Uncertainty-modulated prediction errors in cortical microcircuits

Katharina A. Wilmes[1,*], Mihai A. Petrovici[1], Shankar Sachidhanandam[1], Walter Senn[1]

1 Department of Physiology, University of Bern, Bern, Switzerland
* corresponding author: katharina.wilmes@unibe.ch

May 23, 2023

Abstract

To make contextually appropriate predictions in a stochastic environment, the brain needs to take uncertainty into account. Prediction error neurons have been identified in layer 2/3 of diverse brain areas. How uncertainty modulates prediction error activity and hence learning is, however, unclear. Here, we use a normative approach to derive how prediction errors should be modulated by uncertainty and postulate that such uncertainty-weighted prediction errors (UPE) are represented by layer 2/3 pyramidal neurons. We further hypothesise that the layer 2/3 circuit calculates the UPE through the subtractive and divisive inhibition by different inhibitory cell types. We ascribe different roles to somatostatin-positive (SST), and parvalbumin-positive (PV) interneurons. By implementing the calculation of UPEs in a microcircuit model, we show that different cell types in cortical circuits can compute means, variances and UPEs with local activity-dependent plasticity rules. Finally, we show that the resulting UPEs enable adaptive learning rates.

Introduction

Decades of cognitive research indicate that our brain maintains a model of the world, based on which it can make predictions about upcoming stimuli (Kveraga *et al.*, 2007; Cohen *et al.*, 2011). Predicting the sensory experience is useful for both perception and learning: Perception becomes more tolerant to noise when sensory information and predictions are integrated (Payzan-LeNestour and Bossaerts, 2011). Learning can happen when predictions are compared to sensory information, as the resulting prediction error indicates how to improve the internal model. In both cases, the uncertainties (associated with the sensory information and with the internal model) are an important factor that according to theoretical accounts should determine how much weight we give to the sensory information relative to the predictions. Behavioural and electrophysiological studies indicate that humans indeed estimate uncertainty and adjust their behaviour accordingly (Payzan-LeNestour and Bossaerts, 2011; Walker *et al.*, 2019; Goris *et al.*, 2018; Cannon *et al.*, 2021). The neural mechanisms underlying uncertainty and prediction error computation are, however, less well understood. Recently, the activity of individual neurons of layer 2/3 cortical circuits in diverse cortical areas of mouse brains has been linked to prediction errors (visual, (Keller *et al.*, 2012; Zmarz and Keller, 2016; Fiser *et al.*, 2016; Attinger *et al.*, 2017; Gillon *et al.*, 2021), auditory (Eliades and Wang, 2008; Keller and Hahnloser, 2009), somatosensory (Ayaz *et al.*, 2019)). These neural circuits consist of heterogeneous cell types, most of which are inhibitory. It has been suggested that prediction error activity results from an imbalance of excitatory and inhibitory inputs (Hertäg and Sprekeler, 2020; Hertäg and Clopath, 2022), and that the prediction is subtracted from the sensory input (see e.g. Rao and Ballard, 1999; Attinger *et al.*, 2017), possibly mediated by so-called somatostatin-positive interneurons (SSTs) (Attinger *et al.*, 2017). How uncertainty is influencing these computations has not yet been investigated. Prediction error neurons receive inputs from a diversity of inhibitory cell types (Fig. 1), the role of which is not completely understood. We here hypothesise that one role of inhibition is to modulate the prediction error neuron activity by uncertainty.

Results

Normative theories suggests uncertainty-modulated prediction errors (UPEs)

In an complex, uncertain, and hence partly unpredictable world, it is impossible to avoid prediction errors. Some prediction errors will be the result of this variability or noise, other prediction errors will be the result of a change in the environment or new information. Ideally, only the latter should be used for learning, i.e., updating the current model of the world. The challenge our brain faces is to learn from prediction errors that result from new information, and less from prediction errors that result from noise. Hence, intuitively, if we learned that a kind of stimulus or context is very variable (high uncertainty), then a prediction error should have only little influence on our model. If, on the contrary, we learned that the kind of stimulus or context is not very variable (low uncertainty), a prediction error should have a larger impact on our model. This intuition of modulating prediction errors by the uncertainty associated with the

stimulus or context is supported by both behavioural studies and normative theories of learning.

Before we consider how cortical circuits compute such uncertainty-modulated prediction errors, let us consider the normative solution to a simple association that a mouse can learn, and which involves a simple Gaussian-distributed stimulus such that our theory will be able to make experimentally testable predictions. The setting we consider is to predict a somatosensory stimulus based on an auditory stimulus (Fig. 1A). The auditory stimulus a is fixed, and the subsequent somatosensory stimulus s is variable and sampled from a Gaussian distribution (s $\sim \mathcal{N}(\mu, \sigma)$, Fig. 1B). Framed as an optimisation problem, the goal is to adapt the internal model $\hat{\mu}$ such that the probability of observing samples s from the true distribution of whisker deflections is maximised given this model.
Hence, stochastic gradient descent (SGD) learning on the negative log likelihood

$$-\log L = -\log p(\mathbf{s}|\hat{\mu}, \hat{\sigma}) = -\log \prod_{n=1}^{N} \mathcal{N}(s_n|\hat{\mu}, \hat{\sigma}), \tag{1}$$

shows that with each observation s, the prediction, corresponding to the internal model of the mean, should be updated as follows to approach the maximum likelihood solution:

$$\Delta\hat{\mu} \propto -\frac{\partial}{\partial\hat{\mu}}(-\log L) = \frac{1}{\sigma^2}(s - \hat{\mu}). \tag{2}$$

According to this formulation, the update for the internal model should be the prediction error scaled inversely by the variance σ^2. Therefore, we propose that prediction errors should be modulated by uncertainty.

Computation of UPEs in cortical microcircuits

How can uncertainty modulation be achieved by cortical microcircuits? Prediction errors can be either positive or negative but neuronal firing rates are always positive. Because in addition, baseline firing rates are low in layer 2/3 pyramidal cells, positive and negative prediction errors were suggested to be represented by distinct neuronal populations (Keller and Mrsic-Flogel, 2018), which is in line with the data (Jordan and Keller, 2020). We, therefore, decompose the UPE into a positive UPE$^+$ and a negative UPE$^-$ component (Fig. 1C,D):

$$\text{UPE} := \text{UPE}^+ - \text{UPE}^- = \frac{1}{\sigma^2}\lfloor s - \mu \rfloor^+ - \frac{1}{\sigma^2}\lfloor \mu - s \rfloor^+ \tag{3}$$

It has been suggested that error neurons compute prediction errors by subtracting the prediction from the stimulus input (or vice versa), resulting in an excitation-inhibition balance when they match. To represent a UPE, error neurons need to additionally be divisively modulated by the uncertainty. Depending on synaptic properties, such as reversal potentials, inhibitory neurons can have subtractive or divisive influences on their postsynaptic targets. Therefore, we propose that an inhibitory cell type that divisively modulates prediction error activity represents the uncertainty. Prediction error neurons in layer 2/3 receive subtractive inhibition from somatostatin (SST) and divisive inhibition from parvalbumin (PV) interneurons (Wilson *et al.*, 2012). Therefore, we hypothesise that in positive prediction error circuits, the SSTs represent the mean μ of the prediction (which is then subtracted from the prediction error neuron activity by subtractive inhibition), and that the PVs represent the uncertainty σ^2 of the prediction (which then divisively modulates the prediction error neuron activity (Fig. 1C,D), such that the layer 2/3 pyramidal cell activity reflects the uncertainty-modulated prediction error.

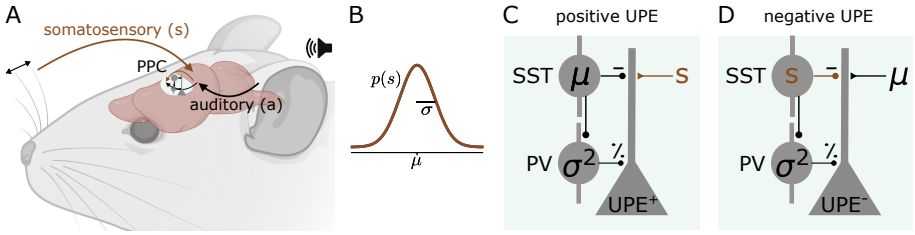

Figure 1: **Computation of UPEs in cortical microcircuits** A: a mouse learns the association between a sound (a) and a whisker deflection (s). The posterior parietal cortex (PPC) receives inputs from both somatosensory and auditory cortex. B: The whisker stimulus intensities are drawn from a Gaussian distribution with mean μ and standard deviation σ. C: Positive prediction error circuit consisting of three cell types: layer 2/3 pyramidal cells (triangle), somatostatin-positive interneurons (SST, circle) and parvalbumin-positive interneurons (PV). SSTs represent the mean prediction, and PVs the variance. D: Negative prediction error circuit, similar to C, SST now represent the stimulus and the mean prediction is an excitatory input.

Local inhibitory cells learn to represent the mean and the variance given an associative cue

As we have discussed above, how much an individual sensory input should contribute to updating our internal model depends on the uncertainty associated with the sensory stimulus. As the estimation of the uncertainty requires multiple stimulus samples, the uncertainty cannot be estimated from a single input. Additionally, the uncertainty may be context-dependent. Therefore, our brain needs to have a mechanism to learn about the uncertainty from past instances of the sensory input. Let us consider the simple example from above, in which a sound stimulus represents a context, which is associated with a particular amount of uncertainty. Here, we investigate whether the presentation of the sound can elicit activity in the PVs that reflects the uncertainty of the situation in addition to the sound causing an expectation of the stimulus intensity. To investigate whether a sound can cause activity in SSTs and PVs that reflects the mean and the variance of the whisker stimulus intensities, respectively, we simulated a rate-based circuit model, consisting of pyramidal cells and the relevant inhibitory cell types.

SSTs can estimate the mean

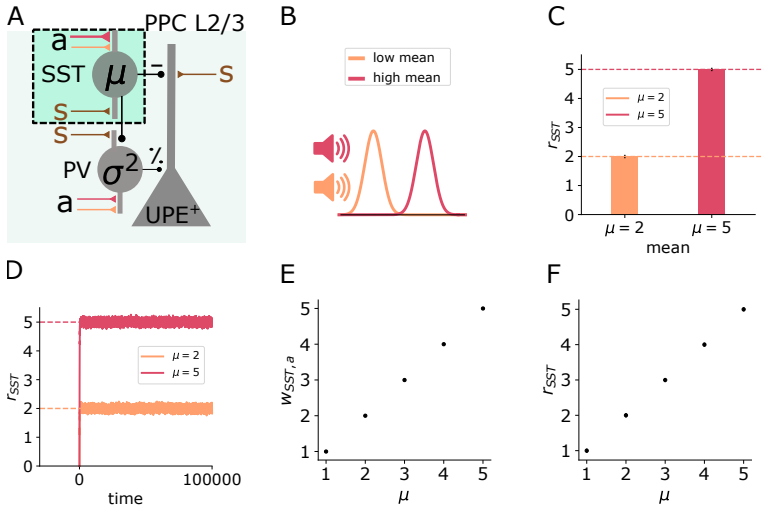

Figure 2: **SSTs learn to represent the mean given an associative cue** Illustration of the changes in the positive prediction error circuit. Thicker lines denote stronger weights. B: Two different tones (red, orange) are associated with two somatosensory stimulus distributions with different means (red: high, orange: low). C: SST firing rates (mean and std) during stimulus input. D: SST firing rates over time for low (orange) and high (red) stimulus means. E: Weights from sound A to SST for different values of μ (mean and std from timestep 100000 to the end of the simulation). F: SST firing rates for different values of μ (mean and std from timestep 100000 to the end of the simulation).

To investigate whether a sound input can cause SSTs to fire proportional to the mean whisker stimulus intensity, we simulate a firing-rate based model of the circuit (Fig. 2A, see Methods), in which SSTs receive whisker stimulus inputs s, drawn from Gaussian distributions (Fig. 2B) and an input from a higher level representation of the sound a (which is either on or off, see Methods) with a weight $w_{\mathrm{SST,a}}$ which is plastic according to the following local activity-dependent plasticity rule (Urbanczik and Senn, 2014):

$$\Delta w_{\mathrm{SST,a}} = \eta(r_{\mathrm{SST}} - \phi(w_{\mathrm{SST,a}}\, r_{\mathrm{a}}))\, r_{\mathrm{a}}, \tag{4}$$

where η is the learning rate, r_{a} is the pre-synaptic firing rate, r_{SST} is the post-synaptic firing rate of the SSTs, $\phi(x)$ is a rectified linear activation function of the SSTs, and the SST activity is determined by

$$\tau_{\mathrm{I}} \frac{dr_{\mathrm{SST}}}{dt} = -r_{\mathrm{SST}} + \phi((1-\beta) w_{\mathrm{SST,a}}\, r_{\mathrm{a}} + \beta s). \tag{5}$$

The SST activity is influenced (nudged with a factor β) by the somatosensory stimuli s, which provide targets for the desired SST activity. The learning rule ensures that the auditory input alone causes SSTs to fire at their target activity. As in the original proposal (Urbanczik and Senn, 2014), the terms in the learning rule can be mapped to local neuronal variables, which could be represented by dendritic ($w_{\mathrm{SST,a}}\, r_{\mathrm{a}}$) and somatic ($r_{\mathrm{SST}}$) activity. After learning, the weight and the average SST firing rate reflect the expected value—the mean—of the presented whisker stimulus intensities (Fig. 2C-F).

PVs can estimate the variance

The PVs could either inherit the representation of the uncertainty from another brain area, or they could estimate it themselves. We next addressed whether a local variance estimate is possible to achieve with PVs. To estimate the

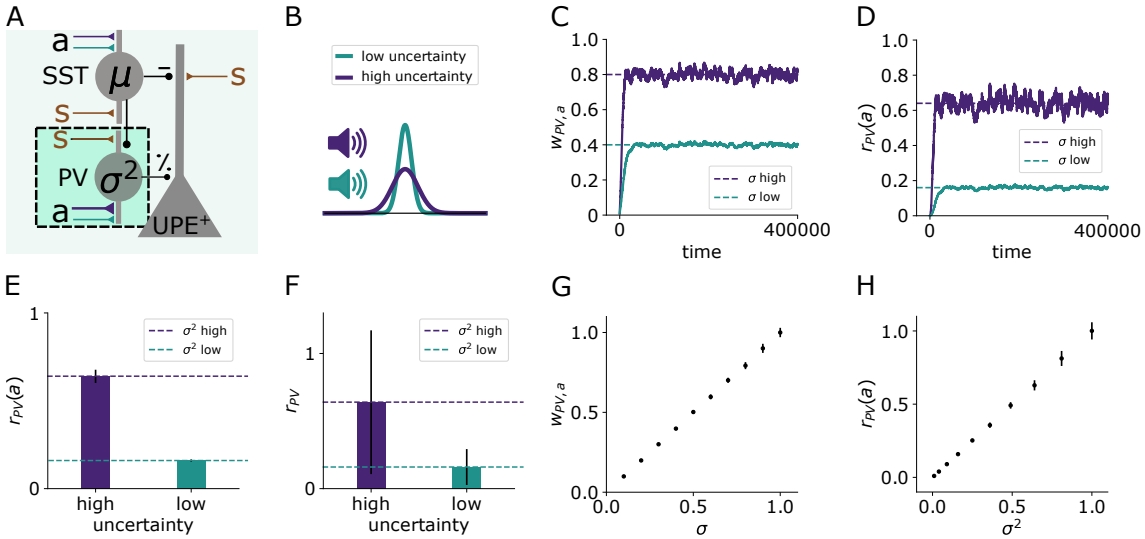

Figure 3: **PVs learn to represent the variance given an associative cue.** A: Illustration of the changes in the positive prediction error circuit. Thicker lines denote stronger weights. B: Two different tones (purple, green) are associated with two somatosensory stimulus distributions with different variances (purple: high, green: low). C: Weights from sound a to PV over time for two different values of stimulus variance (high: $\sigma = 0.8$ (purple), low: $\sigma = 0.4$ (green)). D: PV firing rates over time given sound input (without stimulus input) for low (green) and high (purple) stimulus variance. E: PV firing rates (mean and std) given sound input for low and high stimulus variance. F: PV firing rates (mean and std) during sound and stimulus input. G: Weights from sound a to PV for different values of σ (mean and std from timestep 100000 to the end of the simulation). H: PV firing rates given sound input for different values of σ^2 (mean and std from timestep 100000 to the end of the simulation).

variance of the whisker deflections s, the PVs have to estimate $\sigma^2[s] = \mathbb{E}_s[(s - \mathbb{E}[s])^2] = \mathbb{E}_s[(s-\mu)^2]$. To do so, they need to have access to both the whisker stimulus s and the mean μ. PVs in PPC respond to sensory inputs in diverse cortical areas (S1: Sachidhanandam *et al.*, 2016) and are inhibited by SSTs in layer 2/3, which we assumed to represent the mean. Finally, for calculating the variance, these inputs need to be squared. PVs were shown to integrate their inputs supralinearly (Cornford *et al.*, 2019), which could yield an approximation to the variance. To test whether the PVs can learn to represent the variance of an upcoming whisker stimulus based on an auditory input, we simulated a rate-based model of the circuit (Fig. 3A). Two different auditory inputs (Fig. 3B purple, green) are now paired with two whisker stimulus distributions that differ in their variances (green: low, purple: high). Because the PVs receive the stimulus as well as the mean subtraction from the SSTs, their firing rate is influenced (nudged) by a variable input (target), whose expected value is proportional to the standard deviation of the stimulus:

$$\tau_{\mathrm{I}} \frac{dr_{\mathrm{PV}}}{dt} = -r_{\mathrm{PV}} + \phi_{\mathrm{PV}}((1-\beta) w_{\mathrm{PV,a}}\, r_{\mathrm{a}} + \beta(w_{\mathrm{PV,s}}\, r_{\mathrm{s}} - w_{\mathrm{PV,SST}}\, r_{\mathrm{SST}}))), \tag{6}$$

where $\phi_{PV}(x)$ denotes the activation function (see Methods).
The weight from the auditory input to the PVs was plastic according to the same local activity-dependent plasticity rule as the weight to the SSTs:

$$\Delta w_{\mathrm{PV,a}} = \eta(r_{\mathrm{PV}} - \phi_{\mathrm{PV}}(w_{\mathrm{PV,a}}\, r_{\mathrm{a}}))\, r_{\mathrm{a}}. \tag{7}$$

With this learning rule, the weight onto the PV approaches σ (Fig. 3C), such that the PV firing rate approaches σ^2 on average (Fig. 3D). The average PV firing rate is exactly σ^2 with a square activation function $\phi_{PV}(x)$ (Fig. 3D-F,H) and proportional with other choices of activation functions (Suppl. Fig. 9). This holds both when the sound input is presented alone (Fig. 3D,E,H) or when it is paired with whisker stimulation (Fig. 3F).
To estimate the variance, the mean needs to be subtracted from the stimulus samples. A faithful mean subtraction is only ensured if the weight from the SSTs to the PVs ($w_{\mathrm{PV,SST}}$) matches the weight from the stimuli s to the PVs ($w_{\mathrm{PV},s}$; see Suppl. Figs. 14,15). The weight $w_{\mathrm{PV,SST}}$ can be learned to match the weight $w_{\mathrm{PV},s}$ with a local activity-dependent plasticity rule (see Suppl. Fig. 10 and Suppl. Methods).
The PVs can similarly estimate the uncertainty in negative prediction error circuits (Suppl. Fig. 11). In these circuits, SSTs represent the current sensory stimulus, and the mean prediction is an excitatory input to both negative prediction error neurons and PVs.

Calculation of the UPE in Layer 2/3 error neurons

Layer 2/3 pyramidal cell dendrites are capable of generating NMDA and calcium spikes, which cause a nonlinear integration of inputs. Therefore, the dendritic input may instead be better represented by $I_{\mathrm{dend}} = \lfloor w_{\mathrm{UPE},s}\, s - w_{\mathrm{UPE,SST}}\, r_{\mathrm{SST}} \rfloor^k$

(Fig. 4C-F for $k = 2$,G-J for $k = 2.5$). It turns out that such a nonlinear integration of inputs is convenient when the mean input changes and the current prediction differs strongly from the new mean of the stimulus distribution. In this case, the PVs do not estimate the variance, but the mean-squared error MSE $= \sum_{n=1}^{N} \frac{1}{N}(s_n - \hat{\mu})^2$. The difference is that the true mean μ is used to calculate the variance, and the mean estimate $\hat{\mu}$ is used to calculate the MSE. Because PVs reflect the MSE in this case, their firing rate will increase for larger errors, and inhibit error neurons more strongly than indicated by the learned variance estimate. The nonlinearity compensates this increased inhibition by PVs, such that in the end, layer 2/3 cell activity approximates the uncertainty-modulated prediction error (Fig. 4E) in both negative (Fig. 4A) and positive (Fig. 4B) prediction error circuits. A stronger nonlinearity has an interesting effect: error neurons elicit much larger responses to outliers than stimuli that match the predicted distribution—a cell-intrinsic form of *out-of-distribution detection*.

To ensure a comparison between the stimulus and the prediction, the weights from the SSTs to the UPE neurons need to match the weights from the stimulus s to the UPE neuron and from the mean representation to the UPE neuron, respectively. With inhibitory plasticity, the weights from the SSTs can learn to match the incoming excitatory weights (Suppl. Fig. 12).

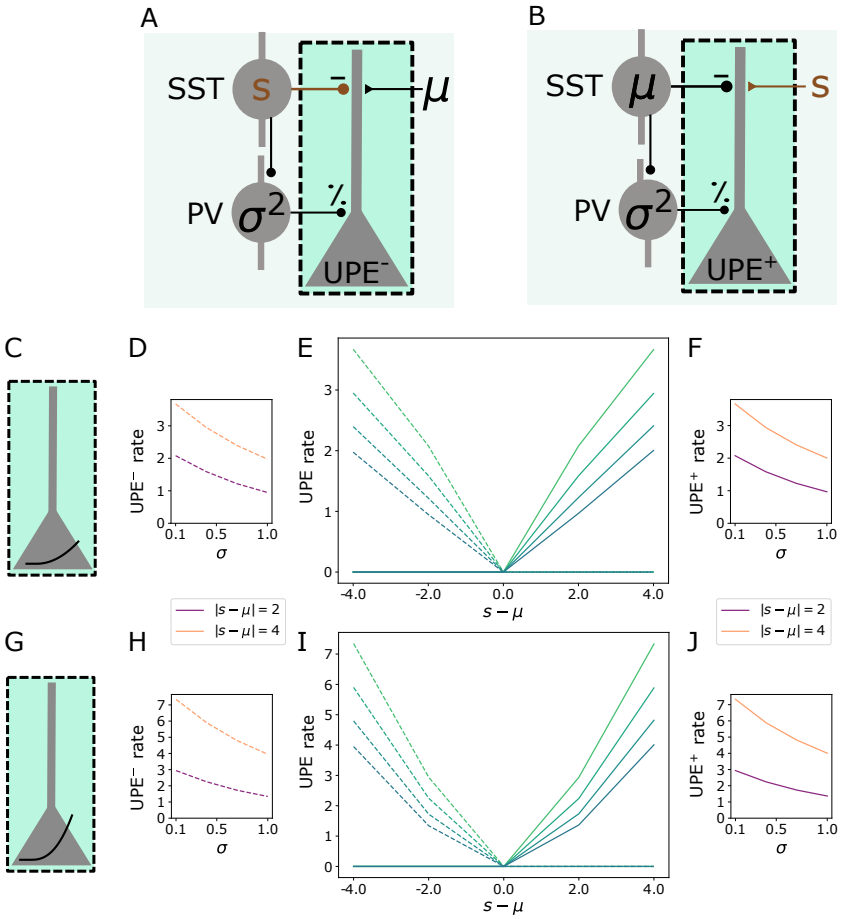

Figure 4: **Calculation of the UPE in layer 2/3 error neurons** A: Illustration of the negative prediction error circuit. B: Illustration of the positive prediction error circuit. C: Illustration of error neuron with a nonlinear integration of inputs ($k = 2$). D: firing rate of the error neuron in the negative prediction error circuit (UPE$^-$) as a function of σ for two values of $|s - \mu|$. E: Rates of both UPE$^+$ and UPE$^-$-representing error neurons as a function of the difference between the stimulus and the mean ($s - \mu$). F: firing rate of the error neuron in the positive prediction error circuit (UPE$^+$) as a function of σ for two values of $|s - \mu|$. G: Illustration of an error neuron with a non-linear activation function $k = 2.5$. H-J: same as D-F for error neurons with $k = 2.5$.

Learning the mean representation with UPEs

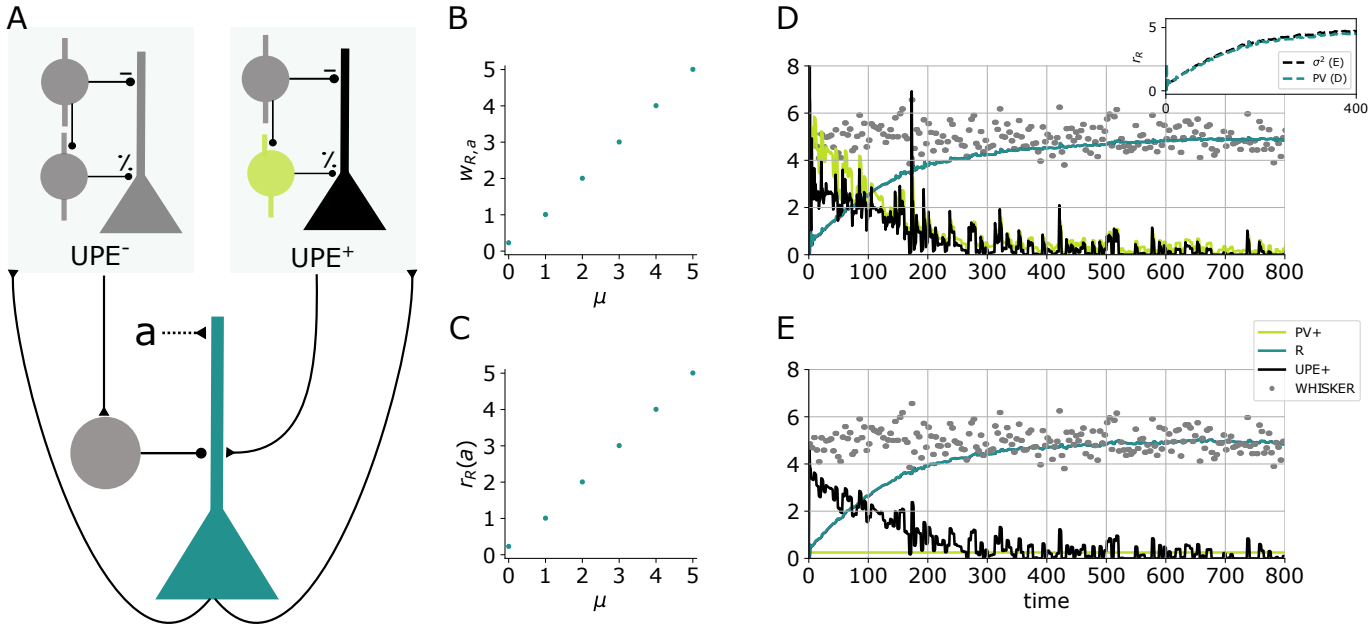

Figure 5: **Learning the mean representation with UPEs** A: Illustration of the circuit. A representation neuron (turquoise, R) receives input from both positive and negative prediction error circuits (UPE^+ and UPE^-) and projects back to them (connectivity is simplified in the illustraiton, see Methods for detailed connectivity structure). The UPE^- has a negative impact on the firing rate of the representation neuron (r_R). A weight $w_{R,a}$ from the higher level representation of the sound a is learned. B: Weights from sound a to R for different values of μ (mean and std from timestep 100000 to the end of the simulation). C: R firing rates given sound input for different values of μ (mean and std from timestep 100000 to the end of the simulation). D: Activity of the different cell types (PV: light green, SST: purple, R: green, UPE: black) and whisker stimulus samples (grey dots) over time. Learning the mean representation with PVs (light green) reflecting the MSE at the beginning, which is compensated by nonlinear activation of L2/3 neurons (black). The evolution of the mean rate of neuron R (turquoise) is similar to the perfect case in E (inset shows comparison). E: Same colour code as in D. Learning the mean representation assuming PVs (light green) perfectly represent the variance.

Here we first show that a neuron that receives inputs from L2/3 error neurons can learn a representation of the mean and hence function as a representation neuron. We implemented a circuit in which a representation neuron receives inputs from positive and negative prediction error neurons and from a higher level representation of the sound a (Fig. 5A). The representation neuron sends its current mean estimate to the error circuits by either targeting the SSTs (in the positive circuit) or the pyramidal cells directly (in the negative circuit). Hence, in this recurrent circuit, the SSTs inherit the mean representation instead of learning it. After learning in this recurrent circuit, the weights from the sound to the representation neuron and the average firing rate of this representation neuron reflects the mean of the stimulus distribution (Fig.5B,C).

A circuit in which pyramidal cells exhibit NMDA spikes (Fig. 5D) approximates learning in an idealised circuit, in which the PV rate perfectly represents the variance (Fig. 5E, see inset for comparison of the two models). Also in this recurrent circuit, PVs learn to reflect the variance, as their activity is proportional to the variance and the weight from the sound representation a is learned to be proportional to σ (Suppl. Fig. 13).

Next, we studied the link between circuit connectivity and biases in the mean representation. We showed that inhibitory plasticity enables to match the connections from SST to PV and from SST to UPE to the connections from the stimulus representation. This ensures that no biases are introduced in the computation. Interestingly, biases in perception have been linked to numerous psychological conditions. To better understand how biases depend on the connectivity in our model, we perturbed connections from the SSTs to the prediction error neurons. A 100% increase in the connection from SSTs to both prediction error neurons leads to an underestimation of the mean, which increases with increasing mean (Suppl. Fig. 16). To better understand, why this is the case, we perturbed this connections in each of the two error circuits individually. Increasing this connection only in the positive prediction error circuit has a very similar effect (Suppl. Fig. 17). while increasing this connection only in the negative prediction error circuit has the opposite effect and leads to an overestimation of the mean (Suppl. Fig. 18). The total effect is hence mostly driven by the perturbation in the positive circuit in this case. This is expected, because the total error is dominated by positive prediction errors as the initial mean prediction was too small. In any case, the learned representation is proportional to the mean.

Effective learning rate is automatically adjusted with UPEs

To investigate how UPEs influence the effective learning rate, we looked at two contexts that differed in the amount of uncertainty, and compared how the mean representation evolves with and without UPEs. (To ensure a fair comparison, the unmodulated control has an effective learning rate that is the mean of the two effective learning rates in the uncertainty-modulated case.) In a low uncertainty setting, the mean representation can be learned faster with UPEs (in comparison to unmodulated, Fig. 6A,B). In the high uncertainty setting, the effective learning rate is slower, and the mean representation is less variable than in the unmodulated case (Fig. 6B,D). The firing rate variation (measured in standard deviation of the firing rate) increases only logarithmically with the standard deviation of the inputs (Fig. 6E). Hence, uncertainty-modulation of prediction errors enables an adaptive learning rate modulation.

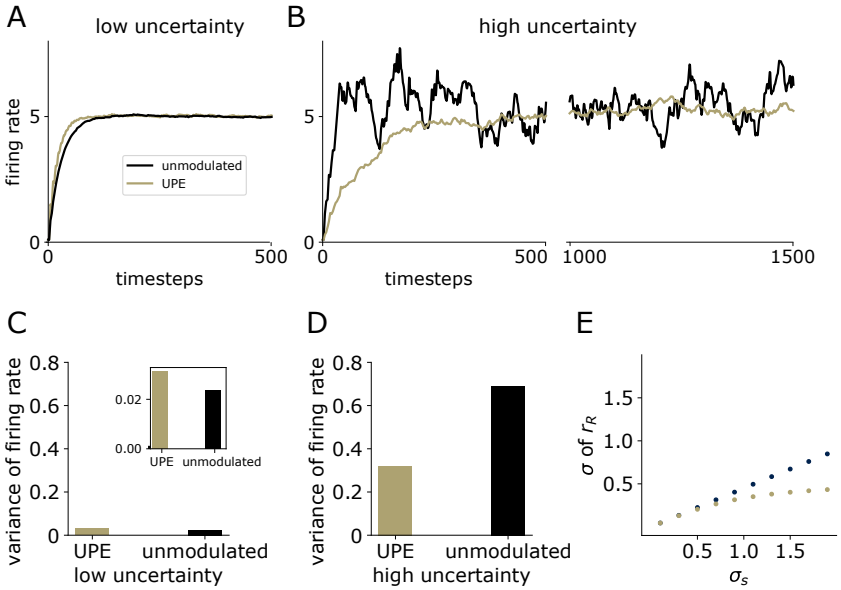

Figure 6: **Effective learning rate is automatically adjusted with UPEs** A,B: Firing rate over time of the representation neuron over time in a circuit with uncertainty-modulated prediction errors (pink) and in a circuit with unmodulated errors (black) in a low uncertainty setting (A) and a high uncertainty setting (B), C: variance of the firing rate of the representation neuron in the low uncertainty setting (inset has a different scale, outer axis scale matches the one in D), D: variance of the firing rate of the representation neuron in the high uncertainty setting. E: Variability (σ) of the firing rate r_R as a function of the standard deviation of the presented stimulus distribution σ.

Predictions for different cell types

Our model makes predictions for the activity of different cell types when a mouse receives whisker stimuli that are larger (A, black) or smaller (B, grey) than expected after a sound that is associated with a certain (green) or an uncertain context (purple). Our model suggests that there are two types of interneurons that provide subtractive inhibition to the prediction error neurons (presumably SSTs): SSTs in the positive prediction error circuit (SST^+) signal the expected value of the whisker stimulus intensity (C,D). SSTs in the negative prediction error circuit (SST^-) signal the whisker stimulus intensity (E,F). We further predict that interneurons that divisely modulate prediction error neuron activity represent the uncertainty (presumably the PVs). Those do not differ between positive and negative circuits, and may even be shared across the two circuits: PVs in both positive and negative prediction error circuits signal the variance (G,H). L2/3 pyramidal cells that encode prediction errors signal uncertainty-modulated positive prediction errors (I) and uncertainty-modulated negative prediction errors (L), respectively. Finally, it has been proposed that there are so-called internal representation neurons (Keller and Mrsic-Flogel, 2018). In our case, those neurons or a population of those neurons represent the predicted mean of the associated whisker deflections. Our model predicts that upon presentation of an unexpected whisker stimulus, those internal representation neurons adjust their activity to represent the new whisker deflection depending on the variability of the associated whisker deflections: they adjust their activity more (given equal deviations from the mean) if the associated whisker deflections are less variable.

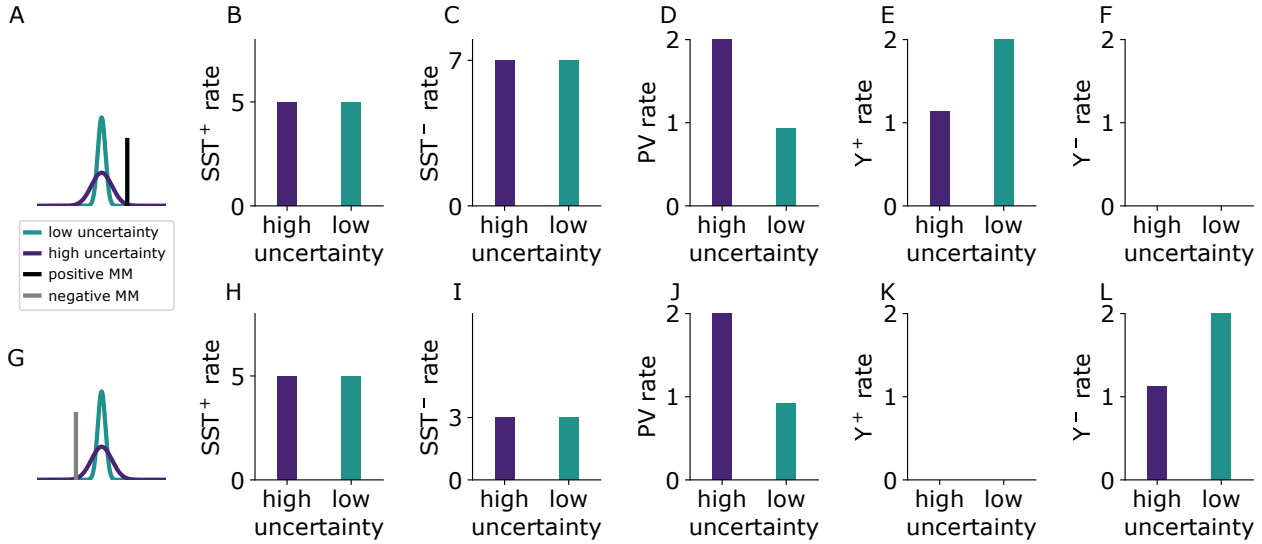

Figure 7: **Cell-type specific experimentally testable predictions** A: Illustration of the two experienced stimulus distributions with different variances that are associated with two different sounds (green, purple). The presented mismatch stimulus (black) is larger than expected (positive prediction error). B-E: Firing rates of different cell types in the positive circuit to the mismatch when a sound associated with high (purple) or low (green) uncertainty is presented. B,H: SSTs in the positive circuit, C,I: SSTs in the negative circuit, D,J: PVs in both circuits, Error neuron in the positive circuit, F,K: Error neuron in the negative circuit. G: Illustration of the two experienced stimulus distributions with different variances that are associated with two different sounds (green, purple). The presented mismatch stimulus (grey) is smaller than expected (negative prediction error). H-K: Firing rates of different cell types in the positive circuit to the mismatch when a sound associated with high (purple) or low (green) uncertainty is presented.

Discussion

Based on normative theories, we propose the existence of uncertainty-modulated prediction errors. In particular, we hypothesise that layer 2/3 prediction error neurons represent prediction errors that are inversely modulated by uncertainty. Here we showed that different inhibitory cell types in layer 2/3 cortical circuits can compute means and variances and thereby enable pyramidal cells to represent uncertainty-modulated prediction errors. We further showed that the cells in the circuit are able to learn to predict the means and variances of their inputs with local activity-dependent plasticity rules. Our study makes experimentally testable predictions for the activity of different cell types, in particular, PV and SST interneurons and L2/3 pyramidal cells. Finally, we showed that circuits with uncertainty-modulated prediction errors enable adaptive learning rates, resulting in fast learning when uncertainty is low and slow learning to avoid unnecessary fluctuations when uncertainty is high.

Here we proposed that uncertainty is locally computed by PVs and stored in the connections to those PVs on each level of the hierarchy. This has the following implications: (1) The same PV cell can store different uncertainties depending on the context, which is encoded in the pre-synaptic activation. (2) Fewer PVs than pyramidal cells are required for the mechanism (compatible with the 80/20 ratio of excitatory to inhibitory cells). (3) It allows for uncertainty estimates that are specific to the processing level of a particular area.

The uncertainty represented by PVs in our theoretical framework corresponds to expected uncertainty that results from noise in the environment, and should therefore decrease the learning rate. The uncertainty-modulation is also related to the Kalman gain in the Kalman filter. The Kalman gain, however, additionally takes into account estimate uncertainty, which decreases with the number of observations, and ideally increases again with unexpected uncertainty. The presence of unexpected rewards, punishments, or surprise, can be viewed as an increase in unexpected uncertainty that should increase the learning rate. VIPs could be responsible for signalling unexpected uncertainty. They were shown to provide potent disinhibition of pyramidal cells, are activated by reward, punishment and surprise (Pi *et al.*, 2013), and inhibit PVs in layer 2/3. Hence, they could increase error activity, which would in turn result in a larger learning signal. In general, interneurons are heavily innervated by different kinds of neuromodulators (Lee *et al.*, 2013; Prönneke *et al.*, 2015). Therefore, neuromodulators could have powerful control over error neuron activity and hence perception and learning.

Uncertainty estimation is relevant for Bayes-optimal integration of different sources of information, e.g., different modalities (Ernst and Banks, 2002; Fetsch *et al.*, 2012) or priors and sensory information. We here present a circuit implementation for weighing sensory information according to its uncertainty. It has previously been suggested that Bayes-optimal multi-sensory integration could be achieved in single neurons (Fetsch *et al.*, 2012; Jordan *et al.*, 2022). Our proposal is

complementary to this solution in that uncertainty-modulated errors can be forwarded to other cortical and subcortical circuits at different levels of the hierarchy where they can be used for integrating sensory information and for learning. It further allows for a context-dependent modulation of inputs.

To provide experimental predictions that are immediately testable, we suggested specific roles for SSTs and PVs, as they were shown to subtractively and divisively modulate pyramidal cell activity, respectively. I principle, our theory more generally posits that any subtractive or divisive inhibition could implement the suggested computations. With the emerging data on inhibitory cell types, subtypes of SSTs and PVs or other cell types may turn out to play the proposed role. To compare predictions and stimuli in a subtractive manner, the encoded prediction/stimulus needs to be translated into a direct variable code. In this framework, we are assuming that this can be achieved by the weight matrix defining the synaptic connections from the neural populations representing predictions and stimuli (possibly in a population code).

Methods

We modelled a circuit consisting of excitatory prediction error neurons in layer 2/3, and two inhibitory populations, corresponding to PV, and SST interneurons.
Layer 2/3 pyramidal cells receive divisive inhibition from PVs (Wilson *et al.*, 2012). We, hence, modelled the activity of prediction error neurons as

$$\tau_{\mathrm{E}} \frac{dr_{\mathrm{UPE}}}{dt} = -r_{\mathrm{UPE}} + \phi\left(\frac{I_{\mathrm{dend}}}{I_0 + w_{\mathrm{UPE,PV}}\, r_{\mathrm{PV}}}\right), \tag{8}$$

where $\phi(x)$ is the activation function, defined below, $I_{\mathrm{dend}} = \lfloor w_{\mathrm{UPE,s}}\, r_{\mathrm{s}} - w_{\mathrm{UPE,SST}}\, r_{\mathrm{SST}} \rfloor^k$ is the dendritic input current to the positive prediction error neuron (see below for the negative prediction error neuron). The nonlinearity in the dendrite is determined by the exponent k, which is by default $k = 2$, unless otherwise specified as in Fig. 4G-J. $I_0 > 1$ is a constant ensuring that the divisive inhibition does not become excitatory.

The PV firing rate is determined by the input from the sound representation ($w_{\mathrm{PV^+,a}}\, r_{\mathrm{a}}$) and the whisker stimuli, from which their mean is subtracted ($w_{\mathrm{PV^+,s}}\, r_{\mathrm{s}} - w_{\mathrm{PV^+,SST^+}}\, r_{\mathrm{SST^+}}$). The mean-subtracted whisker stimuli serve as a target for learning the weight from the sound representation to the PV $w_{\mathrm{PV^+,A}}$. The PV firing rate evoles over time according to:

$$\tau_{\mathrm{I}} \frac{dr_{\mathrm{PV^+}}}{dt} = -r_{\mathrm{PV^+}} + \phi_{\mathrm{PV}}((1-\beta) w_{\mathrm{PV^+,a}}\, r_{\mathrm{a}} + \beta(w_{\mathrm{PV^+,s}}\, r_{\mathrm{s}} - w_{\mathrm{PV^+,SST^+}}\, r_{\mathrm{SST^+}}))) \tag{9}$$

In Fig. 5, we added an internal representation neuron to the circuit with firing rate r_{R}.

The firing rate of each population r_i where $i \in [\mathrm{SST^+}, \mathrm{SST^-}, \mathrm{PV^+}, \mathrm{PV^-}, \mathrm{UPE^+}, \mathrm{UPE^-}, \mathrm{R}]$ evolves over time according to the following neuronal dynamics. ϕ denotes a rectified linear activation function, ϕ_{PV} denotes a rectified square activation function.

Neuronal dynamics

$$\tau_{\mathrm{I}} \frac{dr_{\mathrm{SST^+}}}{dt} = -r_{\mathrm{SST^+}} + \phi(w_{\mathrm{SST^+,R}}\, r_{\mathrm{R}}) \tag{10}$$

$$\tau_{\mathrm{I}} \frac{dr_{\mathrm{PV^+}}}{dt} = -r_{\mathrm{PV^+}} + \phi_{\mathrm{PV}}((1-\beta) w_{\mathrm{PV^+,a}}\, r_{\mathrm{a}} + \beta(w_{\mathrm{PV^+,s}}\, r_{\mathrm{s}} - w_{\mathrm{PV^+,SST^+}}\, r_{\mathrm{SST^+}}))) \tag{11}$$

$$\tau_{\mathrm{E}} \frac{dr_{\mathrm{UPE^+}}}{dt} = -r_{\mathrm{UPE^+}} + \phi\left(\frac{\lfloor w_{\mathrm{UPE,s}}\, r_{\mathrm{s}} - w_{\mathrm{UPE,SST^+}}\, r_{\mathrm{SST^+}} \rfloor^k}{I_0 + w_{\mathrm{UPE,PV^+}}\, r_{\mathrm{PV^+}}}\right) \tag{12}$$

$$\tau_{\mathrm{I}} \frac{dr_{\mathrm{SST^-}}}{dt} = -r_{\mathrm{SST^-}} + \phi(w_{\mathrm{SST^-,s}}\, r_{\mathrm{s}}) \tag{13}$$

$$\tau_{\mathrm{I}} \frac{dr_{\mathrm{PV^-}}}{dt} = -r_{\mathrm{PV^-}} + \phi_{\mathrm{PV}}((1-\beta) w_{\mathrm{PV^-,a}}\, r_{\mathrm{a}} + \beta(w_{\mathrm{PV^+,R}}\, r_{\mathrm{R}} - w_{\mathrm{PV^-,SST^-}}\, r_{\mathrm{SST^-}}))) \tag{14}$$

$$\tau_{\mathrm{E}} \frac{dr_{\mathrm{UPE^-}}}{dt} = -r_{\mathrm{UPE^-}} + \phi\left(\frac{\lfloor w_{\mathrm{UPE,R}}\, r_{\mathrm{R}} - w_{\mathrm{UPE,SST^-}}\, r_{\mathrm{SST^-}} \rfloor^k}{I_0 + w_{\mathrm{UPE,PV^-}}\, r_{\mathrm{PV^-}}}\right) \tag{15}$$

$$\tau_{\mathrm{E}} \frac{dr_{\mathrm{R}}}{dt} = -r_{\mathrm{R}} + \phi(w_{\mathrm{R,a}}\, r_{\mathrm{a}} + w_{\mathrm{R,UPE^+}}\, r_{\mathrm{UPE^+}} - w_{\mathrm{R,UPE^-}}\, r_{\mathrm{UPE^-}})), \tag{16}$$

$$\tag{17}$$

where

$$\phi(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } 0 < x < x_{\max} \\ r_{\max} & \text{if } x \geq x_{\max} \end{cases} \tag{18}$$

and

$$\phi_{\mathrm{PV}}(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x^2 & \text{if } 0 < x < x_{\max} \\ r_{\max} & \text{if } x \geq x_{\max} \end{cases} \tag{19}$$

Parameter	Value	Description
$w_{\mathrm{PV}^+,\mathrm{SST}^+}$	$\sqrt{\frac{2-\beta}{\beta}}$	weight from SST^+ to PV^+
$w_{\mathrm{PV}^+,\mathrm{s}}$	$\sqrt{\frac{2-\beta}{\beta}}$	weight from s to PV^+
$w_{\mathrm{PV}^-,\mathrm{SST}^-}$	$\sqrt{\frac{2-\beta}{\beta}}$	weight from SST^- to PV^-
$w_{\mathrm{PV}^-,\mathrm{R}}$	$\sqrt{\frac{2-\beta}{\beta}}$	weight from R to PV^-
$w_{\mathrm{SST}^+,\mathrm{R}}$	1.0	weight from R to SST^+
$w_{\mathrm{SST}^-,\mathrm{s}}$	1.0	weight from s to SST^-
$w_{\mathrm{UPE}^+,\mathrm{SST}^+}$	1.0	weight from SST^+ to UPE^+
$w_{\mathrm{UPE}^+,\mathrm{s}}$	1.0	weight from s to UPE^+
$w_{\mathrm{UPE}^-,\mathrm{R}}$	1.0	weight from R to UPE^-
$w_{\mathrm{UPE}^-,\mathrm{SST}^-}$	1.0	weight from SST^- to UPE^-
$w_{\mathrm{R},\mathrm{UPE}^+}$	0.1(1.0)	weight from UPE^+ to R (Fig. 6)
$w_{\mathrm{R},\mathrm{UPE}^-}$	0.1(1.0)	weight from UPE^- to R (Fig. 6)
$x_{\max}$	20	limits neuronal activity
β	0.1	nudging parameter

Table 1: **Parameters of the network.**

Parameter	Value	Description
a	1.0	auditory stimulus
s	$\sim \mathcal{N}(\mu, \sigma)$	somatosensory (whisker) stimulus

Table 2: **Inputs.**

Synaptic dynamics / Plasticity rules Synapses from the higher level representation of the sound a to the SSTs, PVs, and to R were plastic according to the following activity-dependent plasticity rules (Urbanczik and Senn, 2014).

$$\Delta w_{\mathrm{SST,a}} = \eta_{\mathrm{SST}}(r_{\mathrm{SST}} - \phi(w_{\mathrm{SST,a}} r_{\mathrm{a}}))r_{\mathrm{a}} \tag{20}$$

$$\Delta w_{\mathrm{PV,a}} = \eta_{\mathrm{PV}}(r_{\mathrm{PV}} - \phi_{\mathrm{Q}}(w_{\mathrm{PV,a}} r_{\mathrm{a}}))r_{\mathrm{a}} \tag{21}$$

$$\Delta w_{\mathrm{R,a}} = \eta_{\mathrm{R}}(r_{\mathrm{R}} - \phi(w_{\mathrm{R,a}} r_{\mathrm{a}}))r_{\mathrm{a}} \tag{22}$$

$$\tag{23}$$

where $\eta_{\mathrm{PV}} = 0.01\eta_R$

Simulation We paired a constant tone with N samples from the whisker stimulus distribution.

Parameter	Value	Description
η_{SST}	0.1	learning rate for $w_{\mathrm{SST}^{+/-},\mathrm{a}}$
η_{PV}	$0.01 * \eta_{\mathrm{R}} = 0.001$	learning rate for $w_{\mathrm{PV}^{+/-},\mathrm{a}}$
η_{R}	0.1	learning rate for $w_{\mathrm{R,a}}$
$w^{\mathrm{initial}}_{\mathrm{SST,a}}$	0.01	initial value for $w_{\mathrm{SST}^{+/-},\mathrm{a}}$
$w^{\mathrm{initial}}_{\mathrm{PV,a}}$	0.01	initial value for $w_{\mathrm{PV}^{+/-},\mathrm{a}}$
$w^{\mathrm{initial}}_{\mathrm{R,a}}$	0.01	initial value for $w_{\mathrm{R,a}}$

Table 3: **Parameters of the plasticity.**

Parameter	Value	Description
N	200000	number of whisker stimulus samples
D	1	stimulus duration
T	$N * D$	simulation time
dt	1.0	simulation time step
τ_E	1.0	excitatory membrane time constant
τ_I	1.0	inhibitory membrane time constant

Table 4: **Parameters of simulations in Figs. 2-5.**

Parameter	Value	Description
N	200000	number of whisker stimulus samples
D	10	stimulus duration
T	$N * D$	simulation time
dt	1.0	simulation time step
τ_E	10.0	excitatory membrane time constant
τ_I	2.0	inhibitory membrane time constant
η_R	0.01	learning rate of $w_{R,A}$

Table 5: **Parameters of the simulation in Fig. 6.**

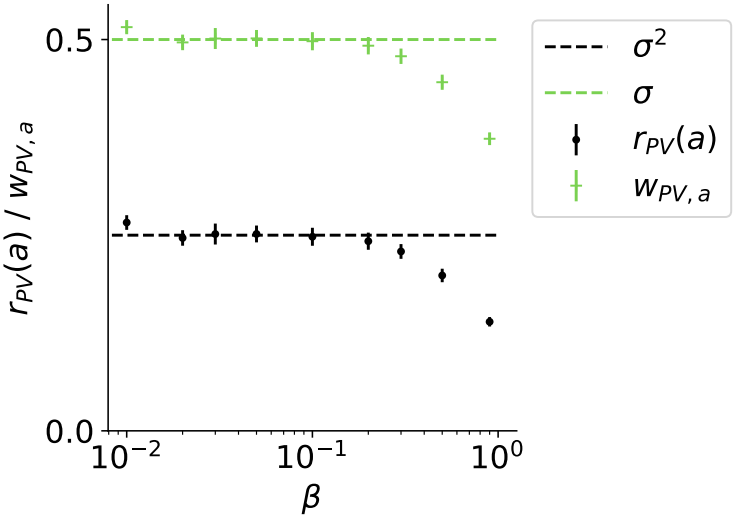

Figure 8: This figure shows that $w_s = \sqrt{\frac{2-\beta}{\beta}}$ is a good approximation for small β. For small values of β, the weight from a to PV approaches σ and the PV firing rate approaches σ^2.

Estimating the variance correctly The PVs estimate the variance of the whisker stimuli from the variance of the teaching input $(s - \mu)$, which nudges the membrane potential of the PVs with a nudging factor β. The nudging factor reduces the effective variance of the teaching input, such that in order to correctly estimate the variance, this reduction

needs to be compensated by larger weights from the SSTs to the PVs ($w_{\mathrm{PV,SST}}$) and from the whisker stimuli to the PVs ($w_{\mathrm{PV},s}$). To determine how strong the weights $w_s = w_{\mathrm{PV,SST}} = w_{\mathrm{PV},s}$ need to be to compensate for the downscaling of the input variance by β, we require that $\langle wa \rangle^2 = \sigma^2$ when the average weight change $\langle \Delta w \rangle = 0$. The learning rule for w is as follows:

$$\begin{aligned} \Delta w &= \eta[r_{\mathrm{PV}} - \phi(wa)]a && (24) \\ &= \eta[\phi((1-\beta)wa + \beta w_s \tilde{s}) - \phi(wa)]a && (25) \\ & && (26) \end{aligned}$$

where $r_{PV} = \phi((1-\beta)wa + \beta w_s \tilde{s})$ and $\tilde{s} = (s - \mu) \sim \mathcal{N}(0, \sigma)$.

Using that $\phi(u) = u^2$, the average weight change becomes:

$$\begin{aligned} \langle \Delta w \rangle &= \langle (1-\beta)^2 w^2 a^2 + \beta^2 w_s^2 \tilde{s}^2 \rangle + 2(1-\beta) wa \beta w_s \tilde{s} - w^2 a^2 \rangle a && (27) \\ &= \langle -2\beta w^2 a^2 + \beta^2 w^2 a^2 + \beta^2 w_s^2 \tilde{s}^2 \rangle a && (28) \\ &= \langle w^2 \rangle a^2 (\beta - 2) + \beta w_s^2 \sigma^2 && (29) \\ & && (30) \end{aligned}$$

Given our objective $\langle wa \rangle^2 = \sigma^2$, we can write:

$$\begin{aligned} \langle \Delta w \rangle &= (\beta - 2 + \beta w_s^2) && (31) \\ 0 &= (\beta - 2 + \beta w_s^2) && (32) \\ w_s &= \sqrt{\frac{2-\beta}{\beta}} && (33) \end{aligned}$$

Here, we assumed that $\phi(u) = u^2$ instead of $\phi(u) = \lfloor u \rfloor^2$. To test how well this approximation holds, we simulated the circuit for different values of β and hence w_s, and plotted the PV firing rate given the sound input a $r_{\mathrm{PV}}(a)$ and the weight from a to PV $w_{\mathrm{PV},a}$ for different values of β (Fig. 8). This analysis shows that the approximation holds for small β up to a value of $\beta = 0.2$.

Supplementary Information

Supplementary Methods

Synaptic dynamics/plasticity rules

$$\Delta w_{\mathrm{PV}^+,\mathrm{SST}^+} = \eta_{\mathrm{PV}}(w_{\mathrm{PV}^+,s} r_{\mathrm{s}} - w_{\mathrm{PV}^+,\mathrm{SST}^+} r_{\mathrm{SST}^+}) r_{\mathrm{SST}^+} \tag{34}$$

$$\Delta w_{\mathrm{PV}^-,\mathrm{SST}^-} = \eta_{\mathrm{PV}}(w_{\mathrm{PV}^-,R} r_{\mathrm{R}} - w_{\mathrm{PV}^-,\mathrm{SST}^-} r_{\mathrm{SST}^-}) r_{\mathrm{SST}^-} \tag{35}$$

$$\Delta w_{\mathrm{UPE}^+,\mathrm{SST}^+} = \eta_{\mathrm{UPE}}(w_{\mathrm{UPE}^+,r_{\mathrm{s}}} r_{\mathrm{s}} - w_{\mathrm{UPE}^+,\mathrm{SST}^+} r_{\mathrm{SST}^+}) r_{\mathrm{SST}^+} \tag{36}$$

$$\Delta w_{\mathrm{UPE}^+,\mathrm{SST}^-} = \eta_{\mathrm{UPE}}(w_{\mathrm{UPE}^-,\mathrm{R}} r_{\mathrm{R}} - w_{\mathrm{UPE}^-,\mathrm{SST}^-} r_{\mathrm{SST}^-}) r_{\mathrm{SST}^-} \tag{37}$$

$$\tag{38}$$

Different choice of supralinear activation function for PV

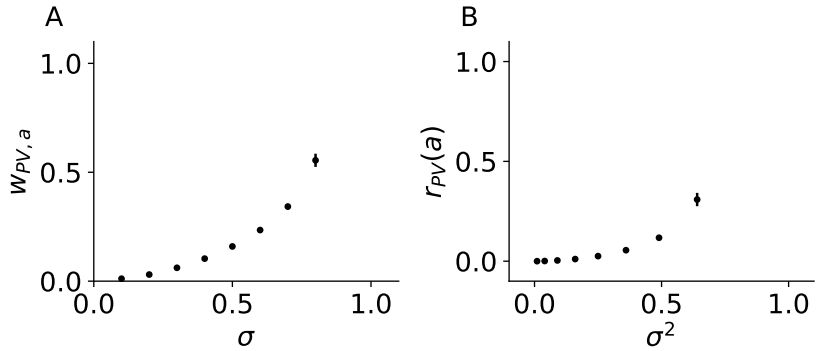

Figure 9: Learning the variance in the positive prediction error circuit with PVs with a power activation function (exponent = 3.0). The figure is analogous to Fig. 3G,H, and the circuit is the same except that the activation function of the PVs ($\phi_{PV}(x)$) has an exponent of 3.0 instead of 2.0.

Plastic weights from SST to PV learn to match weights from s to PV

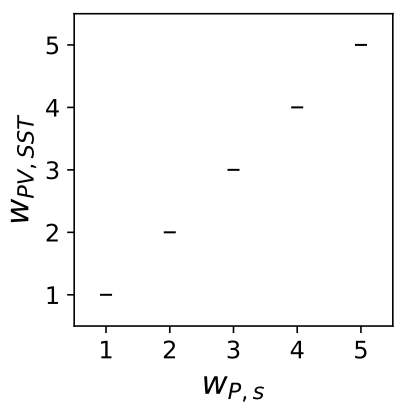

Figure 10: With inhibitory plasticity, weights from SST to PV can be learned. This figure shows that the weight from SST to PV ($w_{\mathrm{PV,SST}}$) is equal to the weight from s to PV ($w_{\mathrm{PV},s}$). The inhibitory plasticity rule is described in the supplementary methods.

PVs learn the variance in the negative prediction error circuit

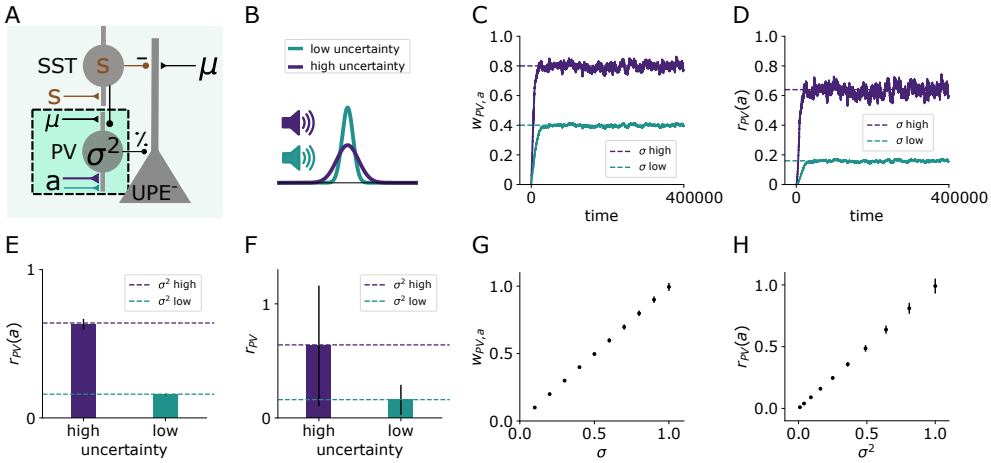

Figure 11: **PVs learn to represent the variance given an associative cue in the negative prediction error circuit.** A: Illustration of the changes in the negative prediction error circuit. Thicker lines denote stronger weights. B: Two different tones (purple, green) are associated with two somatosensory stimulus distributions with different variances (purple: high, green: low). C: Weights from sound a to PV over time for two different values of stimulus variance (high: $\sigma = 0.8$ (purple), low: $\sigma = 0.4$ (green)). D: PV firing rates over time given sound input (without stimulus input) for low (green) and high (purple) stimulus variance. E: PV firing rates (mean and std) given sound input for low and high stimulus variance. F: PV firing rates (mean and std) during sound and stimulus input. G: Weights from sound a to PV for different values of σ (mean and std from timestep 100000 to the end of the simulation). H: PV firing rates given sound input for different values of σ^2 (mean and std from timestep 100000 to the end of the simulation).

Learning SST to E weights

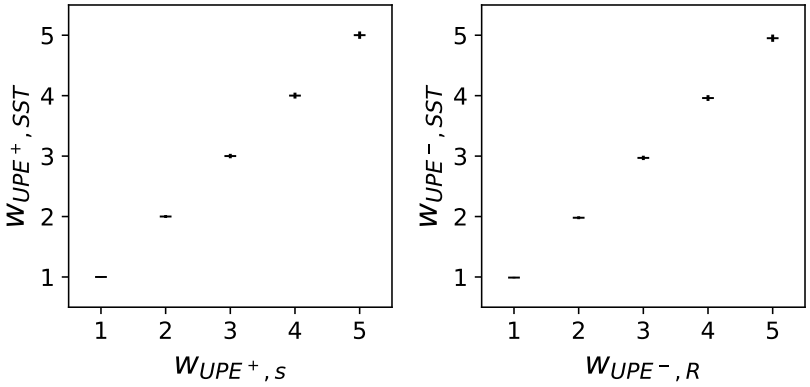

Figure 12: Learning the weights from the SSTs to the UPE neurons. This figure shows that the weights from the SSTs to the UPEs in both the positive (left) and the negative (right) prediction error circuit can be learned with inhibitory plasticity to match the weights from the stimulus representation s to the UPEs. The inhibitory plasticity rule is described in the supplementary methods.

PV activity is proportional to the variance in the recurrent circuit

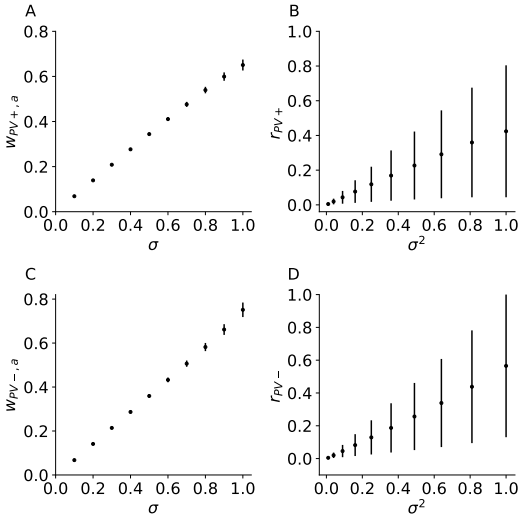

Figure 13: PV firing rates are proportional to the variance in the recurrent circuit model. Weights from a to PV as a function of σ in the positive (A) and negative (C) prediction error subcircuit. PV firing rates as a function of σ^2 in the positive (B) and negative (D) prediction error circuit.

Perturbation experiments

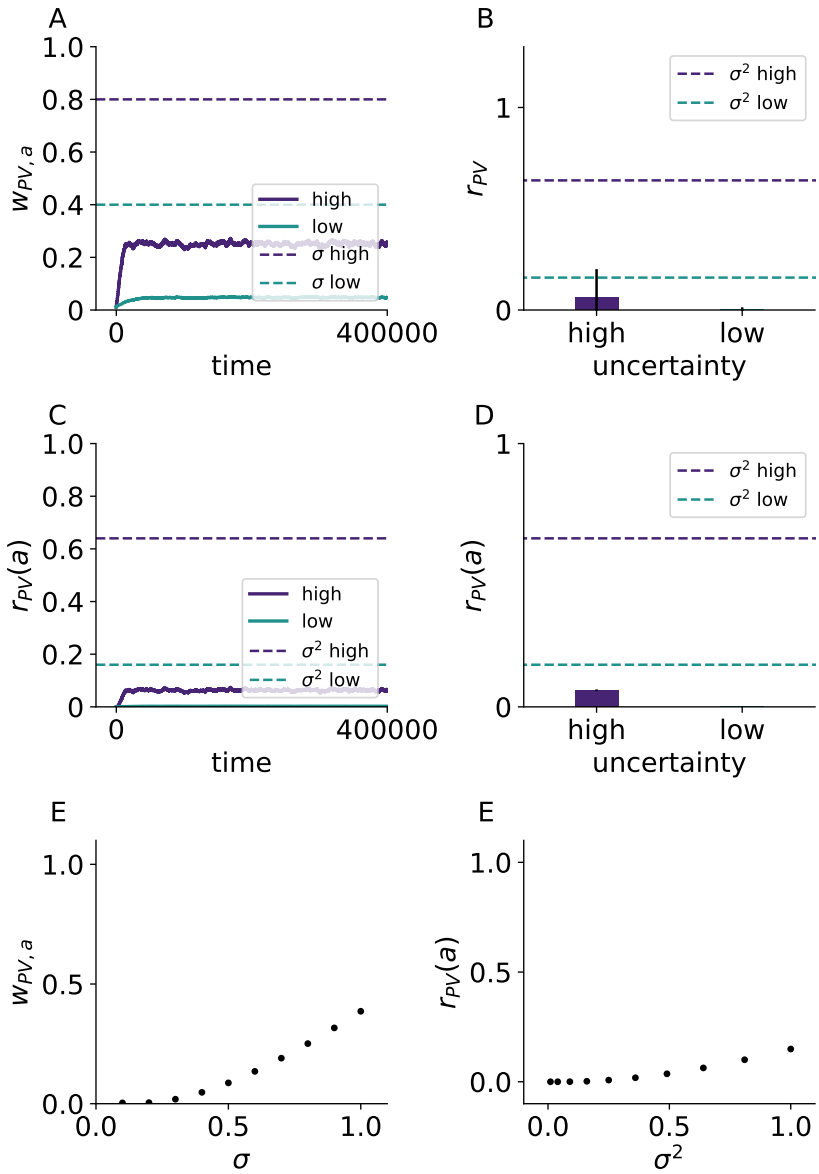

Figure 14: The panels are analogous to the panels in Fig. 3. Here, we used the same positive prediction error circuit as inf Fig. 3, and perturbed the SS-to-PV weight in the positive prediction error circuit by +10%.

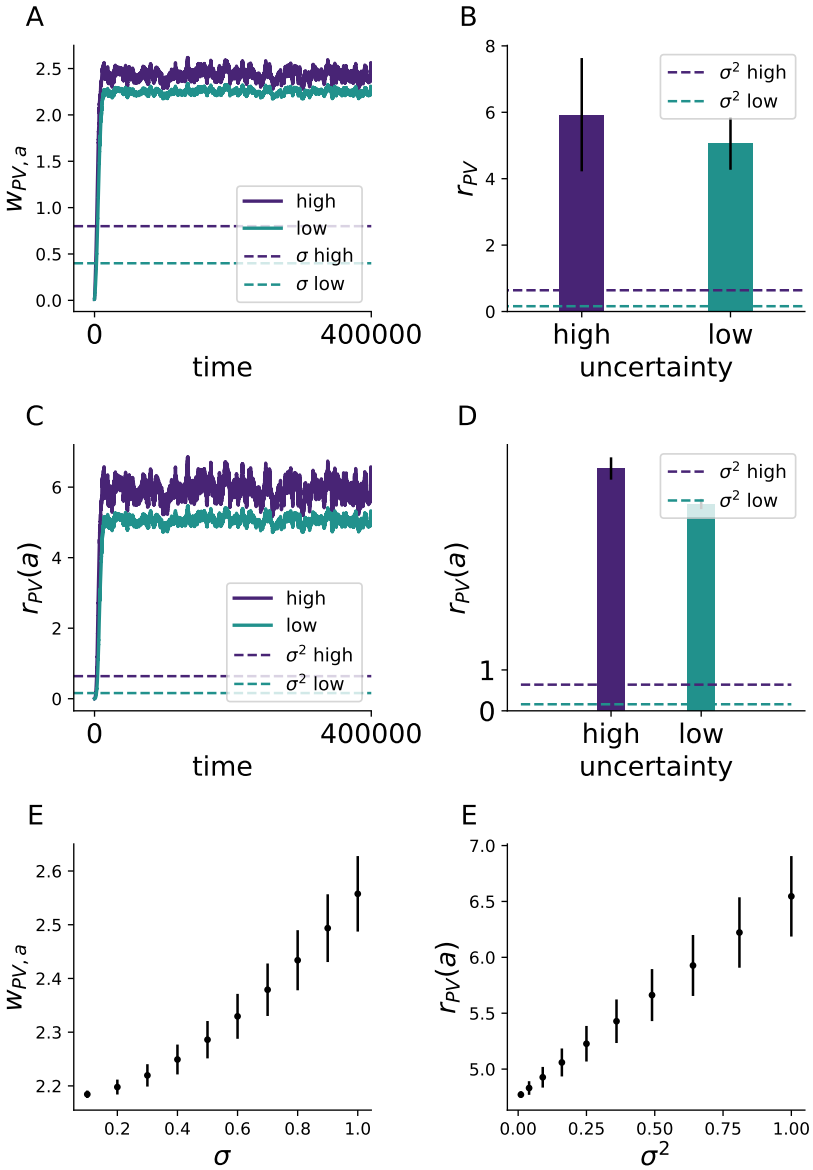

Figure 15: The panels are analogous to the panels in Fig. 3. Here, we used the same positive prediction error circuit as inf Fig. 3, and perturbed the SS-to-PV weight in the positive prediction error circuit by -10%.

.

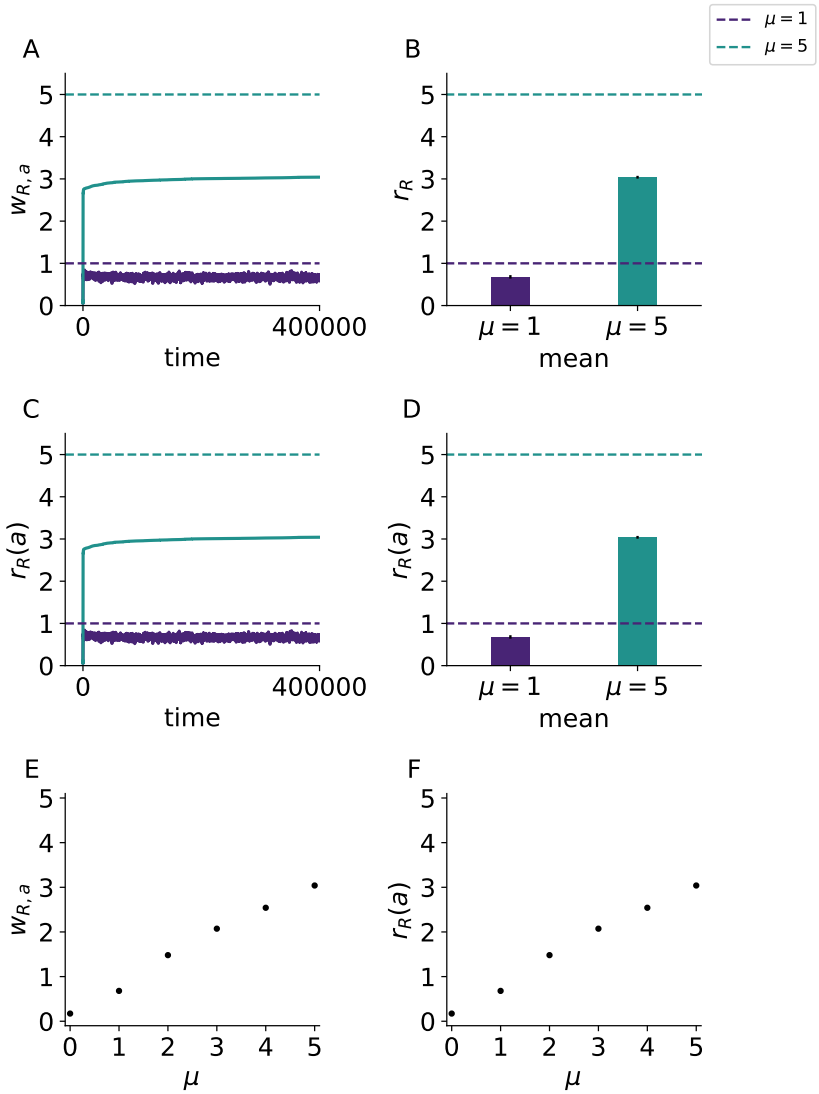

Figure 16: **Perturbation of SST$^+$ to UPE$^+$ and from SST$^-$ to UPE$^-$ weights by +100%.** Here we used the recurrent circuit model from Fig. 5 and perturbed the weights from SST$^+$ to UPE$^+$ and SST$^-$ to UPE$^-$ by +100%. Two different tones (purple, green) are associated with two somatosensory stimulus distributions with different means (purple: small, green: large). A: Weights from sound a to R over time for two different values of stimulus mean (small: $\mu = 1.0$ (purple), large: $\mu = 5.0$ (green)). B: R firing rates (mean and std) given sound input and whisker input for small and large stimulus means. C: R firing rates over time given sound input (without whisker stimulus input) for small (purple) and large (green) stimulus means. D: R firing rates (mean and std) given sound input (without whisker stimulus input). E: Weights from sound a to R for different values of μ (mean and std from timestep 100000 to the end of the simulation). H: R firing rates given sound input (without whisker stimulus input) for different values of μ^2 (mean and std from timestep 100000 to the end of the simulation).

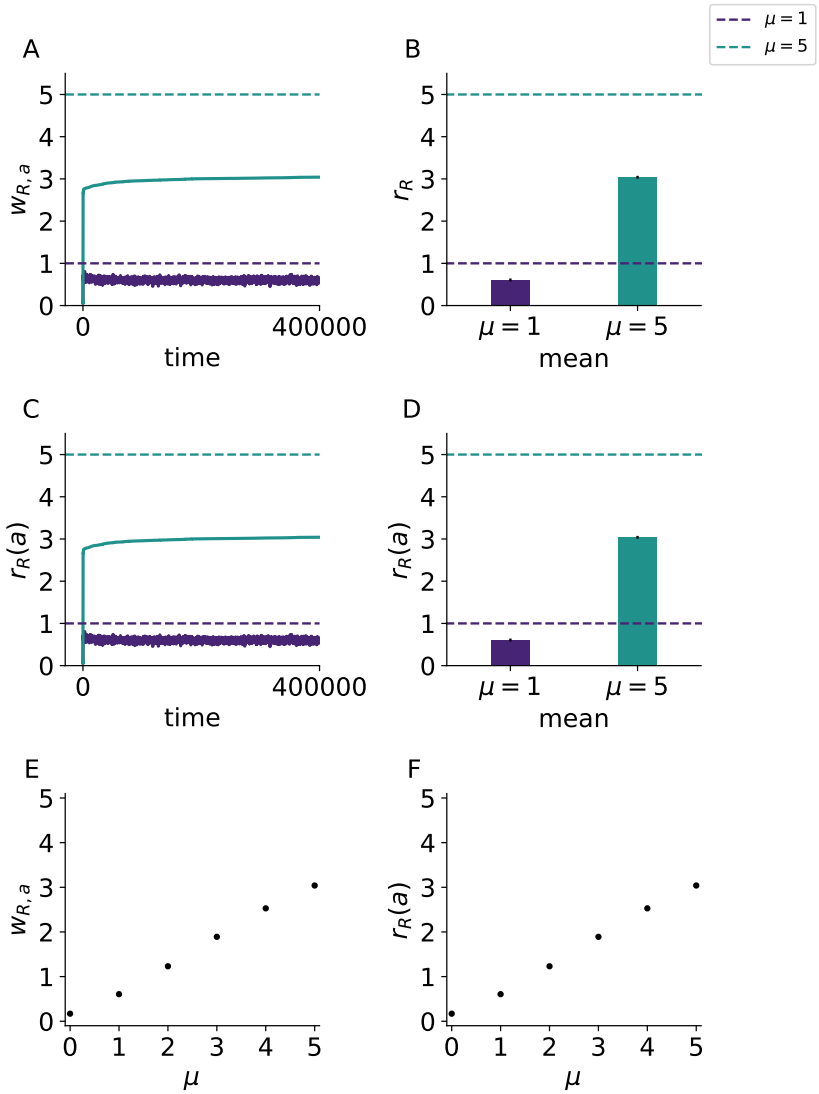

Figure 17: **Perturbation of SST$^+$ to UPE$^+$ weights by +100%.** Here we used the recurrent circuit model from Fig. 5 and perturbed the weights from SST$^+$ to UPE$^+$ by +100%. Two different tones (purple, green) are associated with two somatosensory stimulus distributions with different means (purple: small, green: large). A: Weights from sound a to R over time for two different values of stimulus mean (small: $\mu = 1.0$ (purple), large: $\mu = 5.0$ (green)). B: R firing rates (mean and std) given sound input and whisker input for small and large stimulus means. C: R firing rates over time given sound input (without whisker stimulus input) for small (purple) and large (green) stimulus means. D: R firing rates (mean and std) given sound input (without whisker stimulus input). E: Weights from sound a to R for different values of μ (mean and std from timestep 100000 to the end of the simulation). H: R firing rates given sound input (without whisker stimulus input) for different values of μ^2 (mean and std from timestep 100000 to the end of the simulation).

.

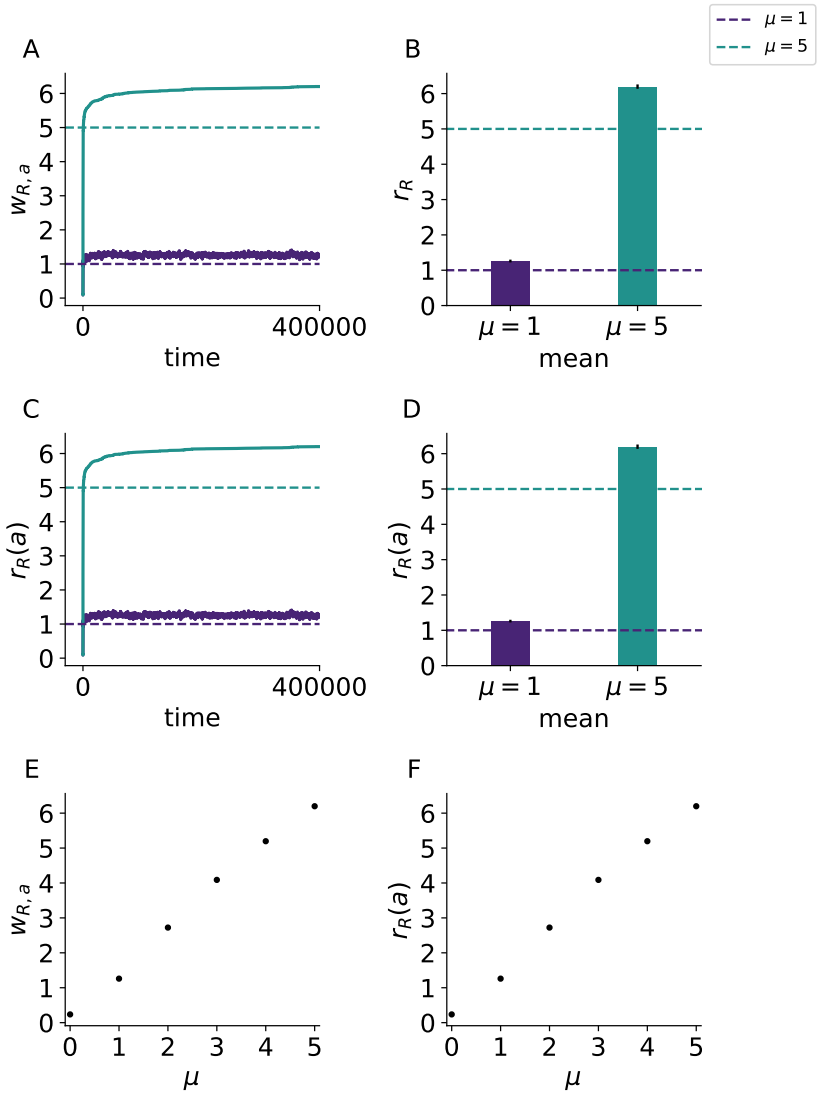

Figure 18: **Perturbation of SST$^-$ to UPE$^-$ weights by +100%.** Here we used the recurrent circuit model from Fig. 5 and perturbed the weights from SST$^-$ to UPE$^-$ weights by +100%. Two different tones (purple, green) are associated with two somatosensory stimulus distributions with different means (purple: small, green: large). A: Weights from sound a to R over time for two different values of stimulus mean (small: $\mu = 1.0$ (purple), large: $\mu = 5.0$ (green)). B: R firing rates (mean and std) given sound input and whisker input for small and large stimulus means. C: R firing rates over time given sound input (without whisker stimulus input) for small (purple) and large (green) stimulus means. D: R firing rates (mean and std) given sound input (without whisker stimulus input). E: Weights from sound a to R for different values of μ (mean and std from timestep 100000 to the end of the simulation). H: R firing rates given sound input (without whisker stimulus input) for different values of μ^2 (mean and std from timestep 100000 to the end of the simulation).

.

References

1. A. Attinger, B. Wang, G. B. Keller, *Cell* **169**, 1291–1302.e14, (https://doi.org/10.1016/j.cell.2017.05.023) (2017).
2. A. Ayaz *et al.*, *Nature Communications* **10**, 2585, (https://doi.org/10.1038/s41467-019-10564-8) (2019).
3. J. Cannon, A. M. O'Brien, L. Bungert, P. Sinha, eng, *Autism Res* **14** (2021).
4. M. X. Cohen, **Wilmes, K.**, I. v. d. Vijver, eng, *Trends Cogn Sci* **15**, 558–566 (2011).
5. J. H. Cornford *et al.*, *eLife* **8**, ed. by M. Bartos, G. L. Westbrook, C.-C. Lien, J. C. Poncer, e49872, (https://doi.org/10.7554/eLife.49872) (2019).
6. S. J. Eliades, X. Wang, *Nature* **453**, 1102–1106, (https://doi.org/10.1038/nature06910) (2008).
7. M. O. Ernst, M. S. Banks, *Nature* **415**, 429–433, (https://doi.org/10.1038/415429a) (2002).
8. C. R. Fetsch, A. Pouget, G. C. DeAngelis, D. E. Angelaki, *Nature Neuroscience* **15**, 146–154, (https://doi.org/10.1038/nn.2983) (2012).
9. A. Fiser *et al.*, *Nature Neuroscience* **19**, 1658 EP, (https://doi.org/10.1038/nn.4385) (Sept. 2016).
10. C. J. Gillon *et al.*, *bioRxiv*, eprint: https://www.biorxiv.org/content/early/2021/01/16/2021.01.15.426915.full.pdf, (https://www.biorxiv.org/content/early/2021/01/16/2021.01.15.426915) (2021).
11. J. Goris *et al.*, eng, *Biol Psychiatry Cogn Neurosci Neuroimaging* **3**, 667–674 (2018).
12. L. Hertäg, C. Clopath, eng, *Proc Natl Acad Sci U S A* **119**, e2115699119 (2022).
13. L. Hertäg, H. Sprekeler, *eLife* **9**, ed. by S. Ostojic, R. B. Ivry, S. Ostojic, e57541, (https://doi.org/10.7554/eLife.57541) (2020).
14. J. Jordan, J. Sacramento, W. A. M. Wybo, M. A. Petrovici, W. Senn, *Learning Bayes-optimal dendritic opinion pooling*, 2022, arXiv: 2104.13238 [q-bio.NC].
15. R. Jordan, G. B. Keller, *Neuron* **108**, 1194–1206.e5, (https://www.sciencedirect.com/science/article/pii/S0896627320307480) (2020).
16. G. B. Keller, T. Bonhoeffer, M. Hübener, *Neuron* **74**, 809 –815, (http://www.sciencedirect.com/science/article/pii/S0896627312003844) (2012).
17. G. B. Keller, R. H. R. Hahnloser, *Nature* **457**, 187–190, (https://doi.org/10.1038/nature07467) (2009).
18. G. B. Keller, T. D. Mrsic-Flogel, *Neuron* **100**, 424 –435 (2018).
19. K. Kveraga, A. S. Ghuman, M. Bar, *Brain and cognition* **65**, 145–168, (https://pubmed.ncbi.nlm.nih.gov/17923222) (Nov. 2007).
20. S. Lee, I. Kruglikov, Z. J. Huang, G. Fishell, B. Rudy, *Nature Neuroscience* **16**, 1662–1670 (2013).
21. E. Payzan-LeNestour, P. Bossaerts, *PLOS Computational Biology* **7**, 1–14, (https://doi.org/10.1371/journal.pcbi.1001048) (Jan. 2011).
22. H.-J. Pi *et al.*, *Nature* **503**, 521–524 (2013).
23. A. Prönneke *et al.*, *Cerebral Cortex* **25**, 4854–4868 (2015).
24. R. P. N. Rao, D. H. Ballard, *Nature Neuroscience* **2**, 79–87, (https://doi.org/10.1038/4580) (1999).
25. S. Sachidhanandam, B. S. Sermet, C. C. Petersen, *Cell Reports* **15**, 700–706, (https://www.sciencedirect.com/science/article/pii/S2211124716303345) (2016).
26. R. Urbanczik, W. Senn, *Neuron* **81**, 521–528, (https://www.sciencedirect.com/science/article/pii/S0896627313011276) (2014).
27. A. R. Walker, D. Luque, M. E. Le Pelley, T. Beesley, *Psychonomic Bulletin & Review* **26**, 1911–1916, (https://doi.org/10.3758/s13423-019-01653-2) (2019).
28. N. R. Wilson, C. A. Runyan, F. L. Wang, M. Sur, eng, *Nature* **488**, 343–348 (2012).
29. P. Zmarz, G. B. Keller, eng, *Neuron* **92**, 766–772 (2016).