

Uncertainty through Sampling: The Correspondence of Monte Carlo Dropout and Spiking in Artificial Neural Networks

Kai Standvoss* (kstandvoss@gmail.com)

Donders Institute for Brain, Cognition, and Behavior, Radboud University, Montessorilaan 3,
6525 HR Nijmegen, The Netherlands

Lukas Grossberger* (lukas.grossberger@de.bosch.com)

Bosch Center for Artificial Intelligence, Robert Bosch GmbH, Robert-Bosch-Campus 1
71272 Renningen, Germany

* both authors contributed equally

Abstract

Any organism that senses its environment only has an incomplete and noisy perspective on the world, which creates a necessity for nervous systems to represent uncertainty. While the principles of encoding uncertainty in biological neural ensembles are still under investigation, deep learning became a popular and effective machine learning method. In these models, sampling through dropout has been proposed as a mechanism to encode uncertainty. Moreover, dropout has previously been linked to variability in spiking networks under specific assumptions. We compare the relationship between dropout and spiking neuron models by means of the variation ratio over their output. We demonstrate that in cases of incomplete world knowledge (epistemic uncertainty) as well as for noisy observations (aleatoric uncertainty) both neuron models show similar uncertainty representations. These findings provide evidence that sampling could play a fundamental role in representing uncertainties in neural systems.

Keywords: neural networks; uncertainty; rate-based; spiking; dropout; variation ratio

Introduction

Any organism that senses its environment only has an incomplete and noisy perspective on the world. Accordingly, representing uncertainty for the state estimation of the world poses a central challenge for nervous systems (von Helmholtz, 1867). Since the initial formulation of those ideas, the encoding of uncertainty in ensembles of neurons is still an open field of research (Vilares & Kording, 2011).

In the field of machine learning, interconnected neurons inspired deep learning, a recently popular and effective family of methods (Goodfellow, Bengio, & Courville, 2016). These methods are not only loosely inspired by nervous systems but have also been used in turn to analyse biological neural networks (Güçlü & van Gerven, 2015). Cichy and Kaiser (2019) have argued that deep learning models can serve as valuable scientific tools for exploratory research in computational neuroscience.

Introduced as regularisation to prevent overfitting and support generalisation, dropout has become a state of the art addition to deep learning models (Srivastava, Hinton, Krizhevsky, Sutskever, & Salakhutdinov, 2014). The idea of “dropping” neurons with a certain probability from a learning episode both intuitively relates to the variability found in its biological counterparts, and has been linked analytically under specific assumptions as well (Baldi & Sadowski, 2014). This link has been used to improve training and model performance of artificial spiking networks (Neftci, Pedroni, Joshi, Al-Shedivat, & Cauwenberghs, 2016). Gal and Ghahramani (2016) showed that dropout cannot only be used to improve model training but can be interpreted as approximate Bayesian inference. Applying dropout for unseen examples allows for obtaining estimates of the model’s uncertainty in its predictions. Given the theoretical relationship between dropout in rate-based deep learning networks and variable firing of neurons in biological spiking networks, the question emerges whether spiking in artificial neural networks can play a similar role in representing uncertainty, to dropout in classical deep learning architectures.

In this work, we empirically investigate the relationship between dropout and artificial spiking neurons in case of an incomplete knowledge of the world (epistemic uncertainty) as well as for noisy observations (aleatoric uncertainty). Since distinguishing the two types of uncertainty by measuring properties of the model output is difficult (Smith & Gal, 2018), we investigate the two uncertainty scenarios separately. We focus on the sampling view of uncertainty encoding in neural ensembles (Vilares & Kording, 2011). This view holds that the variability in neural firing rates represents uncertainty by sampling from the posterior probability distribution (Buesing, Bill, Nessler, & Maass, 2011). Based on this theoretical account we analyse the correspondence of uncertainty representations in spiking network models and Monte Carlo dropout in rate-based neural networks. We classify MNIST digits with a multi layer perceptron where the hidden layer either consists of traditional rate-based units with dropout, or spiking leaky integrate and fire neurons (Figure 1). We quantify the uncertainty over multiple forward passes or time steps of the predictive model output with the variation ratio (Gal & Ghahramani, 2016). Rate-based and spiking neuron models show highly similar uncertainty



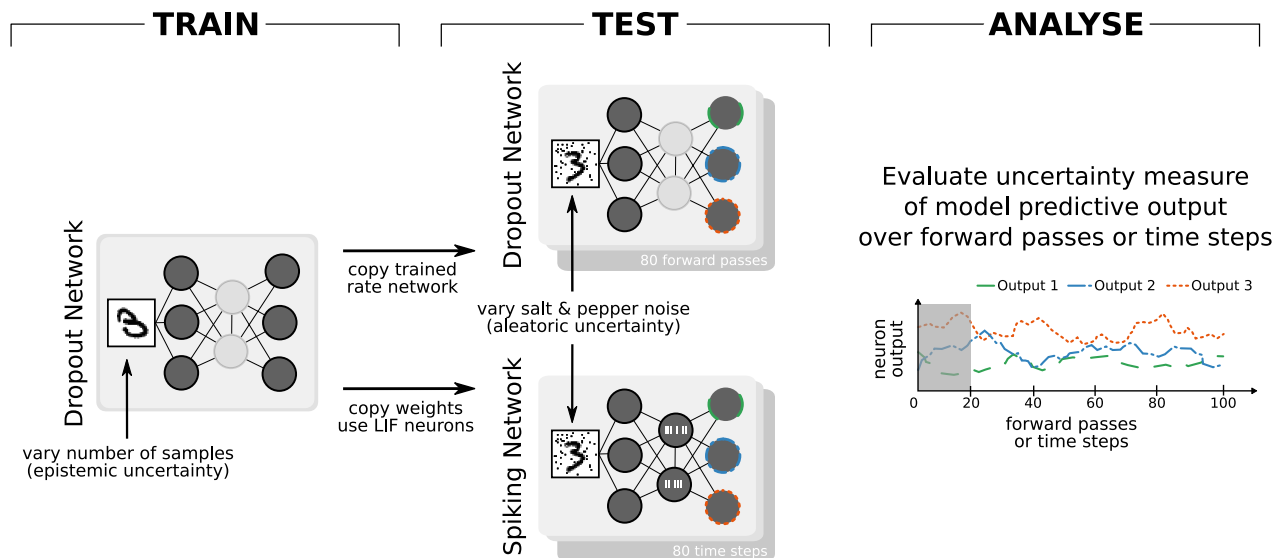


Figure 1: A traditional, rate-based multi layer perceptron (784:512:10) is trained on MNIST with dropout at the hidden layer. The same model is used to sample 80 Monte Carlo dropout predictions per test image. In the bottom case the hidden units are turned into leaky integrate and fire (LIF) neurons, and the predictive output for each input is recorded over 80 time steps. Finally, the variation ratio (Eq. 1) quantifies model predictive uncertainty from the output distributions; while either the size of the training set or the amount of salt & pepper noise on the test set is varied to manipulate epistemic and aleatoric uncertainty respectively.

representations, quantified by the variation ratio, for epistemic uncertainty (Figure 2) as well as for the aleatoric case (Figure 3). Showing that the variability introduced through spiking represents uncertainty similarly to dropout, provides evidence for the relevance of spiking for inference in biological neural networks.

Methods

A simple multi layer perceptron was trained using NengoDL (Rasmussen, 2018), which allows converting rate-based neural networks into their spiking counterparts. The model is trained as a rate-based network by setting the activation functions to a continuous approximation of the leaky integrate and fire (LIF) neuron (Hunsberger & Eliasmith, 2015). Therefore, the model can be run as a regular rate-based network during test time, or alternatively used as a spiking network by replacing the continuous LIF approximation by spiking LIF units. For the task at hand, a model with one fully-connected hidden layer of 512 units with smooth LIF activation function was trained. Neurons were deactivated with dropout probability of 40%. The logit activations of the last layer were normalised using a Softmax function to represent a probability distribution. The network was trained to minimise the cross-entropy between the target and the network output.

To test the representations of both epistemic as well as aleatoric uncertainty in artificial neural networks, two experiments were performed that systematically varied one source of uncertainty individually. Epistemic uncertainty was modulated by training separate networks with a varying number of training

examples, ranging from 2% to 0.1% (i.e. 1200 to 60 random training images). With only a few examples, a model's ability to learn is impaired and it should thus exhibit high epistemic uncertainty. With an increasing amount of data the model should become more confident. In the second experiment, a single model was trained on the full, noise free MNIST training dataset (60000 images). Additionally, for each salt & pepper noise (cf. below) level of 0.1, 0.2, 0.3, 0.4, 0.5, we added a random subset of size 5000 from all training images to provide sufficiently many noisy examples to the model during training. Training the model on the full dataset and noisy exemplars should ensure the absence of epistemic uncertainty. In other words, it prevents the model from leaving the training data manifold when confronted with high levels of noise (i.e. aleatoric uncertainty). Aleatoric uncertainty was then varied by applying an increasing amount of noise to the test data, observing changes in the predictive output of the same underlying model. Noise was introduced to the images by randomly setting individual pixels to either 0 or 1 with varying probability (salt & pepper noise). In this setting the model is expected to show increased predictive uncertainty the more noise is applied to the test data.

Uncertainty was quantified using the variation ratio (Eq. 1) during test time for both the rate-based as well as for the spiking model using the same, initially trained model weights. The uncertainty analysis in the rate setting is based on Monte Carlo dropout as described by Gal and Ghahramani (2016), where 80 Monte Carlo samples were drawn from the approximate posterior using different dropout masks. In this setting the for-

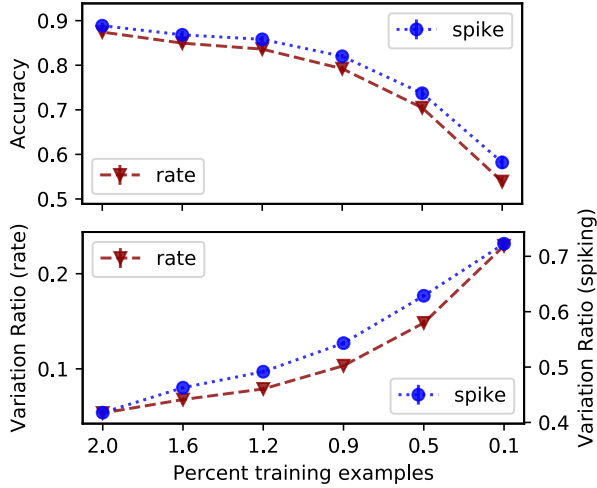


Figure 2: Varying the amount of training examples (epistemic uncertainty) the top plot shows classification accuracy and the bottom plot shows the variation ratio, both on the test set. The rate model results are presented in red, coarsely dashed and triangular markers. Results from the spiking model are presented in blue, fine dashed, round markers. The bottom plot features two individual secondary axes for the variation ratio. The error bars, representing a single standard deviation for 5 independent test runs, are smaller than the markers.

ward passes are independent from each other, stochastically silencing neurons and thereby altering the network activation pattern. In the spiking model, the same input stimulus was propagated through the network for 100 time steps. For every consecutive time step the LIF neurons integrate their input, are silent until their critical activation threshold is reached, and become inactive again during their refractory period. Thereby, despite being deterministic, spiking units introduce variability over time steps. Since roughly 20 time steps are needed to reach a stable firing rate distribution, we only used the last 80 time steps to evaluate the uncertainty.

The variation ratio is defined as the proportion of outcomes that are **not** in the mode category. Accordingly, for S total predictions and a mode (most frequent) prediction occurring s_m times the variation ratio v is given by:

$$v := 1 - \frac{s_m}{S} \quad (1)$$

With higher certainty the network is expected to predict the same class more frequently such that the variation ratio is low and the other way around. This measure does not distinguish, by itself, between aleatoric and epistemic uncertainty.

Results

In figure 2 results are shown for a varying amount of training examples (epistemic uncertainty). The spiking model variant as well as the version with dropout perform highly similar with respect to the accuracy and variation ratio. It can be seen

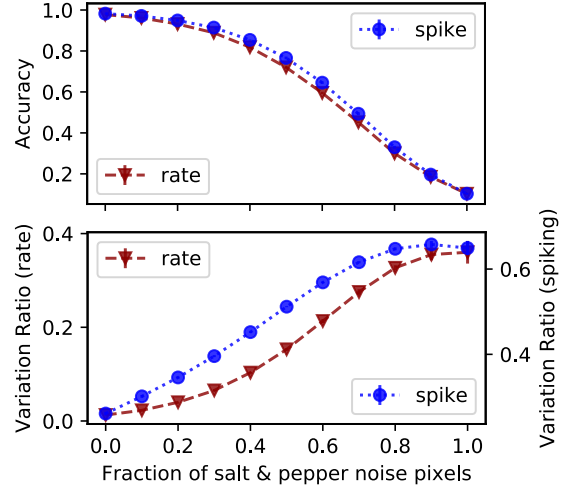


Figure 3: Varying the amount of salt & pepper noise (aleatoric uncertainty) the top plot shows classification accuracy and the bottom plot shows the variation ratio, both on the test set. The rate model results are presented in red, coarsely dashed and triangular markers. Results from the spiking model are presented in blue, fine dashed, round markers. The bottom plot features two individual secondary axes for the variation ratio. The error bars, representing a single standard deviation for 5 independent test runs, are smaller than the markers.

that for both models the accuracy decreases with decreasing amounts of training data. At the same time the model uncertainty, measured by the variation ratio, increases.

Figure 3 shows results for applying an increasing degree of salt & pepper noise (aleatoric uncertainty). The spiking model variant as well as the version with dropout show similar classification test accuracy due to the same network weights being used, displaying monotonically decreasing scores for higher noise levels. The variation ratio is highly correlated (Spearman rank correlation, $r = 1$. and $p < 0.0001$) for both models. The uncertainty metric in both models increases proportional to the applied noise. In both experiments it can be observed that the variation ratio of the spiking network is constantly offset by about 0.3.

Discussion

Representing epistemic uncertainty, that originates from an incomplete view of the world, as well as aleatoric uncertainty, due to noisy senses, is a key challenge for nervous systems. While artificial neural networks from the field of machine learning are partly inspired by brains, the degree of their correspondence and how well they can help to understand their spiking counterpart are open research questions. In this work we compare Monte Carlo dropout with leaky integrate and fire neurons as means to introduce variability and thereby allow for the representation of uncertainty in artificial neural networks.

Our results show a strong correlation between uncertainty quantified by the variation ratio for Monte Carlo dropout and a

network of spiking neurons, even though the variation ratios of the spiking network are constantly offset. For both, epistemic and aleatoric uncertainty, the models' variation ratio increases when trained with fewer samples and increases proportional to the fraction of applied salt & pepper noise, respectively. This means that in both cases the model outputs vary to similar degrees. In cases where uncertainty calibration is relevant, the observed offset might matter. Whereas in cases, where only quantifying relative changes in uncertainty matters, this offset is not important. The offset in the variation ratio for the spiking network originates from more variability due to different inter-activation interval distributions. The inter-activation intervals for spiking neurons follow a Poisson distribution with only few neurons that are very active. In contrast, dropout results in neurons either being inactive or active with a fixed probability resulting in a bimodal distribution, with one mode at zero and the other at one minus the dropout probability. The observed offset is reduced by increasing the dropout probability.

The mechanisms for encoding uncertainty in the brain are still subject to active investigation and debate. Meanwhile, increasing experimental evidence suggests that probability distributions could be neurally encoded by representing their mean and variance through sampling (Orbán, Berkes, Fiser, & Lengyel, 2016), where said variance represents uncertainty. Our results provide another puzzle piece of evidence for this view. With this research we follow a line of studies that use artificial neural network models as scientific tools that allow for easier exploration than large scale human or animal studies while still providing compelling complementary evidence (Cichy & Kaiser, 2019; Güçlü & van Gerven, 2015). Therefore, our results do not only support the sampling view of uncertainty encoding and guide future research about the precise underlying mechanisms, we also reconfirm the value of computational research using artificial neural networks for neuroscience.

These insights provide the basis for further investigation of uncertainty encoding in more complex and biologically more plausible artificial neural networks models. Especially for questions regarding the role such an uncertainty representation might take in the context of hierarchical computation, leveraging uncertainty information to guide agent environment interactions. Such hierarchical models require a shift from the output focused decoding perspective taken by using the variation ratio of the model output, towards investigating uncertainty encoding in latent feature spaces. One approach could be to investigate the variability of such representations, formed in more complex neural networks, with increasing levels of uncertainty. This could provide insights in the mechanisms underlying previous experimental findings (Orbán et al., 2016).

To conclude, we provide evidence that stochasticity introduced by dropout during forward passes displays a similar uncertainty representation as caused by variable firing in spiking LIF networks, following the sampling view of uncertainty encoding (Vilares & Kording, 2011), quantified by the variation ratio. Thereby, we add another building block to describing the relationship between rate-based and spiking artificial neural

networks, bridging between the fields of machine learning and computational neuroscience.

Acknowledgments

We would like to thank B. Ottervanger for supporting the work that preceded this publication, as well as F. P. Battaglia, M. Pfeiffer and S. Quax for their helpful comments on the draft.

References

- Baldi, P., & Sadowski, P. (2014). The dropout learning algorithm. *Artificial intelligence*, *210*, 78–122.
- Buesing, L., Bill, J., Nessler, B., & Maass, W. (2011). Neural dynamics as sampling: a model for stochastic computation in recurrent networks of spiking neurons. *PLoS computational biology*, *7*(11), e1002211.
- Cichy, R. M., & Kaiser, D. (2019). Deep neural networks as scientific models. *Trends in cognitive sciences*.
- Gal, Y., & Ghahramani, Z. (2016). Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *Proceedings of the international conference on machine learning* (pp. 1050–1059).
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.
- Güçlü, U., & van Gerven, M. A. J. (2015). Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. *Journal of Neuroscience*, *35*(27), 10005–10014.
- Hunsberger, E., & Eliasmith, C. (2015). Spiking deep networks with lif neurons. *arXiv preprint arXiv:1510.08829*.
- Neftci, E. O., Pedroni, B. U., Joshi, S., Al-Shedivat, M., & Cauwenberghs, G. (2016). Stochastic synapses enable efficient brain-inspired learning machines. *Frontiers in neuroscience*, *10*, 241.
- Orbán, G., Berkes, P., Fiser, J., & Lengyel, M. (2016). Neural variability and sampling-based probabilistic representations in the visual cortex. *Neuron*, *92*(2), 530–543.
- Rasmussen, D. (2018). NengoDL: Combining deep learning and neuromorphic modelling methods. *arXiv*, *1805.11144*, 1–22.
- Smith, L., & Gal, Y. (2018). Understanding measures of uncertainty for adversarial example detection. *arXiv preprint arXiv:1803.08533*.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, *15*(1), 1929–1958.
- Vilares, I., & Kording, K. (2011). Bayesian models: the structure of the world, uncertainty, behavior, and the brain. *Annals of the New York Academy of Sciences*, *1224*(1), 22–39.
- von Helmholtz, H. (1867). *Treatise on physiological optics* vol. iii.