Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Current Opinion in
Neurobiology

To spike, or when to spike?

Robert Gütig

Abstract

Recent experimental reports have suggested that cortical networks can operate in regimes were sensory information is encoded by relatively small populations of spikes and their precise relative timing. Combined with the discovery of spike timing dependent plasticity, these findings have sparked growing interest in the capabilities of neurons to encode and decode spike timing based neural representations. To address these questions, a novel family of methodologically diverse supervised learning algorithms for spiking neuron models has been developed. These models have demonstrated the high capacity of simple neural architectures to operate also beyond the regime of the well established independent rate codes and to utilize theoretical advantages of spike timing as an additional coding dimension.

Addresses
Max Planck Institute of Experimental Medicine, Hermann-Rein-Str. 3, 37075 Göttingen, Germany

Corresponding author: Gütig, Robert (guetig@em.mpg.de)

Current Opinion in Neurobiology 2014, **25**:134–139

This review comes from a themed issue on **Theoretical and computational neuroscience**

Edited by **Adrienne Fairhall** and **Haim Sompolinsky**

For a complete overview see the Issue and the Editorial

Available online 25th January 2014

http://dx.doi.org/10.1016/j.conb.2014.01.004

Why think about spikes

Spike timing dependent plasticity

By controlling the timing of presynaptic and postsynaptic action potentials in pairs of connected neurons, synaptic physiologists have uncovered an entire family of spike timing dependent synaptic learning rules [1]. In its canonical form spike timing dependent plasticity (STDP) is temporally asymmetric: Causal pairing of presynaptic and postsynaptic spikes induces long-term potentiation, but the reversed, acausal, ordering leads to long-term depression [2,3]. In spite of open questions and concerns [1], the majority of researchers have accepted the notion that the molecular machinery underlying the induction of long-term changes in the efficacies of central nervous synapses can be sensitive also to the relative timing of presynaptic and postsynaptic action potentials down to millisecond precision.

Sparse, spike timing based neural representations

Neural codes are often characterized along two continuous dimensions: The temporal precision with which spikes encode information and the degree to which information is encoded by correlations between spikes fired by the same or different neurons [4]. Most of systems neuroscience, experimental and theoretical, has taken place in the far corner where information is carried by spikes that are independent across neurons and whose temporal precision is coarse. Such independent neuron rate codes have been very successful in describing a large body of neural recordings and have served as the basis of many powerful theoretical models. However, a growing number of experimental reports have described deviations from this idealization. Specifically, simultaneous recordings from populations of sensory neurons have suggested that spikes elicited in response to sensory stimuli can be timed with millisecond precision not only relative to a stimulus event but also relative to other action potentials of the same or other neurons [5–7]. In addition, recently established optogenetic and patch clamp stimulation experiments in vivo have indicated that a small number of spikes within a sparse subset of cortical neurons can be sufficient to drive learning and behavior [8–10].

Encoding and decoding of spike patterns

While their implications for neural coding have remained controversial, the above findings have highlighted two basic questions about neural information processing that were not covered within the traditional focus of firing rate centered theoretical neuroscience.

1. To what extent and how can neurons learn to encode information in precise sequences of spikes?
2. Can neurons learn, and if so how, to decode such spike-timing based representations?

This review covers the last decade of progress that computational neuroscientists have made in answering these two questions for simple, yet biologically plausible neuron models of the spike-response and leaky-integrate-and-fire types [11]. Because of their limited analytical tractability, researchers have translated the above questions into supervised information processing tasks. By endowing neurons with supervised synaptic learning rules, they studied their capabilities to learn given mappings between input spike patterns and desired neural responses. The above questions translate to two conceptually different learning problems: *Encoding tasks*, defined here as scenarios in which the desired neural response specifies a sequence of output spike times, and *decoding tasks* where the desired response specifies the number of output spikes without confining their timing.

Encoding task: learning to generate a desired spike train

A crucial determinant of the computational structure that underlies the learning problem of the encoding task is the fact that the desired output spike times are fixed and given. As a result, the temporal component of the synaptic credit assignment problem is resolved by the voltage kinetics of the considered neuron model which determines the temporal window within which input spikes influence the postsynaptic potential and how strongly they do so. Given the times of desired and erroneous spikes, the remaining credit assignment problem is spatial: The learning has to solve how to scale each input spike's contributions to the postsynaptic membrane potential at the times of desired output spikes such that it reaches the cell's firing threshold but remains below during the desired interspike intervals.

Irrespective of their diverse methodological origins, most proposed learning rules lock the input time dependence of synaptic updates to the times of actual and desired output spikes: Changes in synaptic efficacies receive potentiating contributions from input spikes that arrive within an appropriate time window before the time of a desired output spike and depressing components from input spikes that precede erroneous spikes.

Learning rules

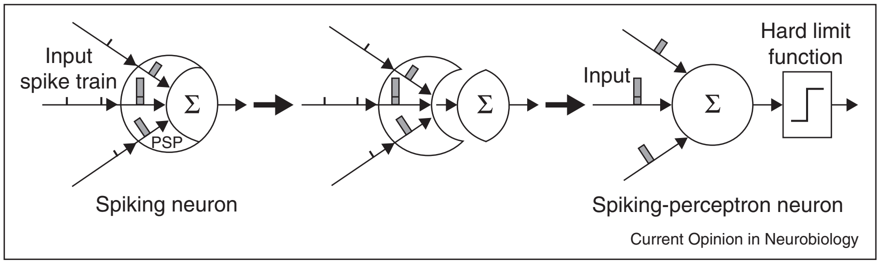
Application of the standard gradient descent formalism to the encoding problem is complicated by the strong nonlinearity of the neural spike generation mechanism: Even small changes in synaptic efficacies that remove or add spikes cause large changes in the output spike train. Addressing the resulting difficulties by an additional set of approximations and learning heuristics, such as the linearization of the membrane potential around output spike times, the clipping of diverging gradients, and appropriate initialization of the synaptic weights, a gradient descent algorithm for spiking neural networks has been developed and used throughout a series of extensions [12–15] of the original SpikeProp algorithm [16]. These algorithms are akin to traditional error-backpropagation learning and based on the Euclidean distance between generated and desired output spike times. Complications associated with unequal numbers of actual and desired spikes were alleviated in [17] by adapting a spike train metric from Victor and Purpura [18] which incorporates missing and extra spikes.

The problems arising from the discontinuous error gradient were circumvented in an alternative gradient based approach that employed a stochastic spike threshold and derived a learning rule by maximizing the smooth likelihood of the neuron to fire at the desired output times [19,20•]. This approach has recently been applied to multi-layer networks [21].

An elegant alternative to the evaluation of an explicit gradient is to exploit the spatial structure of the learning problem and to map it to the well established perceptron framework [22,23••,24] (Figure 1). In the resulting perceptron classification task [25], vectors of the unscaled postsynaptic contributions of each synapse at the times of desired output spikes constitute the positive class while the negative class is sampled from the interspike intervals [23••]. A key strength of this method derives from the perceptron convergence theorem which resolves the confound between the capabilities of the neural architecture and those of the learning rule.

Driven by considerations of biological plausibility several types of learning heuristics were based on empirically observed forms of STDP. Legenstein *et al.* [26] studied the convergence of temporally asymmetric STDP learning dynamics when forced through fixed output spikes. The Remote Supervision Method (ReSuMe) [27•] uses two forms of STDP in parallel: one to strengthen inputs that precede desired output spikes and the other to weaken inputs that precede actual output spikes. When balanced, the two processes result in the convergence of the actual output spike train towards the desired one (Figure 2). The ReSuMe rule has been applied to multilayer networks [28] and has been rederived on the basis of synaptic currents rather than potentials [17].

Figure 1



Current Opinion in Neurobiology

Transformation from a spiking neuron model to a perceptron. For a fixed point in time, the postsynaptic potentials of each synapse can be interpreted as continuous valued inputs to a perceptron.From [23••].