Computational roles of plastic probabilistic synapses

Milton Llera[1], Joao Sacramento[2] and Rui Ponte Costa[1,2,3,*]

1. Computational Neuroscience Unit
Department of Computer Science, Faculty of Engineering
and Bristol Neuroscience
University of Bristol, United Kingdom

2. Department of Physiology
University of Bern, Switzerland

3. Centre for Neural Circuits and Behaviour
Department of Physiology, Anatomy and Genetics
University of Oxford, United Kingdom

Abstract

The probabilistic nature of synaptic transmission has remained enigmatic. However, recent developments have started to shed light on why the brain may rely on probabilistic synapses. Here, we start out by reviewing experimental evidence on the specificity and plasticity of synaptic response statistics. Next, we overview different computational perspectives on the function of plastic probabilistic synapses for constrained, statistical and deep learning. We highlight that all of these views require some form of optimisation of probabilistic synapses, which has recently gained support from theoretical analysis of long-term synaptic plasticity experiments. Finally, we contrast these different computational views and propose avenues for future research. Overall, we argue that the time is ripe for a better understanding of the computational functions of probabilistic synapses.

* Correspondence: rui.costa@bristol.ac.uk

Highlights

- Computational and experimental research suggest that synapses adapt their transmission statistics during learning
- Optimisation of probabilistic synapses occurs jointly in pre- and postsynaptic terminals during synaptic plasticity
- Recent developments in statistical learning point to a reevaluation of the function of probabilistic synapses in cortical circuits
- Insights on the biology of probabilistic synapses may inspire new learning algorithms

1 Introduction

Animals have evolved in uncertain environments. For example, they have adapted to distinguish nutrition sources of different shapes, sizes, colours and tastes. Such perceptual uncertainty should be encoded by the brain to enable accurate decision making (Fiser et al., 2010; Orbán et al., 2016; Haefner et al., 2016). This link between perception and decision is presumably achieved through communication between different brain areas, which ultimately relies on synaptic transmission (Nabavi et al., 2014; Roelfsema and Holtmaat, 2018). Synaptic transmission is inherently stochastic: a presynaptic action potential may or may not trigger neurotransmitter release that in turn binds to postsynaptic receptors (Malagon et al., 2016). For synaptic transmission to successfully trigger a behavioural decision synaptic response statistics should be tuned during learning (Nabavi et al., 2014; Costa et al., 2017b; Roelfsema and Holtmaat, 2018). However, it has remained unclear exactly which aspects of probabilistic synapses should be modified during learning.

There is wide evidence of plasticity occurring at the key components that underlie synaptic transmission statistics. For example, not only does plasticity change the properties and number of postsynaptic receptors, but also the intricate presynaptic machinery responsible for stochastic neurotransmitter release (Padamsey and Emptage, 2014; Costa et al., 2017b). Because synaptic plasticity is believed to underlie learning (Nabavi et al., 2014; Roelfsema and Holtmaat, 2018), this body of experimental work suggests that the brain shapes probabilistic synapses as animals adapt to the environment. This has important theoretical implications (Kappel et al., 2015; Aitchison and Latham, 2015; Blundell et al., 2015; Costa et al., 2015), but most computational models of learning and synaptic plasticity have considered only changes in the mean synaptic weight (e.g. Brea et al. (2016); Bittner et al. (2017); Pereira and Brunel (2018)). Below we review recent experimental and theoretical developments on the plasticity and computation roles of probabilistic synapses.

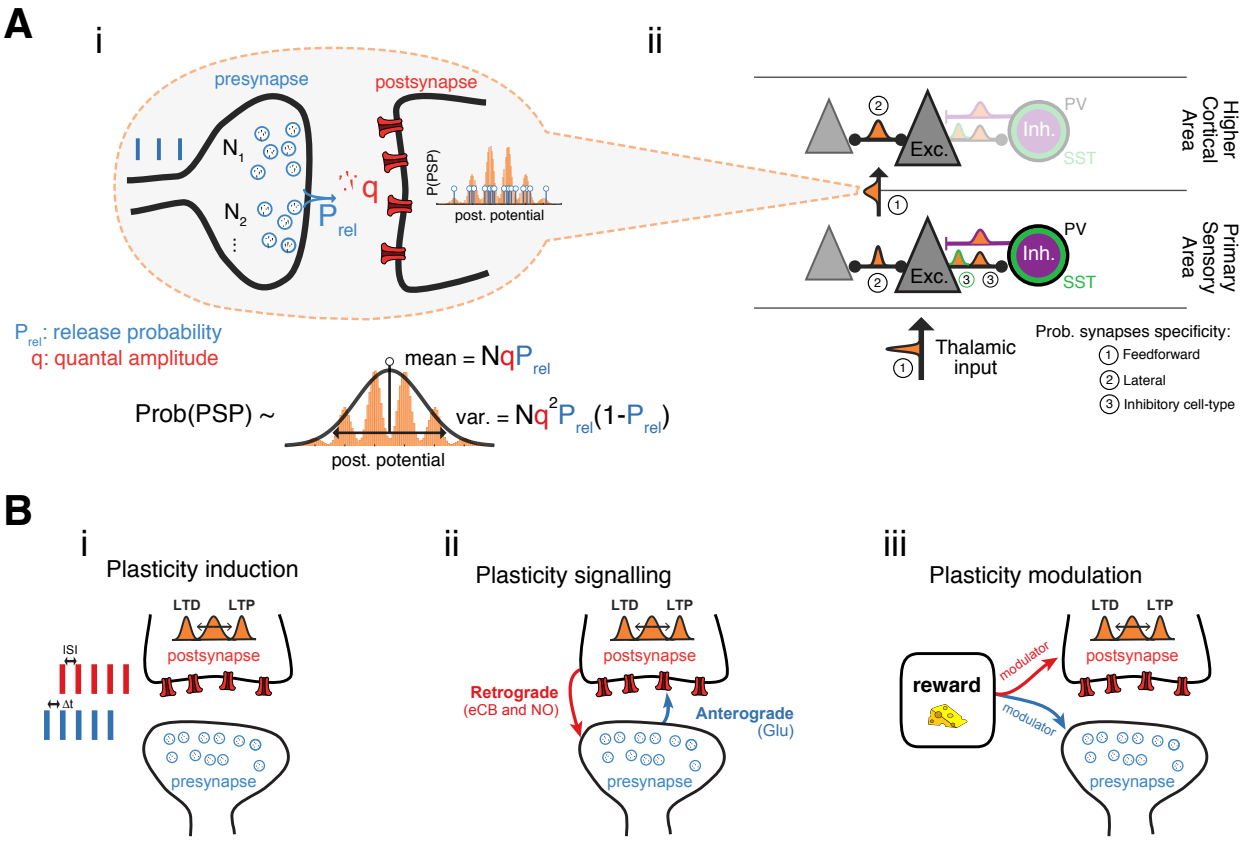

Figure 1: Specificity and plasticity of probabilistic synapses. (A) Throughout the brain virtually every synapse is probabilistic. (i) When a presynaptic spike (blue vertical line on the left) occurs a presynaptic vesicle (blue circles) may release neurotransmitters (red dots) that bind to postsynaptic receptors (red) which elicits a postsynaptic potential (PSP; PSPs of different amplitudes are represented by the small vertical blue lines). The key parameters that determine the statistics of probabilistic synaptic release are the number of presynaptic release sites (N, groups of vesicles in blue; only two release sites are represented, N_1 and N_2, out of the five modelled here), release probability ($P_{\rm rel}$, blue arrows) and quantal amplitude which is proportional to the number of postsynaptic receptors, (q, red). This process is typically modelled as a binomial probability distribution (orange histogram, with N=5, $P_{\rm rel} = 0.5$ and $q = 1$), which in the limit of large N can be approximated as a Gaussian distribution (black line) with mean=$NqP_{\rm rel}$ and variance=$Nq^2P_{\rm rel}(1-P_{\rm rel})$. (ii) Simplified representation of cortical circuits, with both excitatory (black) and inhibitory (purple) synapses and neuron types. Each synaptic connection is stochastic (represented as a Gaussian distribution). Two different inhibitory cell-types are represented: somatostatin (SST, dashed green circle) and parvalbumin (PV, black circle); here these two separate inhibitory cell-types are represented as overlapping circles for simplicity. Note that different connections exhibit statistics of different means and variances (see main text for more details). (B) Long-term plasticity of probabilistic synapses. (i) Different induction protocols have been shown to trigger changes in the probability of postsynaptic responses. Schematic on the left represents pre- and postsynaptic spikes in a spike-timing-dependent plasticity protocol, which depending on the timing between pre- and postsynaptic spikes (Δt) as well as the inter-spike interval (ISI) may lead to long-term potentiation (LTP) or depression (LTD). This in turn changes not only the mean synaptic response, but also its variance. (ii) Modifications to probabilistic synapses during plasticity are known to rely on specific retrograde (e.g. endocannabinoids (eCB) and nitric oxide (NO)) and anterograde signals (glutamate (Glu)). (iii) Behavioural outcomes (e.g. reward) may rely on neuromodulation (e.g. Dopamine) to regulate plasticity at probabilistic synapses.

2 Specificity of synaptic transmission statistics

The probabilistic nature of synaptic transmission has been described as a binomial process (Del Castillo and Katz, 1954; Malagon et al., 2016; Costa et al., 2017b), which is parametrised by the (i) number of synaptic release sites N, (ii) presynaptic release probability P_{rel} and (iii) quantal amplitude q – proportional to the number of postsynaptic receptors [1] (Fig. 1Ai). Together these three parameters define the statistics of synaptic responses, with mean given by NqP_{rel} and variance by $Nq^2P_{\mathrm{rel}}(1-P_{\mathrm{rel}})$ (Fig. 1Ai).

The exact mean and variance of synaptic transmission depends on where the synapse is located. In cortical circuits the statistics (e.g. means and variances) of synaptic responses exhibit a high degree variability that depends on cell-type (Brémaud et al., 2007), connection-type (Brémaud et al., 2007; Blackman et al., 2013; Costa et al., 2013), layer (Brémaud et al., 2007; Thomson, 2007), brain area (Wang et al., 2006), age (Reyes and Sakmann, 1999), and even species (Testa-Silva et al., 2014). For example, excitatory synapses from thalamic projections onto layer-4 granule cells are more reliable (Silver, 2003) than synapses between layer-5 pyramidal cells (Costa et al., 2013). Remarkably, connections from pyramidal cells onto lateral inhibitory cells can also be dramatically different: synapses onto somatostatin-positive interneurons cells communicate with a low basal release probability, whereas synapses onto parvalbumin-positive interneurons are stronger with higher release probability (Blackman et al., 2013; Costa et al., 2013) (Fig. 1Aii). Such high specificity of probabilistic synapses suggests that they are modified during learning.

3 Plasticity of probabilistic synapses

Accumulating evidence suggests that synaptic plasticity underlies learning in the brain (Nabavi et al., 2014; Roelfsema and Holtmaat, 2018). Synaptic plasticity not only modifies the mean synaptic response, but also its variance (Fig. 1B). In particular, it has been shown that long-term synaptic plasticity leads to changes in both the presynapse by modifying P_{rel} and the postsynapse by modifying q (Costa et al., 2017b) (Fig. 1A,Bi). After a decade-long debate, today it is widely accepted that both pre- and postsynaptic physiology can be modified during long-term potentiation (LTP) and depression (LTD) (Padamsey and Emptage, 2014; Costa et al., 2017b). However, exactly how much each component is changed can have a dramatic impact on the synaptic transmission statistics (Costa et al., 2015). Using a synaptic plasticity model tuned to pre- and postsynaptic plasticity Costa et al. (2015) demonstrated that both mean and variance of synaptic responses are regulated both in vitro and in perceptual learning experiments performed in the primary auditory cortex of rats (Froemke et al., 2013). Interestingly, there are homeostatic forms of plasticity at the presynapse that compensate for altered postsynaptic function (Li et al., 2018) and modifications to the number of release sites N during long-term plasticity (Loebel et al., 2013; Tang et al., 2016), which may also shape the synaptic transmission

[1] This is a simplified view of the complicated release machinery. For example, the quantal amplitude q also depends on the amount of neurotransmitter per (presynaptic) vesicle and on the sensitisation of postsynaptic receptors.

statistics.

Although the expression of synaptic plasticity can be presynaptic, its induction depends on postsynaptic activity (Monday and Castillo, 2017). This implies the need for retrograde signals that communicate with the presynapse. Two main signals have been identified: nitric oxide, which is responsible for presynaptic LTP, and endocannabinoids, which mediates (in part) presynaptic LTD (Andrade-Talavera et al., 2016; Monday and Castillo, 2017) (Fig. 1Bii). Interestingly, recent evidence shows that deficits in the retrograde signalling systems of both nitric oxide and endocannabinoids have been implicated in learning and memory impairments, anxiety and depression (Monday and Castillo, 2017). This may be due to a failure in correctly adjusting probabilistic synapses during plasticity (Hebert-Chatelain et al., 2016; Monday and Castillo, 2017).

Synaptic modifications should ultimately lead to more successful behavioural outcomes. Reward-based synaptic plasticity provides a framework in which synapses are modified by specific neuromodulators conveying behaviour relevant information (Frémaux and Gerstner, 2016). One such neuromodulator is dopamine, which is known to correlate with reward (Stauffer et al., 2016). Moreover, dopamine and other neuromodulators regulate long-term synaptic plasticity (Pawlak et al., 2010; Frémaux and Gerstner, 2016), suggesting that they may also control learning at probabilistic synapses (Fig. 1Biii). This is consistent with recent results on neuromodulation of presynaptic long-term plasticity (Monday and Castillo, 2017), which has also been observed *in vivo* in Drosophila (Cohn et al., 2015).

4 Computational roles of probabilistic synapses

Despite the growing body of experimental observations suggesting a precise control of probabilistic synapses, it has remained unclear how these relate to computational functions. Below we highlight three key computational roles of probabilistic synapses and how they may be reconciled with experimental findings.

4.1 Biophysical constraint

It is conceivable that due to high energetic costs associated with neurotransmitter transmission synapses remain unreliable unless necessary (Harris et al., 2012) (Fig. 3A). This view suggest that only synapses that are important for a given neuronal representation or behaviour should become reliable (see Aitchison et al. (2018) for a similar argument at the neuronal level). Consistent with this hypothesis mathematical modelling of slice long-term synaptic plasticity experiments showed that after induction of long-term plasticity synapses become more reliable (Costa et al., 2015). This result was further supported by reanalyses of *in vivo* sensory perception experiments (Froemke et al., 2013; Costa et al., 2015). However, it has remained unclear whether synapses not only become more reliable, but aim for the most reliable state (i.e. minimal variance). Recently, Costa et al. (2017c) put forward a model in which synapses optimise their response statistics during long-term synaptic plasticity towards reliable responses (i.e. with a given mean and

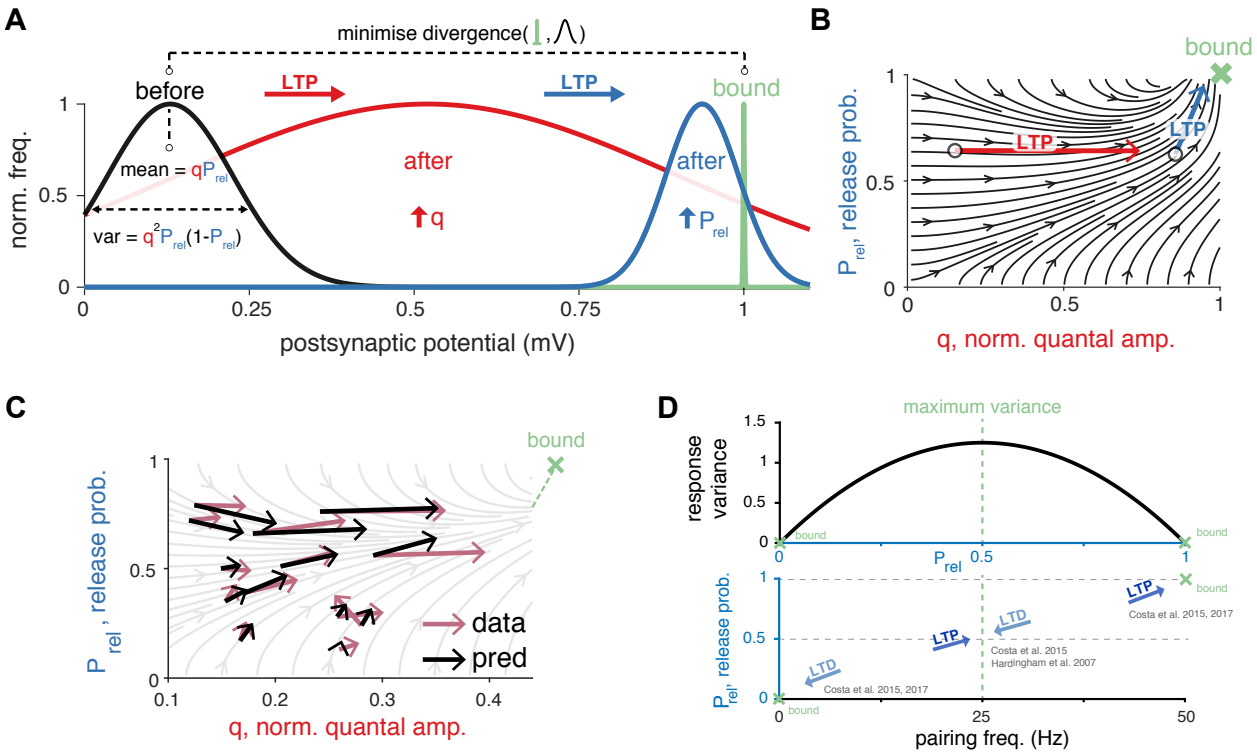

Figure 2: Statistical long-term synaptic plasticity (*stat*LTSP). (A) The theory proposes that during long-term potentiation (LTP) synapses optimise their response statistics towards reliable responses (i.e. they minimise the divergence between their current statistics and a upper bound). This can be achieved by modifying either the postsynapse (through changes in q, red) and the presynapse (through changes in P_{rel}, blue). (B) The *Stat*LTSP proposal predicts a flow field of pre- and postsynaptic changes that depend on the current state of the synapse (given a Euclidean-metric and normalised q). (C) Theory (black) captures single experiment variability (purple) of LTP induction in Hippocampus. Predicted flow field in the background (grey). (D) Frequency-dependent uncertainty encoding of synaptic plasticity. Plasticity experiments suggest that not only synapses aim for reliable responses when stimulated at high frequencies (long-term potentiation, as in (A-C))) or low frequencies (long-term depression) (Costa et al., 2015, 2017c), but also that at intermediate frequencies ($\sim$ 25Hz) synapses aim for maximum unreliability by setting $P_{\text{rel}} \sim 0.5$ (dashed green line, Hardingham et al. (2007); Costa et al. (2015)). Synaptic response variance (top) is calculated using standard binomial release statistics as $Nq^2P_{\text{rel}}(1-P_{\text{rel}})$, with $q = 1$ and $N = 5$. Bottom panel illustrates the different release probability end points as a function of long-term plasticity pairing frequency. Figure partly adapted from Costa et al. (2017c).

minimal variance) referred to as statistical long-term synaptic plasticity (*stat*LTSP; Fig. 2A).

*Stat*LTSP suggests a gradual optimisation process of synaptic transmission towards a reliable target synaptic weight (or bound) that should be triggered with every plastic event (Fig. 2B). This theory can explain a wide range of apparently disparate observations of long-term potentiation (LTP) at hippocampal and visual cortex excitatory synapses. For long-term depression (LTD), *stat*LTSP predicts presynaptic expression of plasticity – i.e. changing P_{rel} more rapidly decreases the synaptic response statistics towards a lower reliable target. Importantly, the model captures changes in the synaptic transmission statistics (pre- and postsynaptic) of individual recordings (Fig. 2C) [2]. How exactly would *stat*LTSP be implemented at synapses remains unclear. Nevertheless, Costa et al. (2017c) identified nitric oxide and

[2]This model is based on standard gradient descent using an euclidean-metric and normalised q, cross-validated using several datasets.

endocannabinoids as retrograde signals (Fig. 1Bii) encoding errors in q and P_{rel} consistent with the predictions.

Taken together this body of work suggest that long-term plasticity aims to reduce synaptic unreliability, consistent with the constraint view of stochastic synapses. However, this does not necessarily imply that synapses end up being reliable. First, in the intact brain during learning a mixture of LTP and LTD is likely to occur, which would maintain or increase response variability; second homeostatic mechanisms may control reliability due to its high energetic costs (as discussed above) and third, there are protocols (typically at intermediate frequencies, ~ 25Hz) that appear to maximise synaptic response variability (i.e. aiming for $P_{\text{rel}} = 0.5$; Hardingham et al. (2007); Costa et al. (2015); Fig. 2D). Interestingly, this last observation suggests a frequency-dependent variance encoding – whether synapses aim for minimal or maximal variance depends on the firing rate of pre- and postsynaptic neuron. The framework discussed here only aims to optimise the synaptic response variability without a clear behaviourally relevant task. But, it should be possible to extend *stat*LTSP to explicitly relate synaptic response variability to task-relevant uncertainty encoding.

4.2 Encoding perceptual uncertainty

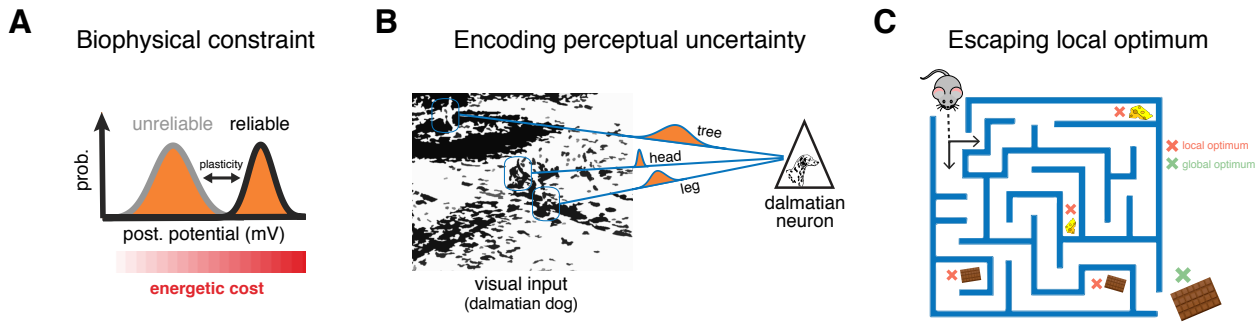

Figure 3: Computational roles of plastic probabilistic synapses. (A) Biophysical constraints, such as limited energy supply (Harris et al., 2012) may only allow reliable synapses to develop if necessary (e.g. during long-term plasticity) due to the high energetic costs (represented by red colour bar) of reliable synaptic transmission. (B) It has been postulated that the brain should also encode sensory statistics. Neurons in the brain responding to specific visual objects (e.g. dalmatian dog) should combine contextual information when inferring the presence or absence of an object. For a dalmatian neuron it would be important to integrate visual features such trees, dog head and animal legs (blue boxes). The uncertainty of the connections representing different features should be proportional to how relevant that feature is for that particular object. (C) Plastic probabilistic synapses have also been suggested to enable neural networks to find better solutions, escaping local optimum. For example, as animals explore an environment adaptive probabilistic synapses might enable animals to find better global paths.

To maximise chances of survival animals should encode perceptual uncertainty associated with the environment in which they live (Fiser et al., 2010) (Fig. 3B). A principled framework often used to describe how the brain may encode perceptual uncertainty is that of Bayesian inference (Berger, 2013). According to the Bayesian inference hypothesis the brain computes the posterior probability over latent variables (e.g. predators) given sensory stimuli (e.g. visual stimuli) $P(\text{latent} \mid \text{stimuli})$, which combines prior beliefs over the latent variables $P(\text{latent})$ with the incoming sensory evidence (likelihood) $P(\text{stimuli} \mid \text{latent})$.

A growing body of work suggests that cortical circuits encode perceptual uncertainty following Bayesian inference ideas (Fiser et al., 2010; Ma and Jazayeri, 2014; Orbán et al., 2016; Haefner et al., 2016). Exactly how such encoding of perceptual uncertainty may be used or learned at the synaptic level has remained unclear. Recent proposals have put forward the notion of synaptic sampling (Aitchison and Latham, 2015; Kappel et al., 2015, 2018), in which each synaptic release or structure (i.e. dendritic spine and axonal bouton) can be interpreted as a sample from a specific posterior distribution. Synaptic sampling can in principle be used by postsynaptic neurons to estimate posterior distributions (and uncertainty) over synaptic weights. For example, Kappel et al. (2015) demonstrated that spine motility (structural dynamics) similar to that observed experimentally (Mongillo et al., 2017) can be interpreted as sampling from a posterior distribution over neural network configurations. This framework was recently extended to reward-based learning, thus adding a behaviourally-relevant component to previous work by the same authors (Kappel et al., 2018). An alternative approach is that of Aitchison and Latham (2015) in which posterior distributions representing task uncertainty are formally encoded over synaptic weights distribution presumably through long-term synaptic plasticity. It should be noted that none of these frameworks consider synapse release (binomial) statistics as introduced above (Fig. 1Ai) and that structural sampling (Kappel et al., 2015, 2018) operates on a slower timescale than synaptic release sampling (Aitchison and Latham, 2015).

Other approaches have built on the framework of Bayesian inference to introduce gradient descent methods to optimise the full distribution over the weights (Blundell et al., 2015). Similarly, Bellec et al. (2018) introduced a network that optimally rewires as needed in a supervised learning setting closely following the synaptic sampling framework discussed above. Generative models are another class of probabilistic models implicitly related to Bayesian inference. Hierarchical variants of such models can learn progressively higher level features and uncertainty representations (Goodfellow et al., 2016), consistent with experimental observations in sensory neuroscience (Fiser et al., 2010; Haefner et al., 2016; Yamins and DiCarlo, 2016). Recently, Neftci et al. (2016) introduced a generative model with stochastic synapses that together with a local synaptic learning rule can be used for image-recognition tasks.

Despite the appealing properties and growing interest on Bayesian inference for uncertainty representation at the synaptic level (Aitchison and Latham, 2015; Kappel et al., 2015, 2018), it has remained unclear whether synapses optimise are modified so as to encode some form of uncertainty. However, recent work by Costa et al. (2017c) provided some of the first evidence suggesting that synapses optimise their response statistics (see above). If mapped onto task-relevant quantities (e.g. probability of predator given auditory stimuli) this line of research may provide the first synaptic basis for uncertainty encoding in the brain. Additionally, there are open issues with the Bayesian hypothesis: first, in sampling-based frameworks it implies the need for a large number of samples to accurately estimate encoded uncertainty, second, full Bayesian inference requires computing a normalisation factor, which is computationally costly (although this can be often relaxed). Lastly, and perhaps more importantly, it is unclear whether alternative views, such as more standard predictive views of sensory coding are not sufficient; but, these views can be understood as special

cases of each other (Aitchison and Lengyel, 2017).

4.3 Escaping local optima in deep neural networks

One recurring aspect of statistical learning is that noise injection may improve the search over the solution space. Simulated annealing is a well-known variation of this idea in which the level of noise added to the network starts out being relatively high[3], but is gradually decreased over learning (Kirkpatrick et al., 1983), allowing the network to escape local optima and converge to a good solution (Fig. 3C). This concept is remarkably similar to the biology of plastic probabilistic synapses, in that synapses also change their level of noise (variance) over learning (see above) (Costa et al., 2015, 2017c).

Similar principles have played an important role in the recent rise of deep learning (Goodfellow et al., 2016; Yamins and DiCarlo, 2016; Hassabis et al., 2017). One of the algorithms that has significantly improved performance of deep neural networks is Dropout (Srivastava et al., 2014). The idea is to randomly drop (i.e. momentarily remove) neurons with some probability during training (but not during testing), which acts as a regulariser on the network (i.e. reduces over-fitting) and enables uncertainty representation akin to Bayesian inference (Gal and Ghahramani, 2016). More recently, inspired on stochastic synaptic transmission this idea was applied at the level of synapses (DropConnect) (Wan et al., 2013), which randomly 'drops' connections instead of units with a predefined probability.

Dropout (or DropConnect) and its implications in machine learning can thus help us understand the functional utility of probabilistic synapses in the brain. One hypothesis is that learnable stochastic synapses could serve as a mechanism by which neural networks achieve better generalisation akin to the simulated annealing algorithm (Neftci et al., 2016; Bowers, 2017). Additionally, learning probabilistic synapses may also provide a good trade-off between exploration and exploitation during reinforcement learning (Seung, 2003; Blundell et al., 2015). Overall, understanding how to best adapt 'drop' probabilities is a open problem in both machine learning and neuroscience, but new developments in statistical learning have started shedding light on this issue (Gal et al., 2017).

5 Conclusions and future directions

Recent technical developments on the measurement of presynaptic and postsynaptic terminals both *in vitro* and *in vivo* is reaching the point at which it will soon be possible to monitor the synaptic response statistics as an animal learns with high spatial and temporal resolution (Rey et al., 2015; Tang et al., 2016). In particular, recent advances in ultrafast glutamate imaging (Helassa et al., 2018) and statistical inference methods (Costa et al., 2013; Bird et al., 2016; Ghanbari et al., 2017) will enable accurate and optical measurements of synaptic transmission statistics. However, despite such fast developments in experimental neuroscience, theoretical neuroscience, with some exceptions (e.g.

[3]Note that keeping the noise high throughout may hinder learning, by preventing the system from exploiting the solution space.

Seung (2003); Costa et al. (2015); Kappel et al. (2015); Aitchison and Latham (2015); Costa et al. (2017c)), has so far largely overlooked the role of probabilistic synapses in neural networks and synaptic plasticity.

Combined theoretical and experimental research has suggested that synapses optimise their response statistics through changes in pre- and postsynaptic components (Costa et al., 2015, 2017c). In future work it would be important to extend these theories to also capture other puzzling experimental observations such as presynaptic homeostatic plasticity (Branco et al., 2008; Li et al., 2018), plastic number of release sites (Loebel et al., 2013; Tang et al., 2016), spine and bouton turnover (Kappel et al., 2015; Jackson et al., 2017; Mongillo et al., 2017), connection-type specificity (Brémaud et al., 2007; Thomson, 2007; Brémaud et al., 2007; Blackman et al., 2013; Costa et al., 2013), dependence on postsynaptic voltage (Sjöström et al., 2001; Branco et al., 2008), variability optimisation (Hardingham et al., 2007; Costa et al., 2015), and the multiple timescales and differential expression of synaptic plasticity (Costa et al., 2017b; Roelfsema and Holtmaat, 2018).

On the functional side several properties may be attributed to probabilistic synapses, such as encoding perceptual uncertainty (Fiser et al., 2010; Aitchison and Latham, 2015), escaping local optimum (Seung, 2003; Blundell et al., 2015; Kappel et al., 2015, 2018), but also reflecting biophysical constraints (Harris et al., 2012; Costa et al., 2015, 2017c) in the addition to contributing to information processing (Zhang and Peskin, 2015; Nolte et al., 2018). It is conceivable that plastic probabilistic synapses enable not just one, but several of these computational functions.

Finally, recent exciting developments have led to deep neural networks that learn to encode uncertainty (Blundell et al., 2015; Gal et al., 2017). These developments together with the recent drive to map deep learning methods onto cortical circuits properties (Hassabis et al., 2017; Guerguiev et al., 2017; Costa et al., 2017a; Sacramento et al., 2017) will help to guide new research into the function of probabilistic synapses. However, the brain still has a remarkable ability to efficiently encode perceptual and task-specific uncertainty in complex environments that far outperforms current machine learning methods (Lake et al., 2016). Therefore, novel, unifying insights into the biology of probabilistic synapses also have the potential to inspire new learning algorithms.

Acknowledgements

We would like to thank Jean-Pascal Pfister, Walter Senn, Jesper Sjöström, Mark van Rossum, and Tim P. Vogels for useful discussions over the years. Finally, we would also like to thank Everton J. Agnes, Jakob Jordan and Jean-Pascal Pfister for useful feedback on this review.

Highlighted references

Costa et al. (2017c)**: The authors propose that synapses optimise their own statistics to become strong and reliable. This work demonstrates that synapses may indeed optimise their response statistics.

Costa et al. (2015)**: In this paper the authors introduced a synaptic plasticity model that captures both pre- and postsynaptic plasticity data. They also show that this model matches *in vivo* perceptual learning data, and proposed a synaptic theory of memory savings.

Cohn et al. (2015)**: Imaging of Kenyon cell output synapses in the intact brain of Drosophila shows reward-dependent modifications at presynaptic terminals. This work suggests that probabilistic synapses are also modified during rewarded events.

Blundell et al. (2015)**: The authors introduced *Bayes by Backprop*, in which the distribution over weights is learnt, not just their means. They show that this improves generalisation and yields a good exploration-exploitation trade off in reinforcement learning tasks.

Tang et al. (2016)**: Using localisation microscopy the authors revealed a new level of synaptic organisation in which release and postsynaptic receptors are aligned at the nano scale. This opens possibility of new level of structure for probabilistic synaptic transmission.

Kappel et al. (2015)*: The authors use a model to suggest that stochastic spine motility (Mongillo et al., 2017) reflects probabilistic inference through sampling over network configurations.

Orbán et al. (2016)*: In this work the authors show that many aspects of neuronal variability reflect perceptual uncertainty encoding.

Malagon et al. (2016)*: The authors use recent developments to estimate binomial release statistics at single glutamatergic synapses.

Bird et al. (2016)*: The authors introduce a new, more complete statistical inference method to infer both synaptic transmission and sort-term synaptic plasticity parameters.

Helassa et al. (2018)*: The authors introduce novel sensors for ultrafast imaging of glutamate release. These sensors offer the promise to measure binomial release parameters before and after long-term plasticity induction.

Jackson et al. (2017)*: Both synaptic boutons and spines exhibit high (and fast) turnover rates, which may be interpreted as a form of probabilistic synapses (Kappel et al., 2015, 2018). The authors made the interesting observations that these two components are regulated differentially in early stages of Alzheimer's disease.

References

Aitchison, L., Hennequin, G., and Lengyel, M. (2018). Sampling-based probabilistic inference emerges from learning in neural circuits with a cost on reliability. *arXiv preprint arXiv:1807.08952*.

Aitchison, L. and Latham, P. E. (2015). Synaptic sampling: A connection between PSP variability and uncertainty explains neurophysiological observations. *arXiv.org*, page 1505.04544v2.

Aitchison, L. and Lengyel, M. (2017). With or without you: predictive coding and bayesian inference in the brain. *Current opinion in neurobiology*, 46:219–227.

Andrade-Talavera, Y., Duque-Feria, P., Paulsen, O., and Rodriguez-Moreno, A. (2016). Presynaptic Spike Timing-Dependent Long-Term Depression in the Mouse Hippocampus. *Cerebral cortex (New York, N.Y. : 1991)*, 26(8):3637–3654.

Bellec, G., Kappel, D., Maass, W., and Legenstein, R. (2018). Deep Rewiring: Training very sparse deep networks. In *International Conference on Learning Representations*.

Berger, J. O. (2013). *Statistical decision theory and Bayesian analysis*. Springer Science & Business Media.

Bird, A. D., Wall, M. J., and Richardson, M. J. E. (2016). Bayesian Inference of Synaptic Quantal Parameters from Correlated Vesicle Release. *Frontiers in Computational Neuroscience*, 10:116.

Bittner, K. C., Milstein, A. D., Grienberger, C., Romani, S., and Magee, J. C. (2017). Behavioral time scale synaptic plasticity underlies CA1 place fields. *Science*, 357(6355):1033–1036.

Blackman, A. V., Abrahamsson, T., Costa, R. P., Lalanne, T., and Sjostrom, P. J. (2013). Target-cell-specific short-term plasticity in local circuits. *Frontiers in Synaptic Neuroscience*, 5:11.

Blundell, C., Cornebise, J., Kavukcuoglu, K., and Wierstra, D. (2015). Weight uncertainty in neural networks. *Proceedings of the 32nd International Conference on Machine Learning*.

Bowers, J. S. (2017). Parallel distributed processing theory in the age of deep networks. *Trends in cognitive sciences*.

Branco, T., Staras, K., Darcy, K. J., and Goda, Y. (2008). Local Dendritic Activity Sets Release Probability at Hippocampal Synapses. *Neuron*, 59(3):475–485.

Brea, J., Gaál, A. T., Urbanczik, R., and Senn, W. (2016). Prospective Coding by Spiking Neurons. *PLOS Computational Biology*, 12(6):e1005003.

Brémaud, A., West, D. C., and Thomson, A. M. (2007). Binomial parameters differ across neocortical layers and with different classes of connections in adult rat and cat neocortex. *Proc. Natl. Acad. Sci. USA*, 104(35):14134–14139.

Cohn, R., Morantte, I., and Ruta, V. (2015). Coordinated and Compartmentalized Neuromodulation Shapes Sensory Processing in Drosophila. *Cell*, 163(7):1742–1755.

Costa, R. P., Assael, Y. M., Shillingford, B., de Freitas, N., and Vogels, T. P. (2017a). Cortical microcircuits as gated-recurrent neural networks. In *Advances in Neural Information Processing Systems*, pages 271–282.

Costa, R. P., Froemke, R. C., Sjostrom, P. J., and van Rossum, M. C. W. (2015). Unified pre- and postsynaptic long-term plasticity enables reliable and flexible learning. *eLife*, 4:e09457.

Costa, R. P., Mizusaki, B. E. P., Sjostrom, P. J., and van Rossum, M. C. W. (2017b). Functional consequences of pre- and postsynaptic expression of synaptic plasticity. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 372(1715):20160153.

Costa, R. P., Padamsey, Z., D'amour, J. A., Emptage, N. J., Froemke, R. C., and Vogels, T. P. (2017c). Synaptic Transmission Optimization Predicts Expression Loci of Long-Term Plasticity. *Neuron*, 96(1):177–189.e7.

Costa, R. P., Sjostrom, P. J., and van Rossum, M. C. W. (2013). Probabilistic inference of short-term synaptic plasticity in neocortical microcircuits. *Frontiers in Computational Neuroscience*, 7:75.

Del Castillo, J. and Katz, B. (1954). Quantal components of the end-plate potential. *The Journal of physiology*, 124(3):560–573.

Fiser, J., Berkes, P., Orbán, G., and Lengyel, M. (2010). Statistically optimal perception and learning: from behavior to neural representations. *Trends in Cognitive Sciences*, 14(3):119–130.

Frémaux, N. and Gerstner, W. (2016). Neuromodulated Spike-Timing-Dependent Plasticity, and Theory of Three-Factor Learning Rules. *Frontiers in Neural Circuits*, 9(172):1178.

Froemke, R. C., Carcea, I., Barker, A. J., Yuan, K., Seybold, B. A., Martins, A. R. O., Zaika, N., Bernstein, H., Wachs, M., Levis, P. A., Polley, D. B., Merzenich, M. M., and Schreiner, C. E. (2013). Long-term modification of cortical synapses improves sensory perception. *Nature Neuroscience*, 16(1):79–88.

Gal, Y. and Ghahramani, Z. (2016). Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning*, pages 1050–1059.

Gal, Y., Hron, J., and Kendall, A. (2017). Concrete dropout. In *Advances in Neural Information Processing Systems*, pages 3584–3593.

Ghanbari, A., Malyshev, A., Volgushev, M., and Stevenson, I. H. (2017). Estimating short-term synaptic plasticity from pre- and postsynaptic spiking. *PLOS Computational Biology*, 13(9):e1005738.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press. `http://www.deeplearningbook.org`.

Guerguiev, J., Lillicrap, T. P., and Richards, B. A. (2017). Towards deep learning with segregated dendrites. *eLife*, 6:1.

Haefner, R. M., Berkes, P., and Fiser, J. (2016). Perceptual Decision-Making as Probabilistic Inference by Neural Sampling. *Neuron*, 90(3):649–660.

Hardingham, N. R., Hardingham, G. E., Fox, K. D., and Jack, J. J. B. (2007). Presynaptic Efficacy Directs Normalization of Synaptic Strength in Layer 2/3 Rat Neocortex After Paired Activity. *Journal of Neurophysiology*, 97(4):2965–2975.

Harris, J. J., Jolivet, R., and Attwell, D. (2012). Synaptic Energy Use and Supply. *Neuron*, 75(5):762–777.

Hassabis, D., Kumaran, D., Summerfield, C., and Botvinick, M. (2017). Neuroscience-Inspired Artificial Intelligence. *Neuron*, 95(2):245–258.

Hebert-Chatelain, E., Desprez, T., Serrat, R., Bellocchio, L., Soria-Gómez, E., Busquets-Garcia, A., Pagano Zottola, A. C., Delamarre, A., Cannich, A., Vincent, P., Varilh, M., Robin, L. M., Terral, G., García-Fernández, M. D., Colavita, M., Mazier, W., Drago, F., Puente, N., Reguero, L., Elezgarai, I., Dupuy, J.-W., Cota, D., Lopez-Rodriguez, M.-L., Barreda-Gómez, G., Massa, F., Grandes, P., Bénard, G., and Marsicano, G. (2016). A cannabinoid link between mitochondria and memory. *Nature*, 539(7630):555–559.

Helassa, N., Dürst, C. D., Coates, C., Kerruth, S., Arif, U., Schulze, C., Wiegert, J. S., Geeves, M., Oertner, T. G., and Torok, K. (2018). Ultrafast glutamate sensors resolve high-frequency release at Schaffer collateral synapses. *Proceedings of the National Academy of Sciences of the United States of America*.

Jackson, J. S., Witton, J., Johnson, J. D., Ahmed, Z., Ward, M., Randall, A. D., Hutton, M. L., Isaac, J. T., O'Neill, M. J., and Ashby, M. C. (2017). Altered Synapse Stability in the Early Stages of Tauopathy. *Cell reports*, 18(13):3063–3068.

Kappel, D., Habenschuss, S., Legenstein, R., and Maass, W. (2015). Network Plasticity as Bayesian Inference. *PLOS Computational Biology*, 11(11):e1004485.

Kappel, D., Legenstein, R., Habenschuss, S., Hsieh, M., and Maass, W. (2018). A Dynamic Connectome Supports the Emergence of Stable Computational Function of Neural Circuits through Reward-Based Learning. *eneuro*, pages ENEURO.0301–17.2018.

Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). Optimization by simmulated annealing. *Science*, 220(4598):671–680.

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., and Gershman, S. J. (2016). Building machines that learn and think like people. *Behavioral and Brain Sciences*, 40:195.

Li, X., Goel, P., Chen, C., Angajala, V., Chen, X., and Dickman, D. K. (2018). Synapse-specific and compartmentalized expression of presynaptic homeostatic potentiation. *eLife*, 7:e34338.

Loebel, A., Le Bé, J.-V., Richardson, M. J. E., Markram, H., and Herz, A. V. M. (2013). Matched pre- and post-synaptic changes underlie synaptic plasticity over long time scales. *The Journal of neuroscience*, 33(15):6257–6266.

Ma, W. J. and Jazayeri, M. (2014). Neural coding of uncertainty and probability. *Annual review of neuroscience*, 37:205–220.

Malagon, G., Miki, T., Llano, I., Neher, E., and Marty, A. (2016). Counting Vesicular Release Events Reveals Binomial Release Statistics at Single Glutamatergic Synapses. *The Journal of neuroscience*, 36(14):4010–4025.

Monday, H. R. and Castillo, P. E. (2017). Closing the gap: long-term presynaptic plasticity in brain function and disease. *Current Opinion in Neurobiology*, 45:106–112.

Mongillo, G., Rumpel, S., and Loewenstein, Y. (2017). Intrinsic volatility of synaptic connections - a challenge to the synaptic trace theory of memory. *Current Opinion in Neurobiology*, 46:7–13.

Nabavi, S., Fox, R., Proulx, C. D., Lin, J. Y., and Tsien, R. Y. (2014). Engineering a memory with LTD and LTP. *Nature*.

Neftci, E. O., Pedroni, B. U., Joshi, S., Al-Shedivat, M., and Cauwenberghs, G. (2016). Stochastic Synapses Enable Efficient Brain-Inspired Learning Machines. *Frontiers in neuroscience*, 10(99):796.

Nolte, M., Reimann, M. W., King, J. G., Markram, H., and Muller, E. B. (2018). Cortical Reliability Amid Noise and Chaos. *bioRxiv*, page 304121.

Orbán, G., Berkes, P., Fiser, J., and Lengyel, M. (2016). Neural Variability and Sampling-Based Probabilistic Representations in the Visual Cortex. *Neuron*, 92(2):530–543.

Padamsey, Z. and Emptage, N. (2014). Two sides to long-term potentiation: a view towards reconciliation. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 369(1633):20130154–20130154.

Pawlak, V., Wickens, J. R., Kirkwood, A., and Kerr, J. N. D. (2010). Timing is not everything: neuromodulation opens the STDP gate. *Frontiers in Synaptic Neuroscience*.

Pereira, U. and Brunel, N. (2018). Attractor Dynamics in Networks with Learning Rules Inferred from In Vivo Data. *Neuron*, 0(0).

Rey, S. A., Smith, C. A., Fowler, M. W., Crawford, F., Burden, J. J., and Staras, K. (2015). Ultrastructural and functional fate of recycled vesicles in hippocampal synapses. *Nature Communications*, 6(1):8043.

Reyes, A. and Sakmann, B. (1999). Developmental switch in the short-term modification of unitary EPSPs evoked in layer 2/3 and layer 5 pyramidal neurons of rat neocortex. *THe Journal of Neuroscience*, 19(10):3827–3835.

Roelfsema, P. R. and Holtmaat, A. (2018). Control of synaptic plasticity in deep cortical networks. *Nature Reviews Neuroscience*, 19(3):166–180.

Sacramento, J., Costa, R. P., Bengio, Y., and Senn, W. (2017). Dendritic error backpropagation in deep cortical microcircuits. *arXiv preprint arXiv:1801.00062*.

Seung, H. S. (2003). Learning in Spiking Neural Networks by Reinforcement of Stochastic Synaptic Transmission. *Neuron*, 40(6):1063–1073.

Silver, R. A. (2003). High-Probability Uniquantal Transmission at Excitatory Synapses in Barrel Cortex. *Science*, 302(5652):1981–1984.

Sjöström, P. J., Turrigiano, G. G., and Nelson, S. B. (2001). Rate, Timing, and Cooperativity Jointly Determine Cortical Synaptic Plasticity. *Neuron*, 32(6):1149–1164.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1):1929–1958.

Stauffer, W. R., Lak, A., Yang, A., Borel, M., Paulsen, O., Boyden, E. S., and Schultz, W. (2016). Dopamine Neuron-Specific Optogenetic Stimulation in Rhesus Macaques. *Cell*, 166(6):1564–1571.e6.

Tang, A.-H., Chen, H., Li, T. P., Metzbower, S. R., MacGillavry, H. D., and Blanpied, T. A. (2016). A trans-synaptic nanocolumn aligns neurotransmitter release to receptors. *Nature*, 536(7615):210–214.

Testa-Silva, G., Verhoog, M. B., Linaro, D., de Kock, C. P. J., Baayen, J. C., Meredith, R. M., De Zeeuw, C. I., Giugliano, M., and Mansvelder, H. D. (2014). High bandwidth synaptic communication and frequency tracking in human neocortex. *PLoS Biology*, 12(11):e1002007.

Thomson, A. M. (2007). Functional maps of neocortical local circuitry. *Frontiers in neuroscience*, 1(1):19–42.

Wan, L., Zeiler, M., Zhang, S., Le Cun, Y., and Fergus, R. (2013). Regularization of neural networks using dropconnect. In *International Conference on Machine Learning*, pages 1058–1066.

Wang, Y., Markram, H., Goodman, P. H., Berger, T. K., Ma, J., and Goldman-Rakic, P. S. (2006). Heterogeneity in the pyramidal network of the medial prefrontal cortex. *Nature Publishing Group*, 9(4):534–542.

Yamins, D. L. K. and DiCarlo, J. J. (2016). Using goal-driven deep learning models to understand sensory cortex. *Nature Neuroscience*, 19(3):356–365.

Zhang, C. and Peskin, C. S. (2015). Improved signaling as a result of randomness in synaptic vesicle release. *Proceedings of the National Academy of Sciences of the United States of America*, 112(48):14954–14959.