# On the precision of neural computation with interaural time differences in the medial superior olive

Petr Marsalek[a,b,c], Pavel Sanda[d], and Zbynek Bures[e,f]

[a]Institute of Pathological Physiology, First Medical Faculty, Charles University in Prague,
U Nemocnice 5/478, 128 53, Praha 2, Czech Republic
[b]Department of Radioelectronics, Faculty of Electrical Engineering, Czech Technical University
in Prague, Technická 2/1902, 166 27, Praha 6, Czech Republic
[d]Institute of Computer Science, Czech Academy of Sciences, Pod Vodárenskou věží 2/271,
182 07, Praha 8, Czech Republic
[e]Czech Institute of Informatics, Robotics and Cybernetics, Czech Technical University in
Prague, Jugoslávských partyzánů 4/1903, 166 36, Praha 6, Czech Republic
[f]College of Polytechnics, Tolstého 16/1556, 586 01, Jihlava, Czech Republic
[c]*Corresponding author:* Petr.Marsalek@LF1.CUNI.CZ

October 11, 2021

## Abstract

Incoming sound is in cochlea and auditory nerve encoded into spike trains. At the third neuron of the auditory pathway, spike trains of the left and right sides are processed in brainstem nuclei to yield sound localization information. Two different localization encoding mechanisms are employed in two "centers" for low and high sound frequencies in the brainstem. The centers are superior olivary nuclei, medial and lateral. This paper contains analytical estimates of parameters needed in description of auditory coding in sound localization neural circuit. Our model spike trains are based on electro-physiological recordings. We arrive to best estimates for neuronal signaling with the use of just noticeable difference of the ideal observer. We describe spike timing jitter and its role in the spike train processing. We study frequency dependence of spike trains on the sound frequency. All parameters are accompanied with detailed estimates of their values and variability. Intervals bounding all the parameter from lower and higher values are discussed.

## Keywords

binaural hearing; coincidence detection; ergodic hypothesis; ideal observer; interaural time difference; just noticeable difference; lateral and medial superior olive; neuronal arithmetic; psychophysics; sound localization; spike timing jitter.

## Abbreviations and symbols

$f_{\mathrm{S}}$, sound frequency; $F_{\mathrm{C}}$, critical sound frequency value; $\varphi$, sound phase; ILD, interaural level difference; IPD, interaural phase difference; ITD, interaural time difference; JND, just noticeable difference, also difference limen, or difference threshold; $K(.)$, $K_{\mathrm{C}}$, $K_S$, $K_X$; $A$, $B$, $C$, ... proportionality constants; $l$, sound level, also rate level of point process; LSO, lateral superior olive; MSO, medial superior olive; $R(.)$, $R_{\mathrm{VS}}$, VS, vector strength; $R_F$, firing rate; $\sigma$, standard deviation; $t_{\mathrm{J}}$ timing jitter; $t$, $\Delta t$, time, time difference; $T$, $T_{\mathrm{X}}$, $T_{(.)}$, sound periods, time constants.

# 1 Introduction

Mammalian sound localization circuits contain two nuclei in the auditory brainstem, the medial and the lateral superior olive, MSO and LSO. Neurons in these nuclei are the first binaural neurons in the auditory pathway connected to both ears. Due to the physical nature of the binaural sound, the MSO neurons process spike timing differences, Interaural Time Differences, ITDs, in the range of tens of microseconds, 10 $\mu$s [16]. The MSO processes low frequency sounds, in human this is from 20 Hz to not more than 2 kHz, and the LSO processes high frequency sounds, in human this is from 1 kHz up to 20 kHz [21]. The LSO uses cues of the Interaural Level Difference, ILD. The overlapping region is known to have a sensitivity drop at 1.5 kHz [22].

*Eptesicus fuscus* catches tidbit larvae of the meal-worm beetle *Tenebrio molitor*. The motor precision is in the range of hundreds of nanoseconds, 100 ns, and auditory separation of ultrasound echoes in the bats sonar sense is in the order of microsecond units, 1 $\mu$s [26]. SHOULD BE MOVED TO DISCUSSION.

Firing rate, first spike latency and individual spike timings are used in neural system coding, especially in the auditory pathway. Performance of the human MSO is in the range of microseconds, 10 $\mu$s and is reported to be improved two- to five-fold after several hours of training [21].

Different neural mechanisms are employed in the two nuclei. It has been reported that computation in the MSO is independent on sound intensity [7]. With higher sound intensity, first spike latency is shortening. Relation of this dependence to ITD and ILD has been described, yet it is difficult to interpret.

In contemporary human the MSO is the larger of the two nuclei and contains approximately 10000 - 11000 neurons and the human LSO contains 5600 neurons [23]. To implement the loudness change is much simpler than to record and implement microsecond time delay. Therefore in sound generation and processing, most of current auditory technology works as if the more important of the two localization cues in *Homo sapiens* were the sound intensity cue [30].

It is generally agreed that the main reason, why the workings of the MSO circuit deteriorate towards higher frequency is lowering of the synchronization of spike trains in the circuit with the sound source phase. The synchronization between two corresponding series of point events can be expressed as a discrete formula of vector strength, defined below in equation (1).

This article presents description of information encoding and neural computation in the MSO obtained mostly with analytical computations. Using the analytical tools we extend quantitative results obtained by numerical computations in [25]. We compare this analytical MSO description to the LSO description in [3] to arrive to unified description of neural circuits in the superior olive. We use this description to find maximum spike timing precision. Apparently, low and high frequency sound localization use different neuronal mechanism, because low and high sound frequencies are encoded by distinctive codes. In a simplified view, low frequencies are encoded by both spike timing and tonotopic organization, and high frequencies are encoded solely by the tonotopic organization.

# 2 Methods

## 2.1 Preliminaries

In the neural circuit model used here spikes, or action potentials, fired in the arbitrarily precise time are individual events of neural computation. Arbitrary precise timing would imply arbitrary high information content. In the model, this is limited by assumptions of intrinsic noise content.

Numerical implementation has been described in [25]. Here we develop combined stochastic and analytical description of the model. Our aim is to arrive at parameters and constants useful in the MSO description.

The neural circuit consists of neurons, functional units exchanging spikes. Incoming sound is sequentially processed in the auditory periphery. All the processing stages, including cochlea, are modeled as all-or-none units with various degree of biological realism. After cochlea, individual neurons correspond more-less to anatomical neuron numbering [15], where the zero order neurons is the whole mechanical-to-electrical cochlear mechanism, neurons of order 1 are in auditory nerve, order 2 are neurons in the cochlear nuclei, order two and a half is the medial nucleus of trapezoid body, we regard this nucleus as a "polarity inverter", and order 3 are the neurons of the MSO itself. This is the binaural part of the circuit.

Before they converge on the MSO the two, left and right, branches process sound from left and right ear. After the sound is encoded by cochleas into spike trains, the rest is the processing of spike trains by neurons. The spike trains are subject to delays and synaptic relying. A remarkable property of the auditory pathway is that both synapses and neurons have the shortest response times and highest time precision in the mammalian brain. If the neurons were represented by RC circuit, or similar equivalent biophysical models, their time constants would be comparable to, or lower than 1 ms. Due to vernier mechanisms known from various parts of peripheral sensory pathways, they can capture time events in the range of tens of microseconds. This capability has been described in human, [22]. In some animal specialists, some responses are in the range of tens of nanoseconds, as it has been shown in experiment on bats by [26]. Several other time constants and frequencies are characteristic for this neural circuit. They are shown in Table 1.

## 2.2   Model of the MSO neural circuit

Our model MSO circuit is based on connected phenomenological neurons. Input sound to left and right sides is transformed by the auditory periphery module into spike trains. Spikes in these trains are point events, only spike times matter and the details of spike numerical implementation do not make any differences in model output. These spike trains converge and diverge into higher order neurons. They are relayed from the auditory nerve and cochlear nucleus through the medial and lateral nuclei of trapezoid body up to the neurons of medial superior olive, which are first binaural neurons. Output of these neurons is the azimuth signal encoded in a spike train. See Figure 1.

## 2.3   ITD readout curve

Let us have a monotonous function with firing rate as an input, which outputs azimuth. We will call it the ITD readout curve. In the paper by [25] this curve is constructed by curve fitting to simulated points. Here we construct the curve based on assumption that the main frequency of sound input exist, is unique and is known. In addition to this known frequency, other parameters of the readout curve are set to make the fitting well posed and to obtain correct position of the curve maximum.

## 2.4   Vector strength

The vector strength has been first used in the context of sound localization by [6]. Its definition follows. Let us have sample spike phases $\varphi_i$, $i = 1, 2, \ldots, N$ relative to phases of a given input

master periodic function, which does not enter the formula. The periodicity of tones making up speech is a perfect example of such stimulus. Discrete sum vector strength of sample $\varphi_1, \ldots, \varphi_N$ attains values from 0 to 1 and is defined as

$$R_{\mathrm{VS}}(\varphi_i) = \frac{1}{N} \sqrt{\left(\sum_{i=1}^{N} \cos \varphi_i\right)^2 + \left(\sum_{i=1}^{N} \sin \varphi_i\right)^2}. \tag{1}$$

## 2.5 JND and ideal observer

A higher variability of firing leads to a lower precision of the rate code. Intuitively, if a repeated presentation of the same stimulus evokes each time different spike count, then to distinguish between two different stimuli, the associated spike count change must be larger than the spike count variability. This way we determine the Just Noticeable Difference (JND) of the rate code. In other words, this is the precision of rate coding.

We will ask whether it is possible to distinguish between two random processes with rates $l_1$ and $l_2$, $l_2 > l_1$. If we count events in a given counting window, we get counts $n_1$ and $n_2$. The probability that the observer obtains a result that $l_2 > l_1$ equals to the probability that $n_2 > n_1$. Let us assume that the random variables $n_i$, $i = 1, 2, \ldots$, have probability distributions $p(n_i)$ with means $\mu_i$ and equal standard deviation $\sigma$. A detection distance is then defined [28] as

$$d' = \frac{\mu_2 - \mu_1}{\sigma}. \tag{2}$$

This definition expresses the fact that the larger is the variance of the spike count, the worse is the detection capability. In psychophysics, a threshold value is commonly defined as that value for which the percentage of correct answers equals 75%. In our case, the examined value is the just-noticeable change of firing rate, $\Delta l = l_2 - l_1$. Assuming that both $p(n_1)$ and $p(n_2)$ are Gaussian (normal) distributed, the 75% probability of $n_2 - n_1 > 0$ corresponds to $d' = 1$. To obtain the JND of firing rate, we scale the detection distance with $\Delta l$ and put $\delta' = d'/\Delta l$. Then, the JND of firing rate is

$$\Delta l_{\mathrm{JND}} = 1/\delta' = \frac{l_2 - l_1}{\mu_2 - \mu_1} \sigma. \tag{3}$$

## 3 Results

We investigate how the MSO circuit output and overall performance depend on sound frequency and sound intensity. Figures from 2 to 5 show numerical simulations and experimental data as black lines and points, analytically described lower estimates as blue lines, upper estimates as red lines and standard errors of measurement and variations are shown in green.

Successive figures show individual steps of sound localization processing. Fig. 2 shows how vector strength in individual units lowers towards high sound frequencies. Fig. 3 shows the range of ITDs in sound localization judging. Fig. 4 shows the JND of the neural circuit in the dependence of the spike timing jitter. Fig. 5 shows synchronization to main sound frequency (in the case when it exists).

Figure 2 shows how vector strength $R_{\mathrm{VS}}$ lowers towards higher frequencies, as it can be observed in the module of the auditory periphery consisting of auditory nerve and cochlear nucleus. The prevailing majority of neurons in the auditory pathway has vector strength spike train statistics

sigmoidally dropping towards higher sound frequencies as it is in this example. In this figure, data originally recorded by [11] at the MSO of domestic cat, were fitted to the sigmoidal curve with the general formula of the Boltzmann function used in [19]. The curve fit of vector strength $R_{\mathrm{VS}}$ in dependence on sound frequency $F_{\mathrm{S}}$ is:

$$R_{\mathrm{VS}} = 1/(1 + \exp(K_{\mathrm{S}} f_{\mathrm{S}} - K_{\mathrm{C}} F_{\mathrm{C}1/2})), \tag{4}$$

where parameters with values are $K_{\mathrm{S}} = 2/0.75 = 2.666$ ms (kHz$^{-1}$), sound frequency coefficient; $K_{\mathrm{C}} = 4/0.75 = 5.333$ ms (kHz$^{-1}$), critical coefficient; and $F_{\mathrm{C}1/2} = 0.75$ kHz, critical half frequency. The numerical values are the proportionality constants of the $R_{\mathrm{VS}}$ upper bound. Note that at sound frequencies from 20 to 100 Hz, there are two branches reflecting the existence of two alternative ways how to compute the lower limit, which is 90 % of the upper limit, see Discussion section.

Figure 3 shows the curves limiting the ITD obtained with the basic parameter set in dependence on the sound frequency. The quadratic curve fit of the JND denoted $\Delta t_{\mathrm{JND}}$ is:

$$\Delta t_{\mathrm{JND}} = A(f_{\mathrm{S}} - F_{\mathrm{C}1/2})^2 + B, \tag{5}$$

Values of these parameters are $A = 10^{-5}$, $B = 0.05$, $F_{\mathrm{C}1/2} = 1$ kHz. The parameters in the Figures were constructed as follows: firstly, splines were fitted to experimental data. Splines parameters were allowed to vary within the 10 % of their original value and were approximated by the rounded off values. These procedures were also used to get parameter values below. For ranges of the audiogram parameters, see also [32]. Analogously to Figure 2, sound frequencies from 20 to 100 Hz exhibit higher spread between lower and upper limits, as the fitting method used, quadratic fit, is the same for both limits.

Figure 4 shows how the JND of ITD depends on timing jitter magnitude $t_{\mathrm{J}}$. Figure purpose is to capture, what is the best JND. There are several time constants, which are defined in relation to physical properties of spatial sound processing. To attain to rounded off parameters as in the other figures, we select individual values of the spike timing jitter and describe their purpose in the localization precision estimation. *Critical* timing jitter is lower estimate of timing jitter captured by spike train of typical mammalian neuron, $T_{\mathrm{JC}} = 0.2$ ms. *Normalized* value timing jitter $T_{\mathrm{JN}} = 1$ ms is the value of timing jitter normalized in relation to the output JND with respect to average firing rate. *Optimal* value timing jitter $T_{\mathrm{JO}} = 1.66$ ms is result of crossing two fits described below.

Simulations show that with lowering timing jitter the circuit output is virtually more and more precise. Yet, when the jitter is lower than critical value $T_{\mathrm{JC}}$, determined by intrinsic noise, duration of coincidence detection window, and by other time constants, the precision lowers again. The two curves fitted to the simulation are: 1. fit of exponential function to simulations, red curve:

$$\Delta t_{\mathrm{JND}} = \exp(A_1(t_{\mathrm{J}} - B_1)) - C_1, \tag{6}$$

where $A_1 = 1.9$, $B_1 = 1.25$ and $C_1 = 0.2$ are fitted parameters. This relation is shown conveniently by the logarithmic y-axis in this Figure.

2. another fit, which also takes into account shot noise in lower jitter values, is to a quadratic function, blue curve,

$$\Delta t_{\mathrm{JND}} = A_2(t_{\mathrm{J}} - B_2)^2 + C_2, \tag{7}$$

where $A_2 = 2.5$, $B_2 = 1$, $C_2 = 1$. There is only one parameter sought by numerical simulation. This is $A$, fitted to data, as the point $(x, y) = (B_2, C_2)$ has been chosen to be a unit. This fit is the normalized fit of the model.

| Parameter | Symbol | Units | Typical Value | Ranges |
|---|---|---|---|---|
| Timing Jitter | $t_{\mathrm{J}}, \sigma$ | ms | 1 | 0.125 - 8 |
| Window of Coincidence Detection | $w_{CD}$ | ms | 0.6 | 0.15 - 1.5 |
| Sound Frequency | $f$ | Hz | 200 | 40-1600 |
| Shortest Perceptual Time | $T_{PT}$ | ms | 20 | 20 - 80 |

Table 1: The basic set of parameters.

Logarithm of the simulated JND lowers with the exponential curve (6), which is concave function of sought jitter $t_{\mathrm{J}}$ as the jitter gets lower. The trend towards higher accuracy diverges from the parabolic fit in equation (7), when jitter reaches critical value between $T_{\mathrm{JC}}$ and $T_{\mathrm{JO}}$.

$$T_{\mathrm{JC}} = 0.2 \lesssim t_{\mathrm{J}} \lesssim T_{\mathrm{JO}} \lesssim 2\mathrm{ms}. \tag{8}$$

Beyond that point towards the lower jitter values, the neural circuit cannot function properly, as too low jitter prevents the interaction of spikes from the left and right side within the coincidence detection mechanism.

This corresponds to the analytical dependency obtained in [24] for a perfect integrator model with several inputs. The mechanism studied thereof is close to the MSO neural mechanism studied here. The firing rate changes in dependency on the input spike timing variability of *partially correlated* input spike trains.

Figure 5 shows the ITD readout curve. The rising slope of this curve is used as a readout function yielding the firing rate in dependency on the ITD, which in turn signals the sound azimuth to the next nuclei of the auditory pathway.

In the numerical model of [25] we reproduced a procedure to obtain azimuth tuning curves based on mammalian data. In this procedure, prior assumption of the existence of the main sound frequency and neural tuning to this frequency was used by experimentalists, but not by the model authors. When we use this assumption, which is stronger than in the numerical model, we obtain a fitting curve which is more coherent. (Has the higher vector strength value.) Two estimate errors are present in this Figure. The first is the mismatch between the use of Gaussian (normal) probability density function, as it is used in the experimental literature, and a circular statistics. More details of the circular statistics use in the sound localization context are explained in the article on the ergodicity assumption by [29]. The first error is shown here as the green curve. The other error, not shown in Figure, would arise from the prior assumption above. It appears that the assumption of the main frequency existence leads to more precise estimates.

## 4   Discussion

In this paper we have revisited numerical simulations by [25]. We have added analytical estimations to the description of the MSO function, which have not been known previously. Our analytical calculations make possible to derive time constants useful in description and design of normal human hearing. The descriptions are valid also for hearing with hearing aids and cochlear implants. All Figures contain analytically expressed upper and lower limits in their transfer functions or other functional descriptions.

Figure 2 contains two lower limit branches at low frequencies (shown by the blue and green curves, respectively). The green curve uses an assumption of lower energy and lower contribution to spike rate in neural units in lower frequencies. Limiting lower bound by two different analytical

functions (branches) can be understood as the estimate uncertainty. A conservative estimate of the lower bound always considers the lower of the two branches. This uncertainty should be recognized as one of original results presented for the first time in this paper. Its existence has been proposed in a doctoral thesis by [2]. To our knowledge this observation has not yet been published elsewhere.

Figure 3 depicts a quadratic fit. Clearly the data cannot be captured by the linear curve. The procedure to obtain the fit is analogous to obtaining other parameters in this paper. Initially splines were used and then their output was rounded to arrive to the quadratic fit. This fit is the simplest analytical way, how to capture nonlinear and band limited span of human hearing range.

Figure 4 calculations use assumptions about intrinsic noise [2]. The simulation data have been obtained by arbitrary precision calculation. Any neural recording cannot reach this precision due to the internal noise of both neurons and recording electronics. In order to capture circuit noisiness, we have used both exponential and quadratic fits. When we attempt to use them as upper and lower bounds, we notice that they exchange their order in region close to the optimal jitter value. In other words, at the lower jitter values the two estimates exchange their ordering. This is the choice of the quadratic fit to obtain a normalized bound together with other data-points. Numerical simulation with the basic set of parameters around the x-axis value of $t_J = 1$ ms lies beyond this point, but close to the exponential fit.

Figure 5 contains better fit of the Sine function, as compared to [25]. As the lower bound we can also use circular normal density function, the difference is negligible, not shown [29]. Comparison of time constants and sound periods in the model presented here will answer a tentative question: What is the highest slope of the ITD interpolation curve, such that it gives the resolution of well known minimum audible angle in the midline (ITD = 0), which is $4°$ in angular degrees? This slope is more steep in higher frequency sounds, its maximum is attained in maximum frequency of the MSO circuit operation, around 1 kHz [19].

Towards the analytical descriptions it is important to note that other periodic functions can be used as the ITD readout curves. In [29] we have compared the *Sine function* with the *circular beta density*, and other alternative functions. To impose periodic and infinite boundaries to the problem, regular normal density and circular normal density have been used and tested in previous versions of our model. No differences between these densities with proper parameters have been shown by common statistical tests when testing differences between two probability densities, ibidem.

[5] use recorded spike trains of several nuclei in the MSO neural circuit to demonstrate that coincidence detection is an essential part of the neuronal arithmetic executed by the neural circuit, see also [1]. These authors show simulations combined with experimental description of MSO workings in line with findings of this paper and with coincidence detection theories. Another MSO model, already studied in 2005 is: [31], this is an example of simplistic model, motivating the neural circuit description presented in this paper. For discussion of neural coding in the auditory nerve, auditory pathway, cochlear implants and brainstem neural circuits see [13].

Our investigation of quantitative properties of the superior olive neural circuit is also motivated by the three LSO experimental papers, which have detailed methodology applicable to LSO, to the overlap of sound frequency ranges between the LSO and the MSO; and also to the MSO range itself; [10]; [11]; [12].

Following his paper from 1948, Lloyd A. Jeffress dedicated lots of efforts to the search of a mechanism, by which microsecond time scale events of directional sound difference can be transformed into a code processed and transmitted by action potentials lasting several microseconds [8]. Historical comments on Jeffress papers are summarized by Cariani in Scholarpedia [4].

A plausible explanation of the microsecond precision of the MSO circuit describes the neural computation by leading edges of action potentials and post-synaptic potentials, [17], [29].

In [17], individual steps of signal processing in the superior olive neural circuits have been investigated. Various synaptic mechanisms have been proposed [18]. Spike timing jitter and spike variability have been systematically analytically investigated by [14].

In this last paragraph of Discussion section, we should mention briefly rest of scientific papers, where we found ideas towards the design of the mechanistic model presented here. Article by [20] discusses interaural phase delays (IPDs; when they exist, their utility is equivalent to that of the ITDs) and cochlear delays. For cochlear delays, very important is to review the ranges of delays in comparison to sound periods and classically described excitatory-excitatory and excitatory-inhibitory responses to binaural inputs in [9]. Paper of Srinivasan, Laback and Majdak cites current progress of ITD encoding by binaural cochlear implants, this is important for model validations and applications to studies with hearing aids and electrical hearing, [27].

## Conclusions

This theoretical paper is continuation of sound localization precision descriptions in the MSO [25] and in the LSO [3]. Major novel results here are at least two: 1) analytical estimates of results obtained previously only by numerical simulation and 2) estimates of auditory parameters od functions bounding from the bottom and from the top known characteristics of the human sound localization circuit.

## Acknowledgments

## Author contributions

Model design: PM, ZB. Methodology: PM, PS. Writing â original draft: PM. Writing â editing: PM, PS, ZB.

### Technical, T<sub>E</sub>X Typesetting, Archiving Notes

Compilation notes: LaTeX compilation date is October 11, 2021. BibTeX output: Circa 30 is the estimated references count (March 2020: 32). Version of sources in octave is 2020.1.4, version of manuscript is 2020.A.2. Complete sources are available upon request. The sources web link /protected by a password/ is here. `http://nemo.lf1.cuni.cz/mlab/ftp/ftp/bures-z` Preview web link: `http://nemo.lf1.cuni.cz/mlab/ftp/ftp/to-be-added`

Things to do to by co-authors:
Petr Marsalek. Clean up and finish first pass writing.
Pavel Sanda is kindly asked to help the author team with tasks: 1) reading Discussion 2) checking up figures. Zbynek Bures: should do internal English language copy-editing. final reading of the manuscript before submission to the journal.

# References

[1] Bures, Z., The stochastic properties of input spike trains control neuronal arithmetic, Biol. Cybern. 106:111–122, 2012.

[2] Bures, Z., Internal Representation and Processing of Acoustic Stimuli in the Nervous System. Habilitation thesis, Brno, Czech Republic: VUTIUM (Publishing House of the Brno University of Technology), 2014, 122 pages.

[3] Bures, Z. and Marsalek, P., On the precision of neural computation with the interaural level difference in the lateral superior olive, Brain Res. 1536:111–122, 2013.

[4] Cariani, P., Jeffress model, Scholarpedia 6:2920, 2011.

[5] Franken, T.P., Bremen, P. and Joris, P.X., Coincidence detection in the medial superior olive: mechanistic implications of an analysis of input spiking patterns, Front. Neural Circuits 8:1–21, 2014.

[6] Goldberg, J.M. and Brown, P.B., 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.

[7] Grothe, B., Pecka, M. and McAlpine, D., Mechanisms of sound localization in mammals, Physiol. Rev. 90:983–1012, 2010.

[8] Jeffress, L.A., Blodgett, H.C. and Deatherage, B.H., Effect of interaural correlation on the precision of centering a noise, J. Acoust. Soc. Am. 34:1122–1123, 1962.

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

[10] Joris, P.X. and Yin, T.C., Envelope coding in the lateral superior olive. I. Sensitivity to interaural time differences, J. Neurophysiol. 73:1043–1062, 1995.

[11] Joris, P.X., 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.

[12] Joris, P.X. and Yin, T.C., Envelope coding in the lateral superior olive. III. Comparison with afferent pathways, J. Neurophysiol. 79:253–69, 1998.

[13] Kerber, S. and Seeber, B.U., Sound localization in noise by normal-hearing listeners and cochlear implant users, Ear Hear. 33:445–457, 2012.

[14] Kostal, L. and Marsalek, P., Neuronal jitter: can we measure the spike timing dispersion differently?, Chinese J. Physiol. 53:454–464, 2010.

[15] Kulesza Jr, R.J., Cytoarchitecture of the human superior olivary complex: medial and lateral superior olive, Hear. Res. 225:80–90, 2007.

[16] Laback, B. and 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.

[17] Marsalek, P., Coincidence detection in the Hodgkin-Huxley equations, Biosystems 58:83–91, 2000.

[18] Marsalek, P. and Kofranek, J., Spike encoding mechanisms in the sound localization pathway, Biosystems 79:191–8, 2005.

[19] Marsalek, P. and Lansky, P., Proposed mechanisms for coincidence detection in the auditory brainstem, Biol. Cybern. 92:445–451, 2005.

[20] Michelet, P., Kovacic, D. and Joris, P.X., Ongoing temporal coding of a stochastic stimulus as a function of intensity: time-intensity trading, J. Neurosci. 32:9517–9527, 2012.

[21] Middlebrooks, J.C. and Green, D.M., Sound localization by human listeners, Annu. Rev. Psychol. 42:135–159, 1991.

[22] Mills, A.W., On the minimum audible angle, J. Acoust. Soc. Am. 30:237–246, 1958.

[23] Moore, J.K., Organization of the human superior olivary complex, Microsc. Res. Tech. 51:403–412, 2000.

[24] Salinas, E. and Sejnowski, T.J., Impact of correlated synaptic input on output firing rate and variability in simple neuronal models, J. Neurosci. 20:6193–6209, 2000.

[25] Sanda, P. and Marsalek, P., Stochastic interpolation model of the medial superior olive neural circuit, Brain Res. 1434:257–265, 2012.

[26] Simmons, J.A., Ferragamo, M.J. and Moss, C.F., Echo-delay resolution in sonar images of the big brown bat, *Eptesicus fuscus*, Proc. Natl. Acad. Sci. USA 95:12647–12652, 1998.

[27] Srinivasan, S., Laback, B., Majdak, P. and Delgutte, B., Introducing short interpulse intervals in high-rate pulse trains enhances binaural timing sensitivity in electric hearing, J. Assoc. Res. Otolaryngol. 19:301–315, 2018.

[28] Tanner Jr., W.P., Physiological implications of psychophysical data, Ann. N. Y. Acad. Sci. 89:752–765, 1961.

[29] Toth, P.G., Marsalek, P. and Pokora, O., Ergodicity and parameter estimates in auditory neural circuits, Biol. Cybern. 112:41–55, 2018.

[30] Vencovský, V. and Rund, F., Using a physical cochlear model to predict masker phase effects in hearing-impaired listeners: a role of peripheral compression, Acta Acust. United Acust. 102:373–382, 2016.

[31] Zhou, Y., Carney, L.H. and Colburn, H.S., 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.

[32] Zwislocki, J. and Feldman, R., Just noticeable differences in dichotic phase, J. Acoust. Soc. Am. 28:860–864, 1956.

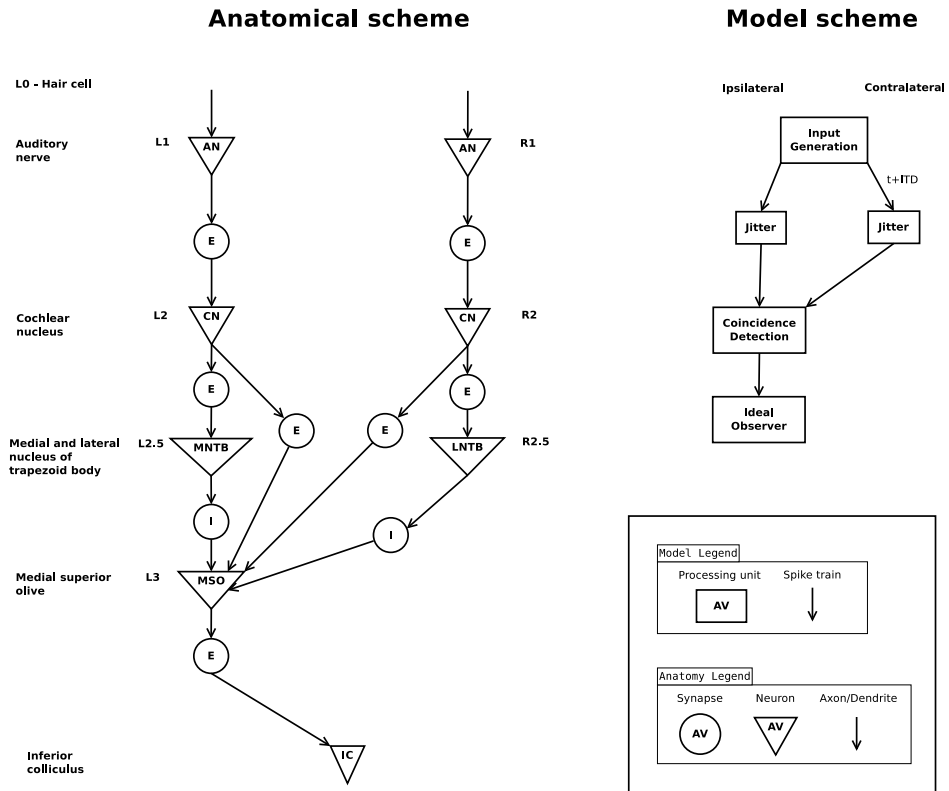**Anatomical scheme**                    **Model scheme**



Figure 1: **Schematic MSO circuit with anatomic and connection layouts.** This Figure with modifications is reprinted with permission from the manuscript of [25] to show the numerical model key features.
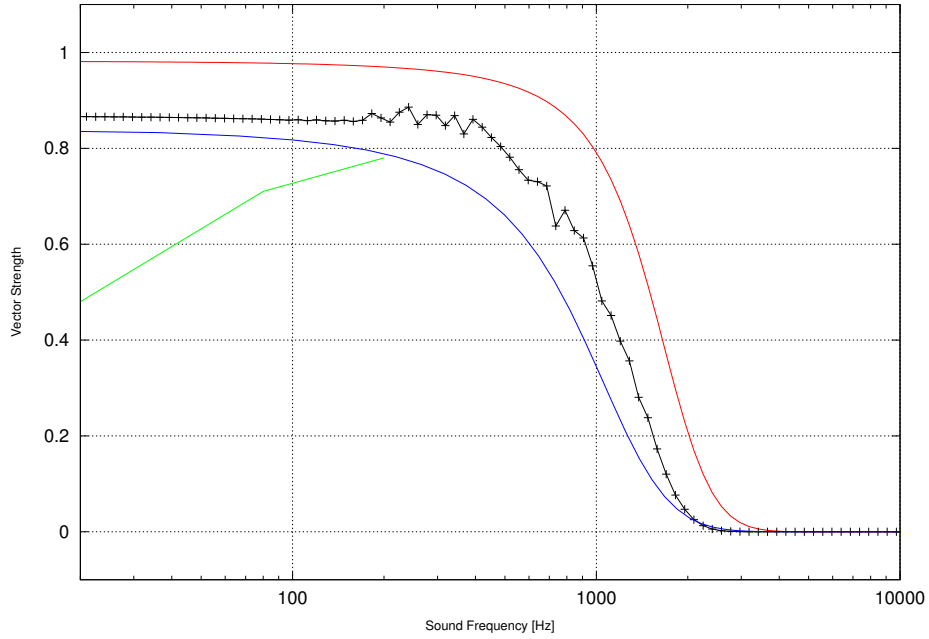
Figure 2: **Vector strength of auditory nerve spike trains in dependence on sound frequency.** X-axis shows sound frequency in Hz in logarithmic scale and y-axis shows the vector strength. Even though in some nuclei up the auditory pathway the synchronization can be maintained towards higher frequencies than shown here, the decrease of the vector strength towards higher frequencies is a general property of all neurons in the auditory pathway. Red curve shows upper theoretical limit, blue curve shows lower limit, green curve shows limit imposed by lower firing rate and lower energy of low frequency sound. Black points are data simulated with the use of point-process spike train generation with the use of the dead time Poisson process. Note that in frequency $f_S$ range from 20 to 200 Hz the lower limit is shown by curve branching to two branches to the left. The upper is the Boltzmann function fit and the lower is decrease of vector strength at low frequencies due to stochastic response of high spontaneous rate neurons.
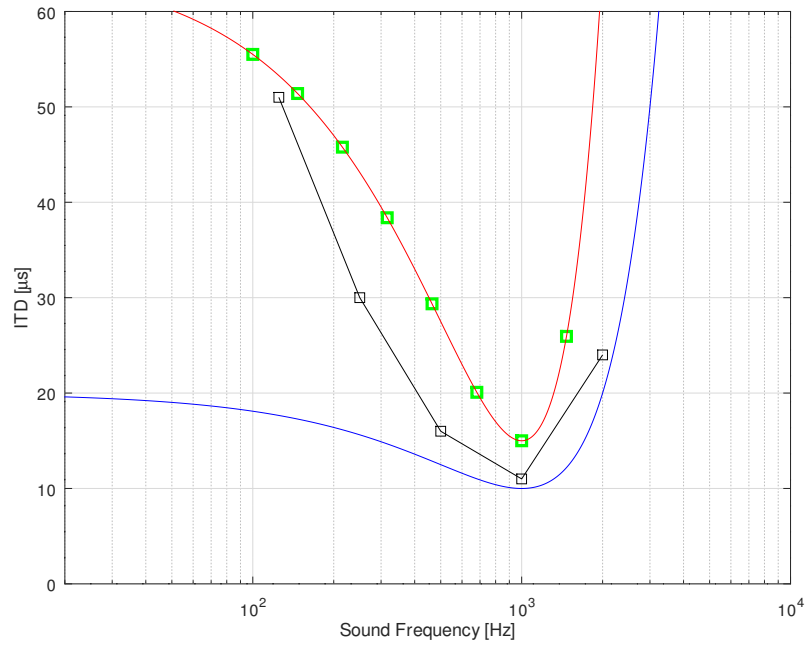
Figure 3: **The shortest JND of ITD detected in the dependence on sound frequency.** X-axis shows sound frequency in Hz in a logarithmic scale and y-axis the shortest JND of ITD in $\mu$s. This is a theoretical prediction based on the analytical model and basic parameter set used in simulations. As in other figures, black line is obtained by simulation and red and blue lines are respectively upper and lower bounds obtained by an analytic fit.
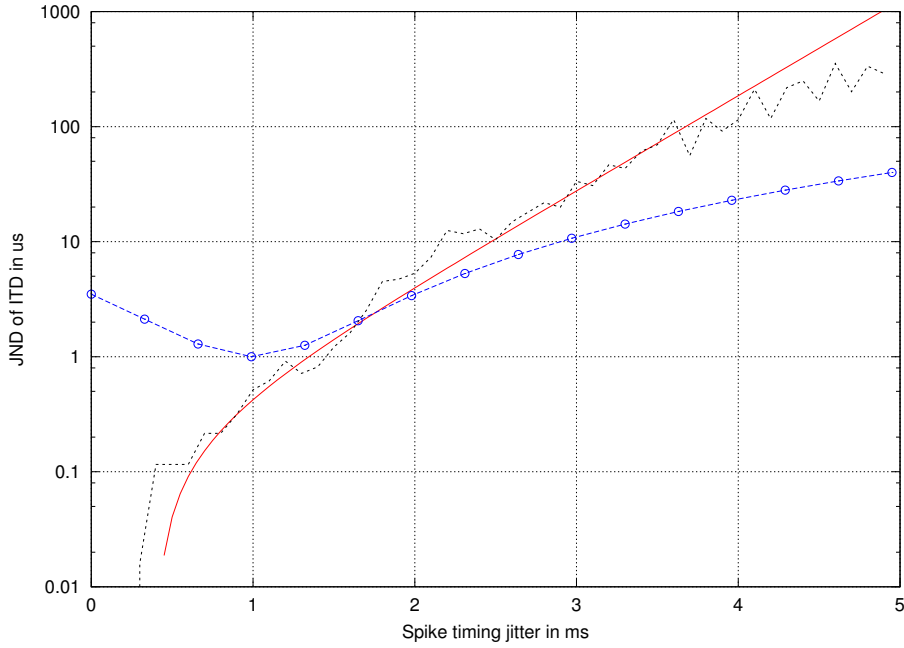
Figure 4: **JND values in basic parameter set in dependence on the spike timing jitter magnitude.** This plot in semi-logarithmic y-scale shows JND (just noticeable difference) of interaural time difference depending on variation of the spike timing jitter. Jagged black line: simulated data, solid line: an exponential fit to the simulations under the assumption of arbitrary time precision in the model circuit, dotted line: a quadratic function estimate of spike timing precision in a system with addition of noise. Note that in this Figure the exponential and quadratic fits cross at $f_S = 1.66$ ms. In order to correspond to other Figures showing the upper and lower bounds of the estimate of stochastic model, the two fits are split into two branches of the same function at this point of $f_S = 1.66$ ms. For lower x-axis $t_J$ values, quadratic fit is larger than the exponential, and vice versa. This is indicated by distinctive data-points. (These are circles and triangles; no data-points and squares, respectively.) Also notice that the curve of the quadratic fit goes through the point [1,1], this is a consequence of using normalized parameter set.
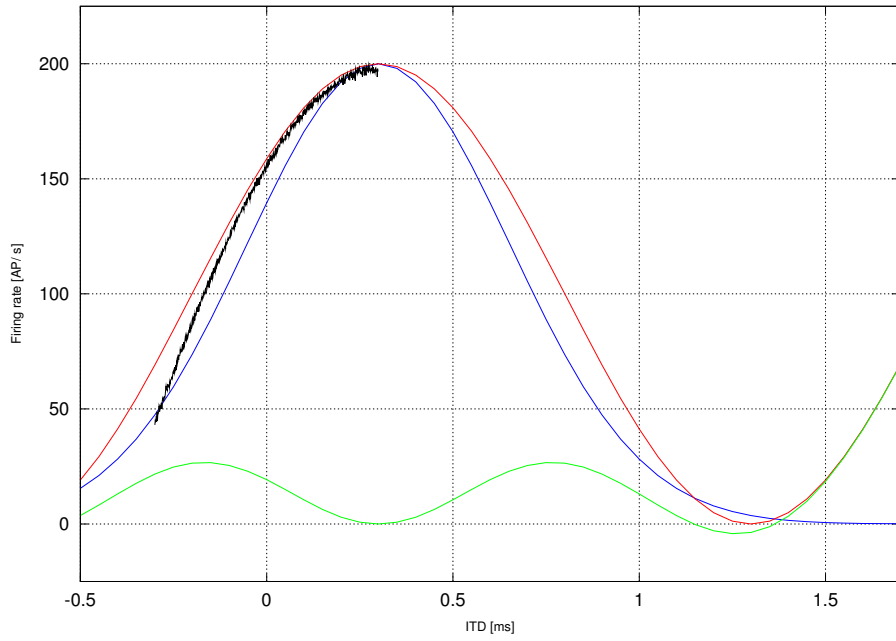
Figure 5: **Fit of example functions to firing rate slope ITD readout curve.** X-axis shows ITD in ms and y-axis shows corresponding firing rate in action potentials per second. Note that the curve peak is offset from the origin of coordinates at $t_{\mathrm{ITD}} = 0$. Red curve is the Sine density function fit and blue curve is fit by the normal density function with the variance set to correspond to the known sound main period. Green curve shows the difference between the red and blue curves.