Synapses with short-term plasticity are optimal estimators of presynaptic membrane potentials

Jean-Pascal Pfister, Peter Dayan, Mate Lengyel

Synapses with short-term plasticity are optimal estimators of presynaptic membrane potentials

Jean-Pascal Pfister[1*], Peter Dayan[2], Máté Lengyel[1]

[1] Computational and Biological Learning Lab, Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK

[2] Gatsby Computational Neuroscience Unit, University College London, 17 Queen Square, London WC1N 3AR, UK

* corresponding author: jean-pascal.pfister@eng.cam.ac.uk

Abstract

The trajectory of the somatic membrane potential of a cortical neuron exactly reflects the computations performed on its afferent inputs. However, the spikes of such a neuron are a very low dimensional and discrete projection of this continually evolving signal. We explore the possibility that the neuron's efferent synapses perform the critical computational step of estimating the membrane potential trajectory from the spikes. We show that short-term changes in synaptic efficacy can be interpreted as implementing an optimal such estimator. Short-term depression arises when pre-synaptic spiking is sufficiently intense as to reduce the uncertainty associated with the estimate; short-term facilitation reflects structural features of the statistics of the pre-synaptic neuron such as up and down states. Our analysis provides a unifying account of a powerful, but puzzling form of plasticity.

Synaptic efficacies can increase (facilitate) or decrease (depress) several-fold in strength on the time scale of single inter-spike intervals[1-3]. This short-term plasticity (STP), which is well captured by simple, but powerful, mechanistic models[1,3,4], is of a regularity and magnitude that argues against its being treated only as wanton variability[5]. There have thus been various important suggestions for the functional significance of STP, including, just to name a few: low-, high-, or band-pass filtering of inputs[3,6] (but see Ref. 7), rendering postsynaptic responses insensitive to the absolute intensity of presynaptic activity[8,9] decorrelating input spike sequences[10], and maintaining working memories in the prefrontal cortex[11].

However, despite the ubiquity of STP in cortical circuits[2], these suggestions are restricted to select neural subsystems[9,11] or forms of STP[5,8-10], and are often limited to feed-forward networks[8-10], or to a firing rate-based description of presynaptic activities[8] thereby

ignoring the fundamentally fast fluctuations in synaptic efficacies due to STP. Worse, the vast bulk of models of neural circuit information processing require synaptic efficacies to be constant over the short-term of a single computation, changing at most very slowly to average across the statistics of input, or changing only in the light of a gating mechanism[12,13]. These would seem to be incompatible with substantial STP. Here, we argue that, far from hindering such circuit computations, STP is in fact a near-optimal solution to a central problem neural circuits face associated with spike-based communication.

Although, as digital quantities, spikes have the mechanistic advantage of allowing regenerative error-correction, they are a significantly impoverished representation of the fast-evolving, analog, membrane potentials of the neurons concerned[14-16]. These analog quantities are normally considered to lie at the heart of computations[17,18], and so it is common to appeal to averages over space (i.e., multiple identical neurons) and/or time (i.e., slow currents) to allow them to be represented by spike trains[18]. However, both sorts of averages are neurobiologically questionable. In many circumstances computations need to be executed in the matter of a few interspike intervals[19-21], precluding extensive averaging over time; and in many circuits neurons represent independent analog quantities, as in recurrent network models of autoassociative memories[22,23], or partially independent quantities, as in surface attractor models of population codes[24]. We make the alternative suggestion that the analog membrane potential of a neuron is being estimated in a statistically appropriate manner by its efferent synapses based on the spike trains that the neuron emits, and that STP is a signature of this solution.

In particular, the informativeness of an incoming spike about the membrane potential varies greatly depending on the uncertainty left by the preceding spike train. This makes the spike's impact significantly context-dependent. We demonstrate that key elements of this context-dependency are realized by synaptic depression and facilitation. Further, since incoming spikes are sparse, the behavior of the optimal estimator critically depends on prior assumptions about presynaptic membrane potential dynamics. Thus our approach allows us to make detailed predictions about how the properties of STP, implementing the optimal estimator, should be matched to the statistics of presynaptic membrane potential fluctuations.

Results

Postsynaptic potentials and the optimal estimator

We first define the optimal estimator of the continuously varying membrane potential u of a presynaptic cell from its past spikes. We then show that it depends on these spikes in the same way as a particular measure of its efferent synapses' contributions to their postsynaptic membrane potentials. Since spikes are discrete, they cannot support recovery of u with absolute certainty, and so the full solution to the estimation task is a posterior probability distribution[20,25-29] $\mathrm{P}\left(u_t \mid s_{0:t}\right)$ over the possible values that the

presynaptic membrane potential at time t, u_t, might take based on all the spikes observed so far, $s_{0:t}$. The mean of this posterior is then the estimator $\hat{u}_t$ that minimizes the squared error[25]. We interpret the local postsynaptic potential at an excitatory synapse v_t as representing this optimal estimate. This local potential is loosely defined as the sum of all EPSPs at this synapse (see **Supplementary Note** for a more precise definition); a filtered version of it is recorded in standard experiments into STP.

To be correct, the estimator must implicitly incorporate a statistically appropriate model of membrane potential fluctuations and spike generation in the presynaptic neuron[25]. For the latter, we adopt the common characterization that a spike is created stochastically whenever u_t exceeds a (soft) threshold[30,31] (**Fig. 1**; see also **Supplementary Fig. 2**). In this case, the occurrence of a spike implies that the membrane potential is likely to be high; the absence of spikes implies the membrane potential is likely to be low. Thus $\hat{u}_t$ should increase following a spike, and decrease during interspike intervals (**Fig. 1a**, **inset**; **Online Methods**). The decrease should be gradual since, even at high membrane potential values, the firing rate of a neuron is limited, and so the absence of a spike is relatively weak evidence that the membrane potential is low. As required by our interpretation, the local post-synaptic potential, v_t, at an excitatory synapse shows the same qualitative characteristics: it increases suddenly at the times of presynaptic spikes, and decays gradually towards a lower baseline between transmission events (**Fig. 1a**, **inset**).

However, this observation is only approximate. As we have mentioned, under STP, the actual size of an excitatory postsynaptic potential (EPSP) depends on the past history of spiking. Such history-dependence is also a hallmark of optimal estimation, because the evidence supplied by a spike needs to be evaluated in the context of the current state of the estimator, i.e., the current posterior[25]. The posterior is computed on-line in a recursive manner, according to standard Bayesian filtering. In this, likelihood information from the current presence or absence of a spike, s_t, $\mathrm{P}(s_t \mid u_t)$, is combined with the posterior computed in the previous timestep, $\mathrm{P}(u_{t-\delta t} \mid s_{0:t-\delta t})$ (**Online Methods**):

$$\mathrm{P}(u_t \mid s_{0:t}) \propto \mathrm{P}(s_t \mid u_t) \int_{-\infty}^{\infty} \mathrm{P}(u_t \mid u_{t-\delta t}) \mathrm{P}(u_{t-\delta t} \mid s_{0:t-\delta t}) du_{t-\delta t} \qquad (1)$$

Thus the impact of an incoming spike on the mean of the posterior, $\hat{u}_t$, will be context-dependent, varying as a function of the posterior propagated from the past.

The precise nature of the context-dependence of the changes in $\hat{u}_t$ will depend on the particular statistical model assumed for the dynamics of u_t, $\mathrm{P}(u_t \mid u_{t-\delta t})$. Below, we consider two increasingly complex, models: one in which u_t performs a random walk around a fixed `resting' (or baseline) membrane potential, u^{rest}; and one in which u_t^{rest} itself also changes in time. This allows us to explore two fundamental factors contributing to the context-dependence of $\hat{u}_t$ that correspond respectively to the effects of synaptic depression and facilitation on v_t.

Synaptic depression and uncertainty

After eliminating the influence of spikes themselves, the simplest approximation to the statistics of the membrane potential u_t of the presynaptic neuron is as an Ornstein-Uhlenbeck (OU) process[32,33]. For this, the total input to the presynaptic cell is assumed to be Gaussian white noise[33], and is subject to leaky integration, decaying towards its resting value, u^{rest} (**Fig. 1a**; **Online Methods** and **Supplementary Fig. 2**). In this case, the posterior distribution $\mathrm{P}(u_t \mid s_{0:t})$ can be well characterized as a Gaussian, with mean $\hat{u}_t$, and variance σ_t^2, which expresses the estimator's current uncertainty about u_t^{rest}. Given the assumption about the way it is generated, observation of a spike provides evidence that u_t is high. However, the quantitative effect on raising $\hat{u}_t$ depends on the uncertainty σ_t. The less the uncertainty, i.e., the better known is u_t already, the less the estimate should be influenced by a spike (**Fig. 1b**) and the lower the synapse's apparent efficacy. The uncertainty is determined by the evidence from past spikes; and thus we should expect the magnitude of the EPSPs to fluctuate according to this history.

Uncertainty in the optimal estimate increases during interspike intervals (**Fig. 1c**), because the absence of a spike is only weak evidence for a low membrane potential. Therefore, spikes that arrive after a longer period of silence should increase $\hat{u}_t$ by more than spikes arriving in quick succession. This closely resembles, in fact all the way down to fine quantitative details, the effect of synaptic depression in a biophysically motivated canonical STP model[5,34] (**Fig. 1**). In that model, depression is mediated by the depletion of a synaptic resource variable x_t which thus behaves as the biophysical analog of estimator uncertainty, σ_t^2.

In fact, for a paired-pulse protocol, the dynamical equations describing the time-evolution of the optimal estimator and its uncertainty, $\hat{u}_t$ and σ_t^2, are formally equivalent to those of the biophysical STP model describing the time-evolution of the postsynaptic membrane potential and synaptic resource variable, v_t and x_t, respectively (**Online Methods**). As a result, higher presynaptic firing rates lead to diminishing postsynaptic responses both in the optimal estimator and in the STP model (**Fig. 3a**).

Synaptic facilitation and "up" state probability

The assumption that the presynaptic membrane potential follows an OU process is often too simplistic[35]. Fortunately, it is straightforward in our framework to incorporate other statistical properties of membrane potential fluctuations and study their effects on the features of STP. For instance, many cortical cells show phasic activation patterns whereby their membrane potential alternates between a resting and a depolarized state, such as "up" and "down" states in the cortex[36,37], or out-of-place-field and within-place-field activity for hippocampal place cells[38]. This can be captured by extending the model for the dynamics of u_t to allow the resting membrane potential u_t^{rest} itself to switch

between two possible values, corresponding to an "up" and a "down" state (**Fig. 2** ; see **Supplementary Note** and **Supplementary Fig. 2**).

In this extended model, the true current value of u_t^{rest} is unknown to the postsynaptic cell, just like u_t. Thus, the full solution to the estimation problem is a posterior distribution expressing joint uncertainty about these two quantities (**Supplementary Note**). In this case, the estimated probability that u_t^{rest} is in its "up" state also reflects recent spikes, and influences $\hat{u}_t$ (ρ_t, **Fig. 2a–c**) . Observing a spike when the current estimate of the membrane potential is compatible with the presynaptic neuron being in its "down" state increases this probability somewhat. Observing then a second spike in a short time window, providing substantial extra evidence for this "up" state, will thus cause a larger increment in $\hat{u}_t$ (**Fig. 2a**, **inset**). EPSPs in the facilitating biophysical STP model and in actual facilitating synapses in a paired-pulse protocol[39] (**Figs. 2d** and **3b**) show the same effect.

Match between STP and the dynamics of the presynaptic cell

Our theory requires the synaptic estimator to be matched to the statistical properties of the presynaptic membrane potential fluctuations which it needs to estimate. Thus, fitting the optimal estimator to experimentally measured STP data (as shown in **Fig. 3a–b**) allowed us to predict, without further parameter fitting, properties of the membrane potential dynamics that the corresponding presynaptic cell type should exhibit. Testing these predictions is challenging because they are about the natural statistics of membrane potential fluctuations and thus require *in vivo* intracellular recordings, preferably from behaving rather than anaesthetized animals.

Nevertheless, starting from our fits to data about synaptic depression and facilitation for cerebellar climbing fibers (**Fig. 3a**) and hippocampal Schäffer collateral inputs (**Fig. 3b**), we predict membrane potential fluctuations in inferior olive neurons (**Fig. 3e**) and hippocampal pyramidal cells (**Fig. 3f**) respectively, that are in broad qualitative agreement with those found *in vivo* in these cell types[38,40] (**Fig. 3c**, **d**). Note that the absence and presence of marked "up" and "down" states in these two cell types respectively is predicted by our theory directly from the absence and presence of facilitation in their corresponding efferent synapses.

The advantage of STP

In order to quantify the advantage that STP brings to synapses for tracking the presynaptic membrane potential, we measured how well $\hat{u}_t$, or its biophysical analog, v_t, performs on the estimation task in terms of the time-averaged squared error between u_t and its estimate (**Online Methods**). We compared the optimal estimator of the presynaptic membrane potential (**Figure 4a**, red line) with the postsynaptic membrane potential occasioned by a synapse undergoing STP (blue line). These two traces are very close. Importantly, a static synapse without STP, whose fixed efficacy is still optimized

for the same estimation task, performs significantly less well (green line).

Our account of synaptic dynamics assumes no transmission failures, or, equivalently, a large number of release sites between the pre- and postsynaptic cells. Therefore, we also ran simulations with stochastic synapses[4] (**Online Methods**) to test a more realistic regime of synaptic communication (**Fig. 4b**). The advantage of a dynamic over a static synapse was maintained when the number of release sites was low even in the face of transmission failures and the added variability of the dynamic synapse entailed by the stochastic restocking of vesicles.

Discussion

We showed that STP arranges for the local postsynaptic membrane potential at a synapse, v_t, to behave as an optimal estimator of the presynaptic membrane potential, u_t. We argued that this is central for a wide swathe of feedforward and recurrent neural circuits. In particular, it allows network computations based on analog quantities encoded in the somatic membrane potentials of neurons to be realized, even though their spikes offer only a low dimensional and discrete projection of those potentials.

As a first step, we concentrated on the interaction between a pair of cells, and on reproducing as close an estimate as possible to the exact presynaptic membrane potential on the postsynaptic side. Of course, v_t can be changed by many factors other than synaptic currents, for example voltage signals propagating from other dendritic compartments or indeed backpropagating action potentials from the soma. However, under most experimental paradigms used to test STP[41,42], the magnitudes of these other factors are minimized, such that the somatic membrane potential recorded reflects, perhaps with some dendritic filtering, the local EPSPs at the stimulated synapse. This allows our theory to be directly applied to data obtained under such conditions (see **Supplementary Note**).

A significant apparent challenge to our theory, as indeed to many previous functional accounts of STP[8-11], is that different efferent synapses of the same cell can express different forms of STP[1]. Two particular classes of factors that can affect v_t may account for this. Both factors have to do with the overall computation performed by the neuron in the network, which need not strictly factorize into individual estimation of each presynaptic membrane potential and combination of these across multiple synapses (see **Supplementary Fig. S1**).

First, computations might be based on estimates of different functions of the presynaptic membrane potentials than just their mean value, for instance their rates of change or higher order temporal derivatives. In this case, the different efferent synapses of a single neuron might estimate different functions, and therefore be different. It will be an interesting extension of our theory to study how synaptic estimation and single-neuron computation can be blended, rather than being performed as separate algorithmic steps.

Second, the long-run efficacies of synapses in a neural circuit also play a key part in the computations it performs[18,43], suggesting that presynaptic membrane potentials scaled by these computational factors should be reproduced instead. In this case, the optimal estimator just scales too. However, if the synapse suffers from saturation or other similar nonlinear constraints, then the form of STP that minimizes the error in the estimate of the scaled presynaptic potential would have to reflect these non-linearities, and change as the long-run weight alters. These changes could take potentially quite complex forms. Since different efferent synapses have different long-term weights, they could exhibit different forms of STP. Indeed, metaplasticity of STP due to long-term potentiation has been observed experimentally[41,44].

We should also note that differences in STP across efferent synapses have been predominantly shown for inhibitory interneurons[1,45,46]. In our theory, estimation of the membrane potential comes in the service of network computations, which are typically posed in terms of the excitatory principal cells rather than inhibitory interneurons. Therefore, the theory does not fully extend to cover inhibitory neurons.

Our theory employs a standard account of the relationship between u_t and the presynaptic spikes, and so is dual to the suggestion[47] that generation is more complex so that estimation can be straightforward even with a static synapse. The latter account is not easy to reconcile with the fluctuations evident in STP. Others have pursued ideas more similar to ours about adaptive gain control mechanisms in neurons generally acting as optimal filters[48], and dynamic synapses specifically acting as estimators of presynaptic firing rates[15] (see also B. Cronin, M. Sur & K. Koerding. *Soc. Neurosci. Abstr.*, 663.6, 2007) or interspike intervals[15]. Some of these studies do not encompass STP; others address depression but not facilitation. Finally, those studies aim mainly at predicting a general advantage (if one exists at all[7]) of dynamic synapses over static ones.

Our approach is unique in suggesting that synaptic dynamics are matched to the statistics of presynaptic membrane potential fluctuations which we were able to demonstrate at least qualitatively (**Fig. 3c–f**). Even such a qualitative match is noteworthy, given that the STP data we fitted were not extensive, were recorded *in vitro* under potentially very different stimulation regimes and neuromodulatory milieux for synaptic dynamics than those relevant *in vivo*, and given our highly simplified statistical model of presynaptic membrane potential dynamics. For example, a common dynamical motif shared by many neurons, including olivary neurons and hippocampal pyramids, and ignored by our model, is the presence of subthreshold membrane potential oscillations[38,40].

The simplicity of our model of presynaptic membrane potential dynamics makes it hard to provide a direct biological interpretation of the optimal parameter values (**Online Methods**) from the fits to data. Further theoretical work would be necessary to incorporate higher-order statistical regularities of these dynamics into the model. Further empirical studies recording STP under more naturalistic conditions and *in vivo* membrane potential recordings from the same identified pair of neurons, or at least the same cell type, will also be required. These would jointly license more direct comparisons and interpretations. Such experiments would of couse be technically challenging. However,

the link to optimal estimation offered by our theory provides them with the potential to test directly an important facet of neural circuit computations.

Acknowledgements

We thank M. Häusser and R. Brown for useful references, and L. Abbott, Sz. Káli and members of the Budapest Computational Neuroscience Forum for valuable discussions. This work was supported by the Wellcome Trust (J-P.P., M.L., P.D.) and the Gatsby Charitable Foundation (P.D.).

Author contributions

J.P.P. and M.L. developed the mathematical framework. J.P.P performed the numerical simulations. All authors wrote the manuscript.

Figures and legends

Figure 1. Estimating the presynaptic membrane potential from spiking information. (**a**) Sample trace (black) of the presynaptic membrane potential generated from an Ornstein-Uhlenbeck process. When the membrane potential exceeds a soft threshold, action potentials (vertical black lines) are generated. The optimal estimator of the presynaptic membrane potential (red line: mean estimate $\hat{u}_t$, red shading: one standard deviation σ_t) closely matches an optimally tuned canonical model of short-term plasticity[11] (blue). Inset shows a magnified section. (**b**) EPSP amplitude of the optimal estimator (red: mean $\pm$ s.t.d.) and of the canonical model of short-term plasticity (blue: mean $\pm$ s.t.d.) as a function of the estimator uncertainty σ^2. Note that EPSP amlitudes in the biophysical model tend to be smaller than those in the optimal estimator which is compensating for a somewhat slower decay in the biophysical model (see (**a**) inset). (**c**) The dynamics of the scaled uncertainty σ^2/σ^2_{max} (red) closely match the resource variable x_t of the canonical model of STP (blue), σ^2.

Figure 2. Estimating the presynaptic membrane potential when the resting membrane potential randomly switches between two different values. (**a**) Presynaptic subthreshold membrane potential with action potentials (black), its optimal mean estimate ($\hat{u}$, red line) with the associated standard deviation (σ, red shading), and the postsynaptic membrane potential in a model synapse[11] with optimally tuned short-term plasticity (blue line). Inset: the (scaled) optimal estimator (red solid line) strongly depends on the estimated probability ρ of being in the "up" state (red dot-dashed line). (**b**) EPSP amplitude in the optimal estimator depends on its uncertainty (x-axis, σ^2), and the change in the estimated probability that the presynaptic cell is in its "up" state (color code, $\Delta\rho$). (**c**) The estimated probability that the presynaptic cell is in its "up" state ρ (red) tracks the state of the presynaptic neuron (black) as it randomly switches between its "up" and "down" states. (**d**) EPSP magnitudes in the optimal estimator against EPSP magnitudes in the model synapse.

Figure 3. The optimal estimator reproduces experimentally observed patterns of synaptic depression and facilitation. (**a**) Synaptic depression in cerebellar climbing fibers (circles: mean $\pm$ s.e.m.; redrawn from Ref. 6) and in the model (solid line), measured as the ratio of the amplitude of the eighth and first EPSP as a function of the stimulation rate during a train of eight presynaptic spikes. (**b**) Synaptic facilitation in hippocampal Schäffer collaterals (circles: mean $\pm$ s.e.m.; redrawn from Ref. 39) and in the model (solid line), measured as the ratio of the amplitude of the second and first EPSP as a function of the interval between a pair of presynaptic spikes. Shading in (**a**) and (**b**) shows the robustness of the fits (**Online Methods**): model predictions when best-fit parameters are perturbed by 5% (dark gray) or 10% (light gray). (**c**) - (**f**) Comparison between in-vivo data and samples traces from the model. (**c**) *In vivo* intracellular recordings from the (anaesthetised) rat inferior olive nucleus (reproduced from Ref. 40). (**d**) *In vivo*

intracellular recordings from hippocampal pyramidal cells of the behaving rat (reproduced from Ref. 38). (**e**) Prediction of the model for the dynamics of the inferior olive neurons. Sample trace generated with parameters fitted to the data about STP in cerebellar climbing fibers shown in (**a**) . (**f**) Prediction of the model for the dynamics of hippocampal pyramidal neurons. Sample trace generated with parameters fitted to the data about STP in Schäffer collaterals shown in (**b**) .

Figure 4. Estimation performance of the presynaptic membrane potential. (**a**) Peformance as a function of the determinism of presynaptic spiking, β (**Online Methods**), in the optimal estimator (red), an optimally fitted dynamic synapse (blue), and an optimally fitted static synapse, without short-term plasticity (green). When $\beta = 0$ (entirely stochastic spiking), spikes are generated independently of the membrane potential, and as a consequence, all models fail to track the membrane potential. As β becomes larger (more deterministic spiking), the dynamic synapse model matches the optimal estimator in performance and significantly outperforms the static synapse. Realistic values for $\beta\sigma_{\mathrm{OU}}$ (here $\sigma_{\mathrm{OU}} = 1$ mV) have been found to be between 2 and 3 in L5 pyramidal cells of somatosensory cortex[49]. (**b**) Estimation performance with stochastic vesicle release as a function of the number of synaptic release sites, N. The dynamic synapse (blue) tracks the performance of the optimal estimator (red) well and outperforms the static synapse (green) at all values of N. The performance of all estimators decreases only when the number of independent release sites becomes very low ($N = 1$ or 2). When N is large ($N \to \infty$), synaptic transmission becomes deterministic, even though spike generation itself remains stochastic (with parameter $\beta = 2$ shown by the arrow in **a**).

Online Methods

The optimal estimator

Our goal is to understand the factors contributing to the mean of the posterior, or the estimate

$$\hat{u}_t = \int_{-\infty}^{\infty} u_t \cdot \mathrm{P}(u_t \mid s_{0:t}) du_t \tag{2}$$

corresponding to the postsynaptic potential under our interpretation, and in particular to the size of change in this estimate in response to an incoming spike, the analogue of the size of an EPSP.

The Ornstein-Uhlenbeck process

The generative model involves two simplifying assumptions. First we assume that presynaptic membrane potential dynamics are discrete time and Markovian (**Supplementary Fig. 1a**)

$$\mathrm{P}(u_t \mid u_{0:t-\delta t}) = \mathrm{P}(u_t \mid u_{t-\delta t}) \tag{3}$$

In particular, we assume that the presynaptic membrane potential evolves as an Ornstein-Uhlenbeck (OU) process, given (in discrete time steps of size δt, and thus as an AR(1) process) by

$$\mathrm{P}(u_t \mid u_{t-\delta t}) = \mathcal{N}\left[u_t;\ u_{t-\delta t} + \frac{1}{\tau}\left(u^{\mathrm{rest}} - u_{t-\delta t}\right)\delta t,\ \sigma_{\mathrm{W}}^2 \delta t\right] \tag{4}$$

where u^{rest} is the resting membrane potential of the presynaptic cell (assumed to be constant here), τ is its membrane time constant, and σ_{W} is the step size for the random walk behavior of its membrane potential. Because both τ and σ_{W} are assumed to be constant, the marginal variance of the presynaptic membrane potential, $\sigma_{\mathrm{OU}}^2 = \sigma_{\mathrm{W}}^2 \tau/2$, is stationary.

The second assumption is that spiking activity at any time only depends on the membrane potential at that time:

$$\mathrm{P}(s_t \mid u_{0:t}) = \mathrm{P}(s_t \mid u_t) \tag{5}$$

In particular, we assume that the spike generating mechanism is an inhomogeneous Poisson process. Thus, at time t, the neuron emits a spike ($s_t = 1$) with probability $g(u_t)\Delta t$, and therefore the spiking probability given the membrane potential can be written as:

$$\mathrm{P}(s_t \mid u_t) = \left[1 - g(u_t)\delta t\right]^{1-s_t}\left[g(u_t)\delta t\right]^{s_t} \tag{6}$$

We further assume that the transfer function is exponential (**Supplementary Fig. 1c**):

$$g(u_t) = g_0 e^{\beta u_t} \tag{7}$$

where g_0 and β are the base rate and determinism of the spike generation process,

respectively.

Since equations 3 and 5 define a hidden Markov model, the posterior distribution over u can be written in a recursive form:

$$\mathrm{P}(u_t \mid s_{0:t}) \propto \mathrm{P}(s_t \mid u_t) \int_{-\infty}^{\infty} \mathrm{P}(u_t \mid u_{t-\delta t}) \mathrm{P}(u_{t-\delta t} \mid s_{0:t-\delta t}) du_{t-\delta t} \tag{8}$$

That is, the posterior at time t, $\mathrm{P}(u_t \mid s_{0:t})$, can be computed by combining information from the current time step with the posterior obtained in the previous time step, $\mathrm{P}(u_{t-\delta t} \mid s_{0:t-\delta t})$. Note that even though inference can be performed recursively, and the hidden dynamics is linear-Gaussian (Eq. 4), the standard (extended) Kalman filter cannot be used for inference because the measurement does not involve additive Gaussian noise, but rather comes from the stochasticity of the spiking process (Eqs. 6-7).

Performing recursive inference (filtering), as described by Eq. 8, under the generative model described by equations 3-7 results in a posterior distribution that is approximately Gaussian[5] with a mean $\hat{u}_t$ and a variance σ_t^2:

$$\begin{aligned} \mathrm{P}(u_t \mid s_{0:t}) &= \left(\frac{g(u_t)}{\gamma_t} \right)^{s_t} \left(\frac{1 - g(u_t)\delta t}{1 - \gamma_t \delta t} \right)^{1-s_t} \cdot \\ &\quad \cdot \mathcal{N}\left[u_t;\ \hat{u}_{t-\delta t} + \left(u^{\mathrm{rest}} - \hat{u}_{t-\delta t} \right) \frac{\delta t}{\tau},\ \sigma_{\mathrm{W}}^2 \delta t + \left(1 - \frac{\delta t}{\tau} \right)^2 \sigma_{t-\delta t}^2 \right] \\ &\simeq \mathcal{N}\left[u_t; \hat{u}_t, \sigma_t^2 \right] \end{aligned} \tag{9}$$

Note that the smaller the bin size δt is, the closer this posterior distribution is to a Gaussian. The expected firing rate γ_t of the presynaptic cell at time t is obtained from the normalization condition $\left\langle \mathrm{P}(u_t \mid s_{0:t}) \right\rangle_{u_t \mid s_{0:t}} = 1$,

$$\gamma_t = \left\langle g(u_t) \right\rangle_{u_t \mid s_{0:t}} = g_0 \exp\left(\beta \hat{u}_t + \frac{\beta^2 \sigma_t^2}{2} \right) \tag{10}$$

The mean and variance of the posterior in Eq. 9 evolve (in continuous time, by taking the limit $\delta t \to 0$) as[5]:

$$\dot{\hat{u}}_t = \frac{1}{\tau}\left[u^{\mathrm{rest}} - \hat{u}_t \right] + \beta \sigma_{t-\varepsilon}^2 \left[S_t - \gamma_t \right] \tag{11}$$

$$\dot{\sigma}_t^2 = \frac{2}{\tau}\left[\sigma_{\mathrm{OU}}^2 - \sigma_t^2 \right] - \gamma_t \beta^2 \sigma_t^4 \tag{12}$$

where $S_t = \lim_{\delta t \to 0} \frac{s_t}{\delta t}$ denotes the spike train of the presynaptic cell represented as a sum of Dirac delta functions, and ε is an arbitrary small positive constant that ensures that at the time of a spike $t = t^{spike}$, the update of $\hat{u}$ is based on the variance just before the spike $\sigma_{t^{spike}-\varepsilon}^2$. (A similar, but not identical, derivation can be found in Ref. 26.)

Equation 11 indicates that each time a spike is observed, the estimated membrane potential should increase proportionally to the uncertainty (variance) about the current estimate. In turn, this estimation uncertainty then decreases each time a spike is observed (Eqs. 10 and 12). Conversely, in the absence of spikes, the estimated membrane potential

decreases while the variance increases back to its asymptotic value. It can be shown[5] that the representation of uncertainty about the membrane potential by σ^2 is self-consistent because it is predictive of the error of the mean estimator, $\hat{u}$.

The dynamics of the membrane potential estimator in equations 11 and 12 is closely related to the dynamics of short-term depression. This can be shown formally by taking the limit when presynaptic spikes are rare. In this case, equations 11-12 can be rewritten[5] as:

$$\dot{\hat{u}}_t \simeq \frac{\hat{v}_0 - \hat{u}_t}{\hat{\tau}_{\mathrm{m}}} + \hat{J}\hat{Y}\bar{\sigma}^2_{t-\varepsilon}S_t \tag{13}$$

$$\dot{\bar{\sigma}}^2_t \simeq \frac{1-\bar{\sigma}^2_t}{\hat{\tau}_{\mathrm{D}}} - \hat{Y}\bar{\sigma}^2_{t-\varepsilon}S_t \tag{14}$$

where $\bar{\sigma}^2 = \sigma^2/\sigma^2_\infty$ is the normalized variance of the estimator. The other constants involved are $\hat{v}_0 = \hat{u}_\infty$, $\hat{\tau}_{\mathrm{m}} = \tau$, $\hat{J} = 1/\left(\tau\gamma_\infty\beta^3\sigma^2_\infty\right)$, $\hat{Y} = \tau\gamma_\infty\beta^4\sigma^4_\infty$, and $\hat{\tau}_{\mathrm{D}} = 1/\left(2/\tau + \gamma_\infty\beta^2\sigma^2_\infty\right)$. $\hat{u}_\infty$, σ^2_∞ and γ_∞ are the stationary posterior mean, variance, and expected firing rate in the optimal estimator (Eqs. 9-10) in the absence of presynaptic spikes. More precisely, from Eqs. 11 and 12, we have

$$0 = \frac{1}{\tau}\left[u^{\mathrm{rest}} - \hat{u}_\infty\right] - \beta\sigma^2_\infty\gamma_\infty \tag{15}$$

$$0 = \frac{2}{\tau}\left[\sigma^2_{\mathrm{OU}} - \sigma^2_\infty\right] - \gamma_\infty\beta^2\sigma^4_\infty \tag{16}$$

where the expected firing rate is $\gamma_\infty = g_0\exp(\beta\hat{u}_\infty + \beta^2\sigma^2_\infty/2)$. Although it is difficult to get an explicit expression for $\hat{u}_\infty$ as a function of the model parameters only, from Eqs. 15 and 16 we can still express it as a function of the stationary variance σ^2_∞:

$$\hat{u}_\infty = u^{\mathrm{rest}} - \frac{2}{\beta}\left(\frac{\sigma^2_{\mathrm{OU}}}{\sigma^2\infty} - 1\right) \tag{17}$$

Equations 13-14 directly map the posterior mean, $\hat{u}$ and (normalized) variance, $\bar{\sigma}^2$ onto the postsynaptic membrane potential, v and the synaptic resource variable x in a canonical, biophysically-motivated model of a synapse undergoing synaptic depression (c.f. Equations 20-21).

The switching Ornstein-Uhlenbeck process

In the previous section, we modeled presynaptic membrane potential fluctuations with an Ornstein-Uhlenbeck process around a constant resting membrane potential. Here we generalize this process by letting the resting potential itself change. In this switching OU process, the resting membrane potential u^{rest}_t is not fixed but randomly switches between two levels, u^+ and u^-, corresponding to "up" and "down" states (**Supplementary Fig. 1b**)

$$\mathrm{P}\left(u_t^{\mathrm{rest}} \mid u_{t-\delta t}^{\mathrm{rest}}\right)=\begin{cases}1-\eta^{-}\delta t & \text{if } u_t^{\mathrm{rest}}=u^{+} \text{ and } u_{t-\delta t}^{\mathrm{rest}}=u^{+}\\ \eta^{-}\delta t & \text{if } u_t^{\mathrm{rest}}=u^{-} \text{ and } u_{t-\delta t}^{\mathrm{rest}}=u^{+}\\ 1-\eta^{+}\delta t & \text{if } u_t^{\mathrm{rest}}=u^{-} \text{ and } u_{t-\delta t}^{\mathrm{rest}}=u^{-}\\ \eta^{+}\delta t & \text{if } u_t^{\mathrm{rest}}=u^{+} \text{ and } u_{t-\delta t}^{\mathrm{rest}}=u^{-}\end{cases} \tag{18}$$

where η^{-} and η^{+} are the rates of switching to the "down" and "up" states, respectively. Similarly to equation 4, the presynaptic membrane potential evolves as an Ornstein-Uhlenbeck (OU) process around the resting potential u_t^{rest} which is now time-dependent:

$$\mathrm{P}\left(u_t \mid u_{t-\delta t}, u_t^{\mathrm{rest}}\right)=\mathcal{N}\left[u_t;\ u_{t-\delta t}+\frac{1}{\tau}\left(u_t^{\mathrm{rest}}-u_{t-\delta t}\right)\delta t,\ \sigma_{\mathrm{W}}^2\delta t\right] \tag{19}$$

Spike generation is described by the same rule as before (Eqs. 6-7). Although we were able to develop some analytical insight into the behavior of the optimal estimator in the case of a switching OU process (**Supplementary Note**), a full analytical treatment remains a challenging task. Thus, the results displayed in **Fig. 2** have been obtained by using standard particle filtering techniques[50] (see below).

Biophysically motivated STP model

The model we used was taken directly from Ref. 11 as a canonical model of a synapse undergoing STP. It describes how the postsynaptic potential v_t, the synaptic resource x_t (responsible for depression) and the utilization factor y_t (responsible for facilitation) co-vary in time:

$$\dot{v}_t=\frac{v_0-v_t}{\tau_{\mathrm{m}}}+J y_{t-\varepsilon} x_{t-\varepsilon} S_t \tag{20}$$

$$\dot{x}_t=\frac{1-x_t}{\tau_{\mathrm{D}}}-y_{t-\varepsilon} x_{t-\varepsilon} S_t \tag{21}$$

$$\dot{y}_t=\frac{Y-y_t}{\tau_{\mathrm{F}}}+Y(1-y_{t-\varepsilon})S_t \tag{22}$$

where v_0 is the postsynaptic resting membrane potential, τ_{m} is the postsynaptic membrane constant, J is (the static part of) synaptic efficacy, Y is the maximal synaptic utilization (and the rate of increase in y in response to a spike), τ_{D} is the time constant of synaptic depression, and τ_{F} is the facilitation time constant. Note that if the facilitation time constant is very short ($\tau_{\mathrm{F}} \to 0$), then y_t can be replaced by Y in Eqs. 20 and 21, resulting in pure depression. Also note that this standard model ignores the finite rise time of EPSPs. However, since rise times are usually about an order of magnitude faster than decay time constants, the effects of this approximation on the estimation performance of a synapse (as shown in **Fig. 4**) are expected to be negligible and in any case affect the static and the dynamic versions of the model equally.

Measuring the performance of estimators

The performance of an estimator, P, was measured as its rescaled root mean squared

error (**Fig. 4**):

$$P = 1 - \frac{1}{\sigma_{\mathrm{OU}}}\left[\frac{1}{T}\int_0^T \left(\hat{u}_t - u_t\right)^2 dt\right]^{\frac{1}{2}} \tag{23}$$

where $\hat{u}_t$ can be substituted with v_t to measure the performance of the biophysical models. Note that this provides a suitably normalized measure of performance as perfect estimation results in $P = 1$, and an estimator that outputs the expected mean presynaptic membrane potential, thereby completely ignoring presynaptic spikes, has $P = 0$.

Stochastic release

Fig. 4 shows the performance of static and depressing synapses and the optimal estimator for the case of stochastic vesicle release. Here, we provide the details of the calculations involved following Ref. 4.

Depressing synapses. Let N denote the total number of independent release sites. Each site can release at most one vesicle each time a presynaptic spike occurs. Each released vesicle gives rise to a quantum $q = J/N$ postsynaptic response, where J is the maximal EPSP amplitude. The postsynaptic membrane potential evolves as

$$\dot{v}_t = -\frac{v_0 - v_t}{\tau_m} + q n_t^r S_t \tag{24}$$

where τ_m is the membrane time constant, v_0 is the postsynaptic resting potential, and S_t is the presynaptic (delta) spike train. The total number of vesicles released at time t in response to a presynaptic spike, n_t^r, depends on the number of vesicles that are ready to fuse, N_t^r, and the utilisation fraction Y. More precisely, at the time of a spike, n_t^r is drawn from a binomial distribution:

$$n_t^r \sim Binomial(N_t^r, Y) \tag{25}$$

The number of ready-to-fuse vesicles N_t^r decreases by n_t^r each time there is a spike, and increases stochastically back to N with a time constant τ_D in the absence of spikes. Formally, the dynamics of N_t^r is given by

$$\dot{N}_t^r = R_t - n_t^r S_t \tag{26}$$

where $R_t = \sum_{t^{stk}} \delta(t - t^{stk})$, and t^{stk} denote the stochastic restocking times produced by an inhomogeneous Poisson process with intensity $(N - N_t^r)/\tau_D$.

It is easy to show that taking the expectation of Eqs. 24 and 26 over the stochastic release and restocking events, and setting $x_t = \left\langle N_t^r \right\rangle / N$, we get back the standard model of short-term depression (see Eqs. 20 and 21).

Static synapses. If we take the limit of a short time constant for depression, i.e. $\tau_D \to 0$, the restocking of the vesicle described by Eq. 26 becomes instantaneous and therefore we have $N_t^r \to N$. As a consequence, the number of released vesicles at the time of a spike

is given by

$$n_t^r \sim Binomial(N, Y) \tag{27}$$

Optimal estimator. In the case of stochastic vesicle release, the variables of the optimal estimator evolve as

$$\dot{\hat{u}}_t = \frac{1}{\tau}\left[u^{\text{rest}} - \hat{u}_t\right] + \beta\sigma_t^2\left[\frac{n_t^r}{NY}S_t - \gamma_t\right] \tag{28}$$

$$\dot{\sigma}_t^2 = \frac{2}{\tau}\left[\sigma_{\text{OU}}^2 - \sigma_t^2\right] - \gamma_t\beta^2\sigma_t^4 \tag{29}$$

where the number of released vesicles n_t^r is given by Eq. 27. Note that if we replace n_t^r by its expectation $\left\langle n_t^r \right\rangle = NY$, we get back the deterministic optimal estimator derived in Eqs. 11-12.

The "utilisation" parameter Y describes the probability of release for a presynaptic spike in all three models. For a fair comparison, it was chosen to be $Y = 0.39$ for all models, which optimized the performance of the dynamic synapse in the deterministic case for $\beta = 2$ (the value we used for the stochastic simulations), and which is consistent with experimentally reported values for the probability of release in cortical pyramidal-to-pyramidal cell connections[42].

Numerical simulations for the optimal estimator

We evaluated the mean posterior $\hat{u}_t$, the conditional means μ^+, μ^- and the conditional variances σ^+, σ^- numerically using a standard particle filtering technique[50]. In practice, we used $N^{part} = 10{,}000$ particles, each of which was two-dimensional $(u_t^{(i)}, u_t^{\text{rest},(i)})$, $i = 1,\ldots,N^{part}$. They evolved according to:

$$u_t^{\text{rest},(i)} \sim \mathrm{P}\left(u_t^{\text{rest},(i)} \mid u_{t-\delta t}^{\text{rest},(i)}\right) \tag{30}$$

$$u_t^{(i)} \sim \mathrm{P}\left(u_t^{(i)} \mid u_{t-\delta t}^{(i)}, u_t^{\text{rest},(i)}\right) \tag{31}$$

where $\mathrm{P}\left(u_t^{\text{rest}} \mid u_{t-\delta t}^{\text{rest}}\right)$ is given by Eq. 18 and $\mathrm{P}\left(u_t \mid u_{t-\delta t}, u_t^{\text{rest}}\right)$ by Eq. 19.

There was an (importance) weight, $w_t^{(i)}$, assigned to each particle, which was updated according to

$$w_t^{(i)} \propto w_{t-\delta t}^{(i)}\mathrm{P}\left(s_t \mid u_t^{(i)}\right), \qquad i = 1,\ldots,N^{part} \tag{32}$$

In each step of the simulation, all weights were renormalised such that $\sum_{i=1}^{N} w_t^{(i)} = 1$. The particles were resampled when the weights became strongly uneven. Formally, if the number of effective particles at time t, defined as

$$N_t^{eff} = \left(\sum_{i=1}^{N}\left(w_t^{(i)}\right)^2\right)^{-1} \tag{33}$$

fell below a given threshold $N_{thresh} = 9{,}000$ then all particles were resampled and the weights were all set back to $w_t^{(i)} = 1/N^{part}$.
The empirical mean and variance of the posterior membrane potential distribution were determined as

$$\hat{u}_t \simeq \sum_{i=1}^{N} u_t^{(i)} w_t^{(i)} \tag{34}$$

$$\sigma_{tot,t}^2 \simeq \sum_{i=1}^{N} \left(u_t^{(i)}\right)^2 w_t^{(i)} - \hat{u}_t^2 \tag{35}$$

The numerical evaluation of μ_t^+, $\left(\sigma_t^+\right)^2$, μ_t^- and $\left(\sigma_t^-\right)^2$ followed the same procedure except that the summation was restricted to the particles that were in the "up" (resp. "down") state and the weights were renormalized accordingly.

Model parameters for simulations

Unless otherwise noted, the presynaptic membrane time constant was set to $\tau = 20$ ms. The spiking determinism parameter was $\beta^{-1} = 3$ mV, and g_0 was set such that $g(-60\,\mathrm{m}V) = 10$ Hz.

Fig. 1 : The standard deviation of the presynaptic membrane potential was $\sigma_{\mathrm{OU}} = 5$ mV, the resting membrane potential was $u^{\mathrm{rest}} = -60$ mV.

Fig. 2 : The resting values were $u^- = -65$ mV, $u^+ = -55$ mV. The transition rates were $\eta^+ = \eta^- = 2$ Hz, and $\sigma_{\mathrm{OU}} = 2$ mV.

Fig. 3a : The fitted parameters were $\beta\sigma_{\mathrm{OU}} = 1.13$, $\tau = 1000$ ms and $g(-60\,\mathrm{m}V) = 10$ Hz. Each one of those 3 parameters were then sequentially perturbed by $\pm 5\%$ (resp. $\pm 10\%$). The area between the minimum and the maximum of those 6 conditions is represented by the dark (resp. light) gray shading.

Fig. 3b : The lower baseline potential was fixed at $u^- = -65$ mV. The fitted parameters were: $\sigma_{\mathrm{OU}} = 0.28$ mV, $\tau = 85.7$ ms, $u^+ = -53.9$ mV, $\eta^- = 1.09$ Hz, $\eta^+ = 1.13$ Hz, $\beta^{-1} = 3$ mV, and g_0 was set such that $g(-60\,mV) = 17.8$ Hz. Each one of those 7 parameters were sequentially perturbed by $\pm 5\%$ (dark gray shading) and by $\pm 10\%$ (light gray shading). Note that in this figure, we do not display the experimental data point for the shortest inter-spike interval (ISI = 5 ms) because our current spiking model does not include the effects of refractoriness and burstiness which may dominate estimation at such short intervals.

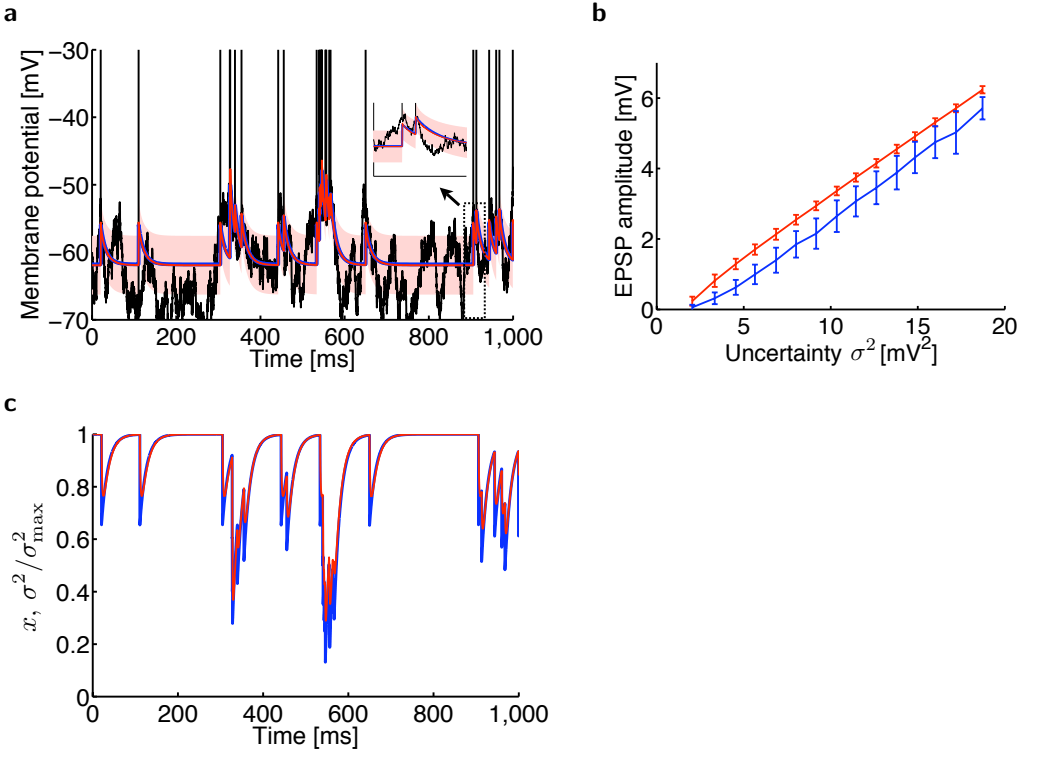
Fig. 4 : We set $u^{\mathrm{rest}} = -60$ mV, $\sigma_{\mathrm{OU}} = 1$ mV, and $\beta^{-1} = 0.5$ mV.

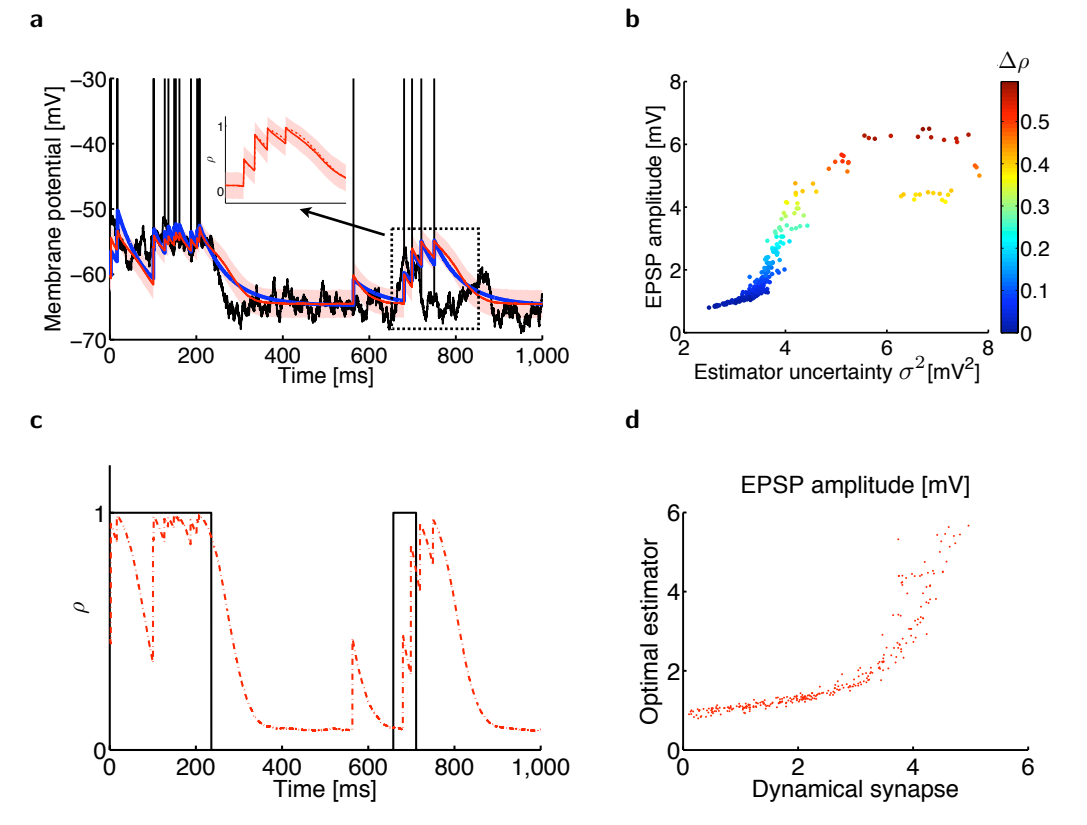
References

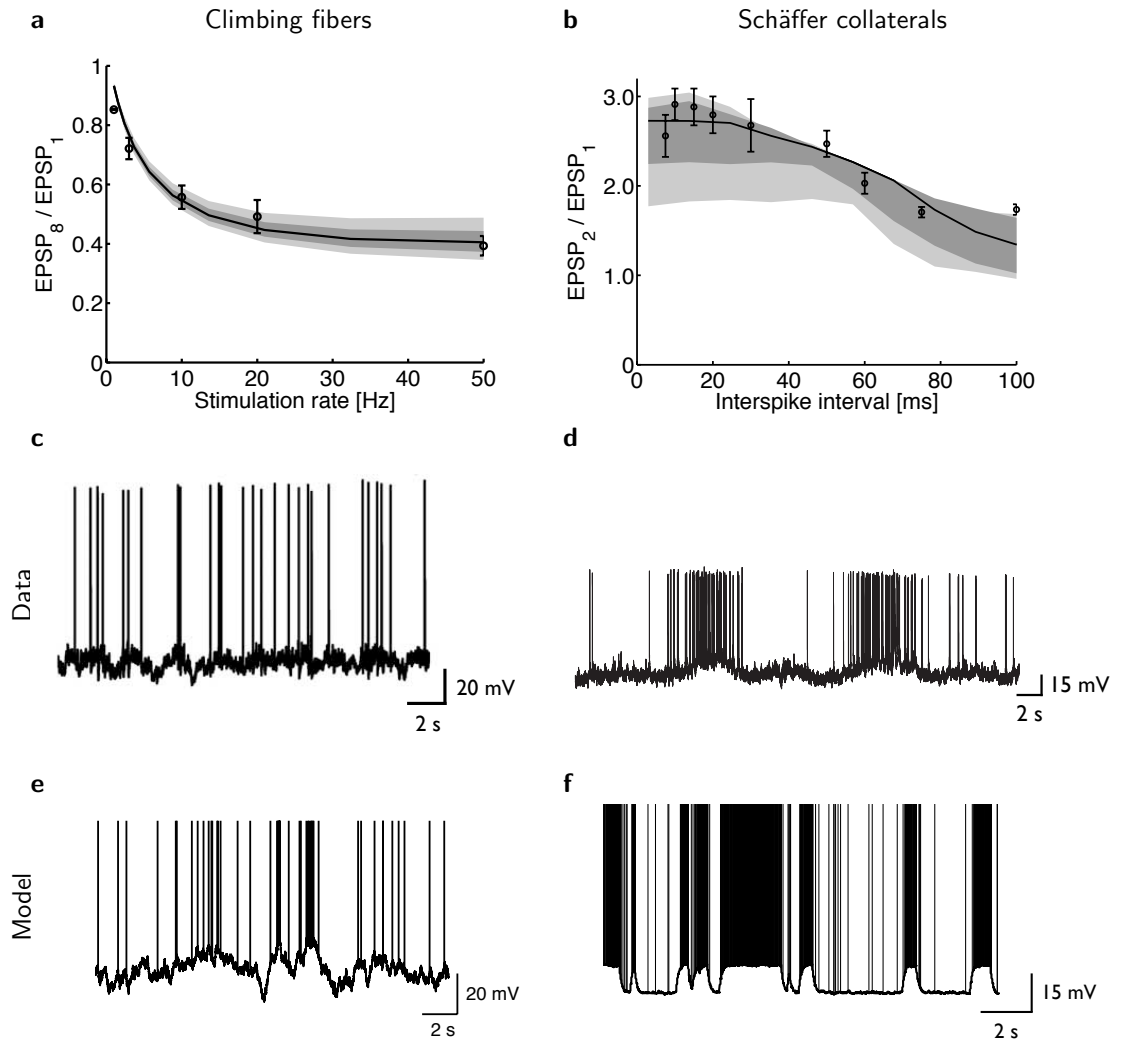
1. Markram, H., Wu, Y. & Tosdyks, M. Differential signaling via the same axon of neocortical pyramidal neurons. *Proc. Natl. Acad. Sci. USA* **95**, 5323–5328 (1998).
2. Zucker, R. & Regehr, W. Short-term synaptic plasticity. *Annu. Rev. Physiol*. **64**, 355–405 (2002).
3. Abbott, L.F. & Regehr, W.G. Synaptic computation. *Nature* **431**, 796–803 (2004).
4. Loebel, A. et al. Multiquantal release underlies the distribution of synaptic efficacies in the neocortex. *Front. Comput. Neurosci.* **3**, 1–13 (2009).
5. Pfister, J.P., Dayan, P. & Lengyel, M. Know thy neighbour: A normative theory of synaptic depression. in *Advances in Neural Information Processing Systems 22* (eds. Bengio, Y., Schuurmans, D., Lafferty, J., Williams, C.K.I. & Culotta, A.) 1464–1472 (2009).
6. Dittman, J., Kreitzer, A. & Regehr, W. Interplay between facilitation, depression, and residual calcium at three presynaptic terminals. *J. Neurosci*. **20**, 1374–1385 (2000).
7. Merkel, M. & Lindner, B. Synaptic filtering of rate-coded information. *Physical Review E* **81**, 41921 (2010).
8. Abbott, L.F., Varela, J.A., Sen, K. & Nelson, S.B. Synaptic depression and cortical gain control. *Science* **275**, 220–224 (1997).
9. Cook, D., Schwindt, P., Grande, L. & Spain, W. Synaptic depression in the localization of sound. *Nature* **421**, 66–70 (2003).
10. Goldman, M., Maldonado, P. & Abbott, L. Redundancy reduction and sustained firing with stochastic depressing synapses. *J. Neurosci*. **22**, 584–591 (2002).
11. Mongillo, G., Barak, O. & Tsodyks, M. Synaptic theory of working memory. *Science* **319**, 1543–1546 (2008).
12. Carpenter, G. & Grossberg, S. Pattern recognition by self-organizing neural networks (The MIT Press, 1991).
13. Hasselmo, M. & Bower, J. Acetylcholine and memory. *Trends Neurosci*. **16**, 218–222 (1993).
14. Zador, A.M. Impact of synaptic unreliability on the information transmitted by spiking neuron. *J. Neurophysiol*. **79**, 1219–1229 (1998).
15. de la Rocha, J., Nevado, A. & Parga, N. Information transmission by stochastic synapses with short-term depression: neural coding and optimization. *Neurocomputing* **44**, 85–90 (2002).
16. Pfister, J.P. & Lengyel, M. Speed versus accuracy in spiking attractor networks. in *Front. Syst. Neurosci. Conference Abstract: Computational and systems neuroscience 2009*, (2009).
17. Wilson, H.R. & Cowan, J.D. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J*. **12**, 1–24 (1972).
18. Dayan, P. & Abbott, L.F. Theoretical Neuroscience (MIT Press, Cambridge, 2001).
19. Thorpe, S., Fize, D. & Marlot, C. Speed of processing in the human visual system. *Nature* **381**, 520–522 (1996).
20. Huys, Q., Zemel, R., Natara jan, R. & Dayan, P. Fast population coding. *Neural*

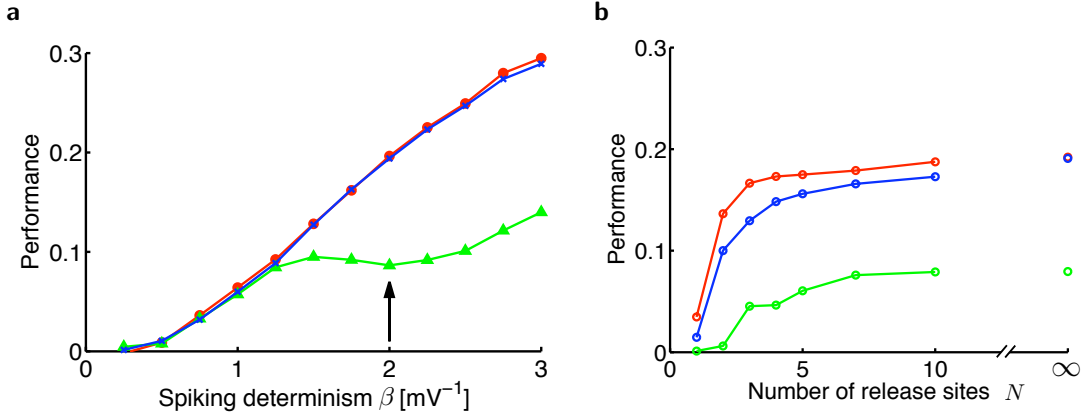
Comput. **19**, 404–441 (2007).
21. Stanford, T., Shankar, S., Massoglia, D., Costello, M. & Salinas, E. Perceptual decision making in less than 30 milliseconds. *Nat. Neurosci*. **13**, 379–385 (2010).
22. Hopfield, J.J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA* **79**, 2554–2558 (1982).
23. Lengyel, M., Kwag, J., Paulsen, O. & Dayan, P. Matching storage and recall: hippocampal spike timing-dependent plasticity and phase response curves. *Nat. Neurosci.* **8**, 1677–1683 (2005).
24. Seung, H.S. & Sompolinski, H. Simple models for reading neuronal population codes. *Proc. Natl. Acad. Sci. USA* 90, 10749–10753 (1993).
25. Anderson, B., Moore, J. & Barratt, J. Optimal filtering (Prentice-Hall Englewood Cliffs, NJ, 1979).
26. Eden, U., Frank, L., Barbieri, R., Solo, V. & Brown, E. Dynamic analysis of neural encoding by point process adaptive filtering. *Neural Comput*. **16**, 971–998 (2004).
27. Paninski, L. The most likely voltage path and large deviations approximations for integrate-and-fire neurons. *J. Comput. Neurosci*. **21**, 71–87 (2006).
28. Bobrowski, O., Meir, R. & Eldar, Y. Bayesian filtering in spiking neural networks: Noise, adaptation, and multisensory integration. *Neural Comput*. **21**, 1277–1320 (2009).
29. Cunningham, J., Yu, B., Shenoy, K. & Sahani, M. Inferring neural firing rates from spike trains using Gaussian processes. in *Advances in Neural Information Processing Systems 20* (eds. Platt, J., Koller, D., Singer, Y. & Roweis, S.) 329–336 (MIT Press, Cambridge, MA, 2008).
30. Gerstner, W. & Kistler, W.K. Spiking Neuron Models (Cambridge University Press, Cambridge UK, 2002).
31. Paninski, L., Pillow, J. & Simoncelli, E. Maximum likelihood estimate of a stochastic integrate-and-fire neural encoding model. *Neural Comput*. **16**, 2533–2561 (2004).
32. Stein, R.B. A theoretical analysis of neuronal variability. *Biophys. J*. **5**, 173–194 (1965).
33. Lansky, P. & Ditlevsen, S. A review of the methods for signal estimation in stochastic diffusion leaky integrate-and-fire neuronal models. *Biol. Cybern*. **99**, 253–262 (2008).
34. Tsodyks, M., Pawelzik, K. & Markram, H. Neural networks with dynamic synapses. *Neural Comput*. **10**, 821–835 (1998).
35. Shinomoto, S., Sakai, Y. & Funahashi, S. The Ornstein-Uhlenbeck process does not reproduce spiking statistics of neurons in prefrontal cortex. *Neural Comput*. **11**, 935–951 (1999).
36. Steriade, M., Nunez, A. & Amzica, F. A novel slow (< 1 Hz) oscillation of neocortical neurons in vivo: depolarizing and hyperpolarizing components. *J. Neurosci.* **13**, 3252–3265 (1993).
37. Cossart, R., Aronov, D. & Yuste, R. Attractor dynamics of network up states in the neocortex. *Nature* **423**, 283–288 (2003).
38. Harvey, C., Collman, F., Dombeck, D. & Tank, D. Intracellular dynamics of hippocampal place cells during virtual navigation. *Nature* **461**, 941–946 (2009).
39. Dobrunz, L., Huang, E. & Stevens, C. Very short-term plasticity in hippocampal synapses. *Proc. Natl. Acad. Sci. USA* **94**, 14843–14847 (1997).

40. Chorev, E., Yarom, Y. & Lampl, I. Rhythmic episodes of subthreshold membrane potential oscillations in the rat inferior olive nuclei in vivo. *J. Neurosci*. **27**, 5043–5052 (2007).
41. Markram, H. & Tsodyks, M. Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature* **382**, 807–810 (1996).
42. Thomson, A. Facilitation, augmentation and potentiation at central synapses. *Trends Neurosci*. **23**, 305–312 (2000).
43. Martin, S., Grimwood, P. & Morris, R. Synaptic plasticity and memory: an evaluation of the hypothesis. *Annu. Rev. Neurosci*. **23**, 649–711 (2000).
44. Manabe, T., Wyllie, D., Perkel, D. & Nicoll, R. Modulation of synaptic transmission and long-term potentiation: effects on paired pulse facilitation and EPSC variance in the CA1 region of the hippocampus. *J. Neurophysiol*. **70**, 1451 (1993).
45. Reyes, A. et al. Target-cell-specific facilitation and depression in neocortical circuits. *Nat. Neurosci*. **1**, 279–285 (1998).
46. Koester, H. & Johnston, D. Target cell-dependent normalization of transmitter release at neocortical synapses. *Science* **308**, 863–866 (2005).
47. Denève, S. Bayesian spiking neurons I: inference. *Neural Comput*. **20**, 91–117 (2008).
48. Wark, B., Fairhall, A. & Rieke, F. Timescales of inference in visual adaptation. *Neuron* **61**, 750–761 (2009).
49. Jolivet, R., Rauch, A., Lüscher, H.R. & Gerstner, W. Predicting spike timing of neocortical pyramidal neurons by simple threshold models. *J. Comput. Neurosci*. **21**, 35–49 (2006).
50. Doucet, A., De Freitas, N. & Gordon, N. Sequential Monte Carlo methods in practice (Springer, New York, 2001).


a
Membrane potential [mV]
−30
−40
−50
−60
−70
0
200
400
600
800
1,000
Time [ms]
b
EPSP amplitude [mV]
0
2
4
6
0
5
10
15
20
Uncertainty σ^2 [mV^2]
c
$x,\ \sigma^2/\sigma^2_{\max}$
0
0.2
0.4
0.6
0.8
1
0
200
400
600
800
1,000
Time [ms]


a
Membrane potential [mV]
−30
−40
−50
−60
−70
0
200
400
600
800
1,000
Time [ms]
1
ρ
0
b
EPSP amplitude [mV]
8
6
4
2
0
2
4
6
8
Estimator uncertainty σ^2[mV^2]
$\Delta\rho$
0.5
0.4
0.3
0.2
0.1
0
c
ρ
1
0
0
200
400
600
800
1,000
Time [ms]
d
EPSP amplitude [mV]
Optimal estimator
6
4
2
0
0
2
4
6
Dynamical synapse


a
Climbing fibers
EPSP8 / EPSP1
Stimulation rate [Hz]
b
Schäffer collaterals
EPSP2 / EPSP1
Interspike interval [ms]
c
Data
20 mV
2 s
d
15 mV
2 s
e
Model
20 mV
2 s
f
15 mV
2 s


a
b
Performance
0.3
0.2
0.1
0
0
1
2
3
Spiking determinism β [mV^{-1}]
Performance
0.3
0.2
0.1
0
0
5
10
∞
Number of release sites N