1 Microscopic Foundation: Neural Dynamics

Before deriving the macroscopic plasticity rules, we must first establish the microscopic dynamics of the individual neurons comprising the network. We employ the Leaky Integrate-and-Fire (LIF) model, a standard reduction of the Hodgkin-Huxley formalism that captures the essential sub-threshold integration and thresholding behavior of cortical neurons [1].

1.1 Membrane Potential Dynamics

The state of a postsynaptic neuron i is described by its membrane potential $V_i(t)$. In the absence of input, the membrane potential relaxes to a resting potential E_L. The evolution of $V_i(t)$ is governed by the conservation of current across the cell membrane, modeled as an RC circuit consisting of a leakage resistor R_m and a membrane capacitor C_m in parallel [2]:

$$\tau_m \frac{dV_i(t)}{dt} = -(V_i(t) - E_L) + R_m I_{syn,i}(t) + R_m I_{ext,i}(t), \tag{1}$$

where $\tau_m = R_m C_m$ represents the membrane time constant. In cortical pyramidal neurons, τ_m typically lies in the range of 10–20 ms [2]. The term $I_{syn,i}(t)$ represents the total synaptic current received from presynaptic neurons, and $I_{ext,i}(t)$ accounts for any external background currents or noise.

1.2 Synaptic Interaction

The synaptic current $I_{syn,i}(t)$ is determined by the activity of the presynaptic population. Let the spike train of a presynaptic neuron j be denoted by $\rho_j(t) = \sum_k \delta(t - t_j^k)$. The specific mathematical definition of these spike times is provided in Section 1.3.

The arrival of a spike from neuron j induces a transient change in the input to neuron i. We utilize the *current-based* approximation, which assumes that synaptic currents are independent of the postsynaptic membrane potential (unlike conductance-based models where $I \propto g(t)(V - E_{syn})$). This approximation is standard for analyzing network-level learning dynamics [1]. The total synaptic current is the linear sum of filtered presynaptic spikes, weighted by the synaptic efficacy w_{ij} (the dynamics of which are derived in Section 2.4):

$$I_{syn,i}(t) = \sum_j w_{ij} \int_{-\infty}^{t} \alpha(t - s)\rho_j(s)\, ds, \tag{2}$$

where $\alpha(t)$ is the postsynaptic current (PSC) kernel. Following Dayan and Abbott [2], this is modeled as an exponential decay function $\alpha(t) = \frac{1}{\tau_s} e^{-t/\tau_s}\Theta(t)$, where $\Theta(t)$ is the Heaviside step function. The synaptic time constant τ_s typically ranges from 2–10 ms for fast AMPA/GABA receptors [3]-.

0

1.3 Spike Generation Mechanism

The continuous voltage dynamics defined above give rise to discrete events. Mathematically, a spike is not a duration but a single point in time t_i^k. This is defined as the moment the membrane

potential crosses a fixed threshold ϑ from below:

$$t_i^k : V_i(t_i^k) = \vartheta \quad \text{and} \quad \left.\frac{dV_i}{dt}\right|_{t=t_i^k} > 0. \tag{3}$$

While a simple condition $V > \vartheta$ is often used in discrete-time simulations, the derivative condition $\frac{dV}{dt} > 0$ is required analytically to distinguish the leading edge of the spike (the event) from the subsequent duration of the pulse or the repolarization phase.

Following the formalism in Dayan and Abbott [2], the absolute refractory period is modeled by interrupting the integration of (1). When V_i reaches threshold at t_i^k, the potential is reset to $V_{reset} < \vartheta$. The dynamics in (1) are then *suspended* for the duration τ_{ref} (typically 2–5 ms). Integration resumes at $t = t_i^k + \tau_{ref}$ with the initial condition $V_i(t_i^k + \tau_{ref}) = V_{reset}$. This mechanism defines the postsynaptic spike train $\rho_i(t) = \sum_k \delta(t - t_i^k)$, which serves as the input to the plasticity equations in the following section.

2 Mathematical Formulation of the Three-Factor Plasticity Model

Having defined the generation of spike times via the LIF dynamics, we now analyze the evolution of the synaptic weights w_{ij}. We focus on a plasticity rule belonging to the class of *three-factor learning rules*, as reviewed by Frémaux and Gerstner [4].

In standard STDP, weight changes depend on two local factors:

- **Factor 1:** Presynaptic spike timing (t_j^k).
- **Factor 2:** Postsynaptic spike timing (t_i^k).

In the three-factor framework, these local correlations are gated by a global third signal, **Factor 3**, representing neuromodulation or reward.

2.1 Neural Activity and Notation

Following the framework established in the previous section, the neural response functions are formally treated as sums of Dirac distributions:

$$\rho_j(t) = \sum_{k=1}^{N_j} \delta(t - t_j^k) \quad \text{and} \quad \rho_i(t) = \sum_{k=1}^{N_i} \delta(t - t_i^k). \tag{4}$$

Here, N_j and N_i denote the total number of spikes fired by each neuron over the simulation interval $t \in [0, T]$, where T represents the total duration of the learning epoch. These counts are determined by the dynamics in (1) and (3). Note that while the equations are deterministic, the spike trains often exhibit Poisson-like variability due to stochastic external currents $I_{ext}(t)$ [2].

2.2 Local Dynamics: The Eligibility Trace

A central feature of this model is that the coincidence of spikes creates a temporary memory trace, $E_{ij}(t)$, known as the *eligibility trace* [4]. This trace allows the synapse to bridge the temporal delay between millisecond-scale neural activity and delayed reward signals. The eligibility

trace evolves according to:

$$\tau_e \frac{dE_{ij}(t)}{dt} = -E_{ij}(t) + S_{ij}(t), \tag{5}$$

where τ_e is the decay time constant. For reinforcement learning tasks, τ_e is typically on the order of hundreds of milliseconds to seconds (e.g., 0.1–1.0 s), significantly longer than the membrane time constant τ_m [1].

The driving term $S_{ij}(t)$ represents the instantaneous induction of Spike-Timing-Dependent Plasticity (STDP). To define $S_{ij}(t)$, we use variables $x_j(t)$ and $y_i(t)$ that track the recent history of presynaptic and postsynaptic activity:

$$\tau_+ \frac{dx_j(t)}{dt} = -x_j(t) + \rho_j(t), \tag{6}$$

$$\tau_- \frac{dy_i(t)}{dt} = -y_i(t) + \rho_i(t), \tag{7}$$

where τ_+ and τ_- are the time constants for the potentiation and depression windows. Experimental measurements suggest these values are typically in the range of 20–40 ms [5].

The STDP induction term $S_{ij}(t)$ combines Long-Term Potentiation (LTP) and Long-Term Depression (LTD):

$$S_{ij}(t) = \underbrace{A_+(w_{ij})x_j(t)\rho_i(t)}_{\text{LTP contribution}} - \underbrace{A_-(w_{ij})y_i(t)\rho_j(t)}_{\text{LTD contribution}}. \tag{8}$$

2.3 Global Dynamics: Neuromodulated Update

The actual change in synaptic weight is governed by a global neuromodulatory signal $M(t)$:

$$\frac{dw_{ij}(t)}{dt} = M(t)E_{ij}(t). \tag{9}$$

The signal $M(t)$ is modeled as a Reward Prediction Error (RPE), calculated as the difference between the instantaneous reward $R(t)$ and a baseline expectation $\bar{R}(t)$:

$$M(t) = R(t) - \bar{R}(t). \tag{10}$$

Here, $\bar{R}(t)$ serves as a reference point, allowing for bidirectional regulation of synaptic weights [4].

2.4 Weight Constraints and Stability

The synaptic weight w_{ij} represents the efficacy of the connection from neuron j to neuron i. To prevent unbounded growth during the learning process, we constrain the weight to a closed interval $w_{ij} \in [0, w_{\max}]$. The parameter $w_{\max}$ represents the physiological saturation limit of synaptic efficacy.

We enforce these bounds using "soft bound" dependencies in the scaling functions:

$$A_+(w_{ij}) = \eta_+(w_{\max} - w_{ij}) \quad \text{and} \quad A_-(w_{ij}) = \eta_- w_{ij}, \tag{11}$$

where η_+, η_- are the learning rates. This linear dependence ensures that the rate of weight change drops to zero as the weight approaches either limit (0 or $w_{\max}$), naturally stabilizing the dynamics [1].

References

[1] Wulfram Gerstner et al. *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge: Cambridge University Press, 2014.

[2] Peter Dayan and L. F. Abbott. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Cambridge, MA: MIT Press, 2001. ISBN: 0-262-04199-5.

[3] Alain Destexhe, Zachary F Mainen, and Terrence J Sejnowski. "Kinetic Models of Synaptic Transmission". In: *Methods in Neuronal Modeling*. Ed. by Christof Koch and Idan Segev. 2nd ed. Cambridge, MA: MIT Press, 1998, pp. 1–25.

[4] Nicolas Frémaux and Wulfram Gerstner. "Neuromodulated Spike-Timing-Dependent Plasticity, and Theory of Three-Factor Learning Rules". In: *Frontiers in Neural Circuits* 9 (2015), p. 85. DOI: 10.3389/fncir.2015.00085.

[5] Guo-qiang Bi and Mu-ming Poo. "Synaptic Modifications in Cultured Hippocampal Neurons: Dependence on Spike Timing, Synaptic Strength, and Postsynaptic Cell Type". In: *The Journal of Neuroscience* 18.24 (1998), pp. 10464–10472.