Stochastic Interpolation Model of the Medial Superior Olive Neural Circuit

Pavel Sanda[a,c,d,*], Petr Marsalek[b,c,d]

[a] *Institute of Physiology, Academy of Sciences of the Czech Republic, Videnska 1083, 142 20, Praha 4, Czech Republic*
[b] *Institute of Pathological Physiology, First Medical Faculty, Charles University of Prague, U Nemocnice 5, 128 53, Praha 2, Czech Republic*
[c] *Max Planck Institute for the Physics of Complex Systems, Noethnitzer Str. 38, 011 87, Dresden, Germany*
[d] *Faculty of Biomedical Engineering, Czech Technical University of Prague, Nam. Sitna 3105, 272 01, Kladno, Czech Republic*

The final version of this article has been published in Brain Research (doi:10.1016/j.brainres.2011.08.048).

Abstract

This article presents a stochastic model of binaural hearing in the medial superior olive (MSO) circuit. This model is a variant of the slope encoding models. First, a general framework is developed describing the elementary neural operations realized on spike trains in individual parts of the circuit and how the neurons converging onto the MSO are connected. Random delay, coincidence detection of spikes, and divergence and convergence of spike trains are operations implemented by the following modules: spike generator, jitter generator, and coincidence detector. Subsequent processing of spike trains computes the sound azimuth in the circuit. The circuit parameters that influence efficiency of slope encoding are studied. In order to measure the overall circuit performance the concept of an ideal observer is used instead of a detailed model of higher relays in the auditory pathway. This makes it possible to bridge the gap between psychophysical observations in humans and recordings taken of small rodents. Most of the results are obtained through numerical simulations of the model.

Key words: coincidence detection, directional hearing, interaural time delay, sound azimuth, interpolation model

*Corresponding author. Tel. +420 241 062 733. Fax. +420 241 062 488.
Email addresses: sanda@biomed.cas.cz (Pavel Sanda)

Abbreviations: AN, auditory nerve; AP, action potential; EE, excitatory-excitatory; EI, excitatory-inhibitory; EPSP, excitatory post-synaptic potentials; IPSP, inhibitory post-synaptic potentials; ITD, interaural time difference; MSO, medial superior olive.

1. Introduction

A model of the mammalian MSO circuit which localizes the direction of the sound source in the horizontal plane (azimuth) at low frequency range is presented here. This model is based upon the **interpolation** of the direction in question using the firing rate code of the MSO binaural neurons. This is proposed as an alternative mechanism for low frequency sound localization in mammals, as the precise mechanism in mammals is not known to date (Dean et al., 2008; Joris et al., 1998, 2006; Karino et al., 2011). The **interpolation model** differs from the model proposing an **array of delay lines** (Jeffress, 1948). In the Jeffress model, the sound direction is encoded using a spatial code which includes the position of peak neural activity within the array. In the interpolation model, the binaural neuron encodes the sound direction using its firing rate. Firing of this neuron constitutes the ITD (interaural time difference) curve. The sound direction is read out by interpolation from this curve.

McAlpine et al. (2001), Brand et al. (2002), Grothe (2003) and others have proposed an ITD readout mechanism which encodes the sound direction through the maximal slope of the response function. This function has the ITD as its argument and the firing rate as its output. In line with the terminology used previously, we can call this mechanism **slope encoding**. Our interpolation model is a variant of this slope encoding mechanism. One of the crucial experimental findings is that abolishing of inhibitory inputs by strychnine moves the maximum slope out of the operational range (Brand et al., 2002). Based on this observation, the slope encoding model has, as one of its key elements, an inhibitory input to the MSO.

Our model reproduces the MSO computations at the level of abstract neural circuit description. Biophysical interactions of the excitatory and inhibitory postsynaptic potentials (EPSPs and IPSPs) at the postsynaptic membrane are not modeled in detail. Concerning the interaction of the IPSP with the EPSP, their time order is important. The IPSP must precede the EPSP to exceed the threshold (Grothe, 2003; Pecka et al., 2008). The PSPs

from the two sides have to meet within a given coincidence detection time window. The window size is one of the key parameters in our model. We stress this aspect of the model because the rise and decay time constants of postsynaptic voltage dependent currents determine the coincidence detection window size. The subcellular details are abstracted in this paper. *In vitro* experiments conducted in the search for the subcellular mechanism responsible for the sub-millisecond precision show the plurality of underlying biophysical mechanisms. To name only two of these one particular mechanism which contributes to setting the duration of the coincidence detection window is implemented by the voltage-activated potassium current at the postsynaptic membrane (Mathews et al., 2010). Another mechanism by which the order of the IPSP and EPSP is distinguished is the post-inhibitory rebound mechanism (Brand et al., 2002).

We also use the ergodic assumption of indistinguishable time and population ensemble encoding of the signals by spike trains. The ergodic assumption in physics was originally introduced in statistical mechanics of gases by Boltzmann in the nineteenth century. According to this assumption the average of a selected quantity is the same for the time average and for the statistical ensemble average (Papoulis, 1991). In other words, it assumes that the same result is obtained when observing a process over a long period of time as when sampling a given number of shorter independent realizations of the same process. This approach is applied in the neural coding description as follows. To receive information about the sound source direction, one can either wait for a given period of time and measure all the spikes in that period on an individual neuron, or one can observe a given number of neurons over a shorter time period and achieve the same precision. The number of spikes required for such a measurement was estimated using the Bernoulli process statistical model in Marsalek and Lansky (2005).

2. Methods

We propose a simple model where the basic processing stages of the neural circuit have been replaced by four modules operating in succession (see Fig. 1): a spike generator that represents the input from the auditory nerve (AN), a jitter generator that adds jitter and delays, occurring in the pathway up to the MSO, a coincidence detector that represents the first binaural neuron in the MSO, and an ideal observer module.

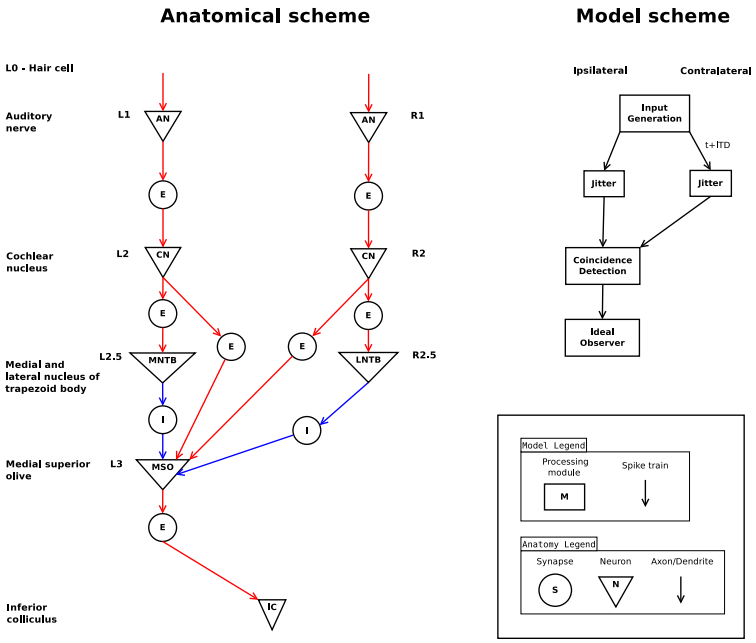

Figure 1: **Schemes of the MSO circuit anatomy and the interpolation model. The anatomical scheme** (left) shows branches of the binaural circuit (only one half of the symmetrical circuit is shown). L1, L2, ..., R1, R2 are left and right neurons. The direction from the auditory periphery to the auditory center is oriented top-down. Each neuron is denoted by its nucleus abbreviation (expanded in the left-hand column). Synapses are either excitatory (E) or inhibitory (I). **The model scheme** (right) employs neural circuit simplifications without losing the implementation of the key mechanisms.

2.1. Spike Generator

In some simulations a constant frequency spike train was used as the input, with two spikes generated in each cycle – one to one side, the other delayed by the ITD (this variable is denoted by D_{ITD}) to the other side.

While the regular spiking input is suitable for analysis of circuit parameter properties, a more biologically plausible model of spike generation was also included. Specifically, the first part of the computational model of the auditory periphery as described in Meddis (2006) was used. Here, the complex acoustic stimuli are passed through successive stages of processing: stapes, basilar membrane, inner hair cell receptor potential, presynaptic calcium currents and transmitter release events, AN (auditory nerve) synapse and finally AN spiking response, including refractory effects. The output – spike train in the auditory nerve fibre with selected characteristic frequency – is sub-

sequently used as an input to our circuit. An extensive overview of all the parameters can be found in the appendix under Meddis (2006). This model was then used without further modifications.

2.2. Jitter Generator

In the next processing stage, jitter is generated and added to the spike times, simulating natural noise occurring in the neural spike transmission. Each spike is shifted in time by a small random delay value defined as $D = T_J(B(a,b) - 0.5)$, where $B(a,b)$ is a random variable from the beta distribution with parameters $a = 2$, $b = 4$ and T_J is the timing jitter magnitude. Values of $a = 2$ and $b = 4$ are the smallest nontrivial values resulting in a nonzero skewness of the beta probability density function, as discussed in Drapal and Marsalek (2010). Note the positive role of this type of noise in the circuit – it is because of this noise that more refined azimuth precision is achieved.

2.3. Coincidence Detector

Whenever two spikes are detected within the short time window, the coincidence detector neuron fires. In the classical Jeffress (1948) model, coincidence detection is realized by the excitatory spikes converging on a binaural neuron from both sides simultaneously in a short time succession. The mere fact of coincidence determines the azimuth defined by the position of the neuron in a delay line. The interpolation model also introduces a short time window for detecting consequent spikes, while distinguishing between two EPSPs and a pair of one EPSP and one IPSP. The preceding application of the timing jitter is of essential importance to the circuit here as it allows a more subtle distribution of coincidences. Instead of a binary answer about two spikes being in the coincidence window, a probabilistic variable is obtained which, in turn, allows a finer distribution of recognized ITD values. Compared to the "all or none" output of the Jeffress model, this means that one single binaural neuron is already capable of determining D_{ITD}. However a longer spike train is needed to achieve the necessary precision of the estimate.

The important property of the ITD curve is its peak shift as observed in an experimental study of mammals (Brand et al., 2002). First, this leads to a shift of possible ITD values inside the physiologically relevant range. Second, it allows the unambiguous interpolation of a particular firing rate to a single ITD value. This is in contrast to the situation where the ITD curve has a

peak at $D_{ITD} = 0$ and the concave shape allows two possible ITDs to be selected. We show that in order to achieve this ITD curve shift, we need an "asymmetric" rule of EPSP and IPSP coincidence detection, allowing incoming spikes to trigger an output spike only if the inhibitory spike precedes the excitatory spike. This corresponds to the detailed biophysics of the PSPs. Pharmacological suppression of the inhibition is equivalent in our model to disabling this asymmetric rule and letting the symmetric coincidence detection operate on the incoming EPSPs only. By "asymmetry" we mean that the order of excitatory and inhibitory spikes matters — in contrast to the excitatory only situation.

2.4. Ideal Observer Module

Compared to the Jeffress model where the azimuth is determined by the position of the firing neuron in the delay line, the interpolation model relies on the value of a firing rate. The firing rate is realized by random variations in the input spike timings, and one way to eliminate this random component is to take an average of the firing rates over a given time. Because the input sound sequences may only last for a short period of time, we would like to know what length of time intervals are needed in order to obtain an azimuth estimate. Our aim here is not to model the higher stages of spatial sound localization, but rather to estimate the information that is available after the MSO processing stage and to connect the physiologically relevant information (firing rate) with experimental azimuth measurements.

The concept of the ideal observer, originally also known as the ideal receiver, is frequently used in the signal detection theory and psychophysics (Tanner Jr., 1961). By implementing the ideal observer module, we used this concept to describe a mechanism which represents the function of the next processing stages of the auditory pathway. It should be noted that this is not strictly an ideal observer analysis. Unlike previous modules, this module does not process the data. Instead, it simply extracts the available information encoded in the output spike train from the binaural neuron, and in particular it reads the sound source direction from the MSO neuron. This is done via the interpolation of the firing rates in the ITD curve. This lookup (calibration) function must be known beforehand and it can be approximated by a sine function $R_{ITD} \approx R_{max} \sin 2\pi f_{in} D_{ITD}$. The meaning of the parameters of this approximation are listed in the Results section.

Since the ideal observer has complete access to perceptual information, the azimuth perception can be obtained. In psychophysics this precision is

called the just noticeable difference, and in azimuthal sound localization in the head-on direction is 4° (angular degrees). This has been observed within a decade variation amongst distant animal species, in humans, Mills (1972), guinea pigs, McAlpine et al. (1996) barn owls, Moiseff and Konishi (1981) and even in parasitic flies *Ormia ochracea,* Miles et al. (1995). Here, different azimuths are distinguished as distinct points on the ITD curve along its slope. Thanks to interpolation the just noticeable difference is obtained from the difference between two spike train firing rates elicited by two stimuli (azimuths) φ_1 and φ_2. The firing rates of spike trains in response to stimuli φ_1 and φ_2 denoted as R_1 and R_2 can be described as random variables with means μ_1 and μ_2 and standard deviation σ (we assume $\sigma_1 = \sigma_2$). Detection distance is defined as: $d = (\mu_2 - \mu_1)/\sigma$. When $t \to \infty$, firing rates converge to their means, $R_1 \to \mu_1$ and $R_2 \to \mu_2$. The computational precision performance of the circuit is dependent upon speed of this convergence. Thus we measure the time to reach the direction estimate T_A. This variable is mainly dependent on the firing rate R_{ITD} and input sound frequency f_{in}, as well as other parameters.

2.5. Simulations and Parameters

The model circuit was simulated numerically with exploration of plausible parameter space. Each of the circuit modules has its own set of parameters which together define the circuit parameters. A default value of jitter of $T_J = 1$ ms was assigned since the experimentally observed values of jitter do not exceed 1 ms as shown for example in the experiment involving the octopus type cells of the cochlear nucleus in a domestic cat (Oertel et al., 2000). The width of the of coincidence window w_{CD} has been assigned as 600 μs. This is in agreement with the value used in a model part of Beckius et al. (1999). Briefly, their model performed well as long as the coincidence detection window was shorter than the period of the fundamental frequency. (We get $1/w_{CD} = 1667$ Hz, therefore it must be $f_{in} < 1667$ Hz.) This latter assumption is also frequently mentioned in other models. The subcellular experimental studies also measured values of relevant postsynaptic current- and voltage- dependent ionic current rise time. The values measured give window width in the sub-millisecond range close to this value (Jercog et al., 2010). Representative fundamental sound frequency in the following results is $f_{in} = 140$ Hz. Besides the fundamental frequency of spike generator f_{in}, we set the duration of most simulations to 500 s in order to reach a steady state. We select the parameters above as a **basic set**, summarized in Table

Parameter	Value
T_J	1 ms
w_{CD}	600 μs
f_{in}	140 Hz
T	500 s

Table 1: The basic set parameters.

1. An extensive list of circuit parameters can be found in the Supplementary Information. In all the simulations we use this basic set of the parameters unless stated otherwise.

The physiological range of D_{ITD} in humans when localizing sound source in one quadrant (0 - 90°) ranges from 0 to 600 μs. This can be calculated from the sound speed in the air and the distance between the ears. Differences in the ear distance of smaller animals have to be taken into account when interpreting the data recorded in various laboratory animals.

3. Results

3.1. Effects of Inhibition

For the basic set of parameters, we obtained the ITD curve shown in Fig. 2. It was checked that disabling the asymmetric rule in the coincidence

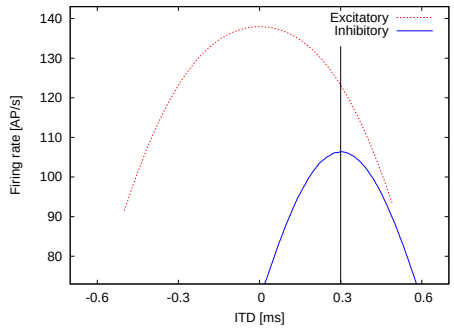

Figure 2: **ITD curves for the basic set of parameters.** This shows a detailed view of the peak with inhibition shifted out of the point, where $D_{ITD} = 0$ as compared to the peak without inhibition. The relative distances of the two curves (30-50 Hz) are given in Grothe (2003). Further, the EE interaction has a higher maximum firing rate than the EI interaction. The basic ITD curve has shift of 0.25 ms, which is an emergent property with no free parameter tuning.

detector affects the output of the circuit. We consider this to be our counterpart of inhibition suppression in the physiological model (see Section 2.3). The result was as expected: in the purely excitatory case, i.e., with no inhibition, there is a higher firing frequency and, due to the symmetric processing of excitatory spikes, the curve peaks at $D_{ITD} = 0$. In all other simulations we used the asymmetric coincidence detection rule only.

3.2. Effects of Varying Time Parameters

After testing the basic parameter set, let us move to exploring the parameter space, beginning with the timing jitter and input sound frequency. When the ITD curves for increasing input frequency f_{in} are generated, it can be observed that the firing rate grows and the ITD curve changes from flat in lower frequencies to steep in higher frequencies (Fig. 3, left). Similar rise of the firing rate with rising input frequency is observed with the use of more detailed Meddis model, with markedly lower output firing rates (Fig. 3, right). Furthermore, the distance between the ITD peaks becomes smaller with increasing f_{in} as the input sound period decreases. The curve offset around the $D_{ITD} = 0$ does not shift while changing f_{in}.

As jitter decreases, the firing rate in the ITD curves increases accordingly.

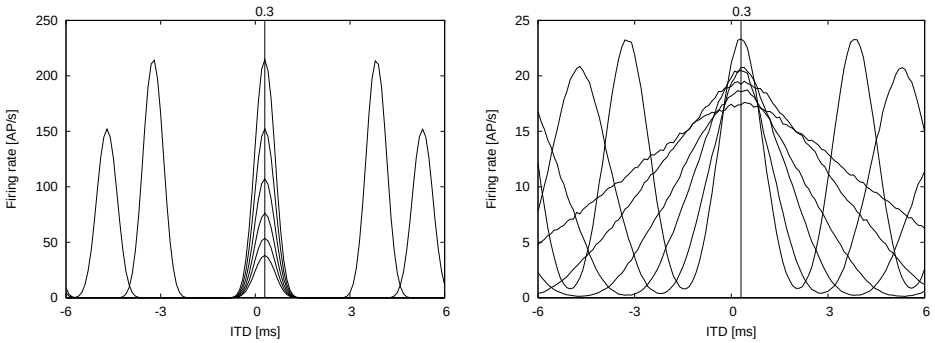

Figure 3: **ITD response curves for different input frequencies f_{in} directly delivered to the circuit (left-hand panel) and delivered through the Meddis model (right-hand panel).** For increasing the sound frequencies with step $\sqrt{2}$ (50, 71, 100, 141, 200, 283 Hz), the maximum firing also increases. In both cases, the period of the ITD curve decreases as the input frequency increases. If cochlear filtering is used, it can be seen that the width of a single peak increases as the input frequency decreases. Secondly, the output firing is approximately ten times lower with the Meddis model involved. For each sound frequency shown, the Meddis model was used with the same characteristic frequency. For cut-off frequencies see Supplementary Information.

See Fig. 4, right. This is due to the fact that the smaller jitter helps spikes to occur in the same coincidence window. For $T_J \rightarrow 0$, we would obtain a rectangular pulse, which is unusable for interpolation. If the jitter becomes larger than the sound period, this would render the response flat (not shown, effects of higher jitter values on ITD curve are thoroughly demonstrated by simulations in Supplementary Information). Thus, the feasible jitter values are bounded on both sides. Changes in jitter T_J values in the jitter generator do not affect the ITD curve shift.

3.2.1. Window Width and ITD Curve Shift

The ITD curve shift in our model is produced by the asymmetric coincidence rule but its particular size is affected by the parameters of the circuit. Therefore each parameter of our circuit was studied to determine which parameters influence the ITD peak shift. As a result of the simulations, it was possible to observe that the parameter that changes the peak shift considerably is the coincidence detection window width w_{CD}. See Fig. 4.

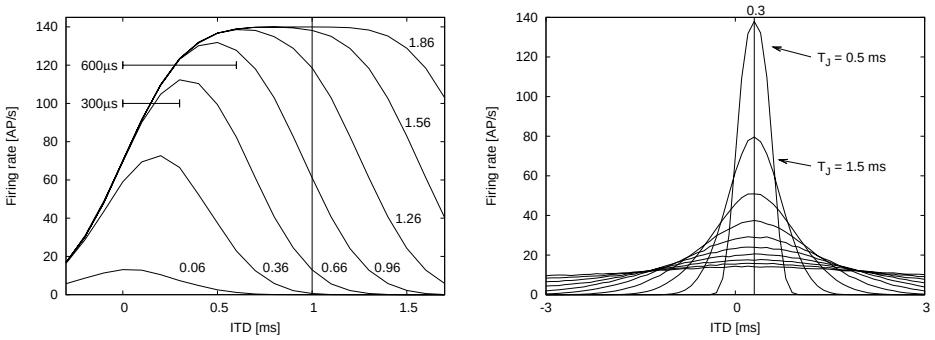

Figure 4: **Dependence of ITD peak shift on the coincidence detection window width w_{CD} (left-hand panel) and different values of jitter T_J (right-hand panel).** Peaks do not exceed 140 Hz due to the input frequency $f_{in} = 140$ Hz in the basic set of parameters. In the left-hand panel a shift of the peak towards higher ITDs can be seen as the window size increases from 60 μs towards 2 ms (with step 300 μs). The two top left intervals show physiologically relevant range in a domestic cat (300 μs) and in a human (600 μs), respectively. From around $w_{CD} = 0.6$ ms, and above that value, the maximum slope is observed and the shape of the ITD function in physiologically relevant range does not change. For further notes about physiological relevance of this observation see the Discussion section. Right-hand panel: Larger jitter causes lower output dynamic range and wider curves. The presence of high jitter values renders the response curve flat. The response increases as the magnitude of jitter decreases towards zero.

Other parameters of our model did not affect the peak shift significantly. See vertical lines in Fig. 3, 4.

3.3. Ideal Observer

The ideal observer module is connected to the output of the coincidence detector. For interpolation from the firing rate back to the D_{ITD} value a previously fitted ITD function is used. The ITD readout slope can be fitted to a sine function: $y = \frac{1}{2}A(1+\sin(Bt+C)+D)$, where A is the multiplicative constant setting up the maximum firing rate, B is dependent on the input frequency, C is the phase shift and D is the spontaneous activity. See Fig. 5.

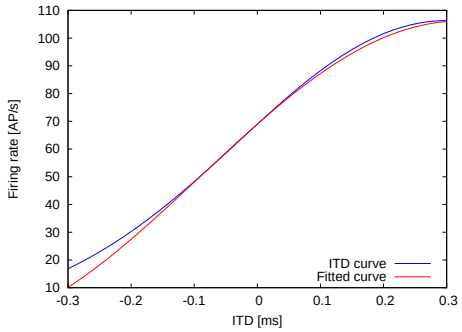

Figure 5: **The interpolation ITD curve in the physiologically relevant range of values.** The ITD curve is shown together with the fitted sinus function used for interpolation in the ideal observer. This function is fixed for the set of parameters which would correspond to the fixed tuning of the neural circuit for selected input signals. In this example, we used the ITD curve, which results from simulation of the circuit for the basic set of parameters. Note that the period of this slope corresponds to the fundamental sound period.

The shape of this function used for interpolation determines how precise the computation of D_{ITD} encoded in the incoming spike train is. As shown in previous sections, two features are important. The first feature is the unambiguous assignment of the unique D_{ITD} value from the firing rate. It is the asymmetric rule together with w_{CD} that plays an important role in determining the readout curve whose fit is shown in Fig. 5. The second feature is the slope of the interpolation curve obtained as the difference $\Delta_{ITD} = \Delta D_{ITD}/\Delta t$ and output range of the ITD curve. Both are decisive for accuracy (or speed) of ITD estimate by the observer. As the output range increases, we obtain a more precise azimuth estimate, which is influenced by all the parameters, with jitter T_J being the main factor, as shown above.

3.3.1. Asymptotic Times

Up to now we have looked only at the circuit output in the terms of the firing rate. At the end, the information contained in the output spike train is used in higher stages of the auditory pathway to estimate the azimuth. Since we do not model higher stages of auditory processing, we use the observer module, which sequentially collects information from the circuit output and predicts resulting ITD from the information given. Naturally, the longer the spike train is measured for, the more accurate the estimate of the source's D_{ITD} which can be assessed. Thus in order to investigate the overall estimate performance of our model circuit, we fed it with a longer input and estimated the time T_A needed for the ideal observer to achieve accuracy 4°, in the head-on direction (Mills, 1972). Fig. 6 shows fast asymptotic convergence towards the source D_{ITD} value when basic jitter value ($T_J = 1$ ms) is applied (left), and a comparison of the asymptotic times for different parameter sets (right). In all cases shown here we set $D_{ITD} = 0$. A comparison between

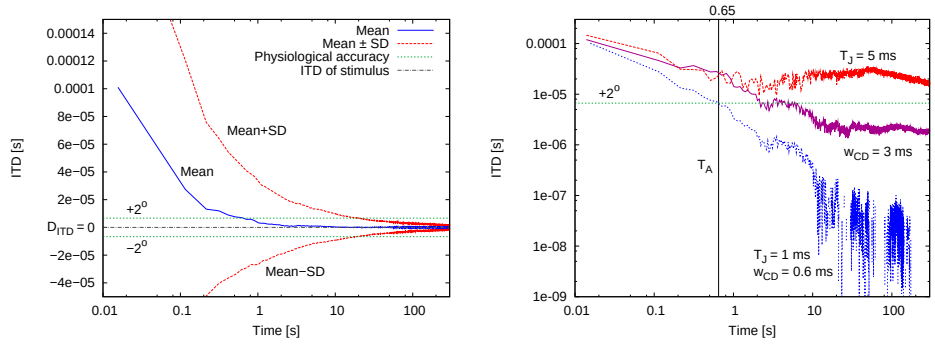

Figure 6: **Time required to reach a reliable estimate (with a given precision) of azimuth T_A by a single binaural neuron.** The x-axis shows time and the y-axis shows the ITD. Note the logarithmic scales used. In the left-hand panel, the middle solid line represents the **mean** value obtained for 1,000 simulation runs, while the dotted lines show standard deviation. The three dotted horizontal lines show D_{ITD} original source and margins of azimuth difference ±2° respectively. In the right-hand panel, the top black line shows mean convergence when using larger jitter ($T_J = 5$ ms), the middle line shows how convergence improves when using a wider window ($w_{CD} = 3$ ms), which is different from the basic parameter set, and the bottom line shows faster convergence for the basic set of parameters ($T_J = 1$ ms). The horizontal lines show the precision obtained in the psychoacoustical experiments. At the point in time T_A (T_A=0.65 ms for basic set of parameters) when the mean value crosses the precision line, enough information has been obtained for our model to be able to report the azimuth with average precision corresponding to that of human listeners ([-2°; 2°]).

a pure tone and a noisy input produced by the Meddis model was added to the Supplementary Information.

The time to reach the asymptotic values T_A, which range from 0.5 s up to 5 s (see Fig. 6), would be too long compared to the typical reaction time. The human minimum integration time is discussed in Middlebrooks and Green (1991) as a range of values: while a stimulus of 50 ms duration is detected with errors, typical azimuth integration time ranges from 150 to 300 ms of sound duration. However, the pathway from the AN to the MSO is highly parallel and consists of up to a few hundred neurons. When the ergodic hypothesis is used, as discussed in Section 4.4, T_A can be divided by the number of neurons. The time thus obtained corresponds to the theoretical minimum time required by human listeners to determine the D_{ITD}, when the entire brainstem circuit is working in parallel and cooperating on the ITD estimation. When the basic set of parameters is used we obtain that approximately 4 parallel neurons would be needed in order to achieve the 150 ms integration time.

4. Discussion

4.1. Type of Neuronal Coding

Jeffress (1948) proposed a theory of how sound is localized by a particular neural circuit. According to this concept, neural representation divides the whole azimuthal space into distinct sections. Each section is handled by a specific neuron and the fact that the sound originates from a given direction is represented by higher firing activity of the corresponding neuron. Jeffress suggested that this circuit is anatomically realized by an array of delay lines. Indeed, experimental evidence supporting this hypothesis was later found in the circuit of the nucleus laminaris of birds by Carr and Konishi (1988).

The responses of higher nuclei such as the colliculus inferior in both birds, Wagner et al. (1987), Wagner et al. (2007) and mammals, Joris et al. (1998, 2006) have been systematically studied until recently. This is because the colliculus inferior is more accessible to recordings and its binaural response is analogous to that of the nucleus laminaris or the MSO, respectively. In the MSO, which is a mammalian anatomical counterpart of the avian circuit, there is little supporting evidence for the existence of a structure similar to such an array of delay lines, Grothe (2003), but see also Joris et al. (1998). This raises the question of whether such a type of neural coding is also to

be found in the mammalian brain stem. Various alternatives have been proposed, for example, delays of the cochlear traveling wave (Shamma et al., 1989; Joris et al., 2006; Jennings and Colburn, 2010). Previous experiments have shown that ITD peaks occur out of the relevant physiological range (McAlpine et al., 2001) and further reproductions of mammalian experiments similar to those have led to the development of a theory showing that the mammalian localization mechanism is different from that of the avian. The importance of the role of synaptic inhibition was also experimentally discovered by Brand et al. (2002). The shift of the ITD curve peak is somewhat counter-intuitive when we assume that the Jeffress model is employed, and thus the slope (rate) coding model has been proposed instead. In this paper a circuit based on the slope coding scheme is studied.

4.2. Asymmetric Rule for Coincidence Detection

The excitatory and inhibitory effects in the sound localization circuit have been extensively discussed in experimental literature. Classical electrophysiological works by Goldberg and Brown (1969) and others have distinguished between the EE and EI type of binaural unit response. The EE (excitatory-excitatory) abbreviation labels units where the sound level increment at both sides has an excitatory effect, while EI stands for the excitatory effect of one side and inhibitory effect of the other side. However, this is a phenomenological classification and does not imply the synaptic mechanisms.

In the mammalian MSO circuit, other inhibitory inputs besides those from the MNTB (Medial Nucleus of Trapezoid Body) are involved. Each sensory pathway, including the auditory pathway, also contains local inhibitory interneurons and often other inhibitory circuits, like, for example the indirect pathway of the basal ganglia. In all these synaptic circuits, the role of inhibition is to dynamically set up the gain of the circuit via feedback connections. Other inhibitory synapses in the MSO circuit contribute to setting the gain, adaptation, and perhaps to further unknown functions of the circuit.

Our model follows anatomical connections of the feedforward sound localization pathway. Although throughout the paper excitatory and inhibitory synapses are referred to according to their anatomical and functional classification, we have shown that instead of synapse polarity, it is possible to use a timing rule for coincidence detection. Thus, we have been able to describe the circuitry independent of the synapses polarity in more abstract and simple terms. We have also demonstrated that our results are in line with the

experimental results related to inhibition blocking (Brand et al., 2002; Pecka et al., 2008).

4.3. Properties of the Circuit

Having studied the parameters that affect important properties of the ITD curve, we have found that the parameter significantly affecting peak position is the coincidence window width. One consequence of shifting the peak by employing a wider window is a steeper ITD function in the physiological range of ITDs. Bringing steepness to its maximum value would require $w_{CD} \approx 0.6$ ms which is higher than the observed values (Jercog et al., 2010). Using smaller jitter values proved much more effective in increasing the output range and steepness of the ITD curve. As noted in Section 2.3, on the one hand, jitter plays a positive role for more subtle differentiation of azimuthal space, but on the other hand, higher jitter reduces coding efficiency. This indicates the possible existence of an optimal jitter value which merits further research.

Asymptotic behavior of the sound direction "ideally observed" by the model was also investigated. This gives us overall precision and time performance for the whole circuit. It could be seen that an ideal observer located on a single fiber performs in the range of seconds, which is acceptable considering that the auditory pathway works in parallel using tens to hundreds of neurons. Using an ergodic assumption for the ideal observer, we obtain a hundreds of milliseconds time range. This allows us to connect physiological and psychoacoustical observations, which are normally two separate areas of research.

Besides the regular input, the Meddis (2006) model of the auditory periphery was used to compare the performance of the two input models. Generally, the Meddis model causes a fall in the firing rate delivered as the input to our model, compared to the regular spiking input we used. In addition, only phase-locked spikes will be registered by coincidence detection. This reduces both the available dynamic range of values of the ITD curve and the computational performance of the circuit. The question remains how closely this model corresponds to the real input to the MSO, since the model produces only the spike train occurring at the auditory fiber. The cochlear nucleus can improve the phase-locking due to summation of more AN fiber inputs.

The timing jitter also plays a role in deteriorating the phase-locking precision. Most of the auditory processing stages blur the phase locking. The

comparison with the Meddis model indicates that introducing more processing stages into the circuit does not necessarily blur the spike trains beyond the state when the circuit would no longer function and would no longer possess a functioning coincidence detection mechanism. Indeed, Reed et al. (2002) have shown that the phase locking in a model neuron within a neural circuit can deteriorate, can be sharpened, or can stay unaltered, depending upon the connection divergence and other simple properties of the neuron as a relay within the sensory pathway.

4.4. Ergodic Assumption

The performance of a single binaural neuron has been shown for several parameters and inputs. Time (T_A) to reach the estimate of D_{ITD} by single neuron is higher than in psychoacoustic experiments. It is known that the auditory pathway uses AN fibers and subsequent neurons in parallel so the brainstem contains several parallel circuits cooperating on the computation of azimuth. We used the ergodic hypothesis to divide T_A by the number of parallel neurons. This hypothesis used in interpretation of Fig. 6 states that the same result is obtained when counting spikes for a long time as when sampling them across a given number of parallel fibers. This is also consistent with the actual ITD processing in the MSO reported in several species, which produced the same ITD tuning curves by several repeats of sound stimulus, as shown, for instance by Goldberg and Brown (1969).

The exact number of circuits which work in parallel is difficult to estimate. It concerns not only the number of the binaural MSO neurons but also the number and connectivity of the spherical bushy cells of the cochlear nucleus. About 7000 of these cells connect to the MSO in a domestic cat. Since their axons branch, the average number of connections to the MSO is difficult to obtain with the current methodology (P.X. Joris, personal communication). Next, the exact spectral content of incoming acoustic stimuli affects the number of involved fibers due to the tonotopic organization. For example the broadband stimuli would be expected to perform better than pure tone stimuli due to higher number of parallel fibers transmitting the signal. All these problems were not studied deeper in the current paper but they pose interesting open questions for further work.

4.5. Stochastic coding of ITD

Finally, we would like to point out the use of the built-in random number generator in our model. By means of these stochastic simulations we have

obtained realizations of a random variable, namely of the timing jitter. The results obtained for the ideal observer are shown together with their statistical description. A previous paper by Marsalek and Lansky (2005) presented a stochastic model of a coincidence detector and theoretically calculated average detection times. In this paper a further step has been taken, as the entire MSO model circuit has been studied. The precision of the ITD detection assessed by our model requires further verification by future experiments employing psychoacoustical data.

Acknowledgments

Supported by research initiatives MSM 0021620806 and MSM 6840770012 by the Ministry of Education, Youth and Sports, by MPO FR-TI3/ 869 by the Ministry of Industry and Trade of the Czech Republic, by P103/11/0282 by the Grant Agency of the Czech Republic, and by the Max Planck Society. Thanks to Philip Joris, Petr Lansky, David McAlpine and Raymond Meddis for discussions and to Deborah A. M. James, Martina Missikova and Linda Jayne Turner for copy-editing.

References

Beckius, G. E., Batra, R., Oliver, D. L., 1999. Axons from anteroventral cochlear nucleus that terminate in medial superior olive of cat: observations related to delay lines. J Neurosci 19 (8), 3146–3161.

Brand, A., Behrend, O., Marquardt, T., McAlpine, D., Grothe, B., 2002. Precise inhibition is essential for microsecond interaural time difference coding. Nature 417 (6888), 543–547.

Carr, C. E., Konishi, M., 1988. Axonal delay lines for time measurement in the owl's brainstem. Proc Natl Acad Sci USA 85 (21), 8311–8315.

Dean, I., Robinson, B. L., Harper, N. S., McAlpine, D., 2008. Rapid neural adaptation to sound level statistics. J Neurosci 28 (25), 6430–6438.

Drapal, M., Marsalek, P., 2010. Stochastic model shows how cochlear implants process azimuth in real auditory space. Chinese J Physiol 53 (6), 439–446.

Goldberg, J. M., Brown, P. B., 1969. Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: Some physiological mechanisms of sound localization. J Neurophysiol 32 (4), 613–636.

Grothe, B., 2003. New roles for synaptic inhibition in sound localization. Nat Rev Neurosci 4 (7), 540–50.

Jeffress, L. A., 1948. A place theory of sound localization. J Comp Physiol Psychol 41 (1), 35–39.

Jennings, T. R., Colburn, H. S., 2010. Models of the Superior Olivary Complex. In: Meddis, R., Lopez-Poveda, E. A., Fay, R. R., Popper, A. N. (Eds.), Computational Models of the Auditory System. Springer, New York, pp. 65–96.

Jercog, P. E., Svirskis, G., Kotak, V. C., Sanes, D. H., Rinzel, J., 2010. Asymmetric excitatory synaptic dynamics underlie interaural time difference processing in the auditory system. PLoS Biol 8 (6), e1000406, 1–9.

Joris, P. X., Smith, P. H., Yin, T. C. T., 1998. Coincidence detection in the auditory system: 50 years after Jeffress. Neuron 21 (6), 1235–1238.

Joris, P. X., Van de Sande, B., Louage, D. H., van der Heijden, M., 2006. Binaural and cochlear disparities. Proc Natl Acad Sci USA 103 (34), 12917–12922.

Karino, S., Smith, P., Yin, T., Joris, P., 2011. Axonal branching patterns as sources of delay in the mammalian auditory brainstem: A re-examination. J Neurosci 31 (8), 3016–3031.

Marsalek, P., Lansky, P., 2005. Proposed mechanisms for coincidence detection in the auditory brainstem. Biol Cybern 92 (6), 445–451.

Mathews, P., Jercog, P., Rinzel, J., Scott, L., Golding, N., 2010. Control of submillisecond synaptic timing in binaural coincidence detectors by K(v)1 channels. Nat Neurosci 13 (5), 601–609.

McAlpine, D., Jiang, D., Palmer, A., 1996. Interaural delay sensitivity and the classification of low best-frequency binaural responses in the inferior colliculus of the guinea pig. Hear Res 97 (1-2), 136–152.

McAlpine, D., Jiang, D., Palmer, A. R., 2001. A neural code for low-frequency sound localization in mammals. Nat Neurosci 4 (4), 396–401.

Meddis, R., 2006. Auditory-nerve first-spike latency and auditory absolute threshold: A computer model. J Acoust Soc Am 119 (1), 406–417.

Middlebrooks, J. C., Green, D. M., 1991. Sound localization by human listeners. Annu Rev of Psychol 42 (1), 135–159.

Miles, R., Robert, D., Hoy, R., 1995. Mechanically coupled ears for directional hearing in the parasitoid fly *Ormia ochracea*. J Acoust Soc Am 98 (6), 3059–3070.

Mills, A. W., 1972. Auditory localization. In: Tobias, J. V. (Ed.), Foundations of Modern Auditory Theory. Academic Press, New York, pp. 303–348.

Moiseff, A., Konishi, M., 1981. Neuronal and behavioral sensitivity to binaural time differences in the owl. J Neurosci 1 (1), 40–48.

Oertel, D., Bal, R., Gardner, S. M., Smith, P. H., Joris, P. X., 2000. Detection of synchrony in the activity of auditory nerve fibers by octopus cells of the mammalian cochlear nucleus. Proc Natl Acad Sci USA 97 (22), 11773–11779.

Papoulis, A., 1991. Probability, Random Variables, and Stochastic Processes. McGraw-Hill, New York.

Pecka, M., Brand, A., Behrend, O., Grothe, B., 2008. Interaural time difference processing in the mammalian medial superior olive: The role of glycinergic inhibition. J Neurosci 28 (27), 6914–6925.

Reed, M. C., Blum, J. J., Mitchell, C. C., 2002. Precision of neural timing: effects of convergence and time-windowing. J Comput Neurosci 13 (1), 35–47.

Shamma, S., Shen, N., Gopalaswamy, P., 1989. Stereausis: Binaural processing without neural delays. J Acoust Soc Am 86 (3), 989–1006.

Tanner Jr., W. P., 1961. Physiological implications of psychophysical data. Ann NY Acad Sci 89 (5), 752–765.

Wagner, H., Asadollahi, A., Bremen, P., Endler, F., Vonderschen, K., von Campenhausen, M., 2007. Distribution of interaural time difference in the barn owl's inferior colliculus in the low-and high-frequency ranges. J Neurosci 27 (15), 4191–4200.

Wagner, H., Takahashi, T., Konishi, M., 1987. Representation of interaural time difference in the central nucleus of the barn owl's inferior colliculus. J Neurosci 7 (10), 3105–3116.