A probabilistic population code based on neural sampling

Ankani Chattoraj, Shu Chen Wu, Richard D. Lange, Ralf M. Haefner

Brain and Cognitive Sciences, University of Rochester

swu32@u.rochester.edu, {achattor,rlange,rhaefne2}@ur.rochester.edu

Summary Visual processing is often characterized as implementing probabilistic inference: networks of neurons compute posterior beliefs over unobserved causes given the sensory inputs. How these beliefs are computed and represented by neural responses is much-debated (Fiser et al. 2010, Pouget et al. 2013), often focusing on whether the code is "explicit," with neurons representing concrete features of the outside world (e.g. Gabor-shaped patches for V1 neurons), or whether it is "implicit", with responses representing beliefs about abstract variables such as orientation (Pitkow & Angelaki 2017). A second debate concerns whether neural responses represent samples of latent variables (Hoyer & Hyvarinnen 2003) or parameters of their distributions (Ma et al. 2006). We present a model that exhibits key characteristics of both sampling and parametric codes, and that agrees with classic empirical data about neural responses in V1. We propose that V1 spikes represent binary samples from a linear model of the image. The spike rate in such a code is proportional to the marginal posterior probability over the variable represented by the neuron, making this at once a sampling-based as well as a parametric code. Previous work has shown that learning natural images in such a model yields localized, oriented, bandpass, receptive fields in agreement with empirical data (Bornschein et al. 2013). Correspondingly, neural responses in this model show tuning to external stimulus parameters like orientation and spatial frequency. Surprisingly, those tuning curves are approximately contrast-invariant with spike counts that are approximately Poisson-distributed – compatible with a linear probabilistic population code (PPC) for orientation. Finally, extending work by Buesing et al. (2011), we translate the sampling equations into a network of integrate-and-fire neurons that compute conditional probabilities in dendritic nonlinearities and generate spikes according to the correct posterior distribution given an input image.

Additional Detail

Relationship to previously proposed classes of neural codes: Pouget et al. (2013) distinguish between three types of neural representations of probabilities: 'linear probabilistic population codes' (linear PPC), 'codes proportional to probabilities', and 'sampling-based codes'. The code we propose, binary sampling from a linear model of the image (Fig. 1), can be interpreted as each of these three: its spike counts form a linear PPC with respect to orientation, its firing rate is proportional to the probability of Gabor-shaped features being present in the image, and its sequence of spikes converted into zero's and one's can be interpreted as a binary sampling code. In Fig. 2a+b we show that orientation tuning curves scale with the contrast of the presented grating for a toy-example of 50 identical rotated copies of the same projective field (Fig 2a) as well as for a full model based on 248 neurons (Fig 2b) with projective fields covering a range of orientations, spatial frequencies and positions (adopted from Orban et al. 2016). Fig. 2a demonstrates the close qualitative match between the actual sampling-derived tuning curves (dashed lines) and an approximation that we have derived in fixpoint form from the Gibbs sampling equations that allows us intuitive and analytical insights into the source of this surprising contrast-invariance:

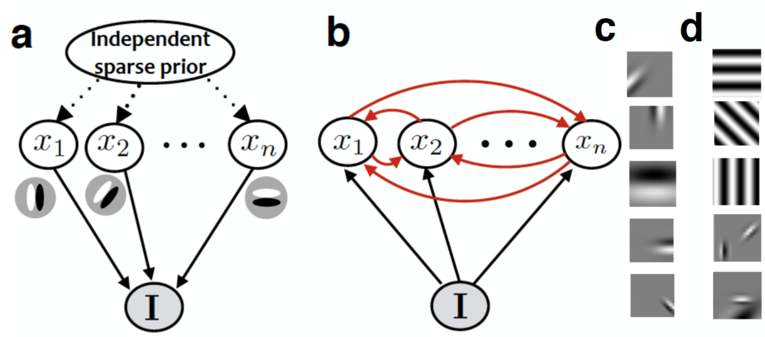
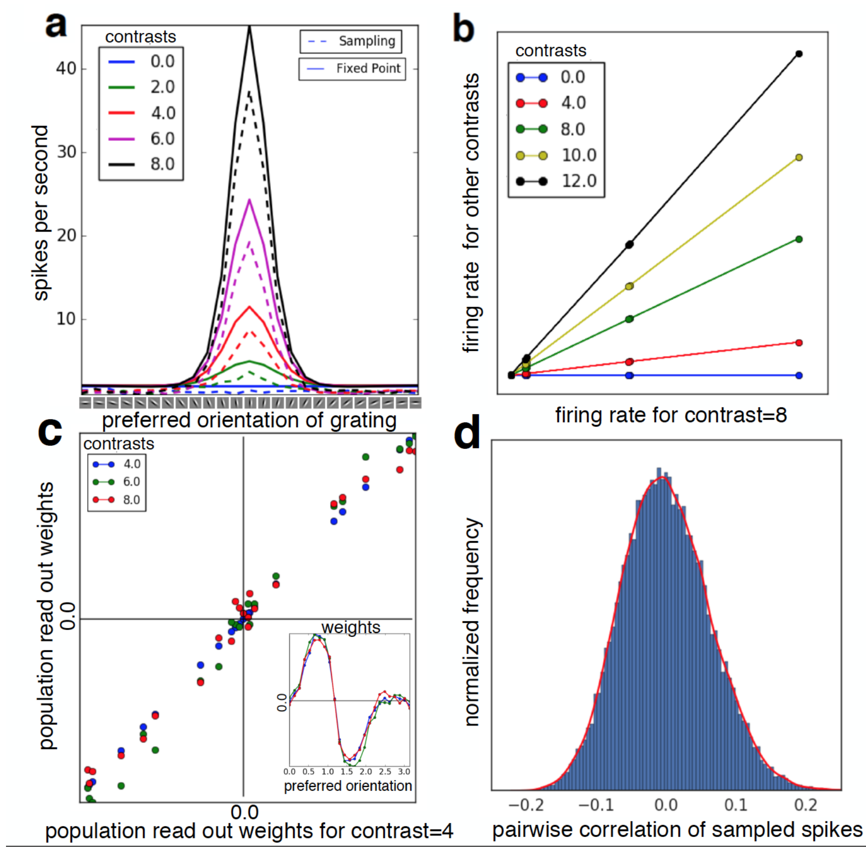


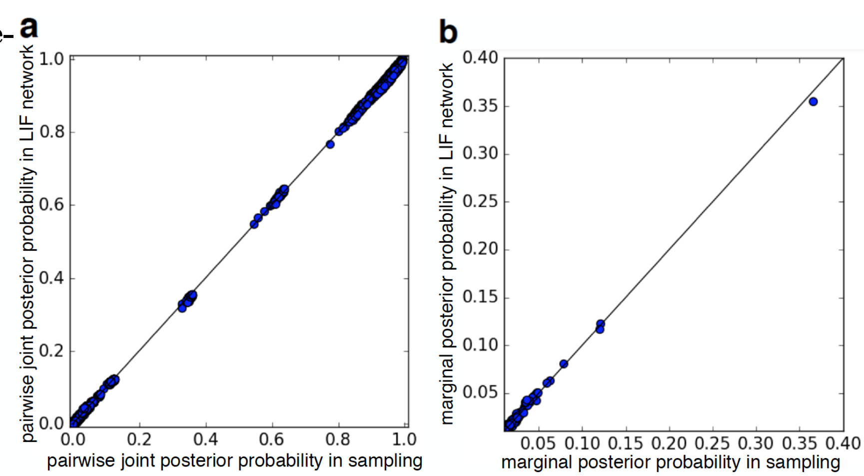
Figure 1: Setup. **a:** $p(\mathbf{I}|\mathbf{x}) = \mathcal{N}(\sum_i \mathbf{PF}_i x_i, 1)$. **b:** Information flow during inference. **c:** Example projective fields. **d:** Example images for inference.

$$\mathbb{E}[p(x_k)] \;=\; \text{sigmoid}\left(\frac{\mathbf{I}^\top \mathbf{PF}_k + \sum_{i,i\neq k} \mathbf{PF}_k^\top \mathbf{PF}_i \mathbb{E}[p(x_i)] + \frac{1}{2}\mathbf{PF}_k^\top \mathbf{PF}_k \mathbb{E}[p(x_k)]}{\sigma^2}\right).$$

Since spike counts in our model are approximately Poisson distributed (sums of Bernoulli variables with small probability, p), their variance scales with the mean and decoding weights for locally estimating orientation from neural responses are invariant to the nuisance variable contrast (Beck et al. 2008) (Fig. 2c) – a condition for our code to be interpretable as a linear PPC for orientation. Previously, direct probability codes had been thought to be in contradiction with direct probability codes since the latter were assumed to violate the contrast invariance of orientation tuning. It is therefore important to note that our code is a direct probability code over features, but *not* over orientation. Finally, Fig. 2d shows the distribution over noise correlation coefficients in our model – distributed around zero in agreement with published data (Ecker et al. 2010).



Relationship to prior work: Our code is closely related to the work of Hoyer & Hyvarinnen (2003), Orban et al. (2016), and Haefner et al. (2016) with the difference that latents are assumed to be binary rather than continuous, solving the problem that downstream neurons have to base their computations on the absence and presence of the actually received spikes rather than firing rates that need to be inferred over longer time scales, or inferred membrane potentials. It builds on the work by Buesing et al. (2011) who showed how probabilistic inference using binary variables may be implemented in networks of asynchronously spiking stochastic units. It also benefits from prior work by Bornschein et al. (2013) who showed that learning the distribution over natural images in the generative model used by us leads to receptive fields in agreement with empirical observations.



Implementation using leaky integrate-and-fire (LIF) neurons: Finally, we show how the Gibbs sampling equations can be straightforwardly translated into a network of LIF neurons whose spikes – while more regular than Poisson – map out the correct posterior conditioned on an input image. Simulating a network of 248 neurons with projective fields as in Orban et al. (2016), and comparing the spiking output to Gibbs samples from the same model, we find excellent agreement both for marginal and for the joint pairwise posterior probabilities (Fig. 3b+a). (Also see Petrovici et al. (2016) for an implementation different from ours based on LIF neurons in a high-conductance state.)