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]

[a]Department of Natural Sciences, Faculty of Biomedical Engineering, Czech Technical University in Prague, nám. Sítná 3105, 272 01 Kladno, Czech Republic
[b]Institute of Pathological Physiology, First Medical Faculty, Charles University in Prague, U Nemocnice 5/478, 128 53, Praha 2, 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
[c]*Corresponding author:* Petr.Marsalek@LF1.CUNI.CZ

December 30, 2021

Abstract

In the auditory nerve and the following auditory pathway, incoming sound is encoded into spike trains – series of neural action potentials. At the third neuron of the auditory pathway, spike trains of the left and right sides converge and are processed to yield sound localization information. Two different localization encoding mechanisms are employed for low and high sound frequencies in two dedicated nuclei in the brainstem: the medial and lateral superior olivary nuclei. Building upon our previous computational model of medial superior olive (MSO), this paper brings analytical estimates of parameters needed to describe auditory coding in the MSO circuit. We arrive to best estimates for neuronal signaling with the use of just noticeable difference and the ideal observer concepts. We describe spike timing jitter and its role in the spike train processing. We study the dependence of sound localization precision on the sound frequency. All parameters are accompanied with detailed estimates of their values and variability. Intervals bounding all the parameters 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_{S}, sound frequency; F_{C}, critical sound frequency value; φ, 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_{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_{VS}, VS, vector strength; R_F, firing rate; σ, standard deviation; t_{J} timing jitter; t, Δt, time, time difference; T, T_{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, respectively. Neurons in these nuclei are the first binaural neurons in the auditory pathway, processing the information arriving from both ears. 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 [Middlebrooks and Green(1991)]. Due to the physical nature of the binaural stimulus, the MSO neurons process spike timing differences, Interaural Time Differences, ITDs. The ITDs emerge as a result of different distance from the sound source to each ear. The ITDs take values given by the distance between the ears (in human this is circa 16 cm/ 466 μs) and are detected with precision of tens of microseconds, 10 μs [Laback and Majdak(2008)]. The LSO detects the Interaural Level Difference, ILD, given by the acoustic shadow of the head. The frequency region, where these two mechanisms overlap (around 1.5 kHz) is known to have a drop in localization sensitivity [Mills(1958)].

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

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

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 [Sanda and Marsalek(2012)]. We compare this analytical MSO description to the LSO description in [Bures and Marsalek(2013)] to arrive to unified description of neural circuits in the superior olive. We use this description to find the performance limits of the MSO circuit in dependence on sound frequency and intensity. Apparently, low and high frequency sound localization use different neuronal mechanisms. Sound features in low and high sound frequencies are encoded by distinctive codes and this multitude of codes also affects binaural processing.

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 [Sanda and Marsalek(2012)]. 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 [Kulesza Jr(2007)], 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, [Mills(1958)]. In some animal specialists, some responses are in the range of tens of nanoseconds, as it has been shown in experiment on bats by [Simmons et al.(1998)Simmons, Ferragamo, and Moss]. 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, where 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 the binaural neurons is the azimuth signal encoded in a spike train. However, components of our computational model can not be identified with individual neurons and their anatomical connectivity. Rather, the model consists of functional block units representing the crucial properties of the explored neural mechanism, such as intrinsic noise, temporal delay and coincidence detection. This functional block model layout provides access to descriptive parameters, which are setting the circuit's performance limits, see Figure 1. In our MSO model, the interaural delay (ITD) present at the MSO input is represented by firing rate at the MSO output. The relationship between the input ITD and output spike rate is called ITD readout curve, see the next section.

2.3 ITD readout curve

Let us have a monotonous function with firing rate as an input, which outputs azimuth. We call it the ITD readout curve. In the paper by [Sanda and Marsalek(2012)] this curve is constructed by curve fitting to simulated points. Several assumptions about the sound objects have to be fulfilled to assure the existence of the readout curve. Here we improved construction of this curve. Now it is based on assumption that the main (i.e. dominant) 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 [Kessler et al.(2021)Kessler, Carr, Kretzberg, and Ashida] has been first used in the context of sound localization by [Goldberg and Brown(1969)]. It is commonly used in auditory neuroscience to quantify how well a spike train phase-locks, or synchronizes with periodic stimulus phases. Its definition follows. Let us have sample spike phases φ_i, $i = 1, 2, \ldots, N$ relative to phases of a given input master periodic function. Only the phases enter the formula. The periodicity of vowel tones making up speech is a perfect example of such stimulus. Discrete sum vector strength of samples $\varphi_1, \ldots, \varphi_N$ attains values from 0 to 1 and is defined as

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

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

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 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 μ_i and equal standard deviation σ. A detection distance is then defined [Tanner Jr.(1961)] 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 Δl and put $\delta' = d'/\Delta l$. Then, the JND of firing rate is

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

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

3 Results

We investigate how the MSO circuit output and overall performance depend on sound frequency and sound intensity. Figures from2 to 5 show results of numerical simulations obtained using the computational model from [Sanda and Marsalek, 2012] side to side to 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, which are shown in green.

Successive figures show individual steps of sound localization signal 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 precision. Fig. 4 shows the JND of the neural circuit in the dependence on the spike timing jitter. Fig. 5 shows synchronization to main sound frequency (in the case when it exists).

Figure2 shows how vector strength R_{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 [Joris(1996)] at the MSO of domestic cat, were fitted to the sigmoidal curve with the general formula of the Boltzmann function used in [Marsalek and Lansky(2005)]. The curve fit of vector strength R_{VS} in dependence on sound frequency F_{S} is:

$$R_{\mathrm{VS}} = 1/(1 + \exp(K_{\mathrm{S}} f_{\mathrm{S}} - K_{\mathrm{C}} F_{\mathrm{C1/2}})), \tag{4}$$

where named 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 $F_{\mathrm{C1/2}} = 0.75$ kHz, critical half frequency. The numerical values are the proportionality constants of the R_{VS} upper bound curve, denoted red. 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. The first, blue curve corresponds to quantile between 80 and 90 % of the upper limit. The second, green, lower curve is based on estimate of probability of spike coming from both sides, which progressively lowers towards lower frequencies, calculated as in [Marsalek and Lansky(2005)], 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 Δt_{JND} is:

$$\Delta t_{\mathrm{JND}} = A(f_{\mathrm{S}} - F_{\mathrm{C1/2}})^2 + B, \tag{5}$$

Values of these parameters are $A = 10^{-5}$, $B = 0.05$, $F_{C1/2} = 1$ kHz. The parameters in the Figures were constructed as follows: firstly, splines were fitted to experimental data, published by [Mills 1958] and reproduced by his followers. 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 [Zwislocki and Feldman(1956)]. 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_{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_{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.

Logarithm of the simulated JND lowers with the exponential curve (6), which is concave function of sought jitter t_{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_{JC} and T_{JO}.

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

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 [Salinas and Sejnowski(2000)] 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 [Sanda and Marsalek(2012)] we reproduced a procedure to obtain azimuth tuning curves based on hypotheses, howmammalian neural circuits work [Grothe et al, 2010]. In this procedure, prior assumption of the existence of the main (dominant) sound frequency and neural tuning to this frequency was used by experimentalists, but not by the authors of the numerical only model in (Sanda and Marsalek, 2012). When we use this assumption, which is stronger than in the numerical only model, in the calculations presented in this paper, 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 some of the experimental literature, and Sine function. Sine function is simpler circular statistics counterpart of the Gaussian, where the rigorous circular statistics choice is the von Mises distribution. More details of the circular statistics use in the sound localization context are explained in the article on the ergodicity assumption by [Toth et al.(2018)Toth, Marsalek, and Pokora]. The difference between

Parameter	Symbol	Units	Typical Value	Ranges
Timing Jitter	t_{J}, σ	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.

the two functions is shown here as the green curve. The other error, not shown in Figure, would arise from the not using the prior assumption as in the numerical only model mentioned above. It appears that the assumption of the main frequency existence leads to more precise estimates. This should be, because the assumption adds more information besides the fitted data.

4 Discussion

In this paper we have revisited numerical simulations by [Sanda and Marsalek(2012)]. 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 of normal human hearing. Because of minimalistics set of assumptions about stimuli, we hypothesize that the descriptions are valid also for hearing with hearing aids and cochlear implants. All Figures contain model parameters and analytically expressed upper and lower limits of model transfer functions.

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 [Bures(2014)]. 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 [Bures(2014)]. The simulation data have been obtained by arbitrary precision calculation. Any neural data 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 $t_{\mathrm{J}} = 1$ ms lies beyond this point, but close to the exponential fit.

Figure 5 contains better fit of the Sine function, as compared to [Sanda and Marsalek(2012)]. As the lower bound we can also use circular normal density (von Mises) function, the difference is negligible, not shown [Toth et al.(2018)Toth, Marsalek, and Pokora]. 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 the 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 [Marsalek and Lansky(2005)].

Towards the analytical descriptions it is important to note that other periodic functions can be used as the ITD readout curves. In [Toth et al.(2018)Toth, Marsalek, and Pokora] 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 (von Mises) 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.

[Franken et al.(2014)Franken, Bremen, and Joris] 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 [Bures(2012)]. 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: [Zhou et al.(2005)Zhou, Carney, and Colburn], this is an example of simplistic model, motivating the neural circuit description presented also in this paper. For discussion of neural coding in the auditory nerve, auditory pathway, cochlear implants and brainstem neural circuits see [Kerber and Seeber(2012)].

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; [Joris and Yin(1995)]; [Joris(1996)]; [Joris and Yin(1998)] in experiments on the domestic cat.

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 [Jeffress et al.(1962)Jeffress, Blodgett, and Deatherage]. Historical comments on Jeffress papers are summarized by Cariani in Scholarpedia [Cariani(2011)]. A plausible explanation of the microsecond precision of the MSO circuit can be based on descriptions of the neural computation using leading edges of action potentials and post-synaptic potentials, [Marsalek(2000)], [Toth et al.(2018)Toth, Marsalek, and P

Comparative physiology is useful in showing how the same mechanism works in other animals and what alternatives are found in the phylogenetic tree. Just two examples here: *Ormia ochracea* (fly parasiting on cricket) has the same precision of motor response in ITD processing as human. The 'computational unit' there is chitin lever connecting two eardrums. Much higher precision demonstrated by behavioral response is described in bats. *Eptesicus fuscus* (brown bat) 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 μs [Simmons et al.(1998)Simmons, Ferragamo, and Moss].

In contemporary human, the MSO is the larger of the two nuclei and contains approximately 10000 - 11000 neurons, while the human LSO contains 5600 neurons [Moore(2000)]. 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 [Vencovsky and Rund(2016)], even though the evidence is far not definitive.

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

In this last paragraph of Discussion section, we should mention briefly rest of scientific papers, where we found ideas towards the design of the analytically tractable model presented here. Article by [Michelet et al.(2012)Michelet, Kovacic, a discusses interaural phase delays (IPDs; when they exist, their utility is equivalent to that of the ITDs) and cochlear delays. In the case of 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 [Joris et al.(2006)Joris, Van de Sande, Louage, and van der Heijden]. 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, [Srinivasan et al.(2018)Srinivasan, Laback, Majdak, and Delgutte].

Conclusions

This theoretical paper is continuation of sound localization precision descriptions in the MSO [Sanda and Marsalek(2012)] and in the LSO [Bures and Marsalek(2013)]. 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

This project was in part funded by Charles University graduate students research program, acronym SVV, No. 260 519/ 2020-2022, to Petr Marsalek. We acknowledge uploading this manuscript first to ArXiv repository prior to impacted journal submission. Special thanks to Marek Hajny for reading and commenting the manuscript.

Author contributions

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

Technical, TEX Typesetting, Archiving Notes

Compilation notes: LATEX compilation date is December 30, 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

[Middlebrooks and Green(1991)] Middlebrooks JC, Green DM. Sound localization by human listeners. *Annu. Rev. Psychol.* **42** (1991) 135–159.

[Laback and Majdak(2008)] Laback B, Majdak P. Binaural jitter improves interaural time-difference sensitivity of cochlear implantees at high pulse rates. *Proc. Natl. Acad. Sci. USA* **105** (2008) 814–817.

[Mills(1958)] Mills AW. On the minimum audible angle. *J. Acoust. Soc. Am.* **30** (1958) 237–246.

[Grothe et al.(2010)Grothe, Pecka, and McAlpine] Grothe B, Pecka M, McAlpine D. Mechanisms of sound localization in mammals. *Physiol. Rev.* **90** (2010) 983–1012.

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

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

[Kulesza Jr(2007)] Kulesza Jr RJ. Cytoarchitecture of the human superior olivary complex: medial and lateral superior olive. *Hear. Res.* **225** (2007) 80–90.

[Simmons et al.(1998)Simmons, Ferragamo, and Moss] Simmons JA, Ferragamo MJ, Moss CF. Echo-delay resolution in sonar images of the big brown bat, *Eptesicus fuscus*. *Proc. Natl. Acad. Sci. USA* **95** (1998) 12647–12652.

[Kessler et al.(2021)Kessler, Carr, Kretzberg, and Ashida] Kessler D, Carr CE, Kretzberg J, Ashida G. Theoretical relationship between two measures of spike synchrony: correlation index and vector strength. *Front. Neurosci.* **15** (2021) 761826. 1-14.

[Goldberg and Brown(1969)] 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** (1969) 613–636.

[Tanner Jr.(1961)] Tanner Jr WP. Physiological implications of psychophysical data. *Ann. N. Y. Acad. Sci.* **89** (1961) 752–765.

[Joris and Yin(1995)] Joris PX, Yin TC. Envelope coding in the lateral superior olive. I. Sensitivity to interaural time differences. *J. Neurophysiol.* **73** (1995) 1043–1062.

[Joris(1996)] 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** (1996) 2137–2156.

[Joris and Yin(1998)] Joris PX, Yin TC. Envelope coding in the lateral superior olive. III. Comparison with afferent pathways. *J. Neurophysiol.* **79** (1998) 253–69.

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

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

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

[Toth et al.(2018)Toth, Marsalek, and Pokora] Toth PG, Marsalek P, Pokora O. Ergodicity and parameter estimates in auditory neural circuits. *Biol. Cybern.* **112** (2018) 41–55.

[Bures(2014)] 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.

[Franken et al.(2014)Franken, Bremen, and Joris] Franken TP, Bremen P, Joris PX. Coincidence detection in the medial superior olive: mechanistic implications of an analysis of input spiking patterns. *Front. Neural Circuits* **8** (2014) 1–21.

[Bures(2012)] Bures Z. The stochastic properties of input spike trains control neuronal arithmetic. *Biol. Cybern.* **106** (2012) 111–122.

[Zhou et al.(2005)Zhou, Carney, and Colburn] Zhou Y, Carney LH, Colburn HS. 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** (2005) 3046–3058.

[Kerber and Seeber(2012)] Kerber S, Seeber BU. Sound localization in noise by normal-hearing listeners and cochlear implant users. *Ear Hear.* **33** (2012) 445–457.

[Jeffress et al.(1962)Jeffress, Blodgett, and Deatherage] Jeffress LA, Blodgett HC, Deatherage BH. Effect of interaural correlation on the precision of centering a noise. *J. Acoust. Soc. Am.* **34** (1962) 1122–1123.

[Cariani(2011)] Cariani P. Jeffress model. *Scholarpedia* **6** (2011) 2920. doi:10.4249/scholarpedia.2920.

[Marsalek(2000)] Marsalek P. Coincidence detection in the Hodgkin-Huxley equations. *Biosystems* **58** (2000) 83–91.

[Moore(2000)] Moore JK. Organization of the human superior olivary complex. *Microsc. Res. Tech.* **51** (2000) 403–412.

[Vencovsky and Rund(2016)] Vencovsky V, 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** (2016) 373–382.

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

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

[Michelet et al.(2012)Michelet, Kovacic, and Joris] Michelet P, Kovacic D, Joris PX. Ongoing temporal coding of a stochastic stimulus as a function of intensity: time-intensity trading. *J. Neurosci.* **32** (2012) 9517–9527.

[Joris et al.(2006)Joris, Van de Sande, Louage, and van der Heijden] Joris PX, Van de Sande B, Louage DH, van der Heijden M. Binaural and cochlear disparities. *Proc. Natl. Acad. Sci. USA* **103** (2006) 12917–12922.

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

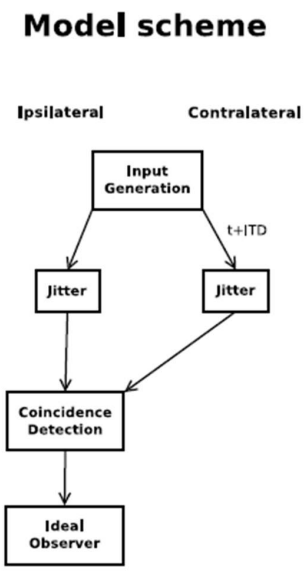

Figure 1: **Schematic MSO circuit with anatomic and connection layouts.** This Figure with modifications is reprinted with permission from the manuscript of [Sanda and Marsalek(2012)] to show the numerical model key features.

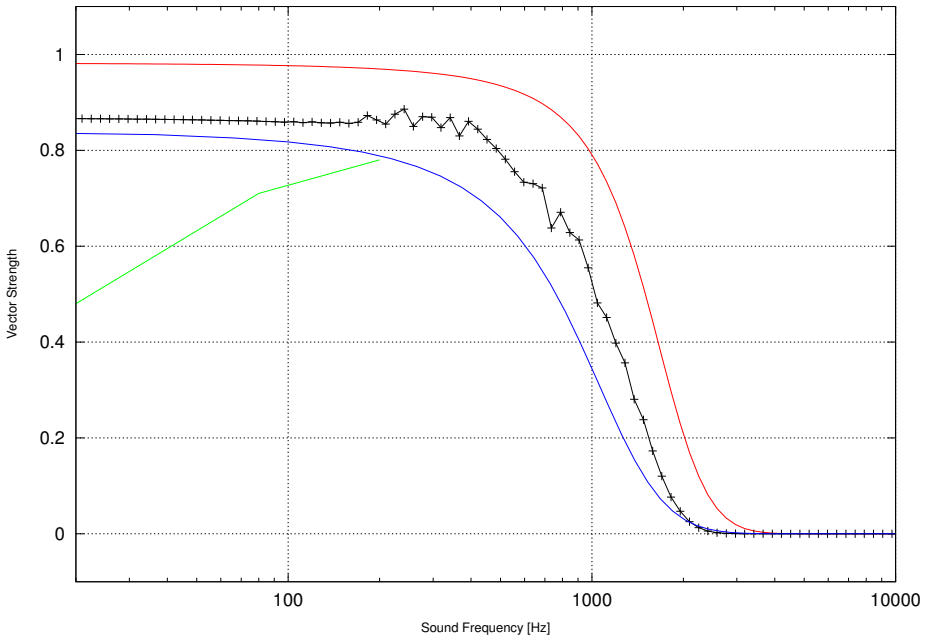

Figure 2: **Vector strength of auditory nerve spike trains in dependence on sound frequency.** X-axis shows sound frequency in Hz in logarithmic scale and y-axis shows the vector strength. Even though in some nuclei up the auditory pathway the synchronization can be maintained towards higher frequencies 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.

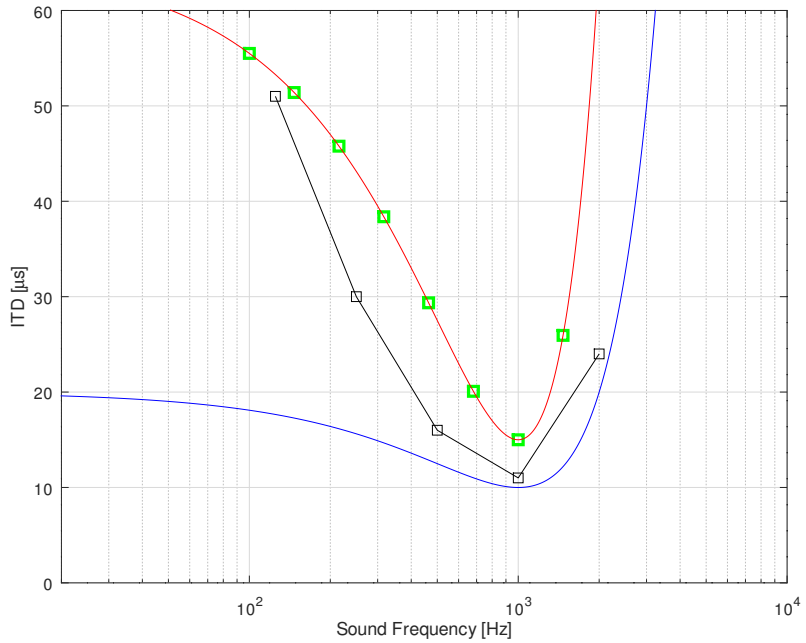

Figure 3: **The shortest JND of ITD detected in the dependence on sound frequency.** X-axis shows sound frequency in Hz in a logarithmic scale and y-axis the shortest JND of ITD in μs. This is a theoretical prediction based on the analytical model and basic parameter set used in simulations. As in other figures, black line is obtained by simulation and red and blue lines are respectively upper and lower bounds obtained by an analytic fit.

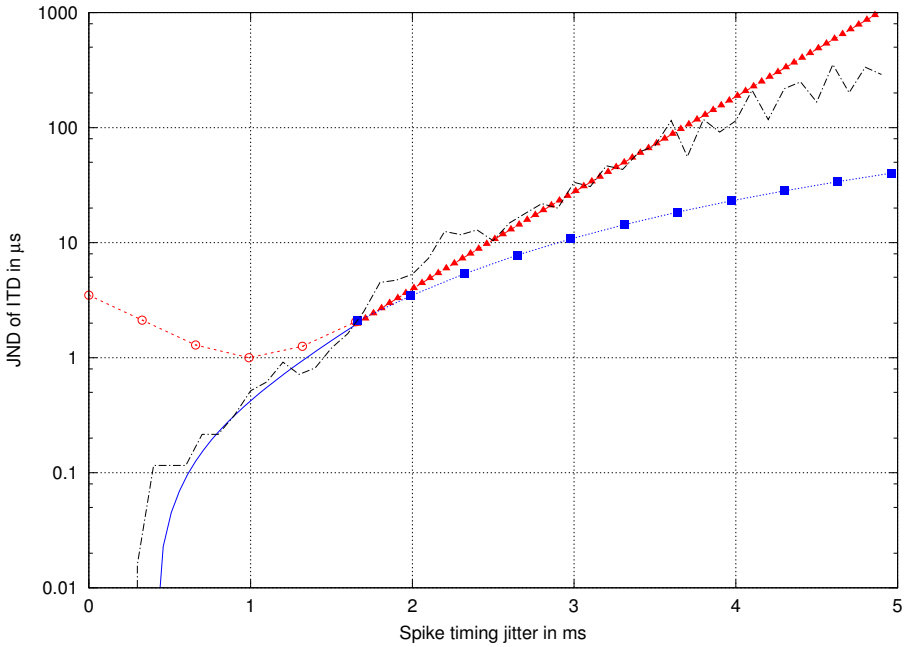

Figure 4: **JND values in basic parameter set in dependence on the spike timing jitter magnitude.** This plot in semi-logarithmic y-scale shows JND (just noticeable difference) of interaural time difference depending on variation of the spike timing jitter. Jagged black line: simulated data, solid line: 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 $f_{\mathrm{S}} = 1.66$ 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 $f_{\mathrm{S}} = 1.66$ 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.

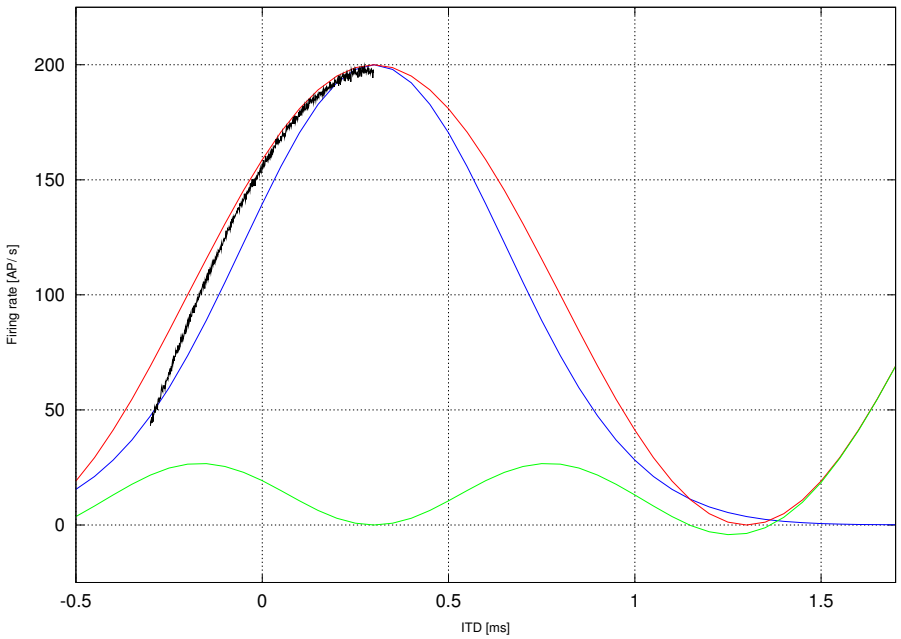

Figure 5: **Fit of example functions to firing rate slope ITD readout curve.** X-axis shows ITD in ms and y-axis shows corresponding firing rate in action potentials per second. Note that the curve peak is offset from the origin of coordinates at $t_{\mathrm{ITD}} = 0$. Red curve is the Sine density 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.