B. Chklovskii, "Highly nonrandom features of synaptic connectivity in local cortical circuits," *PLoS Biol.*, vol. 3, no. 3, pp. 507–519, 2005.

[13] X. Jiang *et al.*, "Principles of connectivity among morphologically defined cell types in adult neocortex," *Science*, vol. 350, no. 6264, p. aac9462, 2015.

[14] J. J. Jun *et al.*, "Fully integrated silicon probes for high-density recording of neural activity," *Nature*, vol. 551, no. 7679, pp. 232–236, Nov. 2017.

[15] M. Yang *et al.*, "MATRIEX imaging: Multiarea two-photon real-time in vivo explorer," *Light Sci. Appl.*, vol. 8, no. 1, p. 109, Nov. 2019.

[16] J. K. Liu *et al.*, "Inference of neuronal functional circuitry with spike-triggered non-negative matrix factorization," *Nat. Commun.*, vol. 8, no. 1, p. 149, Jul. 2017.

[17] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, Oct. 1999.

[18] X. Zhao *et al.*, "Scalable linear visual feature learning via online parallel nonnegative matrix factorization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 12, pp. 2628–2642, Dec. 2016.

[19] M. Ye, Y. Qian, and J. Zhou, "Multitask sparse nonnegative matrix factorization for joint spectral–spatial hyperspectral imagery denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2621–2639, May 2015.

[20] K. Kwon, J. W. Shin, and N. S. Kim, "NMF-based speech enhancement using bases update," *IEEE Signal Process. Lett.*, vol. 22, no. 4, pp. 450–454, Apr. 2015.

[21] N. Guan, D. Tao, Z. Luo, and B. Yuan, "NeNMF: An optimal gradient method for nonnegative matrix factorization," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2882–2898, Jun. 2012.

[22] B. Gao, W. L. Woo, and B. W. Ling, "Machine learning source separation using maximum a posteriori nonnegative matrix factorization," *IEEE Trans. Cybern.*, vol. 44, no. 7, pp. 1169–1179, Jul. 2014.

[23] X. Pei, T. Wu, and C. Chen, "Automated graph regularized projective nonnegative matrix factorization for document clustering," *IEEE Trans. Cybern.*, vol. 44, no. 10, pp. 1821–1831, Oct. 2014.

[24] T. Blumensath, "Directional clustering through matrix factorization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 10, pp. 2095–2107, Oct. 2016.

[25] H. Wang, F. Nie, H. Huang, and F. Makedon, "Fast nonnegative matrix tri-factorization for large-scale data co-clustering," in *Proc. Int. Joint Conf. Artif. Intell. (IJCAi)*, vol. 22, 2011, p. 1553.

[26] K. Devarajan, "Nonnegative matrix factorization: An analytical and interpretive tool in computational biology," *PLoS Comput. Biol.*, vol. 4, no. 7, Jul. 2018, Art. no. e1000029.

[27] K. Gold, C. Havasi, M. Anderson, and K. C. Arnold, "Comparing matrix decomposition methods for meta-analysis and reconstruction of cognitive neuroscience results," in *Proc. 24th Int. FLAIRS Conf.*, 2011, pp. 1–6.

[28] R. Maruyama *et al.*, "Detecting cells using non-negative matrix factorization on calcium imaging data," *Neural Netw.*, vol. 55, pp. 11–19, Jul. 2014.

[29] M. Beyeler, N. Dutt, and J. L. Krichmar, "3D visual response properties of MSTd emerge from an efficient, sparse population code," *J. Neurosci.*, vol. 36, no. 32, pp. 8399–8415, 2016.

[30] E. A. Pnevmatikakis *et al.*, "Simultaneous denoising, deconvolution, and demixing of calcium imaging data," *Neuron*, vol. 89, no. 2, pp. 285–299, 2016.

[31] P. Zhou *et al.*, "Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data," *eLife*, vol. 7, Feb. 2018, Art. no. e28728.

[32] P. O. Hoyer, "Non-negative matrix factorization with sparseness constraints," *J. Mach. Learn. Res.*, vol. 5, pp. 1457–1469, Nov. 2004.

[33] J. Eggert and E. Korner, "Sparse coding and NMF," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, vol. 4. Budapest, Hungary, 2004, pp. 2529–2533.
[34] 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. 607–679, 1996.
[35] P. O. Hoyer, "Modeling receptive fields with non-negative sparse coding," *Neurocomputing*, vols. 52–54, pp. 547–552, Jun. 2003.
[36] T. Gollisch and M. Meister, "Modeling convergent on and off pathways in the early visual system," *Biol. Cybern.*, vol. 99, nos. 4–5, pp. 263–278, 2008.
[37] R. A. Sandler and V. Z. Marmarelis, "Understanding spike-triggered covariance using wiener theory for receptive field identification," *J. Vis.*, vol. 15, no. 9, p. 16, Jul. 2015.
[38] J. L. Gauthier *et al.*, "Receptive fields in primate retina are coordinated to sample visual space more uniformly," *PLoS Biol.*, vol. 7, no. 4, Apr. 2009, Art. no. e1000063.
[39] C. H. Ding, T. Li, and M. I. Jordan, "Convex and semi-nonnegative matrix factorizations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 1, pp. 45–55, Jan. 2010.
[40] H. Kim and H. Park, "Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis," *Bioinformatics*, vol. 23, no. 12, pp. 1495–1502, May 2007.
[41] Y. Li and A. Ngom, "The non-negative matrix factorization toolbox for biological data mining," *Source Code Biol. Med.*, vol. 8, no. 1, p. 10, 2013.
[42] A. Onken, J. K. Liu, P. C. R. Karunasekara, I. Delis, T. Gollisch, and S. Panzeri, "Using matrix and tensor factorizations for the single-trial analysis of population spike trains," *PLoS Comput. Biol.*, vol. 12, no. 11, Nov. 2016, Art. no. e1005189.
[43] J. K. Liu and T. Gollisch, "Spike-triggered covariance analysis reveals phenomenological diversity of contrast adaptation in the retina," *PLoS Comput. Biol.*, vol. 11, no. 7, Jul. 2015, Art. no. e1004425.
[44] A. Onken and S. Panzeri, "Mixed vine copulas as joint models of spike counts and local field potentials," in *Advances in Neural Information Processing Systems*. Red Hook, NY, USA: Curran, 2016, pp. 1325–1333.
[45] L. An *et al.*, "Coding capacity of Purkinje cells with different schemes of morphological reduction," *Front. Comput. Neurosci.*, vol. 13, p. 29, May 2019.
[46] L. An *et al.*, "Intrinsic and synaptic properties shaping diverse behaviors of neural dynamics," *Front. Comput. Neurosci.*, vol. 14, p. 26, Apr. 2020.
[47] V. Zampini, J. K. Liu, M. A. Diana, P. P. Maldonado, N. Brunel, and S. Dieudonné, "Mechanisms and functional roles of glutamatergic synapse diversity in a cerebellar circuit," *eLife*, vol. 5, Sep. 2016, Art. no. e15872.
[48] D. S. Bassett and O. Sporns, "Network neuroscience," *Nat. Neurosci.*, vol. 20, no. 3, pp. 353–364, 2017.
[49] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *J. Physiol.*, vol. 117, no. 4, pp. 500–544, 1952.
[50] T. Baden, P. Berens, M. Bethge, and T. Euler, "Spikes in mammalian bipolar cells support temporal layering of the inner retina," *Current Biol.*, vol. 23, no. 1, pp. 48–52, 2013.
[51] N. Maheswaranathan, D. B. Kastner, S. A. Baccus, and S. Ganguli, "Inferring hidden structure in multilayered neural circuits," *PLoS Comput. Biol.*, vol. 14, no. 8, 2018, Art. no. e1006291.
[52] J. J. Moore *et al.*, "Dynamics of cortical dendritic membrane potential and spikes in freely behaving rats," *Science*, vol. 355, no. 6331, p. eaaj1497, 2017.
[53] M. Li, F. Liu, H. Jiang, T. S. Lee, and S. Tang, "Long-term two-photon imaging in awake macaque monkey," *Neuron*, vol. 93, no. 5, pp. 1049–1057, 2017.
[54] R. Ding *et al.*, "Targeted patching and dendritic Ca^{2+} imaging in nonhuman primate brain in vivo," *Sci. Rep.*, vol. 7, no. 1, p. 2873, 2017.
[55] D. L. K. Yamins and J. J. Dicarlo, "Using goal-driven deep learning models to understand sensory cortex," *Nat. Neurosci.*, vol. 19, no. 3, pp. 356–365, 2016.
[56] P. J. Vance *et al.*, "Bioinspired approach to modeling retinal ganglion cells using system identification techniques," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 5, pp. 1796–1808, May 2018.
[57] A. H. Marblestone, G. Wayne, and K. P. Kording, "Toward an integration of deep learning and neuroscience," *Front. Comput. Neurosci.*, vol. 10, p. 94, Sep. 2016.
[58] V. Z. Marmarelis, "Identification of nonlinear biological systems using laguerre expansions of kernels," *Ann. Biomed. Eng.*, vol. 21, no. 6, pp. 573–589, Nov. 1993.
[59] W. X. Li, R. H. Chan, W. Zhang, R. C. Cheung, D. Song, and T. W. Berger, "High-performance and scalable system architecture for the real-time estimation of generalized Laguerre–Volterra MIMO model from neural population spiking activity," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 1, no. 4, pp. 489–501, Dec. 2011.
[60] R. Guyonneau, R. VanRullen, and S. J. Thorpe, "Temporal codes and sparse representations: A key to understanding rapid processing in the visual system," *J. Physiol. Paris*, vol. 98, nos. 4–6, pp. 487–497, Jul.–Nov. 2004.
[61] Q. Yu, R. Yan, H. Tang, K. C. Tan, and H. Li, "A spiking neural network system for robust sequence recognition," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 3, pp. 621–635, Mar. 2016.
[62] H. Demis, K. Dharshan, S. Christopher, and B. Matthew, "Neuroscience-inspired artificial intelligence," *Neuron*, vol. 95, no. 2, pp. 245–258, Jul. 2017.
[63] A. J. Kell, D. L. Yamins, E. N. Shook, S. V. Norman-Haignere, and J. H. McDermott, "A task-optimized neural network replicates human auditory behavior, predicts brain responses, and reveals a cortical processing hierarchy," *Neuron*, vol. 98, no. 3, pp. 630–644, 2018.
[64] T. Gollisch and M. Meister, "Eye smarter than scientists believed: Neural computations in circuits of the retina," *Neuron*, vol. 65, no. 2, pp. 150–164, Jan. 2010.
[65] Y. Li, "Advances in multi-view matrix factorizations," in *Proc. Int. Joint Conf. Neural Netw. Neural Netw. (IJCNN)*, Vancouver, BC, Canada, 2016, pp. 3793–3800.
[66] Q. Yan, Z. Yu, F. Chen, and J. K. Liu, "Revealing structure components of the retina by deep learning networks," in *Proc. NIPS Symp. Interpretable Mach. Learn.*, 2017, pp. 1–6.
[67] Q. Yan *et al.*, "Revealing fine structures of the retinal receptive field by deep-learning networks," *IEEE Trans. Cybern.*, early access, Mar. 10, 2020, doi: 10.1109/TCYB.2020.2972983.
[68] J.-H. Ahn, S. Choi, and J.-H. Oh, "A multiplicative up-propagation algorithm," in *Proc. 21st Int. Conf. Mach. Learn.*, 2004, p. 3.
[69] A. Cichocki and R. Zdunek, "Multilayer nonnegative matrix factorization using projected gradient approaches," *Int. J. Neural Syst.*, vol. 17, no. 06, pp. 431–446, Dec. 2007.
[70] S. Lyu and X. Wang, "On algorithms for sparse multi-factor NMF," in *Advances in Neural Information Processing Systems*. Red Hook, NY, USA: Curran, 2013, pp. 602–610.
[71] H. A. Song and S.-Y. Lee, "Hierarchical representation using NMF," in *Proc. Int. Conf. Neural Inf. Process.*, 2013, pp. 466–473.
[72] G. Trigeorgis, K. Bousmalis, S. Zafeiriou, and B. W. Schuller, "A deep matrix factorization method for learning attribute representations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 3, pp. 417–429, Mar. 2017.
[73] H.-W. Tseng, M. Hong, and Z.-Q. Luo, "Combining sparse NMF with deep neural network: A new classification-based approach for speech enhancement," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Brisbane, QLD, Australia, 2015, pp. 2145–2149.
[74] J. T. Geiger, F. Weninger, J. F. Gemmeke, M. Wöllmer, B. Schuller, and G. Rigoll, "Memory-enhanced neural networks and NMF for robust ASR," *IEEE/ACM Trans. Audio, Speech Language Process.*, vol. 22, no. 6, pp. 1037–1046, Jun. 2014.
[75] D. S. Williamson, Y. Wang, and D. Wang, "Deep neural networks for estimating speech model activations," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Brisbane, QLD, Australia, Apr. 2015, pp. 5113–5117.
[76] H. Zhang, H. Liu, R. Song, and F. Sun, "Nonlinear non-negative matrix factorization using deep learning," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Vancouver, BC, Canada, Jul. 2016, pp. 477–482.
[77] T. G. Kang, K. Kwon, J. W. Shin, and N. S. Kim, "NMF-based target source separation using deep neural network," *IEEE Signal Process. Lett.*, vol. 22, no. 2, pp. 229–233, Feb. 2015.

Shanshan Jia received the master's degree from the Smart City College, Beijing Union University, Beijing, China, in 2018.

She is currently a Research Assistant with the National Engineering Laboratory for Video Technology, School of Electronics Engineering and Computer Science, Peking University, Beijing. Her current research interests include artificial intelligence, brain-inspired computing, and computational neuroscience.

Zhaofei Yu (Member, IEEE) received the B.S. degree from the Hong Shen Honors School, College of Optoelectronic Engineering, Chongqing University, Chongqing, China, in 2012, and the Ph.D. degree from the Automation Department, Tsinghua University, Beijing, China, in 2017.

He is currently an Assistant Professor with the Institute for Artificial Intelligence, Peking University, Beijing. His current research interests include artificial intelligence, brain-inspired computing, and computational neuroscience.

Arno Onken received the Ph.D. degree from the Institute of Software Engineering and Theoretical Computer Science, Technical University of Berlin, Berlin, Germany, in 2009.

He is currently a Lecturer with the Data Science for Life Sciences, University of Edinburgh, Edinburgh, U.K. His area of research includes computational neuroscience and machine learning.

Yonghong Tian (Senior Member, IEEE) received the Ph.D. degree from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2005.

He is currently a Full Professor with the School of Electronics Engineering and Computer Science, Peking University, Beijing, and the Deputy Director of AI Research Center, Peng Cheng Laboratory, Shenzhen, China. He has authored or coauthored over 170 technical articles in refereed journals and conferences. His research interests include computer vision, multimedia big data, and brain-inspired computation.

Prof. Tian was the recipient of the Chinese National Science Foundation for Distinguished Young Scholars in 2018, two National Science and Technology Awards, and Three Ministerial-Level Awards in China, and obtained the 2015 EURASIP Best Paper Award for the *EURASIP Journal on Image and Video Processing*, and the Best Paper Award of IEEE BigMM 2018. He was an Associate Editor of IEEE TRANSACTIONS ON MULTIMEDIA from August 2014 to August 2018, and he has been an Associate Editor of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY and *IEEE Multimedia Magazine* since January 2018, and IEEE ACCESS since January 2017. He coinitiated IEEE International Conference on Multimedia Big Data (BigMM) and served as a TPC Co-Chair of BigMM 2015, and also served as the Technical Program Co-Chair of IEEE ICME 2015, IEEE ISM 2015, and IEEE MIPR 2018/2019, the Organizing Committee Member of ACM Multimedia 2009, IEEE MMSP 2011, IEEE ISCAS 2013, and IEEE ISM 2015/2016. He has been the Steering Member of IEEE ICME since 2018 and IEEE BigMM since 2015, and is a TPC Member of more than ten conferences, such as CVPR, ICCV, ACM KDD, AAAI, ACM MM, and ECCV. He is a member of ACM and a senior member of CIE and CCF.

Tiejun Huang (Senior Member, IEEE) received the bachelor's and master's degrees in computer science from the Wuhan University of Technology, Wuhan, China, in 1992 and 1995, respectively, and the Ph.D. degree in pattern recognition and intelligent systems from Huazhong (Central China) University of Science and Technology, Wuhan, in 1998.

He is currently a Professor with the School of Electronic Engineering and Computer Science, Peking University, Beijing, China, where he is also the Director of the Institute for Digital Media Technology. He has authored or coauthored over 100 peer-reviewed papers and three books. His research area includes video coding, image understanding, digital right management, and digital library.

Prof. Huang is a Member of the Board of Director for Digital Media Project, the Advisory Board of the IEEE Computing Society, and the Board of the Chinese Institute of Electronics.

Jian K. Liu received the Ph.D. degree in mathematics from UCLA, Los Angeles, CA, USA, in 2009.

He is currently a Lecturer with the Centre for Systems Neuroscience, University of Leicester, Leicester, U.K. His area of research includes computational neuroscience and brain-like computation.