Nature **384**, 76–78, (5. September 1996).

A neuronal learning rule for sub-millisecond temporal coding

Wulfram Gerstner[1,+], **Richard Kempter**[1],
J. Leo van Hemmen[1], **and Hermann Wagner**[2,*]

(1) Physik-Department der TU München
(2) Fakultät für Chemie und Biologie, TU München
D-85747 Garching bei München, Germany

(+) present address: Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland
(*) present address: RWTH Aachen, Institut für Biologie II, Kopernikusstr. 16, D-52074 Aachen, Germany

corresponding author: Leo.van.Hemmen@physik.tu-muenchen.de

An unresolved paradox exists in auditory and electrosensory neural systems [1, 2]: they encode behaviourally relevant signals in the range of a few microseconds with neurons that are at least one order of magnitude slower. We take the barn owl's auditory system as an example and present a modeling study based on computer simulations of a neuron in the laminar nucleus. Three observations resolve the paradox. First, spiking of an integrate-and-fire neuron driven by excitatory postsynaptic potentials (EPSPs) with a width at half maximum of 250 μs has an accuracy of 25 μs if the presynaptic signals arrive coherently. Second, the necessary degree of coherence in the signal arrival times can be attained during ontogenetic development by virtue of an unsupervised Hebbian learning rule. Learning selects connections with matching delays from a broad distribution of axons with random delays. Third, the learning rule also selects the correct delays from two independent groups of inputs, for example, from the left and right ear.

The importance of temporal coding in neural information processing is being debated vigorously [3, 4, 5, 6, 7, 8]. One of the central issues in this debate is the question of whether neuronal firing can be more precise than the time constants of the neuronal processes involved [9]. We address this problem and study the barn owl's auditory system.

Barn owls use interaural time difference for azimuthal sound localisation [10, 11]. They can locate sounds, and hence prey, with a precision of 1-2 degrees [12], a capability which requires a time resolution of less than 5 μs [11]. The key process in the auditory system is 'phase locking': spikes occur preferentially at a certain phase, termed 'mean phase', of a stimulating tone at frequencies up to 8 kHz [13]. In order to understand the high degree of temporal precision in phase locking, we have to consider both the typical duration τ_s of synaptic input current and the membrane time constant τ_m. For auditory neurons in the chicken [14, 15], time constants of synaptic input are in the range of 200 μs. Even though the passive membrane time constant is about 2 ms [14], the *effective* membrane time constant (cf. Methods, Fig. 1) of magnocellular and laminar neurons is shorter than 200 μs because there is an outward rectifying current which is activated above resting potential [14]. Such outward rectifying currents are commonly found in phase-locking neurons of the auditory system [16, 17]. As a result of the short time constants, the width of single EPSPs in chicken is about 500 - 800 μs at half maximum amplitude [15]. We assume that laminar neurons in an auditory specialist like the barn owl are twice as fast as those in chicken and model EPSPs with a width of 250 μs (Fig. 1*a*, inset).

We concentrate on a single frequency channel and stimulate with a pure tone of, e.g., 5kHz (period $T = 200$ μs). The input from magnocellular neurons to our integrate-and-fire model neuron in the laminar nucleus is described as a sequence of spikes (Fig. 1*b*). Presynaptic spikes occur preferentially around a specific mean phase of the tone [13]. A jitter of 40 μs takes care of internal sources of variability and noise along the auditory pathway as well as of the bandwidth of frequency tuning of auditory neurons. The mean firing rate is relatively low (667 Hz) so that each presynaptic spike train skips most of the cycles of the 5 kHz tone.

EPSPs of spikes from more than 100 presynaptic neurons that arrive coherently with a common mean phase lead to an oscillating membrane potential $\mathcal{V}$, that builds up, reaches the threshold in the *rising* stage of the cycle and in this way produces a slightly better phase locking than that of the input (Fig. 1a). If presynaptic spikes arrive with random phases, however, $\mathcal{V}$ shows aperiodic fluctuations (Fig. 1a), and the output spikes have a *uniform* phase distribution (Fig. 2a). We expect this to be the initial, embryonic, condition before the connections between the magnocellular and the laminar nucleus are tuned during a sensitive period [19]. In adult owls, transmission delays from the ear to the laminar nucleus differ greatly. Their mean value is between 2 and 3 ms, and the standard deviation is about 200–240 μs [18]. Without a tuning of the delays, any phase information would be lost. A Hebbian learning rule proves to be an efficient tuning mechanism.

In Hebbian learning, a synaptic weight is changed by a small amount ϵ if presynaptic spike arrival and postsynaptic firing 'coincide'. This simultaneity constraint is implemented by a learning window $W(s)$ where s is the difference between the arrival time of a presynaptic spike and the postsynaptic firing. In our model, $W(s)$ has two regimes (Fig. 2d). For $s < 0$, $W(s)$ is positive. Thus, the efficacy of synapses which are repeatedly active shortly before a postsynaptic spike occurs is increased [20, 21]. The weights of synapses which are active shortly *after* the postsynaptic spike are *de*creased [22, 23]. As depolarisation is known to induce potentiation of active synapses [21] and as the neuron is depolarised most of the time between two spikes, we also strengthen each active synpase by a small amount γ, even if no postsynaptic spike occurs. The procedure of continuously strengthening and weakening the synapses automatically leads to a normalization of the total input strength to the postsynaptic neuron in a competitive self-organised process.

As a result of learning, a clearly structured distribution of synapses evolves (Figs. 2b,c). Delays of the remaining synapses differ approximately by multiples of the period T of the stimulating tone. We emphasize that the period is not previously known to the cell. The final synaptic pattern of a neuron stimulated by a 2 kHz signal during learning is different from that of a 5 kHz neuron (Figs. 2b,c).

The evolution of synaptic weights during learning does not depend on the specific shape of the learning window W but only on some generic properties. In particular, the learning window can be nonzero in a domain which extends over several milliseconds [23] and therefore is large compared with the period T of the sound (Fig. 2d). Efficient learning depends on the temporal relationship between the process of strengthening the synaptic weights and that of weakening them. The maximum of the window function $W(s)$ should be at a location $\tilde{s} \approx -\tau_s/2$ where τ_s is the rise time of the postsynaptic potential (Fig. 2d). This can be understood as follows. Let us assume that learning has led to a sharply peaked distribution with all synapses having a common delay Δ (mod T). Coherent input arriving at the synapses can trigger a postsynaptic spike with a mean delay of approximately $\tau_s/2$. Because of learning, all synapses which are active slightly *before* the postsynaptic firing will be strengthened. If $\tilde{s} = -\tau_s/2$, the maximal increase of synaptic weights occurs at those synapses which are already strongest.

So far we have considered monaural input, but laminar neurons also exhibit a sensitivity to the *interaural* time difference (ITD) [18, 24, 25]. We divide the synapses into two groups, viz., with input from the left and the right ear. During learning, both ears are stimulated by the same signal and with a fixed ITD. The learning rule selects synaptic connections so that spikes arrive coherently for exactly this ITD. If the same ITD is used *after* learning, the neuron is driven optimally and emits phase-locked spikes. If the ITD does not match, phase locking of the output spikes breaks down and the mean firing rate decreases (Fig. $3a, b$).

Temporal information conveyed by a *single* laminar neuron is limited. The temporal precision of phase locking is about 20-25 μs (Fig. 2) and the ITD tuning curve is only weakly modulated (Fig. $3a$). Nevertheless, barn owls achieve a behavioural performance corresponding to a temporal precision of 5 μs. How can we understand this remarkable accuracy? The barn owl has a reaction time of about 100 ms before it initiates a head movement [12]. During 100 ms, the firing pattern of a *population* of laminar neurons contains enough information to resolve ITDs with a precision of 5 μs, as will now be shown.

Let us consider a *group* of laminar neurons with ITD tuning curves as in Fig. $3a$, but shifted along the ITD axis so that the optimal responses occur at different ITDs. Neurons are stimulated by a tone with fixed, but unknown, ITD. We estimate the ITD from the neuronal firing pattern by a 'population vector' decoding scheme [26, 27]: Each neuron 'votes' for the ITD corresponding to its optimal response and the votes are weighted according to the number of spikes that occur in a time window of 100 ms. We find that the activity of about 100 independent neurons provides enough information to estimate the ITD with a precision of 5 μs (Fig. 4) – apart from ambiguities that are due to the periodicity of the signal. Weak correlations from common inputs rescale the absolute valuess but do not alter the conclusions.

The firing precision of 20-25 μs of *single* neurons has been achieved despite the fact that input spikes evoke EPSPs which are ten times broader. In an additional set of simulations, we have systematically varied the effective membrane time constant τ_m and, hence, the width of the EPSP. We find that the temporal precision of neuronal spiking depends only weakly on τ_m (Fig. $3d$). This is possible because firing always occurs during the *rise* time of a postsynaptic potential. On the other hand, for a 5 kHz signal, ITD tuning breaks down rapidly, if τ_m exceeds 0.1 ms (Fig. $3c$). Thus modulation of ITD tuning does indeed require short EPSPs. In fact, very short time constants have been measured by Reyes et al. [14, 15].

Our results show that the temporal precision of output spikes need not be limited by the membrane time constant. Spike timing can achieve a resolution shorter than the rise time of an EPSP, given *coherent* spike arrival – and spike arrival times can be tuned accurately by a Hebbian learning rule. It is tempting to apply the same ideas to information processing in the hippocampus [3], the cerebellum [28], or the cortex [5]. If we increase time constants by a factor of about 100 and reinterpret our results, then we are led to the conclusion that in areas where effective membrane time constants are in the range order of 10 - 20 ms [29], a temporal code with an accuracy of 1-3 ms would be possible.

Acknowledgment: We thank D. Kautz, C. Köppl, and A. Reyes for helpful discussions, and R.G. Palmer for a critical reading of the manuscript. W.G. and R.K. have been supported by the DFG.

[1] Carr, C. E. *Ann. Rev. Neurosci.* **16**, 223-243 (1993)
[2] Heiligenberg, W. *Neural nets in electric fish*, Cambridge, Mass. (1991)
[3] Hopfield, J. J. *Nature* **376**, 33-36 (1995).
[4] Bialek, W., Rieke, F., de Ruyter van Steveninck, R. R. & Warland, D. *Science* **252**, 1854–1857 (1991).
[5] Abeles, M., in *Models of Neural Networks II* (eds. E. Domany, J. L. van Hemmen & K. Schulten) 121–140 (Springer, New York, 1994).
[6] Shadlen, M. N. & Newsome, W. T. *Curr. Opin. Neurobiol.* **4**, 569–579 (1994) and **5**, 248–250 (1995).
[7] Softky, W. R. *Curr. Opin. Neurobiol.* **5**, 239–247 (1995).
[8] Mainen, Z. F. & Sejnowski, T. J. *Science* **268**, 1503–1506 (1995).
[9] Softky, W. & Koch, C. *J. Neurosci.* **13**, 334–250 (1993).
[10] Jeffress, L. A. *J. Comp. Physiol. Psychol.* **41**, 35–39 (1948).
[11] Moiseff, A. & Konishi, M. *J. Comp. Physiol. A* **144**, 299–304 (1981).
[12] Knudsen, E. I., Blasdel G. G., and Konishi, M. *J. Comp. Physiol.* **133**, 1-11 (1979).
[13] Sullivan, W. E. & Konishi, M. *J. Neurosci.* **4**, 1787–1799 (1984).
[14] Reyes, A. D., Rubel, E. W., and Spain W. J. *J. Neurosci.* **14**, 5352 - 5364 (1994).
[15] Reyes, A. D., Rubel, E. W., and Spain W. J. *J. Neurosci.* **16**, 993-1007 (1996).
[16] Manis, P. B. & Marx, S. O. *J. Neurosci.* **11**, 2865-2880 (1991).
[17] Oertel, D. *J. Neurosci.* **3**, 2043-2053 (1983).
[18] Carr, C. E. & Konishi, M. *J. Neurosci.* **10**, 3227-3246 (1990).
[19] Carr, C. E. *Advances in Hearing Research* (eds. G. A. Manley, G. M. Klump et al.) 24-20 (World Scientific, Singapur, 1995).
[20] Hebb, D. O. *The organization of behavior.* (Wiley, New York, 1949).
[21] Bliss, T. V. P. & Collingridge G. L. *Nature* **361**, 31–39 (1993).
[22] Debanne, D., Gähwiler, B. H. & Thompson, S. M. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1148–1152 (1994).
[23] Markram, H. & Sakmann, B. *Soc. Neurosci. Abstr.* **21**, 2007 (1995).
[24] Goldberg, J. M. & Brown, P. B. *J. Neurophysiol.* **32**, 613–636 (1969).
[25] Yin T. C. T. & Chan, J. C. K *J. Neurophysiol.* **64**, 465–488 (1990).
[26] Georgopoulos, A. P., Schwartz, A. B., & Kettner, R. E. *Science* **233**, 1416-1419 (1986).
[27] Salinas, E. & Abbott, L. F. *J. Comput. Neurosci.* **1**, 89-107 (1994).
[28] Braitenberg, V. *Network* **4**, 11-17 (1993).
[29] Bernander, Ö., Douglas, R. J., Martin K. A. C. & Koch C. *Proc. Natl. Acad. Sci. USA* **88**, 11569-11571 (1991).

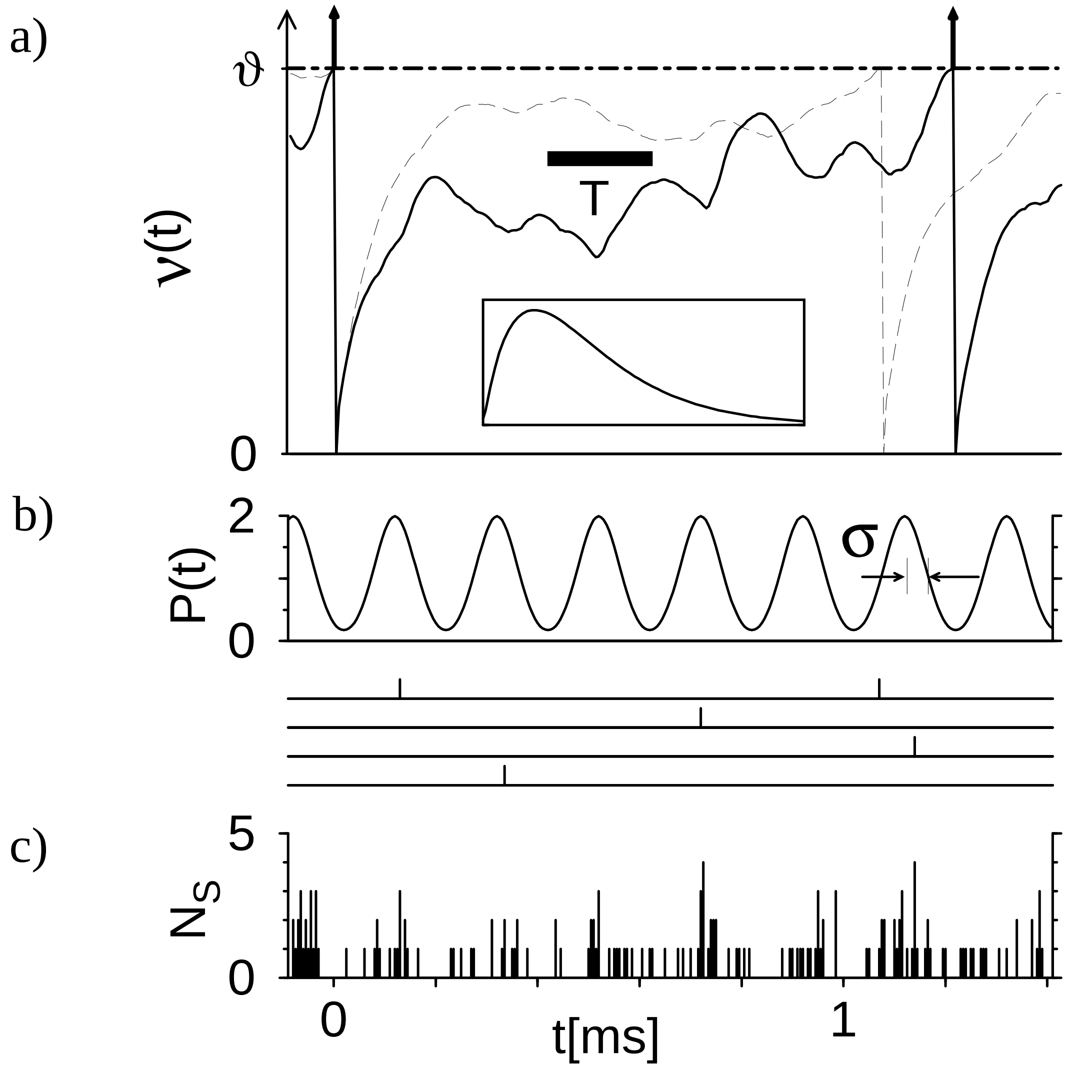

FIG. 1 *a* Membrane potential $\mathcal{V}$ of an integrate-and-fire neuron as a function of time. *b* Probability density $P(t)$ of presynaptic firing during 5 kHz stimulation and four samples of input spike trains (vertical bars). The model neuron receives input from 154 presynaptic neurons [1, 18] in volleys of phase-locked spikes with a jitter of 40 μs driven by a 5 kHz tone. Spikes are generated by a stochastic process with periodically modulated rate (solid line in *b*). A histogram of spike arrival times (number of spikes N_s in bins of 5 μs) summed over all 154 synapses is shown in *c*. Each input spike evokes an excitatory postsynaptic potential (EPSP) shown on an enlarged voltage scale (same time scale) in the inset of *a*. The EPSPs from all neurons are added linearly and yield the membrane voltage $\mathcal{V}$ (*a*, main figure). With the spike input shown in *c* the membrane voltage exhibits oscillations (solid line). The model neuron fires (arrow), if $\mathcal{V}$ reaches a threshold ϑ. Firing must always occur during the time when $\mathcal{V}$ increases so that, in the case of coherent input, output spikes are phase-locked as well. If input spikes arrive incoherently, $\mathcal{V}$ follows a trajectory with statistical fluctuations, but no systematic oscillations (dashed line) and the output spikes are not phase locked. Times scales in a-c are identical; the period of $T = 200$ μs is indicated by a horizontal bar. Voltage in a: arbitrary units; the threshold ϑ is 36 times the amplitude of a single EPSP. Rate in b in kHz.

METHODS. The differential equation $d\mathcal{V}/dt = -\mathcal{V}/\tau_m + I(t)$ has been integrated (time step 5 μs) with an *effective* membrane time constant $\tau_m = 100$ μs. An input spike arriving at time t_j^f at a synapse j evokes an input current $I_j^f(t) = (1/\tau_s)\exp[-(t-t_j^f)/\tau_s]$ for $t > t_j^f$ with $\tau_s = 100$ μs. The total input is $I(t) = \sum_{j,f} J_j I_j^f(t)$ where J_j is the efficacy of synapse j and the sum runs over all synapses and firing times preceding t. For each input channel j, firing times are generated with probability density $P(t) = (\nu T/\sigma\sqrt{2\pi})\sum_{m=-\infty}^{\infty}\exp[-(t-mT-\Delta_j)^2/2\sigma^2]$ where $\nu = 1$ kHz is a rate, $T = 200$ μs is the period of the stimulation tone, Δ_j is the transmission delay from the ear to the laminar nucleus, and $\sigma = 40$ μs is the temporal jitter. We impose absolute refractoriness through a minimal interspike interval of 0.5 ms. For incoherent spike arrival (dashed line in a) the transmission delays Δ_j are drawn from the broad distribution of Fig. 2a, for coherent spike arrival (solid line in a) the delay distribution is that of Fig. 2c. To define the effective time constant τ_m of the integrate-and-fire model, we started from a more detailed model with an explicit description of the outward rectifying current [14, 15].

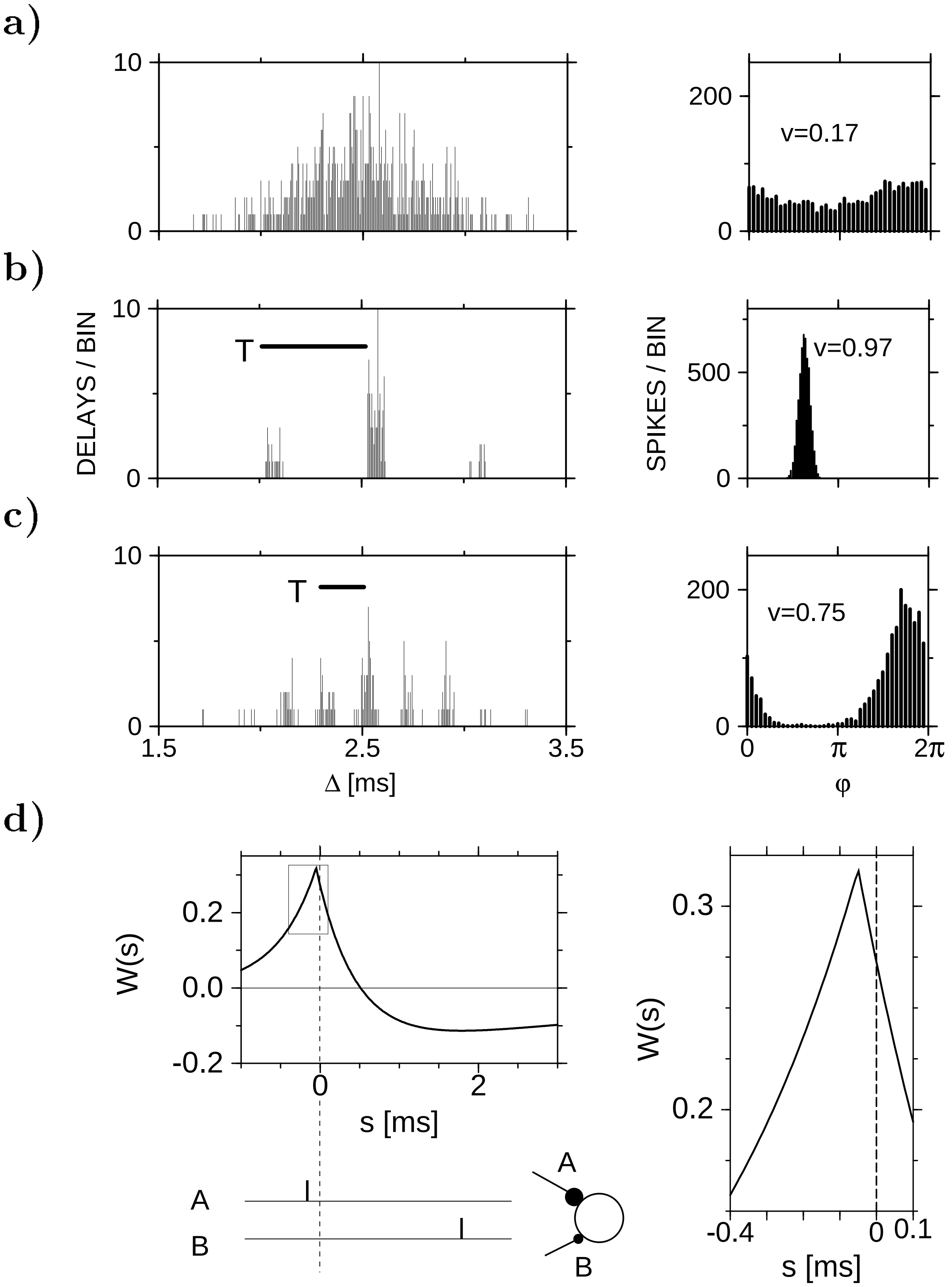

FIG. 2 Hebbian learning. In *a–c*, the left-hand graph shows synapses binned as a function of the signal transmission delay Δ. On the right-hand side the distributions of output phases are shown in period histograms (bin width 5 μs).

a, Before learning, there are 600 synapses with a broad distribution of signal transmission delays (left) drawn from a Gaussian distribution (2.5 ± 0.3 ms). The output of the laminar model neuron shows *no* phase locking to a 2 kHz (not shown) or a 5 kHz signal (right).

b, After a period of learning while being stimulated by a 2 kHz signal. The 105 synapses which survive learning have delays which differ by *multiples* of the period $T = 500$ μs (scaling bar). The output spikes exhibit phase locking with vector strength [24] $v = 0.97$ corresponding to a temporal precision of 20 μs (right).

c, Same plot as in *b*, after learning a 5 kHz input signal. 154 saturated synapses survive. The output spikes exhibit phase locking ($v = 0.75$) corresponding to a temporal precision of 25 μs.

d, Learning window W as a function of the delay s between postsynaptic firing and presynaptic spike arrival. The graph on the right-hand side shows the boxed region around the maximum on an expanded scale. If $W(s)$ is positive (negative) for some s, the synaptic efficacy is increased (decreased). The postsynaptic firing occurs at $s = 0$ (vertical dashed line). Learning is most efficient, if presynaptic spikes arrive shortly before the postsynaptic neuron starts firing as in synapse A. Another synapse B which fires *after* the postsynaptic spike is weakened.

METHODS. Before learning, all synapses have identical weights $J_j = 1$. During learning weights are changed according to $\Delta J_j = \epsilon \sum_f [\gamma + \sum_n W(t_j^f - t^n)]$ with factors $\epsilon = 0.002, \gamma = 0.1$. The sum runs over all spike arrival times $t_j^f < t$ and all postsynaptic firings $t^n < t$. We take $W(s) = 0.3 \exp[(s + 0.05)/0.5]$ for $s < -0.05$ and $W(s) = 0.5 \exp[-(s + 0.05)/0.5] - 0.2 \exp[-(s + 0.05)/5]$ for $s \geq -0.05$ where s is a time in ms. Learning occurs during 3000 seconds in time steps of 5 μs. The synaptic strength saturates at a maximum of 3. A synapse whose weight vanishes is removed. The efficiency of phase locking is quantified by the vector strength [24]. Spike input and electrical time constants are as in Fig. 1.

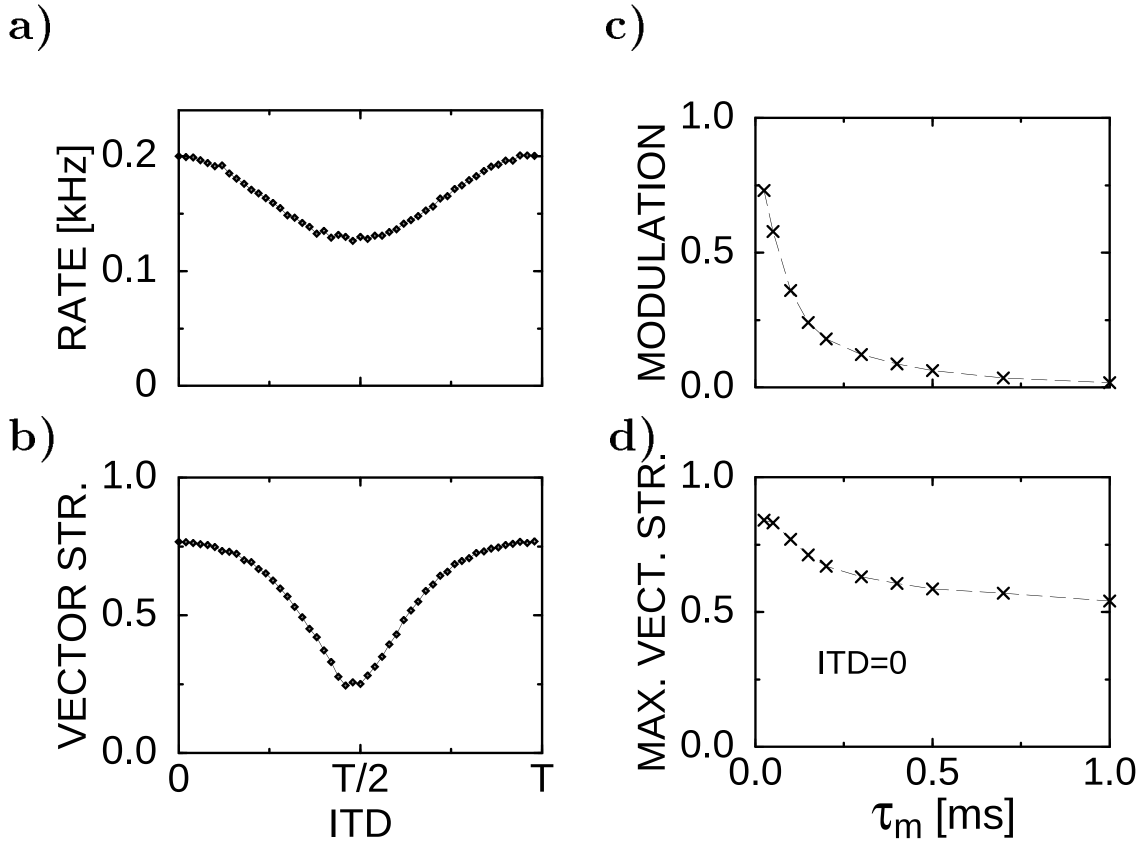

FIG. 3 *a* and *b*. Tuning to interaural time difference (ITD). The output rate (*a*) and the vector strength (*b*) of a laminar neuron are shown as a function of ITD. The model neuron has been tuned to a 5 kHz signal ($T = 200\ \mu$s) as described in the text and the legend of Fig. 2. Half of the 154 synapses which survive learning transmit signals from the left ear, the others from the right ear. The neuron exhibits best phase locking and maximal output rate $f_{\max}$ for the interaural delay used during learning (ITD = 0). The rate has a minimum $f_{\min}$ for ITD = $T/2$. *c* and *d*. Effect of τ_m. We define the modulation as $(f_{\max} - f_{\min})/f_{\max}$ and plot the modulation as a function of the effective membrane time constant τ_m in *c*. The vector strength at ITD = 0 is shown as a function of τ_m in *d*.

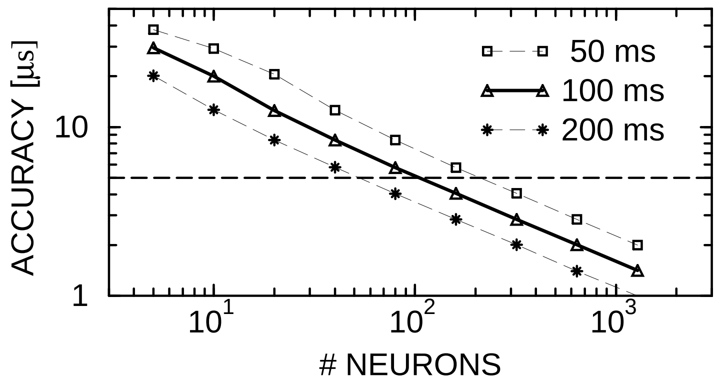

FIG. 4 Population coding. The temporal precision that can be achieved in a population of N independent neurons is shown as a function of N under the assumption of population vector coding [26, 27]. The ITD is estimated based on spikes occurring in a time window of $\tau = 50$ ms (squares), 100 ms (triangles), or 200 ms (stars). The accuracy is defined as the average deviation of the estimate from the actual ITD of the stimulus. The horizontal line at 5 μs is the boundary given by the behavioural performance of 1-2 degree of azimuthal angle [12]. For large N, the accuracy scales as $1/\sqrt{\tau\, N}$ as expected from the central limit theorem.

METHODS: We took N neurons $1 \leq k \leq N$ with ITD tuning curves similar to Fig. 3a, approximated by $f_k(x) = \{164 + 36\cos[(x - x_k)2\pi/T]\}$ Hz where x is the ITD and T the period of the stimulation, x_k is the optimal ITD of neuron k and f_k its mean firing rate in Hz. The values of x_k have a uniform distribution between 0 and T. Spikes of neuron k are generated by a Poisson process with mean rate f_k. In an additional simulation (data not shown) we have confirmed that Poisson statistics give a reasonable assumption and the results of Fig. 4 do not change, if we generate spikes by a model neuron as in Fig. 1. The ITD of the stimulus is estimated as [26] $x^{\rm est} = (T/2\pi)\arg[\sum_k n_k\, e^{i2\pi x_k/T}]$ where n_k is the number of spikes of neuron k measured in a time window of length τ. We plot the accuracy $\langle |x^{\rm est} - x|^2\rangle^{1/2}$ averaged over 20 000 trials.