

Supplementary Methods

Mean field approximation

The mean field approximation provides a simplification of the integrate-and-fire dynamics considered in the spiking simulations, which are computationally expensive and are therefore not suitable for exhaustive parameter explorations. The details of the complete mean field analysis of neural networks with conductance-based synaptic inputs can be found in Brunel and Wang [1]. In this work, we have used the mean field approximation therein derived to scan the parameter space $\{\omega_I, \omega_+\}$. Following this approximation, the stationary dynamics of each neural population is described by the population transduction function. This provides the average population rates after a period of dynamical transients as a function of the average input current. Such stationary rates can be found as the stable solutions of a self-consistency condition and correspond to the stable attractors of the structured network.

In this formulation the potential of a neuron is calculated as:

$$\tau_x \frac{dV(t)}{dt} = -V(t) + \mu_x + \sigma_x \sqrt{\tau_x} \eta(t) \quad (\text{S.1})$$

where $V(t)$ is the membrane potential, x labels the populations (i.e. activated selective subpopulations, non-activated selective subpopulations, nonselective cells and inhibitory cells), τ_x is the effective membrane time constant, μ_x is the mean value the membrane potential would have in the absence of spiking and fluctuations, σ_x measures the magnitude of the fluctuations and η is a Gaussian process with absolute exponentially decaying correlation function and time constant τ_{AMPA} . The magnitudes μ_x and σ_x are given by the following equations:

$$\mu_x = \frac{(T_{\text{ext}}\nu_{\text{ext}} + T_{\text{AMPA}}n_x^{\text{AMPA}} + \rho_1 n_x^{\text{NMDA}})V_E + \rho_2 n_x^{\text{NMDA}}\langle V \rangle + T_I n_x^{\text{GABA}}V_I + V_L}{S_x} \quad (\text{S.2})$$

$$\sigma_x^2 = \frac{g_{\text{AMPA,ext}}^2 (\langle V \rangle - V_E)^2 N_{\text{ext}}\nu_{\text{ext}}\tau_{\text{AMPA}}^2 \tau_x}{g_m^2 \tau_m^2} \quad (\text{S.3})$$

In these equations ν_{ext} corresponds to the external incoming spiking rate, ν_I is the spiking rate of the inhibitory population, and $\tau_m = C_m/g_m$ (whose value depends on whether excitatory or inhibitory cells are considered, see Table 1 in main text). The rest of variables are given by:

$$S_x = 1 + T_{\text{ext}}\nu_{\text{ext}} + T_{\text{AMPA}}n_x^{\text{AMPA}} + (\rho_1 + \rho_2)n_x^{\text{NMDA}} + T_I n_x^{\text{GABA}} \quad (\text{S.4})$$

$$\tau_x = \frac{C_m}{g_m S_x} \quad (\text{S.5})$$

$$T_{\text{ext}} = \frac{g_{\text{AMPA,ext}}\tau_{\text{AMPA}}}{g_m} \quad (\text{S.6})$$

$$T_{\text{AMPA}} = \frac{g_{\text{AMPA,rec}}N_E\tau_{\text{AMPA}}}{g_m} \quad (\text{S.7})$$

$$T_I = \frac{g_{\text{GABA}}N_I\tau_{\text{GABA}}}{g_m} \quad (\text{S.8})$$

$$n_x^{\text{AMPA}} = \sum_{j=1}^p f_j \omega_{jx}^{\text{AMPA}} \nu_j \quad (\text{S.9})$$

$$n_x^{\text{GABA}} = \sum_{j=1}^p f_j \omega_{jx}^{\text{GABA}} \nu_j \quad (\text{S.10})$$

$$n_x^{\text{NMDA}} = \sum_{j=1}^p f_j \omega_{jx}^{\text{NMDA}} \Psi(\nu_j) \quad (\text{S.11})$$

$$\Psi(\nu) = \frac{\nu \tau_{\text{NMDA}}}{1 + \nu \tau_{\text{NMDA}}} \left(1 + \frac{1}{1 + \nu \tau_{\text{NMDA}}} \sum_{n=1}^{\infty} \frac{(-\alpha \tau_{\text{NMDA, rise}})^n T_n(\nu)}{(n+1)!} \right) \quad (\text{S.12})$$

$$T_n(\nu) = \sum_{k=0}^n (-1)^k \binom{n}{k} \frac{\tau_{\text{NMDA, rise}} (1 + \nu \tau_{\text{NMDA}})}{\tau_{\text{NMDA, rise}} (1 + \nu \tau_{\text{NMDA}}) + k \tau_{\text{NMDA, decay}}} \quad (\text{S.13})$$

$$\tau_{\text{NMDA}} = \alpha \tau_{\text{NMDA, rise}} \tau_{\text{NMDA, decay}} \quad (\text{S.14})$$

$$\rho_1 = \frac{g_{\text{NMDA}} N_{\text{E}}}{g_m J} \quad (\text{S.15})$$

$$\rho_2 = \beta \frac{g_{\text{NMDA}} N_{\text{E}} (\langle V_x \rangle - V_{\text{E}}) (J - 1)}{g_m J^2} \quad (\text{S.16})$$

$$J = 1 + \gamma \exp(-\beta \langle V_x \rangle) \quad (\text{S.17})$$

$$\langle V_x \rangle = \mu_x - (V_{\text{thr}} - V_{\text{reset}}) \nu_x \tau_x \quad (\text{S.18})$$

where p denotes the number of selective excitatory subpopulations (i.e. eight throughout this study), f_x is the coding level or fraction of neurons in each excitatory subpopulation x , ω_{jx} is the synaptic weight of the connections from population x to population j , ν_x is the spiking rate of the excitatory subpopulation x , $\gamma = [\text{Mg}^{2+}]/3.57$, $\beta = 0.062$, and the average membrane potential $\langle V_x \rangle$ has a value in the range $[-55, -50]$ mV. In this study, the synaptic weights described in Fig. 2 have been used. In particular, in those cases when $\omega_{jx}^{\text{GABA}}$ is different from zero it always takes the value ω_I . Furthermore, $\omega_{jx}^{\text{AMPA}} = \omega_{jx}^{\text{NMDA}}$, which may take the values ω_+ , ω_- or $\omega = 1$, depending on the neural populations x and j which are connected.

The average firing rates of the different populations in the network can be calculated in the mean field approximation from the solution of a set of n nonlinear equations which depend on μ_x and σ_x as follows:

$$\nu_x = \Phi(\mu_x, \sigma_x) \quad x = 1, \dots, n \quad (\text{S.19})$$

where Φ is the transduction function of population x (i.e. the f-I curve), which enables the calculation of the output rate of population x in terms of its inputs, which in turn also depend on the spiking rates of all the rest of the populations. The transduction function is given by:

$$\Phi(\mu_x, \sigma_x) = \left(\tau_{\text{TP}} + \tau_x \int_{\beta(\mu_x, \sigma_x)}^{\alpha(\mu_x, \sigma_x)} du \sqrt{\pi} \exp(\mu^2) [1 + \text{erf}(u)] \right)^{-1} \quad (\text{S.20})$$

$$\alpha(\mu_x, \sigma_x) = \left(\frac{V_{\text{thr}} - \mu_x}{\sigma_x} \right) \left(1 + 0.5 \frac{\tau_{\text{AMPA}}}{\tau_x} \right) + 1.03 \sqrt{\frac{\tau_{\text{AMPA}}}{\tau_x}} - 0.5 \frac{\tau_{\text{AMPA}}}{\tau_x} \quad (\text{S.21})$$

$$\beta(\mu_x, \sigma_x) = \frac{V_{\text{reset}} - \mu_x}{\sigma_x} \quad (\text{S.22})$$

A set of first-order differential equations (Eq. S.23) whose fixed points correspond to solutions of Eq. S.19 are used to solve the entire set of equations.

$$\tau_x \frac{d\nu_x}{dt} = -\nu_x + \Phi(\mu_x, \sigma_x) \quad (\text{S.23})$$

The network dynamics could converge to different attractors states corresponding to multi-item memory states or m -memory states, with m being the number of coactivated subpopulations. The working point was selected based on a systematic scan of the parameter space $\{\omega_I, \omega_+\}$ with the aim to reproduce the experimental data reported in the literature. To this end, a set of initial conditions obtained by varying the number of initially coactivated populations both in the saliency and no saliency condition were considered (see Fig. 2). In order to exhaustively search the possible fixed points for the selected working point $\{\omega_I = 1.15, \omega_+ = 2.2\}$, we integrated Eq. S.23 with 1000 different initial conditions (selected randomly except from those already included in the previous study) for the population firing rates. The same upper working memory capacity limit (i.e. 4 items) found in Fig. 2 was obtained in this study.

References

1. Brunel N, Wang XJ (2001) Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *J Comput Neurosci* 11: 63-85.