

# Stochastic inference with deterministic spiking neurons\*

Mihai A. Petrovici<sup>1\*</sup>, Johannes Bill<sup>2\*</sup>, Ilja Bytschok<sup>1</sup>, Johannes Schemmel<sup>1</sup>, Karlheinz Meier<sup>1</sup>

<sup>1</sup>Kirchhoff Institute for Physics, University of Heidelberg

<sup>2</sup>Institute for Theoretical Computer Science, Graz University of Technology

(Dated: November 14, 2013)

The seemingly stochastic transient dynamics of neocortical circuits observed *in vivo* have been hypothesized to represent a signature of ongoing stochastic inference. *In vitro* neurons, on the other hand, exhibit a highly deterministic response to various types of stimulation. We show that an ensemble of deterministic leaky integrate-and-fire neurons embedded in a spiking noisy environment can attain the correct firing statistics in order to sample from a well-defined target distribution. We provide an analytical derivation of the activation function on the single cell level; for recurrent networks, we examine convergence towards stationarity in computer simulations and demonstrate sample-based Bayesian inference in a mixed graphical model. This establishes a rigorous link between deterministic neuron models and functional stochastic dynamics on the network level.

PACS numbers: xxx-xxx

## Introduction

In responding to environmental sensory stimuli, brains have to deal with what is typically limited, noisy and ambiguous data. Based on such imperfect information, animals need to predict and react to changes in their environment. The recent hypothesis that the brain handles these challenges by performing Bayesian, rather than logical inference [1–3], has been strengthened by electrophysiological data which identified neural correlates of the involved computations [4, 5] and theoretical work on spiking network implementations [6–8].

In probabilistic inference, the potential values of a quantity are described by a random variable (RV)  $z_k$  and all knowledge about dependencies between random variables is stored in a joint probability distribution  $p(z_1, \dots, z_K)$ . The Bayesian belief about a set of unobserved RVs  $\{z_1, \dots, z_M\}$  given an observed set of RVs is represented by the posterior distribution  $p(z_1, \dots, z_M | z_{M+1}, \dots, z_K)$ . In particular, the posterior contains information not only on the most likely conclusion and potential alternatives, but also on the level of uncertainty of the outcome.

Theoretical work [2] has argued in favor of sample-based representations of probability distributions in the brain. In this representation, instead of providing the entire distribution at any point in time, only samples  $\mathbf{z}^{(t)} \sim p(z_1, \dots, z_K)$  are used as a proxy. When modeling large systems, this offers three important advantages. First, approximate solutions can be provided at any time, with increasingly reliable results as the calculation progresses (“anytime computing”). Secondly, in a sample-based representation marginalization comes at no cost, as  $p(z_k)$  can be determined by simply neglecting the values of all other RVs. Thirdly, some sampling algorithms such as Gibbs sampling support a high degree of parallelization. In particular, the resulting algorithmic structure is reminiscent of neural networks [9].

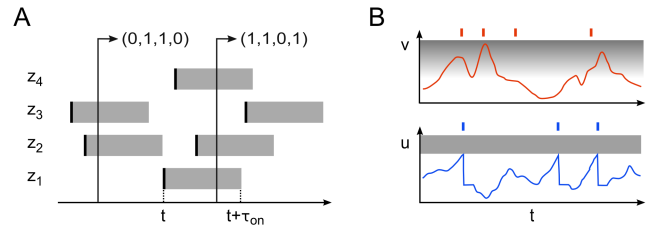


FIG. 1. (A) Interpretation of spike patterns as samples of a binary random vector  $\mathbf{z}$ . The variable  $z_k$  is active for duration  $\tau_{\text{on}}$  (grey bar) after a spike of neuron  $k$ . (B) In stochastic neuron models, internal state variables modulate the instantaneous firing probability (red). In contrast, deterministic integrate-and-fire neurons elicit a spike when the membrane potential crosses a threshold voltage (blue). The probability of firing as a function of the respective internal variable is represented by the grayscale in the background.

Recently, a theory has been suggested which combines these advantages by implementing Markov chain Monte Carlo sampling in networks of abstract model neurons [7]. In this framework, spike patterns are interpreted as samples of binary RVs as follows (see Fig. 1A):

$$z_k^{(t)} = 1 \Leftrightarrow \text{Neuron } k \text{ fired in } (t - \tau_{\text{on}}, t] \quad . \quad (1)$$

The duration  $\tau_{\text{on}}$  of the active state following a spike is a free parameter; in cortex,  $\tau_{\text{on}} \approx 10$  ms is a good estimate for the timescale on which a spiking neuron affects the membrane potential of downstream cells. The neuron model underlying [7] is inherently stochastic (Fig. 1B top), with an instantaneous firing rate defined by

$$\begin{aligned} r_k(t) &= \lim_{\Delta t \rightarrow 0} \frac{p(\text{spike in } [t, t + \Delta t))}{\Delta t} \\ &= \begin{cases} \frac{1}{\tau} \exp(v_k) & \text{if } z_k = 0 \\ 0 & \text{if } z_k = 1 \end{cases} \quad , \end{aligned} \quad (2)$$

where  $v_k$  represents an abstract membrane potential.

In contrast to this approach, *in vitro* experiments have demonstrated the largely deterministic nature of single

neurons [10]. Similarly, microscopic models of neural circuits typically rely on deterministic dynamics of their constituents. An often-used mechanistic model is the leaky integrate-and-fire (LIF) neuron

$$C_m \frac{du_k}{dt} = g_l(E_l - u_k) + I \quad , \quad (3)$$

with capacitance  $C_m$ , membrane potential  $u_k$ , leak potential  $E_l$ , leak conductance  $g_l$  and input current  $I$ . The spiking condition is deterministic as well: when  $u_k$  crosses a threshold  $\vartheta$  from below, a spike is emitted and  $u_k$  is reset to  $\varrho$  for a refractory period  $\tau_{\text{ref}}$  (see Fig. 1B bottom). For conductance-based synapses, the synaptic input current to a neuron is typically modelled as

$$\frac{dI_k^{\text{syn}}}{dt} = -\frac{I_k^{\text{syn}}}{\tau_{\text{syn}}} + \sum_{\text{syn}} \sum_{i \text{ spk } s} w_{ki} (E_i^{\text{rev}} - u_k) \delta(t - t_s) \quad , \quad (4)$$

with the synaptic time constant  $\tau_{\text{syn}}$ , the synaptic weight  $w_{ki}$  and the reversal potential of the  $i$ th synapse  $E_i^{\text{rev}}$ .

The aim of this letter is to demonstrate how a network of deterministic neurons in a biologically plausible spiking noisy environment can quantitatively reproduce the stochastic dynamics required for sampling from a well-defined distribution  $p(z_1, \dots, z_K)$  and perform inference given observations. We start by calculating the dynamics of a single LIF neuron in a spiking noisy environment and derive its activation function by describing the spike response as a first passage time (FPT) problem. This establishes an equivalence to the abstract, inherently stochastic units (2). On the network level, we show how biologically realistic conductance-based synapses (4) approximate the interaction for sampling from a well-defined target distribution. We complement our study with a demonstration of probabilistic inference by implementing the posterior of a small graphical model for handwritten digit recognition in a recurrent network of LIF neurons.

### Deterministic neurons in a noisy environment

The total input current  $I_k$  to a neuron can be formally partitioned into recurrent synaptic input, diffuse synaptic noise and additional external currents:  $I_k = I_k^{\text{rec}} + I_k^{\text{noise}} + I_k^{\text{ext}}$ . While the synaptic currents  $I_k^{\text{rec}}$  and  $I_k^{\text{noise}}$  obey eqn. (4), the current  $I_k^{\text{ext}}$  captures additional current stimuli. We start by considering a single neuron that receives diffuse synaptic noise  $I_k^{\text{noise}}$  in the form of random spikes from its surrounding. The capacity of recurrent networks to produce such noise has been shown in [11]. Throughout the following analysis of individual neurons we omit the index  $k$  and set  $I^{\text{rec}} = 0$ .

When a conductance-based LIF neuron receives strong synaptic stimulation, it enters a so-called high-conductance state (HCS, [12]), characterized by accelerated membrane dynamics. It is therefore convenient to

rewrite (3) as

$$\tau_{\text{eff}} \frac{du}{dt} = u_{\text{eff}} - u \quad , \quad (5)$$

where the membrane time constant  $\tau_m = C_m/g_l$  is replaced by a smaller effective time constant  $\tau_{\text{eff}} = C_m/g_{\text{tot}}$ , with the total conductance  $g_{\text{tot}}$  subsuming both leakage and synaptic conductances. In a HCS,  $\tau_{\text{eff}}$  governs the decay towards an effective leak potential  $u_{\text{eff}} = (g_l E_l + \sum_i g_i^{\text{noise}} E_i^{\text{rev}} + I^{\text{ext}})/g_{\text{tot}}$ , where  $g_i^{\text{noise}}$  represents the total conductance at the  $i$ th synapse. In a high input rate regime,  $\sqrt{\text{Var}(g_{\text{tot}})}/\langle g_{\text{tot}} \rangle \rightarrow 0$  and the equation governing the membrane potential can be written as

$$\tau_{\text{eff}} \frac{du}{dt} = \frac{I^{\text{ext}} + g_l E_l}{\langle g_{\text{tot}} \rangle} + \frac{\sum_i g_i^{\text{noise}} E_i^{\text{rev}}}{\langle g_{\text{tot}} \rangle} - u \quad , \quad (6)$$

with  $\langle \cdot \rangle$  denoting the mean. In a first approximation,  $\tau_{\text{eff}}$  can be considered very small in the HCS, resulting in  $u \approx u_{\text{eff}}$ , with the effective potential  $u_{\text{eff}}$  simply being a linear transformation of the synaptic noise input.

Using methods similar to [13], it can be shown that, if stimulated by a large number of uncorrelated spike sources, the synaptic current  $I^{\text{noise}}$  – and therefore, also  $u_{\text{eff}}$  – can be described as an Ornstein-Uhlenbeck (OU) process:

$$du(t) = \theta \cdot (\mu - u(t)) + \Sigma \cdot dW(t) \quad , \quad (7)$$

with parameters

$$\theta = \tau_{\text{syn}} \quad (8)$$

$$\mu = \frac{I^{\text{ext}} + g_l E_l + \sum_i \nu_i w_i^{\text{noise}} E_i^{\text{rev}} \tau_{\text{syn}}}{\langle g_{\text{tot}} \rangle} \quad (9)$$

$$\Sigma^2 = \sum_i \nu_i [w_i^{\text{noise}} (E_i^{\text{rev}} - \mu)]^2 \tau_{\text{syn}} / \langle g_{\text{tot}} \rangle \quad . \quad (10)$$

where  $\nu_i$  represents the input rate at the  $i$ th noise synapse and  $w_i$  its weight.

### The activation function as an FPT problem

The inherently stochastic neuron model (2) leads to a logistic activation function for constant potential  $v$ :

$$p(z = 1) = \sigma(v) := [1 + \exp(-v)]^{-1} \quad . \quad (11)$$

In the following, we derive the activation function of the deterministic LIF neuron in a spiking noisy environment. Similarly to the abstract model [7], we define the refractory state of a neuron as  $z = 1$ .

An example of membrane potential dynamics with reset is shown in Fig. 2A. Two modes of firing can be observed: the ‘‘bursting’’ mode, where the effective membrane potential after the refractory period is still above

threshold, and the freely evolving mode, where the neuron does not spike again immediately after the refractory period. Denoting the relative occurrence of burst lengths  $n$  by  $P_n$  and the average duration of the freely evolving mode that follows an  $n$ -spike-burst by  $T_n$ , we can identify the following relation:

$$p(z = 1) = \frac{\sum_n P_n \cdot n \cdot \tau_{\text{on}}}{\sum_n P_n \cdot (n \cdot \tau_{\text{on}} + T_n)} . \quad (12)$$

Given the parameters of the associated OU process, we can derive a recursive expression for  $P_n$  and  $T_n$ , thereby ultimately allowing the calculation of  $p(z = 1)$ :

$$\begin{aligned} P_n &= p(u_n < \vartheta, u_{n-1} \geq \vartheta, \dots, u_1 \geq \vartheta | u_0 = \vartheta) \quad (13) \\ &= \left(1 - \sum_{i=1}^{n-1} P_i\right) \int_{\vartheta}^{\infty} du_{n-1} p(u_{n-1} | u_{n-1} \geq \vartheta) \\ &\quad \left[ \int_{-\infty}^{\vartheta} du_n p(u_n | u_{n-1}) \right] , \\ T_n &= \int_{\vartheta}^{\infty} du_{n-1} p(u_{n-1} | u_{n-1} \geq \vartheta) \quad (14) \\ &\quad \left[ \int_{-\infty}^{\vartheta} du_n p(u_n | u_n < \vartheta, u_{n-1}) \langle T(\vartheta, u_n) \rangle \right] . \end{aligned}$$

Fig. 2B displays an intuitive picture of the integrals in (13) and (14). The transfer function  $p(u_n | u_{n-1})$  is the Green's function of the OU process for  $t = \tau_{\text{ref}}$ :

$$p(u_n | u_{n-1}) = C e^{-\frac{\theta}{\Sigma^2} \left[ \frac{(u_n - (u_{n-1} - \mu) \exp(-\theta \tau_{\text{ref}}) - \mu)^2}{1 - \exp(-2\theta \tau_{\text{ref}})} \right]} , \quad (15)$$

with the normalization  $C = \sqrt{\theta/\pi \Sigma^2 (1 - e^{-2\theta \tau_{\text{ref}}})}$ .  $T(u_b, u_a)$  denotes the time the membrane needs to reach  $u_b$  starting from  $u_a$ . This FPT problem has been extensively discussed in literature and its moments can be given in closed form [14].

To improve the prediction of the activation function, we further take into account small, but finite  $\tau_{\text{eff}}$ , in which case the membrane potential no longer directly follows the input current, but is a low-pass-filtered version thereof. By using an expansion in  $\sqrt{\tau_{\text{eff}}/\tau_{\text{syn}}}$ , a first-order correction to the FPT can be calculated [15]:

$$\langle T(\vartheta, u) \rangle = \tau_{\text{syn}} \sqrt{\pi} \int_{\frac{u-\mu}{\sigma}}^{\frac{\vartheta_{\text{eff}}-\mu}{\sigma}} dx \exp(x^2) [\text{erf}(x) + 1] , \quad (16)$$

with the effective threshold  $\vartheta_{\text{eff}} \approx \vartheta - \zeta \left(\frac{1}{2}\right) \sqrt{\frac{\tau_{\text{eff}}}{2\tau_{\text{syn}}}}$ , where  $\zeta$  denotes the Riemann Zeta function. A comparison of the predicted  $p(z = 1)$  with results from a numerical simulation is shown in Fig. 2C. Here, the average effective potential  $\bar{u} = \langle u_{\text{eff}} \rangle$  was established through an external current  $I^{\text{ext}}$ . For the translation from the LIF domain to

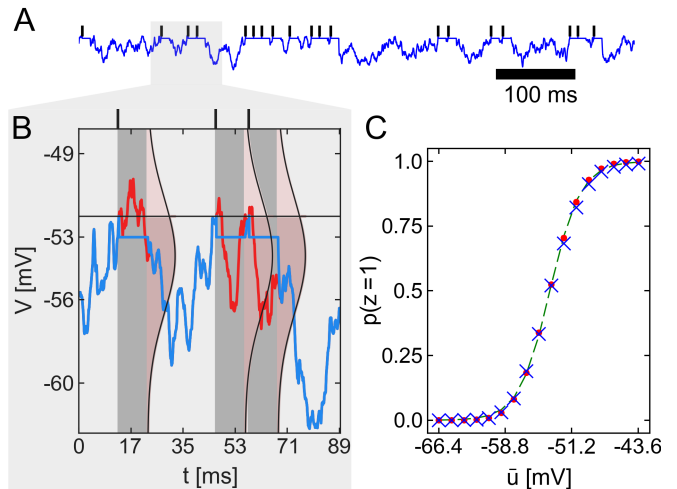


FIG. 2. (A) Membrane potential  $u(t)$  (blue) and resulting spike activity (black) of a LIF neuron in a spiking noisy environment. (B) Overlay of  $u$  (blue) and  $u_{\text{eff}}$  (red). Due to the small  $\tau_{\text{eff}}$  in the HCS, the two curves are nearly identical when the neuron is not refractory. At the end of each refractory state (corresponding to  $z = 1$ , grey), the predicted probability distribution for  $u_{\text{eff}}$  is plotted in pink. The normalized subthreshold area (dark pink) is used for the propagation in (13). (C) Theoretical prediction (red) vs. simulation results (blue); errors are smaller than the symbol size. A logistic function  $\sigma(\bar{u})$  (green) has been fitted to the prediction.

the abstract stochastic model (2) we identify

$$v = \frac{\bar{u} - \bar{u}^0}{\alpha} , \quad (17)$$

where  $\bar{u}^0$  denotes the value of  $\bar{u}$  for which  $p(z = 1) = \sigma(0) = \frac{1}{2}$  and  $\alpha$  represents a scaling factor between the two domains. We conclude that a single LIF neuron in a spiking noisy environment can closely reproduce the activation function (11).

### Sampling via recurrent networks of LIF neurons with conductance-based synapses

We next connect the neurons to a recurrent network. In addition to noise stimuli, a LIF neuron receives synaptic currents  $I_k^{\text{ec}}$  from other neurons in the ensemble. Synaptic interaction introduces correlations among the spike response of different neurons, i.e. the random variables  $z_1, \dots, z_K$  are not independent. For certain connectivity structures, it is possible to specify a target distribution [7, 16] for the network states  $\mathbf{z}^{(t)}$  that occur under the network dynamics.

In the following, we use the emulation of Boltzmann machines as an example case. The joint distribution reads:

$$p_{\text{B}}(\mathbf{z}) = \frac{1}{Z} \exp\left(\frac{1}{2} \mathbf{z}^{\text{T}} \mathbf{W} \mathbf{z} + \mathbf{z}^{\text{T}} \mathbf{b}\right) , \quad (18)$$

where  $\mathbf{W}$  is a symmetric zero-diagonal weight matrix,  $\mathbf{b}$  is a bias vector,  $\mathbf{z}^\top \mathbf{W} \mathbf{z}$  and  $\mathbf{z}^\top \mathbf{b}$  denote bilinear forms over  $\mathbf{W}$  and  $\mathbf{1}$  respectively, and  $Z = \sum_{\mathbf{z}'} \exp\left(\frac{1}{2} \mathbf{z}'^\top \mathbf{W} \mathbf{z}' + \mathbf{z}'^\top \mathbf{b}\right)$  is the partition function that ensures correct normalization. This probabilistic model underlies state-of-the-art machine learning algorithms for image [17] and speech recognition [18]. For the abstract neuron model (2), it had been proven [7] that a membrane potential of the form

$$v_k = b_k + \sum_{j=1}^K W_{kj} z_j \quad (19)$$

leads to the desired target distribution (18) of network states  $\mathbf{z}^{(t)}$  for  $t \rightarrow \infty$ . This finding uses the fact that individual neurons can sample from the conditionals  $p(z_k = 1 | \mathbf{z}_{\setminus k}) = \sigma(v_k)$ , with  $\mathbf{z}_{\setminus k} = \{z_j | j \neq k\}$ , in a Gibbs sampling inspired updating scheme.

As shown above, LIF neurons in a spiking noisy environment closely approximate this logistic activation function if the synaptic currents  $I^{\text{rec}}$  shift the mean membrane potential  $\bar{u}_k$  according to the linear interaction (19). Using the linear transformation (17) between  $v_k$  and  $\bar{u}_k$ , and estimating the impact of a pre-synaptic spike on the post-synaptic neuron through conductance-based synapses of weight  $w_{kj}$ , we arrive at the following parameter translation between the abstract and the LIF domain:

$$b_k = (\bar{u}_k^b - \bar{u}_k^0) / \alpha \quad (20)$$

$$W_{kj} = \frac{1}{\alpha C_m} \frac{w_{kj} (E_{kj}^{\text{rev}} - \mu)}{1 - \frac{\tau_{\text{syn}}}{\tau_{\text{eff}}}} \left[ \tau_{\text{syn}} (e^{-1} - 1) - \tau_{\text{eff}} \left( e^{-\frac{\tau_{\text{syn}}}{\tau_{\text{eff}}}} - 1 \right) \right] \quad (21)$$

where  $\bar{u}_k^b$  is the mean free potential  $\bar{u}_k$  in Fig. 2C that establishes  $p(z_k = 1 | \mathbf{z}_{\setminus k} = \mathbf{0}) = \sigma(b_k)$ , and  $E_{kj}^{\text{rev}}$  denotes the reversal potential for synapse  $w_{kj}$ . The idea behind (21) is to match the integrals of individual postsynaptic potentials (PSPs) on  $v_k$  and  $\bar{u}_k$ . We furthermore employ short-term synaptic depression to approximate the theoretically optimal rectangular PSP shape also in case of consecutive spikes (bursts).

The sampling process with networks of LIF neurons was examined in computer simulations. Fig. 3A shows the spike pattern of a recurrent network of  $K = 5$  LIF neurons that sample from a randomly generated Boltzmann machine. The parameters  $b_k$  and  $W_{kj}$  were drawn from the interval  $[-0.6, 0.6]$ . The thus defined target distribution  $p_B(\mathbf{z})$  is approximated by the distribution  $p_N(\mathbf{z})$  of network states when the spike pattern is interpreted as samples  $\mathbf{z}^{(t)}$  by convolution with a  $\tau_{\text{on}} = 10$  ms box kernel. Fig. 3B shows the average network distribution  $p_N(\mathbf{z})$  (blue bars) after  $T = 10$  s simulation time alongside the target values  $p_B(\mathbf{z})$  (red lines) calculated

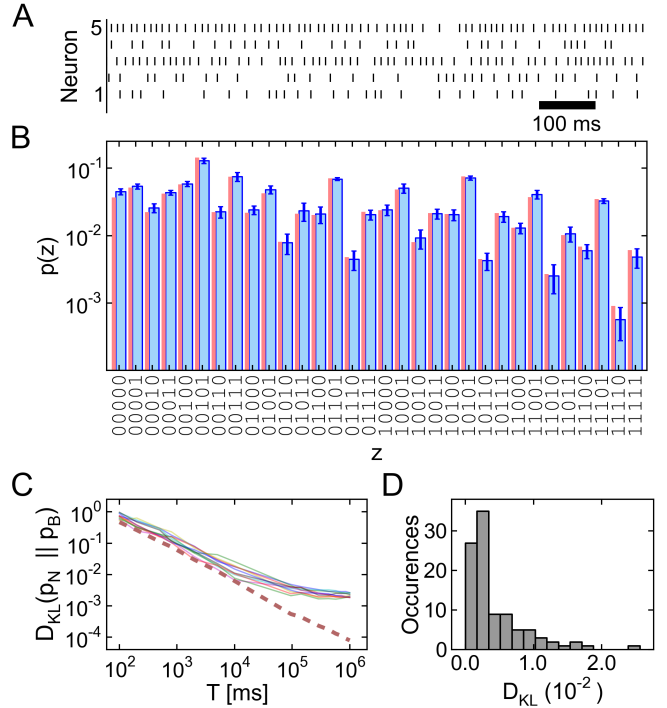


FIG. 3. (A) Spike pattern of a recurrent network of LIF neurons during sampling from a randomly generated Boltzmann machine. (B) Sampled distribution  $p_N(\mathbf{z})$  of network states (blue bars) and analytically calculated target distribution  $p_B(\mathbf{z})$  (in red). (C) Kullback-Leibler divergence between the sampled distribution  $p_N$  and the target distribution  $p_B$  as a function of integration time  $T$  for 10 trials (thin lines). The red dotted line shows convergence for a network of intrinsically stochastic (theoretically optimal) neurons that is guaranteed to converge to  $p_B$  for  $T \rightarrow \infty$ . (D) Kullback-Leibler divergence  $D_{\text{KL}}(p_N || p_B)$  when sampling for  $T = 10^3$  s from 100 different randomly generated target distributions.

from (18). Sampled probabilities  $p_N(\mathbf{z})$  depict the mean over 10 independent simulation runs, errorbars reflect stochastic variations between individual runs. The chosen integration time  $T = 10$  s displays a conservative estimate of the maximum duration a neuronal ensemble experiences stable stimulus conditions in a behaving organism and can thus be expected to sample from a stationary distribution. We find that the recurrent network of LIF neurons accurately encodes the target distribution over several orders of magnitude, within the precision imposed by the sample-based representation.

Fig. 3C shows how the network distribution becomes increasingly more reliable as more samples are considered. After few samples, the network has generated a coarse approximation of  $p_B(\mathbf{z})$  that could serve as an “educated guess” in online computation tasks. For simulation times  $T$  well beyond biologically relevant timescales, systematic errors in  $p_N(\mathbf{z})$  become apparent: The KL divergence saturates on a non-zero value, while the (theoretically ideal) abstract model [7] further converges to-

wards the target distribution (red dotted line). Fig. 3D shows that the sampling quality holds for a variety of similarly generated target distributions.

### Demonstration of probabilistic inference

We conclude our investigation of sampling in recurrent networks of LIF neurons with an example of Bayesian inference based on incomplete observations. A fully connected Boltzmann machine of  $K = 144$  neurons, aligned on a 12x12 grid, was trained as an associative network [19] to store three prototypic patterns, namely handwritten digits 0, 3 and 4. In the network, each cell is assigned to one pixel of the image. Statistical correlations between the pixels as well as their mean intensities were encoded in the weights  $W_{kj}$  and biases  $b_k$  of the corresponding joint distribution  $p(\mathbf{z})$ . This distribution reflects ‘‘prior knowledge’’ imprinted in the network.

The probabilistic model is augmented by adding real-valued input channels for each pixel. Like the network variables  $z_k$ , input channels are associated with random variables  $y_k \in \mathbb{R}$ ,  $1 \leq k \leq K$ . The resulting graphical model is sketched in Fig. 4A and entails the following structure for a full probabilistic model:

$$p(\mathbf{y}, \mathbf{z}) = p(\mathbf{z}) \cdot \prod_{k=1}^K p(y_k | z_k) \quad . \quad (22)$$

The full model  $p(\mathbf{y}, \mathbf{z})$  connects the network variables  $z_k$  of the prior to inputs  $y_k$  by means of the likelihood  $p(y_k | z_k)$ . We have chosen a Gaussian likelihood with unit variance (see Fig. 4B):

$$p(y_k | z_k) = \mathcal{N}(y_k; \mu = z_k - \frac{1}{2}, \sigma^2 = 1) \quad . \quad (23)$$

The task for the network is to implement the posterior distribution that follows from Bayes’ rule:  $p(\mathbf{z} | \mathbf{y}) \propto p(\mathbf{z}) \cdot p(\mathbf{y} | \mathbf{z})$ . The posterior combines two sources of information: The likelihood  $p(y_k | z_k)$  tends to align the network state with the observation, i.e.  $z_k = 1$  for  $y_k > 0$ , while the prior  $p(\mathbf{z})$  reconciles the observations with knowledge on consistent activation patterns  $\mathbf{z}$ . In this way, the posterior  $p(\mathbf{z} | \mathbf{y})$  evaluates all possible outcomes  $\mathbf{z} | \mathbf{y}$  simultaneously by assigning a belief to each of them, and thus captures the model’s (un-)certainty about different solutions. A short derivation shows that the posterior  $p(\mathbf{z} | \mathbf{y})$  is a Boltzmann machine for any input  $\mathbf{y}$ , thus being compatible with the sampling dynamics of spiking networks. More specifically, we obtain the following abstract membrane potential:

$$v_k = b_k + y_k + \sum_j W_{kj} z_j \quad . \quad (24)$$

In the LIF domain, the sum  $b_k + y_k$  is equivalent to an effective bias (20) and corresponds to an external current  $I_k^{\text{ext}} = I_k^b + I_k^y$  that shifts  $\bar{u}_k$  appropriately. Thus,

a network neuron receives synaptic input from recurrent connections and noise sources, as well as an external current, i.e.,  $I_k = I_k^{\text{rec}} + I_k^{\text{noise}} + I_k^b + I_k^y$ .

In case of  $I_k^y = 0 \forall k$ , the network samples from the prior distribution  $p(\mathbf{z}) = p(\mathbf{z} | \mathbf{y} = \mathbf{0})$ . A two-dimensional projection of network states  $\mathbf{z}^{(t)} \sim p(\mathbf{z})$  is shown in Fig. 4C. The sampled distribution exhibits three distinct peaks that correspond to the three handwritten digits stored in the recurrent weight matrix. A closer look at the network trajectory reveals how the system stays in one mode for some duration, traverses the state space and then samples from a different mode of the distribution. These dynamics also reflect in the marginals of the network variables under a 20 ms box filter,  $\bar{z}_k(t) = \frac{1}{20 \text{ ms}} \int_{t-20 \text{ ms}}^t z_k^{(t')} dt'$ , shown in the color maps.

A typical scenario of stochastic inference on incomplete observations is shown in Fig. 4D. Four input channels, located at the center of the grid, were picked to inject positive currents  $I_k^y > 0$  to the network while all other inputs remained uninformative. Positive currents  $I_k^y$  encode positive values of the respective input pixels  $y_k$  and were chosen such that the observation appeared incompatible with the digit 0, and remained ambiguous with respect to digits 3 and 4. Accordingly, in the posterior distribution  $p(\mathbf{z} | \mathbf{y})$  the 0-mode is significantly suppressed while uncertainty about the provided cue is expressed by two distinct modes in the 3 and 4 directions.

### Discussion

We have shown how recurrent networks of deterministic neurons in a spiking noisy environment can perform probabilistic inference through sampling from a well-defined posterior distribution. Our approach builds on theoretical work by Buesing et al. [7] and extends Bayesian spiking network implementations to deterministic neuron models widely used in computational neuroscience. For the analytical derivation and the computer simulations we have employed leaky integrate-and-fire neurons with conductance-based synapses. However, the analysis can be readily transferred to other neuron models [20]. The essential diffusion approximation relies on high-frequency spiking inputs that could be provided by the surrounding network and lead to strong synaptic conductances and fast membrane dynamics. Thereby, our derivation identifies a potential functional role of biologically observed high-conductance states within a normative framework of brain computation.

For mathematical tractability, simplifying modeling assumptions had to be made. The neuron model only uses an absolute refractory time  $\tau_{\text{ref}}$ , which matches the activation time constant  $\tau_{\text{on}}$ , and neglects any additional gradual recovery effects after a spike. On the network level, we have assumed statistically indepen-

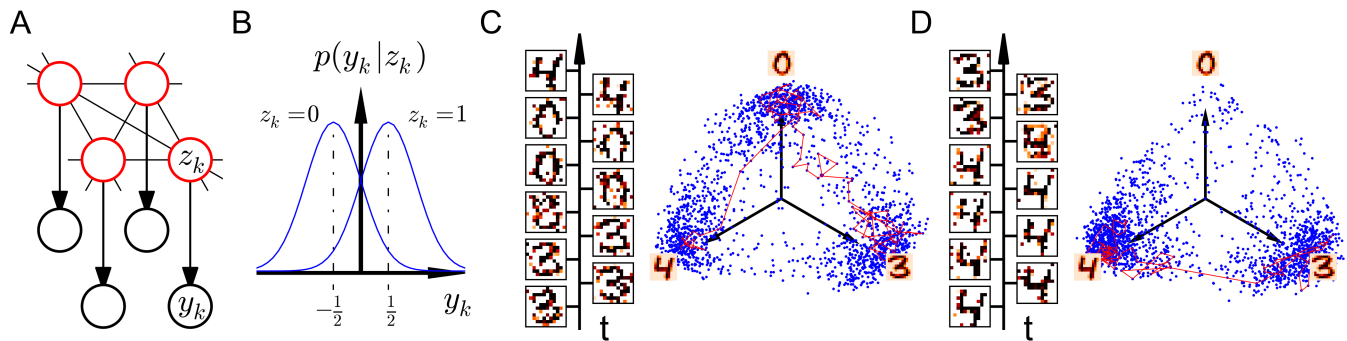


FIG. 4. (A) Graphical model used for the probabilistic inference task. The network implements a Markov random field over latent variables  $\mathbf{z}$ . Observables  $y_k$  are conditionally independent given the network state. (B) A mixture-of-Gaussians likelihood model provides input to the sampling neurons. (C) Two dimensional projection of sampled states  $\mathbf{z}^{(t)}$  when sampling from the prior  $p(\mathbf{z})$ . The network preferentially spends time in modes close to the stored hand-written digits 0, 3, 4. Solid line: network trajectory over 200 ms. Color maps: Marginals  $\bar{z}_k(t)$  averaged over 20 ms. The time arrow covers the duration of the red trajectory and consecutive snapshots are 20 ms apart. (D) As in (C) when sampling from the posterior  $p(\mathbf{z} | \mathbf{y})$  with incomplete observations  $\mathbf{y}$ . The provided input  $\mathbf{y}$  is incompatible with digit 0 and ambiguous with respect to digits 3 and 4.

dent noise sources and instantaneous axonal transmission. Furthermore, post-synaptic potentials mediated through conductance-based synapses differ from the theoretically optimal rectangular shape and can lead to deviations from the target distribution outside of the high noise regime [21]. However, computer simulations indicate that in most scenarios the above approximations are not critical.

Beyond neuroscience, the ability to perform probabilistic inference with deterministic neurons displays a promising computing paradigm for neuromorphic hardware systems. Originally designed as neuroscientific modeling tools, these systems typically implement a physical model of integrate-and-fire neurons [22, 23], which renders the application of the proposed networks straightforward. In particular, the distributed nature of the sampling algorithm allows to exploit the inherent parallelism of neuromorphic architectures, fostering an application of neuromorphic hardware to online data evaluation and robotics.

---

\* The first two authors contributed equally to this work. We thank W. Maass for his essential support and Stefan Habenschuss for helpful comments. This research was supported by EU grants #269921 (BrainScaleS), #237955 (FACETS-ITN), the Austrian Science Fund FWF #I753-N23 (PNEUMA) and the Manfred Stärk Foundation.

- [1] K. P. Körding and D. M. Wolpert, *Nature* **427**, 244 (2004).
- [2] J. Fiser, P. Berkes, G. Orbán, and M. Lengyel, *Trends in cognitive sciences* **14**, 119 (2010).
- [3] K. Friston, J. Mattout, and J. Kilner, *Biological cybernetics* **104**, 137 (2011).
- [4] T. Yang and M. N. Shadlen, *Nature* **447**, 1075 (2007).

- [5] P. Berkes, G. Orbán, M. Lengyel, and J. Fiser, *Science* **331**, 83 (2011).
- [6] S. Deneve, *Neural computation* **20**, 91 (2008).
- [7] L. Buesing, J. Bill, B. Nessler, and W. Maass, *PLoS computational biology* **7**, e1002211 (2011).
- [8] R. P. Rao, in *Advances in neural information processing systems* (2004) pp. 1113–1120.
- [9] P. O. Hoyer and A. Hyvarinen, *Advances in neural information processing systems*, 293 (2003).
- [10] Z. F. Mainen and T. J. Sejnowski, *Science* **268**, 1503 (1995).
- [11] N. Brunel, *Journal of computational neuroscience* **8**, 183 (2000).
- [12] A. Destexhe, M. Rudolph, and D. Pare, *Nature Reviews Neuroscience* **4**, 739 (2003).
- [13] L. M. Ricciardi and L. Sacerdote, *Biological Cybernetics* **35**, 1 (1979).
- [14] L. M. Ricciardi and S. Sato, *Journal of Applied Probability* **25**, 43 (1988).
- [15] N. Brunel and S. Sergi, *Journal of Theoretical Biology* **195**, 87 (1998).
- [16] D. Pecevski, L. Buesing, and W. Maass, *PLoS computational biology* **7**, e1002294 (2011).
- [17] R. Salakhutdinov and G. E. Hinton, in *International Conference on Artificial Intelligence and Statistics* (2009) pp. 448–455.
- [18] A.-R. Mohamed, G. E. Dahl, and G. Hinton, *Audio, Speech, and Language Processing, IEEE Transactions on* **20**, 14 (2012).
- [19] G. Hinton, P. Dayan, B. Frey, and R. Neal, *Science* **268**, 1158 (1995).
- [20] R. Brette and W. Gerstner, *Journal of neurophysiology* **94**, 3637 (2005).
- [21] W. Gerstner and J. L. van Hemmen, *Network: Computation in Neural Systems* **3**, 139 (1992).
- [22] S. Mitra, S. Fusi, and G. Indiveri, *Biomedical Circuits and Systems, IEEE Transactions on* **3**, 32 (2009).
- [23] D. Brüderle, M. A. Petrovici, B. Vogginger, M. Ehrlich, T. Pfeil, S. Millner, A. Grübl, K. Wendt, E. Müller, M.-O. Schwartz, and et al., *Biological Cybernetics* **104**, 263 (2011).