Check for updates

Information representation in an oscillating neural field model modulated by working memory signals

William H. Nesse[1]*, Kelsey L. Clark[2] and Behrad Noudoost[2]

[1]Department of Mathematics, University of Utah, Salt Lake City, UT, United States, [2]Department of Ophthalmology, University of Utah, Salt Lake City, UT, United States

We study how stimulus information can be represented in the dynamical signatures of an oscillatory model of neural activity—a model whose activity can be modulated by input akin to signals involved in working memory (WM). We developed a neural field model, tuned near an oscillatory instability, in which the WM-like input can modulate the frequency and amplitude of the oscillation. Our neural field model has a spatial-like domain in which an input that preferentially targets a point—a stimulus feature—on the domain will induce feature-specific activity changes. These feature-specific activity changes affect both the mean rate of spikes and the relative timing of spiking activity to the global field oscillation—the phase of the spiking activity. From these two dynamical signatures, we define both a spike rate code and an oscillatory phase code. We assess the performance of these two codes to discriminate stimulus features using an information-theoretic analysis. We show that global WM input modulations can enhance phase code discrimination while simultaneously reducing rate code discrimination. Moreover, we find that the phase code performance is roughly two orders of magnitude larger than that of the rate code defined for the same model solutions. The results of our model have applications to sensory areas of the brain, to which prefrontal areas send inputs reflecting the content of WM. These WM inputs to sensory areas have been established to induce oscillatory changes similar to our model. Our model results suggest a mechanism by which WM signals may enhance sensory information represented in oscillatory activity beyond the comparatively weak representations based on the mean rate activity.

KEYWORDS

neural coding, phase, information theory, working memory, computational model

Introduction

Signals from cortical areas reflecting cognitive states, including working memory (WM), have been shown to modulate responses to incoming sensory stimuli (Desimone and Duncan, 1995; Humphreys et al., 1998; Lee et al., 2005; Mitchell et al., 2007; Fries, 2009, 2015; Churchland et al., 2010; Bosman et al., 2012; Vinck et al., 2013; van Kerkoerle et al., 2014; Womelsdorf and Everling, 2015; Engel et al., 2016; Michalareas et al., 2016; Moore and Zirnsak, 2017). Several modeling studies have explored these effects (Brunel and Wang, 2001; Ardid et al., 2007; Lakatos et al., 2008; Kopell et al., 2011; Lee et al., 2013; Kanashiro et al., 2017). Firing rates of sensory cortical neurons are not modulated, or modulated only weakly, by WM alone (Lee et al., 2005; Zaksas and Pasternak, 2006; Mendoza-Halliday et al., 2014). In the presence of a sensory signal, firing rates of neurons in these sensory areas are

enhanced when the stimulus in their receptive field is the target of WM (Merrikhi et al., 2018). However, these observed enhancements to firing rate are modest, and it is unknown how such small changes could account for the cognitive enhancements that WM produces (Bahmani et al., 2018). In addition to changes in mean activity, WM also induces changes to oscillations in local field potentials (LFPs) (Siegel et al., 2008; Liebe et al., 2012; Daume et al., 2017) and the timing of spikes relative to these oscillations (Lee et al., 2005; Bahmani et al., 2018). These oscillatory effects suggest a potential role of oscillatory phase of spikes in encoding information about a stimulus.

Prior results by our group (Bahmani et al., 2018) have shown that in a WM task, when a receptive field of an extrastriate visual neuron is in a remembered location, recordings show increased LFP oscillatory power and peak-power frequency in the α-β band (8–25 Hz), increasing spike phase locking (SPL) to the LFP oscillation. Stimuli present in the visual field also have been shown to increase LFP oscillatory power and peak-power frequency as a function of stimulus contrast (Roberts et al., 2013). Notably, cortical beta oscillations exhibit a diffuse relationship with individual spiking cells and show weak-to-moderate phase locking to the local field. Modeling work by our group (Nesse et al., 2021a) has also identified how WM signals to recurrent spiking networks incorporating excitatory (e) and inhibitory (i) cells could modulate emergent network oscillations and SPL, consistent with experimental findings in the study by Bahmani et al. (2018). While this prior modeling study identified the role of e- and i-cells coordinating these oscillatory effects, it suggest two further questions. First, is there a dynamical mechanism that gives rise to such oscillations and their modulation by WM signals? Second, does stimulus-driven changes to the timing of spike activity relative to the LFP phase serve as an effective means of encoding sensory information, and can it be enhanced by WM input?

To address these aforementioned questions, we devised an oscillatory mean field type or neural mass model (Wilson and Cowan, 1972) that exhibits oscillations. We find that the constellation of oscillatory effects due to WM (top down) input and stimulus (bottom up) input observed in previous studies (Roberts et al., 2013; Bahmani et al., 2018) are consistent with a dynamical system tuned near a supercritical Hopf instability, in which external input—either WM-like or sensory-like input– serves as a bifurcation parameter. Such instabilities give rise to low amplitude oscillations whose amplitude can be modulated by input, but whose amplitude is continuous with zero at the point of bifurcation.

Based on the hypothesis of supercritical Hopf instability, we identified a parameter regime in our model that exhibits such a tuning and verified that our neural mass model shows good correspondence with a large-N spiking model network. Using this parameter tuning, we then expanded our neural mass model to a neural field model by including a spatial-like domain, representing a stimulus feature parameter (Ben-Yishai et al., 1997; Shriki et al., 2003; Coombes et al., 2014). The activity of the neural field model can vary its responses to stimulus input features and can represent stimulus information in the mean rate response–the rate code—and the timing of activity relative to the phase of the global oscillation– the phase code. Using these two encodings, we then measured stimulus information coding performance of bottom-up inputs.

We establish that the phase-based coding performance is *positively* enhanced by increased WM-like signals that increase oscillation amplitude and frequency, whereas the rate coding performance is reduced by the same WM input. Additionally, we show phase coding performance as a function of stimulus contrast that exhibits increased gain in the presence of WM input. Conversely, rate coding performance shows reduced stimulus contrast gain in the presence of WM input.

Furthermore, we find that phase code performance is roughly two orders of magnitude larger than the rate code. This large coding disparity can be explained by the nature of relaxation oscillations of local e-i network dynamics. The oscillations arise from the interplay between excitatory rising phase activity being curtailed by the delayed response of the inhibitory downward phase of the cycle. By definition, mean firing rate representations average over this oscillating cycle. As such, the adaptive inhibitory response partly obscures the information contained stimulus-induced activity, whereas the time course of e-i cycle of activity retains the underlying stimulus information more effectively (Nesse et al., 2021b).

Methods

Neural mass model

To assess sensory coding performance, we used the neural field model which was constructed from a spatial continuum of neural mass models connected through a spatial kernel. As such, we must first construct the neural mass model. We developed a neural mass model inspired by Wilson and Cowan (1972) representing a population-average activity of a recurrently connected network (i.e., a neural mass), in which $u(t)$ and $v(t)$ represent the mean synaptic activity of e- and i- cells, respectively.

$$\tau_e \frac{du}{dt} = -u + f_\sigma(w_{ee}u - w_{ei}v + I_e),$$
$$\tau_i \frac{dv}{dt} = -v + f_\sigma(w_{ie}u - w_{ii}v + I_i), \tag{1}$$

where the function $f_\sigma(I)$ in Eq. 1 is a "F-I" function representing the firing rate of a neural mass as a function of the average input current I. The F-I curve is parameterized by $\sigma > 0$ that controls the overall slope or firing rate gain, resulting from input current changes. The F-I curve is derived as the inverse of the mean first passage time of a leaky integrate and fire (LIF) model driven by input current I and uncorrelated noise with strength σ (see below). The w_{jk} is synaptic weight between and within cell types, and I_j is an offset current for $j, k = e, i$. We chose $\tau_e = 5$ ms and $\tau_i = 15$ ms monoexponential decay timescales, consistent with common estimates of effective excitatory AMPA receptor and inhibitory GABA$_A$ decay timescales, respectively (Kapur et al., 1997; Xiang et al., 1998). We have chosen synaptic weights where the inter-population connections—w_{ei}, w_{ie}—were roughly double that of the intra-population connections—w_{ee}, w_{ii} (see Table 1). This choice is consistent with the findings that e- and i-cell populations exhibit strong inhibition-mediated

stabilization of the network (Sanzeni et al., 2020; Sadeh and Clopath, 2021). A complete list of parameter values is presented in Table 1.

The neural mass model equilibria and stability of equilibria were computed over a wide parameter scan of input I and firing rate gain σ parameters. Both e- and i-cells received input controlled by a single I parameter, where $I_e = I$ and $I_i = I - 0.5$, where the inhibitory offset "-0.5" in I_i reflects higher spike thresholds for inhibitory cells that are typically used in neural mass models (Angelucci et al., 2017). Equilibria (u^*, v^*) were solved using a standard iterative Newton's method $(u_{n+1}, v_{n+1}) = (u_n, v_n) - \mathbf{J}^{-1}((u_n, v_n))\mathbf{F}((u_n, v_n))$, where $\mathbf{F}$ is the differential equation (DE) right hand side and $\mathbf{J}$ is the Jacobian matrix linearization of the DE. Stability of the equilibrium was obtained by computing the eigenvalues of the linearized model (i.e., the Jacobian matrix). Where Hopf instabilities were observed (see Results), we computed the first Lyapunov coefficient ℓ_1 to determine the type of Hopf bifurcation—supercritical ($\ell_1 < 0$) or subcritical ($\ell_1 > 0$)—according to the standard formula (Kuznetsov, 1998).

F-I curve derivation

The firing rate as a function of input current curve $f_\sigma(I)$ was derived from a LIF model stochastic differential equation (SDE) and subjected to white noise $\sigma \xi(t)$ in which $\langle \xi(t) \rangle = 0$ and $\langle \xi(t)\xi(t + \tau) \rangle = \delta_{t,t+\tau}$ (Stein, 1965) are given by

$$C\frac{dV}{dt} = g(V_l - V) + I + \sigma\xi(t), \tag{2}$$

with $V(t)$ being the membrane voltage (mV); $C = 1$ the membrane capacitance in micro-Farads; $g = 1/15$ the leak conductance in micro Siemens; $V_l = -65$ mV is reversal leak potential; and I is the input current. With these parameters, a $I = 1$ nA of non-fluctuating input will induce the membrane to reach the spike threshold voltage $V_t = -50$. With the addition of fluctuating white noise input $\sigma\xi(t)$ and assuming a reflecting boundary for very large negative voltages, the mean first passage time $T(V; I, \sigma)$ for our LIF model (Eq. 2) to reach spiking threshold, while initially starting at $V(0) = V_r$, solves the following differential equation with absorbing spike threshold boundary condition (Gardiner, 2009):

$$(I + gV_l - gV)\frac{\partial T}{\partial V} + \frac{\sigma^2}{2}\frac{\partial^2 T}{\partial V^2} = -1, \tag{3}$$
$$T(V_t : I, \sigma) = 0.$$

The solution of (Eq. 3) is expressed as follows:

$$\psi(V) = e^{\frac{2V(I+gV_l)-gV^2}{\sigma^2}}, \tag{4}$$

$$T(V; I, \sigma) = \frac{2}{\sigma^2}\int_V^{V_t} \psi(s)^{-1} \int_{-\infty}^s \psi(\xi)d\xi \ ds. \tag{5}$$

The mean firing rate $f_\sigma(I)$ of the LIF model is then given by the inverse of the first passage time

$$f_\sigma(I) \equiv \frac{1,000}{T(V_l; I, \sigma)}, \tag{6}$$

expressed in units of Hertz.

The above mean first passage time is the spiking time of the LIF model in the diffusion limit of temporally uncorrelated synaptic δ-impulses (see Stein, 1965; Sanzeni et al., 2020).

Obtaining the mean field model from a spiking LIF network

Considering a network of LIF model neurons as in Eq. 2, consisting of $j = 1, 2, \ldots, N_e$ and $\ell = 1, 2, \ldots, N_i$ numbers of e- and i-cell constituents, each cell emits spikes at times $t_{j,k}$ and $t_{\ell,k}$ for the spike timing index k, respectively. For simplicity, we assume that the network is all-to-all connected with uniform synaptic strength, in which the input current I to each e- and i-cell models is

$$I_e = I_{0e} + w_{ee}\hat{u}(t) + w_{ei}\hat{v}(t),$$
$$I_i = I_{0i} + w_{ie}\hat{u}(t) + w_{ii}\hat{v}(t)$$

respectively, where the synaptic currents $\hat{u}$ and $\hat{v}$ are given as follows:

$$\tau_e\frac{d\hat{u}}{dt} = -\hat{u} + \frac{\Omega}{N_e}\sum_{j=1}^{N_e}\sum_k \delta(t - t_{j,k}),$$
$$\tau_i\frac{d\hat{v}}{dt} = -\hat{v} + \frac{\Omega}{N_i}\sum_{\ell=1}^{N_i}\sum_k \delta(t - t_{\ell,k}), \tag{7}$$

where δ is the Dirac delta functional and Ω is a synaptic strength parameter. We define the integral Z of either above sums in (Eq. 7) over a small time interval $\Delta t \ll \tau_e, \tau_i$ to be

$$Z = \int_t^{t+\Delta t} \frac{\Omega}{N_e}\sum_{j=1}^N\sum_k \delta(s - t_{j,k})ds. \tag{8}$$

In the limit of large N_e and N_i, and the assumption of Poisson spike emission, the integral (Eq. 8) will have an expectation approaching the mean firing rate multiplied by Δt, and the variance of the integral will limit to zero:

$$\langle Z \rangle \sim f_\sigma(I(t))\Delta t,$$
$$Var(Z) \sim 0. \tag{9}$$

Owing to Eq. 9, the sums in Eq. 7 are approximated by mean firing rates of the e- and i-cells, respectively, in which $\hat{u} \to u$ and $\hat{v} \to v$ in probability. Therefore, Eq. 1 will approximate the mean of Eq. 7. Note that the synaptic strength can be set by the w_{jk} weights; so without loss of generality, we set the other strength parameter $\Omega = 1,000$ so that the numerical values of u and v track the mean firing rate, expressed in spikes per second (Hz).

Notably, in our model, for large N_e and N_i, the variability of Eq. 7 limits to zero, and consequently, the external Weiner noise input $\sigma\xi(t)dt$ accounts for all of the noisy fluctuating input to each cell. We choose this limiting case so that we can study emergent oscillatory dynamics by retaining the standard deviation σ of the noise as a control parameter. This limiting case contrasts with a

different limiting case, in which, in the absence of any external fluctuating input, a recurrent neuronal network can self-generate its own variability *via* deterministic chaos. These networks are characterized by variance and covariance order parameters, whose values are determined by self-consistency conditions in which mean and (co)variability of input of each cell must match its output (Moreno-Bote et al., 2008; di Volo et al., 2019; Sanzeni et al., 2020). We have chosen our approach, in which input fluctuations are not enforced to be self-consistent with output fluctuations of network elements to allow flexibility to assume that input fluctuations may be dictated by external inputs, and in addition, we can use the noise parameter σ, an exploratory tuning parameter, to discover possible regimes that support oscillatory solutions.

We simulated networks of $N_e, N_i = 20,000$ cells to verify that the spiking network model's population firing rates are a decent match to that of the mean field model over 10-s simulations. We defined the LFP proxy of the network as the average of all currents to excitatory neurons $I_{LFP}(t) = w_{ee}u(t) - w_{ei}v(t)$ (Mazzoni et al., 2015).

Neural field model

The neural field model is a neural mass model extended over a spatial domain (Coombes et al., 2014), in which nearby regions on the spatial domain are preferentially connected. We have chosen a ring topology spatial domain, parameterized by $\theta \in [0, \pi)$, which is intended to represent a hypercolumn-like population, with each θ-value referring to sub-population of the neural field that is selectively responsive to a θ-parameterized stimulus feature akin to a familiar bar visual stimulus with orientation angle θ (Ben-Yishai et al., 1997; Shriki et al., 2003). The neural field model is described by $u(\theta, t), v(\theta, t)$, the e-activity and i-activity for every θ location on the ring, that solves the integro-differential equation as follows:

$$\tau_e \frac{\partial u}{\partial t} = - u + f_\sigma(\mathcal{W} * [w_{ee}u - w_{ei}v] + I_e)$$
$$\tau_i \frac{\partial v}{\partial t} = - v + f_\sigma(\mathcal{W} * [w_{ie}u - w_{ii}v] + I_i). \tag{10}$$

The integral convolution "$*$" in Eq. 10 is over the θ-domain— $\mathcal{W}*h = \int_0^\pi \mathcal{W}(\theta-\theta')h(\theta')d\theta'$, for any h—and the $\mathcal{W}(\theta)$ von-Mises periodic weight kernel:

$$\mathcal{W}(\theta) = \frac{e^{\kappa \cos(2\theta)}}{\pi I_0(\kappa)}. \tag{11}$$

With this weight kernel (Eq. 11), two points on the ring at θ and θ' will be connected with weight $\mathcal{W}(\theta - \theta')d\theta$. The parameter κ is an inverse variance-like scale parameter, where increasing κ produces a narrower distribution, and $I_0(\kappa)$ is the order-zero modified Bessel function of the first kind, which serves as the normalization constant. We chose the spatial scale of connection to be sufficiently broad (see Figure 3A top left sub panel) so as to support spatial stability across the ring. Moreover, the same spatial scale κ was used for e- and i-cell

populations which also typically result in spatial stability across the ring (see Ali et al., 2016), giving rise to, for example, a uniform bulk oscillation (as in Figure 3A). Having a broader distribution footprint for e- relative to i-cell connection or a narrower overall scale for each subpopulation (larger κ) is typically associated with spatial instabilities (Coombes et al., 2014).

Stimulus inputs are orientation-tuned, given similarly by a Von-Mises-like distribution function

$$\mathcal{S}(\theta) = e^{\kappa_s \cos(2(\theta-\theta_0))-\kappa_s}, \tag{12}$$

in which the peak strength (set to unity) of the stimulus located at orientation $\theta_0 = \pi/2$ and the input spread set by κ_s (see red curve, Figure 3A bottom left sub panel).

The LFP proxy is defined as the α-β-band filtered e-cell total input current (10–30 Hz notch filter), averaged over the ring $I_{e,tot}(t) = \langle \mathcal{W} * [w_{ee}u - w_{ei}v] + I_e \rangle_\theta$ (see Mazzoni et al., 2015), where $\langle \cdot \rangle_\theta$ is the ring average. The phase $\phi(t)$ of this resulting oscillating signal is given by computing the Hilbert transform $\mathfrak{H}$ of $I_{e,tot}(t)$

$$\mathfrak{H}(I_{e,tot}(t)) = I_{e,tot}(\theta, t) + iA(\theta, t), \tag{13}$$

and then taking the angle argument of the resulting analytic signal:

$$\phi(t) \equiv Arg(I_{e,tot} + iA)(t). \tag{14}$$

The phase has an inverse function we defined to be $\tilde{g}(\phi) \equiv t$.

On simulations that were analyzed to assess information coding performance, we included slow-timescale Ornstein-Ulenbeck noise $y(t)$ to both e- and i-cell input currents I_e and I_i globally to the entire network (uniformly across all θ-values). While the origins of the oscillatory dynamics in our model arises from a deterministic Hopf instability, the purpose of this additional noise was to introduce variability in the oscillation amplitude and period, such as those observed in real cortical tissues, as well as to sample oscillatory dynamics for a fluctuating input near the oscillation threshold. However, we have not chosen to implement independent noise inputs across the neural field domain (non-uniform input). Such non-uniform noise would naturally lead to degraded coding performance for both phase and rate representations of activity. The relative degree of degradation for the two codes is an important question but is not addressed in this article. The dynamics of y are given by the SDE

$$\tau_y \frac{dy}{dt} = -y + \sigma_y \sqrt{\tau_y}\xi(t), \tag{15}$$

where $\xi(t)$ is uncorrelated zero-mean, unit-variance gaussian white noise. This SDE (Eq. 15) results in a normal stationary distribution of currents q with zero-mean, standard deviation σ_y, and a temporal autocorrelation decay timescale $\tau_y = 50$ ms so that the network oscillations, which were typically in the $\sim$20 Hz β-band range ($\sim$50 ms cycles), showed robust cycle to cycle variability but little long-timescale multi-cycle correlation.

We wish to study how input modulations akin to WM signals can modulate the response properties—both rate responses and

TABLE 1 A list of parameter values of the neural mass and field models.

Parameter	Value
V_t	-50 mV
V_r	-65 mV
V_l	-65 mV
C	1 mF
τ_i	15 ms
τ_e	5 ms
g	1/15 mS
σ_0	5.5
I_{e0}	-2.31
I_{i0}	-3.81
Δ_{stim}	0.018
Δ_{WM}	0.015
w_{ee}	0.9
w_{ii}	1.9
w_{ie}	1
w_{ei}	2
κ	5.0625
κ_s	20
σ_y	0.02
τ_y	50 ms

oscillatory responses—of neural masses within the neural field. In the absence of any external input, the neural field model parameters were set to a point near a supercritical Hopf instability where oscillations were small-amplitude, with an oscillation frequency in the β-band approximately 15–20 Hz. We ran simulations from this set point over four levels of a global WM input and four levels of orientation-selective stimulus input, starting from zero, which models the effective contrast of the stimulus. We call these WM 0, 1, 2, and 3 levels and stimulus contrast levels 0, 1, 2, and 3. Altogether, the input to cells can be represented by

$$I_e(\theta, t) = I_{e0} + a\Delta_{stim}\mathcal{S}(\theta) + b\Delta_{WM} + y(t) \quad (16)$$
$$I_i(\theta, t) = I_{i0} + a\Delta_{stim}\mathcal{S}(\theta) + b\Delta_{WM} + y(t),$$

where Δ_{stim} and Δ_{WM} are the current increments for the respective input levels of stimulus $a = 0, 1, 2, 3$ and WM $b = 0, 1, 2, 3$. For the neural field simulations, we fixed the noise parameter to $\sigma = 5.5$. The complete list of parameter values of the neural mass and field models are given in Table 1.

Information-theoretic measures of phase and rate coding

We wanted to test the ability of the neural field ring model to represent distinct sensory stimuli, so we tested the discriminability of the neural field activity across distinct stimuli positions on the ring. We gave elevated input centered at $\theta_0 = \pi/2$ according to Eq. 12 and assigned this to be the favored stimulus region, whereas elsewhere were less stimulated regions.

Each oscillation cycle is characterized by a phase decomposition $\phi(t)$ of the LFP proxy (see above) in which $\phi \in (-\pi, \pi]$. A typical cell at a location θ fires at the instantaneous mean rate described by the F-I curve $f_\sigma(I(\theta, t))$ as a function of the input current $I(\theta, t)$. Spikes are distributed over the oscillation cycle period T in proportion to the instantaneous firing rate $f_\sigma(I(\theta, t))$. To obtain a phase distribution from the instantaneous firing rate over the cycle period, we performed a change of variables from the time domain to the phase domain given by the function $\tilde{g}(\phi) = t$ (see Eqs. 13, 14). The phase density obtained from this change of variables is

$$p_\theta(\phi) = \frac{f_\sigma(I(\theta, \tilde{g}(\phi)))}{\int_0^T f_\sigma(I(\theta, t))dt} \frac{d\tilde{g}}{d\phi}(\phi), \quad (17)$$

where $\frac{d\tilde{g}}{d\phi}(\phi)$ is the change-of-variables Jacobian factor, and the integral in the denominator is the normalization ensuring $\int_{-\pi}^{\pi} p_\theta(\phi)d\phi = 1$. We assume the typical cell at a location θ emits a Poisson-distributed number of spikes per unit time, with poisson parameter $\lambda_\theta \equiv T^{-1}\langle f_\sigma(I(\theta, t))\rangle_T$, where $\langle\cdot\rangle$ represents a temporal average over a cycle of time-length T. Analogously, we define spikes per cycle as $\hat{\lambda}_\theta = \lambda_\theta T$ to be emitted by a typical cell, yielding a discrete Poisson distribution for the number of spikes n in the same oscillation cycle window T: $r_\theta(n) = \frac{\hat{\lambda}_\theta^n e^{-\hat{\lambda}_\theta}}{n!}$.

We computed two measures of coding performance. The first examines the discriminability of two θ-stimuli points on the ring. Without loss of generality, we chose $\theta = \pi/2, \pi/4$ for our two representative points. The discriminable information or information gain (IG) is defined as the Kullback–Libler divergence (in nats) $D_{KL}(q_{\pi/2}||q_{\pi/4}) = \sum_z p(z)\ln(q_{\pi/2}(z)/q_{\pi/4}(z))$, where $z = \phi$ and $q = p$ for the phase code and $z = n$ and $q = r$ for the rate code:

$$IG_\phi(\pi/4 \to \pi/2) = D_{KL}(p_{\pi/2}||p_{\pi/4})$$
$$IG_r(\pi/4 \to \pi/2) = D_{KL}(r_{\pi/2}||r_{\pi/4}). \quad (18)$$

The IG (Eq. 18) measures the amount of information learned about the data distributed according to $p_{\pi/2}$ given one scores data according to $p_{\pi/4}$, minus the true entropy of the data.

The second measure of coding performance we computed was the mutual information (MI) between stimulus feature θ and either the oscillation phase variable ϕ or the spike count rate variable n. For concreteness, we chose to represent the stimulus feature across eight points, which evenly distributed about the ring $\theta_k = \frac{\pi k}{8}$, for $k = 1, 2, \ldots 8$. We assume on a "trial" that each of the eight stimuli was equally likely to occur, in which case, the probability of each stimuli was 1/8. The MI then can be constructed to be the Kullback–Libler divergence from the joint phase distribution $p_{\theta_k} = \frac{1}{8}\sum_{j=1}^{8} p_{\theta_j}(\phi)\delta_{jk}$ (where δ_{jk} is the Kronecker delta) to the marginal distribution over the θ_k stimuli: $\bar{p}(\phi) = \frac{1}{8}\sum_{j=1}^{8} p_{\theta_j}(\phi)$ and similarly for the rate code. This formula can be generalized to any number of points on the ring. For each code, the MI is also expressible in terms of the entropy $\mathcal{H}$ of the average phase (rate)

code across the stimuli, minus the average entropy of each phase (rate) code at a given stimulus point:

$$MI_{\phi}(\pi/2; \pi/4) = D_{KL}(p_{\theta_k}||\bar{p}) = \mathcal{H}(\bar{p}) - \frac{1}{n}\sum_{k=1}^{8}\mathcal{H}(p_{\theta_k})$$

$$MI_{r}(\pi/2; \pi/4) = D_{KL}(r_{\theta_k}||\bar{r}) = \mathcal{H}(\bar{r}) - \frac{1}{8}\sum_{k=1}^{8}\mathcal{H}(r_{\theta_k}). \tag{19}$$

We simulated the neural field over a 10-s duration time window over each of the 16 WM and stimulus input value combinations. For each simulation, using the notch filtered LFP (10–30 Hz), we segmented the LFP into individual cycles with individual periods. On each cycle, we computed phase and rate encoding distributions and computed the aforementioned information performance measures IG and MI and examined the average and variability across cycle trials and conditions.

Numerical methods

All numerical simulations and calculations were performed using custom-written Matlab (Natick, MA) code. Simulations of DE models utilized a fixed step time of $\Delta t = 0.02$ ms. Simulations of the two-variable neural mass model without noise were performed using a standard Matlab "ode23" DE solver. Notably, the F-I curve we used involved solving the mean first passage time (Eqs. 4, 5) that consisted of nested integrals, which would be numerically inefficient to function-call on each time step. Instead, we computed these curves once over a fine mesh and used a cubic spline fit in lieu of the exact solution.

Stochastic simulations of the LIF spiking model network model were computed using the Euler–Maruyama method, and an explicit 2nd-order Heun's method was used for all deterministic variables, with $\Delta t = 0.02$

The neural field model was simulated by partitioning the orientation tuning domain in 360 equally spaced grid points along the $\theta \in [0, \pi)$ domain. An explicit 2nd-order Heun's method was used for all deterministic variables, while noise input was computed each using a first-order explicit Euler–Maruyama method (Gardiner, 2009).

Results

Analysis of neural mass model

The main question we address in this article is how oscillatory and non-oscillatory sensory representations in neural populations can be modulated by top-down input. In order to do so, we first investigate a e- and i-cell recurrent network that supports oscillations in isolation, termed the neural mass model. Then, in the following section we construct a neural field model consisting of many of these e- and i-cell subpopulations that will support sensory coding.

We derived a mean field model of these spiking networks characterized by a 2D nonlinear system of DEs (Eq. 1) that describe the mean activity of the e- and i-cell populations, in which the mean firing rate is dictated by a firing rate function—the F-I curve—as a function of input current. The F-I curve $f_{\sigma}(I)$ was defined as the firing rate of an LIF model cell subject to noise input with strength σ (see Eqs. 4–6), and are shown in Figure 1A for several σ-values. The curves show good correspondence to the firing rates obtained from Monte Carlo simulations of the LIF cells. Increasing σ increases firing rates overall, resulting in an leftward translation of the curves. In addition, increasing σ has the effect of reducing the firing rate gain $\frac{df_{\sigma}}{dI}$ at any fixed firing rate f_{σ} (Figure 1B; see for instance Chance et al., 2002; Ferguson and Cardin, 2020).

The we searched for oscillatory instabilities in our neural mass model (Eq. 1) by performing a wide parameter scan over I-σ combinations. For each I-σ combination, we located equilibria and computed linear stability. Figure 1C shows there is an unstable region in I-σ-space where an eigen-analysis returns positive real part eigenvalues of the linearized system. Parameter pairs (I, σ) that move upward to the right produce increased mean e-cell firing rate (Figure 1D). Entering the left-side of the unstable region in I-σ-space from left to right, yields a supercritical Hopf bifurcation (see magenta-dashed line, where the Lyapunov coefficient is negative). Moving left to right through the bifurcation (see example brown-line trajectory) also results in oscillations with frequencies in the β-range 12–30 Hz (Figure 1E), whose amplitudes increase continuously from zero (Figure 1F).

The neural mass model is was derived as a simplified representation of the dynamics of a large-N spiking model consisting of excitatory and inhibitory LIF model neurons. The mean field model predicts that oscillations can emerge from increasing mean input I (for a fixed $\sigma = 5.5$), depicted by the brown line in Figure 1F. Sample solutions of the neural mass model are shown for three input levels ($I = -2.45, -2.3, -2.15$) in Figure 2A, one below the bifurcation where a equilibrium solution is stable, one slightly past the bifurcation point where small-amplitude oscillations emerge, and one further past the bifurcation where larger-amplitude oscillations are observed. Using the same input levels and shared parameter settings, we also simulated a LIF spiking network model ($N_e, N_i = 20,000$). These stochastic simulations show a correspondence to the neural mass model solutions. At the lower input level, weak fluctuations exist, reflecting the fluctuation-driven spiking in the finite-N networks. At the predicted point of the neural field bifurcation, the LIF network exhibits emergence of more regular, intermediate-amplitude oscillations; while increasing input further leads to larger amplitude oscillations, in a manner similar to the neural mass model.

Supercritical Hopf bifurcations yield oscillation magnitudes that are continuous with zero at the point of bifurcation. Commensurate with the mean field model, we expect that the large-N LIF spiking network exhibits LFP power spectra that show an increase in oscillatory power with input. We also expect the central peak of LFP power to shift to higher frequencies as input is increased through the predicted bifurcation point. Figure 2C shows this predicted increase oscillation power and shift in frequency, both centered in the β-band, over seven input levels (we interleaved four more input levels between the three depicted in Figure 2B). This predicted increase in mean LFP oscillation frequency and oscillatory power in the LIF model network is

FIGURE 1
An unstable equilibrium, in which oscillations emerge, occurs in a protruding region in I-σ parameter space. **(A)** F-I curve solutions $f_\sigma(I)$ to Eqs. (4–6), parameterized by the diffusion noise σ parameter. **(B)** Increasing σ reduces the overall firing rate gain at a given firing rate $df_\sigma/dI|_{f_\sigma}$. **(C)** The real part of the smallest-magnitude eigenvalue $\mathrm{Re}(\lambda_k)$ parameterized by the input I (abscissa), and F-I curve gain parameter σ (ordinate). There is a subregion where the eigenvalues are complex $\omega = \mathrm{Im}(\lambda_k) \neq 0$, in which crossing into the instability region produces a supercritical Hopf bifurcation (magenta dashed line). **(D)** Moving from left to right and from low, to high, the mean firing rate increases in a continuous manner, both outside and inside the unstable region. **(E)** Inside the unstable region, the oscillation frequency increases going diagonally up and to the right. **(F)** The supercritical Hopf bifurcation elicits oscillation amplitudes that emerge continuously from zero on the upper part of the unstable region. In **(D–F)** the brown line from left to right indicates a input parameter path of interest where oscillations emerge via supercritical Hopf in which mean firing rate, oscillation frequency, and weak amplitude oscillations increase continuously from zero.

summarized in Figures 2D, E, respectively, and shows a strong correspondence to those computed for the neural mass model. Note, however, that the LIF network model exhibits non zero power for input levels below the bifurcation point; whereas the neural mass model oscillatory power is necessarily zero. This is naturally due to the LIF network being a stochastic simulation in which case we expect small fluctuations. Such sub-threshold transient oscillatory behavior has been suggested to endow networks with useful function (Palmigiano et al., 2017). Importantly, just above the bifurcation point, there is good correspondence between the oscillatory frequency and power of the neural mass and LIF spiking model.

Phase and rate representations of stimulus inputs in the neural field model

The neural field model is a spatially extended neural mass model (see Coombes et al., 2014) with a ring connection topology, designed to represent an orientation tuning domain in which nearby regions on the ring are preferentially connected. We tuned each point on the ring to be near the supercritical Hopf instability that we identified in the neural mass model. In the absence of any θ-dependent input, the neural field exhibits a bulk oscillation

(Figure 3A top panel; see also Mosheiff et al., 2023) due to the lateral connections in the network producing oscillation syncrony across the ring. The synaptic weight function $\mathcal{W}(\theta)$ (see Eq. 11, centered at $\pi/2$ for convenience) is also depicted in Figure 3A (left sub panel of top panel). Three types of external inputs were given to the neural field: stimulus-like inputs, WM-like inputs, and temporally fluctuating random inputs (see Eqs. 15, 16).

Stimulus-like inputs are tuned to preferentially activate one θ-tuned region of the ring over other regions. These θ-tuned stimulus inputs induce changes to the time course of the bulk oscillation as well as mean firing rates across the ring. For example, Figure 3A (bottom panel) shows e-cell activity with the $\pi/2$-tuned input producing a temporal bend in the oscillation across the ring arising from elevated input in regions centered about $\theta = \pi/2$. The synaptic input function $\mathcal{S}(\theta)$ (Eq. 12) is depicted along with $\mathcal{W}(\theta)$ in Figure 3A (left sub panel of bottom panel). This input induced the oscillation to initiate sooner in regions near $\pi/2$ relative to elsewhere on the ring.

To illustrate the effect of featured tuned inputs on the relative timing of oscillatory activity across the ring, we chose two stimulus features on the ring, a favored feature $\theta = \pi/2$ from which we discriminate from the unfavored feature $\theta = \pi/4$ (see Figure 3 top panel, yellow, red-dash, respectively), and from which we have defined a two-point discrimination task (see below). At these two

FIGURE 2

The emergence of oscillations through increased input in the neural mass model corresponds similar emergent oscillations in the large-N LIF spiking network model from which the neural mass model was derived. **(A)** Sample solution trajectories of the neural mass model at three ascending input levels ($I = -2.45, -2.3, -2.15$) one below the bifurcation (lower panel), one just slightly above (middle), and one further beyond (top). **(B)** The large-N ($N_e, N_i = 20,000$) LIF spiking network driven at the same inputs and parameter settings as the neural mass model in **(A)**. **(C)** The power spectrum of the large-N LIF model over seven input levels (four more levels interleaved between the three illustrated in **(A, B)** shows that increasing input I produces greater oscillatory power, and a shift toward higher frequencies within the β-oscillatory band. **(D, E)** The mean oscillation frequency and oscillatory power, respectively, computed from the power spectra in **(C)** for the LIF spiking network, and the neural mass model.

positions on the ring, the mean instantaneous firing rate curves $f_\sigma(I(t))$ are shown in Figure 3B for no-stimulus (top panel) in which the two features exhibit the exact same rate dynamics as the bulk oscillation. When stimulus centered at $\theta = \pi/2$ is given, the firing rate of the favored feature exhibits a larger amplitude peak that occurs earlier than peak at the unfavored location (bottom panel).

The oscillatory phase angle of the neural field is defined by the phase decomposition $\phi(t)$ of the LFP proxy signal (via Hilbert transform, see Methods). Here, LFP is defined as the mean e-cell input over the entire ring (Mazzoni et al., 2015). Figure 3C shows the phase-distributions at the $\theta = \pi/2$ and $\pi/4$ locations (see Eq. 18). Consistent with the firing rate time courses (in Figure 3B) the phase distributions reflect that without stimulus input, the phase distributions are identical (top). A stimulus input at $\theta = \pi/2$ produces more spikes to occur earlier in the neural field model's oscillatory cycle and is more sharply peaked than the phase distribution at the unfavored feature $\theta = \pi/4$.

Note that while the phase signatures of the stimulus are salient to the eye in Figure 3C, the changes to the average firing rate over the cycle are very weak. Figure 3D shows the differences in the spike count distributions that defines the rate code in the stimulus and no-stimulus conditions. Naturally, the no-stimulus case has zero difference between spike count distributions; however, even while the peak firing rate during the oscillation cycle differs by about 3 Hz between the favored and unfavored locations (see Figure 3B bottom), the spike count distributions over the entire cycle differ by at most $\sim 1 \times 10^{-2}$, which is small relative to the phase distribution differences.

In contrast with stimulus inputs, WM-like inputs are global, uniform over the ring. Increases to global WM input (with WM input random fluctuations, see Methods) to the neural field ring produces increases in LFP peak power and frequency in the beta band (Figure 3E). Finally, WM input has a positive effect on mean firing rate of the ring (Figure 3F), and mean spike phase locking (SPL; Figure 3G; SPL is 1 minus the circular variance). These effects observed in the neural field model are consistent with the constituent neural mass model, and are also consistent with experimentally observed WM-driven changes observed in extrastriate areas (Bahmani et al., 2018).

Finally, temporally fluctuating random inputs were included to simulations that we used to perform our sensory coding analyses (see below). We included these noise inputs to elicit a broad sampling of oscillatory frequencies and amplitudes, particularly near the deterministic bifurcation.

Sensory coding performance of phase and rate codes

We sought a rigorous way to discriminate representations of activity between stimulus features on the ring θ by representing the neural field model solution dynamics based on two coding schemes: an oscillatory code, and a spike rate code. Specifically, we wanted to determine how stimulus discrimination performance was affected by stimulus contrast level as well as the WM-input levels. For both of the coding schemes, we computed two measures

FIGURE 3
The neural field model replicates oscillatory behaviors observed in sensory areas during WM tasks. **(A)** The neural field e-cell oscillatory dynamics over its feature space in the presence of no stimulus (top), depicted along with the weight kernel $\mathcal{W}$ (top left sub panel; see Eq. 11). In the presence of a stimulus $\mathcal{S}$ (bottom, left sub panel red curve; see Eq. 12), the e-cells exhibit an oscillation with a bend located at the stimulus location. Two example stimulus feature locations $\theta = \pi/2, \pi/4$ are shown on the ring in the top panel, yellow, red-dash, respectively. Neural field model example simulations exhibit phase locked oscillations across the ring, but preferred stimulus exhibits earlier activity. **(B)** e-cell firing rate on ring at the two example locations over 100ms simulation time for increasing stimulus contrast from no-stimulus (top) and with-stimulus (bottom). **(C)** The associated spike phase distributions for the same two example stimulus features as in **(B)** in no-stimulus and with-stimulus (top and bottom). **(D)** The difference in spike count distributions between the two example stimulus features in no-stimulus and with-stimulus conditions shows weak ($\sim 10^{-3}$) average spike count changes due to stimulus, while peak firing rate during an oscillatory cycle exhibits 3Hz differences between stimuli [see **(B)**, bottom]. **(E–G)** Increases in WM input produce LFP peak power and LFP frequency increases, mean firing rate increases, and increased in SPL, respectively.

of coding performance. First, we computed the IG [see Methods (Eq. 18), also termed the relative entropy or Kullback-Libler divergence], measuring the discrimination information between two representative points on the ring: $\theta = \pi/2$ and $\pi/4$. Second, we computed MI (Eq. 19) between the stimulus ring location variable θ and the neural activity representation variable—either phase or rate representations. We computed these measures over each oscillation cycle of the model-generated data from 10 s of simulation, to generate averages and standard deviations for each performance measure.

The IG is a simple means of assessing amount of information one gains from observing data obtained from the $\pi/2$-parameterized distribution if one had a initially assumed the data came from the $\pi/4$-parameterized source. Greater IG underlies greater statistical discrimination power from which one can formulate a decision threshold to detect the $\pi/2$ stimulus over $\pi/4$. Figure 4A shows the IG measure for both the phase code (warm colors) and spike rate code (cooler colors) on ordinate log scale. Naturally, increasing contrast improves performance across all WM-input levels; although the zero contrast discrimination performance was, of course, zero, and was not shown on the log scale. Interestingly, WM-input increases produce positive enhancements in the phase code across all non zero contrast

levels. Conversely, the spike rate code shows the opposite effect—a reduction of coding performance results from WM input increases. The insets show the same data as a function of contrast level in linear ordinate scale, which reveals that increasing WM input enhances stimulus contrast gain in the phase code, while simultaneously reducing the spike rate code gain. Note also that the phase code over all contrasts and WM levels was about two orders of magnitude larger than that of the spike rate code. These effects were significant: over three factors of the choice of code—phase or rate—as well as the 0–3 input levels for WM, and the 0–3 stimulus contrast levels all produced significant isolated effects on coding performance, as well as significant interaction effects when tested in a three-factor ANOVA ($p = 0$ for all factors and interactions).

The finding that information coding performance is reduced for the spike rate code in the presence of WM signals can be attributed to the manner in which Poisson-based spike codes depend on baseline firing rate λ_0. WM inputs are global across the neural field and slightly increase the overall baseline firing rate: $\Delta\lambda_0$. Likewise, the stimulus in our setup produces another small change $\Delta\lambda_s$ in the average firing rate between the two stimuli-locations θ on the neural field. It is a standard result in information theory that Poisson information gain scales as $IG \approx \frac{1}{2}\frac{\Delta\lambda_s^2}{\lambda_0}$. So, the information gain is inversely proportional to baseline firing λ_0.

Thus, any elevation of baseline $\lambda_0 \rightarrow \lambda_0 + \Delta\lambda_0$ by WM inputs undercuts discrimination performance: $\frac{\Delta\lambda_s^2}{\lambda_0 + \Delta\lambda_0} < \frac{\Delta\lambda_s^2}{\lambda_0}$. That is, this change in baseline firing $\Delta\lambda_0$ due to WM input makes any stimulus driven change $\Delta\lambda_s$ less discriminable through a divisive normalization by λ_0 (Reynolds and Heeger, 2009).

The MI is a complementary means of assessing coding performance that measures the amount of information shared between the stimulus features over multiple points across the ring and the neural field readout of the phase and count distributions—an MI value of 1 would mean that spike data observations (phase or count) completely determine which stimulus is present. We used eight evenly spaced points across the ring to achieve broad coverage. The MI measure results are qualitatively similar in all respects to the IG measure (Figure 4B). Just as with IG, the MI measure depended significantly on the three factors: choice of code, WM level, and contrast level, as well as significant interaction effects between all factors when tested in a three-factor ANOVA ($p = 0$, for all factors and interactions).

Discussion

In this article we sought to address how oscillatory changes observed in sensory areas during WM input can possibly enhance oscillatory representations of sensory activity—the phase code—given that mean rate changes associated with WM are weak or nonexistent (Bahmani et al., 2018). To achieve a realistic phase code, we chose a model—our neural mass model—that satisfied the requirement that cortical spikes be only modestly phase locked to the LFP β-oscillation, and the model oscillation must be modulated by WM-like input in a similar fashion to experimental findings. We were able to satisfy to these requirements, and achieve the requisite enhancement of the phase code in response to elevated WM input, by having our model tuned at the point of a certain type of oscillatory instability: a supercritical Hopf point. At this instability point, WM input produces a weak amplitude oscillation continuous with zero. Such weak amplitude oscillations are consistent with experimental findings that cells exhibit weak or intermediate phase locking to the β-band of the LFP, in contrast with large-amplitude oscillations that would necessarily produce strong phase locking and synchrony. The model was derived from, and well-described by, a large-N LIF recurrent network of e- and i-cell populations driven by noise input.

By joining a continuum neural mass models into a neural field model across a spatial domain θ that represented a stimulus feature, we could preferentially stimulate areas of the domain to create sensory representations in the neural activity (Ben-Yishai et al., 1997; Shriki et al., 2003; Coombes et al., 2014). Both oscillatory and mean rate based representations—the phase and rate codes—were assessed using information theoretic measures as a function of stimulus contrast and WM-input level. The main effect we found was that increasing WM input *increased* phase coding performance while rate coding performance *decreased*. WM also produced an enhancement of contrast gain for the phase code, and a decrease in contrast gain for the rate code. The phase coding performance, owing to the larger information capacity of the continuous phase variable relative to the discrete-valued rate code, was also about two orders of magnitude larger than the rate coding performance.

The mechanism by which WM enhances phase code representations is due to the connection between oscillation amplitude and phase certainty. Specifically, the phase code $p_\theta(\phi)$ is a one-to-one transformation with the e-cell firing rate curve $f_\sigma(I_e(t, \theta))$ defined over the oscillation cycle $t \in [0, T]$, at a feature θ. This one-to-one map from $\phi \in [-\pi, \pi]$ to $t \in [0, T]$ provided a link between oscillation amplitude and phase certainty: WM-driven oscillations with larger amplitude tended to produce spikes in a more concentrated period of time, yielding larger spike phase locking, which in turn, translated to greater probabilistic certainty (reduced variance) in the phase variable. Naturally, reduction of variance makes any two θ-value-indexed densities $p_\theta(\phi)$ discriminated more readily in the presence of a stimulus.

In contrast to phase coding, the information coding decline of the rate code as a function of WM can be accounted for by the nature of Poisson spiking: information discrimination of small firing rate changes from λ_0 to $\lambda_0 + \Delta\lambda$ is inversely proportional to the baseline firing λ_0; therefore, WM input that slightly elevates baseline firing λ_0 has the effect of undercutting coding performance.

This link between increased oscillation amplitude and increased phase certainty and enhanced information gain was independent of mean firing rate over the oscillation cycle. In our representation of neural activity, the spike rate code is a distinct channel of information from the phase code because the neural activity is readily segmented into the constant and oscillating components of the cycle activity in a Fourier decomposition, respectively: $f_\sigma(I_e(t, \theta)) = \lambda + \left(\sum_{n=1}^{\infty} c_n e^{i\frac{n\pi}{T}t} + CC \right)$, where λ is the mean firing rate (i.e., the DC mode of the firing rate oscillatory cycle). While we can associate two distinct codes as distinct Fourier features of the neural oscillation, how exactly global WM input can modulate both mean cycle activity λ and the oscillatory features c_n in a coordinated fashion is not fully understood. Specifically, more study is needed to understand how discriminable information is allocated to these distinct channels, as well as how this allocation could be controlled by inputs. In the model we studied here, the overwhelming majority share of discriminable information was allocated to the phase code channel while the mean rate λ exhibited comparably little change in response to stimulus or WM inputs. That our neural field model exhibits a relaxation oscillation offers a plausible explanation for this divergent information allocation: the e-cell rising phase of the cycle is stopped by the i-cell-driven falling phase that partly offsets e-cell firing activity, yielding less average firing rate gain, and so less rate coding performance from an input. This information about an input is still available in the phase variable, encoded the timing of the rise-then-fall phase relative to other parts of the neural field (see Nesse et al., 2021b for a similar perspective on rate coding performance loss through intracellular adaptation in single neuron spike coding).

In our model, we have relied on noise driven spiking of the constituent LIF cells whose average synaptic dynamics is described by our neural mass and field models. Endogenous network fluctuations in the finite-N networks were not explicitly accounted for in cell inputs, but clamped to a specific white noise level σ. This noise-clamping enabled us to locate supercritical-Hopf oscillatory instabilities in the I-σ parameter space of the neural mass model. This limiting case, where noise is a fixed exogenous input, is distinct from network models in which mean

FIGURE 4
Divergent sensory discrimination performance is exhibited by the phase and rate codes as a function of WM input. **(A)** IG measure shows enhancement for the phase code across all non-zero stimulus contrasts while the rate code is reduced, measured on a log ordinate scale. Insets of **(A)** The same data as the main panel but plotted on a ordinate linear scale as a function of contrast, which shows increasing WM input enhances the coding gain as a function of contrast while reducing rate coding gain for the same WM increases. **(B)** Qualitatively similar results to **(A)** but for the MI coding performance measure. All error bars reflect standard deviations.

and co-variability of cells is self-consistently solved for to obtain asynchronous steady states (Moreno-Bote et al., 2008; di Volo et al., 2019; Sanzeni et al., 2020). It is unknown to us if networks derived under such self-consistency constraints can support the emergent oscillations that we studied here.

The results of this modeling study have applications to sensory areas of the brain that are known to receive signals that reflect the content of WM. Consistent with the present modeling results, previous work by our group in areas V4 and MT of the Macaque, has identified that WM signals induce changes in both the oscillatory power, peak-power frequency of LFPs in the α-β band, as well as modulating the occurrence of spikes at certain phases of the LFP oscillation. Also consistent with our results here, it was simultaneously observed that such WM signals to V4-MT elicit small changes to mean firing rate that may not readily account for the observed cognitive benefits produced by WM (Bahmani et al., 2018). The mechanism by which WM modulates phase certainty through amplitude modulations, as well as the prospect of storing vastly more stimulus information in the phase code relative to the rate code, suggests a possible way in which the cognitive benefits of WM are mediated. Note however, that our rate representations of stimuli assumed uncorrelated Poisson spiking that does not account for neuronal correlations that have been shown to be modulated by cognitive states (Cohen and Maunsell, 2009) and it is unknown how such a correlation modulation compares to

the phase code modulations we have studied here. Additionally, the potential we have identified for greater sensory information to be contained oscillatory representations begs the question of how such representations are read out by downstream targets. We do not address that question here. Certainly, read out of phase coded information by downstream brain areas is more complicated than the transfer of rate-coded information, but potential mechanisms have been proposed (Comeaux et al., 2023).

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

Author contributions

WN designed all analyses, simulations, and wrote the article. KC provided input on text composition, citations, and research design. BN helped administer research design and text composition. All authors contributed to the article and approved the submitted version.

Funding

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ali, R., Harris, J., and Ermentrout, B. (2016). Pattern formation in oscillatory media without lateral inhibition. *Phys. Rev. E* 94, 012412. doi: 10.1103/PhysRevE.94.012412

Angelucci, A., Bijanzadeh, M., Nurminen, L., Federer, F., Merlin, S., Bressloff, P. C., et al. (2017). Circuits and mechanisms for surround modulation in visual cortex. *Annu. Rev. Neurosci.* 40, 425–451. doi: 10.1146/annurev-neuro-072116-031418

Ardid, S., Wang, X.-J., and Compte, A. (2007). An integrated microcircuit model of attentional processing in the neocortex. *J. Neurosci.* 27, 8486–8495. doi: 10.1523/JNEUROSCI.1145-07.2007

Bahmani, Z., Daliri, M., Merrikhi, Y., Clark, K., and Noudoost, B. (2018). Working memory enhances cortical representations via spatially specific coordination of spike times. *Neuron* 97, 967–979. doi: 10.1016/j.neuron.2018.01.012

Ben-Yishai, R., Hansel, D., and Sompolinsky, H. (1997). Traveling waves and the processing of weakly tuned inputs in a cortical network module. *J. Comput. Neurosci.* 4, 57–77. doi: 10.1023/A:1008816611284

Bosman, C. A., Schoffelen, J.-M., Brunet, N., Oostenveld, R., Bastos, A. M., Womelsdorf, T., et al. (2012). Attentional stimulus selection through selective synchronization between monkey visual areas. *Neuron* 75, 875–888. doi: 10.1016/j.neuron.2012.06.037

Brunel, N., and Wang, X. (2001). Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *J. Comput. Neurosci.* 11, 63–85. doi: 10.1023/A:1011204814320

Chance, F. S., Abbott, L. F., and Reyes, A. D. (2002). Gain modulation from background synaptic input. *Neuron* 35, 773–782. doi: 10.1016/S0896-6273(02)00820-6

Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., et al. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat. Neurosci.* 13, 369–378. doi: 10.1038/nn.2501

Cohen, M. R., and Maunsell, J. H. R. (2009). Attention improves performance primarily by reducing interneuronal correlations. *Nat. Neurosci.* 12, 1594–1600. doi: 10.1038/nn.2439

Comeaux, P., Clark, K., and Noudoost, B. (2023). A recruitment through coherence theory of working memory. *Prog. Neurobiol.* 228, 102491. doi: 10.1016/j.pneurobio.2023.102491

Coombes, S., Beim Graben, P., Potthast, R., and Wright, J. (2014). *Neural Fields: Theory and Applications*. Berlin: Springer. doi: 10.1007/978-3-642-54593-1

Daume, J., Gruber, T., Engel, A. K., and Friese, U. (2017). Phase-amplitude coupling and long-range phase synchronization reveal frontotemporal interactions during visual working memory. *J. Neurosci.* 37, 313–322. doi: 10.1523/JNEUROSCI.2130-16.2016

Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222. doi: 10.1146/annurev.ne.18.030195.001205

di Volo, M., Romagnoni, A., Capone, C., and Destexhe, A. (2019). Biologically realistic mean-field models of conductance-based networks of spiking neurons with adaptation. *Neural Comput.* 31, 653–680. doi: 10.1162/neco_a_01173

Engel, T. A., Steinmetz, N. A., Gieselmann, M. A., Thiele, A., Moore, T., Boahen, K., et al. (2016). Selective modulation of cortical state during spatial attention. *Science* 354, 1140–1144. doi: 10.1126/science.aag1420

Ferguson, K. A., and Cardin, J. A. (2020). Mechanisms underlying gain modulation in the cortex. *Nat. Rev. Neurosci.* 21, 80–92. doi: 10.1038/s41583-019-0253-y

Fries, P. (2009). Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annu. Rev. Neurosci.* 32, 209–224. doi: 10.1146/annurev.neuro.051508.135603

Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron* 88, 220–235. doi: 10.1016/j.neuron.2015.09.034

Gardiner, C. (2009). *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Berlin: Springer-Verlag.

Humphreys, G. W., Duncan, J., Treisman, A., and Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 353, 1245–1255. doi: 10.1098/rstb.1998.0280

Kanashiro, T., Ocker, G. K., Cohen, M. R., and Doiron, B. (2017). Attentional modulation of neuronal variability in circuit models of cortex. *Elife* 6, e23978. doi: 10.7554/eLife.23978.019

Kapur, A., Pearce, R., Lytton, W., and Haberly, L. (1997). Gabaa-mediated ipscs in piriform cortex have fast and slow components with different properties and locations on pyramidal cells. *J. Neurophysiol.* 78, 2531–2545. doi: 10.1152/jn.1997.78.5.2531

Kopell, N., Whittington, M. A., and Kramer, M. A. (2011). Neuronal assembly dynamics in the beta1 frequency range permits short-term memory. *Proc. Nat. Acad. Sci.* 108, 3779–3784. doi: 10.1073/pnas.1019676108

Kuznetzov, Y. (1998). *Elements of Applied Bifurcation Theory*. Berlin: Springer.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., and Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320, 110–113. doi: 10.1126/science.1154735

Lee, H., Simpson, G. V., Logothetis, N. K., and Rainer, G. (2005). Phase locking of single neuron activity to theta oscillations during working memory in monkey extrastriate visual cortex. *Neuron* 45, 147–156. doi: 10.1016/j.neuron.2004.12.025

Lee, J., Whittington, M., and Kopell, N. (2013). Top-down beta rhythms support selective attention via interlaminar interaction: a model. *PLoS Comput. Biol.* 9, e1003164. doi: 10.1371/journal.pcbi.1003164

Liebe, S., Hoerzer, G. M., Logothetis, N. K., and Rainer, G. (2012). Theta coupling between v4 and prefrontal cortex predicts visual short-term memory performance. *Nat. Neurosci.* 15:456-462, S1–S2. doi: 10.1038/nn.3038

Mazzoni, A., Lindén, H., Cuntz, H., Lansner, A., Panzeri, S., Einevoll, G. T., et al. (2015). Computing the local field potential (LFP) from integrate-and-fire network models. *PLoS Comput. Biol.* 11, 1–38. doi: 10.1371/journal.pcbi.1004584

Mendoza-Halliday, D., Torres, S., and Martinez-Trujillo, J. C. (2014). Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. *Nat. Neurosci.* 17, 1255–1262. doi: 10.1038/nn.3785

Merrikhi, Y., Clark, K., and Noudoost, B. (2018). Concurrent influence of top-down and bottom-up inputs on correlated activity of macaque extrastriate neurons. *Nat. Commun.* 9, 5393. doi: 10.1038/s41467-018-07816-4

Michalareas, G., Vezoli, J., van Pelt, S., Schoffelen, J.-M., Kennedy, H., and Fries, P. (2016). Alpha-beta and gamma rhythms subserve feedback and feedforward influences among human visual cortical areas. *Neuron* 89, 384–397. doi: 10.1016/j.neuron.2015.12.018

Mitchell, J. F., Sundberg, K. A., and Reynolds, J. H. (2007). Differential attention-dependent response modulation across cell classes in macaque visual area v4. *Neuron* 55, 131–141. doi: 10.1016/j.neuron.2007.06.018

Moore, T., and Zirnsak, M. (2017). Neural mechanisms of selective visual attention. *Annu. Rev. Psychol.* 68, 47–72. doi: 10.1146/annurev-psych-122414-033400

Moreno-Bote, R., Renart, A., and Parga, N. (2008). Theory of input spike auto-and cross-correlations and their effect on the response of spiking neurons. *Neural Comput.* 20, 1651–1705. doi: 10.1162/neco.2008.03-07-497

Mosheiff, N., Ermentrout, B., and Huang, C. (2023). Chaotic dynamics in spatially distributed neuronal networks generate population-wide shared variability. *PLoS Comput. Biol.* 19, 1–20. doi: 10.1371/journal.pcbi.1010843

Nesse, W. H., Bahmani, Z., Clark, K., and Noudoost, B. (2021a). Differential contributions of inhibitory subnetwork to visual cortical modulations identified via computational model of working memory. *Front. Comput. Neurosci.* 15, 632730. doi: 10.3389/fncom.2021.632730

Nesse, W. H., Maler, L., and Longtin, A. (2021b). Enhanced signal detection by adaptive decorrelation of interspike intervals. *Neural Comput.* 33, 341–375. doi: 10.1162/neco_a_01347

Palmigiano, A., Geisel, T., Wolf, F., and Battaglia, D. (2017). Flexible information routing by transient synchrony. *Nat. Neurosci.* 20, 1014–1022. doi: 10.1038/nn.4569

Reynolds, J. H., and Heeger, D. J. (2009). The normalization model of attention. *Neuron* 61, 168–185. doi: 10.1016/j.neuron.2009.01.002

Roberts, M. J., Lowet, E., Brunet, N. M., Wal, M. T., Tiesinga, P., Fries, P., et al. (2013). Robust gamma coherence between macaque v1 and v2 by dynamic frequency matching. *Neuron* 78, 523–536. doi: 10.1016/j.neuron.2013.03.003

Sadeh, S., and Clopath, C. (2021). Inhibitory stabilization and cortical computation. *Nat. Rev. Neurosci.* 22, 21–37. doi: 10.1038/s41583-020-00390-z

Sanzeni, A., Histed, M. H., and Brunel, N. (2020). Response nonlinearities in networks of spiking neurons. *PLoS Comput. Biol.* 16, 1–27. doi: 10.1371/journal.pcbi.1008165

Shriki, O., Hansel, D., and Sompolinsky, H. (2003). Rate models for conductance-based cortical neuronal networks. *Neural Comput.* 15, 1809–1841. doi: 10.1162/08997660360675053

Siegel, M., Donner, T. H., Oostenveld, R., Fries, P., and Engel, A. K. (2008). Neuronal synchronization along the dorsal visual pathway reflects the focus of spatial attention. *Neuron* 60, 709–719. doi: 10.1016/j.neuron.2008.09.010

Stein, R. B. (1965). A theoretical analysis of neuronal variability. *Biophys. J.* 5, 173–194. doi: 10.1016/S0006-3495(65)86709-1

van Kerkoerle, T., Self, M. W., Dagnino, B., Gariel-Mathis, M.-A., Poort, J., van der Togt, C., et al. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proc. Nat. Acad. Sci.* 111, 14332–14341. doi: 10.1073/pnas.1402773111

Vinck, M., Womelsdorf, T., Buffalo, E., Desimone, R., and Fries, P. (2013). Attentional modulation of cell-class-specific gamma-band synchronization in awake monkey area v4. *Neuron* 80, 1077–1089. doi: 10.1016/j.neuron.2013.08.019

Wilson, H., and Cowan, J. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* 12, 1–24. doi: 10.1016/S0006-3495(72)86068-5

Womelsdorf, T., and Everling, S. (2015). Long-range attention networks: circuit motifs underlying endogenously controlled stimulus selection. *Trends Neurosci.* 38, 682–700. doi: 10.1016/j.tins.2015.08.009

Xiang, Z., Huguenard, J. R., and Prince, D. A. (1998). Gabaa receptor-mediated currents in interneurons and pyramidal cells of rat visual cortex. *J. Physiol.* 506, 715–730. doi: 10.1111/j.1469-7793.1998.715bv.x

Zaksas, D., and Pasternak, T. (2006). Directional signals in the prefrontal cortex and in area MT during a working memory for visual motion task. *J. Neurosci.* 26, 11726–42. doi: 10.1523/JNEUROSCI.3420-06.2006