Stochastic model explains role of excitation and inhibition in binaural sound localization in mammals

M. DRAPAL[1], P. MARSALEK[1, 2, 3, 4]

[1] Department of Pathological Physiology, First Medical Faculty, Charles University of Prague, Czech Republic

[2] Faculty of Biomedical Engineering, Czech Technical University of Prague, Kladno, Czech Republic

[3] Institute of Physiology, Academy of Sciences of the Czech Republic, Prague, Czech Republic

[4] Max Planck Institute for the Physics of Complex Systems, Dresden, Germany

Corresponding author:

Petr Marsalek, Max Planck Institute for the Physics of Complex Systems,

Noethnitzer Str. 38, D-011 87, Dresden, Germany.

Tel: +49-351 871 1217 and +420-910 806 204. E-mail: marsalek@pks.mpg.de

Submitted December 30, 2009

Revision of June 15, 2010

Second revision of August 23, 2010

Short title: Excitation and inhibition in binaural sound localization

Key words: Coincidence detection, Directional hearing, Interaural time delay, Medial superior olive, Stochastic neuronal model.

Summary

Interaural time differences (ITDs), the differences of arrival time of the sound at the two ears, provide a major cue for low-frequency sound localization in the horizontal plane. The first nucleus involved in the computation of ITDs is the medial superior olive (MSO). We model the neural circuit of the MSO using a stochastic description of spike timing. The inputs to the circuit are stochastic spike trains with a spike timing distribution described by a given probability density function (beta density). The outputs of the circuit reproduce the empirical firing rates found in experiment in response to the varying ITD. The outputs of the computational model are calculated numerically and these numerical simulations are also supported by analytical calculations. We formulate a simple hypothesis concerning how sound localization works in mammals. According to this hypothesis, there is no array of delay lines as in the Jeffress' model, but instead the inhibitory input is shifted in time as a whole. This is consistent with experimental observations in mammals.

1 Introduction

The auditory system of humans and many other animals is able to localize sound sources with amazing precision. This ability is partially possible with only one ear (monoaural hearing), yet for localization in the horizontal plane two ears are necessary (binaural hearing). Sound source localization can be enhanced when the source and the receiver move relative to each other (Phillips and Brugge 1985). In this paper, however, we limit the task of sound localization to static sources using binaural cues at low frequencies. Binaural cues determine the azimuth defined as follows. The vector from a listener to a sound source is projected perpendicularly onto the horizontal plane. The angle between the projected vector and a reference vector, forming the intersection of the horizontal plane with the plane of head symmetry, oriented to the front, is called the azimuth. One of the parameters influencing binaural sound source localization is the fundamental sound

frequency. In mammals, for low fundamental frequencies (below 1,500 Hz) or for broadband sounds, the interaural time difference (ITD) is the dominant sound localization cue. For high fundamental frequencies, the interaural intensity difference (IID) is used. While this paper deals with lower frequency bands, it is possible that higher frequencies are processed with the use of similar neuronal algorithms, as we proposed in earlier studies by Marsalek and Kofranek (2004 and 2005).

This paper presents a theory of how binaural sound localization for low frequencies might be realized in mammals and particularly in humans. The theory of Jeffress (1948) is one of the first well-known attempts to explain how neuronal circuitry achieves this. His prescient work is still frequently cited (Joris *et al.* 1998). Jeffress' visionary hypothesis asserted that the ITD is converted to a binary signal in a higher order neuron through an array of delay lines of fibers in lower order neurons from both sides. Pioneering experiments by Carr and Konishi (1988) showed that Jeffress was correct in case of birds. As far as we know, the existence of an analogous delay line in mammals remains an open question (Grothe 2003, Joris and Yin 2007, McAlpine and Grothe 2003). What other neural circuit mechanism might be responsible for calculating the azimuth from the ITD? In this paper we propose an alternative to the delay line array model based on recent physiological evidence. This alternative is a stochastic delay of a very small number of broadly tuned channels (McAlpine and Grothe 2003).

The amazing time precision (Joris *et al.* 1998) in the range of tens of microsecond points towards another statement of Jeffress that the neurons of the circuit should be located among the lower order neurons of the auditory pathway. The lowest order suitable neuron is the first binaural neuron.

The information about the sound source location contained in the ITD is implicitly encoded by spike trains of lower order neurons. The first binaural neurons function as encoders of the ITD. The circuit has to make the information accessible, in other words make it explicit within another spike train in higher level neurons of the auditory pathway. The function of the circuit is to convert the information implicit in the ITD into the explicit neural code for the ITD. The definition of implicit and explicit neural coding can

be found in Koch (2004). The binaural neurons of the circuit can function either as a delay line (Jeffress 1948), or as broadly tuned channels (McAlpine and Grothe 2003).

Our model is based on one time delay in the neuronal circuit. The access to a continuum of responses to various azimuth locations is accomplished through stochastic variations of action potential times as processed by the model circuit. Our novel finding demonstrates that stochastic spike timing can be used by neurons as an instrument for computing the sound azimuth. Model circuit connections and properties after mathematical simplification of their connectivity are still consistent with the neuro-anatomical description of the wiring of the medial superior olive (MSO) circuit in mammals (Beckius *et al.* 1999, Young 1998, McAlpine *et al.* 2001).

We have designed and improved a model description of how the neural circuit in the auditory brain stem calculates the direction of incoming sound. This model is an alternative to the classical theory of delay lines. We present a stochastic description of the output spike train and spike timing within the model. Both the analytical calculations and numerical simulations give qualitatively similar results to those of experimental recordings from the rodent auditory brain stem. Our results are also comparable with recordings from brainstems in gerbils by Brand *et al.* (2002). Other authors (McAlpine *et al.* 2001) have found similar tuning curves in response to the changing ITD in different (higher) neurons of the auditory pathway (colliculus inferior) of another animal (guinea pig). We find the results to be robust with respect to variations of the time window size and spike timing jitter.

2 Model

2.1 Anatomical connections and their simplification

The notation of the mathematical formulation of the model follows conventions used in Marsalek and Lansky (2005) and Marsalek and Drapal (2008), where the excitatory-excitatory (EE) interaction is called excitatory coincidence detection (ECD) and the excitatory-inhibitory (EI) interaction is called inhibitory coincidence detection

(ICD). The medial superior olive (MSO) works mostly with low frequencies and the lateral superior olive (LSO) deals mostly with high frequencies.

Our model assumes different connections to the MSO neurons than those assumed in the Jeffress model. The division of neuronal fibers based on their excitatory and inhibitory effect is important. We show in the Results section that inhibitory fibers phase shift the tuning curve of ITD, as compared to that obtained without inhibition. The inhibitory connection to the MSO results from the inversion of synaptic polarity in the medial nucleus of the trapezoid body. The MNTB receives excitatory inputs from the contra-lateral cochlear nuclei and sends inhibitory inputs to the ipsi-lateral MSO. The same MSO also receives an inhibitory connection from the ipsi-lateral side. This inhibitory activity originates from the lateral nucleus of trapezoid body (LNTB), which receives its excitation from the ipsi-lateral cochlear nucleus. To complete the picture, the MNTB sends further inhibitory connections to the LSO. This intricate anatomy is nicely summarized by Young (1998).

Though the design of our model is based on knowledge of these anatomical connections, we have to simplify the model wiring to extract the functional core of the neural circuit. The delays in the real system are present in both ipsi- and contra-lateral pathways (Joris 1996, Beckius *et al.* 1999). For the purpose of simplification, however, relative delay on one side suffices. This relative delay represents the net delay difference. Furthermore, one inhibitory branch from one side is enough to model the net inhibition from both sides. This leads us to the schematics shown in Figure 1.

Note for editor: Insert *Figure 1* **around here.**

2.2 Operating conditions and constraints

Before proceeding with a formal description of the variables in the model, we should briefly mention the coincidence detector (CD). Without detection of the leading edges of incoming post-synaptic potentials, extraction of a signal from delays in the microsecond range would not be possible. Regardless of which of the two theories we propose, they all must use this microsecond precision. The element detecting the leading edge is called the coincidence detector. The anatomical substrate of the CD is believed to be within the

MSO neurons. The location of right and left CDs in our model circuit in Figure 1 does not reflect all the detailed connections found in anatomy, however. The algorithm realized by a single delay in the model circuit is computationally equivalent to the original circuit, regardless of the actual succession of coincidence detectors, delays and the polarity change from excitatory to inhibitory signals. For details see the Discussion section.

We can shuffle the order of delays among selected points in our model without loss of generality. This is based on observation that selected neural operations are commutative. An example of commutative additions of delays is shown in Equation (1) below. The first processing elements of our model are random delays, which have a specific probability density function (PDF) of synaptic input to neuron in time. Since there is chain of delays in both synaptic chains from the left and from the right ear, we can suppose that we have *n* ipsilateral delays in the ipsilateral (A) branch of the pathway $\Delta_{A1}, \Delta_{A2}, ..., \Delta_{An}$ and *m* contralateral delays in the contralateral (B) branch of the pathway $\Delta_{B1}, \Delta_{B2}, ..., \Delta_{Bm}$. For the EE interaction $m = 3$ and $n = 3$, for the EI interaction of the MNTB $m = 4$ and $n = 3$ and for the EI interaction of the LNTB $m = 3$ and $n = 4$. Numbers *m* and *n* include the first synapses (ribbon synapses from hair cells). The total delay difference between branches (A) and (B), denoted $\Delta_{JA} - \Delta_{JB}$ is given by:

$$\Delta_{JA} - \Delta_{JB} = \sum_{i=1}^{n} \Delta_{Ai} - \sum_{j=1}^{m} \Delta_{Bj}. \quad (1)$$

We assume that all these random delays with (timing) jitter (subscript J) on sides A and B (left and right) are mutually independent and identically distributed non-negative random variables. Δ_{JA} and Δ_{JB} have a maximum of Δ_{max}. The constraints imposed on them are given in Equation (2). The coincidence detection (time) window Δ_{W} must be shorter than or equal to the maximum delay and the sound period, *T*, must be greater than or equal to the maximum delay:

$$
\begin{aligned}
&0 \le \Delta_{JA} \le \Delta_{max}, \\
&0 \le \Delta_{JB} \le \Delta_{max}, \\
&0 \le \Delta_{W} \le \Delta_{max}, \\
&\Delta_{max} \le T.
\end{aligned}
\tag{2}
$$

In excitatory coincidence detection, the spike is generated only when the two spikes from sides (A) and (B) meet in a time interval shorter than Δ_W. In other words, the two spike delays Δ_{JA} and Δ_{JB} must satisfy:

$$
| \Delta_{JA} - \Delta_{JB} | \le \Delta_W . \tag{3}
$$

To model inhibitory coincidence detection, a modification of the condition expressed above in Equation (3) is used. Spikes must arrive in proper succession. The excitation from side A must come after the inhibition from side B. This is formulated as:

$$
0 \le \Delta_{JA} - \Delta_{JB} \le \Delta_W . \tag{4}
$$

Using the model with one delay, one obtains the continuum of detected ITDs through stochastic variation of the random interaural time delay $\Delta_{JA} - \Delta_{JB}$. The output, the spike rate, clearly depends on the choice of input PDF of the synaptic delays. After summarizing the properties of the model, connections between neurons, and constraints imposed on the parameters and random variables, we can turn to the outputs of the model in the following two subsections.

2.3 Input distribution of the coincidence detector

The output of the model is dependent on the proper choice of a PDF of random variables.

Firstly, the range of the PDF is defined such that its support is over one sound period. This is on the time interval [0, *T*], which we normalize to the interval [0, 1]. Therefore probability densities spanning one or both tails from minus infinity to plus infinity, such as gamma density, where its support is on [0, ∞), or normal density, where its support is on ($-\infty$, ∞), are not particularly useful. This justifies the choice of beta density, which is nonzero only within the range [0, 1] and is close to gamma density in this range as well. This circumvents the need to normalize the corresponding cumulative distribution function to unity and makes the calculation more transparent without loss of

generality. Another useful property of beta density is its simple polynomial definition. The formula for beta density appears in the Abbreviations and symbols section at the end of this paper. We used beta density for the description of spike timing distribution also in (Marsalek *et al.* 1997).

Secondly, the PDF shape, specifically the skewness and the kurtosis, influences the shape of the output tuning curve. We have experimented with both uniform and triangular densities, (Marsalek and Lansky 2005), which are special trivial cases of the beta density with parameters $a = 1$, $b = 1$, and $a = 1$, $b = 2$, respectively, and make calculation simpler. However, the corresponding output functions analogous to Equation (7) are not as satisfactory as the output resulting from a non-trivial beta density. These output functions are not shown here and the resulting function itself is discussed in detail at the beginning of the Results section.

Thirdly, the mean (output) activity of the model obtained with inhibition must not drop below the zero line. This is corrected by adding half the height of the span of the output range to the function. This way all output values are positive. They correspond to the neural spike rates which cannot be negative. After eventual normalization, so that the density integral over the whole function range equals unity, the output spike rates express output probability.

Note for editor: Insert *Figure 2* **around here.**

Finally, the input beta density has parameters $a = 2$ and $b = 4$ and we denote it B_{24}. To obtain a smooth function on the interval (0, 1) we must have $a > 1$ and $b > 1$, and to obtain nonzero skewness we must have $a \neq b$. The whole numbers $a = 2$ and $b = 4$ are the second smallest non-trivial values of parameters (after $a = 2$ and $b = 3$). Figure 2 shows the result of the analytical calculation of the output, which is the function denoted q_{24}.

2.4 Output distribution from the coincidence detector

The output function q_{24} is obtained as follows. Let us denote the difference of the two delays in Equation (1) as:

$$Z = \Delta_{JA} - \Delta_{JB}. \tag{5}$$

The probability density function of this new, compound random variable Z is obtained with a convolution integral formula for the difference of the two random variables. The PDFs of the Δ_{JA} and the Δ_{JB} are denoted $f(x)$ and $g(y)$, respectively, and the output PDF of the variable Z is denoted $q(z)$. We substitute $f = B_{24}$ and $g = B_{42}$ into the convolution formula $q(z) = \int_{-\infty}^{\infty} f(x)g(x-z)dx$. Now, since $B_{24}(x)$ and $B_{42}(y)$ are nonzero only for x between 0 and 1, we must evaluate the integral piecewise within the respective ranges, such that $q(z)$ becomes:

$$q(z) = \int_{\max(z-1,0)}^{\min(z,1)} f(x)g(z-x)dx = \int_{z\in[-1,0]} f(x)g(z-x) + \int_{z\in[0,1]} f(x)g(z-x). \tag{6}$$

Obviously, $q(z)$ is an even function, satisfying $q(z) = q(-z)$. Therefore, we can change the sign of the argument $z-x$ without loss of validity. We substitute the polynomial densities $B_{24}(x)$ and $B_{42}(y)$ into the integrals. For two 4th degree polynomials, B_{24} and B_{42}, we obtain the sum of two 9th degree polynomials in two variables x and z. We use the Symbolic Math Toolbox in the Matlab software to avoid tedious manual computation. The source code for the symbolic calculation in the Matlab script language is available upon request.

In summary, using Equation (3) and assuming that Δ_{JA} and Δ_{JB} are distributed with beta densities $B_{24}(x)$ and $B_{42}(y)$ we find that $Z = \Delta_{JA} - \Delta_{JB}$ has density:

$$q_{24}(z) = \sum_{i=0}^{i=9} s_i c_i z^i, \tag{7}$$

where $S = [s_0;\ s_1;\ s_2;\ s_3;\ s_4;\ s_5;\ s_6;\ s_7;\ s_8;\ s_9]$ takes the following values:

$$\begin{aligned}
S = &\ [1;\ 0;\ -1;\ -1;\ -1;\ 0;\ 0;\ -1;\ 0;\ 1] && \text{for } x \in [-1;0], \\
S = &\ [1;\ 0;\ -1;\ 1;\ -1;\ 0;\ 0;\ 1;\ 0;\ -1] && \text{for } x \in [0;1], \text{ and} \\
S = &\ [0;\ 0;\ 0;\ 0;\ 0;\ 0;\ 0;\ 0;\ 0;\ 0] && \text{otherwise.}
\end{aligned} \tag{8}$$

By this prescription, $q_{24}(z)$ is a smooth, piecewise polynomial function. It has four parts defined consecutively on intervals $(-\infty, -1]$, $[-1, 0]$, $[0, 1]$ and $[1, \infty)$. In defining the coefficients $s_i c_i$ of these four parts, we use signum function S valued -1, 0 and 1 listed above in Equation (8). The numerical absolute values of the polynomial coefficients are $c_0 = 1.5873,\quad c_2 = 17.1429,\quad c_3 = 33.33,\quad c_4 = 20,\quad c_7 = 2.8571,\quad c_9 = 0.6349$ and $c_1 = c_5 = c_6 = c_8 = 0$. Signs of the coefficients confirm that this piecewise polynomial function is an even function $q(z) = q(-z)$. We also recently found an alternative way to obtain the analytical results. A more straightforward way to calculate the indefinite integral $q(z)$ makes use of the Laplace transform. In using the transform, we obtain the convolution by using the inverse transform of the product of the two function images (Drapal and Marsalek 2010). We prefer to show here our original piecewise evaluation, because this integral evaluation method gives more insight into how the output density behaves and how it is obtained. The output function is similar to the normal density, although close inspection reveals subtle differences. Comparison of the output and normal probability density functions is shown in Figure 2. Note the difference between the two tails.

Note for editor: Insert *Figure 3* **around here.**

3 Results

Figure 3 shows the neural response curves of the MSO neuron with varying ITD. The curve in Figure 3 with the maximum at ITD = 0 forms the output of the circuit with only the excitatory branch plugged in. This corresponds to the output of the MSO in experiments, when the inhibitory branch is blocked by the application of strychnine, as in Brand *et al.* (2002). The other curve in Figure 3 with the maximal slope at ITD = 0 shows the output of the full model. This corresponds to the normal, control recordings from the MSO. Brand *et al.* (2002) also modeled some properties of the recorded cells using a detailed biophysical model with explicit representations of the voltage sensitive ion channels of the neuron, obtaining results very similar to those presented here.

Figures 3, 4, and 5 show results of numerical simulations of the model, which are in agreement with the analytical calculation. Both analytical and numerical computations were performed using Matlab software. Sound is represented as discrete samples of pure tones digitized using standard sound digitization (44 kHz sampling rate, 128 intensity levels). Spikes in response to pure tones were triggered by the leading edges of sound waves. In Drapal and Marsalek (2010) we used clicks and complex sounds, such as speech, processed by an cochlear implant emulator, with no qualitative difference from pure tone stimulus (not shown). Here we use pure tones only. The sound period, *T*, used in the model is that of the fundamental sound frequency. Spike trains were represented as trains of unitary events without any details of action potential shape. Synaptic integration was implemented with the use of time windows. Spontaneous activity is not a free parameter, but is added to the system to obtain positive values of responses. The activity in this and subsequent figures is a dimensionless quantity, which can be interpreted as the spike rate multiplied by a constant.

Note for editor: Insert *Figure 4* **around here.**

The size of the time window Δ_W is the key parameter of coincidence detection. Figure 4 shows the model output function for Δ_W = 200, 80, 50 and 20 μs. In Figure 4 only the excitatory curves are shown, since the inhibitory curves in the full model are analogous to those in Figures 3 and 5. The size of this window in the real neuronal circuit is set by the ionic currents giving rise to the generation of postsynaptic potentials (Svirskis *et al.* 2003, Szalisznyo 2006). Oertel *et al.* (2000) give the upper estimate of the window size for neurons in the auditory pathway specialized in coincidence detection. They give values in the range lesser than 300-1000 μs, with the smaller values for *in vitro* and larger values for *in vivo* preparations.

Note for editor: Insert *Figure 5* **around here.**

In general, the magnitude of timing jitter (denoted Δ_{JA} and Δ_{JB}) in a neuronal nucleus is dependent on the degree of neuronal convergence in the previous stages of

processing. In spike trains propagating to a higher order nucleus, one of three jitter changes may occur: the wave of spikes can sharpen (the jitter decreases), the jitter may not change, or the spike volley can become blurred (the jitter increases). All three variants can be obtained with the perfect integrator neuronal model in the regime of coincidence detection, for different sets of parameters (Reed *et al.* 2002). All three variants were shown to exist in experimental recordings (Gerstner *et al.* 1996, Joris 1996, Marsalek *et al.* 1996). We do not discuss these differences further here. The output activity shown in Figure 5 is dependent on ITD for two different magnitudes of timing jitter. Figure 5 shows two functions of the full model only (with both excitation and inhibition) for two values of timing jitter, 100 μs and 400 μs. Oertel *et al.* (2000) give jitter values of 200 μs for *in vivo* and values of 20-40 μs for *in vitro* preparations.

A related question was raised in review: How many ipsi- and how many contra-lateral synaptic connections (and of what proportion of excitatory and inhibitory connections) would a real MSO neuron need in order to achieve the accuracy observed physiologically in these neurons? Probabilistic spiking transmits information only with a given probability (reliability). In a previous investigation we calculated the time to spike required by the coincidence detection circuit for two given probabilities of achieving the decisive spike (reliabilities), $p_1 = 50\ \%$ and $p_2 = 95\ \%$ (Marsalek and Lansky, 2005). These probabilities were calculated based on the number, K, of unitary events. The calculation procedure, known as a Bernoulli process, is described in the 2005 paper, however, the values of K are not tabulated in the paper, since they are only intermediate results of the calculation. For the whole range of sound frequencies relevant for the human MSO, the number of events, K, attained values from 1 up to 50, higher values for higher frequencies. However, this number of unitary events can be achieved either by waiting for N sound periods, or by parallel processing by several neurons. In the latter case, the MSO nucleus must consist of at least M copies of the circuit, but may even have N tonotopic channels, such that $MN > K$. The numbers K, M, N represent the minimal values required for a functioning circuit. Numbers of neurons and connections in the

MSO are probably higher, since most nuclei show a redundancy in the number of their neurons and connections.

4 Discussion

Binaural sound localization is achieved with remarkable precision throughout the animal kingdom. The timing precision of individual spikes in most neurons in the auditory pathways of various species is lower than animals' behavioral assessment of the ITD. Given that this precision is important for survival and easily demonstrated, the means by which it is achieved by the neural circuits forms a fascinating and as yet unresolved question. In this paper we have presented a stochastic model employing both excitatory and inhibitory synaptic inputs to address this question in the neural circuit of the mammalian medial superior olive.

In recent years it has been demonstrated in experiments on gerbils (Brand *et al.* 2002) and also on other mammals (Grothe 2003) that synaptic inhibition plays a critical role in the sound localization circuit. The original theory of delay lines array by Jeffress applies in birds (Carr and Konishi 1988). It is not clear whether it also applies in mammals. Recordings in cats do not show the same ITD tuning curve slopes as those in rodent recordings (Joris and Yin 2007, Yin and Chan 1990). Harper and McAlpine (2004) present a theoretical explanation of these differences, including the discussion of human data. Psychophysical experiments studying the circuit in humans using subjective response might resolve this question in the future. These experiments range from normal hearing (Middlebrooks and Green 1991), to electrical hearing sense in cochlear implantees (Laback and Majdak 2008). Of course it is possible that the mammalian circuit uses some entirely different mechanism to those presented here and in previous investigations.

Another observation in the abovementioned experiments shows that the maximal response does not occur at the best ITD, but that the best ITD occurs where the slope of the response curve is maximal. As early as the 1970s and 1980s, some authors discussed the possibility that the maximal slope of the response of a coincidence detector measuring

the ITD between spike trains may relay the information (Goldberg and Brown 1969, Phillips and Brugge 1985). The possibilities to test this proposal numerically were limited at that time. Here we exploit the advancement of computational tools in a numerical study of the problem.

Most neurons do not achieve high spike timing precision, but in neurons in the periphery, the information about timing must be somehow encoded and preserved before reaching the first binaural neurons. Such timing precision is enabled by cellular and sub-cellular mechanisms and is best studied using deterministic equations, specifically differential equations. One can compare the work of Svirskis *et al.* (2003), which studies the state space of the nonlinear dynamics of ion currents in the MSO neurons, with that of (Szalisznyo 2006), in which the state space of the nonlinear ion currents in the LSO is studied. The authors of these two papers also performed experimental recordings. In both the circuits of the MSO and the LSO, the nonlinearities are necessary for the proper function of coincidence detectors. Other models (including this work) are based on stochastic spiking, or cast the neural circuit in terms of a logical circuit. See for example a review of Colburn (1996), or (Marsalek 2000). Experimental investigations can also be supported by conceptual models (Brand *et al.* 2002). It is argued frequently that the ITD tuning curve slope offset is due to the optimized wiring of the circuit (Grothe 2003, Harper and McAlpine 2004).

Most theories which include inhibitory mechanisms claim that inhibition is precisely timed and brief. The models assume high precision of spike times in all neurons in the circuit. Not all spikes are so precise, as is shown in experiments of Batra *et al.* (1997) and Joris (1996). The spatial organization of the neurons themselves is important. Agmon-Snir *et al.* (1998) and Zhou *et al.* (2005) suggest models based on a spatial organization of neurons in the MSO.

The parameter space of the binaural model is limited at higher frequencies, where ambiguities arise due to the relatively short sound wavelength in comparison with the ITD. In the mammalian brainstem, the circuit of the MSO processes the low frequencies and another twin circuit, the LSO, processes the high frequencies. The two circuits are

developed in parallel in human, cat, dog and other experimental species although in some mammals only one of them is present (Grothe 2003). The relative importance of the ITD and the IID within these respective sound frequency ranges in human was elucidated by means of confusing these two cues in psychophysical experiments (Wightman and Kistler 1992). Breebaart *et al.* (2001a and 2001b) brought a wider perspective towards the possible mechanisms used in binaural processing by using the excitatory-inhibitory mechanism where previous models dealt only with the excitatory-excitatory interaction. Their model uses the canonical structure of a grid of parallel tonotopic pathways together with delay lines and covers both the low frequency (ITD) and high frequency (IID) ranges with similar mechanisms. The model of Breebaart *et al.* (2001) comprises of building blocks of the signal processing circuits. Colburn (1996) and his numerous collaborators developed a series of binaural circuit models of varying complexity. They stress the importance of the coincidence detector within the circuit. Stern and Trahiotis (1996) review existing models and their own circuit implementations are close to delay line concepts.

Our model differs from those mentioned above in that it employs randomness in the spike arrival time between synapses. Random delay and spike timing jitter might seem like an impediment, especially in models with precisely timed inhibition. The statistical properties of spike trains average out errors in individual spikes and enable the neural computation of azimuth at the same time.

When two spikes from opposite sides arrive at the first binaural neuron, their coincidence must be detected with higher precision. Jeffress (1948) was first to notice this. All subsequent sound localization models have to assume that these particular neurons are coincidence detectors. Neurons as coincidence detectors have been used frequently in sound localization models (Gerstner *et al.* 1996, Marsalek 2000, Zhou *et al.* 2005). Let us also mention the classical model describing neural coincidence detectors in visual circuits in invertebrates, starting with the Reichardt model (Srinivasan and Bernard 1976, Zanker *et al.* 1999).

The first binaural neuron must use coincidence detection to extract information from interaural sound timing disparity. This neuron must indeed be precise as a coincidence detector. Such neuron outputs spike only if left and right inputs coincide within a short time window, in the time range of microseconds. In our model, stochastic properties make use of timing jitter, which at first glance is merely signal deteriorating noise.

From the point of view of implementation of the model, it does not matter exactly where the net ITD is placed in the two (left and right side) converging branches of the neural pathway. If an additional delay is added in both sides, it cancels out. The net delay between the two sides only matters when compared with the ITD, as written in Equation (1).

In a follow-up to this paper (Drapal and Marsalek 2010), we show that our numerical model can be connected to cochlear implant software emulator to give similar results to those observed by psychophysical methods in implanted volunteers (Laback and Majdak, 2008). As the possibilities to experiment with binaurally implanted and bimodal hearing subjects expand, it is possible that some psychophysical experiment will determine, whether delay line, stochastic delay, or both mechanisms are used in human hearing, even before definitive electrophysiological experiments on mammals are performed (Joris and Yin 2007).

In this paper we present a model for a neural algorithm performed by a circuit in the MSO. This is part of a general quest to capture the multitude of neural algorithms serving specific purposes. Let us give two closing examples of these algorithms. One is the case of spatial maps in the auditory brainstem of birds (Peña 2003). Peña (2003) shows how the brainstem circuit implements spike rate multiplication in order to calculate the location of a sound source. An analogous case from the 1980s concerns the neocortical visual circuit in the higher order visual areas which calculates the location of the illusory contour (von der Heydt *et al.* 1984). These authors demonstrate that a higher neocortical visual projection area can respond to a virtual object, an illusory contour, which is a result of a neural algorithm, as if it were a real solid object in a visual scene. Both these models were first proposed as hypotheses of a specific neural computation. The existence of

neural circuits implementing the respective neural computations was subsequently confirmed by a targeted experimental recording.

Abbreviations and symbols

CD – coincidence detection, **ECD** – excitatory coincidence detection, **ICD** – inhibitory coincidence detection, **EE** – excitatory–excitatory interaction, **EI** – excitatory–inhibitory interaction, **IID** – interaural intensity difference, **IPD** – interaural phase difference, **ISD** – interaural spectral difference, **ITD** – interaural time difference, **LSO** – lateral superior olive, **MSO** – medial superior olive, **PDF** – probability density function, Δ_J – time jitter (delay random variable), Δ_W – coincidence detection window, T – sound period and Δ_{max} – maximum delay.

The **beta density** is a probability density function written in a standard form as: $B_{ab}(x) = x^{b-1}(1-x)^{a-1}B(a,b)^{-1}$, for $x \in [0,1]$, and $B_{ab}(x) = 0$ otherwise, where the parameters a, $b > 0$ and $B(a,b)$ is the (Euler) beta function.

Conflict of Interests

The authors declare no conflict of interests.

Acknowledgments

Thanks to Michal Andrlik, Nick Dorrell, Abigail Klopper, Lubomir Kostal, Benjamin Lindner, Jose Louis Peña, Daniel Šuta and to four anonymous referees. The research was supported by two grants from the Czech Ministry of Education (MSMT), Research Initiatives no. 6840770012 and no. 0021620806.

References

AGMON-SNIR H, CARR CE, RINZEL, J: The role of dendrites in auditory coincidence detection. *Nature* **393**: 268–272, 1998.

BATRA R, KUWADA S, FITZPATRICK D: Sensitivity to interaural temporal disparities of low-and high-frequency neurons in the superior olivary complex. II. Coincidence detection. *J Neurophysiol* **78**: 1237–1247, 1997.

BECKIUS GE, BATRA R, OLIVER DL: Axons from anteroventral cochlear nucleus that terminate in medial superior olive of cat: observations related to delay lines. *J Neurosci* **19**: 3146–3161, 1999.

BRAND A, BEHREND O, MARQUARDT T, MCALPINE D, GROTHE B: Precise inhibition is essential for microsecond interaural time difference coding. *Nature* **417:** 543–547, 2002.

BREEBAART J, VAN DE PAR S, KOHLRAUSCH A: Binaural processing model based on contralateral inhibition. I. Model structure. *J Acoust Soc Am* **110**: 1074–1088, 2001a.

BREEBAART J, VAN DE PAR S, KOHLRAUSCH A: Binaural processing model based on contralateral inhibition. II. Dependence on spectral parameters. *J Acoust Soc Am* **110**: 1089–1103, 2001b.

CARR CE, KONISHI M: Axonal delay lines for time measurement in the owl's brainstem. *Proc Natl Acad Sci USA* **85**: 8311–8315, 1988.

COLBURN HS: Binaural models. In: *Auditory Computation*. HL HAWKINS, TA MCMULLEN, AN POPPER, RR FAY (eds), Springer Verlag, New York, 1996, pp 332-400.

DRAPAL M, MARSALEK P: Stochastic model shows how cochlear implants process azimuth in real auditory space. *Chinese J Physiol* **53**: In press, 2010.

GERSTNER W, KEMPTER R, VAN HEMMEN JL, WAGNER H: A neuronal learning rule for sub-millisecond temporal coding. *Nature* **383**: 76–78, 1996.

GOLDBERG JM, BROWN PB: Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: Some physiological mechanisms of sound localization. *J Neurophysiol* **32**: 613–636, 1969.

GROTHE B: New roles for synaptic inhibition in sound localization. *Nat Rev Neurosci* **4**: 540–50, 2003.

HARPER NS, MCALPINE D: Optimal neural population coding of an auditory spatial cue. *Nature* **430:** 682–686, 2004.

JEFFRESS LA: A place theory of sound localization *J Comp Physiol Psychol* **41:** 35–39, 1948.

JORIS P, YIN TCT: A matter of time: internal delays in binaural processing. *Trends Neurosci* **30**: 70-78, 2007.

JORIS PX, SMITH PH, YIN TC: Coincidence detection in the auditory system: 50 years after Jeffress. *Neuron* **21**: 1235–1238, 1998.

JORIS PX: Envelope coding in the lateral superior olive. II. Characteristic delays and comparison with the responses in the medial superior olive. *J Neurophysiol* **76**: 2137–2156, 1996.

KOCH C: *The Quest for Consciousness: A Neurobiological Approach*. Roberts and Co, Englewood, 2004.

LABACK B, MAJDAK P: Binaural jitter improves interaural time–difference sensitivity of cochlear implantees at high pulse rates. *Proc Natl Acad Sci USA* **105:** 814–817, 2008.

MARSALEK P: Coincidence detection in the Hodgkin-Huxley equations. *Biosystems* **58**: 83–91, 2000.

MARSALEK P, DRAPAL M: Mechanisms of coincidence detection in the auditory brainstem: Examples. In: *Mathematical Modeling of Biological Systems,* A DEUTSCH, R BRAVO DE LA PARRA, R DE BOER, O DIEKMANN, P JAGERS, E KISDI, M KRETZSCHMAR, P LANSKY, H METZ (eds), vol. 2. Birkhaeuser, Boston, 2008, pp 255–264.

MARSALEK P, KOCH C, MAUNSELL J: On the relationship between synaptic input and spike output jitter in individual neurons. *Proc Natl Acad Sci USA* **94**: 735-740, 1997.

MARSALEK P, KOFRANEK J: Sound localization at high frequencies and across the frequency range. *Neurocomputing* **58–60:** 999–1006, 2004.

MARSALEK P, KOFRANEK J: Spike encoding mechanisms in the sound localization pathway. *Biosystems* **79:** 191–198, 2005.

MARSALEK P, LANSKY P: Proposed mechanisms for coincidence detection in the auditory brainstem. *Biol Cybern* **92**: 445–451, 2005.

MCALPINE D, GROTHE B: Sound localization and delay lines – do mammals fit the model? *Trends Neurosci* **26**: 347–350, 2003.

MCALPINE D, JIANG D, PALMER A: A neural code for low-frequency sound localization in mammals. *Nat Neurosci* **4**: 396–401, 2001.

MIDDLEBROOKS J, GREEN D: Sound localization by human listeners. *Ann Rev Psychol* **42**: 135–159, 1991.

OERTEL D, BAL R, GARDNER SM, SMITH PH, JORIS PX: Detection of synchrony in the activity of auditory nerve fibers by octopus cells of the mammalian cochlear nucleus. *Proc Natl Acad Sci USA* **97**: 11773–11779, 2000.

PEÑA JL: Binaural processing in the synthesis of auditory spatial receptive fields. *Biol Cybern* **89**: 371–377, 2003.

PHILLIPS DP, BRUGGE JF: Progress in neurophysiology of sound localization. *Ann Rev Psychol* **36***:* 245–274, 1985.

REED MC, BLUM JJ, MITCHELL CC: Precision of neural timing: Effects of convergence and time-windowing. *J Comput Neurosci* **14:** 35–47, 2002.

SRINIVASAN MV, BERNARD GD: A proposed mechanism for multiplication of neural signals. *Biol Cybern* **21**: 227–236, 1976.

STERN RM, TRAHIOTIS C: Models of binaural perception. In: *Binaural and Spatial Hearing in Real and Virtual Environments.* R GILKEY, TR ANDERSON (eds), Lawrence Erlbaum Associates, New York, 1996, pp 499-531.

SVIRSKIS G, DODLA R, RINZEL J: Subthreshold outward currents enhance temporal integration in auditory neurons. *Biol Cybern* **89**: 333–340, 2003.

SZALISZNYO K: Role of hyperpolarization-activated conductances in the lateral superior olive: a modeling study. *J Comput Neurosci* **20:** 137–152, 2006.

VON DER HEYDT R, PETERHANS E, BAUMGARTNER G: Illusory contours and cortical neuron responses. *Science* **224:** 1260–1262, 1984.

WIGHTMAN FL, KISTLER DJ: The dominant role of low-frequency interaural time differences in sound localization. *J Acoust Soc Am* **91**: 1648–1661, 1992.

YIN T, CHAN J: Interaural time sensitivity in medial superior olive of cat. *J Neurophysiol* **64**: 465–488, 1990.

YOUNG ED: Cochlear nucleus. In *Synaptic Organization of the Brain, 4th edition*, GM SHEPHERD (ed) Oxford University Press, New York, 1998, pp. 121–157.

ZANKER JM, SRINIVASAN MV, EGELHAAF M: Speed tuning in elementary motion detectors of the correlation type. *Biol Cybern* **80:** 109–116, 1999.

ZHOU Y, CARNEY L, COLBURN H: A model for interaural time difference sensitivity in the medial superior olive: interaction of excitatory and inhibitory synaptic inputs, channel dynamics, and cellular morphology. *J Neurosci* **25:** 3046–3058, 2005.

Note for editor: All figures are single column.

Figure legends:

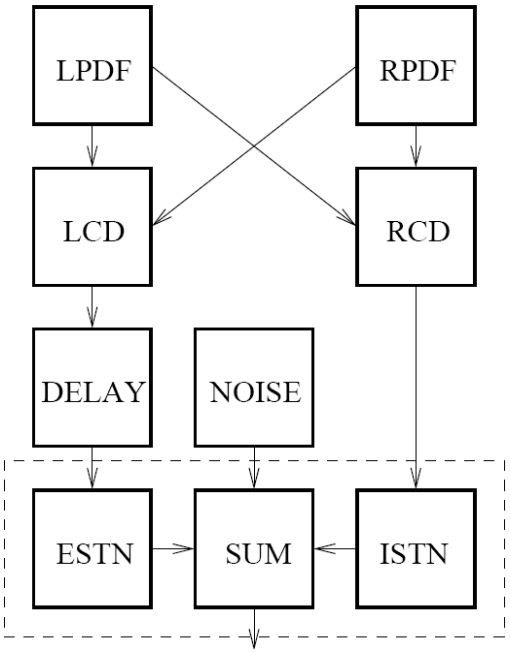

Figure 1: **Flow chart of the neuronal circuitry.** (LPDF) and (RPDF) represent input spikes on left and right sides. Left (Right) Probability Density Function in time governs spike times in the last monaural neuron. (LCD) and (RCD) are Left and Right Coincidence Detectors, realized by the first binaural neurons on the ipsi- and contra-lateral sides, respectively. Next is the (DELAY), which is shown here at the excitatory branch of the circuit only, because only the relative delay matters. Spontaneous spiking acts as an additional (NOISE) source. (ISTN) and (ESTN) represent Inhibitory and Excitatory Synaptic Transmissions, respectively. (SUM) adds excitatory inputs, inhibitory inputs and noise together. From the point of view of functionality, the exact sequence of operations (delay, sign change due to the inhibitory synapse, coincidence detection) in the feed forward pathway does not matter. The dashed line box encompasses processing, which can be performed in one binaural neuron. Note that individual boxes in this chart neither necessarily correspond to individual neurons nor represent the description of the mammalian brain stem circuit in the text.

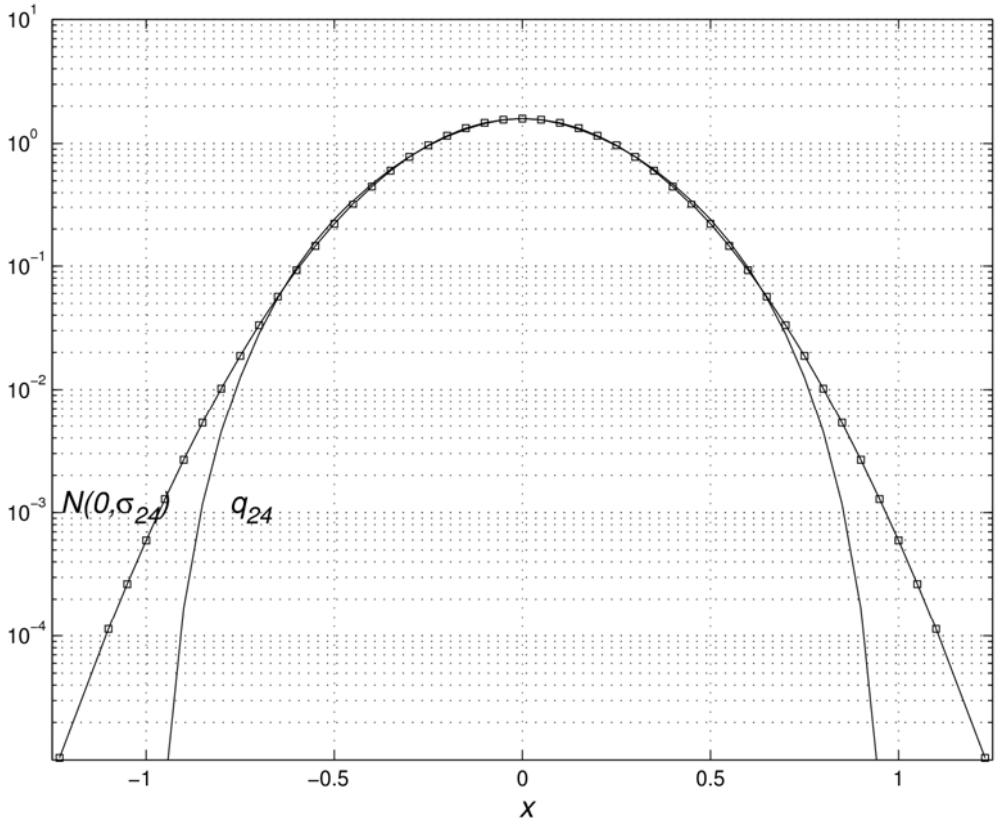

Figure 2: **Resulting probability density function q_{24}.** We use semi-logarithmic scale in the y-axis. On a linear scale (not shown) the resulting function q_{24} cannot be separated from the probability density of the normal distribution $N(0,\sigma_{24})$, where its standard deviation σ_{24} was chosen to match standard deviation of the function q_{24}. The semi-logarithmic scale used here clearly shows the divergence of the two tails of the normal and of the resulting density, when the tail of the normal density $N(0,\sigma_{24})$, marked with squares, is nonzero outside of interval [-1, 1]. The other function q_{24} possesses two discontinuities at points $q_{24}(-1) = 0$ and $q_{24}(1) = 0$. On the semi-logarithmic scale in y-axis these cannot be shown. At these points, the tails of the q_{24} function are cut off by the impulse (Heaviside) function. Semi-logarithmic plotting on the y-axis of the PDF shows normal density as parabolic in these coordinates.

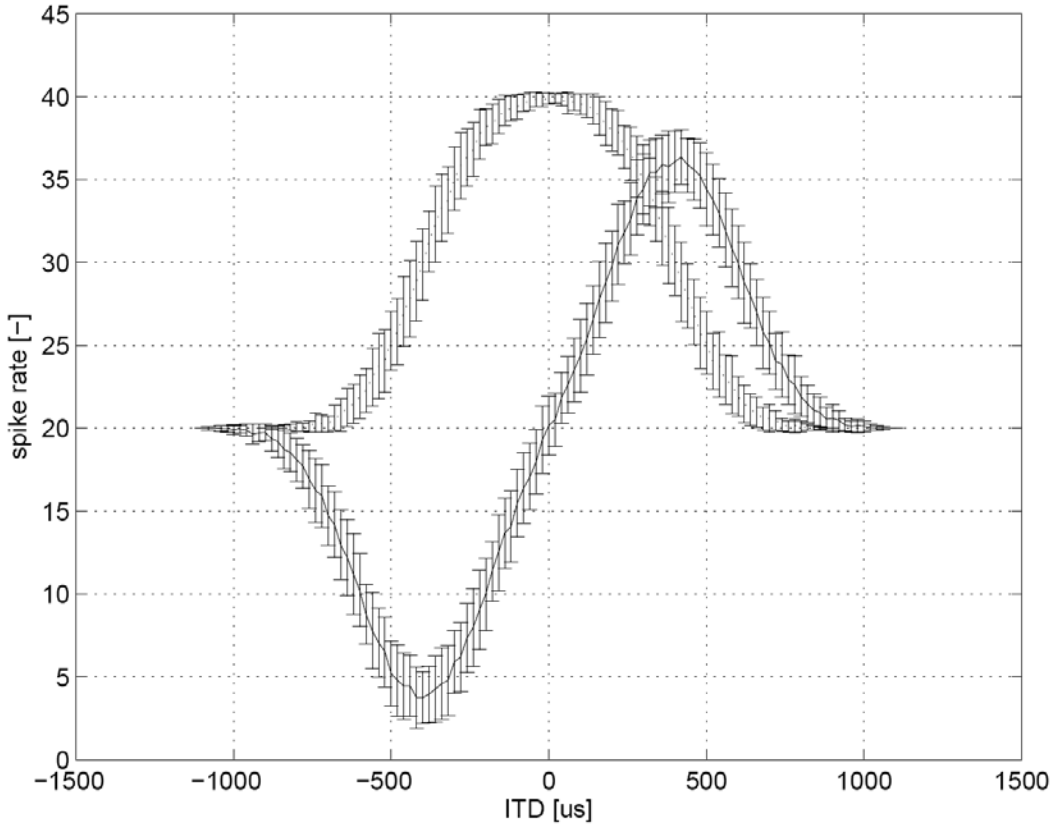

Figure 3: **Model response to pure tone stimulation.** ITD is on the x-axis in microseconds. Neuronal output activity is on the y-axis. The solid curve is the output of the full model with both excitation and inhibition. The dotted curve is the model without inhibition. Only one period of the output is shown here for clarity, but the output is periodic, so the x-axis maps to interaural phase difference. The error bars are sample standard deviations obtained numerically in simulations. In both curves they are the result of 100 trials. In other words, these error bars do not represent the level of the noise in the system. Instead, the noise is introduced into the system via the randomness of random variables and the magnitude of timing jitter.

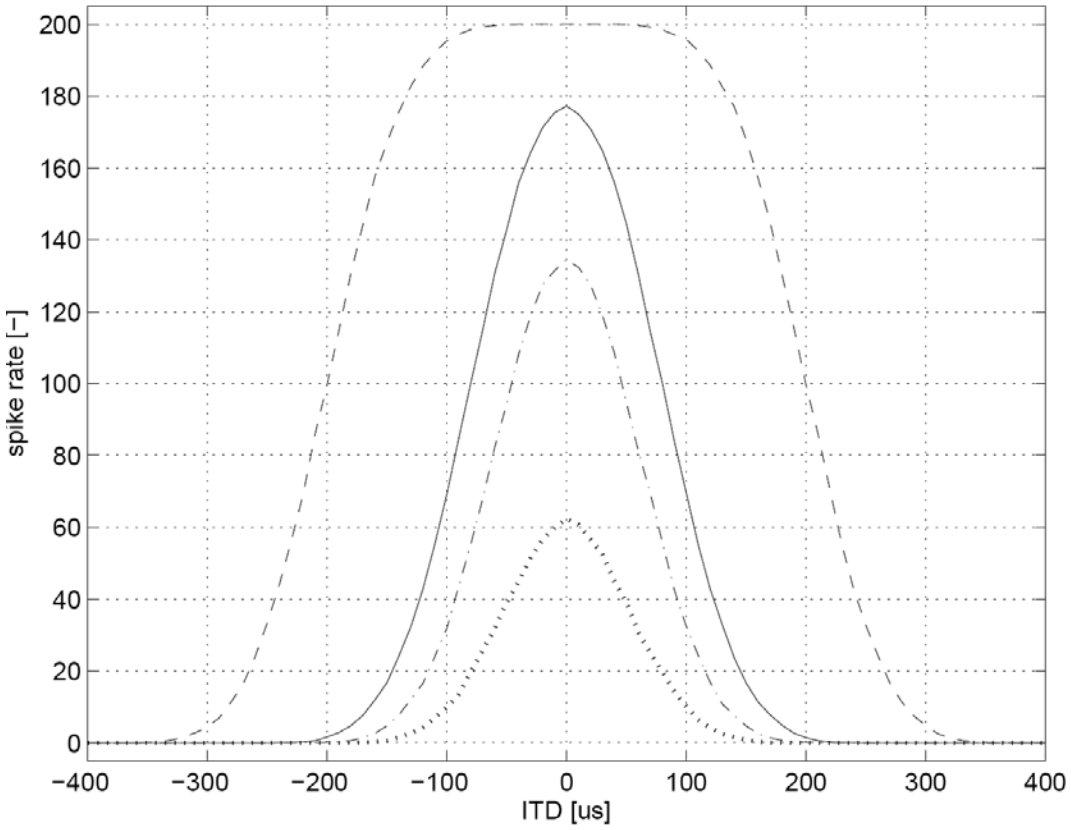

Figure 4: **Effect of the size of the time window.** The width of the time windows for CDs sets the range of ITDs detectable by the circuit. Four widths are compared here: Δ_W = 200 μs (--), 80 μs (-), 50 μs (-.) and 20 μs (…). Plotted output activity is dependent on ITD, as in Figure 3. Only the excitatory part is active in this plot. From an engineering point of view, an optimal curve uses the full dynamic range of available output activities, Δ_W = 50 μs.

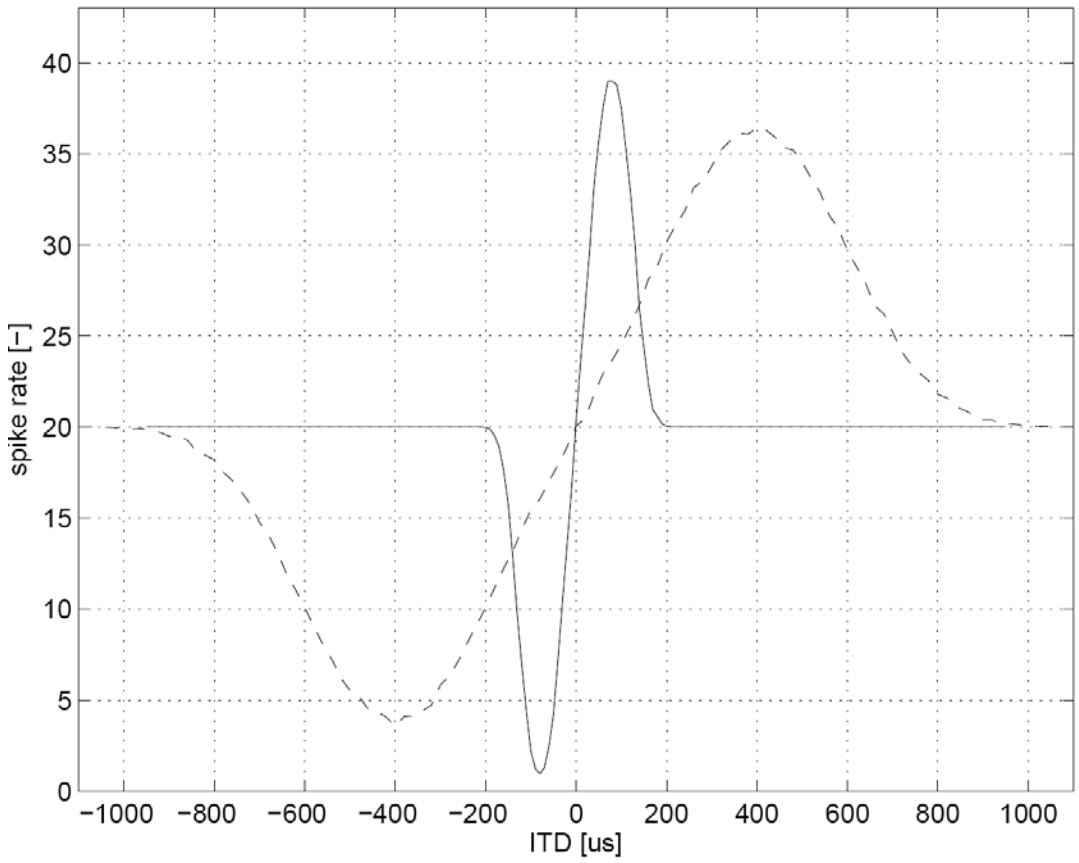

Figure 5: **Effect of the timing jitter magnitude.** Two magnitudes of timing jitter Δ_{JA} and Δ_{JB} are shown here, 100 μs as a solid line (-) and 400 μs as a broken line (--). They also set the slope of the model output curve and thus the range of ITDs detected. Together with the previous Figure 4, this shows the robustness of the model with respect to variation of parameters. Figures 4 and 5 illustrate the range of parameter values accessible to the model.