Distributed Sampling-based Bayesian Inference in Coupled Neural Circuits

Wen-Hao Zhang[1], **Tai Sing Lee**[2], **Brent Doiron**[1,3], **Si Wu**[4]
wenhao.zhang@pitt.edu; tai@cnbc.cmu.edu; bdoiron@pitt.edu; siwu@pku.edu.cn
[1]Department of Mathematics, University of Pittsburgh.
[2]Computer Science Department and Neuroscience Institute, Carnegie Mellon University.
[3]Departments of Neurobiology and Statistics, Grossman Center for Quantitative Biology and Human Behavior, University of Chicago.
[4]School of Electronics Engineering & Computer Science, IDG/McGovern Institute for Brain Research, Peking-Tsinghua Center for Life Sciences, Peking University.

Abstract

The brain performs probabilistic inference to interpret the external world, but the underlying neuronal mechanisms remain not well understood. The stimulus structure of natural scenes exists in a high-dimensional feature space, and how the brain represents and infers the joint posterior distribution in this rich, combinatorial space is a challenging problem. There is added difficulty when considering the neuronal mechanics of this representation, since many of these features are computed in parallel by distributed neural circuits. Here, we present a novel solution to this problem. We study continuous attractor neural networks (CANNs), each representing and inferring a stimulus attribute, where attractor coupling supports sampling-based inference on the multivariate posterior of the high-dimensional stimulus features. Using perturbative analysis, we show that the dynamics of coupled CANNs realizes Langevin sampling on the stimulus feature manifold embedded in neural population responses. In our framework, feedforward inputs convey the likelihood, reciprocal connections encode the stimulus correlational priors, and the internal Poisson variability of the neurons generate the correct random walks for sampling. Our model achieves high-dimensional joint probability representation and Bayesian inference in a distributed manner, where each attractor network infers the marginal posterior of the corresponding stimulus feature. The stimulus feature can be read out simply with a linear decoder based only on local activities of each network. Simulation experiments confirm our theoretical analysis. The study provides insight into the fundamental neural mechanisms for realizing efficient high-dimensional probabilistic inference.

1 Introduction

The theory that the brain performs probabilistic inference to interpret the external world [1–3] has been supported by extensive human behavioral [4–6] and animal neurophysiological studies [7, 8]. Yet, exactly how neural circuits in the brain realize Bayesian inference remains poorly understood. To address this question, we need to answer how probabilistic information is represented in neural responses and what inference algorithms are adopted by neural circuits. An added difficulty is that signals in the world are high-dimensional. The brain often extracts and represents multiple stimulus variables in a parallel, distributed, and hierarchical fashion using separate neural circuits. Examples include: integrating multisensory cues [8], resolving the consistency of local percepts (e.g. nose and eyes) and global percepts (e.g. face) in a hierarchy [1, 9], and grouping orientation edge neurons into contour [10, 11]. The brain needs to represent and infer the joint posterior distribution of all these

Preprint. Under review.

variables in different locations, different levels, or even different sensory systems. This fundamental challenge must be addressed in the context of other properties and constraints of Bayesian inference in the brain.

The neural mechanics of inference has been an active research area, with a number of existing computational models. A popular proposal is that of probabilistic population coding (PPC), whereby a deterministic network integrates feedforward Poissonian inputs to parametrically infer and represent the posterior of a stimulus feature [12]. This model was later extended to process high-dimensional stimulus features by coupling networks together, where the coupling encodes the correlation prior between stimulus features [13]. Such a model scales linearly, rather than combinatorially, with the number of stimulus features. This allows each feature's marginal distribution to be computed and readout from each network. Despite its conceptual appeal, the interaction in such a coupled PPC network is highly nonlinear, requiring a complex circuit with multiplicative and divisive operations. An alternative paradigm to represent the joint posterior of multiple variables is using non-parametric sampling (e.g., [14–20]), which can be mediated by linear stochastic dynamics (e.g., [16]). However, these studies typically consider sampling the posterior of neuronal responses, without specifying how stimulus features embedded in neuronal responses are sampled (e.g., [14–16]), or they considered sampling stimulus features but did not specify a concrete neural circuit model to enact the sampling [18, 21]. Moreover, when sampling occurs in the very high-dimensional neural response space, involving neurons in different hypercolumns, visual areas, and even different sensory systems, the sampling timescale becomes a serious obstacle [16, 22], which must be overcome to produce perception in a biologically realistic time frames.

In this study, we propose a novel model to represent and infer the posterior of high-dimensional stimulus features. Our model combines the strengths of PPC and sampling-based codes in a unified framework. It decomposes the high dimensional parameter space into distinct variables, represented individually by distinct continuous attractor neural networks (CANNs). In our model feedforward inputs convey the likelihood via PPC, excitatory reciprocal connections between the attractors encode the stimulus correlational priors, and the internal Poisson variability of the neurons generate the correct variability for sampling on the stimulus feature space. We analytically show that the dynamics of coupled CANNs in fact realizes Langevin sampling on the stimulus feature manifold embedded in neural population responses. The model achieves sampling-based Bayesian inference in a distributed attractor network, each of which infers the marginal posterior of the corresponding stimulus feature, and its local activities allow readout of the stimulus feature using a linear decoder. Simulation results confirmed our theoretical analysis of the model.

2 The generative model and sampling-based inference

2.1 The probabilistic generative model

We consider a linear Gaussian generative model (Fig. 1A), in which the observed stimulus features $\mathbf{x} = \{x_m\}_{m=1}^M \in \mathbb{R}^M$ are independently generated by the external latent features $\mathbf{s} = \{s_m\}_{m=1}^M \in \mathbb{R}^M$. These features can be any stimulus features that are extracted by the cortex, such as line orientation, moving direction etc. The likelihood function $p(\mathbf{x}|\mathbf{s})$ and the prior $p(\mathbf{s})$ of the generative model are

$$p(\mathbf{x}|\mathbf{s}) = \mathcal{N}(\mathbf{x}|\mathbf{s}, \mathbf{\Lambda}^{-1}), \quad p(\mathbf{s}) \propto \mathcal{N}(\mathbf{s}|0, \mathbf{L}^{-1}), \tag{1}$$

where $\mathcal{N}(\cdot)$ denotes a Gaussian distribution. $\mathbf{\Lambda}$ is the precision matrix (the inverse of the covariance matrix) of the likelihood function, which is assumed to be diagonal, i.e., $\mathbf{\Lambda} = \text{diag}(\Lambda_1, \Lambda_2, \cdots, \Lambda_M)$. This implies that each observed feature x_m is independently generated by s_m, giving $p(\mathbf{x}|\mathbf{s}) = \prod_{m=1}^M p(x_m|s_m) = \prod_{m=1}^M \mathcal{N}(x_m|s_m, \Lambda_m^{-1})$.

The precision matrix $\mathbf{L}$ of the prior $p(\mathbf{s})$ is a generalized Laplacian matrix, with $\mathrm{L}_{mm} = -\sum_n \mathrm{L}_{mn}$ and $\mathrm{L}_{mn} = \mathrm{L}_{nm} \leq 0$ $(m \neq n)$. For example, in the case of two-dimensional features ($M = 2$), the prior $p(\mathbf{s}) = p(s_1, s_2)$ is written as,

$$p(s_1, s_2) = \frac{\sqrt{\mathrm{L}_{12}}}{\sqrt{2\pi}\mathrm{w}_s} \exp\left[-\frac{\mathrm{L}_{12}}{2}\mathbf{s}^\top \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \mathbf{s}\right] = \frac{\sqrt{\mathrm{L}_{12}}}{\sqrt{2\pi}\mathrm{w}_s} \exp\left[-\frac{\mathrm{L}_{12}}{2}(s_1 - s_2)^2\right],$$

where w_s is the width of stimulus feature space, e.g., $\mathrm{w}_s = 2\pi$ if s is direction of stimulus motion. Priors with a Laplacian precision matrix have been *implicitly* considered in the modelling studies for

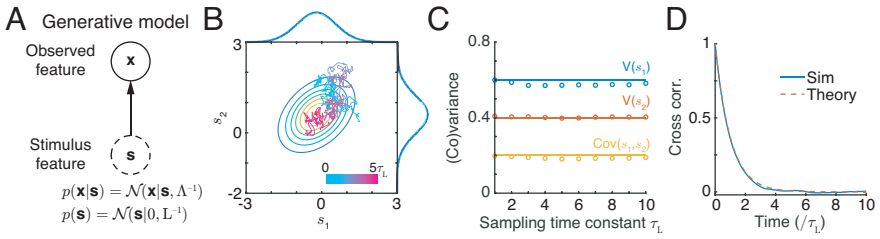

Figure 1: A probabilistic generative model and its inference by Langevin sampling. (A) A linear Gaussian generative model, where $\mathbf{s}$ and $\mathbf{x}$ are the latent and observed stimulus features, respectively. $\mathbf{\Lambda}$ and $\mathbf{L}$ are precision matrices of the likelihood function and the prior, respectively. (B) Langevin sampling for approximating the posterior of two-dimensional features. Solid ellipses: the posterior. Colors of the trajectory indicate the time elapsed. Marginal plot: marginal posterior (empirical: solid line; theory: shaded line). (C) The equilibrium variance of samples vs. the sampling time constant τ_L (empirical: circle; theory: solid line). (D) The temporal correlation of sampling over time.

multisensory cue integration [23–25], and it quantifies the co-occurrence of stimulus features, e.g., L_{12} characterizes the correlation between s_1 and s_2. Notably, according to Eq. 1, the marginal prior of each stimulus feature is uniform, i.e., $p(s_m) = 1/\mathrm{w}_s$, a property coming from that the determinant of the Laplacian matrix is zero, i.e., $|\mathbf{L}| = 0$.

According to Bayes' theorem, the posterior distribution of the latent variables $\mathbf{s}$ given the observed features $\mathbf{x}$ are calculated by inverting the generative model (Eq. 1),

$$p(\mathbf{s}|\mathbf{x}) \propto p(\mathbf{x}|\mathbf{s})p(\mathbf{s}) = \mathcal{N}(\mathbf{s}|\boldsymbol{\mu}_\mathbf{s}, \mathbf{\Omega}^{-1}), \tag{2}$$

which is a multivariate Gaussian distribution, with the precision matrix $\mathbf{\Omega}$ and mean $\boldsymbol{\mu}_\mathbf{s}$ given by,

$$\mathbf{\Omega} = \mathbf{\Lambda} + \mathbf{L}, \quad \boldsymbol{\mu}_\mathbf{s} = \mathbf{\Omega}^{-1}\mathbf{\Lambda}\mathbf{x}. \tag{3}$$

2.2 Langevin sampling for approximate inference

Numerous psychophysical studies suggest that the brain infers the external world via Bayesian inference (e.g., [2, 3, 26]), but it remains unresolved how the inference is implemented by cortical circuits in the brain. Here, we propose that the brain employs a sampling strategy to approximate the inference of stimulus features, and later we demonstrate that this is feasible in a biologically plausible neural circuit. The sampling strategy we consider is Langevin sampling, which is a Markov chain Monte Carlo (MCMC) algorithm widely used to numerically approximate the posterior [16, 27, 28]. The dynamics of Langevin sampling performs stochastic gradient ascent on the manifold of the log-posterior of stimulus features, which is written as,

$$\frac{d\mathbf{s}_t}{dt} = (2\tau_L)^{-1}\frac{d\ln p(\mathbf{s}|\mathbf{x})}{d\mathbf{s}} + \sqrt{\tau_L^{-1}}\boldsymbol{\xi}_t = -(2\tau_L)^{-1}\left[(\mathbf{L}+\mathbf{\Lambda})\mathbf{s}_t - \mathbf{\Lambda}\mathbf{x}\right] + \sqrt{\tau_L^{-1}}\boldsymbol{\xi}_t, \tag{4}$$

where τ_L is the time constant of sampling. $\boldsymbol{\xi}_t$ are multivariate independent Gaussian-white noises, satisfying $\langle \boldsymbol{\xi}_t \boldsymbol{\xi}_{t'}^\top \rangle = \mathbf{I}\delta(t-t')$, with $\mathbf{I}$ the identity matrix and $\delta(t-t')$ the Dirac delta function, and they induce fluctuations of $\mathbf{s}_t$ necessary for sampling. It can be checked that the equilibrium distribution of $\mathbf{s}_t$ has the same form as the posterior (Eqs. 2-3), since its equilibrium mean and covariance satisfy (see details in Supplementary Information (SI.) Sec. 2),

$$\bar{\mathbf{s}} \equiv \langle \mathbf{s}_t \rangle = \boldsymbol{\mu}_\mathbf{s}, \quad \mathbf{\Sigma}_\mathbf{s} \equiv \left\langle (\mathbf{s}_t - \bar{\mathbf{s}})(\mathbf{s}_t - \bar{\mathbf{s}})^\top \right\rangle = \mathbf{\Omega}^{-1}, \tag{5}$$

where $\langle \cdot \rangle$ denotes averaging over trials. Thus, in the dynamics of Langevin sampling, the instant value of $\mathbf{s}_t$ can be regarded as a sample from the posterior specified by Eqs. (2-3). Fig. 1B shows an example trajectory of Langevin sampling of two-dimensional stimulus features $\mathbf{s}_t$ over time, demonstrating that the sampled $\mathbf{s}$ indeed satisfies the posterior as expected. The sampling time constant τ_L doesn't affect the equilibrium covariance (Eq. 5), but only affects the temporal correlation of sampling (i.e., the normalized cross-correlation of samples at time t_1 and t_2, which is given by $\rho_\mathbf{s}(t_1, t_2) = e^{-|t_1-t_2|/2\tau_L}$), and these two theoretical predictions are also confirmed by simulations (Fig. 1C-D).

3 Distributed sampling-based inference in coupled attractor networks

3.1 Coupled continuous attractor neural networks

We explore how a canonical neural circuit model implements Langevin sampling to approximate the posterior of high-dimensional stimulus features. The model we consider is composed of M reciprocally connected continuous attractor neural networks (CANNs), and they interact with each other to achieve inference in a distributed manner, such that each network m infers the *marginal* posterior of one stimulus feature $p(s_m|\mathbf{x})$. For simplicity, we consider that each network m has the same number of N neurons with preferred stimulus values $\{\theta_j\}_{j=1}^N$, with θ_j being the preference of the j-th neuron. We take $\{\theta_j\}_{j=1}^N$ to uniformly cover the feature space of s_m, where $\theta \in (-\pi, \pi]$ satisfying the periodic boundary condition.

CANNs are a canonical neural circuit model widely used to elucidate the network mechanism underlying many brain functions (see e.g., [29–31]). In the continuum limit ($\theta_j \to \theta$), the dynamics of coupled CANNs is written as,

$$\tau \frac{\partial \mathbf{u}_{mt}(\theta)}{\partial t} = -\mathbf{u}_{mt}(\theta) + \rho \sum_n \mathbf{W}^{\mathsf{r}}_{mn}(\theta) * \mathbf{r}_{nt}(\theta) + \rho \mathbf{W}^{\mathsf{f}}_{m}(\theta) * \mathbf{I}^{\mathsf{f}}_{m}(\theta) + \sqrt{\tau \mathsf{F} \mathbf{u}_{mt}(\theta)}\, \boldsymbol{\xi}_{mt}(\theta), \quad (6)$$

where $\mathbf{u}_{mt}(\theta)$ and $\mathbf{r}_{mt}(\theta)$ represent, respectively, the synaptic inputs and firing rates of neurons at time t in network m whose preferred value of stimulus feature s_m is θ. τ is the time constant, and $\rho = N/\mathsf{w}_s$ is the neuronal density covering the stimulus feature space. F is the Fano factor of the internal Poisson-like variability. $\mathbf{W}^{\mathsf{r}}_{mm}$ denotes the recurrent connection kernel between neurons in the same network, $\mathbf{W}^{\mathsf{r}}_{mn}$ for $m \neq n$ is the reciprocal connection kernel between neurons from network n to network m, and $\mathbf{W}^{\mathsf{f}}_{m}$ is the feedforward connection kernel. These kernels have the form of Gaussian function and are written as,

$$\mathbf{W}^{\mathsf{r}}_{mn}(\theta) = \mathsf{w}^{\mathsf{r}}_{mn} \boldsymbol{g}(\theta), \quad \mathbf{W}^{\mathsf{f}}_{m}(\theta) = \mathsf{w}^{\mathsf{f}} \boldsymbol{g}(\theta), \quad \boldsymbol{g}(\theta) = (\sqrt{2\pi}a)^{-1} \exp\left(-\theta^2/2a^2\right). \quad (7)$$

The symbol $*$ denotes the convolution, i.e., $\mathbf{W}(\theta) * \mathbf{r}(\theta) = \int \mathbf{W}(\theta - \theta')\mathbf{r}(\theta')d\theta'$, which implies the translation-invariant property of the connection pattern between neurons over the feature space, a key characteristic of CANNs. For simplicity, we assume the feedforward connection weight w^{f} is the same in different networks.

The relationship between the synaptic input and firing rate of neurons is modelled as divisive normalization [32, 33], which is given by

$$\mathbf{r}_{mt}(\theta) = \frac{[\mathbf{u}_{mt}(\theta)]_+^2}{1 + k\rho \int [\mathbf{u}_{mt}(\theta')]_+^2 d\theta'}. \quad (8)$$

Here k determines the global inhibition strength and $[\cdot]_+$ is negative rectification. Divisive normalization is a canonical operation widely observed in the cortex and could be implemented via parvalbumin (PV) inhibitory neurons [34].

Feedforward input encoding the stimulus likelihood

In our model each network m receives one feedforward input $\mathbf{I}^{\mathsf{f}}_m$, which conveys information about the latent stimulus feature s_m (Fig. 2A). Given the stimulus feature s_m, $\mathbf{I}^{\mathsf{f}}_m$ is modeled as independent Poisson spikes with Gaussian tuning (mean firing rate) (Fig. 2C). Mathematically, the probability of observing a particular value of $\mathbf{I}^{\mathsf{f}}_m$ given s_m is,

$$\mathbf{I}^{\mathsf{f}}_m | s_m \sim \prod_{j=1}^N \text{Poisson}[\lambda_{m,j}(s_m)], \quad \lambda_{m,j}(s_m) = \mathrm{I}^{\mathsf{f}}_m \exp[-(\theta_j - s_m)^2/2a^2], \quad (9)$$

where $\lambda_{m,j}(s_m)$ is the mean firing rate of the input component $\mathbf{I}^{\mathsf{f}}_{m,j}$ received by j-th neuron in network m, with $\mathrm{I}^{\mathsf{f}}_m$ and a characterizing the peak input rate and tuning width, respectively. The above feedforward input has been widely used in previous neural coding studies, and satisfies the linear probabilistic population code proposed in [12]. Based on the Gaussian tuning and Poisson variability (Eq. 9), the likelihood of s_m given an observed $\mathbf{I}^{\mathsf{f}}_m$ is also a Gaussian distribution (Fig. 2E), i.e., $p(\mathbf{I}^{\mathsf{f}}_m|s_m) = p(x_m|s_m) = \mathcal{N}(x_m|s_m, \Lambda_m^{-1})$, whose mean and precision are linear over $\mathbf{I}^{\mathsf{f}}_m$ and are calculated to be

$$x_m = \frac{\sum_j \mathbf{I}^{\mathsf{f}}_{m,j}\theta_j}{\sum_j \mathbf{I}^{\mathsf{f}}_{m,j}}, \quad \Lambda_m = a^{-2} \sum_j \mathbf{I}^{\mathsf{f}}_{m,j}. \quad (10)$$

Comparing to Eq. 1, we see that the feedforward input $\mathbf{I}^{\mathsf{f}}_m$ conveys the likelihood of the latent stimulus feature s_m parametrically.

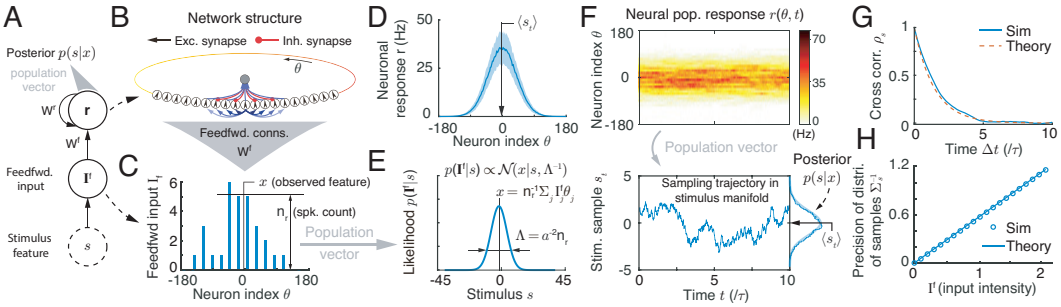

Figure 2: Langevin sampling of one dimensional stimulus feature in a CANN. (A) The information flow of the generative model. (B) The structure of a CANN. (C) The feedforward input is modeled as continuous approximation of Poisson spikes with Gaussian tuning over the stimulus feature. (D) The time averaged neural population responses, with the shaded region denoting standard deviation. (E) The likelihood is encoded by the feedforward input parametrically. (F) Neural population responses of the CANN over time (top); stimulus feature values sampled by the network dynamics (bottom left), whose distribution gives the posterior (bottom right; empirical: solid line, theory: shaded line). The theoretical value of the posterior is obtained from the feedforward input $\mathbf{I}^f$ by using Eq. (10) (see details in SI. Sec. 4.3). (G) Temporal correlation of sampling over time. (H) The precision (inverse of variance) of samples in equilibrium with the input intensity. Parameters can be found in SI. Sec. 4.

Internal Poisson-like variability for reliable sampling

Implementing Langevin sampling in a network requires the network to generate additional internal variability not inherited from the feedforward inputs, which produces random walk on the log-posterior surface of the stimulus features (Eq. 4). Moreover, this internal variability should be Gaussian distributed with a white spectrum, so that the variance of samples (Eq. 5) matches the variance of the posterior ($\mathbf{\Omega}^{-1}$ in Eq. 3).

Encoding the variance of the posterior is necessary for the brain to perceive the uncertainty of inputs [12, 17, 35]. To achieve Langevin sampling in our model, we consider that each network generates independent internal Poisson-like variability at the single neuron level (the last term in Eq. 6, e.g., [36–38]), which makes neural responses $\mathbf{u}_m$ and $\mathbf{r}_m$ exhibit the Poisson-like variability. As will be shown later, this Poisson-like variability in neural activities contributes to the required Gaussian-white variability in the stimulus features space (embedded in neural population responses), ensuring the realization of Langevin sampling. The Poisson-like variability provides a reliable source of variability to conduct sampling, as the Fano factor of neuronal responses only changes mildly across stimulus conditions [7, 12, 39, 40]. In reality, this internal Poisson-like (spiking) variability can naturally arise from the chaotic state of cortical circuit dynamics [40–44].

Network responses and decoding

Given the feedforward input $\mathbf{I}^f$, our theoretical analyses and simulations reveal that in the equilibrium, the mean neuronal responses of each network m have a Gaussian shape (Fig. 2D, i.e., [45, 46]),

$$\langle \mathbf{u}_m(\theta) \rangle = \mathsf{U}_m \exp\left[-(\theta - \bar{s}^{\mathbf{r}}_m)^2/4a^2\right], \quad \langle \mathbf{r}_m(\theta) \rangle = \mathsf{R}_m \exp\left[-(\theta - \bar{s}^{\mathbf{r}}_m)^2/2a^2\right]. \tag{11}$$

Here, U_m and R_m denotes the peak values of synaptic inputs and firing rate respectively. $\bar{s}^{\mathbf{r}}_m$ is the position of the Gaussian center on the stimulus feature space (Fig. 2D), which is the mean of stimulus feature samples in the network dynamics (see below). Because of the Gaussian tuning (Eq. 11) and the Poisson-like variability, given an instantaneous neuronal response $\mathbf{r}_{mt}(\theta)$ in network m, an instantaneous value of the stimulus feature can be efficiently read out using population vector [12, 47], i.e.,

$$s^{\mathbf{r}}_{mt} = \textstyle\sum_j \mathbf{r}_{mt}(\theta_j)\theta_j / \sum_j \mathbf{r}_{mt}(\theta_j), \tag{12}$$

which is regarded as a sample of stimulus s_m (Eq. 14). Notably, the stimulus feature $s^{\mathbf{r}}_{mt}$ is read out based only on local neuronal responses in network m. Moreover, due to the divisive normalization operation in the network dynamics (Eq. 8), the Gaussian profile of neural responses (Eq. 11) can be well maintained even if the input disparity $|x_m - x_n|$ is large (Eq. 9), which ensures that stimulus features can always be efficiently read out by population vector.

3.2 Langevin sampling of stimulus features in coupled networks

We now elucidate how coupled CANNs can realize Langevin sampling of the posterior of high-dimensional stimulus features in a distributed manner. We apply perturbative analysis to derive the dynamics of stimulus features embedded in neural population responses (see details in SI. Sec. 3.2). For each network m, we consider that its instantaneous neural response is perturbed from its equilibrium mean, i.e., $\mathbf{u}_{mt}(\theta) = \langle \mathbf{u}_m(\theta) \rangle + \delta \mathbf{u}_{mt}(\theta)$. The (unnormalized) eigenfunction of the perturbation $\delta \mathbf{u}_{mt}(\theta)$ corresponding to the change of stimulus feature s_m is derived as,

$$\boldsymbol{\phi}(\theta | s^{\mathbf{r}}_m) = a^{-1}(\theta - s^{\mathbf{r}}_m) \exp\left[-(\theta - s^{\mathbf{r}}_m)^2/4a^2\right]. \tag{13}$$

Previous studies have shown this eigenfunction has the largest eigenvalue for the perturbation of CANN state [45]. We project the dynamics of each network m onto the corresponding eigenfunction (Eq. 13), where the projection is simply the inner product between the eigenfunction and the network response, i.e., $\langle \boldsymbol{\phi}, \mathbf{u}_{mt} \rangle = \int \boldsymbol{\phi}(\theta) \mathbf{u}_{mt}(\theta) d\theta$. After projection, we obtain the dynamics on the manifold of stimulus features embedded in neural responses (see details in SI 3.2),

$$\frac{d\mathbf{s}^{\mathbf{r}}_t}{dt} = -\frac{\rho}{\sqrt{2}} (\tau \mathbf{D}_{\mathsf{U}})^{-1} \left[(\mathbf{L}^{\mathbf{r}} + \mathsf{w}^{\mathsf{f}} \mathbf{D}_{\mathsf{f}}) \mathbf{s}^{\mathbf{r}}_t - \mathsf{w}^{\mathsf{f}} \mathbf{D}_{\mathsf{f}} \mathbf{x} \right] + \sigma_s \sqrt{(\tau \mathbf{D}_{\mathsf{U}})^{-1}} \boldsymbol{\xi}_t, \tag{14}$$

where $\mathbf{s}^{\mathbf{r}}_t$ denotes the instantaneous positions of neural activity bumps at time t, which is regarded as a sample of stimulus features by the coupled networks (Eq. 12). $\mathbf{x}$ is the observed stimulus feature (Eq. 1) conveyed by the feedforward inputs (Eq. 9, Fig. 2E). $\mathbf{L}^{\mathbf{r}}$ is a generalized Laplacian matrix with $\mathrm{L}^{\mathbf{r}}_{mn} = -\mathsf{w}^{\mathsf{r}}_{mn} \mathsf{R}_n$, $\mathrm{L}^{\mathbf{r}}_{mm} = -\sum_{n \neq m} \mathrm{L}^{\mathbf{r}}_{mn}$, where R_n is the peak firing rate of network n (Eq. 11). w^{f} is a scalar variable denoting the feedforward connection strength (Eq. 7). $\mathbf{D}_{\mathsf{U}} = \mathrm{diag}(\mathsf{U}_1, \mathsf{U}_2, \cdots, \mathsf{U}_M)$ and $\mathbf{D}_{\mathsf{f}} = \mathrm{diag}(\mathrm{I}^{\mathsf{f}}_1, \mathrm{I}^{\mathsf{f}}_2, \cdots, \mathrm{I}^{\mathsf{f}}_M)$ are diagonal matrices, denoting the peak value of the mean synaptic input, and the mean feedforward input in each network respectively (Eqs. 9 and 11). $\mathbf{D}_{\mathsf{U}}$ can be analytically solved in our model (Eqs. S13-S14). Notably, after projection, the internal Poisson-like variability of neuronal activities (Eq. 6) becomes the Gaussian-white variability of stimulus features (the last term in Eq. 14), as required by Langevin sampling. The strength of Gaussian-white variability in the stimulus feature manifold is $\sigma_s^2 = 8a\mathsf{F}/(3\sqrt{3\pi})$, which is unchanged with respect to the feedforward input and network responses.

It is interesting to compare the dynamics on stimulus features (Eq. 14) with that of Langevin sampling (Eq. 4): both equations contain a Laplacian matrix ($\mathbf{L}$ or $\mathbf{L}^{\mathbf{r}}$) representing the stimulus prior, a diagonal matrix ($\boldsymbol{\Lambda}$ or $\mathbf{D}_{\mathsf{f}}$) representing the precision matrix of the likelihood, and Gaussian-white noises for producing a random walk. Thus, the dynamics of coupled CANNs in effect realizes Langvein sampling in the manifold of stimulus features. The equilibrium distribution of stimulus features $\mathbf{s}^{\mathbf{r}}$ in Eq. (14) is a multivariate Gaussian, denoted as $\mathcal{N}(\mathbf{s}^{\mathbf{r}} | \bar{\mathbf{s}}^{\mathbf{r}}, \Sigma^{\mathbf{r}}_{\mathbf{s}})$, whose mean $\bar{\mathbf{s}}^{\mathbf{r}}$ and covariance of $\Sigma^{\mathbf{r}}_{\mathbf{s}}$ satisfy two conditions (for details, see SI. Sec. 2),

$$(\mathbf{L}^{\mathbf{r}} + \mathsf{w}^{\mathsf{f}} \mathbf{D}_{\mathsf{f}}) \, \bar{\mathbf{s}}^{\mathbf{r}} = \mathsf{w}^{\mathsf{f}} \mathbf{D}_{\mathsf{f}} \mathbf{x}, \tag{15}$$

$$(\mathbf{L}^{\mathbf{r}} + \mathsf{w}^{\mathsf{f}} \mathbf{D}_{\mathsf{f}}) \boldsymbol{\Sigma}^{\mathbf{r}}_{\mathbf{s}} + \boldsymbol{\Sigma}^{\mathbf{r}}_{\mathbf{s}} (\mathbf{L}^{\mathbf{r}} + \mathsf{w}^{\mathsf{f}} \mathbf{D}_{\mathsf{f}})^{\top} = \sqrt{2} \sigma_s^2 \rho^{-1}, \tag{16}$$

where $\bar{\mathbf{s}}^{\mathbf{r}} = \langle \mathbf{s}^{\mathbf{r}}_t \rangle$, and $\Sigma^{\mathbf{r}}_{\mathbf{s}} = \left\langle (\mathbf{s}^{\mathbf{r}}_t - \bar{\mathbf{s}}^{\mathbf{r}})(\mathbf{s}^{\mathbf{r}}_t - \bar{\mathbf{s}}^{\mathbf{r}})^{\top} \right\rangle$. By setting the equilibrium mean and covariance of stimulus features sampled by the network to be equal to that of the posterior of latent stimulus features, i.e., $\bar{\mathbf{s}}^{\mathbf{r}} = \boldsymbol{\mu}_{\mathbf{s}}$ and $\boldsymbol{\Sigma}^{\mathbf{r}}_{\mathbf{s}} = \boldsymbol{\Omega}^{-1}$ (see Eq. 3), we get the required network connections. It can be checked that $\mathbf{L}^{\mathbf{r}} + \mathsf{w}^{\mathsf{f}} \mathbf{D}_{\mathsf{f}} = \sigma_s^2 (\sqrt{2}\rho)^{-1} \boldsymbol{\Omega}$ can make $\boldsymbol{\Sigma}^{\mathbf{r}}_{\mathbf{s}} = \boldsymbol{\Omega}^{-1}$ hold. Substituting this into Eqs. (15-16), we get the network connections for realizing Langevin sampling of the posterior of stimulus features, which are,

$$\mathsf{w}^{\mathsf{r}}_{mn} = \frac{a \mathsf{w}^{\mathsf{f}}}{\sqrt{2\pi}\rho} \frac{\mathrm{L}_{mn}}{\mathsf{R}_n}, \quad \mathsf{w}^{\mathsf{f}} = \frac{\sqrt{\pi}}{a} \sigma_s^2 = \left(\frac{2}{\sqrt{3}} \right)^3 \mathsf{F}. \tag{17}$$

Interestingly, we see that the reciprocal connections $\mathsf{w}^{\mathsf{r}}_{mn}$ and $\mathsf{w}^{\mathsf{r}}_{nm}$ between networks n and m encode the prior (the correlation) L_{mn} between stimulus features s_m and s_n (Eq. 1). Moreover, although the prior precision matrix is symmetric, i.e., $\mathrm{L}_{mn} = \mathrm{L}_{nm}$, the couplings $\mathsf{w}^{\mathsf{r}}_{mn}$ and $\mathsf{w}^{\mathsf{r}}_{nm}$ can be asymmetric, which is biologically more plausible. When $\mathrm{L}_{mn} = 0$ for $m \neq n$, the prior $p(\mathbf{s})$ degenerates into uniform, and the reciprocal connections $\mathsf{w}^{\mathsf{r}}_{mn} = 0$, for $m \neq n$. In such a case, each network individually samples its marginal posterior which is proportional to the likelihood.

In summary, we have shown that the model of coupled CANNs with appropriate feedforward inputs (conveying the likelihoods, Eq. 9), appropriate reciprocal connections (encoding the prior, Eq. 17),

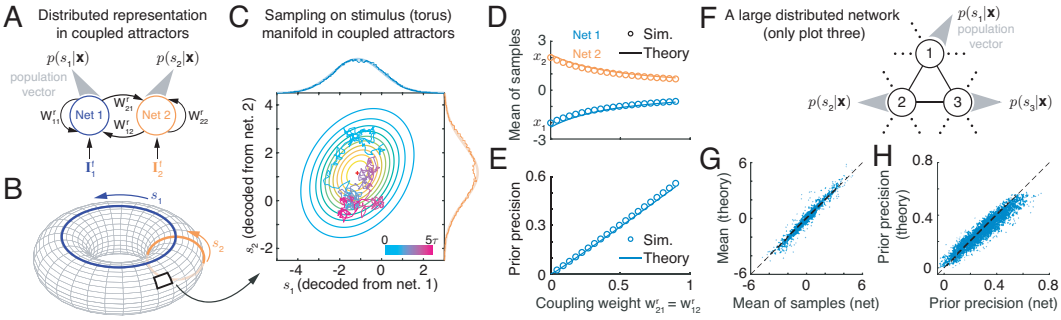

Figure 3: Distributed Langevin sampling in coupled CANNs. (A) The structure of two-coupled CANNs. Each network receives a feedforward input conveying the likelihood of stimulus feature and meanwhile is reciprocally connected to other networks. (B) The torus manifold of two-dimensional stimulus features embedded in two coupled CANNs. (C) The model of two-coupled CANNs implements Langevin sampling of the posterior of stimulus features, for each network inferring the corresponding marginal posterior. Solid ellipses: the posterior by theory. Colors of the trajectory indicate the time elapsed. Marginal plot: marginal posterior (empirical: solid line; theory: shaded line). (D) The mean of samples in each network is consistent with the mean of the posterior. (E) The coupling strengths encode the precision of the prior (for detailed calculations, see SI. Sec. 4.4). (F) A large network contains up to ten coupled CANNs. (G-H) The comparisons of the mean of sampling (G) and the prior precision stored in the network (H) with theoretical predictions. Each point is a result obtained under a combination of different inputs, different connection weights, and different realizations of the network numbers. Parameters can be found in SI. Sec. 4.

and appropriate internal variability (producing random walks, Eq. 14) can implement Langevin sampling on the manifold of high-dimensional latent stimulus features. Further, this is done in a distributed manner, in term of that each network infers the marginal posterior of one stimulus feature. Finally, the sampled stimulus features can be read out easily by using a linear decoder (population vector) and separately from local neural activities at each network (see Eq. 12).

4 Simulation experiments

We carry out simulations to validate the above theoretical analysis. We first consider a single CANN (equivalent to an uncoupled CANN; Fig. 2A-B). In this case each network m independently infers a posterior $p(s_m|x_m)$ based on the likelihood $p(x_m|s_m)$ and the uniform stimulus prior. We present a constant feedforward input $\mathbf{I}^{\mathrm{f}}_m(\theta)$ generated by Eq. 9 to the network, and then evaluate the sampling performance of the network (more details can be found in SI. Sec. 4). Firstly, we observe that due to the internal variability, the neural population responses fluctuate over time (Fig. 2F), whose time-averaged profile has a Gaussian shape and individual neuronal responses exhibit Poisson-like variability (Eq. 11, Fig. 2D). This guarantees that the instant sample of stimulus feature $s^{\mathbf{r}}_t$ can be efficiently read out from the instantaneous neuronal response $\mathbf{r}_t$ using population vector (Eq. 12). Secondly, we observe that the network performs Langevin sampling. As shown in Fig. 2F (bottom), the stimulus feature $s^{\mathbf{r}}_t$ sampled by the network shows random walk behavior over time, a characteristic of Langevin sampling. Moreover, the temporal correlation of sampling agrees with the theoretical prediction of Langevin sampling, i.e., $\rho_s(\Delta t) = \exp(-\rho\Delta t/\sqrt{2}\tau \mathrm{U})$ (Fig. 2G). Thirdly, we observe that by setting the feedforward weight as in Eq. 17, the equilibrium distribution of sampled stimulus feature (i.e., the distribution of sequential samples in equilibrium, Fig. 2F, bottom) is consistent with the posterior (Fig. 2H). Thus, a single CANN with internal Poisson-like variability can implement Langevin sampling on the stimulus feature manifold.

Next, we demonstrate that coupled CANNs implement Langevin sampling of the posterior of high-dimensional stimulus features. We first consider a model of two coupled networks (Fig. 3A), in which the stimulus manifold is a torus, with each circle representing a stimulus feature (Fig. 3B). Similar to the case of a single CANN, neuronal responses at each network exhibits Langevin sampling behaviors due to the internal variability. The instant stimulus estimate $s^{\mathbf{r}}_{mt}$ can be read out using population vector based only on the local instantaneous neural activity $\mathbf{r}_{mt}$ in network m (Eq. 12, Fig. 3A). This

implies that a neural decoder only needs to access neural activity locally rather than over distributed brain areas. We find that the network dynamics is performing Langevin sampling on the manifold of torus (Fig. 3C), and that the equilibrium distribution of samples agrees with the theoretical prediction (Fig. 3D, see details in SI. Sec. 4.3). In particular, the distribution of sampled feature values at each individual network agrees with the marginal posterior of the corresponding stimulus feature, demonstrating that our model is performing inference in a distributed manner. Furthermore, we confirm that the reciprocal coupling between two networks encodes the prior of stimulus features (i.e., the precision matrix, Fig. 3E, Eq. 17).

Our model is robust and scalable to arbitrary number of networks for processing arbitrary dimension of stimulus features. To demonstrate this, we simulate a model containing up to ten coupled CANNs (Fig. 3F), and verify that indeed each network is able to implement Langevin sampling of the marginal posterior of the corresponding stimulus feature (Fig. 3G-H).

5 Conclusions and Discussions

In this study, we propose that coupled CANNs with appropriate structures are able to implement distributed, sampling-based inference to approximate the multivariate posterior of high-dimensional stimulus features. We elucidate that the dynamics of coupled CANNs in effect realizes Langevin sampling in the manifold of stimulus features, where the feedforward inputs convey the likelihood, the reciprocal connections encode the priors of stimulus association, and the internal Poisson variability of the neurons produce random walks for sampling. Our model achieves sampling-based inference in a distributed manner, in term of that each network infers the marginal posterior of the corresponding stimulus feature. Furthermore, the sampled stimulus features can be read out easily using population vector based only on local neural activities at each network.

Our model is different from others in the literature for implementing Bayesian inference. Compared to PPC and its extended version to coupling networks [12, 13], our model is distinguished in that: 1) the inference and representation of the posterior in our model is sampling-based, rather than parametrically represented; 2) our model is stochastic, which generates internal Poisson variability (not as inherited from the noisy feedforward inputs as in PPC) to trigger random walks of sampling stimulus features; 3) the marginalization via sampling in our model collectively emerges from the stochastic dynamics of coupled networks, rather than through message-passing mediated by nonlinear couplings between networks [13]. Compared to other sampling-based inference models (e.g., [14–20]), our model considers that sampling occurs on the manifold of stimulus features embedded in neuronal population responses, where stimulus features are the variables of interest. Furthermore, we develop a concrete neural circuit model to implement this inference, and demonstrate that the Poisson (spiking) variability of the neurons contributes to the Gaussian-white variability of stimulus features necessary for Langevin sampling. A previous study also employed coupled CANNs to realize distributed multisensory integration [46], but their way of realizing Bayesian inference is not sampling-based, but rather being a statistical result by averaging over input trials. Overall, we have proposed a novel model which integrates the strengths of PPC and sampling-based codes in a unified framework to achieve high-dimensional Bayesian inference efficiently.

Nevertheless, there are some detailed issues needing to be addressed in the future. Langevin sampling is known to be slow especially in a high-dimensional space [48]. In our model, sampling occurs on the manifold of stimulus features, whose dimension (equalling to the number of coupled networks) is already smaller by orders of magnitude compared to the dimension of neuronal responses (equalling to the number of neurons), but still the sampling speed may be not quick enough in practice. Our analysis shows that the temporal correlation of sampling in our model is $\rho_s(\Delta t) = \exp(-\rho\Delta t/\sqrt{2}\tau \mathsf{U})$ (Fig. 2G), where the decaying time constant $\sqrt{2}\tau \mathsf{U}/\rho$ (whose inverse quantifying the sampling speed) increases with the neural activity U (Eq. 11). To further speed up sampling, a potential strategy is to include short-term synaptic plasticity [49], which increases the transition probability between network states [50] and has the potential to generate effective anti-symmetric connections between networks to speed up sampling as proposed in [16]. We will study this issue in future work.

Acknowledgments

This work is supported by National Science Foundation (1816568), the National Institutes of Health (Grants 1U19NS107613-01 and R01EB026953), the Vannevar Bush Faculty (Fellowship N00014-18-1-2002), and the Simons Foundation Collaboration on the Global Brain.

References

[1] Tai Sing Lee and David Mumford. Hierarchical bayesian inference in the visual cortex. *JOSA A*, 20(7):1434–1448, 2003.

[2] David C Knill and Alexandre Pouget. The bayesian brain: the role of uncertainty in neural coding and computation. *TRENDS in Neurosciences*, 27(12):712–719, 2004.

[3] Alan Yuille and Daniel Kersten. Vision as bayesian inference: analysis by synthesis? *Trends in cognitive sciences*, 10(7):301–308, 2006.

[4] Marc O Ernst and Martin S Banks. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870):429–433, 2002.

[5] Konrad P Körding and Daniel M Wolpert. Bayesian integration in sensorimotor learning. *Nature*, 427(6971):244–247, 2004.

[6] Wei Ji Ma, Vidhya Navalpakkam, Jeffrey M Beck, Ronald Van Den Berg, and Alexandre Pouget. Behavior and neural basis of near-optimal visual search. *Nature neuroscience*, 14(6):783, 2011.

[7] Yong Gu, Dora E Angelaki, and Gregory C DeAngelis. Neural correlates of multisensory cue integration in macaque mstd. *Nature Neuroscience*, 11(10):1201–1210, 2008.

[8] Christopher R Fetsch, Gregory C DeAngelis, and Dora E Angelaki. Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. *Nature Reviews Neuroscience*, 14(6):429–442, 2013.

[9] Elias B Issa, Charles F Cadieu, and James J DiCarlo. Neural dynamics at successive stages of the ventral visual stream are consistent with hierarchical error signals. *Elife*, 7:e42870, 2018.

[10] David J Field, Anthony Hayes, and Robert F Hess. Contour integration by the human visual system: evidence for a local "association field". *Vision research*, 33(2):173–193, 1993.

[11] Wu Li, Valentin Piëch, and Charles D Gilbert. Contour saliency in primary visual cortex. *Neuron*, 50(6):951–962, 2006.

[12] Wei Ji Ma, Jeffrey M Beck, Peter E Latham, and Alexandre Pouget. Bayesian inference with probabilistic population codes. *Nature Neuroscience*, 9(11):1432–1438, 2006.

[13] Rajkumar Vasudeva Raju and Zachary Pitkow. Inference by reparameterization in neural population codes. In *Advances in Neural Information Processing Systems*, pages 2029–2037, 2016.

[14] Patrik O Hoyer and Aapo Hyvärinen. Interpreting neural response variability as monte carlo sampling of the posterior. In *Advances in neural information processing systems*, pages 293–300, 2003.

[15] Lars Buesing, Johannes Bill, Bernhard Nessler, and Wolfgang Maass. Neural dynamics as sampling: a model for stochastic computation in recurrent networks of spiking neurons. *PLoS computational biology*, 7(11):e1002211, 2011.

[16] Guillaume Hennequin, Laurence Aitchison, and Máté Lengyel. Fast sampling-based inference in balanced neuronal networks. In *Advances in neural information processing systems*, pages 2240–2248, 2014.

[17] József Fiser, Pietro Berkes, Gergő Orbán, and Máté Lengyel. Statistically optimal perception and learning: from behavior to neural representations. *Trends in cognitive sciences*, 14(3):119–130, 2010.

[18] Ralf M Haefner, Pietro Berkes, and József Fiser. Perceptual decision-making as probabilistic inference by neural sampling. *Neuron*, 90(3):649–660, 2016.

[19] Shangqi Guo, Zhaofei Yu, Fei Deng, Xiaolin Hu, and Feng Chen. Hierarchical bayesian inference and learning in spiking neural networks. *IEEE transactions on cybernetics*, 49(1):133–145, 2017.

[20] Ying Fang, Zhaofei Yu, Jian K Liu, and Feng Chen. A unified neural circuit of causal inference and multisensory integration. *Neurocomputing*, 358:355–368, 2019.

[21] Sabyasachi Shivkumar, Richard Lange, Ankani Chattoraj, and Ralf Haefner. A probabilistic population code based on neural samples. In *Advances in Neural Information Processing Systems*, pages 7070–7079, 2018.

[22] Laurence Aitchison and Máté Lengyel. The hamiltonian brain: efficient probabilistic inference with excitatory-inhibitory neural circuit dynamics. *PLoS computational biology*, 12(12), 2016.

[23] Jean-Pierre Bresciani, Franziska Dammeier, and Marc O Ernst. Vision and touch are automatically integrated for the perception of sequences of events. *Journal of Vision*, 6(5):2, 2006.

[24] Neil W Roach, James Heron, and Paul V McGraw. Resolving multisensory conflict: a strategy for balancing the costs and benefits of audio-visual integration. *Proceedings of the Royal Society of London B: Biological Sciences*, 273(1598):2159–2168, 2006.

[25] Yoshiyuki Sato, Taro Toyoizumi, and Kazuyuki Aihara. Bayesian inference explains perception of unity and ventriloquism aftereffect: identification of common sources of audiovisual stimuli. *Neural Computation*, 19(12):3335–3355, 2007.

[26] Daniel Kersten, Pascal Mamassian, and Alan Yuille. Object perception as bayesian inference. *Annu. Rev. Psychol.*, 55:271–304, 2004.

[27] Max Welling and Yee W Teh. Bayesian learning via stochastic gradient langevin dynamics. In *Proceedings of the 28th international conference on machine learning (ICML-11)*, pages 681–688, 2011.

[28] Radford M Neal et al. Mcmc using hamiltonian dynamics. *Handbook of markov chain monte carlo*, 2(11):2, 2011.

[29] R Ben-Yishai, R Lev Bar-Or, and H Sompolinsky. Theory of orientation tuning in visual cortex. *Proceedings of the National Academy of Sciences*, 92(9):3844–3848, 1995.

[30] James J Knierim and Kechen Zhang. Attractor dynamics of spatially correlated neural activity in the limbic system. *Annual review of neuroscience*, 35:267–285, 2012.

[31] Si Wu, Kosuke Hamaguchi, and Shun-ichi Amari. Dynamics and computation of continuous attractors. *Neural Computation*, 20(4):994–1025, 2008.

[32] Sophie Deneve, Peter E Latham, and Alexandre Pouget. Reading population codes: a neural implementation of ideal observers. *Nature Neuroscience*, 2(8):740–745, 1999.

[33] Matteo Carandini and David J Heeger. Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, 13(1):51–62, 2012.

[34] Cristopher M Niell. Cell types, circuits, and receptive fields in the mouse visual cortex. *Annual review of neuroscience*, 38:413–431, 2015.

[35] Alexandre Pouget, Jeffrey M Beck, Wei Ji Ma, and Peter E Latham. Probabilistic brains: knowns and unknowns. *Nature neuroscience*, 16(9):1170, 2013.

[36] Jonathan W Pillow, Jonathon Shlens, Liam Paninski, Alexander Sher, Alan M Litke, EJ Chichilnisky, and Eero P Simoncelli. Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*, 454(7207):995, 2008.

[37] Volker Pernice, Benjamin Staude, Stefano Cardanobile, and Stefan Rotter. How structure determines correlations in neuronal networks. *PLoS computational biology*, 7(5):e1002059, 2011.

[38] James Trousdale, Yu Hu, Eric Shea-Brown, and Krešimir Josić. Impact of network structure and cellular response on spike time correlations. *PLoS computational biology*, 8(3):e1002408, 2012.

[39] David J Tolhurst, J Anthony Movshon, and Andrew F Dean. The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision research*, 23(8):775–785, 1983.

[40] Brent Doiron, Ashok Litwin-Kumar, Robert Rosenbaum, Gabriel K Ocker, and Krešimir Josić. The mechanics of state-dependent neural correlations. *Nature neuroscience*, 19(3):383, 2016.

[41] Carl Van Vreeswijk and Haim Sompolinsky. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*, 274(5293):1724–1726, 1996.

[42] Michael London, Arnd Roth, Lisa Beeren, Michael Häusser, and Peter E Latham. Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex. *Nature*, 466(7302):123–127, 2010.

[43] Ashok Litwin-Kumar and Brent Doiron. Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nature neuroscience*, 15(11):1498, 2012.

[44] Francesca Mastrogiuseppe and Srdjan Ostojic. Linking connectivity, dynamics, and computations in low-rank recurrent neural networks. *Neuron*, 99(3):609–623, 2018.

[45] C. C Alan Fung, K. Y. Michael Wong, and Si Wu. A moving bump in a continuous manifold: A comprehensive study of the tracking dynamics of continuous attractor neural networks. *Neural Computation*, 22(3):752–792, 2010.

[46] Wen-Hao Zhang, Aihua Chen, Malte J Rasch, and Si Wu. Decentralized multisensory information integration in neural systems. *The Journal of Neuroscience*, 36(2):532–547, 2016.

[47] Apostolos P Georgopoulos, Andrew B Schwartz, and Ronald E Kettner. Neuronal population coding of movement direction. *Science*, 233(4771):1416–1419, 1986.

[48] Christopher M Bishop. *Pattern recognition and machine learning*. springer, 2006.

[49] Misha Tsodyks, Klaus Pawelzik, and Henry Markram. Neural networks with dynamic synapses. *Neural computation*, 10(4):821–835, 1998.

[50] CC Alan Fung, KY Michael Wong, He Wang, and Si Wu. Dynamical synapses enhance neural information processing: gracefulness, accuracy, and mobility. *Neural computation*, 24(5):1147–1185, 2012.