

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

Petr Marsalek<sup>1,2,\*</sup>, Pavel Sanda<sup>3</sup> and Zbynek Bures<sup>2,4</sup>

<sup>1</sup> *Institute of Pathological Physiology, First Medical Faculty, Charles University in Prague, U Nemocnice 5/478, 128 53, Praha 2, Czech Republic*

<sup>2</sup> *Department of Natural Sciences, Faculty of Biomedical Engineering, Czech Technical University in Prague, nám. Sítná 3105, 272 01 Kladno, Czech Republic*

<sup>3</sup> *Institute of Computer Science, Czech Academy of Sciences, Pod Vodárenskou věží 2/271, 182 07, Praha 8, Czech Republic*

<sup>4</sup> *Czech Institute of Informatics, Robotics and Cybernetics, Czech Technical University in Prague, Jugoslávských partyzánů 4/1903, 166 36, Praha 6, Czech Republic*

Correspondence\*:

Petr Marsalek

Petr.Marsalek@cvut.cz

## 2 ABSTRACT

3 In the auditory nerve and the following auditory pathway, incoming sound is encoded into spike  
4 trains – series of neural action potentials. At the third neuron of the auditory pathway, spike trains  
5 of the left and right sides converge and are processed to yield sound localization information. Two  
6 different localization encoding mechanisms are employed for low and high sound frequencies in  
7 two dedicated nuclei in the brainstem: the medial and lateral superior olivary nuclei. Building upon  
8 our previous computational model of medial superior olive (MSO), this paper brings analytical  
9 estimates of parameters needed to describe auditory coding in the MSO circuit. We arrive to  
10 best estimates for neuronal signaling with the use of just noticeable difference and the ideal  
11 observer concepts. We describe spike timing jitter and its role in the spike train processing. We  
12 study the dependence of sound localization precision on the sound frequency. All parameters  
13 are accompanied with detailed estimates of their values and variability. Intervals bounding all the  
14 parameters from lower and higher values are discussed.

15 **KEYWORDS:** binaural hearing, coincidence detection, ideal observer, interaural time difference, just noticeable difference, lateral  
16 and medial superior olive, sound localization, spike timing jitter

## 17 Abbreviations and symbols

18  $f_S$ , sound frequency;  $F_C$ , critical sound frequency value;  $\varphi$ , sound phase; ILD, interaural level difference;  
19 IPD, interaural phase difference; ITD, interaural time difference; JND, just noticeable difference, also  
20 difference limen, or difference threshold;  $K(\cdot)$ ,  $K_C$ ,  $K_S$ ,  $K_X$ ;  $A$ ,  $B$ ,  $C$ , ... proportionality constants;  
21  $l$ , sound level, also rate level of point process; LSO, lateral superior olive; MSO, medial superior olive;  
22  $R(\cdot)$ ,  $R_{VS}$ , VS, vector strength;  $R_F$ , firing rate;  $\sigma$ , standard deviation;  $t_J$  timing jitter;  $t$ ,  $\Delta t$ , time, time  
23 difference;  $T$ ,  $T_X$ ,  $T_{(\cdot)}$ , sound periods, time constants.

## 1 INTRODUCTION

24 Mammalian sound localization circuits contain two nuclei in the auditory brainstem - the medial and the  
25 lateral superior olive, MSO and LSO, respectively. Neurons in these nuclei are the first binaural neurons  
26 in the auditory pathway, processing the information arriving from both ears. The MSO processes low  
27 frequency sounds, in human this is from 20 Hz to not more than 2 kHz, and the LSO processes high  
28 frequency sounds, in human this is from 1 kHz up to 20 kHz Middlebrooks and Green (1991). Due to the  
29 physical nature of the binaural stimulus, the MSO neurons process spike timing differences, Interaural  
30 Time Differences, ITDs. The ITDs emerge as a result of different distance from the sound source to each  
31 ear. The ITDs take values given by the distance between the ears (in human this is circa 16 cm/ 466  $\mu$ s  
32 sound takes to travel this distance in the air) and are detected with precision of tens of microseconds, 10  $\mu$ s  
33 Laback and Majdak (2008). The LSO detects the Interaural Level Difference, ILD, given by the acoustic  
34 shadow of the head. The frequency region, where these two mechanisms overlap (around 1.5 kHz) is known  
35 to have a drop in localization sensitivity Mills (1958).

36 Firing rate, first spike latency and individual spike timings are used in neural system coding, especially  
37 in the auditory pathway. Human MSO can detect binaural spike timing shifts as small as 10  $\mu$ s and this  
38 ability is reported to be improved two- to five-fold after several hours of training Middlebrooks and Green  
39 (1991). Considering that duration of a neuronal action potential is approximately 1 ms, such performance is  
40 striking. Furthermore, to be etologically useful, the MSO must provide stable operation over a wide range  
41 of sound frequencies and intensities and listening conditions. While it has been reported that computation  
42 in the MSO is independent on sound intensity Grothe et al. (2010), it is known that its performance drops  
43 as the stimulus frequency increases.. Relation of this dependence to binaural cues has been described, yet it  
44 is difficult to interpret.

45 It is generally agreed that the main reason, why the precision of the MSO circuit deteriorates towards  
46 higher frequency is lowering of the synchronization of spike trains with the sound phase. However, exact  
47 mechanisms are not completely understood. One of our aims is thus to explore the conditions and limits of  
48 MSO circuit operation using the computational modeling approach.

49 This article presents description of information encoding and neural computation in the MSO obtained  
50 mostly with analytical computations. Using the analytical tools we extend quantitative results obtained by  
51 numerical computations in Sanda and Marsalek (2012). We compare this analytical MSO description to  
52 the LSO description in Bures and Marsalek (2013) to arrive to unified description of neural circuits in the  
53 superior olive. We use this description to find the performance limits of the MSO circuit in dependence  
54 on sound frequency and intensity. Apparently, low and high frequency sound localization use different  
55 neuronal mechanisms. Sound features in low and high sound frequencies are encoded by distinctive codes  
56 and this multitude of codes also affects binaural processing.

## 2 METHODS

### 57 2.1 Preliminaries

58 In the neural circuit model used here spikes, or action potentials, fired in the arbitrarily precise time are  
59 individual events of neural computation. Arbitrary precise timing would imply arbitrary high information  
60 content. In the model, this is limited by assumptions of intrinsic noise content. Numerical implementation  
61 has been described in Sanda and Marsalek (2012). Here we develop combined stochastic and analytical  
62 description of the model. Our aim is to arrive at parameters and constants useful in the MSO description.

63 The neural circuit consists of neurons, functional units exchanging spikes. Incoming sound is sequentially  
64 processed in the auditory periphery. All the processing stages, including cochlea, are modeled as all-or-none  
65 units with various degree of biological realism. After cochlea, locations of individual neurons in model  
66 correspond nearly to anatomical neuron numbering Kulesza Jr (2007). The zero order neuron is the whole  
67 mechanical-to-electrical cochlear mechanism, neurons of order 1 are in auditory nerve, order 2 are neurons  
68 in the cochlear nuclei, order two and a half is the medial nucleus of trapezoid body, we regard this nucleus  
69 as a "polarity inverter", and order 3 are the neurons of the MSO itself. The order 3 neurons are the binaural  
70 part of the circuit.

71 Before they converge on the MSO the two, left and right, branches process sound from left and right  
72 ear. After the sound is encoded by cochleas into spike trains, the rest is the processing of spike trains  
73 by neurons. The spike trains are subject to delays and synaptic relying. A remarkable property of the  
74 auditory pathway is that both synapses and neurons have the shortest response times and highest time  
75 precision in the mammalian brain. If the neurons were represented by RC circuit, or similar equivalent  
76 biophysical models, their time constants would be comparable to, or lower than 1 ms. Due to vernier  
77 mechanisms known from various parts of peripheral sensory pathways, they can capture time events in  
78 the range of tens of microseconds. This capability has been described in human, Mills (1958). In some  
79 animal species – localization specialists, responses are in the range of tens of nanoseconds, as it has been  
80 shown in experiment on bats by Simmons et al. (1998). Several other time constants and frequencies are  
81 characteristic for this neural circuit. They are shown in Table 1.

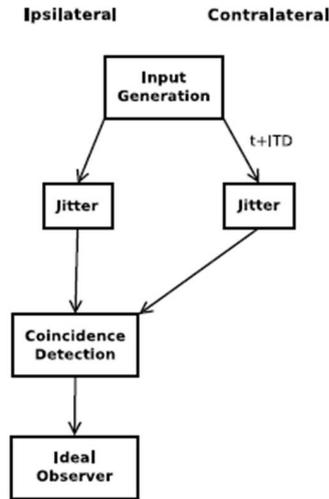
## 82 2.2 Model of the MSO neural circuit

83 Our model MSO circuit is based on connected phenomenological neurons. Input sound to left and right  
84 sides is transformed by the auditory periphery module into spike trains. Spikes in these trains are point  
85 events, where only spike times matter and the details of spike numerical implementation do not make any  
86 differences in model output. These spike trains converge and diverge into higher order neurons. They are  
87 relayed from the auditory nerve and cochlear nucleus through the medial and lateral nuclei of trapezoid  
88 body up to the neurons of medial superior olive, which are first binaural neurons. Output of the binaural  
89 neurons is the azimuth signal encoded in a spike train. However, components of our computational model  
90 can not be identified with individual neurons and their anatomical connectivity. Rather, the model consists  
91 of functional block units representing the crucial properties of the explored neural mechanism, such as  
92 intrinsic noise, temporal delay and coincidence detection. This functional block model layout provides  
93 access to descriptive parameters, which are setting the circuit's performance limits, see Figure 1. In our  
94 MSO model, the interaural time delay (ITD) present at the MSO input is represented by firing rate at the  
95 MSO output. The relationship between the input ITD and output spike rate is called ITD readout curve, see  
96 the next section.

## 97 2.3 ITD readout curve

98 Let us have a monotonous function with firing rate as an input, which outputs azimuth. We call it the ITD  
99 readout curve. In the paper by Sanda and Marsalek (2012) this curve was constructed by curve fitting to  
100 simulated points. Several assumptions about the sound objects had to be fulfilled to assure the existence  
101 of the readout curve. Here we simplified therefore improved construction of this curve. Now it is based  
102 on assumption that the main (i.e. dominant) frequency of sound input exist, is unique and is known. In  
103 addition to this known frequency, other parameters of the readout curve are set to make the fitting well  
104 posed and to obtain correct position of the curve maximum.

## Model scheme



**Figure 1. Schematic MSO circuit with anatomic and connection layouts.** This Figure with modifications is reprinted with permission from the manuscript of Sanda and Marsalek (2012) to show the numerical model key features.

### 105 2.4 Vector strength

106 The vector strength Kessler et al. (2021) has been first used in the context of sound localization by  
 107 Goldberg and Brown (1969). It is commonly used in auditory neuroscience to quantify how well a spike  
 108 train phase-locks, or synchronizes with periodic stimulus phases. Its definition follows. Let us have sample  
 109 spike phases  $\varphi_i$ ,  $i = 1, 2, \dots, N$  relative to phases of a given input master periodic function. Only the  
 110 phases enter the formula. The periodicity of vowel tones making up speech is a perfect example of such  
 111 stimulus. Discrete sum vector strength of samples  $\varphi_1, \dots, \varphi_N$  attains values from 0 to 1 and is defined as

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

112 Here we use the vector strength to analyze temporal acuity of inputs to the MSO model and the consequences  
 113 how the degree of synchronization affects the MSO precision.

### 114 2.5 JND and ideal observer

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

120 We will ask whether it is possible to distinguish between two random processes with rates  $l_1$  and  $l_2$ ,  
 121  $l_2 > l_1$ . If we count events in given counting window, we get counts  $n_1$  and  $n_2$ . The probability that the  
 122 observer obtains a result that  $l_2 > l_1$  equals to the probability that  $n_2 > n_1$ . Let us assume that the random  
 123 variables  $n_i, i = 1, 2, \dots$ , have probability distributions  $p(n_i)$  with means  $\mu_i$  and equal standard deviation  
 124  $\sigma$ . A detection distance is then defined Tanner Jr. (1961) as

$$d' = \frac{\mu_2 - \mu_1}{\sigma}. \quad (2)$$

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

$$\Delta l_{\text{JND}} = 1/\delta' = \frac{l_2 - l_1}{\mu_2 - \mu_1} \sigma. \quad (3)$$

131 We apply these tools to the MSO output to study the precision of the sound localization circuit to discuss  
 132 its dependence on input sound frequency.

### 3 RESULTS

133 We investigate how the MSO circuit output and overall performance depend on sound frequency and sound  
 134 intensity. Figures from 2 to 5 show results of numerical simulations obtained using the computational model  
 135 from Sanda and Marsalek (2012) side to side to experimental data as black lines and points, analytically  
 136 described lower estimates as blue lines, upper estimates as red lines and standard errors of measurement  
 137 and variations, which are shown in green.

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

142 Figure 2 shows how vector strength  $R_{\text{VS}}$  lowers towards higher frequencies, as it can be observed in the  
 143 module of the auditory periphery consisting of auditory nerve and cochlear nucleus. The prevailing majority  
 144 of neurons in the auditory pathway has vector strength spike train statistics sigmoidally dropping towards  
 145 higher sound frequencies as it is in this example. In this figure, data originally recorded by Joris (1996) at  
 146 the MSO of domestic cat, were fitted to the sigmoidal curve with the general formula of the Boltzmann  
 147 function used in Marsalek and Lansky (2005). The curve fit of vector strength  $R_{\text{VS}}$  in dependence on sound  
 148 frequency  $F_{\text{S}}$  is:

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

149 where named parameters with values are  $K_{\text{S}} = 2/0.75 = 2.666 \text{ ms (kHz}^{-1}\text{)}$ , sound frequency coefficient;  
 150  $K_{\text{C}} = 4/0.75 = 5.333 \text{ ms (kHz}^{-1}\text{)}$ , critical coefficient; and  $F_{\text{C}1/2} = 0.75 \text{ kHz}$ , critical half frequency. The  
 151 numerical values are the proportionality constants of the  $R_{\text{VS}}$  upper bound curve, denoted red. Note that at  
 152 sound frequencies from 20 to 100 Hz, there are two branches reflecting the existence of two alternative

153 ways how to compute the lower limit, which is 90 % of the upper limit. The first, blue curve corresponds  
 154 to quantile between 80 and 90 % of the upper limit. The second, green, lower curve is based on estimate  
 155 of probability of spike coming from both sides, which progressively lowers towards lower frequencies,  
 156 calculated as in Marsalek and Lansky (2005), see Discussion section.

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

$$\Delta t_{\text{JND}} = A(f_S - F_{C1/2})^2 + B. \quad (5)$$

159 Values of these parameters are  $A = 10^{-5}$ ,  $B = 0.05$ ,  $F_{C1/2} = 1$  kHz. The parameters in the Figures were  
 160 constructed as follows: firstly, splines were fitted to experimental data, published by [Mills 1958] and  
 161 reproduced by his followers. Splines parameters were allowed to vary within the 10 % of their original value  
 162 and were approximated by the rounded off values. These procedures were also used to get parameter values  
 163 below. For ranges of the audiogram parameters, see also Zwislocki and Feldman (1956). Analogously to  
 164 Figure 2, sound frequencies from 20 to 100 Hz exhibit higher spread between lower and upper limits, as  
 165 the fitting method used, quadratic fit, is the same for both limits.

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

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

178 **1.** fit of exponential function to simulations, red curve,

$$\Delta t_{\text{JND}} = \exp(A_1(t_J - B_1)) - C_1, \quad (6)$$

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

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

$$\Delta t_{\text{JND}} = A_2(t_J - B_2)^2 + C_2, \quad (7)$$

183 where  $A_2 = 2.5$ ,  $B_2 = 1$ ,  $C_2 = 1$ . There is only one parameter sought by numerical simulation. This is  $A$ ,  
 184 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  
 185 the model.

186 Logarithm of the simulated JND lowers with the exponential curve (6), which is concave function of  
 187 sought jitter  $t_J$  as the jitter gets lower. The trend towards higher accuracy diverges from the parabolic fit in

Parameter	Symbol	Units	Typical Value	Ranges
Timing Jitter	$t_J, \sigma$	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.

188 equation (7), when jitter reaches critical value between  $T_{JC}$  and  $T_{JO}$ .

$$T_{JC} = 0.2 \lesssim t_J \lesssim T_{JO} \lesssim 2\text{ms}. \quad (8)$$

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

192 This corresponds to the analytical dependency obtained in Salinas and Sejnowski (2000) for a perfect  
 193 integrator model with several inputs. The mechanism studied thereof is close to the MSO neural mechanism  
 194 studied here. The firing rate changes in dependency on the input spike timing variability of *partially*  
 195 *correlated* input spike trains.

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

199 In the numerical model of Sanda and Marsalek (2012) we reproduced a procedure to obtain azimuth  
 200 tuning curves based on hypotheses, how mammalian neural circuits work Grothe et al. (2010). In this  
 201 procedure, prior assumption of the existence of the main (dominant) sound frequency and neural tuning  
 202 to this frequency was used by experimentalists, but not by the authors of the numerical only model in  
 203 (Sanda and Marsalek, 2012). When we use this assumption, which is stronger than in the numerical only  
 204 model, in the calculations presented in this paper, we obtain a fitting curve which is more coherent. (Has  
 205 the higher vector strength value.) Two estimate errors are present in this Figure. The first is the mismatch  
 206 between the use of Gaussian (normal) probability density function, as it is used in some of the experimental  
 207 literature, and Sine function. Sine function is simpler circular statistics counterpart of the Gaussian, where  
 208 the rigorous circular statistics choice is the von Mises distribution. More details of the circular statistics  
 209 use in the sound localization context are explained in the article on the ergodicity assumption by Toth  
 210 et al. (2018). The difference between the two functions is shown here as the green curve. The other error,  
 211 not shown in Figure, would arise from the not using the prior assumption as in the numerical only model  
 212 mentioned above. It appears that the assumption of the main frequency existence leads to more precise  
 213 estimates. This should be, because the assumption adds more information besides the fitted data.

## 4 DISCUSSION

214 In this paper we have revisited numerical simulations by Sanda and Marsalek (2012). We have added  
 215 analytical estimations to the description of the MSO function, which have not been known previously.  
 216 Our analytical calculations make possible to derive time constants useful in description of normal human  
 217 hearing. Because of minimalist set of assumptions about stimuli, we hypothesize that the descriptions are

218 valid also for hearing with hearing aids and cochlear implants. All Figures contain model parameters and  
219 analytically expressed upper and lower limits of model transfer functions.

220 Figure 2 contains two lower limit branches at low frequencies (shown by the blue and green curves,  
221 respectively). The green curve uses an assumption of lower energy and lower contribution to spike rate in  
222 neural units in lower frequencies. Limiting lower bound by two different analytical functions (branches)  
223 can be understood as the estimate uncertainty. A conservative estimate of the lower bound always considers  
224 the lower of the two branches. This uncertainty should be recognized as one of original results presented  
225 for the first time in this paper. Its existence has been proposed in a doctoral thesis by Bures (2014). To our  
226 knowledge this observation has not yet been published elsewhere.

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

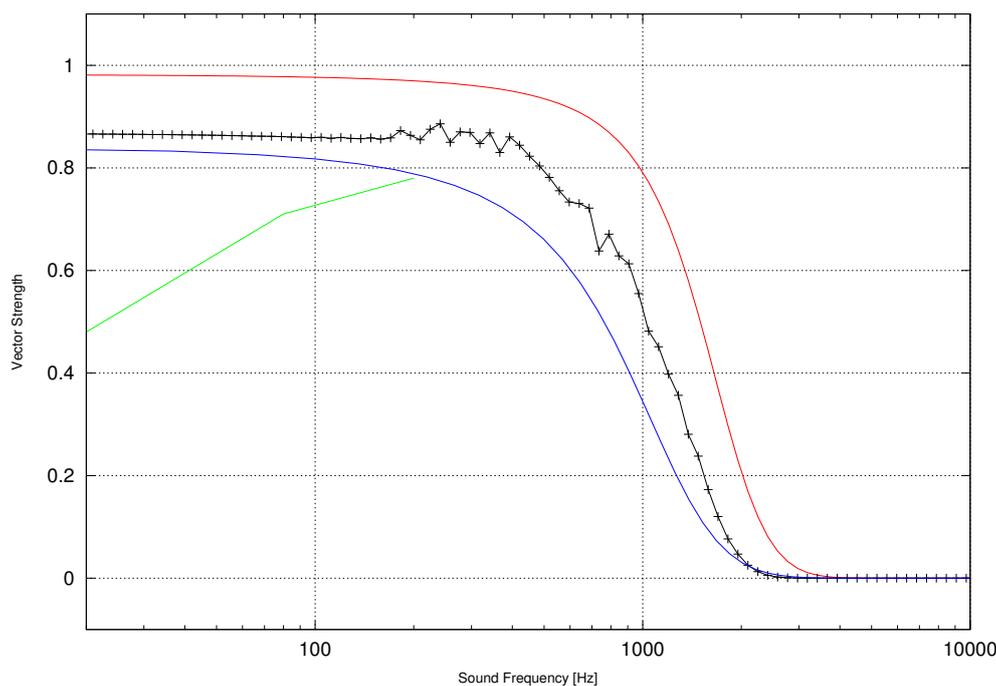
231 Figure 4 calculations use assumptions about intrinsic noise Bures (2014). The simulation data have been  
232 obtained by arbitrary precision calculation. Any neural data recording cannot reach this precision due to  
233 the internal noise of both neurons and recording electronics. In order to capture circuit noisiness, we have  
234 used both exponential and quadratic fits. When we attempt to use them as upper and lower bounds, we  
235 notice that they exchange their order in region close to the optimal jitter value. In other words, at the lower  
236 jitter values the two estimates exchange their ordering. This is the choice of the quadratic fit to obtain a  
237 normalized bound together with other data-points. Numerical simulation with the basic set of parameters  
238 around the x-axis value of  $t_J = 1$  ms lies beyond this point, but close to the exponential fit.

239 Figure 5 contains better fit of the Sine function, as compared to Sanda and Marsalek (2012). As the  
240 lower bound we can also use circular normal density (von Mises) function, the difference is negligible,  
241 not shown Toth et al. (2018). Comparison of time constants and sound periods in the model presented  
242 here will answer a tentative question: What is the highest slope of the ITD interpolation curve, such that it  
243 gives the resolution of the well known minimum audible angle in the midline (ITD = 0), which is  $4^\circ$  in  
244 angular degrees? This slope is more steep in higher frequency sounds, its maximum is attained in maximum  
245 frequency of the MSO circuit operation, around 1 kHz Marsalek and Lansky (2005).

246 Towards the analytical descriptions it is important to note that other periodic functions can be used as  
247 the ITD readout curves. In Toth et al. (2018) we have compared the *Sine function* with the *circular beta*  
248 *density*, and other alternative functions. To impose periodic and infinite boundaries to the problem, regular  
249 normal density and circular normal (von Mises) density have been used and tested in previous versions of  
250 our model. No differences between these densities with proper parameters have been shown by common  
251 statistical tests when testing differences between two probability densities, *ibidem*.

252 Franken et al. (2014) use recorded spike trains of several nuclei in the MSO neural circuit to demonstrate  
253 that coincidence detection is an essential part of the neuronal arithmetic executed by the neural circuit,  
254 see also Bures (2012). These authors show simulations combined with experimental description of MSO  
255 workings in line with findings of this paper and with coincidence detection theories. Another MSO model,  
256 already studied in 2005 is: Zhou et al. (2005), this is an example of simplistic model, motivating the  
257 neural circuit description presented also in this paper. For discussion of neural coding in the auditory nerve,  
258 auditory pathway, cochlear implants and brainstem neural circuits see Kerber and Seeber (2012).

259 Our investigation of quantitative properties of the superior olive neural circuit is also motivated by the  
260 three LSO experimental papers, which have detailed methodology applicable to LSO, to the overlap of



**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.

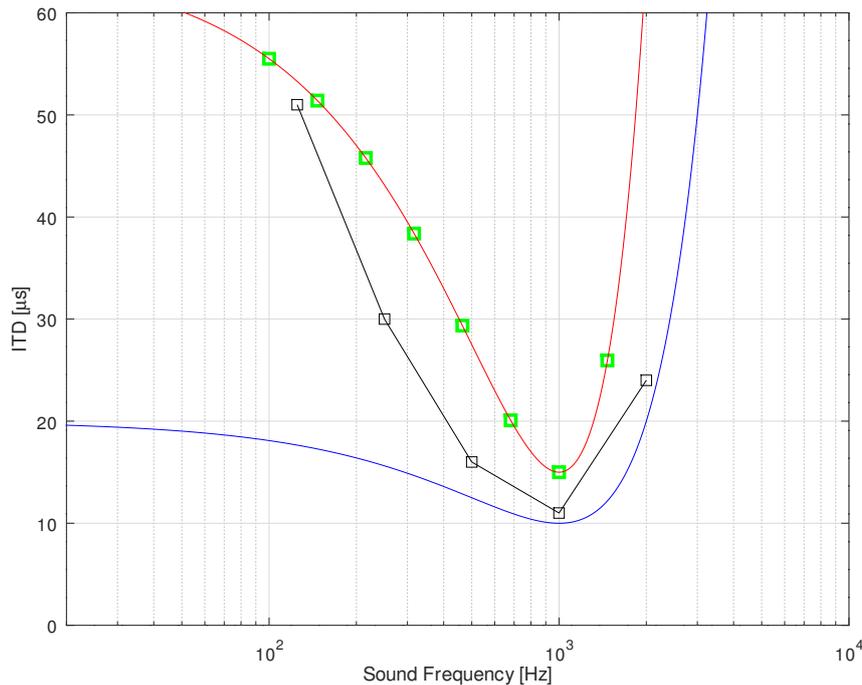
261 sound frequency ranges between the LSO and the MSO; and also to the MSO range itself; Joris and Yin  
 262 (1995); Joris (1996); Joris and Yin (1998) in experiments on the domestic cat.

263 Following his (notorious) paper from 1948, Lloyd A. Jeffress dedicated lots of efforts to the search of a  
 264 mechanism, by which microsecond time scale events of directional sound difference can be transformed  
 265 into a code processed and transmitted by action potentials lasting several microseconds Jeffress et al. (1962).  
 266 Historical comments on Jeffress papers from 1948 and 1962 are summarized by Cariani in Scholarpedia  
 267 Cariani (2011). A plausible explanation of the microsecond precision of the MSO circuit can be based on  
 268 descriptions of the neural computation using leading edges of action potentials and post-synaptic potentials,  
 269 Marsalek (2000), Toth et al. (2018).

270 Comparative physiology is useful in showing how the same mechanism works in other animals and what  
 271 alternatives are found in the phylogeny tree. Let us mention just two examples here:

272 **1.** *Ormia ochracea* (fly parasiting on cricket, *Gryllus campestris*) has the same precision of motor response  
 273 in ITD processing as human. The ‘computational unit’ there is chitin lever connecting two eardrums.

274 **2.** Much higher precision demonstrated by behavioral response is described in bats. *Eptesicus fuscus* (brown  
 275 bat) catches tidbit larvae of the meal-worm beetle *Tenebrio molitor*. The motor precision is in the range of  
 276 hundreds of nanoseconds, 100 ns, and auditory separation of ultrasound echoes in the bats sonar sense is in  
 277 the order of microsecond units, 1  $\mu$ s Simmons et al. (1998).

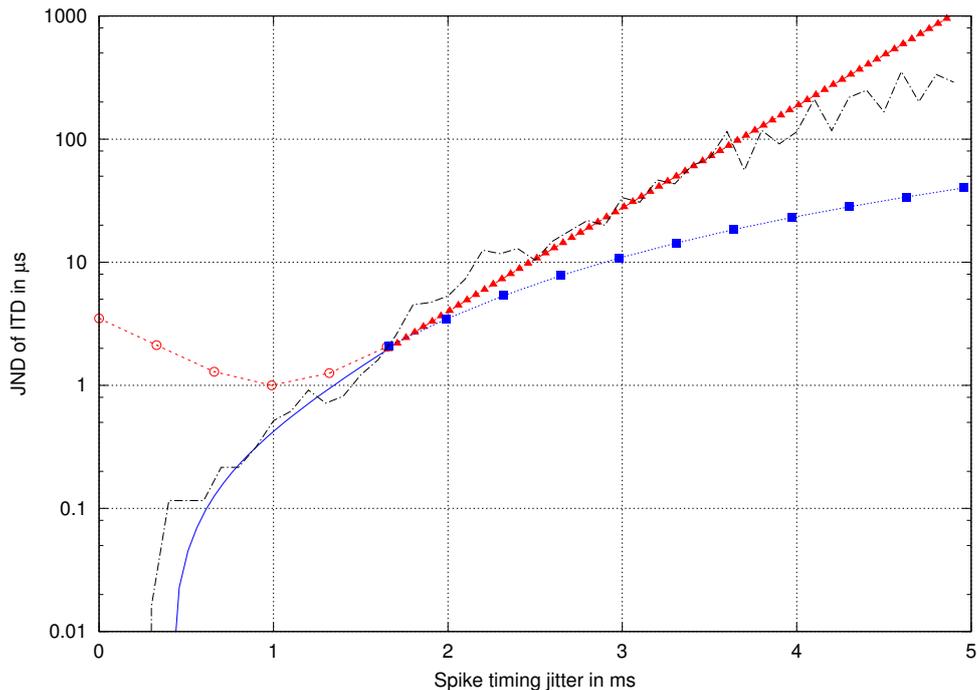


**Figure 3. The shortest JND of ITD detected in the dependence on sound frequency.** X-axis shows sound frequency in Hz in a logarithmic scale and y-axis the shortest JND of ITD in  $\mu\text{s}$ . This is a theoretical prediction based on the analytical model and basic parameter set used in simulations. 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.

278 In contemporary human, the MSO is the larger of the two nuclei and contains approximately 10000 -  
 279 11000 neurons, while the human LSO contains 5600 neurons Moore (2000). To implement the loudness  
 280 change is much simpler than to record and implement microsecond time delay. Therefore in sound  
 281 generation and processing, most of current auditory technology works as if the more important of the two  
 282 localization cues in *Homo sapiens* were the sound intensity cue Vencovsky and Rund (2016), even though  
 283 the evidence is far not definitive.

284 In Marsalek (2000), individual steps of signal processing in the superior olive neural circuits have been  
 285 investigated. Various synaptic mechanisms have been proposed Marsalek and Kofranek (2005). Spike  
 286 timing jitter and spike variability have been systematically analytically investigated by Kostal and Marsalek  
 287 (2010).

288 In this last paragraph of Discussion section, we should mention briefly rest of scientific papers, where we  
 289 found ideas towards the design of the analytically tractable model presented here. Article by Michelet et al.  
 290 (2012) discusses interaural phase delays (IPDs; when they exist, their utility is equivalent to that of the  
 291 ITDs) and cochlear delays. In the case of cochlear delays, very important is to review the ranges of delays  
 292 in comparison to sound periods and classically described excitatory-excitatory and excitatory-inhibitory  
 293 responses to binaural inputs in Joris et al. (2006). Paper of Srinivasan, Laback and Majdak cites current  
 294 progress of ITD encoding by binaural cochlear implants, this is important for model validations and  
 295 applications to studies with hearing aids and electrical hearing, Srinivasan et al. (2018).



**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.

## CONCLUSIONS

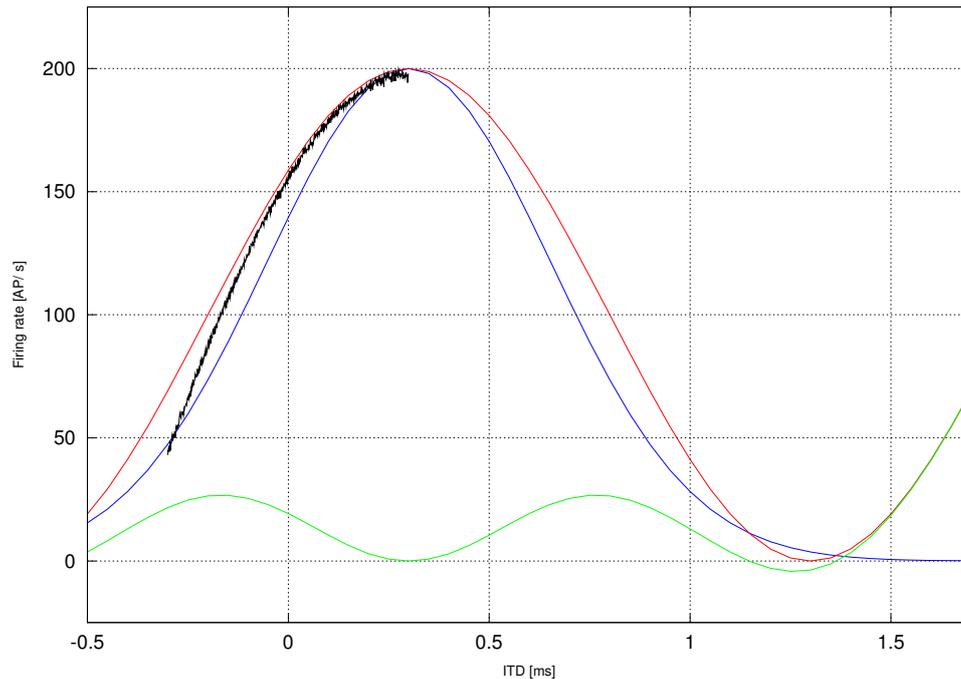
296 This theoretical paper is continuation of sound localization precision descriptions in the MSO Sanda  
 297 and Marsalek (2012) and in the LSO Bures and Marsalek (2013). Major novel results here are two: 1)  
 298 analytical estimates of results obtained previously only by numerical simulation and 2) estimates of auditory  
 299 parameters and functions bounding from the bottom and from the top known characteristics of the human  
 300 sound localization circuit.

## CONFLICT OF INTEREST STATEMENT

301 The authors declare that the research was conducted in the absence of any commercial or financial  
 302 relationships that could be construed as a potential conflict of interest.

## AUTHOR CONTRIBUTIONS

303 Model design: PM, ZB. Methodology: PM, PS. Writing original draft: PM. Writing and editing: PM, PS,  
 304 ZB.



**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  $t_{ITD} = 0$ . 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.

## FUNDING

305 This project was in part funded by Charles University graduate students research program, acronym SVV,  
 306 No. 260 519/ 2020-2022, to Petr Marsalek.

## ACKNOWLEDGMENTS

307 We acknowledge uploading this manuscript first to ArXiv repository prior to impacted journal submission.  
 308 Special thanks to Marek Hajny for reading and commenting on the manuscript.

## DATA AVAILABILITY STATEMENT

309 The datasets and source codes of programs used in this study can be found at the web pages of PM, see  
 310 link <http://patf-onyx.lf1.cuni.cz/mlab/ftp/>.

## REFERENCES

- 311 Bures, Z. (2012). The stochastic properties of input spike trains control neuronal arithmetic. *Biol. Cybern.*  
 312 106, 111–122. doi:10.1007/s00422-012-0483-9  
 313 Bures, Z. (2014). *Internal Representation and Processing of Acoustic Stimuli in the Nervous System.*  
 314 *Habilitation thesis* (Brno, Czech Republic: VUTIUUM (Publishing House of the Brno University of  
 315 Technology)). 122 pages

- 316 Bures, Z. and Marsalek, P. (2013). On the precision of neural computation with the interaural level  
317 difference in the lateral superior olive. *Brain Res.* 1536, 111–122. doi:10.1016/j.brainres.2013.05.008
- 318 Cariani, P. (2011). Jeffress model. *Scholarpedia* 6, 2920. doi:10.4249/scholarpedia.2920
- 319 Franken, T. P., Bremen, P., and Joris, P. X. (2014). Coincidence detection in the medial superior olive:  
320 mechanistic implications of an analysis of input spiking patterns. *Front. Neural Circuits* 8, 1–21
- 321 Goldberg, J. M. and Brown, P. B. (1969). Response of binaural neurons of dog superior olivary complex  
322 to dichotic tonal stimuli: some physiological mechanisms of sound localization. *J. Neurophysiol.* 32,  
323 613–636
- 324 Grothe, B., Pecka, M., and McAlpine, D. (2010). Mechanisms of sound localization in mammals. *Physiol.*  
325 *Rev.* 90, 983–1012
- 326 Jeffress, L. A., Blodgett, H. C., and Deatherage, B. H. (1962). Effect of interaural correlation on the  
327 precision of centering a noise. *J. Acoust. Soc. Am.* 34, 1122–1123
- 328 Joris, P. X., Van de Sande, B., Louage, D. H., and van der Heijden, M. (2006). Binaural and cochlear  
329 disparities. *Proc. Natl. Acad. Sci. USA* 103, 12917–12922
- 330 Joris, P. X. and Yin, T. C. (1995). Envelope coding in the lateral superior olive. I. Sensitivity to interaural  
331 time differences. *J. Neurophysiol.* 73, 1043–1062
- 332 Joris, P. X. (1996). Envelope coding in the lateral superior olive. II. Characteristic delays and comparison  
333 with the responses in the medial superior olive. *J. Neurophysiol.* 76, 2137–2156
- 334 Joris, P. X. and Yin, T. C. (1998). Envelope coding in the lateral superior olive. III. Comparison with  
335 afferent pathways. *J. Neurophysiol.* 79, 253–69
- 336 Kerber, S. and Seeber, B. U. (2012). Sound localization in noise by normal-hearing listeners and cochlear  
337 implant users. *Ear Hear.* 33, 445–457
- 338 Kessler, D., Carr, C. E., Kretzberg, J., and Ashida, G. (2021). Theoretical relationship between two  
339 measures of spike synchrony: correlation index and vector strength. *Front. Neurosci.* 15, 761826. 1-14
- 340 Kostal, L. and Marsalek, P. (2010). Neuronal jitter: can we measure the spike timing dispersion differently?  
341 *Chinese J. Physiol.* 53, 454–464
- 342 Kulesza Jr, R. J. (2007). Cytoarchitecture of the human superior olivary complex: medial and lateral  
343 superior olive. *Hear. Res.* 225, 80–90
- 344 Laback, B. and Majdak, P. (2008). Binaural jitter improves interaural time-difference sensitivity of cochlear  
345 implantees at high pulse rates. *Proc. Natl. Acad. Sci. USA* 105, 814–817
- 346 Marsalek, P. (2000). Coincidence detection in the Hodgkin-Huxley equations. *Biosystems* 58, 83–91
- 347 Marsalek, P. and Kofranek, J. (2005). Spike encoding mechanisms in the sound localization pathway.  
348 *Biosystems* 79, 191–8
- 349 Marsalek, P. and Lansky, P. (2005). Proposed mechanisms for coincidence detection in the auditory  
350 brainstem. *Biol. Cybern.* 92, 445–451
- 351 Michelet, P., Kovacic, D., and Joris, P. X. (2012). Ongoing temporal coding of a stochastic stimulus as a  
352 function of intensity: time-intensity trading. *J. Neurosci.* 32, 9517–9527
- 353 Middlebrooks, J. C. and Green, D. M. (1991). Sound localization by human listeners. *Annu. Rev. Psychol.*  
354 42, 135–159
- 355 Mills, A. W. (1958). On the minimum audible angle. *J. Acoust. Soc. Am.* 30, 237–246
- 356 Moore, J. K. (2000). Organization of the human superior olivary complex. *Microsc. Res. Tech.* 51, 403–412
- 357 Salinas, E. and Sejnowski, T. J. (2000). Impact of correlated synaptic input on output firing rate and  
358 variability in simple neuronal models. *J. Neurosci.* 20, 6193–6209
- 359 Sanda, P. and Marsalek, P. (2012). Stochastic interpolation model of the medial superior olive neural  
360 circuit. *Brain Res.* 1434, 257–265. doi:10.1016/j.brainres.2011.08.048

- 361 Simmons, J. A., Ferragamo, M. J., and Moss, C. F. (1998). Echo-delay resolution in sonar images of the  
362 big brown bat, *Eptesicus fuscus*. *Proc. Natl. Acad. Sci. USA* 95, 12647–12652
- 363 Srinivasan, S., Laback, B., Majdak, P., and Delgutte, B. (2018). Introducing short interpulse intervals in  
364 high-rate pulse trains enhances binaural timing sensitivity in electric hearing. *J. Assoc. Res. Otolaryngol.*  
365 19, 301–315
- 366 Tanner Jr., W. P. (1961). Physiological implications of psychophysical data. *Ann. N. Y. Acad. Sci.* 89,  
367 752–765
- 368 Toth, P. G., Marsalek, P., and Pokora, O. (2018). Ergodicity and parameter estimates in auditory neural  
369 circuits. *Biol. Cybern.* 112, 41–55
- 370 Vencovsky, V. and Rund, F. (2016). Using a physical cochlear model to predict masker phase effects in  
371 hearing-impaired listeners: a role of peripheral compression. *Acta Acust. United Acust.* 102, 373–382
- 372 Zhou, Y., Carney, L. H., and Colburn, H. S. (2005). A model for interaural time difference sensitivity in  
373 the medial superior olive: interaction of excitatory and inhibitory synaptic inputs, channel dynamics, and  
374 cellular morphology. *J. Neurosci.* 25, 3046–3058
- 375 Zwislowski, J. and Feldman, R. S. (1956). Just noticeable differences in dichotic phase. *J. Acoust. Soc. Am.*  
376 28, 860–864