# Synaptic Plasticity Determines the Character of Interaural-Time-Difference Representation

Christian Leibold and J. Leo van Hemmen
Physik Department, Technische Universität München,
D-85747 Garching bei München, Germany

## Abstract

Neuronal activity of cells in the first binaural brainstem nucleus, in mammals the MSO, depends on interaural time differences (ITDs). The neurons thereby resolve disparities of about $10\,\mu$s. Based on a computational model for the development of synaptic couplings, we give restrictions for a recently proposed coding hypothesis based on a monotonic dependence of the firing rate upon the ITD. We find that synaptic plasticity at MSO neurons with high best frequencies favours a rate-place code with a high variability of the tuning curves, as in the classical Jeffress model, whereas low best frequencies give rise to monotonic ITD dependence.

**Keywords:** Spike-timing-dependent synaptic plasticity; Jeffress model; Psychophysical modeling; Sound localization

## INTRODUCTION

At a neuronal level, azimuthal sound localization operates through determining and representing (coding) interaural time differences (ITDs). Recently McAlpine et al. [10] have proposed an ITD representation mechanism for low-frequency cells. In contrast to the classical model of Jeffress [5], it is not the ITD at which a neuron fires at maximal rate and in this way represents stimulus direction but instead it is the *amount of firing activity in an*

*assembly* that encodes the direction. Low rates thereby represent ipsilateral positions, high rates contralateral ones. The number of action potentials in a population of cells is thus a reliable measure of the actual stimulus azimuth $\varphi$. A prerequisite for this so-called *rate-gradient code* is that the tuning curves of all cells have a maximum at similar contralateral leading ITDs.

We have simulated a spike-timing-dependent learning algorithm [1] with $1/f$-periodic input for excitatory synapses, which suffices to explain ITD tuning curves at a sub-millisecond time scale, and find that the resulting distribution of best interaural phase differences (IPD) varies in dependence upon input frequency. Here IPD is defined as ITD times $2\pi f$ where $f$ stands for frequency. Low-frequency inputs yield synaptic delay distributions that are in accordance with the hypothesis of McAlpine et al. [10] whereas high-frequency inputs favor a Jeffress-like interpretation [5]. As a result we provide a framework relating the coding concept to biological restrictions that are typical to the animal, such as head diameter and the evolutionary necessity of lateralization acuity.

## PSYCHOPHYSICAL MODEL

McAlpine et al. [10] suggested that the amount of neuronal activity within one population of cells accounts for the sensation of a sound-source azimuth. Since a brain is symmetric, there are two inverse rate representations corresponding to the two hemispheres. The simplest assumption possible is that the difference $\Xi(\varphi)$ between the activities of right and left hemisphere is the key quantity for perception. We therefore approximate the rate difference $\Xi$ by a linear function of the ITD $\tau$ so that $\Xi(\tau) = \gamma\,\tau$. This turns out to be a good estimate, at least in the range of naturally occurring ITDs. The respective ITDs depend on the sound speed $c_S$ and the head diameter $d$ through $\tau(\varphi) = [\varphi + \sin(\varphi)]\,d/(2c_S)$.

In order to test the above hypothesis regarding $\Xi(\varphi)$, we have modeled a simple psychoacoustic two-alternative forced-choice experiment where a subject has to discern two possible sound sources at angles $\varphi$ and $\varphi+\delta\varphi$. Defining a discrimination criterion that is optimal in a

sense of minimal decision error based on Gaussian statistics we obtain an analytic expression for the probability $P_c$ of correct answers, viz.,

$$P_c(\varphi, \delta\varphi) = \frac{1}{2}\Big[1 + \mathrm{erf}\Big(\frac{\gamma}{\sqrt{2}\Sigma}\frac{\tau(\varphi) - \tau(\varphi + \delta\varphi)}{2}\Big)\Big] \ , \tag{1}$$

where $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \ \mathrm{d}y \, \exp(-y^2)$ is the error function. The above formula relates the psychophysical quantity $P_c$ to the average tuning slope $\gamma$, a neuro-physiological model parameter, and $\Sigma$, representing the sum of two standard deviations, viz., of (i) the population rate $\Xi$ and (ii) a reading out of $\Xi$ by higher order brain stations. We define a system's time constant $t_S := \Sigma/\gamma$, which sets the psychophysical time resolution in (1) and serves as a fit parameter to psychoacoustical data. A comparison between theory and cat [3] as well as gerbil [4] data yields $t_S \approx 15\,\mu$s.

## LOCALIZATION ERROR AND BIOLOGICAL RESTRICTIONS

For a given threshold criterion for $P_c$ of getting $P_c > p$ for some given $p$ (generic values are $p = 0.75$ [3] or $p = 0.85$), Eq. (1) yields a localization error $E(\varphi, p)$ that is implicitly defined as $p = P_c(\varphi, E)$, meaning that for angles $\delta\varphi > E$ the probability of correct answers exceeds $p$. We can find an analytic expression for $E$, if we approximate the error function by a Fermi-distribution with approximate inverse $\mathrm{erf}^{-1}(x) \approx 2.406^{-1} \ln\left[(1+x)/(1-x)\right]$. In this way we obtain

$$E(\varphi, p) = \frac{t_S}{d}\,\frac{2.35\,c_S}{1 + \cos(\varphi)} \ln\left(\frac{p}{1-p}\right) \ . \tag{2}$$

See Fig. 1 for a plot of (2). Given a time resolution $t_S$, the error $E$ is inversely proportional to the head diameter $d$. If, on the other hand, both head diameter and an upper limit for $E$ are given, the latter incorporating the evolutionary necessity of lateralization acuity, we have obtained a rigorous restriction of $t_S$ based on the assumption of a rate-gradient code as suggested by McAlpine et al. [10]; see Introduction.

As an illustration of Eq. (2), we calculate the temporal resolution $t_S$ for a few species, provided they implement a rate-gradient code. The great titmouse [8], a small bird belonging

to the family *Paridae*, whose head diameter is $d \approx 2$ cm, has a temporal resolution of $t_S \approx 25\,\mu$s. Although barn owls have a head diameter $d$ that is about three times as large, their localization precision is so good ($E \approx 2^\circ$) [9] that temporal resolution should be as small as $t_S \approx 4\,\mu$s. Also humans have a rather high lateralization acuity [3]. At least for angles $\varphi \approx 0$ Eq. (2) yields a time constant $t_S \approx 8\,\mu$s.

The requirement of a high localization precision can thus be balanced by a short system time constant $t_S = \Sigma/\gamma$. Since the latter is defined to be the fraction of noise level $\Sigma$ and the slope of the population tuning $\gamma$, a small $t_S$ can be achieved by either reducing noise or increasing the steepness of tuning curves. If we roughly estimate the noise portion that is induced by the standard deviation of the population rate under the assumption of Poissonian spike generation we find that it is far too small in order to explain the psychophysically obtained system time constants $t_S$. The main portion of $\Sigma$ is therefore realized by the discrimination reliability of the system, i.e., how reliable the "rest" of the brain "reads out" the rate information provided by the cell population under consideration.

Being a question beyond today's knowledge, we refrain from a discussion of how to improve the reliability of reading out a rate-gradient code and concentrate instead on how to achieve steeper slopes $\gamma$. To this end, we again refer to the paper of McAlpine et al. [10], who find that tuning slopes increase with higher best frequency of the cell. A strategy of reducing localization error is thus the recruitment of higher-frequency components for ITD representation. The next section, however, will show that, for frequencies exceeding an upper bound, this strategy is incompatible with a rate-gradient code, if we regard tuning curves as originating from spike-timing-induced synaptic plasticity.

## SYNAPTIC DELAY SELECTION

In order to understand how ontogenetic development and selection of afferent delay distributions give rise to tuning curves that vary on a time scale of $10\,\mu$s, we propose that synaptic weights at binaural cells are subject to a spike-timing-dependent learning rule

[1]. Synaptic strengths are increased (decreased), if presynaptic spikes arrive before (after) the postsynaptic neuron fires. This temporal constraint is implemented by means of a learning window; cf. Fig. 2. Its amplitude is measured in units of the learning parameter $\eta = 1/4000$, adjusting the speed of synaptic change. Furthermore, we use a weight increment of $w^{\rm in} = \eta/50$ following every presynaptic spike and a decrement $w^{\rm out} = -\eta/4$ of all synaptic strengths after each postsynaptic spike. There is a lower bound 0 and an upper bound 2 for each individual synaptic weight. We assume that the width of the learning window in the *medial superior olive* (MSO) is in agreement with the membrane time constants [2] so that it is significantly smaller than that observed in tectal [11] neurons. This spike-based Hebbian learning rule leads to a normalization of the mean output rate of a postsynaptic neuron and a selection of *synaptic delays* in a competitive self-organized process [6].

We have simulated the temporal evolution of solely excitatory synaptic efficacies using two populations (left, right) of input spike trains generated by periodically (period length $1/f$) fluctuating Poisson processes until the synaptic configurations at the integrate-and-fire units had reached a stable state [1]. Firing rates are then obtained by averaging over 10 s. The maximum of a tuning curve defines the best phase difference. For each out of a given set of input frequencies $f \in \{1.0, 1.1, 1.2, 1.3, 1.4, 1.5\}$ kHz, we have performed several runs with changing random seeds and plotted the distribution of best interaural phase differences (IPDs); see Fig. 3 a.

The initial distribution of synaptic weights was Gaussian with mean 1.0 and standard deviation 0.3. A transmission delay from ear to synapse is assigned to each weight. The distribution of transmission delays is also taken to be Gaussian with standard deviation 0.3 ms and mean 1 ms. The best IPD of $\pi/4$, as found by experiment [10] is achieved by an appropriate shift of the synaptic delays after learning; see Discussion. During learning, IPDs have been drawn randomly every 100 ms of formal time from an interval corresponding to ITDs in the range of $\pm 0.2$ ms. In order to quantify the distribution of best ITDs $\Phi_k$, $1 \leq$

$k \leq K$, we define the population vector strength $r$ to be

$$r = \left| \sum_{k=1}^{K} \exp(i\Phi_k) \right| . \tag{3}$$

Figure 3 b shows a phase transition in dependence upon the input frequency. Low frequencies yield narrow distributions with best IPDs in a small range whereas distributions stemming from simulations with higher input frequencies correspond to randomly drawn best IPDs. As a consequence, the rate-gradient code of Ref. [10] breaks down as high-frequency spectral components are used for the neuronal ITD representation, and a Jeffress representation [5] takes over.

## DISCUSSION

We have shown that a rate-gradient representation of interaural time differences is restricted to the low-frequency range $f < f_{\max} \approx 1.5$ kHz and that, as long as this representation is valid, the minimal localization error has a lower bound given by (2), with $t_S \propto 1/f_{\max}$. Synaptic development at higher input frequencies spawns a broad distribution of best IPDs and thus naturally favors a rate-place coding concept as proposed by Jeffress.

The asymmetry of tuning curves with preferred best IPDs at about $\pi/4$ contralateral leading phase is realized by shifting the means of the delay distributions originating from the contralateral ear to longer delays after learning. In MSO cells this tuning asymmetry can be induced by synaptic plasticity of *inhibitory* synapses.

# REFERENCES


[1] W. Gerstner, R. Kempter, J.L. van Hemmen and H. Wagner, A Neuronal Learning Rule for Sub-Millisecond Temporal Coding, Nature 383 (1996) 76–78.

[2] B. Grothe and D.H. Sanes, Synaptic Inhibition Influences the Temporal Coding Properties of Medial Superior Olivary Neurons: An *in vitro* Study, J. Neurosci. 14 (1994) 1701–1709.

[3] R.S. Heffner and H.E. Heffner, Sound localization acuity in the cat: Effect of azimuth, signal duration, and test procedure, Hearing Res. 36 (1988) 221–232.

[4] R.S. Heffner and H.E. Heffner, Sound Localization and Use of Binaural Cues by the Gerbil, Behavioral Neurosci. 102 (1988) 422–428.

[5] L.A. Jeffress, A Place Theory of Sound Localization, J. Comp. Physiol. Psychol. 41 (1948) 35–39.

[6] R. Kempter, W. Gerstner, and J.L. van Hemmen, Intrinsic Stabilization of Output Rates by Spike-Based Hebbian Learning, Neural Comput. 13 (2001) 2709–2741.

[7] R. Kempter, C. Leibold, H. Wagner and J.L. van Hemmen, Formation of Temporal-Feature Maps by Axonal Propagation of Synaptic Learning, Proc. Natl. Acad. Sci. USA 98 (2001) 4166–4171.

[8] G.M. Klump, Sound localization studies in non-specialized birds, in: G.M. Klump, R.J. Dooling, R.R. Fay and W.C. Stebbins, eds., Methods in Comparative Psychoacoustics (Birkhäuser Verlag, Basel/Switzerland, 1995) 171–182.

[9] E.I. Knudsen, G.G. Blasdel and M. Konishi, Sound Localization by the Barn Owl (Tyto alba) Measured with the Search Coil Technique, J. Comp. Physiol. 133 (1979) 1–11.

[10] D. McAlpine, D. Jiang and A.R. Palmer, A Neural Code for Low-Frequency Sound Localization in Mammals, Nat. Neurosci. 4 (2001) 396–401.

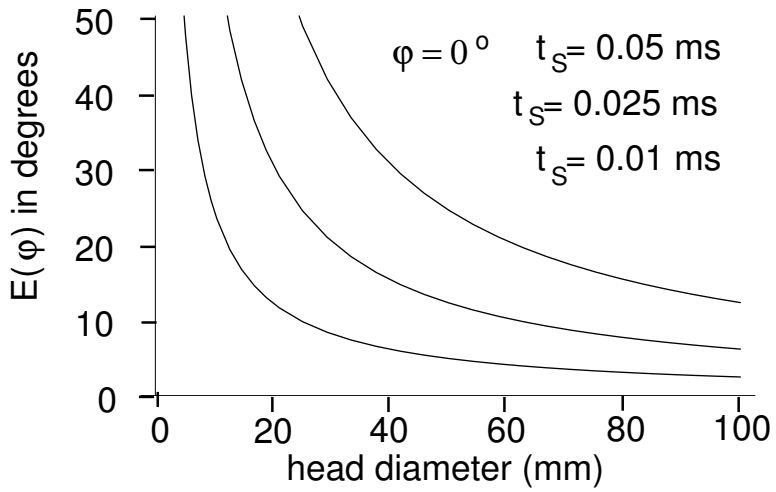
[11] L.I. Zhang, H.W.Tao, C.E. Holt, W.A. Harris and M.-m. Poo, A Critical Window for Cooperation and Competition among Developing Retinotectal Synapses, Nature 395 (1998) 37–44.

## FIGURES

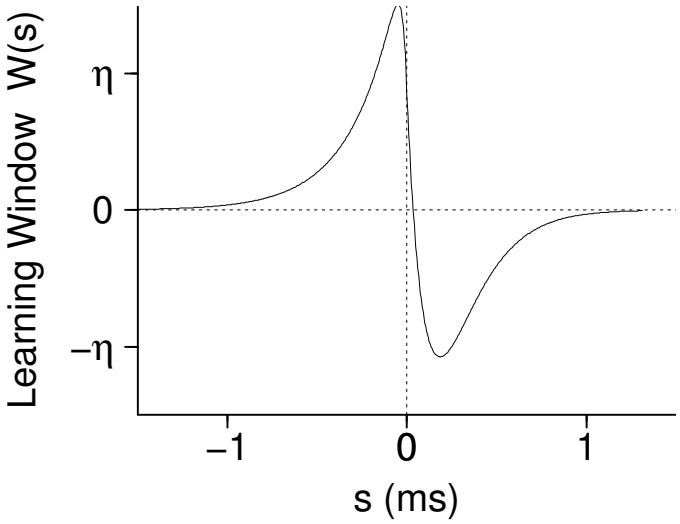
FIG. 1. Localization error $E(\varphi = 0, p = 0.75)$ from Eq. (2) as a function of head diameter $d$ for various values of $t_S$ (top graph $t_S = 50\,\mu$s, middle $t_S = 25\,\mu$s, bottom $t_S = 10\,\mu$s).

FIG. 2. Changes of excitatory synaptic strengths are governed by the time difference $s$ between pre- and postsynaptic spike, $s < 0$ meaning that a presynaptic spike arrives before the postsynaptic neuron fires. The learning window $W$ represents the idea of time-dependent synaptic plasticity [1] in dependence upon $s$; parameters have been taken from [7].

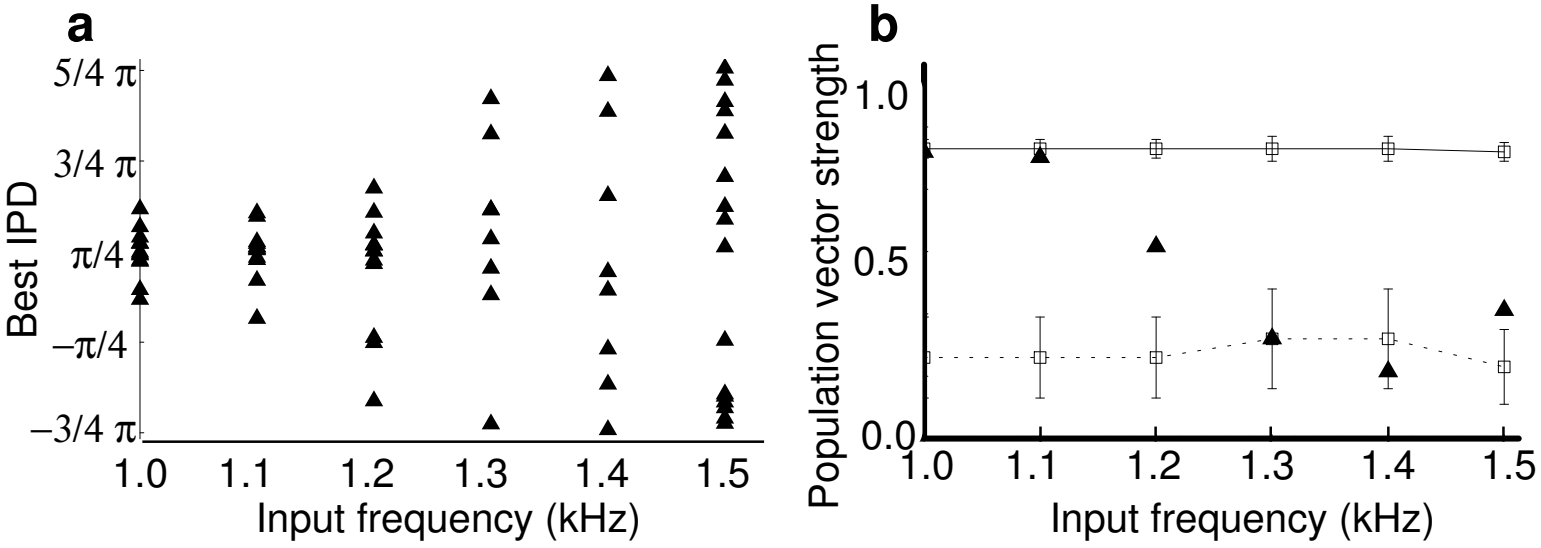
FIG. 3. (a) The distribution of best interaural phase differences (best IPDs) depends on the input frequency. (b) Population vector strength (solid triangles) due to Eq. (3). For frequencies below 1.2 kHz, the vector strength corresponds to the hypothesis of a best-IPD distribution with mean $\pi/4$ and standard deviation as for input frequency 1.0 kHz (solid line). For frequencies beyond 1.2 kHz, the hypothesis of randomly distributed IPDs (dashed line) is more suitable. Means (squares) and standard deviations are obtained through 20000 repetitions of the respective stochastic experiment.




Leibold, Fig. 1




Leibold, Fig. 2




Leibold, Fig. 3