

Error-driven learning supports Bayes-optimal multisensory integration via conductance-based dendrites

Jakob Jordan, João Sacramento, Mihai A Petrovici, Walter Senn

Summary Animals collect information about their environment through a variety of senses that need to be integrated into a coherent perspective. Since all sensory information is both incomplete and corrupted by noise, this integration has the main goal of increasing the information otherwise obtained from only a single sense. The Bayes-optimal estimate for the most likely stimulus under the assumption of independent Gaussian noise is obtained by averaging estimates from different modalities while weighting each with its respective reliability. It was previously demonstrated in behavioral experiments that animals and humans combine multisensory stimuli in this optimal manner (Ernst and Banks, 2002; Fetsch et al., 2009; Nikbakht et al., 2018). What type of neuronal circuitry is able to perform such sensory integration? We present a neuron model capable of implementing the required computations by exploiting the biophysical dynamics of conductance-based neurons with dendritic compartments. Furthermore, a plausible error-driven plasticity rule enables neurons to learn not only input-output mappings, but to also simultaneously represent the respective reliabilities of each input that are necessary for a Bayes-optimal integration. In addition, the model supports dynamic reweighting of modalities and can thereby react to changes in stimulus reliabilities on a much shorter time scale than the one of synaptic plasticity. While both neuron and synapse dynamics are derived from a probabilistic description of neuronal processing, the model does not require a Bayes-optimal teacher but only input-output samples, allowing efficient learning. To illustrate our model, we present a feed-forward circuit receiving input from two different modalities with different associated reliabilities and show that after learning, the circuit optimally takes into account the respective reliabilities when processing new information. Finally, we discuss extensions of our model to non-linear dendritic compartments and to multi-layered cortical circuits that learn continuous input-output mappings (Dold et al., 2018).

Additional Detail The Bayesian estimation of a random variable x from multiple observations x_i proceeds by computing the posterior distribution $p(x|x_0, \dots, x_n)$, via a suitably chosen generative model $p(x_i|x)$ and prior distribution $p(x)$. Similarly, here we seek the posterior distribution of output rates of a neuronal circuit, given activities in multiple input populations projecting to separate dendritic compartments: $p(\mathbf{r}_1|\mathbf{r}_0) = p(\mathbf{r}_1|\mathbf{r}_0^0, \dots, \mathbf{r}_0^D)$. To compute this posterior, we define a generative model $p(\mathbf{r}_0^d|\mathbf{r}_1) := \frac{1}{Z_{\mathbf{r}_0}(\mathbf{r}_1)} e^{-\frac{1}{2} \mathbf{g}_1^{d,T}(\mathbf{r}_0)(\rho^{-1}(\mathbf{r}_1) - \mathbf{u}_1^d(\mathbf{r}_0))^2}$. For ease of notation we define the somatic potential as a deterministic mapping from rates via the inverse transfer function: $\mathbf{u}_1 := \rho^{-1}(\mathbf{r}_1)$. Choosing an appropriate prior and assuming rates of neurons projecting to different dendrites to be conditionally independent we can apply Bayes' theorem to arrive at:

$$p(\mathbf{r}_1|\mathbf{r}_0) = \frac{1}{Z} e^{-\frac{1}{2} g^L |\mathbf{u}_1 - E^L|^2} \prod_{d=1}^D e^{-\frac{1}{2} \mathbf{g}_1^{d,T}(\mathbf{r}_0)(\mathbf{u}_1 - \mathbf{u}_1^d(\mathbf{r}_0))^2}, \quad (1)$$

with the square in the second term acting elementwise and $Z = \prod_{i=1}^n \sqrt{2\pi/(g^L + \sum_{d=1}^D g_{1,i}^d)}$, where $n := \dim(\mathbf{r}_1)$. The typical somatic potential represents a compromise between the dendritic potentials according to their respective reliabilities, similar to products of experts (Hinton, 2002). Dendritic potentials (\mathbf{u}_1^d) are determined from presynaptic rates (\mathbf{r}_0) and synaptic weights (\mathbf{w}), where we assume excitatory projections from the input layer directly onto modality-specific dendrites and via branch-specific inhibitory interneurons with transfer function ρ_I . The dendritic conductance is hence given by $\mathbf{g}_1^d := g^L + \mathbf{w}^E \mathbf{r}_0 + \mathbf{w}^I \rho_I(\mathbf{r}_0)$ and the dendritic potentials by $\mathbf{u}_1^d := \frac{g^L E^L + \mathbf{w}^E E^E \mathbf{r}_0 + \mathbf{w}^I E^I \rho_I(\mathbf{r}_0)}{g^L + \mathbf{w}^E \mathbf{r}_0 + \mathbf{w}^I \rho_I(\mathbf{r}_0)}$. We define the energy function $E(\mathbf{u}_1, \mathbf{r}_0) := -\log p(\mathbf{u}_1|\mathbf{r}_0)$ and propose that neuron and weight dynamics minimize this energy by gradient descent (Rao and Ballard, 1999; Scellier and Bengio, 2017; Sacramento et al., 2018). This leads to the following neuron dynamics, resembling a leaky integrator with conductance-based synapses: $c_m \frac{d\mathbf{u}_1}{dt} = -g^L(\mathbf{u}_1 - E^L) - \sum_{d=1}^D \mathbf{g}_1^d(\mathbf{u}_1 - \mathbf{u}_1^d)$. In the stationary case ($\frac{d\mathbf{u}_1}{dt} = 0$), the somatic potential is given by a convex combination of leak and dendritic potentials $\mathbf{u}_1 = \frac{g^L E^L + \sum_{d=1}^D \mathbf{g}_1^d \mathbf{u}_1^d}{g^L + \sum_{d=1}^D \mathbf{g}_1^d}$. Note the similarity to the maximum-a-posterior (MAP) estimate for combining observations from n sources

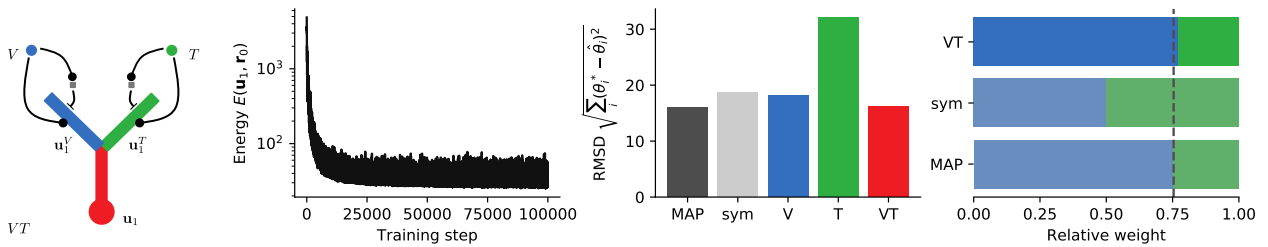


Figure 1: Multisensory integration of visual and tactile information in conductance-based neurons with dendrites. (a) Sketch of multimodal neuron receiving inputs from unimodal populations. (b) Energy $E(\mathbf{u}_1, \mathbf{r}_0)$ as a function of training step. (c) Root-mean-square deviation (RMSD) of true orientations θ_i^* and estimates $\hat{\theta}_i$. See text for abbreviations. (d) Relative weight of visual (blue) and tactile (green) input. For the different conditions, the visual weight is given by: $\frac{1}{\sigma_V^2} / (\frac{1}{\sigma_V^2} + \frac{1}{\sigma_T^2})$ ("MAP"), 0.5 ("sym"), $(\mathbf{w}_V^E + \mathbf{w}_V^I) / (\mathbf{w}_V^E + \mathbf{w}_V^I + \mathbf{w}_T^E + \mathbf{w}_T^I)$ ("VT"). Corrupting noise on the visual modality 42% smaller than on the tactile modality.

with corrupting Gaussian noise of different magnitudes: $\hat{x} = 1 / \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right)$. For appropriate relative ratios of \mathbf{g}_1^d , the stationary somatic potential can hence be interpreted as the Bayes-optimal combination of dendritic potentials. The dynamics of excitatory weights is given by:

$$\frac{d\mathbf{w}^E}{dt} = \eta \left((\mathbf{u}_1 - \mathbf{u}_1^n)(E^E - \mathbf{u}_1^n) - \frac{1}{2} \left((\mathbf{u}_1 - \mathbf{u}_1^n)^2 - \frac{1}{\mathbf{g}_1^n} \right) \right) \mathbf{r}_0^T, \quad (2)$$

with $\mathbf{g}_1^n := g^L + \sum_{d=1}^D \mathbf{g}_1^d$, and $\mathbf{u}_1^n := \frac{g^L E^L + \sum_{d=1}^D \mathbf{g}_1^d \mathbf{u}_1^d}{g^L + \sum_{d=1}^D \mathbf{g}_1^d}$. Inhibitory weights follow identical dynamics with $E^E \rightarrow E^I$. While the first term matches average somatic and potentials similar to a previously proposed somatodendritic error reduction rule (Urbanczik and Senn, 2014), the second term matches the total conductance with the variability of the total dendritic potential with respect to the somatic potential.

To illustrate this model, we consider a cue-combination task from visual and tactile orientation information. Each modality is represented by a homogeneous population with von Mises tuning curves in which each neuron has a different preferred orientation. All neurons in one population project to a specific dendritic branch on each of the neurons in the output population where they elicit dendritic potentials \mathbf{u}_1^d (figure 1a). For simplicity, we assume large and symmetric transfer conductances between the somatic and dendritic compartments and a linear transfer function for interneurons. Training consist in generating ground-truth orientations θ^* which are used to generate noisy modality-specific orientations represented by the visual and tactile population with modality specific reliabilities: $\theta^{V/T} \sim \mathcal{N}(\theta^*, \sigma_{V/T}^2)$. For each output neuron, θ^* determines a specific target potential \mathbf{u}_1 which in combination with the input-driven dendritic potentials \mathbf{u}_1^d is used to compute synaptic weight updates according to equation (2) and the corresponding rule for inhibitory synapses. From the output population the estimated angle $\hat{\theta}$ is reconstructed using a population-vector readout (see, e.g., Herz et al., 2017). Over the course of learning synaptic plasticity continuously decreases the energy (figure 1b). We compare the root-mean-square deviation between the ground-truth orientation θ^* , and the Bayes-optimal maximum a posteriori estimate ("MAP"), a naive estimate that equally weights visual and tactile input ("sym"), the estimate reconstructed from the output population providing only visual input ("V"), only tactile input ("T"), or both visual and tactile input ("VT") (figure 1c). While the naive and single modal estimates perform significantly worse than the MAP estimate, the estimate reconstructed from the population when both modalities are active achieves similar error levels. The output population indeed learned the reliabilities associated with each modality (figure 1d) and takes these into account when combining information from the two channels.

We presented a neuron model capable of performing multisensory integration based on learned modality-specific reliabilities and demonstrated that it achieves similar performance to a Bayes-optimal observer. On a conceptual level, our approach makes use of divisive normalization, similar to previously proposed models of multisensory integration (Ohshiro et al., 2011) but without the need for additional recurrent circuitry. Due to this reduced complexity our model can be trained using an extended error-reducing learning rule, making it particularly suitable to be integrated into multi-layer cortical microcircuits in which errors can be backpropagated across layers (Sacramento et al., 2018). In particular, this would support a plastic hierarchical representation of reliabilities for various sensory streams.