

# Philipps-Universität Marburg

Johannes Gille

Fachbereich 17

AG Allgemeine und Biologische Psychologie

AE Theoretische Kognitionswissenschaft

Learning in cortical microcircuits with multi-compartment pyramidal neurons

## **Supervisors:**

Prof. Dr. Dominik Endres, Philipps-Universität Marburg

Dr. Johan Kwisthout, Radboud University

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Motivation . . . . .	3
1.2	The Backpropagation of errors algorithm . . . . .	4
1.3	Concerns over biological plausibility . . . . .	4
1.3.1	Local error representation . . . . .	4
1.3.2	The weight transport problem . . . . .	5
1.3.3	Neuron models . . . . .	5
1.4	Overcoming biological implausibility . . . . .	6
1.4.1	Dendrites as computational elements . . . . .	7
1.5	Cortical microcircuits . . . . .	8
1.6	Model selection . . . . .	8
1.6.1	Predictive coding . . . . .	9
1.6.2	The Dendritic error model . . . . .	10
<b>2</b>	<b>Methods</b>	<b>12</b>
2.1	The dendritic error model . . . . .	12
2.1.1	Network architecture . . . . .	12
2.1.2	Neuron models . . . . .	14
2.2	Urbanczik-Senn Plasticity . . . . .	16
2.2.1	Derivation . . . . .	16
2.3	The self-predicting network state . . . . .	17
2.4	Training the network . . . . .	19
2.5	The NEST simulator . . . . .	20
2.6	Transitioning to spiking communication . . . . .	21
2.7	Event-based Urbanczik-Senn plasticity . . . . .	22
2.7.1	Integrating weight changes . . . . .	23
	this notation seems slightly abusive, as neither side considers $t$ , but it is taken precisely from Stapmanns et al. (2021) . . . . .	24
	That proof is currently part of the appendix, as the original proof (Stapmanns et al., 2021) has errors. Should I keep it or refer to the original paper? .	25

2.8	Latent Equilibrium . . . . .	25
2.9	Implementational details . . . . .	28
2.9.1	Neuron model Adaptations . . . . .	30
<b>3</b>	<b>Results</b>	<b>32</b>
3.1	The self-predicting state . . . . .	32
3.2	Apical compartment capacitance . . . . .	32
3.3	Presentation times with latent equilibrium . . . . .	32
3.4	Separation of synaptic polarity . . . . .	33
3.5	In search of plausible spike frequencies . . . . .	36
3.6	Resilience to imperfect connectivity . . . . .	36
3.7	direct feedback connections to interneurons . . . . .	36
3.8	Performance of the different implementations . . . . .	36
	Is there a more conventional notation? . . . . .	39
3.9	Response to unpredicted stimuli . . . . .	39
<b>4</b>	<b>Discussion</b>	<b>41</b>
4.1	Contribution . . . . .	41
4.1.1	criteria for biological plausibility . . . . .	41
4.2	Limitations . . . . .	41
4.3	Future directions . . . . .	42
4.3.1	Additional Backprop concerns . . . . .	42
4.4	Should it be considered pre-training? . . . . .	42
4.5	Relation to energy minimization . . . . .	42
4.6	Correspondence of the final model to cortical circuitry . . . . .	42
4.6.1	improvements to the neuron models . . . . .	43
4.7	Conclusion . . . . .	43
<b>5</b>	<b>Appendix</b>	<b>44</b>
5.1	Somato-dendritic coupling . . . . .	44
5.2	Integration of the spike-based Urbanczik-Senn plasticity . . . . .	44
5.3	Dendritic leakage conductance . . . . .	48
5.4	Plasticity in feedback connections . . . . .	49
5.5	Presentation times and Latent Equilibrium . . . . .	49
<b>6</b>	<b>additional notes and questions</b>	<b>52</b>
6.1	Questions . . . . .	52
6.2	TODOS . . . . .	52
6.2.1	Interneurons and their jobs . . . . .	53
6.2.2	tpres . . . . .	53

changes. As the history of dendritic error applies equally to all incoming synapses, it only needs to be recorded once at the neuron. Alongside each entry in the history, a counter is stored and incremented whenever a synapse has read the history at that time step. Once all synapses have read out an entry, it is deleted. Thus, the history dynamically grows and shrinks during simulation and is only ever as long as the largest inter-spike interval (ISI) of all presynaptic neurons. This approach proves to be more efficient in terms of computation time, since fewer calls to the update function are required per synapse. It does come at the cost of memory consumption, as the history can grow particularly large for simulations with low in-degrees or large ISI<sup>5</sup>. During testing, the Event-based scheme proved substantially more efficient for many network types. This did however introduce the challenge of retroactively computing weight changes from the time of the last spike upon arrival of a new spike.

### 2.7.1 Integrating weight changes

Stapmanns et al. describe the Urbanczik-Senn plasticity rule based on the general equation for weight changes, while omitting obsolete parameters:

$$\dot{w}_{ij}(t) = F(s_j^*(t), V_i^*(t)) \quad (2.21)$$

where the change in weight  $\dot{w}_{ij}$  of a synapse from neuron  $j$  to neuron  $i$  at time  $t$  is given by a function  $F$  that depends on the postsynaptic membrane potential  $V_i^*$  and the presynaptic spiketrain  $s_j^*$ . The  $*$  operator denotes a causal function, indicating that a value  $V_i^*(t)$  potentially depends on all previous values of  $V_i(t' < t)$ . One can formally integrate Equation 2.21 in order to obtain the weight change between two arbitrary time points  $t$  and  $T$ :

$$\Delta w_{ij}(t, T) = \int_t^T dt' F[s_j^*, V_i^*](t') \quad (2.22)$$

This integral forms the basis of computing the change in weight between two arriving spikes. Thus, at the implementational level,  $t$  is usually the time of the last spike that traversed the synapse, and  $T$  is the current `biological_time`<sup>6</sup>. For spiking neurons, it is necessary to

---

<sup>5</sup>It should also be noted that in this approach requires redundant integration of the history by every synapse. Stapmanns et al. propose a third solution, in which this integration is performed once whenever a spike is transmitted through any incoming connection, with the resulting weight change being applied to all synapses immediately. This approach proved to be even more efficient for some network configurations, but is incompatible with simulations where incoming synapses have heterogeneous synaptic delays due to the way that these delays are processed by the NEST simulator. See Section 3.1.3 in Stapmanns et al. (2021) for a detailed explanation.

<sup>6</sup>This term is adopted from the NEST convention, where it describes the time in *ms* which the simulator has computed. In other words, it is the number of simulation steps times  $\Delta t$ , not to be confused with a simulation's hardware-dependent runtime (sometimes also called *wall clock time* (Van Albada et al., 2018)).

approximate the presynaptic rate ( $r_j = \phi(u_j)$ ). For this, a well established solution is to transform the spiketrain  $s_j$  into a decaying trace using an exponential filter kernel  $\kappa$ :

$$\kappa(t) = H(t) \frac{1}{t} e^{\frac{-t}{\tau_\kappa}} \quad (2.23)$$

$$H(t) = \begin{cases} 1 & \text{if } t > 0 \\ 0 & \text{if } t \leq 0 \end{cases} \quad (2.24)$$

$$(f * g)(t) = \int_{-\infty}^{\infty} f(t') g(t - t') dt' \quad (2.25)$$

$$s_j^* = \kappa_s * s_j. \quad (2.26)$$

with filter time constant  $\tau_\kappa$ . The trace is computed by convolving (Equation 2.25) the spiketrain with the exponential filter kernel  $\kappa$ . The filter uses the Heaviside step function  $H(t)$ , and is therefore only supported on positive values of  $t$  (also called a one-sided exponential decay kernel). This property is important, as integration limits of the convolution can be truncated when  $f$  and  $g$  are both only supported on  $[0, \infty)$ :

$$(f * g)(t) = \int_0^t f(t') g(t - t') dt' \quad (2.27)$$

Since spikes naturally only occur for  $t > 0$ , this simplified integral allows for a much more efficient computation of the convolution. The Function  $F$  on the right-hand side of Equation 2.21 can therefore be rewritten as:

$$F[s_j^*, V_i^*] = \eta \kappa * (V_i^* s_j^*) \quad (2.28)$$

$$V_i^* = (\phi(u_i^{som}) - \phi(\hat{v}_i^{dend})) \quad (2.29)$$

**this notation seems slightly abusive, as neither side considers  $t$ , but it is taken precisely from Stapmanns et al. (2021)** with learning rate  $\eta$ .  $V_i^*$  then is the dendritic error of the dendrite that the synapse between  $j$  and  $i$  is located at<sup>7</sup>. Writing out the convolutions in Equation 2.22 explicitly, we obtain

---

<sup>7</sup>The dendritic error here is defined as the difference between two hypothetical rates based on the arbitrary function  $\phi$ . The original implementation uses the difference between the actual postsynaptic spiketrain and this dendritic prediction ( $V_i^* = (s_i - \phi(\hat{v}_i^{dend}))$ ). Furthermore, Stapmanns et al. show that generating a spiketrain from the dendritic potential ( $V_i^* = (s_i - s_i^{dend})$ ) also results in successful learning, although at the cost of additional training time. The rate-based variant was chosen in order to not hinder learning performance any more than necessary.

$$\Delta w_{ij}(t, T) = \int_t^T dt' F[s_j^*, V_i^*](t') \quad (2.30)$$

$$= \int_t^T dt' \eta \int_0^{t'} dt'' \kappa(t' - t'') V_i^*(t'') s_j^*(t'') \quad (2.31)$$

Computing this Equation directly is inefficient due to the nested integrals. Yet, it is possible to break up the integrals into two simpler computations and rewrite the weight change as:

$$\Delta W_{ij}(t, T) = \eta \left[ I_1(t, T) - I_2(t, T) + I_2(0, t) \left( 1 - e^{-\frac{T-t}{\tau_\kappa}} \right) \right] \quad (2.32)$$

$$I_1(a, b) = \int_a^b dt V_i^*(t) s_j^*(t) \quad (2.33)$$

$$I_2(a, b) = \int_a^b dt e^{-\frac{b-t}{\tau_\kappa}} V_i^*(t) s_j^*(t) \quad (2.34)$$

$$(2.35)$$

See Section **TODO: ref** for a rigorous proof that this is in fact the desired integral. **That proof is currently part of the appendix, as the original proof (Stapmanns et al., 2021) has errors. Should I keep it or refer to the original paper?** The resulting equations allow for a rather efficient computation of weight changes compared to the complex integral described in Equation 2.31. This integration is performed whenever a spike traverses a synapse. It generalizes to all special cases in Equations 2.13-2.16, as long as the appropriate dendritic error is stored by the postsynaptic neuron.

## 2.8 Latent Equilibrium

The most significant drawback of the Sacramento model is the previously mentioned requirement for long stimulus presentation times and appropriately low learning rates. This makes the network prohibitively inefficient for the large networks required for complex learning tasks. Sacramento et al. developed a steady-state approximation of their network which models the state of the network after it has balanced out in response to a stimulus-target pair. It does not suffer from these issues and shows that their model can in principle solve more demanding learning tasks such as MNIST. Yet these types of approximation are much further detached from biological neurons than the original model and thus do not lend themselves well to an investigation of biological plausibility (Gerstner and Naud, 2009). Furthermore, the approximation is unsuitable for an investigation of spike-based communication, since the steady state of both network ideally would be the same. Thus, neither the fully modeled neuron dynamics nor the steady-state approximation are suited for complex learning tasks. A substantial

# Bibliography

- Abbott, L. F. and Nelson, S. B. (2000). Synaptic plasticity: taming the beast. *Nature neuroscience*, 3(11):1178–1183.
- Adams, R. A., Friston, K. J., and Bastos, A. M. (2015). Active inference, predictive coding and cortical architecture. *Recent Advances on the Modular Organization of the Cortex*, pages 97–121.
- Ahmad, N., van Gerven, M. A., and Ambrogioni, L. (2020). Gait-prop: A biologically plausible learning rule derived from backpropagation of error. *Advances in Neural Information Processing Systems*, 33:10913–10923.
- Barranca, V. J., Bhuiyan, A., Sundgren, M., and Xing, F. (2022). Functional implications of dale’s law in balanced neuronal network dynamics and decision making. *Frontiers in Neuroscience*, 16.
- Bartunov, S., Santoro, A., Richards, B., Marris, L., Hinton, G. E., and Lillicrap, T. (2018). Assessing the scalability of biologically-motivated deep learning algorithms and architectures. *Advances in neural information processing systems*, 31.
- Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., and Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, 76(4):695–711.
- Bellec, G., Scherr, F., Hajek, E., Salaj, D., Legenstein, R., and Maass, W. (2019). Biologically inspired alternatives to backpropagation through time for learning in recurrent neural nets. *arXiv preprint arXiv:1901.09049*.
- Bellec, G., Scherr, F., Subramoney, A., Hajek, E., Salaj, D., Legenstein, R., and Maass, W. (2020). A solution to the learning dilemma for recurrent networks of spiking neurons. *Nature communications*, 11(1):1–15.
- Bengio, Y. (2014). How auto-encoders could provide credit assignment in deep networks via target propagation. *arXiv preprint arXiv:1407.7906*.
- Bengio, Y., Lee, D.-H., Bornschein, J., Mesnard, T., and Lin, Z. (2015). Towards biologically plausible deep learning. *arXiv preprint arXiv:1502.04156*.

- Bengio, Y., Mesnard, T., Fischer, A., Zhang, S., and Wu, Y. (2017). Stdp-compatible approximation of backpropagation in an energy-based model. *Neural computation*, 29(3):555–577.
- Brea, J. and Gerstner, W. (2016). Does computational neuroscience need new synaptic learning paradigms? *Current opinion in behavioral sciences*, 11:61–66.
- Chavlis, S. and Poirazi, P. (2021). Drawing inspiration from biological dendrites to empower artificial neural networks. *Current opinion in neurobiology*, 70:1–10.
- Crick, F. (1989). The recent excitement about neural networks. *Nature*, 337(6203):129–132.
- Davies, M., Srinivasa, N., Lin, T.-H., Chinya, G., Cao, Y., Choday, S. H., Dimou, G., Joshi, P., Imam, N., Jain, S., et al. (2018). Loihi: A neuromorphic manycore processor with on-chip learning. *Ieee Micro*, 38(1):82–99.
- Doron, G., Shin, J. N., Takahashi, N., Drüke, M., Bocklisch, C., Skenderi, S., de Mont, L., Toumazou, M., Ledderose, J., Brecht, M., et al. (2020). Perirhinal input to neocortical layer 1 controls learning. *Science*, 370(6523):eaaz3136.
- Eppler, J. M., Helias, M., Muller, E., Diesmann, M., and Gewaltig, M.-O. (2009). Pynest: a convenient interface to the nest simulator. *Frontiers in neuroinformatics*, 2:12.
- Eyal, G., Verhoog, M. B., Testa-Silva, G., Deitcher, Y., Benavides-Piccione, R., DeFelipe, J., De Kock, C. P., Mansvelder, H. D., and Segev, I. (2018). Human cortical pyramidal neurons: from spines to spikes via models. *Frontiers in cellular neuroscience*, 12:181.
- Felleman, D. J. and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex (New York, NY: 1991)*, 1(1):1–47.
- Friston, K. (2008). Hierarchical models in the brain. *PLoS computational biology*, 4(11):e1000211.
- Friston, K. and Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical transactions of the Royal Society B: Biological sciences*, 364(1521):1211–1221.
- George, D. and Hawkins, J. (2009). Towards a mathematical theory of cortical micro-circuits. *PLoS computational biology*, 5(10):e1000532.
- Gerstner, W. and Naud, R. (2009). How good are neuron models? *Science*, 326(5951):379–380.
- Gewaltig, M.-O. and Diesmann, M. (2007). Nest (neural simulation tool). *Scholarpedia*, 2(4):1430.
- Gidon, A., Zolnik, T. A., Fidzinski, P., Bolduan, F., Papoutsi, A., Poirazi, P., Holtkamp, M., Vida, I., and Larkum, M. E. (2020). Dendritic action potentials and computation in human layer 2/3 cortical neurons. *Science*, 367(6473):83–87.



- Ginzburg, I. and Sompolinsky, H. (1994). Theory of correlations in stochastic neural networks. *Physical review E*, 50(4):3171.
- Grossberg, S. (1987). Competitive learning: From interactive activation to adaptive resonance. *Cognitive science*, 11(1):23–63.
- Guerguiev, J., Lillicrap, T. P., and Richards, B. A. (2017). Towards deep learning with segregated dendrites. *ELife*, 6:e22901.
- Haeusler, S. and Maass, W. (2007). A statistical analysis of information-processing properties of lamina-specific cortical microcircuit models. *Cerebral cortex*, 17(1):149–162.
- Haider, P., Ellenberger, B., Kriener, L., Jordan, J., Senn, W., and Petrovici, M. A. (2021). Latent equilibrium: A unified learning theory for arbitrarily fast computation with arbitrarily slow neurons. *Advances in Neural Information Processing Systems*, 34:17839–17851.
- Häusser, M. and Mel, B. (2003). Dendrites: bug or feature? *Current opinion in neurobiology*, 13(3):372–383.
- Hertäg, L. and Clopath, C. (2022). Prediction-error neurons in circuits with multiple neuron types: Formation, refinement, and functional implications. *Proceedings of the National Academy of Sciences*, 119(13):e2115699119.
- Hines, M. L. and Carnevale, N. T. (1997). The neuron simulation environment. *Neural computation*, 9(6):1179–1209.
- Ishizuka, N., Cowan, W. M., and Amaral, D. G. (1995). A quantitative analysis of the dendritic organization of pyramidal cells in the rat hippocampus. *Journal of Comparative Neurology*, 362(1):17–45.
- Izhikevich, E. M. (2007). Solving the distal reward problem through linkage of stdp and dopamine signaling. *Cerebral cortex*, 17(10):2443–2452.
- Kandel, E., Koester, J., Mack, S., and Siegelbaum, S. (2021). *Principles of Neural Science, Sixth Edition*. McGraw-Hill Education.
- Kawaguchi, Y. (2001). Distinct firing patterns of neuronal subtypes in cortical synchronized activities. *Journal of Neuroscience*, 21(18):7261–7272.
- Khaligh-Razavi, S.-M. and Kriegeskorte, N. (2014). Deep supervised, but not unsupervised, models may explain it cortical representation. *PLoS computational biology*, 10(11):e1003915.
- Kubilius, J., Bracci, S., and Op de Beeck, H. P. (2016). Deep neural networks as a computational model for human shape sensitivity. *PLoS computational biology*, 12(4):e1004896.
- Larkum, M. E. (2022). Are dendrites conceptually useful? *Neuroscience*, 489:4–14.

- Larkum, M. E., Petro, L. S., Sachdev, R. N., and Muckli, L. (2018). A perspective on cortical layering and layer-spanning neuronal elements. *Frontiers in neuroanatomy*, page 56.
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *nature*, 521(7553):436–444.
- Lee, C., Sarwar, S. S., Panda, P., Srinivasan, G., and Roy, K. (2020). Enabling spike-based backpropagation for training deep neural network architectures. *Frontiers in neuroscience*, page 119.
- Lee, D.-H., Zhang, S., Fischer, A., and Bengio, Y. (2015). Difference target propagation. In *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2015, Porto, Portugal, September 7-11, 2015, Proceedings, Part I 15*, pages 498–515. Springer.
- Lee, J. H., Delbruck, T., and Pfeiffer, M. (2016). Training deep spiking neural networks using backpropagation. *Frontiers in neuroscience*, 10:508.
- Legenstein, R., Pecevski, D., and Maass, W. (2008). A learning theory for reward-modulated spike-timing-dependent plasticity with application to biofeedback. *PLoS computational biology*, 4(10):e1000180.
- Liao, Q., Leibo, J., and Poggio, T. (2016). How important is weight symmetry in backpropagation? In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30.
- Lillicrap, T. P., Cownden, D., Tweed, D. B., and Akerman, C. J. (2014). Random feedback weights support learning in deep neural networks. *arXiv preprint arXiv:1411.0247*.
- Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., and Hinton, G. (2020). Backpropagation and the brain. *Nature Reviews Neuroscience*, 21(6):335–346.
- Magee, J. C. and Grienberger, C. (2020). Synaptic plasticity forms and functions. *Annual review of neuroscience*, 43:95–117.
- Mancilla, J. G., Lewis, T. J., Pinto, D. J., Rinzel, J., and Connors, B. W. (2007). Synchronization of electrically coupled pairs of inhibitory interneurons in neocortex. *Journal of Neuroscience*, 27(8):2058–2073.
- Marblestone, A. H., Wayne, G., and Kording, K. P. (2016). Toward an integration of deep learning and neuroscience. *Frontiers in computational neuroscience*, page 94.
- Mazzoni, P., Andersen, R. A., and Jordan, M. I. (1991). A more biologically plausible learning rule for neural networks. *Proceedings of the National Academy of Sciences*, 88(10):4433–4437.
- McCormick, D. A., Connors, B. W., Lighthall, J. W., and Prince, D. A. (1985). Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. *Journal of neurophysiology*, 54(4):782–806.

- Meulemans, A., Carzaniga, F., Suykens, J., Sacramento, J., and Grewe, B. F. (2020). A theoretical framework for target propagation. *Advances in Neural Information Processing Systems*, 33:20024–20036.
- Millidge, B., Seth, A., and Buckley, C. L. (2021). Predictive coding: a theoretical and experimental review. *arXiv preprint arXiv:2107.12979*.
- Millidge, B., Tschantz, A., and Buckley, C. L. (2022). Predictive coding approximates backprop along arbitrary computation graphs. *Neural Computation*, 34(6):1329–1368.
- Millidge, B., Tschantz, A., Seth, A. K., and Buckley, C. L. (2020). Activation relaxation: A local dynamical approximation to backpropagation in the brain. *arXiv preprint arXiv:2009.05359*.
- O’Reilly, R. C. (1996). Biologically plausible error-driven learning using local activation differences: The generalized recirculation algorithm. *Neural computation*, 8(5):895–938.
- Poirazi, P. and Papoutsi, A. (2020). Illuminating dendritic function with computational models. *Nature Reviews Neuroscience*, 21(6):303–321.
- Potjans, T. C. and Diesmann, M. (2014). The cell-type specific cortical microcircuit: relating structure and activity in a full-scale spiking network model. *Cerebral cortex*, 24(3):785–806.
- Rao, R. P. and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, 2(1):79–87.
- Richards, B. A. and Lillicrap, T. P. (2019). Dendritic solutions to the credit assignment problem. *Current opinion in neurobiology*, 54:28–36.
- Roelfsema, P. R. and Holtmaat, A. (2018). Control of synaptic plasticity in deep cortical networks. *Nature Reviews Neuroscience*, 19(3):166–180.
- Sacramento, J., Ponte Costa, R., Bengio, Y., and Senn, W. (2018). Dendritic cortical microcircuits approximate the backpropagation algorithm. *Advances in neural information processing systems*, 31.
- Scellier, B. and Bengio, Y. (2017). Equilibrium propagation: Bridging the gap between energy-based models and backpropagation. *Frontiers in computational neuroscience*, 11:24.
- Schiess, M., Urbanczik, R., and Senn, W. (2016). Somato-dendritic synaptic plasticity and error-backpropagation in active dendrites. *PLoS computational biology*, 12(2):e1004638.
- Schmidt, M., Bakker, R., Hilgetag, C. C., Diesmann, M., and van Albada, S. J. (2018). Multi-scale account of the network structure of macaque visual cortex. *Brain Structure and Function*, 223(3):1409–1435.

- Seung, H. S. (2003). Learning in spiking neural networks by reinforcement of stochastic synaptic transmission. *Neuron*, 40(6):1063–1073.
- Sjostrom, P. J., Rancz, E. A., Roth, A., and Hausser, M. (2008). Dendritic excitability and synaptic plasticity. *Physiological reviews*, 88(2):769–840.
- Sporea, I. and Grüning, A. (2013). Supervised learning in multilayer spiking neural networks. *Neural computation*, 25(2):473–509.
- Spruston, N. (2008). Pyramidal neurons: dendritic structure and synaptic integration. *Nature Reviews Neuroscience*, 9(3):206–221.
- Stapmanns, J., Hahne, J., Helias, M., Bolten, M., Diesmann, M., and Dahmen, D. (2021). Event-based update of synapses in voltage-based learning rules. *Frontiers in neuroinformatics*, page 15.
- Thomson, A. M. and Bannister, A. P. (2003). Interlaminar connections in the neocortex. *Cerebral cortex*, 13(1):5–14.
- Urbanczik, R. and Senn, W. (2014). Learning by the dendritic prediction of somatic spiking. *Neuron*, 81(3):521–528.
- van Albada, S., Morales-Gregorio, A., Dickscheid, T., Goulas, A., Bakker, R., Bludau, S., Palm, G., Hilgetag, C.-C., and Diesmann, M. (2022). Bringing anatomical information into neuronal network models. In *Computational Modelling of the Brain*, pages 201–234. Springer.
- Van Albada, S. J., Rowley, A. G., Senk, J., Hopkins, M., Schmidt, M., Stokes, A. B., Lester, D. R., Diesmann, M., and Furber, S. B. (2018). Performance comparison of the digital neuromorphic hardware spinnaker and the neural network simulation software nest for a full-scale cortical microcircuit model. *Frontiers in neuroscience*, 12:291.
- van Elburg, R. A. and van Ooyen, A. (2010). Impact of dendritic size and dendritic topology on burst firing in pyramidal cells. *PLoS computational biology*, 6(5):e1000781.
- Vaughn, M. J. and Haas, J. S. (2022). On the diverse functions of electrical synapses. *Frontiers in Cellular Neuroscience*, 16.
- Vezoli, J., Falchier, A., Jouve, B., Knoblauch, K., Young, M., and Kennedy, H. (2004). Quantitative analysis of connectivity in the visual cortex: extracting function from structure. *The Neuroscientist*, 10(5):476–482.
- Werfel, J., Xie, X., and Seung, H. (2003). Learning curves for stochastic gradient descent in linear feedforward networks. *Advances in neural information processing systems*, 16.