Patterns, Volume *4*

Supplemental information

BrainCog: A spiking neural network based, brain-inspired cognitive intelligence engine for brain-inspired AI and brain simulation

Yi Zeng, Dongcheng Zhao, Feifei Zhao, Guobin Shen, Yiting Dong, Enmeng Lu, Qian Zhang, Yinqian Sun, Qian Liang, Yuxuan Zhao, Zhuoya Zhao, Hongjian Fang, Yuwei Wang, Yang Li, Xin Liu, Chengcheng Du, Qingqun Kong, Zizhe Ruan, and Weida Bi

1. SUPPLEMENTAL EXPERIMENTAL PROCEDURES S1: UNSUPERVISED STDP-BASED SNN

Unsupervised learning is an essential cognitive function of the brain. STDP is a widespread rule of synaptic weight modification in the brain. It updates the synaptic weights according to the temporal relationship of the pre-and post-synaptic spikes. In BrainCog, we design an unsupervised STDP-based spiking neural network model [1] as a functional module incorporated to BrainCog. The architecture of unsupervised STDP-based SNN is shown in Fig. S1.

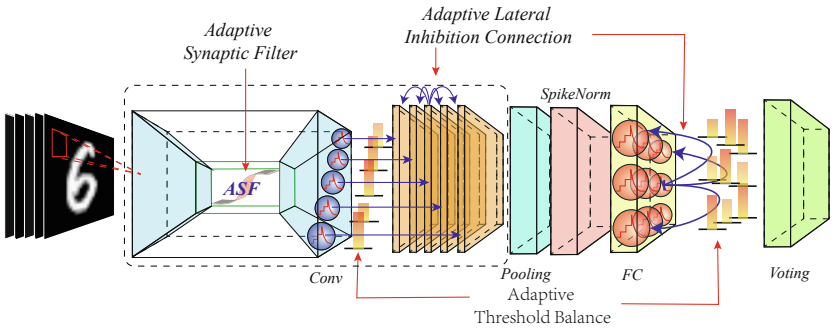

Fig. S1. The framework of the unsupervised STDP-based spiking neural network model.

Unsupervised STDP-based SNN contains various adaptive mechanisms to improve the self-organization ability of the overall network. STP is another synaptic learning mechanism that exists in the brain. Based on the *STDP* and *STP* in BrainCog, we designed an adaptive synaptic filter (ASF) that integrates input currents through nonlinear units, and an adaptive threshold balance (ATB) that dynamically changes the threshold of each neuron to avoid excessively high or low firing rates. The combination of the two controls the firing balance of neurons. We also address the problem of coordinating neurons within a single layer with an adaptive lateral inhibitory connection (ALIC). The mechanism has different connection structures for different input samples. Finally, in order to solve the problem of low efficiency of STDP training, we designed a sample temporal batch STDP. It combines the information between temporal and samples to uniformly update the synaptic weights, as shown by the following formula.

$$\frac{dw_j^{(t)^+}}{dt} = \sum_{m=0}^{N_{batch}} \sum_{n=0}^{T_{batch}} \sum_{f=1}^{N} W(t_i^{f,m} - t_j^{n,m}) \tag{S1}$$

Where $W(x)$ is the function of STDP, N_{batch} is the batchsize of the input, T_{batch} is the batch of time step, N is the number of neurons.

We verified the unsupervised learning ability of BrainCog on MNIST and Fashion-MNIST, achieving 97.9% and 87.0% accuracy, respectively. To the best of our knowledge, these are the state-of-the-art results for unsupervised SNNs based on STDP. To better illustrate the power of our model on small sample training, we tested the model with small samples and found that this model has stronger small sample processing ability than ANN with similar structures, as shown in Tab. S1.

Table S1. The performance of unsupervised SNN compared with ANN on MNIST dataset with different number of training samples.

samples	200	100	50	10
ANN	79.77%	71.40%	68.72%	47.12%
Ours	81.45%	75.44%	72.88%	51.45%
Difference	1.68%	4.04%	4.16%	4.33%

2. SUPPLEMENTAL EXPERIMENTAL PROCEDURES S2: SNN WITH GLOBAL FEEDBACK CONNECTIONS

Most SNNs are based on feedforward structures, while the importance of brain-inspired feedback structures has been ignored. The feedback connections carry out the predictions from the top layer to cooperate with the local plasticity rules to formulate the learning and inference in the brain. In BrainCog, we introduce the global feedback connections and the local differential learning rule [2] in the training of SNNs.

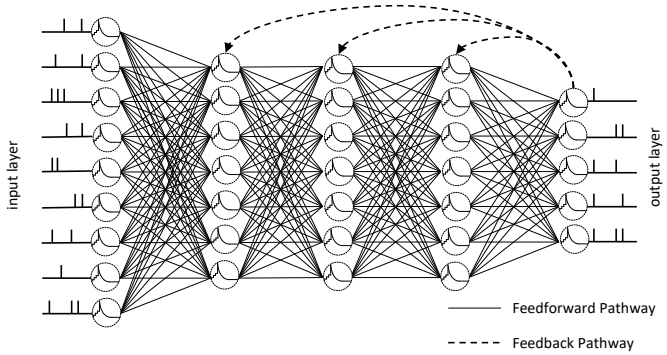

Fig. S2. The feedforward and feedback pathway in the SNN model. The global feedback pathway propagates the target of the hidden layer.

We use the LIF spiking neuron model in the BrainCog to simulate the dynamical process of the membrane potential $V(t)$. We use the mean firing rates S_l of each layer to denote the representation of the l_{th} layer in the forward pathway, and the corresponding target is denoted as $\hat{S}_l$. Here we use the mean squared loss (MSE) as the final loss function. $\hat{S}_{L-1}$ denotes the target of the penultimate layer, and is calculated as shown in Eq. S2, W_{L-1} denotes the forward weight between the $(L-1)_{th}$ layer and the L_{th}. η_t represents the learning rate of the target.

$$\hat{S}_{L-1} = S_{L-1} - \eta_t \Delta S = S_{L-1} - \eta_t W_{L-1}^T (S_{out} - S^T) \tag{S2}$$

The target of the other hidden layer can be obtained through the feedback connections:

$$\hat{S}_l = S_l - G_l (S_{out} - S^T) \tag{S3}$$

By combining the feedforward representation and feedback target, we compute the local MSE loss. We can compute the local update of the parameters with the surrogate gradient. We have conducted experiments on the MNIST and Fashion-MNIST datasets, and achieved 98.23% and 89.68% test accuracy with three hidden layers. Each hidden layer is set with 800 neurons. The details are shown in Fig. S3.

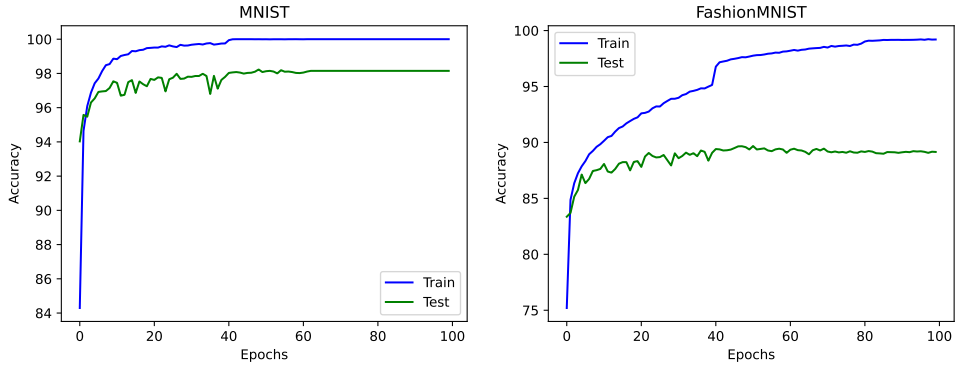

Fig. S3. The test accuracy on MNIST and Fashion-MNIST datasets of the SNNs with global feedback connections.

3. SUPPLEMENTAL EXPERIMENTAL PROCEDURES S3: BIOLOGICAL-BP SNN

For backpropagation-based SNN training methods, BrainCog provides a biologically plausible spatio-temporal adjustment [3], as shown in Fig. S4, which can correctly assign credit according to the contribution of the neuron to the membrane potential at each moment.

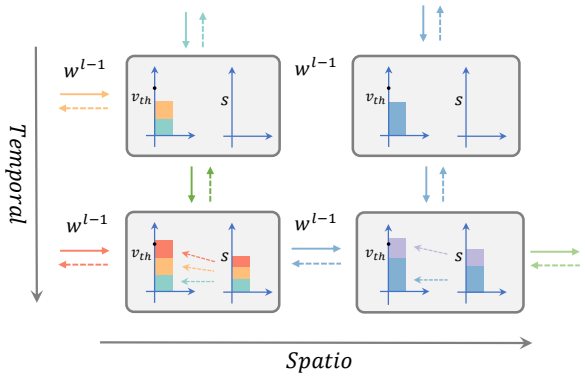

Fig. S4. The forward and backward process of biological BP-based SNNs for BrainCog.

Based on LIF spiking neuron, the direct input encoding strategy, the MSE loss function and the surrogate gradient function supplied in BrainCog, we propose a Biologically Plausible Spatio-Temporal Adjustment (BPSTA) to help the BP algorithm with more reasonable error adjustment in the spatio-temporal dimension [3]. The algorithm realizes the reasonable adjustment of the gradient in the spatial dimension, avoids the unnecessary influence of the neurons that do not generate spikes on the weight update, and extracts more important features. By applying the temporal residual pathway, our algorithm helps the error to be transmitted across multiple spikes and enhances the temporal dependency of the BP-based SNNs. Compared with SNNs and ANNs with the same structure that only used the BP algorithm, our model greatly improves the performance of SNNs on the DVS-CIFAR10 and DVS-Gesture datasets, while also greatly reducing the energy consumption and decay of SNNs, as shown in Tab. S2.

Table S2. The energy efficiency study. The former represents our method, the latter represents the baseline.

Dataset	Accuracy	Firing-rate	EE = $\frac{E_{ANN}}{E_{SNN}}$
MNIST	99.58%/99.42%	0.082/0.183	35.1x/15.7x
N-MNIST	99.61%/99.32%	0.097/0.176	29.6x/16.3x
CIFAR10	92.33%/89.49%	0.108/0.214	26.6x/13.4x
DVS-Gesture	98.26%/93.92%	0.083/0.165	34.6x/17.4x
DVS-CIFAR10	77.76%/71.40%	0.097/0.177	29.5x/16.2x

4. SUPPLEMENTAL EXPERIMENTAL PROCEDURES S4: ANN-SNN CONVERSION WITH BURST SPIKES

Conversion methods can produce spike neural networks with performance comparable to artificial neural networks while minimizing training costs. However, traditional conversion methods normalize weights based on activation values obtained from sampled data, resulting in a trade-off between speed and accuracy for SNNs. Furthermore, previous work commonly uses AvgPooling instead of MaxPooling because selecting neurons with maximum firing rate in the time dimension is challenging.

To address the issue of the residual membrane potential of neurons, we introduce a bursting mechanism, as illustrated in Fig. S5, that allows neurons to send multiple spikes between two

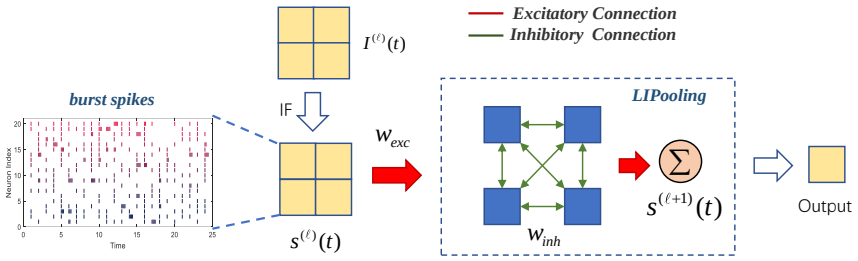

Fig. S5. The forward process of SNN with burst spikes and LIPooling.

time steps depending on the current membrane potential. This approach enables neurons to transmit the remaining information by sending spikes between two time steps, leading to an increased firing rate of the SNN and passing on the remaining membrane potential to the next layer of neurons. Inspired by the lateral inhibition mechanism, we propose LIPooling to transform the maximal ensemble layer, as shown in Fig. S5. Operationally, the winner in LIPooling inhibits other neurons, potentially preventing the neuron with the highest firing rate from spiking due to the suppression of other neurons in the history. LIPooling aggregates the output of all neurons during the simulation, utilizing competition between neurons to obtain an accurate sum (equal to the actual maximum) rather than merely picking winners.

We tested our approach in VGG16 and ResNet16 networks on CIFAR100 datasets, achieving significant improvements in conversion speed and accuracy. Our method achieved a performance of 77.93% on VGG16 using 84 time steps and 80.17% on ResNet20 using 93 time steps, as demonstrated in the Tab. S3.

Table S3. The classification accuracy of converted SNNs on CIFAR100 with BrainCog.

Method	Model	ANN Acc(%)	SNN Best Acc(%)	Time Step
baseline	VGG16	77.75	72.13	256
	ResNet20	80.25	72.01	256
ours	VGG16	77.75	**77.93**	**84**
	ResNet20	80.25	**80.17**	**93**

5. SUPPLEMENTAL EXPERIMENTAL PROCEDURES S5: QUANTUM SUPERPOSITION INSPIRED SNN

In the microscopic size, quantum mechanics dominates the rules of operation of objects which reveals the probabilistic and uncertainties of the world. New technologies based on quantum theory like quantum computation and quantum communication provide an alternative to information processing. Researches show that biological neurons spike at random and the brain can process information with huge parallel potential like quantum computing.

Inspired by this, we propose the Quantum Superposition Inspired Spiking Neural Network (QS-SNN) [4], complementing quantum image (CQIE) method to represent image in the form of quantum superposition state and then transform this state to spike trains with different firing rate and phase. The effort tries to incorporate the quantum superposition mechanism into SNNs as a new form of encoding strategy for BrainCog, and the model finally shows its capability on robustness for learning. The proposed QS-SNN model is tested on background inverted MNIST datasets. Using the quantum superposition encoding module (QSEncoder) provided by BrainCog, the background reverse images are represented by quantum superposition state and then transferred to spike sequences of different frequencies and phases.

Furthermore, we use two-compartment spiking neural networks to process these spike trains. The SNN model is built by multi-compartment neuron integrated in BrainCog, which contains

dendrite and soma computing parts. The dendrite process the input spike signal. The soma integrates the dendritic potential and generates spike output. The proposed QS-SNN model is trained by using dendrite prediction and proximal gradient method.

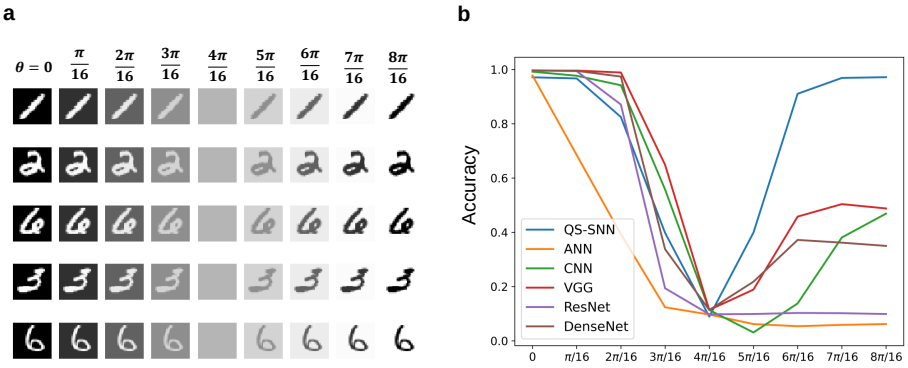

Fig. S6. QS-SNN classifies background reverse MNIST images. (a) Background reverse MNIST images. (b) Comparison of QS-SNN and other models on MNIST dataset.

We compare the QS-SNN model with other convolutional models. The result in Fig. S6 shows that our QS-SNN model overtakes other convolutional neural networks in recognizing background-inverted image tasks. The traditional fully-connected ANNs and convolutional models struggle to handle large changes in image properties, such as background inversion, even if the spatial features of the image remain unchanged. Compared with other models, the QS-SNN model can keep the recognition performance basically unchanged when recognizing the background reversed image. The VGG16, ResNet101 and DenseNet121 models with deep architectures also face drastic performance degradation in the face of image background inversion. In contrast, although the QS-SNN model also suffered from performance degradation when the image information was blurred, the QS-SNN model regained its recognition ability when the digital features of the background-reversed image became clear.

6. SUPPLEMENTAL EXPERIMENTAL PROCEDURES S6: BRAIN-INSPIRED SEQUENCE PRODUCTION SNN

Sequence production is an essential function for AI applications. Components in BrainCog enable the community to build SNN models to handle this task. In this paper, we introduce the brain-inspired symbol sequences production spiking neural network (SPSNN) model [5] that has been incorporated in BrainCog. SPSNN incorporates multiple neuroscience mechanisms including Population Coding, STDP, Reward-Modulated STDP, and Chunking Mechanism, mostly covered and provided by BrainCog. After reinforcement learning, the network can complete the memory of different sequences and production sequences according to different rules.

For Population Coding, this model utilizes populations of neurons to represent different symbols. The whole neural loop of SPSNN is divided into Working Memory Circuit, Reinforcement Learning Circuit, and Motor Neurons [5]. The Working Memory Circuit is mainly responsible for completing the memory of the sequence. The Reinforcement Learning Circuit is responsible for acquiring different rules during the reinforcement learning process. The Motor Neurons can be regarded as the network's output.

In the working process of the model, the Working Memory Circuit and the Reinforcement Learning Circuit cooperate to complete the memory and production of different sequences [5]. It is worth mentioning that with the increase of background noise, the recall accuracy of symbols at different positions in a sequence gradually decreases, and the overall change trend follows the "U-shaped accuracy", which is consistent with experiments in psychology and neuroscience [6]. The results are highly consistent due to the superposition of primacy and recency effects. Our model provides a possible explanation for both effects from a computational perspective.

7. SUPPLEMENTAL EXPERIMENTAL PROCEDURES S7: COMMONSENSE KNOWLEDGE REPRESENTATION GRAPH SNN

BrainCog provides commonsense knowledge representation SNN (CKR-SNN) to explore SNN-based commonsense knowledge representation and reasoning. Inspired by the population coding mechanism, this module encodes the entities and relations of commonsense knowledge graph into different populations of neurons. Via the spiking timing-dependent plasticity (STDP) learning principle, the synaptic connections between neuron populations are formed after guiding the sequential firings of corresponding neuron populations [7]. Neuron populations together constructed the giant graph spiking neural networks, which contain the commonsense knowledge. In this module, CKR-SNN represents a subset of Commonsense Knowledge Graph ConceptNet [8]. After training, CKR-SNN can complete conceptual knowledge generation and other cognitive tasks [7].

8. SUPPLEMENTAL EXPERIMENTAL PROCEDURES S8: CAUSAL REASONING SNN

In BrainCog, we constructed causal reasoning SNN, as an instance to verify the feasibility of spiking neural networks to realize deductive and inductive reasoning. Specifically, Causal Reasoning Spiking Neural Network (CRSNN) module contains a brain-inspired causal reasoning spiking neural network model [9]. This model explores how to encode a static causal graph into a spiking neural network and implement subsequent reasoning based on a spiking neural network.

The CRSNN module adopts the population coding mechanism and uses neuron populations to represent nodes and relationships in the causal graph. Each node indicates different events in the causal graph. By giving current stimulation to different neuron populations in the spiking neural networks and combining the STDP learning rule [10], CRSNN can encode the topology between different nodes in a causal graph into a spiking neural network. Furthermore, according to this network, CRSNN completes the subsequent deductive reasoning tasks. Then, by introducing an external evaluation function, we can grasp the specific reasoning path in the working process of the network according to the firing patterns of the model, which gives the CRSNN more interpretability compared to traditional ANN models [9].

9. SUPPLEMENTAL EXPERIMENTAL PROCEDURES S9: *DROSOPHILA*-INSPIRED DECISION-MAKING SNN

Drosophila decision-making consists of value-based nonlinear decision and perception-based linear decision, where the nonlinear decision could help to amplify the subtle distinction between conflicting cues and make winner-takes-all choices [11]. Based on the *learningrule, Customlinear and node* component in BrainCog, we build *Drosophila* nonlinear and linear decision-making pathways as shown in Fig. S7a-b. The entire model consists of a training phase and a testing phase as same as [12]. In the training phase, a two-layer SNN with LIF neurons is trained by reward-modulated STDP, which combines local STDP synaptic plasticity with global dopamine regulation. The training phase learns the safe pattern (upright-green T) and the punished pattern (inverted-blue T) [12]. Therefore, it is safe for green color and upright T shape factors, while blue color and inverted T shape are dangerous.

Two cues (color and shape) are restructured during the testing phase, requiring linear and nonlinear pathways to make a choice between inverted-green T and upright-blue T, as shown in Fig. S7c. The linear decision directly uses the knowledge acquired during the training phase to make decisions. The nonlinear network models the recurrent loop of the DA-GABA-MB circuit [11, 13, 14]: KC activates the anterior posterior lateral (APL) neurons, which in turn releases GABA transmitter to inhibit the activity of KC. KC also provides mushroom body output neuron (MBON) with exciting input in order to generate behavioral choices. When faced with conflicting cues, the level of DA increases rapidly and produces mutual inhibition with APL, thereby producing a disinhibitory effect on KC. The excitatory connection between DA and MBON also helps speed up decision-making.

To verify the consistency of *drosophila*-inspired decision-making SNN with the conclusions from neuroscience [11], we count the behavior paradigm of our model under different color intensities over a period of time. First, we run the network for 500 steps to count the time t_1 of selecting behavior 1 (avoiding) and the time t_2 of selecting behavior 2 (approaching). Then we calculate prefer index (PI) values under different color intensity: $PI = \frac{|t_1 - t_2|}{|t_1 + t_2|}$. From Fig. S7d, we find that nonlinear circuits could achieve a gain-gating effect to enhance relative salient

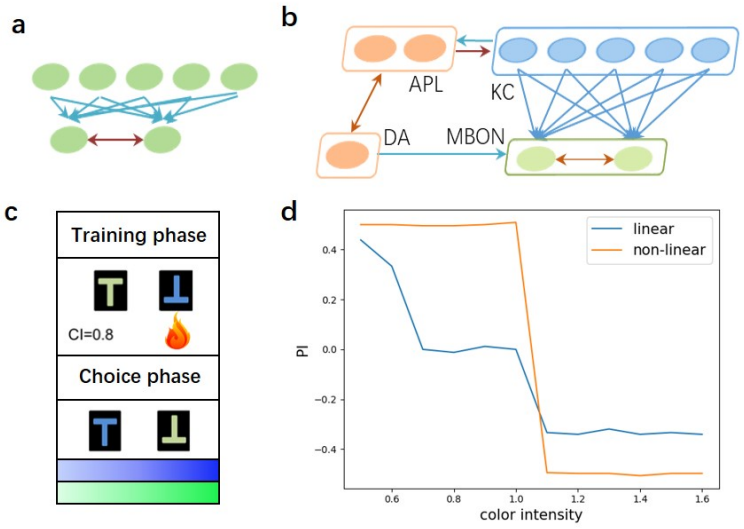

Fig. S7. *Drosophila*-inspired decision-making SNN. (a) Linear Pathway. (b) Nonlinear Pathway. (c) Experiments for training and choice phases. (d) Experimental results of linear and nonlinear networks under the dilemma. The X-axis refers to the color density, and the Y-axis represents the PI values.

cue and suppress less salient cue, thereby displaying the nonlinear sigmoid-shape curve [12]. However, the linear network couldn't amplify the difference between conflicting cues, thus making an ambiguous choice (linear-shape curve) [12]. This work proves that drawing on the neural mechanism and structure of the nonlinear and linear decision-making of the *Drosophila* brain, the brain-inspired computational model implemented by BrainCog could obtain consistent conclusions with the *Drosophila* biological experiment [11].

10. SUPPLEMENTAL EXPERIMENTAL PROCEDURES S10: EMOTION-DEPENDENT ROBOTIC MUSIC COMPOSITION AND PLAYING.

Stylistic Music Composition. Inspired by the mechanisms of the human brain, we introduce a brain-inspired spiking neural network (SNN) that contains several collaborative subnetworks to learn and generate music melodies. The architecture is shown in Fig. S8. Inspired by the function of brain areas which is related to music learning and human creativity (Fig. S8a), the SNN contains three subsystems: 1) The knowledge subsystem is responsible for encoding and learning the background information of a musical piece. As is shown in Fig. S8b, the genre cluster stores the genre of classical music pieces, the composer cluster represents the names of famous composers, and the title cluster mainly encodes the titles of musical pieces. Connections between layers are dynamically generated and updated. Furthermore, neurons in the upper layers project to those in the lower layers. This hierarchical structure helps the model organize and learn complex background information. 2) The emotion subsystem is a neural cluster that simply expresses four emotional types. 3) The sequential memory subsystem, consisting of the pitch subnetwork and duration subnetwork, is critical for music learning and generating. The pitch subnetwork mainly encodes musical pitches and learns the ordered relationships among musical notes. The duration subnetwork represents the temporal intervals of musical notes. Both subnetworks are composed of numerous minicolumns in which neurons with the same preferences are organized in vertical structures. Synaptic connections with different transmission delays are dynamically generated and modulated during the learning process [15].

Learning Process. First, neurons in these subnetworks receive the background knowledge and musical notes as the external stimulus. Since neurons have their preferences for this information, we use the Gaussian filters to simulate these specific neural properties and transform these external stimuli into the input currents. Then neurons that preferred the input information emit spikes continuously. Meanwhile, synaptic connections between these activated neurons are generated and the synaptic weights are updated by the STDP learning rule [15, 16].

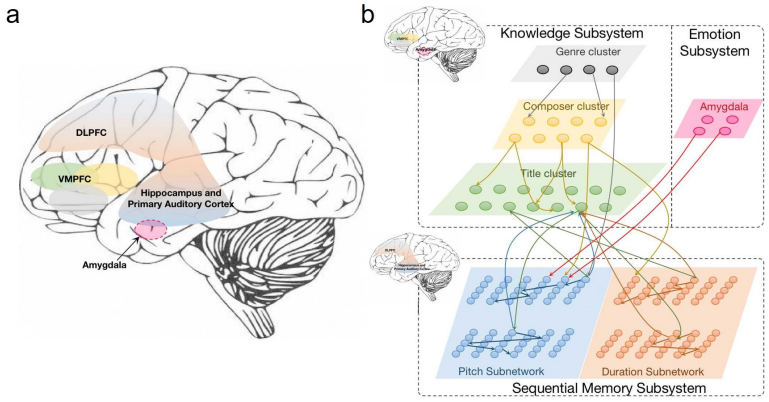

Fig. S8. The architecture of the music learning and composition model, (a) shows the brain areas related to the music memory and stylistic composition, and (b) describes the architecture of the entire spiking neural network.

Generating Process. After the learning process, the model can generate melodies with different styles. Three conditions should be specified first: music style, beginning notes, and the melody length (number of the notes) to be generated. Music style refers to the emotional types, musical genre styles or the composers' characteristics. Based on the learning process, neurons that encode initial information are activated and fire. Since the synaptic weights are modulated during the learning process, different neurons in pitch and duration subnetworks are triggered and emit spikes.

Experiments. A classical piano dataset, including 331 western classical piano works recorded in MIDI format, is employed here to train the model. The musical works in this dataset are labeled with different styles: four emotional types (joyful, sorrowful, peaceful and passionate), three genre types (Baroque, Classical and Romantic), and twenty-five composers' styles). We also invited 41 human listeners to evaluate the qualities of generated melodies and score the generated melodies. The score ranges from 1 (which means the piece has no apparent features) to 5 (the melody style is quite similar to the specified type). In fact, the evaluation results depend on subjective preferences and social experiences, the experiments are relatively complex. The final experiments indicate that the generated melodies with different genres get high scores. However, those with various composers' styles depend on the musicians' characteristics. Similarly, the generated melodies with joyful, sorrowful and peaceful types get high scores [16].

REFERENCES

1. Dong, Y., Zhao, D., Li, Y., and Zeng, Y. (2022). An unsupervised spiking neural network inspired by biologically plausible learning rules and connections. arXiv preprint arXiv:2207.02727. 10.48550/arXiv.2207.02727.
2. Zhao, D., Zeng, Y., Zhang, T., Shi, M., and Zhao, F. (2020). GLSNN: A multi-layer spiking neural network based on global feedback alignment and local STDP plasticity. Front. Comput. Neurosci. *14*. 10.3389/fncom.2020.576841.
3. Shen, G., Zhao, D., and Zeng, Y. (2022). Backpropagation with biologically plausible spatiotemporal adjustment for training deep spiking neural networks. Patterns *3*, 100522. 10.1016/j.patter.2022.100522.
4. Sun, Y., Zeng, Y., and Zhang, T. (2021). Quantum superposition inspired spiking neural network. Iscience *24*, 102880. 10.1016/j.isci.2021.102880.
5. Fang, H., Zeng, Y., and Zhao, F. (2021). Brain inspired sequences production by spiking neural networks with reward-modulated STDP. Front. Comput. Neurosci. *15*, 612041. 10.3389/fncom.2021.612041.
6. Jiang, X., Long, T., Cao, W., Li, J., Dehaene, S., and Wang, L. (2018). Production of supra-regular spatial sequences by macaque monkeys. Current Biology *28*, 1851–1859. 10.1016/j.cub.2018.04.047.
7. Fang, H., Zeng, Y., Tang, J., Wang, Y., Liang, Y., and Liu, X. (2022). Brain-inspired graph

spiking neural networks for commonsense knowledge representation and reasoning. arXiv preprint arXiv:2207.05561. 10.48550/arXiv.2207.05561.

8. Speer, R., Chin, J., and Havasi, C. Conceptnet 5.5: An open multilingual graph of general knowledge. In: *Proceedings of the AAAI conference on artificial intelligence* vol. 31 (2017):10.1609/aaai.v31i1.11164.
9. Fang, H., and Zeng, Y. A brain-inspired causal reasoning model based on spiking neural networks. In: *2021 International Joint Conference on Neural Networks (IJCNN)*. IEEE (2021):(1–5). 10.1109/IJCNN52387.2021.9534102.
10. Dan, Y., and Poo, M.-m. (2004). Spike timing-dependent plasticity of neural circuits. Neuron *44*, 23–30. 10.1016/j.neuron.2004.09.007.
11. Tang, S., and Guo, A. (2001). Choice behavior of drosophila facing contradictory visual cues. Science *294*, 1543–1547. 10.1126/science.1058237.
12. Zhao, F., Zeng, Y., Guo, A., Su, H., and Xu, B. (2020). A neural algorithm for drosophila linear and nonlinear decision-making. Sci. Rep *10*, 1–16. 10.1038/s41598-020-75628-y.
13. Zhang, K., Guo, J. Z., Peng, Y., Xi, W., and Guo, A. (2007). Dopamine-mushroom body circuit regulates saliency-based decision-making in drosophila. Science *316*, 1901–1904. 10.1126/science.1137357.
14. Zhou, M., Chen, N., Tian, J., Zeng, J., Zhang, Y., Zhang, X., Guo, J., Sun, J., Li, Y., Guo, A. et al. (2019). Suppression of gabaergic neurons through d2-like receptor secures efficient conditioning in drosophila aversive olfactory learning. PNAS *116*, 5118–5125. 10.1073/pnas.1812342116.
15. Liang, Q., Zeng, Y., and Xu, B. (2020). Temporal-sequential learning with a brain-inspired spiking neural network and its application to musical memory. Front. Comput. Neurosci *14*, 51. 10.3389/fncom.2020.00051.
16. Liang, Q., and Zeng, Y. (2021). Stylistic composition of melodies based on a brain-inspired spiking neural network. Front. Syst. Neurosci. *15*, 639484. 10.3389/fnsys.2021.639484.