Gamma and Beta Band Oscillation in Working Memory given Sequential or Concurrent Multiple Items: A Spiking Network Model

Shukuo Zhao, Jinpu Zhou, Yongwen Zhang, Da-Hui Wang

1. School of Systems Science, Beijing Normal University, 19th Xinjiekouwai Street, Haidian, Beijing 100875, China
2. State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, 19th Xinjiekouwai Street, Haidian, Beijing 100875, China
3. Beijing Key Laboratory of Brain Imaging and Connectomics, Beijing Normal University, Beijing 100875, China
4. These authors contributed equally

DH designed research, SZ, JZ, YZ and DH performed research, DH wrote the paper.

DaHui Wang: wangdh@bnu.edu.cn

This work was supported by NSFC under Grant No. 32171094. Authors thank Prof. XiaoJing Wang for his
helpful suggestions.
14. Conflict of Interest: No
A. No (State ‘Authors report no conflict of interest’)

2
Gamma and Beta Band Oscillation in Working Memory given Sequential or Concurrent Multiple Items: A Spiking Network Model

Abstract

Working memory can maintain sequential and concurrent information, and the load enhances the gamma-band oscillation during the delay period. To provide a unified account for these phenomena in working memory, we investigated a continuous network model consisting of pyramidal cells, high-threshold fast-spiking interneurons (FS), and low-threshold non-fast-spiking interneurons (nFS) for working memory of sequential and concurrent directional cues. Our model exhibits the gamma (30-100Hz) and beta (10-30Hz) band oscillation during the retention of both concurrent cues and sequential cues. We found that the beta oscillation results from the interaction between pyramidal cells and nFS, whereas the gamma oscillation emerges from the interaction between pyramidal cells and FS due to the strong excitation elicited by cue presentation, shedding light on the mechanism underlying the enhancement of gamma power in many cognitive executions.

Significance Statement

We constructed a spiking network to perform working memory tasks with sequentially or concurrently presented items. The model exhibits the coexistence of beta (10-30Hz) and gamma (30-100Hz) band oscillations during the delay period. We found that gamma and beta-band oscillations recruit separate neural circuits. The low-threshold nFS
neurons are involved in the beta-band oscillation, whereas the high-threshold FS neurons are involved in the gamma-band oscillation. Our results shed light on the well-known phenomenon that cognitive tasks enhance gamma-band oscillations.

**Keywords**

Working memory; Oscillation; Sequential/concurrent items; Spiking network;

**Introduction**

Working memory (WM), the ability to actively maintain and manipulate information in the absence of stimuli, plays a crucial role in cognitive function and executive control of behavior (Baddeley, 2003). The information maintained in WM can enter the brain concurrently or sequentially, meaning that WM can not only hold multiple pieces of information arriving simultaneously as in visuospatial WM tasks (Bays and Husain, 2008; Zhang and Luck, 2008; Gorgoraptis et al., 2011), but also information presented sequentially, as in visual (Gorgoraptis et al., 2010) or speech processing (Cowan, 2001).

The neural implementation of information maintenance is not well understood. The oscillatory model proposes that the information of one item is represented in WM by the reactivation of neurons in the gamma cycle within nested gamma-theta oscillations, mediated by a slow after-depolarization (ADP) with a time constant that should match the theta oscillation (Lisman and Idiart, 1995; Jesen and Lisman, 1998). The dynamic model posits that memory is maintained by item-specific patterns of synaptic plasticity and that neurons exhibit a nonstationary and short-lived attractor activity, in which only one memory representation can be active at a time, but successive reactivations of neuronal pools memorize different items (Mi et al., 2017; Lundqvist et al., 2018 and 2016; Mongill et
The persistent activity model asserts that persistent activity of neurons induced by N-methyl-D-aspartate receptor (NMDAR)-mediated recurrent synaptic current encodes the corresponding items during the delay period (Compte et al., 2000; Wei et al., 2012; Edin et al., 2009). The ability of WM to maintain concurrent and sequential information challenges these models. On the one hand, the oscillatory model and dynamic models can use oscillatory activity to maintain sequential information, but it is difficult for them to represent concurrent information (Constantinidis et al., 2018); on the other hand, the persistent models can manipulate concurrent information but cannot address sequential stimuli and often focus on the persistent activity without oscillation. Experiments have shown that gamma-band oscillations are involved in WM maintenance, in particular, that WM load enhances the gamma-band oscillations but suppresses the alpha-band oscillations. Here, we aim to construct a biophysically plausible network model to implement the storage of concurrent and sequential information and to investigate how WM load alters oscillations in the network.

The biological brain consists of microcircuits with three types of neurons that play key roles in generating oscillations: pyramidal cells, low-threshold nFS, and high-threshold neuronal FS (Chen et al, 2017; Veit et al, 2017; Domhof and Tiesinga, 2021; Hahn et al, 2022). The pyramidal cells are excitatory, while the FS and nFS are GABAergic inhibitory. Pyramidal cells synapse on nFS and FS. In turn, nFS neurons synapse back onto the dendrites of the pyramidal cell, while FS neurons synapse back onto the soma of the pyramidal cell, creating a dynamic feedback loop that regulates excitatory and inhibitory activities. FS and nFS also inhibit each other, resulting in a competitive relationship.
rest, the nFS is more active, inhibiting the FS and forming a beta oscillation with the pyramidal cells. When a perceptual stimulus is presented, the FS cells are activated, inhibiting the nFS cells and creating a gamma oscillation with the pyramidal cells (Chen et al., 2017, Hahn et al., 2022). Given the importance of oscillations in higher cognitive functions, we aim to explore the role of oscillations and microcircuitry in working memory.

We proposed a spiking neural network to implement WM with concurrent and sequential directional information (Gorgoraptis et al., 2011). The network consists of two-compartment pyramidal cells, nFS, and FS. These cells and compartments are interconnected as in the biological brain. The cells are uniformly arranged in a ring according to their preferred direction (Compte et al., 2000). When activated, pyramidal cells activate pyramidal cells with similar preferences as well as nearby FS cells and nFS cells on the ring. FS and nFS inhibit neighbouring neurons on the ring. We found that, regardless of whether the directions are presented concurrently or sequentially, the network elicits corresponding localized activities that persist throughout the delay period. The persistent activity suggests that the model can successfully maintain multiple stimuli presented simultaneously or sequentially in working memory. Furthermore, prior to cue presentation, the interaction between pyramidal cells and nFS dominates the activity of the network and causes a low-band oscillation (10-30Hz); the cue presentation induces strong excitation and recruits FS into the network, and the interaction between pyramidal cells and FS dominates the activity of the network, enhancing the gamma oscillation (35-100 Hz), which persists throughout the delay period.
Methods and Models

Model architecture

Our model has 4096 excitatory pyramidal cells, 512 FS cells, and 512 nFS cells. We have two reasons to choose the number of neurons. Firstly, the ratio of the number of excitatory neurons over that of inhibitory neurons is approximately 4:1 (Braitenberg and Schütz, 1991). Thus, we chose 4906 pyramidal neurons and 1024 inhibitory interneurons (FS+nFS, 1024). Secondly, a network size of $2^n$ neurons is desired for the Fast Fourier Transform (FFT) which recursively divides the input data into smaller subsets, and conducts Fourier transform computations on subsets. These neurons are evenly distributed in a ring configuration and are connected by AMPA, NMDA, and GABA synapses, forming an interconnected spiking network. According to our model, the strength of connections between neurons and the angular differences in their distribution across the ring configuration follows a Gaussian distribution. Neurons that are closer in angle on the ring have stronger connections. The footprint of the connections can be described as equation (1).

\[
\begin{aligned}
W(\theta_i - \theta_j) &= J^- + (J^+ - J^-) \exp \left( -\frac{(\theta_i - \theta_j)^2}{2\sigma^2} \right) \quad \theta_i \in E \\
W(\theta_i - \theta_j) &= \exp \left( -\frac{(\theta_i - \theta_j)^2}{4\pi^2\sigma^2} \right) \quad \theta_i \in I
\end{aligned}
\]  

(1)

\[J^+_{s-s} = 6.5, J^+_{s-fs} = 35, J^+_{s-nfs} = 70.\] By normalizing the footprint \(\frac{1}{360} \int_0^{360} W(\theta_i - \theta_j) d\theta_j = 1\), we obtain \(J^-\) in equation (1). The standard deviations of the Gaussian distributions for connectivity footprint in our model were as follows:
From pyramidal cells soma to pyramidal cells soma ($\sigma_{s\rightarrow s}$): 12.76 degrees

- From pyramidal cells soma to FS cell ($\sigma_{s\rightarrow fs}$): 7.05 degrees

- From pyramidal cells soma to nFS cell ($\sigma_{s\rightarrow nfs}$): 1.41 degrees

- From FS cell to pyramidal cells soma ($\sigma_{fs\rightarrow s}$): 8.46 degrees

- From FS cell to nFS cell ($\sigma_{fs\rightarrow nfs}$): 1.41 degrees

- From nFS cell to pyramidal cells dendrite ($\sigma_{nfs\rightarrow d}$): 5.02 degrees

The value of parameters $J^+_{s\rightarrow s}$ and $\sigma_{s\rightarrow s}$ used in our study were adapted from Compte et al (2000), Wei et al (2012). We further introduced structured excitatory projections from pyramidal cells to FS and nFS and inhibitory projections from FS to pyramidal soma and nFS cells or from nFS to pyramidal dendrite, referencing local circuit properties described by Wang et al. (2004, 2018). FS belongs to the large basket cell that projects to broader region, while nFS has narrow dendritic and axonal arbors and serves local inhibition. Generally, excitatory neurons project to a larger radius (around 200μmphysiologically; as shown by Fitzpatrick, 1985) compared to inhibitory interneurons (around 100μm; as demonstrated by Spiro, 1999) based on biological experiments. Nevertheless, the lateral innervation width of interneurons may be similar to that of excitatory neurons. It is because excitatory neurons project to a larger radius, activating inhibitory neurons within that
radius to locally innervate inhibition (as assumed by Compte et al). The synaptic weights were calculated by the production of footprint and the maximum conductance as follows:

\[ g_{ij} = W(\theta_i - \theta_j)G_{\text{syn}} \]  

(2)

The \( G_{\text{syn}} \) parameters for these connections are specified in Table 1. To simplify the calculations, we have assumed that the remaining inhibitory connections in our model (e.g. inhibitory connection from nFS to FS or recurrent connection from FS to FS) follow a uniform distribution (A simple diagram is shown in Fig.1).

Neuron models

Pyramidal cells and inhibitory cells follow the leaky integrate-and-fire model (Tuckwell, 1988). The neurons generate spikes when their membrane potentials reach threshold \( V_{th} = -50\text{mV} \). And their membrane potential is reset to \( V_{\text{reset}} = -60\text{mV} \) immediately after a spike for \( \tau_s = 2\text{ms} \); \( \tau_d = 2\text{ms} \); \( \tau_{fs} = 1\text{ms} \); \( \tau_{nfs} = 1\text{ms} \) as refractory periods. The Pyramidal cells have two compartments (Wang, 1998): soma(s) and dendrite(d). The membrane potential of the somatic compartment \( (V_s) \) and dendrite compartment \( (V_d) \) of pyramidal cells obey equations (3-4).

\[
C_m \frac{dV_s}{dt} = -I_{Ca} - g_{SL}(V_s - V_0) - \frac{g_{c1}(V_s - V_d)}{p_1} - I_{\text{syn}} 
\]

(3)

\[
C_m \frac{dV_d}{dt} = -g_{dl}(V_d - V_0) - \frac{g_{c2}(V_d - V_s)}{p_2} - I_{\text{syn}} 
\]

(4)

\( V_0 = -70\text{ mV} \), \( C_m = 0.5 \text{nF} \), \( g_{SL} = 0.025 \text{nS} \), \( g_{dl} = 0.025 \text{nS} \), \( g_{c1} = 0.25 \text{nS} \) (Troyer and Miller, 1997; Wei, 2012). The parameters \( p1 \) and \( p2 \) characterize the difference of current
intensity resulting from the identical current in soma and dendrite due to the difference in membrane surface area. Specifically, $p_1$ represents the ratio of somatic area to the total neuronal area, and $p_2$ represents the ratio of dendritic area to the total area (Wang, 1998). In the current model, $p_1$ is set to 0.6, and $p_2$ is set to $1-p_1$, which is 0.4. FS and nFS cells contain single compartments, and their membrane potentials obey equation (5-6). Please note that when the neuron (soma) fires, the membrane potential of the dendrite component is not reset to -60mV. The dendrite component generates its own spikes and resets the membrane potential after each spike.

$$C_{mf,s} \frac{dV_{fs}}{dt} = -g_{fsL}(V_{fs} - V_{fs0}) - I_{syn} \quad (5)$$

$$C_{m,ps} \frac{dV_{nfs}}{dt} = -I_{Ca} - g_{nfsL}(V_{nfs} - V_{nfs0}) - I_{syn} \quad (6)$$

$$C_{mf,s} = 0.2 \ nF, C_{m,ns} = 0.8 \ nF, g_{fsL} = 0.020 \ nS, g_{nfsL} = 0.016 \ nS, V_{fs0} = -86 \ mV, V_{nfs0} = -76 \ mV \ (Troyer \ and \ Miller, \ 1997; \ Wei, \ 2012). \ The \ term \ "Isyn" \ represents \ the \ total \ synaptic \ current. \ The \ ion \ current \ I_{Ca} \ in \ equations \ (3) \ and \ (6) \ represents \ a \ high-threshold \ calcium \ current \ following \ equations \ (7-8; \ Wang,1998).$$

$$I_{Ca} = g_{Ca}m_{\infty}^2(V - V_{Ca}) \quad (7)$$

$$m_{\infty}^2(V) = \frac{1}{1+e^{-\frac{(V+20)/9}{}}} \quad (8)$$

$$g_{Ca-s} = 0.0015 \ nS, g_{Ca-nfs} = 0.001 \ nS. \ V \ stands \ for \ membrane \ potential \ (equals \ V_s \ or \ V_{nfs}).$$
Synapse models

Synaptic currents are mediated by AMPA, NMDA, and GABA transmissions. Three types of synaptic currents follow the equations:

\[ I_{i,\text{AMPA}} = (V_i - V_{\text{AMPA}}) \sum_j g_{ij,\text{AMPA}} s_{j,\text{AMPA}} \] (9)

\[ I_{i,\text{NMDA}} = (V_i - V_{\text{NMDA}}) \sum_j \frac{g_{ij,\text{NMDA}}}{1 + [\text{Mg}^2+] \exp(-0.026V_i/3.57)} \] (10)

\[ I_{i,\text{GABA}} = (V_i - V_{\text{GABA}}) \sum_j g_{ij,\text{GABA}} s_{j,\text{GABA}} \] (11)

The conductance "g" in our calculations primarily follows the equation (2), where \( V_{\text{AMPA}} = V_{\text{NMDA}} = 0 \text{mV}, V_{\text{GABA}} = -70 \text{mV}, \) \([\text{Mg}^2+]\) in equation (10) equals 1 mM (Jahr and Stevens, 1990). The gating variable \( s \) for AMPA or GABA \( (s_{j,\text{AMPA}} \text{ and } s_{j,\text{GABA}} \text{ in equation (9) and (11)}) \) follows the dynamic in equation (12), and the dynamic of \( s \) \( (s_{j,\text{NMDA}} \text{ for NMDA was modelled as equation (13-14).})\)

\[ \frac{ds(t)}{dt} = -s(t) \frac{1}{\tau_s} + \sum_k \delta(t - t_k) \] (12)

\[ \frac{ds(t)}{dt} = -\frac{s(t)}{\tau_{s,\text{NMDA}}} + \alpha_s x(t) (1 - s(t)) \] (13)

\[ \frac{dx(t)}{dt} = -\frac{x(t)}{\tau_{x,\text{NMDA}}} + \sum_k \delta(t - t_k) \] (14)
t_k is the spike sequence of the neuron. The rising time constant for NMDA (τ_x,NMDA) is 2ms, and the decay time constant τ_s is 2ms, 10ms, and 100ms for AMPA, GABA, and NMDA, respectively (Wei, 2012).

Synapses from Pyrs to nFS mediated by NMDAR were facilitated in short term and follow the equation:

\[
\frac{du}{dt} = \frac{U-u}{\tau_{F,NMDA}} + k(1-u)\delta(t-t_k)
\]

(15)

U is set to 0.4 in our simulation (Mongillo et al, 2008).

Simulation protocol

We simulated two types of WM tasks: concurrent-cue tasks (CCTs) and sequential-cue tasks (SCTs), similar to the experiments in (Gorgoraptis et al., 2011). In CCTs, cues are presented to neurons concurrently; each pyramidal cell indexed by θ received a current I_{ext}(θ) during the concurrent cue presentation period from 500 to 700ms (Fig. 1B).

I_{ext}(θ) = \sum_{k=1}^{n} I_0 \frac{1}{\sqrt{2\pi}\sigma_s} \exp \left( -\frac{(\theta - \theta_{in,k})^2}{2\sigma_s^2} \right)

(16)

where θ_{in,k} is the kth direction and n is the number of directions in the cue array. I_0 = 0.5 nA, σ_s = 2 degree.
In SCTs, pyramidal cells received multiple currents sequentially. Each current can be formulated as:

$$I_{\text{ext}}(\theta) = \frac{I_0}{\sqrt{2\pi}\sigma_s} \exp\left(-\frac{(\theta - \theta_{\text{in},k})^2}{2\sigma_s^2}\right) \quad (17)$$

The $k$th current is applied at $500 + 300(k - 1)$ to $700 + 300(k - 1)$ ms (Fig. 1B) during the presentation of the $k$th cue (Wei et al., 2012).

**Noise implementation**

Noise is applied to pyramidal cells and inhibitory cells via AMPA synapses:

$$I_{\text{noise}} = (V_i - V_{\text{AMPA}})g_{\text{noise}}s_{\text{noise}}$$

$$g_{\text{noise-\text{pyr}}} = 0.0039 \ nS, \ g_{\text{noise-\text{pv}}} = 0.0019 \ nS, \ g_{\text{noise-\text{cb}}} = 0.0022 \ nS,$$

and $s_{\text{noise}}$ follows:

$$\frac{ds(t)}{dt} = -\frac{s(t)}{\tau_s} + \eta(t) \quad (19)$$

$$\tau_s = 2 \ \text{ms}, \ \text{and} \ \eta(t) \ \text{is a Poisson process}. $$

$$p(\eta(t) = k) = \frac{e^{-\lambda dt}((\lambda dt)^k)}{k!}, \ k = 0,1... \quad (20)$$

where $\lambda$ is the arrival rate of noise. We set $\lambda = 1kHz$, which is equivalent to 1000 synaptic input at 1 Hz (Wei et al., 2012).
Starting parameters

To establish a more realistic and stable initial condition, we carried out a recording procedure to obtain a consistent value of multiple variables for each type of neurons. The following variables were utilized: the membrane potential ("$V_m$"), the gating variables ("s" and "x") for synapses, the time of the last firing activity ("LastTimeEachFired" minus 250 ms), and the AMPA synaptic current ("$I_{AMP A}$") for the pyramidal soma (over a period of 10 ms), and these recorded values were then used as initial conditions for the algorithms. Initially, we performed a model run with constant initialization values of -51 mV for the membrane potential and 1e-34 for the "s" parameter without these recordings.

Numeric integration and code accessibility

The model was programmed in Matlab code and the gating variables were integrated using a second-order Runge-Kutta (RK2) algorithm.

$$x_{n+1} = x_n + \frac{1}{2} dt (K_1 + K_2)$$

$$K_1 = f(x_n, t_n)$$

$$K_2 = f(x_n + dtK_1, t_n + dt)$$

The membrane potential is integrated with Euler scheme as equation 3-6. This different choice in the integration method is because changes in gating variables during neuronal firing are more complex and usually non-linear, whereas changes in the subthreshold membrane potential are approximately linear. We conducted a convergence test on two neurons with AMPA connections and provided a 2 mV constant stimulus to the
presynaptic neuron for 2000 ms to examine the difference of spike time of postsynaptic neuron between Euler-RK2 and RK2 methods. We found that the spike time errors generated by the two methods are almost negligible for sufficiently small dt (such as 0.02, as in the experiments). By variation of the timesteps, we found the convergence order of integration method $R = \frac{\log_2 |e_{\text{new}}| \cdot |\Delta t_{\text{new}}| - \log_2 |e_{\text{old}}| \cdot |\Delta t_{\text{old}}|}{\log_2 |\Delta e_{\text{new}}| \cdot |\Delta t_{\text{new}}| - \log_2 |\Delta e_{\text{old}}| \cdot |\Delta t_{\text{old}}|}$ equals 2.95, indicating that the error decays to 0.129 of its original error if the timestep is decreased by half.

This approach reduces the computational cost without compromising accuracy. The code/software described in the paper is freely available online at https://github.com/scientific-lab/Gamma-and-Beta-Band-Oscillation-in-Working-Memory.

We performed our research using GPU(K80) with Ubuntu 18.04.6 LTS, Microsoft Surface Pro 5 with Windows 10 operating system, and MacBook Pro 11,1 with Windows 10 operating system.

**Data analysis**

Firing rate calculation: We counted the spikes of a certain type of neuron along the time in a single trial and approximated the firing rate by applying a one-dimensional filter with a window size of 4 ms to the spike train.

Local field potential and spectrograms: We approximated the local field potential (LFP) using the total synaptic currents onto pyramidal cells following the method proposed by Mazzoni et al.(2015):
LFP(t) = $\sum_{\text{pyr}}[I_{\text{AMPA}}(t - \tau_{\text{AMPA}}) + I_{\text{NMDA}}(t) - 1.65 \times I_{\text{GABA}}(t)]$

where $\tau_{\text{AMPA}} = 6\text{ms}$. We applied a one-dimension filter with a window size of 4ms to the LFP and calculated the power spectrogram of LFP using continuous wavelet transform with Morlet wavelet. The background energy of our model demonstrates the 1/f characteristic for high frequency oscillation (refer to Miller et al., 2006; Wang, 2010). We first computed the power-frequency correlation of our model without external input shown in Fig. 3A. We observed that the power for low-frequency oscillation remains almost constant, but decays in accordance with a power-law trend. The power-frequency data was smoothed utilizing an average of 250 points and represented by the red line to signify the background power of oscillation. We subtracted the value of the red line from the simulated power. Figure 3B illustrates the power of LFP, which has been normalised by the red curve in Panel A. This normalisation is applied in subsequent spectral analysis to obtain the normalised power.

Burst rate: A burst in a particular frequency is defined as an interval when spectral power exceeds twice the standard deviation above the average value of that frequency and lasts for at least three cycles (Lundqvist et al., 2016). After we located the bursts in a single trail, we applied the same process to multiple trails and calculated the frequency of the bursts in the beta (10-30 Hz) or gamma (35-100Hz) band over time. Similar to how we extracted the firing rate, we filtered the calculated burst frequency with a one-dimensional Gaussian filter, with a window size of 60ms, to approximate the burst rate for a certain frequency range.
Statistical comparison: In our study, we employed independent sample t-tests and ANOVA to compare between groups after conducting tests for homogeneity of variance and making necessary corrections. We also provided the effect size and confidence interval for each significant test result. When necessary, we conducted multiple comparison analyses to offer comprehensive results. All statistical analyses were performed using Matlab. Our model generated substantial data from multiple simulations with high power and narrow confidence intervals. Based on the results from the online power and sample size calculator (https://www.gigacalculator.com/calculators/), the power for all statistical tests were close to 100%.

Results

The continuous spiking network can hold concurrent and sequential cues

We considered a spiking neural network consisting of one population of 4098 pyramidal neurons, one population of 512 FS, and one population of 512 nFS. Pyramidal cells are uniformly distributed on a ring according to their preferred cue angles (Fig. 1A) (Compte et al., 2000), mimicking the columnar organization of the monkey PFC (Goldman-Rakic, 1995; Rao et al., 1999; Constantinidis et al., 2001). Using this continuous spiking network model, we investigated how WM manipulates sequential and concurrent cues. We first presented one direction to the network for a brief period and then withdraw the cue to explore the spatiotemporal pattern of neural activity elicited by the cue presentation. As shown in Fig. 2A, pyramidal cells showed relatively sparse and regular discharges in the absence of stimulus from 0-500 ms. At 500-700 ms, the stimulation elicited an intense
firing in the cued direction. When the cue was withdrawn, this intense discharge in the preferred direction persisted and showed oscillatory activity throughout the 700-2000 ms delay period. We call this sustained oscillatory activity in the cued direction "oscillatory activity bump" and believe it maintains information about the cue. Pyramidal cells in other directions were unaffected during and after the cue presentation and continued to discharge sparsely and regularly. The inhibitory interneurons nFS (blue in Fig. 2B) and FS (red in Fig. 2B) showed opposite trends. In the pre-stimulus period (0-500 ms), the nFS discharged as sparsely and regularly as the pyramidal cell, whereas the FS barely discharged. Within 500-700 ms of stimulus presentation, the nFS stopped firing in the preferred direction, whereas the FS began to fire strongly and intensely in the preferred direction. This firing pattern continued throughout the 700-2000 ms delay period after stimulus withdrawal. The stimulus had little effect on the non-preferred nFS and FS, and their spikes remained sparse or absent. Fig. 4B provides additional information about the average membrane potential of the cells. It also shows examples of pyramidal cells, FS, and nFS in the preferred and non-preferred directions. We then presented two or three directional cues to the network sequentially or concurrently. Given concurrent cues, the network simultaneously elicits two or three distinct oscillatory activity bumps that persist throughout the delay period (Fig. 2C and D), and each distinct activity bump maintains the information of one cued direction. For sequential cues, the network elicits distinct oscillatory activity bumps one after another (Fig. 2E and F). The later presented cue evokes a new persistent oscillatory activity bump and does not disrupt the persistent oscillatory activity bump elicited by the previous cue. These results show that our continuous network can store not only concurrent information but also sequential
Labor division of interneurons in Gamma and Beta oscillation

To investigate the oscillatory behavior in WM, we calculated the firing rate of pyramidal cells and the LFP given one directional cue in a single trial (Fig. 3). The population firing rate and LFP oscillate before and after the cue presentation and throughout the delay period (Fig. 3C and E). The spectrogram of the LFP shows that the network oscillates in the beta band before the cue presentation but oscillates in the beta and gamma bands during the delay period (Fig.3D and F). Fig. 4 and 5 provide an overview of the average membrane potential (Fig. 4A) and gating variables (Fig. 5A) during a single trial for different neurons or compartments. In addition, these figures provide examples of individual neurons in both preferred and non-preferred directions (Fig. 4B, 5B). Overall, the mean membrane potential of pyramidal cells (Mpre=-55.054, Maft=-54.918, t=-43.399, p<0.001, d=0.430, cl=[-0.142,-0.130], Fig. 4A, where Mpre is the average membrane potential prior to cue presentation and Maft represents that after the cue) and FS (Mpre=-67.212, Maft=-67.117, t=-24.981, p<0.001, d=0.224, cl=[-0.103,-0.088], Fig. 4A) increased, and that of nFS cells decreased after stimulus presentation (Mpre=-53.999, Maft=-54.598, t=542.997, p<0.001, d=3.464, cl=[0.596,0.601], Fig. 4A). We calculated the total synaptic currents received by the interneurons (Fig.6). We found that the cue presentation increases the synaptic input to FS close to the presented direction (Mpre=0.057, Maft=0.330, t=-271.328, p<0.001, d=1.360, cl=[-0.274,-0.270], Fig. 6A) and activates these FS (Fig. 2B) during the cue presentation and throughout the delay period. Cue presentation has little effect on the synaptic input to FS far away from the cued
direction (M_pre=0.049, M_ait=0.041, t=13.920, p<0.001, d=0.092, cl=[0.007,0.009], Fig.6B) and cannot activate them (Fig. 2B). We found that synaptic currents input to nFS near the cued direction decrease due to the inhibition from active FS during cue presentation and the delay period (M_pre=0.081, M_ait=-0.215, t=188.834, p<0.001, d=0.820, cl=[0.293,0.299], Fig.6C), and these nFS are suppressed (Fig. 2B). However, the input to nFS far away from the cued direction does not change much (M_pre=0.079, M_ait=0.080, t=-3.175, p=0.002, d=0.021, cl=[-0.002,-0.000], Fig.6D). In summary, before the cue presentation, weak excitation is insufficient to activate FS, and the interaction between pyramidal cells and nFS dominates the network activity, resulting in beta band oscillation. The cue presentation induces strong excitation and activates FS, and the interaction between pyramidal cells and FS leads to gamma band oscillation in the network. Therefore, these results show a division of labor between FS and nFS in gamma- and the beta-band oscillations, as shown in Fig. 6E.

**Gamma power increases in the cued direction**

Experiments (Pesaran et al. 2002) showed that the increase in gamma band power was sustained during the delay period in the preferred direction, while gamma band power did not change in the anti-preferred direction. We calculated the LFP based on the pyramidal cells close to the cued direction, which is the preferred direction, and the pyramidal cells far away from the cued direction, which is the anti-preferred direction. We found that in the preferred direction, the LFP mainly consists of beta band oscillation before the cue presentation, and a strong gamma band oscillation appears after the cue presentation in a single trial (Fig.7A) or averaged over trials (Fig.7C). The power of gamma band
increases significantly (Mpre=23.942, Maft=26.616, t=-80.835, p<0.001, d=0.784, cl=[-2.739,-2.609]), but the power of beta band decreases (Mpre=37.157, Maft=35.381, t=76.696, p<0.001, d=0.737, cl=[1.731,1.822], Fig.7E). There is little increase in gamma band oscillation (Mpre=26.370, Maft=26.741, t=-54.777, p<0.001, d=0.479, cl=[-0.385,-0.358], Fig.7B) for the anti-preferred direction, and the beta band oscillation is not decreased (Mpre=37.752, Maft=37.450, t=47.799, p<0.001, d=0.348, cl=[0.289,0.314], Fig.7F). Note that there was an early increase in gamma and beta band power in Fig.7E and F. We think this was an artifact of the cwtft algorithm. CWTFT calculates the power of a signal by locally integrating the multiplication of signals with shiftable symmetric wavelets of different window sizes and center frequencies. In the same wavelet window, wavelet analysis is isotropic, thus unable to distinguish components before and after stimulus presentation. When the signal frequency undergoes significant changes within a certain window, the window that contains the changing signal cannot differentiate between the pre-change and post-change components, showing an earlier increase in power (Similar to Kaiju et al., 2017). In short, cue presentation increases the power of the gamma oscillation and decreases beta oscillation in the preferred direction (Fig.7E and F), which are consistent with the experimental observations (Pesaran et al. 2002).

**WM load enhances the Gamma power during the delay period**

We further analyzed the spectrogram of LFP given concurrent or sequential cues. The average spectrogram of LFP over 500 trials exhibits several characteristics. First, as previously shown, the network exhibits beta band oscillation before the cue presentation,
and the beta power tend to decrease after cue presentation (two concurrent cues: $M_{pre}=107.469$, $M_{ aft}=102.083$, $t=39.865$, $p<0.001$, $d=0.412$, $cl=[5.121,5.651]$; three concurrent cues: $M_{pre}=117.394$, $M_{ aft}=108.605$, $t=25.929$, $p<0.001$, $d=0.272$, $cl=[8.124,9.453]$; two sequential cues: $M_{pre}=100.995$, $M_{ aft}=97.639$, $t=83.571$, $p<0.001$, $d=0.549$, $cl=[3.277,3.435]$; three sequential cues: $M_{pre}=98.238$, $M_{ aft}=98.990$, $t=-15.451$, $p<0.001$, $d=0.085$, $cl=[-0.847,-0.657]$). Second, the network elicits higher gamma-band oscillation (around 60-70Hz) upon cue presentation (two concurrent cues: $M_{pre}=62.763$, $M_{ aft}=68.745$, $t=-43.986$, $p<0.001$, $d=0.450$, $cl=[-6.249,-5.715]$; three concurrent cues: $M_{pre}=68.370$, $M_{ aft}=75.840$, $t=-21.950$, $p<0.001$, $d=0.233$, $cl=[-8.137,-6.803]$; two sequential cues: $M_{pre}=58.800$, $M_{ aft}=66.297$, $t=-133.050$, $p<0.001$, $d=1.081$, $cl=[-7.607,-7.386]$; three sequential cues: $M_{pre}=58.202$, $M_{ aft}=70.287$, $t=-193.409$, $p<0.001$, $d=1.299$, $cl=[-12.207,-11.963]$) and gamma-band oscillation maintains during the delay period.

Third, the gamma power remains almost constant during the delay period for concurrent cues (Fig. 8A and B). In contrast, the gamma power increases along with the arrival of new cues in sequential situations and the power is maintained throughout the delay period (two sequential cues: $F=13057.29$, $p<0.001$; $\eta^2=0.225$, three sequential cues: $F=24175.79$, $p<0.001$, $\eta^2=0.446$, with multiple comparisons $p<0.001$. Fig. 8C and D).

We show the average gamma power in sequential trials in Fig 8E and F. The trajectory of gamma power is similar to the results shown in Fig. 6 of Lundqvist et al. (2016), albeit with a slight distinction: the gamma band exhibited an increase following the presentation of the second stimulus prior to the presentation of the third stimulus.
We performed a Fourier analysis on the LFP given concurrent cues, and the results are consistent with the spectrogram analysis. Gamma power increases with the WM load (F=71.06, p<0.001, Fig.9). Three cues lead to the highest gamma power and one cue results in the lowest gamma power. The mean of the gamma power are M1cue = 0.493, M2cue = 1.060, and M3cue = 2.357 for 1, 2, and 3 cues, respectively (multiple comparisons p12=0.002, p13<0.001, p23<0.001, Fig.9). The dependence of gamma power on WM load can be explained by network activity. Strong excitation induced by cue presentation activates high-threshold FS and the interaction between pyramidal cells and FS results in gamma oscillation. One cue elicits one distinct activity bump with gamma oscillation and more cues elicit more distinct activity bumps with gamma oscillation (Fig. 2), suggesting that more neurons participate in gamma oscillation and result in higher gamma power.

**Brief bursts of narrow-band Gamma and Beta oscillation in a single trial**

Although the raster plots of spiking in Fig. 2 show oscillatory activity, the single spectrogram of LFP (Fig. 3B and Fig. 7A, B) shows a different scenario: gamma and beta oscillations occurred irregularly in the form of short bursts that are narrow and variable in frequency. This observation is similar to a previous monkey experiment (Lundqvist et al., 2016). Here, we identified the gamma and beta bursts using the algorithm proposed in the reference (Lundqvist et al., 2016). We identified each burst in 500 trials with sequential cues and plotted each burst as a point in Fig. 10A. We also calculated the beta and gamma burst rates (Fig. 10B). The beta burst rate decreases slightly after cue presentation (Mpre=0.0207, M aft=0.0196, t=21.08, p<0.001, Fig.7F) and remains almost constant during the delay period (Fig.10B). The gamma burst rate increases when the
cue is presented (Mpre=0.0570, Maft=0.0685, t=-48.93, p<0.001, Fig.7F) and increases with increasing WM load (F=229.69, p<0.001, multiple comparisons p<0.001, Fig.10B). During the delay period, the gamma burst shows a slightly increasing trend, which is consistent with the experimental observation in Fig. 8B of Lundqvist et al. (2016).

Discussion
We have developed a working memory model that utilizes a continuous attractor structure based on the bistable excitatory-inhibitory microcircuit consisting of FS, nFS, and pyramidal cells in PFC. The main contributions of our work are threefold. First, the proposed model can maintain the information of sequentially presented cues and concurrently presented cues. Previous oscillatory and dynamic models collapse in the presence of concurrent information because these models rely on the oscillatory phase or timing to encode the information (Constantinidis et al., 2018). The persistent activity model has not been used to manipulate sequential information (Edin et al., 2009; Wei et al., 2012). Second, we identified the mechanism underlying the enhancement of gamma power in the WM task. Cue presentation elicits strong excitatory recurrent synaptic currents that activate high-threshold FS. The interaction between pyramidal cells and FS leads to gamma oscillations and enhances gamma power during the delay period. Third, we demonstrated the division of labor between high-threshold FS and nFS for oscillation in the WM. The interaction between nFS and pyramidal cells leads to beta oscillation, while the interaction between FS and pyramidal cells leads to gamma oscillation.
We showed that the transient and sparkling gamma and beta burst in a single spectrogram of LFP can be approximated from the persistent oscillatory activity of the network (Fig.3B), and the gamma burst rate increases with the WM load (Fig.8B). The discrete and sparkling gamma and beta burst in a single LFP spectrogram has been considered as evidence that WM is manifested through transient or discrete oscillatory dynamics, rather than sustained firing (Lundqvist et al. 2016). However, our results indicate that transient and sparkling gamma and beta bursts in a single LFP spectrogram may indeed result from persistent oscillatory firing. Therefore, our results favor the theory of persistent activity underlying WM.

In this work, we have demonstrated the division of labor between FS and nFS in generating gamma and beta oscillation. The interplay between high-threshold FS and pyramidal cells results in the gamma oscillation, whereas the interaction between nFS and pyramidal cells leads to the beta oscillation. Before the cue presentation in WM tasks, the weak recurrent excitation from spontaneous discharges of pyramidal cells can activate low-threshold nFS but not the high-threshold FS, resulting in beta oscillation. Cue presentation evokes strong activity of pyramidal cells and strong recurrent excitation to activate FS. As a result, the network exhibits gamma oscillation. This observation is consistent with experimental results that optogenetic activation of FS induces gamma rhythm (Cardin et al., 2009) and the spiking of somatostatin (SOM) and parvalbumin (PV) cells differentially correlates with beta and gamma oscillations and activation of PV cells enhances gamma oscillation in V1 area (Chen et al., 2017).
We computationally identified that gamma power can be enhanced by recruiting more high-threshold FS due to the stronger excitation evoked by the cue presentation in WM. This mechanism may be common to other cognitive processes, such as attention and sensory information processing in perceptual decision-making. It has been extensively observed in cognitive processes such as WM (Jensen et al., 2007; Yamamoto et al., 2014), attention (Engel et al., 2001; Fries et al., 2001; Jensen et al., 2007; Fries 2009; Gregoriou et al., 2009; Kim et al., 2016), and sensory information processing in perceptual decision making (Frien et al., 2000; Siegel and König, 2003; Kayser and König, 2004; Henrie and Shapley, 2005; Liu and Newsome, 2006; Berens et al., 2008). Studies had observed that the selected neuronal population increased their firing rate when attention was directed to their receptive field (Luck et al., 1997; McAdams and Maunsell, 1999; Bichot et al., 2015; Thiele et al., 2016), implying strong excitation and potentially recruiting more FS neurons.

Gamma oscillations in the sensory cortex are often considered as a proxy for the encoding of sensory evidence during perceptual decision-making due to visual gamma-band activity and its dependence on stimulus strength and various stimulus features (Frien et al., 2000; Siegel and König, 2003; Kayser and König, 2004; Henrie and Shapley, 2005; Liu and Newsome, 2006; Berens et al., 2008). Based on these, we hypothesize that sensory input would evoke a stronger excitation in the sensory cortex, which could recruit more FS and increase the gamma power.

Worth noting that the synapses from pyramidal cells to the nFS are dynamic synapses with short-term potentiation (STP). STP plays an essential role in the modulation of gamma and beta bursts. In the biological brain, there is short-term synaptic facilitation
between pyramidal cells and nFS and short-term synaptic depression between pyramidal
cells and FS. These short-term plasticities allow the brain to respond differently to
afferents of different durations. Brief currents cause gamma oscillations, while prolonged
currents cause beta oscillations. Based on this mechanism, these observations are
reproduced in Feng et al. 2021 (Feng et al., 2021). In our simulation, we only briefly
present the stimulus to the model, resulting in oscillations at the gamma frequency, and
short-term synaptic facilitation between pyramidal cells and nFS helps control excitability
in our model. The synapses equipped with STP can temporarily increase inhibition by
increasing the effective excitation from the pyramidal cells to the nFS and expanding the
range of parameters to keep excitation and inhibition in balance. Too much excitation can
lead to the merging of different information (Wei et al, 2012) or false memories (Edin et
al, 2008), while too much inhibition can prevent the network from maintaining information.
Suppose the synapses from pyramidal cells to the nFS are fixed without STP; the network
is prone to produce spurious bursts of activity due to too much excitation or to forget the
information due to too much inhibition. At present, we have not considered the effects of
prolonged stimulus. It would be interesting to investigate how the brain adapts and
responds to prolonged stimulus in the future.
Figure legend:

Figure 1. Model structure and simulation protocol. A) The network consists of two-compartment pyramidal cells, fast-spiking (FS) and non-fast-spiking (nFS) interneurons. Pyramidal and inhibitory cells are uniformly placed on a ring and labeled by their preferred directions. B) Simulation protocol. In concurrent-cue tasks, an array of cues is presented to the network from 500ms to 700ms, followed by a delay period of 1300 ms (upper panel). In sequential-cue tasks, cues were presented one after another with an interval of 100ms, and each cue lasted for 200 ms.

Figure 2. Raster plots of spiking of neurons in the network. A) Raster plots of spiking of pyramidal cells given one cue. B) Raster plots of spiking of low-threshold nFS (blue) and high-threshold FS (red) given one cue. C) Two direction cues S1+S2 were concurrently presented to the network during the cue period (500-700 ms). D) Three directional cues S1+S2+S3 were concurrently presented to the network during the cue period. E) Two directional cues (S1 and S2) were sequentially presented to the network. F) Three directional cues (S1, S2, and S3) were sequentially presented to the network. The horizontal axis is the time in ms, and the vertical axis is the preferred direction of the pyramidal cell.

Figure 3. The oscillation in the network. A) The power of raw simulated LFP of the network. The black dots plot the power against frequency on a logarithmic scale for LFP and has
a power-law distribution. The red curve represents the 250-point average smooth curve of the log-scaled power and the log-scaled frequency. B) The power of LFP normalized by the red curve in panel A. C) Example of the population firing rate of pyramidal cells in a single trial. D) Example of the spectrogram of normalized LFP shown in panel C based on Morlet wavelet. E) Approximated LFP in single trial as in panel A. F) Zoom-in on the approximated local field potential around the time of gamma and beta burst as seen in the spectrogram in panel B. Blue curve shows the local field potential filtered at 42Hz and red curves shows local field potential filtered at 17Hz.

Figure 4. Voltage traces from neurons. A) Average membrane potential for pyramidal dendrites (top left), somas (top right), FS cells (bottom left) and nFS cells (bottom right) in response to one cue. Dotted lines indicate the presence of the stimulus. B) Example of single-cell membrane potential for pyramidal dendrites (top left), somas (top right), FS cells (bottom left) and nFS cells (bottom right) in the preferred and non-preferred direction in response to a cue. The voltage trace of neurons in the preferred direction is represented by the black line and those in the non-preferred direction by the blue line.

Figure 5. Gating variables form neurons. A) Average AMPA, NMDA, and GABA gating variables associated with presynaptic pyramidal dendrites, somas, FS cells and nFS cells in response to one cue. Dashed lines indicate the presence of the stimulus. B) Example of single cell gating variable in response to one cue. The gating variables of individual neurons are shown at the same positions as the corresponding average gating variables.
in A). The gating variables of neurons in the preferred direction are represented by the black line and those in the non-preferred direction by the blue line.

Figure 6. Labor division of FS and nFS interneuron. A) The total synaptic input to FS close to the cued direction. B) The total synaptic input to FS far away from the cued direction. C) The total synaptic input to nFS close to the cued direction. D) The total synaptic input to nFS far away from the cued direction. E) The schematic diagram of labor division of interneurons in the gamma band and beta band oscillation. Interaction between pyramidal cell and nFS results in beta band oscillation, while interaction between pyramidal cell and FS leads to gamma oscillation.

Figure 7. Oscillation power in the preferred direction and anti-preferred direction. A) The spectrogram for LFP approximated from the pyramidal cells preferring the cued direction in a single trial. B) The spectrogram for LFP approximated from the pyramidal cells far away from the cued direction in a single trial. C) The average of spectrogram for preferred direction over 900 trials. D) The average of the spectrogram for anti-preferred direction over 900 trials. E) The average normalized power (beta band: 10-30Hz, gamma band: 35-100Hz) for the preferred direction. The red line denotes the gamma power, and the blue line denotes the beta power. F) The average normalized power for non-preferred direction. The red line denotes the gamma power, and the blue line denotes the beta power.

Figure 8. The average spectrogram of LFP given concurrent or sequential cues over 500 trials (A-D) and the average normalized gamma power for sequential cue conditions (E-
Figure 9. The normalized power of local field potential. Gamma power induced by two concurrent cues (red line) is higher than that induced by one cue (blue line) but lower than that resulting from three concurrent cues (black line).

Figure 10. The oscillatory bursts. A) The raster plot shows the occurrence of oscillatory bursts in 500 trials of simulation. B) The rate of beta and gamma bursts.

Table legend:

| Table 1. $g_{syn}$ Parameters for Soma, Dendrite, FS, and nFS Synaptic Connections |
|-----------------------------------------------|----------|----------|----------|----------|
| presynaptic neurons                          | Pyr-soma | FS       | nFS      |          |
|                                              | AMPA     | NMDA     | GABA     | GABA     |
| post synaptic neurons                        |          |          |          |          |
| Pyr-soma                                     | 0.0001125| 0.0000750| 0.0003420| -        |
| Pyr-dendrite                                 | -        | -        | -        | 0.0018210|
| FS                                           | 0.0001600| 0.0000600| 0.0000600| 0.0000900|
| nFS                                          | 0.0002400| 0.0002400| 0.0105000| -        |

This table specifies the parameters used for calculating the AMPA, NMDA, and GABA synaptic connections between soma, dendrite, FS, and nFS (Adapted from Compte et al 2000 and Wei et al 2012).

References


