Author Accepted Manuscript (contents identical to the full published version). This paper was presented at the 132nd Convention of the Audio Engineering Society 2012, as paper number 8677. The full published version can be found at `http://www.aes.org/e-lib/browse.cfm?elib=16315`.

Headphone Selection for Binaural Synthesis with Blocked Auditory Canal Recording

Florian Völk

AG Technische Akustik, MMK, Technische Universität München, Arcisstraße 21, 80333 München, Germany

Correspondence should be addressed to Florian Völk (`florian.voelk@mytum.de`)

ABSTRACT

Binaural synthesis aims at eliciting the reference scene hearing sensations by recreating the sound pressures at the eardrums, typically using headphones. If all transfer functions involved are approximated based on eardrum probe microphone or traditional artificial head measurements, the headphones have been shown not to influence the synthesis. It is also possible to achieve correct binaural synthesis with transfer functions measured at the entrances to the blocked auditory canals. Then, the headphones may influence the results. In this paper, a blocked auditory canal headphone selection criterion (HPSC) for binaural synthesis is proposed. Further, a procedure is derived, which allows to evaluate the HPSC for specific circum-aural headphones based on four measurements using a specifically designed artificial head.

1. INTRODUCTION

It is the goal of binaural synthesis to elicit the hearing sensations of a reference scene using headphones (HPs) in the playback situation [1]. The procedure is based on recreating the reference scene ear signals, the sound pressures at the eardrums. Typically, transfer functions (TFs) are recorded or modeled between the sources to be simulated and a reference position in the auditory canal of a human or an artificial head [2]. Using the adequate procedure, it is possible to equalize the overall binaural synthesis TFs based on artificial head (AH) or probe microphone recordings and corresponding headphone impulse responses (HPIRs) regardless of the specific HP model employed [2]. For binaural synthesis using blocked auditory canal recordings and blocked auditory canal HPIR based equalization on the contrary, the equations 74 and 80 given by [2] can be interpreted in that this procedure possibly results in erroneous ear signals.

In this paper, the question is addressed whether HPs can be found that allow for the synthesis of the reference scene ear signals. Appropriate HPs in this context do according to [2] not alter the reference scene sound pressure TF between the blocked auditory canal entrance and the eardrum with the auditory canal open. Since these TFs are conceptual TFs that do not exist in reality, they cannot be measured. This results from the fact that no sound pressure at the eardrum is present if the auditory canal is blocked. Here, a blocked auditory canal headphone selection criterion (HPSC) is proposed and formulated in such a way that it can be derived based on four TF measurements on an AH.

The paper is structured as follows: Starting from a description of the existing approach to address the suitability of HPs for binaural synthesis with blocked auditory canal recording, the shortcomings of this method are identified and discussed, motivating the formulation of a revised HP selection criterion. On that basis, the HPSC is introduced for human as well as artificial head situations. An exemplary AH evaluation for two HPs shows the general applicability of the procedure. In addition, the stability and repeatability of the HPSC and its implications on the overall binaural synthesis TFs are discussed. For that purpose, a binaural synthesis quality criterion is defined. A summary concludes the paper.

The HPSC is evaluated here for two exemplary, randomly selected HP specimens, which are consequently not representative for the corresponding HP models. A large scale study allowing for more general conclusions by including different HP models and several specimens per model is currently being prepared at our institute. HP manufacturers interested in having their circum- and supra-aural HP models included in the study are encouraged to contact the author.

2. PREVIOUS APPROACH

[1] models the auditory canal as acoustic transmission line, for a point source in the free-field described by the open-circuit pressure spectrum P_2 and the radiation impedance Z_{ra} at the entrance to the canal. Then, the pressure spectrum P_3 at the ear canal entrance in general is computed using Thévenin's theorem [3] by the spectral pressure division

$$\frac{P_3}{P_2} = \frac{Z_{\mathrm{ec}}}{Z_{\mathrm{ec}} + Z_{\mathrm{ra}}} \tag{1}$$

of the open-circuit pressure between the radiation impedance and the auditory canal impedance Z_{ec}. The transmission line model is assumed to be valid up to a frequency in the range of 10 kHz.

To determine the open circuit-pressure, [1] employs a miniature microphone built in a foam earplug and inserted in the auditory canal so that the canal is blocked and the microphone is positioned at the canal entrance. For the measurement of the pressure at the open entrance, [1] proposes probe microphone measurement despite the associated reduction of the signal to noise ratio (SNR) compared to miniature microphones, since the latter would disturb the sound field in the auditory canal.

Based on equation 1 and further assuming Z_{ra} being independent of the source position, [1] splits the TF from the point source pressure spectrum P_1 to the eardrum pressure spectrum P_4 in different partial TFs as given by

$$\frac{P_4}{P_1} = \frac{P_4}{P_3}\frac{P_3}{P_2}\frac{P_2}{P_1}. \tag{2}$$

Considering HP reproduction, [1] splits up the transmission from the HP input voltage spectrum U_{hp} to the eardrum pressure spectrum P_7 using the pressure spectrum P_6 at the entrance to the auditory canal and its open-circuit pendant P_5, resulting in

$$\frac{P_7}{U_{\mathrm{hp}}} = \frac{P_7}{P_6}\frac{P_6}{P_5}\frac{P_5}{U_{\mathrm{hp}}}. \tag{3}$$

The transmission from P_5 to P_6 is given based on the transmission line model using the radiation impedance Z_{hp} at the entrance to the auditory canal for HP playback by

$$\frac{P_6}{P_5} = \frac{Z_{\mathrm{ec}}}{Z_{\mathrm{ec}} + Z_{\mathrm{hp}}}. \tag{4}$$

Here, Z_{hp} includes the transfer characteristics of the volume enclosed by the HP and the HP's mechanical and electrical subsystems.

The transmission line model suggests identical transmission from the auditory canal entrance to the eardrum for free-field and HP listening, that is

$$\frac{P_7}{P_6} = \frac{P_4}{P_3}. \tag{5}$$

Equation 5 holds not true if the transmission line model fails.

The relation of the equations 1 and 4 describing the transfer between the sound pressure at the entrance to the auditory canal and its open-circuit pendant is given by

$$\frac{P_3/P_2}{P_6/P_5} = \frac{Z_{\mathrm{ec}} + Z_{\mathrm{hp}}}{Z_{\mathrm{ec}} + Z_{\mathrm{ra}}}. \tag{6}$$

This ratio is later [4] referred to as the *pressure division ratio (PDR)*. The PDR is approximately unity if $Z_{\mathrm{hp}} \approx Z_{\mathrm{ra}}$ or $Z_{\mathrm{ec}} \gg Z_{\mathrm{hp}}$ and $Z_{\mathrm{ec}} \gg Z_{\mathrm{ra}}$ hold. Then, equation 6 simplifies to

$$\frac{P_3}{P_2} \approx \frac{P_6}{P_5}, \tag{7}$$

and the corresponding HPs are by [1] referred to as *open headphones*. Later [4], the term *free-air equivalent coupling to the ear (FEC) headphones* is introduced to identify HPs that fulfill equation 7.

For binaural synthesis based on recording using a miniature microphone with the TF $M_1 = U_{\mathrm{M}}/P_{\mathrm{M}}$ at the entrance to the blocked auditory canal, [1] concludes that a correction filter with the TF

$$\begin{aligned} G_{\mathrm{C}} &= \frac{P_4/P_3}{P_7/P_6}\frac{P_3/P_2}{P_6/P_5}\frac{1}{M_1 P_5/U_{\mathrm{hp}}} \\ &= \frac{Z_{\mathrm{ec}}+Z_{\mathrm{hp}}}{Z_{\mathrm{ec}}+Z_{\mathrm{ra}}}\frac{1}{M_1 P_5/U_{\mathrm{hp}}} \end{aligned} \tag{8}$$

is required for the binaural synthesis to equal the reference scene: *“The equalizing filter should include extra terms, if recording is made outside a blocked ear canal* [...]*. The extra terms are not required, when an open headphone is used for reproduction.”* This way, [1] assumes that the PDR defined by equation 6 can be equalized for non FEC HPs. This assumption is not necessarily true, because Z_{hp} is defined to include the transfer characteristics of the volume enclosed by the HP and the HP’s mechanical and electrical subsystems. For that reason, possibly occurring resonance and modal effects in the HP-head system are also included. This may affect the equalization process, since equalization of modal structures or resonances is not possible completely correct [5, 6].

2.1. Measurement Procedure

To determine PDRs with the motivation of addressing HPs’ FEC characteristics, [4] recorded the impulse responses (IRs) corresponding to the TFs

$$H_5 = \frac{P_5'}{U_{\mathrm{ls}}} = \frac{P_5 M_2}{U_{\mathrm{ls}}} \quad \text{and} \quad H_6 = \frac{P_6'}{U_{\mathrm{ls}}} = \frac{P_6 M_2}{U_{\mathrm{ls}}} \tag{9}$$

from the loudspeaker (LS) input voltage spectra U_{ls} to the pressure spectra P_5' and P_6' at a probe microphone with the TF M_2, stating: *“Even though efforts were undertaken* [...]*, a small leak could arise between the headphone cushion and the head surface* [... and the] *presence of the probe tube could also cause minor changes in the position and orientation of the headphone capsule”* [4]. For that reason, the same probe microphone was used for both measurements, assuming *“this way the capsule displacement and the leak would have the same influence on P_5 and P_6, and the influence on the pressure division was eliminated.”* In addition, a carefully designed measurement sequence was used, consisting of first recording P_5 with a probe microphone directly in front of the blocked entrance to the auditory canal, then removing the earplug with *“as little disturbance of the probe microphone as possible,”* and recording P_6 at the open canal entrance immediately afterwards. Using the data acquired that way, [4] computed the PDRs

$$\frac{P_3/P_2}{H_5/H_6} = \frac{P_3/P_2}{P_5/P_6} \tag{10}$$

(cf. equation 6) using the free-field pressure divisions P_3/P_2 reported for the same subjects and equipment by [7]. However, [4] remark that *“small changes in microphone and headphone positions between measurements with open and blocked ear canals and between free-air and headphone measurements* [...] *make*[s] *PDRs unreliable above approximately 7* kHz *and thus they are not reported”*. Based on the PDRs, [4] address the question whether specific HPs are suitable for binaural synthesis with blocked auditory canal recording and headphone transfer function (HPTF) measurement. Suitability is assumed if the HPs show FEC characteristics.

2.2. Shortcomings of the Approach

Being based on a transmission line model, the procedure proposed by [1] is valid in principle only for wavelengths large compared to the auditory canal diameter. According to [1], this leads for typical parameters to an upper limiting frequency of approximately 10 kHz. Above this frequency, the assumption of equal sound pressure transfer from the auditory canal entrance to the eardrum in the HP and free-field situations is no longer valid.

The necessity of using probe microphones for the measurement of the PDR results in reduced SNR compared to miniature microphone measurements [4]. Further, the procedure requires the assumption of Z_{ra} being independent of the source position, which is not necessarily fulfilled exactly for all configurations, especially in the frequency range above some 5 kHz [8, 9].

However, the most significant shortcoming is of procedural nature: assuming the disturbance of the probe microphone during the measurements of P_5 and P_6 with intermediate HP repositioning and earplug removal to remain identical and to cancel out that way when computing the PDR is questionable. Comparable issues evolve regarding the probe microphone position. Both of these procedural shortcomings are especially likely to result in high frequency errors, while leakage effects may also occur at frequencies at the lower end of the HP transmission range.

Questioning the validity of the results especially at high frequencies is further supported by the decision of [7] not to report the resulting PDRs for frequencies above 7 kHz. This shortcoming appears severe taking into account the relevance of high frequency information for the elevation localization [10, 11, 12] and the sharpness perception [13, 14]. This becomes even more important considering sharpness to be of major influence on auditory pleasantness [15] and timbre [16]. Consequently, a selection criterion covering the whole audible frequency range is desirable and will be introduced in the following.

3. SELECTION CRITERION

The terminology in the remainder of this paper is based on [2], with lower case letters indicating signals s and systems h in the time domain, while upper case letters denote spectra S and TFs H. Signals or systems representing the left and right channels of a two-channel system are summed up by vectors and therefore set in bold fonts, in their time and frequency domain representations. Consequently, the division or multiplication of TFs is used as shorthand for element wise computation if not denoted otherwise, and the division of TFs is idealistically assumed to be valid [2]. Subscripts are employed to differentiate between signals and systems, while superscripts provide additional information on the specific signal or system. For example are systems depending on individual characteristics indicated by the upper index *ind*, AH related systems by *ah*, systems involving HPs by *h*, and systems including blocked auditory canals by *b*.

3.1. Human Subjects

Based on the system theoretical binaural synthesis framework introduced by [2], the requirements for HPs to be appropriate for binaural synthesis with blocked auditory canal recording (*rec*) and HPTFs (*hptf*) as well as human head playback (*play*) without further equalization can be derived from equation 80 (of [2]). Consequently, appropriate HPs are required to show frequency independent TFs

$$\mathbf{H}^{\mathrm{ind}}_{\mathrm{hpsc}}(\mathbf{x}_{\mathrm{m}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{play}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}}) = \\ = \frac{\mathbf{H}^{\mathrm{ind}}_{\mathbf{p}_{\mathrm{m}_{\mathrm{b}}},\mathbf{p}_{\mathrm{e}},\mathrm{ls}}(\mathbf{x}_{\mathrm{m}_{\mathrm{rec}}})}{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{p}_{\mathrm{m}_{\mathrm{b}}},\mathbf{p}_{\mathrm{e}},\mathrm{hp}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}})} \overset{!}{=} \mathbf{1}. \tag{11}$$

Equation 11 is proposed as the blocked auditory canal headphone selection criterion (HPSC, index *hpsc*) here [17]. The HPSC can be read descriptively in that the conceptual TFs

$$\mathbf{H}^{\mathrm{ind}}_{\mathbf{p}_{\mathrm{m}_{\mathrm{b}}},\mathbf{p}_{\mathrm{e}},\mathrm{ls}}(\mathbf{x}_{\mathrm{m}_{\mathrm{rec}}}) \tag{12}$$

for LS (index *ls*) and

$$\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{p}_{\mathrm{m}_{\mathrm{b}}},\mathbf{p}_{\mathrm{e}},\mathrm{hp}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}}) \tag{13}$$

for HP (index *hp*) playback, relating the sound pressure spectra $\mathbf{P}_{\mathrm{m}_{\mathrm{b}}}$ at the miniature microphones (index *m*) in the blocked auditory canals to the ear signal spectra $\mathbf{P}_{\mathrm{e}}$, have to be identical.

The conceptual TFs defined by the equations 12 and 13 are here referred to as the complex valued blocking factors. The blocking factors are not measurable directly because the sound pressures at the eardrums are not present if the auditory canals are blocked. However, an alternate formulation of the HPSC can be given using the equations 61 and 35 of [2], which formulate the blocking factors dependent on the TFs between the LS and HP input voltage spectra U_{ls} and $\mathbf{U}_{\mathrm{hp}}$ and the different sound pressure spectra under consideration. This way, the HPSC may be rewritten by

$$\mathbf{H}^{\mathrm{ind}}_{\mathrm{hpsc}}(\mathbf{x}_{\mathrm{m}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{play}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}}) = \\ = \frac{\mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}})}{\mathbf{H}^{\mathrm{ind,b}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{rec}}})} \cdot \\ \cdot \frac{\mathbf{H}^{\mathrm{ind,h,b}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}})}{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}})} \overset{!}{=} \mathbf{1}. \tag{14}$$

Equation 14 confirms that the complex valued blocking factors, which are the relations of the sound pressure spectra at the eardrums and at the entrances to the blocked auditory canals, have to be identical with LS and HP reproduction.

The LS and HP situations differ in the active sound sources and in the fact that the HPs are present in the HP playback and HPTF measurement situations, in contrast to the LS based recording situation and reference scene (index *ref*). Therefore, the differences in the blocking factors for LS versus HP reproduction may arise from the different sound sources or from the pure HP presence.

3.2. Human Subjects, Headphone Reference

The contributions of both the latter differences to the overall effect can be separated by additionally considering the binaural synthesis system with regard to a modified reference scene, the so-called headphone reference scene. In contrast to the standard binaural synthesis reference scene defined as a subject listening to a LS [2], the reference scene for binaural synthesis (BS) with HP reference is defined as a subject wearing non-operational headphones listening to a LS [18].

Using BS with HP reference, the HP and LS situations are identical apart from the active sound source. Consequently, the blocked auditory canal headphone selection criterion with headphone reference predicts the differences in the blocking factors between LS and HP playback due to the different sound sources. This is formulated using the index *hpsc,s* (the additional subscript *s* indicates consideration of the source effects) by

$$\begin{aligned}
&\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{hpsc,s}}(\mathbf{x}_{\mathrm{m}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{play}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}}) = \\
&= \frac{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{p}_{\mathrm{m}_{\mathrm{b}}},\mathbf{p}_{\mathrm{e}},\mathrm{ls}}(\mathbf{x}_{\mathrm{m}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{ref}}})}{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{p}_{\mathrm{m}_{\mathrm{b}}},\mathbf{p}_{\mathrm{e}},\mathrm{hp}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}})} \\
&= \frac{\mathbf{H}^{\mathrm{ind,h}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{ref}}})}{\mathbf{H}^{\mathrm{ind,b,h}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{rec}}})} \cdot \\
&\quad \cdot \frac{\mathbf{H}^{\mathrm{ind,h,b}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}})}{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}})}.
\end{aligned} \tag{15}$$

Here, identical HP positions $\mathbf{x}_{\mathrm{hp}_{\mathrm{ref}}} = \mathbf{x}_{\mathrm{hp}_{\mathrm{rec}}}$ in the headphone reference scene and the headphone recording situation are assumed. The influences of the HP presence (additional superscript *p*) on the HPSC may then be assessed mathematically by

$$\begin{aligned}
&\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{hpsc,p}}(\mathbf{x}_{\mathrm{m}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{play}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}}) = \\
&= \frac{\mathbf{H}^{\mathrm{ind}}_{\mathrm{hpsc}}(\mathbf{x}_{\mathrm{m}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{play}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}})}{\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{hpsc,s}}(\mathbf{x}_{\mathrm{m}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{play}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}})}.
\end{aligned} \tag{16}$$

3.3. Artificial Heads

Equation 14 further shows that evaluation of the HPSC is not possible completely thorough on a human head, since it involves the measurement of the ear signals, which is in general not possible completely correct because it would require to take into account exactly the sound pressure detected by the eardrum [19]. However, the HPSC may without loss of generality be evaluated using an AH. Since the effects under consideration are HP characteristics, they are necessarily independent of the specific evaluation head (cf. also section *Hardware Influences*).

If an AH is used that is designed in such a way that the microphones can be positioned *reproducibly* at the entrances to the blocked auditory canals and at the eardrum locations, equation 11 simplifies to

$$\begin{aligned}
&\mathbf{H}^{\mathrm{ah}}_{\mathrm{hpsc}}(\mathbf{x}_{\mathrm{m}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{play}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}}) = \\
&= \frac{\mathbf{H}^{\mathrm{ah}}_{\mathbf{p}_{\mathrm{m}_{\mathrm{b}}},\mathbf{p}_{\mathrm{ahm}_{\mathrm{e}}},\mathrm{ls}}(\mathbf{x}_{\mathrm{m}_{\mathrm{rec}}})}{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{p}_{\mathrm{m}_{\mathrm{b}}},\mathbf{p}_{\mathrm{ahm}_{\mathrm{e}}},\mathrm{hp}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}})} \\
&= \frac{\mathbf{H}^{\mathrm{ah}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}})}{\mathbf{H}^{\mathrm{ah,b}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{rec}}})} \\
&\quad \cdot \frac{\mathbf{H}^{\mathrm{ah,h,b}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}})}{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}})} \overset{!}{=} \mathbf{1}.
\end{aligned} \tag{17}$$

Based on the equations 23, 25, 37, and 39 given by [2], it is possible to rewrite equation 17 in a practically more applicable way, solely dependent on the HPTFs and the recording situation TFs by

$$\begin{aligned}
&\mathbf{H}^{\mathrm{ah}}_{\mathrm{hpsc}}(\mathbf{x}_{\mathrm{m}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{play}}}, \mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}}) = \\
&= \frac{\mathbf{H}^{\mathrm{ah}}_{\mathrm{rec}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}})}{\mathbf{H}^{\mathrm{ah,b}}_{\mathrm{rec}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{rec}}})} \cdot \\
&\quad \cdot \frac{\mathbf{H}^{\mathrm{ah,h,b}}_{\mathrm{hptf}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}}, \mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}})}{\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{hptf}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}})}.
\end{aligned} \tag{18}$$

This formulation of the HPSC requires four measurements on a specifically designed AH to evaluate the suitability of HPs for binaural synthesis with blocked auditory canal recording. The AH must allow for positioning of the microphone reproducibly at the eardrum position and at the entrance to the blocked auditory canal. For measurement with such an AH, the HPSC according to equation 18 is depicted in Figure 1 for the binaural synthesis of an exemplary LS reference scene in reverberant laboratory environment, implemented with an RME Fireface 400 audio interface and a Klein + Hummel Studio Monitor Loudspeaker O 98 for two different HP specimens (black Sennheiser HD 800, gray Stax λ pro NEW).

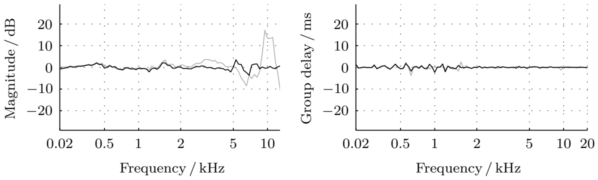

Fig. 1: Spectrally smoothed blocked auditory canal headphone selection criterion (HPSC) for artificial head recording, headphone transfer function measurement, and evaluation using the same head. Sennheiser HD 800 (black) and Stax λ pro NEW headphones (gray).

Figure 1 reveals frequency dependence of the HPSC for both HP specimens under consideration. While the criterion for the HPs indicated by the black line shows a magnitude spectrum frequency independent within ±3 dB, the criterion magnitude spectrum for the model represented by the gray curve exhibits spectral peaks up to 20 dB at frequencies above about 6 kHz. Since the frequency dependencies of the magnitude spectrum in the frequency range below about 3 kHz are comparable for both HPs, they are most likely caused by the prototypical measurement setup itself. The HPSC group delay stays frequency independent on average, while showing narrow peaks and dips, especially in the mid and low frequency range. However, based on the criterion introduced here, the HP specimen represented by the black curve appears based on the prototypical measurement setup discussed here to be more appropriate for application in binaural synthesis with blocked auditory canal recording.

3.4. Artificial Heads, Headphone Reference

Comparable to equation 17, the blocked auditory canal headphone selection criterion with headphone reference given for human subjects by equation 15 for binaural synthesis with headphone reference scene is adapted to the AH case by

$$\begin{aligned}\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{hpsc,s}}(\mathbf{x}_{\mathrm{m_{rec}}},\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{m_{hptf}}}) &= \\ = \frac{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{p}_{\mathrm{m_b}},\mathbf{p}_{\mathrm{ahm_e}},\mathrm{ls}}(\mathbf{x}_{\mathrm{m_{rec}}},\mathbf{x}_{\mathrm{hp_{rec}}},\mathbf{x}_{\mathrm{hp_{ref}}})}{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{p}_{\mathrm{m_b}},\mathbf{p}_{\mathrm{ahm_e}},\mathrm{hp}}(\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{m_{hptf}}})}.\end{aligned} \tag{19}$$

Based on the voltage to sound pressure transfer functions, equation 19 can be simplified to

$$\begin{aligned}\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{hpsc,s}}(\mathbf{x}_{\mathrm{m_{rec}}},\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{m_{hptf}}}) &= \\ = \frac{\mathbf{H}^{\mathrm{ah,h}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{hp_{ref}}})}{\mathbf{H}^{\mathrm{ah,b,h}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{m_{rec}}},\mathbf{x}_{\mathrm{hp_{rec}}})} &\cdot \\ \cdot \frac{\mathbf{H}^{\mathrm{ah,h,b}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{m_{hptf}}})}{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp_{play}}})}.\end{aligned} \tag{20}$$

Assuming identical HP positions $\mathbf{x}_{\mathrm{hp_{ref}}} = \mathbf{x}_{\mathrm{hp_{rec}}}$ in the headphone reference scene and the recording situation, equation 19 may further be given in a simplified manner, solely dependent on the recording situation and headphone transfer functions by

$$\begin{aligned}\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{hpsc,s}}(\mathbf{x}_{\mathrm{m_{rec}}},\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{m_{hptf}}}) &= \\ = \frac{\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{rec_{ahm}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{hp_{ref}}})}{\mathbf{H}^{\mathrm{ah,b,h}}_{\mathrm{rec_{m}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{m_{rec}}},\mathbf{x}_{\mathrm{hp_{rec}}})} &\cdot \\ \cdot \frac{\mathbf{H}^{\mathrm{ah,h,b}}_{\mathrm{hptf_{m}}}(\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{m_{hptf}}})}{\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{hptf_{ahm}}}(\mathbf{x}_{\mathrm{hp_{play}}})}.\end{aligned} \tag{21}$$

Figure 2 shows the HPSC according to equation 21 for the binaural synthesis of an exemplary headphone reference scene implemented using an RME Fireface 400 audio interface and a Klein + Hummel Studio Monitor Loudspeaker O 98 with two different HP specimens (black Sennheiser HD 800, gray Stax λ pro NEW). The data of Figure 2 are qualitatively and quantitatively comparable to those given by Figure 1. Consequently, for the HP specimens under consideration, the source effect is assumed to be the major influence factor on the artifacts visible for the specimen indicated by the gray curve. Again, this specimen is assumed to be less suited for the application in binaural synthesis with blocked

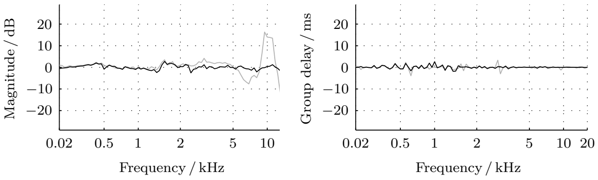

Fig. 2: Spectrally smoothed blocked auditory canal headphone selection criterion with headphone reference (blocked auditory canal headphone selection criterion with headphone reference) for artificial head recording, headphone transfer functions, and evaluation. Sennheiser HD 800 (black), Stax λ pro NEW (gray).

auditory canal recording. The validity of this prognosis is verified by means of loudness comparisons in [18] and [20].

The influences of the HP presence on the HPSC are for the AH situation given analogously to the human head situation (equation 16) by

$$\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{hpsc,p}}(\mathbf{x}_{\mathrm{m_{rec}}}, \mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{m_{hptf}}}) = \\ = \frac{\mathbf{H}^{\mathrm{ah}}_{\mathrm{hpsc}}(\mathbf{x}_{\mathrm{m_{rec}}}, \mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{m_{hptf}}})}{\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{hpsc,s}}(\mathbf{x}_{\mathrm{m_{rec}}}, \mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{m_{hptf}}})}. \tag{22}$$

4. STABILITY AND REPEATABILITY

To address the repeatability of the HPSC evaluation, seven sets of measurements according to equation 18 were carried out using a custom made AH, an RME Fireface 400, a Klein + Hummel Studio Monitor Loudspeaker O 200, and Stax λ pro NEW HPs. Since the measurements include different LS positions, the effects of the sound incidence direction can be assessed based on the data in addition. Figure 3 shows the HPSC for each single measurement (gray) and the arithmetic mean (black) of magnitude spectrum and group delay.

The data depicted in Figure 3 reveal the existence of the prominent spectral characteristics in all magnitude spectra, while the individual group delay measurements show spectral peaks and dips, which appear to be less pronounced on average, but still lie in the same frequency regions.

Considering the magnitude spectrum, it appears valid to use a single measurement to predict the average characteristics. It may further be concluded, that possible influences of the sound incidence direction on the HPSC are within the accuracy of the prototypical measurement system employed here.

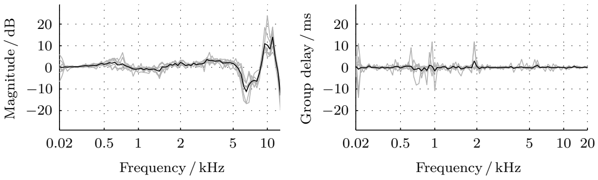

Fig. 3: Spectrally smoothed blocked auditory canal headphone selection criterion (HPSC) for artificial head recording, headphone transfer functions, and evaluation for Stax λ pro NEW headphones. Results for eight different sound incidence directions in the same laboratory room (gray) and average (black).

5. HARDWARE INFLUENCES

Figure 4 shows the HPSC according to equation 18 measured using an RME Fireface 400 and a Klein + Hummel Studio Monitor Loudspeaker O 200 in a different room and with a different LS compared to Figure 1.

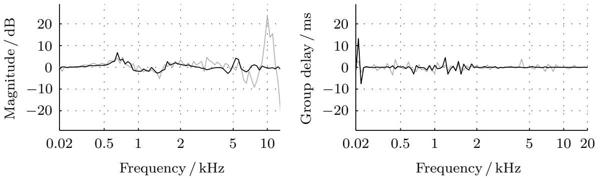

Fig. 4: Spectrally smoothed blocked auditory canal headphone selection criterion (HPSC) for artificial head recording, headphone transfer function measurement, and evaluation using the same head. Sennheiser HD 800 (black) and Stax λ pro NEW headphones (gray). Loudspeaker and measurement room different from the situation shown in Figure 1.

The similarity between the Figures 1 and 4 confirms that the global characteristics of the magnitude spectrum remain constant in the different reverberant laboratories and for the LSs considered here. It is important to note that the HPSC predictions are valid only in the frequency range where the LS employed is able to provide energy to the measurement room. Influences of the specific AH on the HPSC are for circum-aural HPs based on the headphone transfer function variability data given by [21] not expected. However, a study addressing the influence of the AH employed on the HPSC is currently carried out at our laboratory.

6. RELATION TO THE OVERALL BINAURAL SYNTHESIS TRANSFER FUNCTION

According to [2], the audio signal to be presented is convolved in the actual binaural synthesis situation with the IRs recorded according to section 3 of [2] and appropriate equalization filters as defined in section 6 of [2]. The convolution products are presented to a listener by HPs (section 4). The ear signal spectra of the equalized binaural synthesis situation are then computed by

$$\begin{aligned}&\mathbf{P}^{\mathrm{ind,h}}_{\mathrm{e_{bs}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{mic_{rec}}},\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{mic_{hptf}}})=\\&\quad=\mathbf{H}^{\mathrm{ind}}_{\mathrm{eq}}(\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{mic_{rec}}},\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{mic_{hptf}}})\cdot S_{\mathrm{ls}}\cdot\\&\qquad\cdot\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{play}}(\mathbf{x}_{\mathrm{hp_{play}}})\cdot\mathbf{H}^{\mathrm{ind}}_{\mathrm{rec_{mic}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{mic_{rec}}}),\end{aligned}\tag{23}$$

resulting in the equalized binaural synthesis TFs

$$\begin{aligned}&\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{bs}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{mic_{rec}}},\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{mic_{hptf}}})=\\&=\frac{\mathbf{P}^{\mathrm{ind,h}}_{\mathrm{e_{bs}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{mic_{rec}}},\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{mic_{hptf}}})}{S_{\mathrm{ls}}}\\&=\mathbf{H}^{\mathrm{ind}}_{\mathrm{eq}}(\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{mic_{rec}}},\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{mic_{hptf}}})\cdot\\&\qquad\cdot\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{play}}(\mathbf{x}_{\mathrm{hp_{play}}})\cdot\mathbf{H}^{\mathrm{ind}}_{\mathrm{rec_{mic}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{mic_{rec}}}).\end{aligned}\tag{24}$$

For system development and evaluation purposes, AH playback is desirable. While not exactly representing human head playback, AH evaluation allows to assess all system theoretical aspects of binaural synthesis without loss of generality. When transferring the results to human listeners, physical differences as for example the soft human tissue in contrast to the typically harder AH surface or the missing hair have to be considered [22, 23, 24].

Relating the spectra of the digital sequences corresponding to the AH microphone output signals in the evaluation situation to the LS driving spectrum results in the TFs

$$\begin{aligned}&\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{bs_{ver}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{mic_{rec}}},\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{mic_{hptf}}})=\\&=\frac{\mathbf{S}^{\mathrm{ah,h}}_{\mathrm{e_{bs}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{mic_{rec}}},\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{mic_{hptf}}})}{S_{\mathrm{ls}}},\end{aligned}\tag{25}$$

describing the equalized artificial head binaural synthesis situation.

To achieve a quantitative accuracy measure for the ear signals generated, an analytic criterion for the quality of binaural synthesis systems, referred to as binaural synthesis quality criterion (BSQC) is derived based on AH validation measurements in the following. A binaural synthesis system generates the reference scene ear signals if the binaural synthesis TFs according to equation 24 equal the reference scene TFs defined by equation 4 of [2]. For AH recording, equalization, and validation, this can be formulated mathematically using the equations 15 and 16 of [2] in combination with equation 25 given here by

$$\begin{aligned}&\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{bs_{ver}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{mic_{rec}}},\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{mic_{hptf}}})\overset{!}{=}\\&\qquad\overset{!}{=}\mathbf{H}^{\mathrm{ah}}_{\mathrm{ref_{ahm}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}})\cdot\mathbf{H}_{\mathrm{i_{ah}}}\cdot\mathbf{H}_{\mathrm{ahm}}\\&\qquad=\mathbf{H}^{\mathrm{ah}}_{\mathrm{rec_{ahm}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}}).\end{aligned}\tag{26}$$

In other words: the binaural synthesis and the recording situation TFs acquired with the same AH must be frequency independently identical for the binaural synthesis outputs to equal the reference scene ear signals. This requirement is defined as the BSQC (index *bsqc*) by

$$\begin{aligned}&\mathbf{H}^{\mathrm{ah}}_{\mathrm{bsqc}}(\mathbf{x}_{\mathrm{mic_{rec}}},\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{mic_{hptf}}})\overset{!}{=}\mathbf{1}\\&=\frac{\mathbf{H}^{\mathrm{ah}}_{\mathrm{rec_{ahm}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}})}{\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{bs_{ver}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{mic_{rec}}},\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{mic_{hptf}}})}.\end{aligned}\tag{27}$$

A comparison of the HPSC defined by equation 17 to the overall error of a binaural synthesis system with blocked auditory canal recording according to the corresponding BSQC as defined by equation 27 may be used to reveal error contributions not covered by the HPSC. These remaining error contributions are consequently defined mathematically by

$$\begin{aligned}&\mathbf{H}^{\mathrm{ah}}_{\mathrm{rem}}(\mathbf{x}_{\mathrm{m_{rec}}},\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{m_{hptf}}})=\\&\quad\frac{\mathbf{H}^{\mathrm{ah}}_{\mathrm{hpsc}}(\mathbf{x}_{\mathrm{m_{rec}}},\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{m_{hptf}}})}{\mathbf{H}^{\mathrm{ah}}_{\mathrm{bsqc}}(\mathbf{x}_{\mathrm{m_{rec}}},\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{m_{hptf}}})}.\end{aligned}\tag{28}$$

The evaluation of the remaining error according to equation 28 is shown by Figure 5 for the exemplary situation considered here. The deviations indicated by the HPSC selection criterion are confirmed to cover the error in binaural synthesis with blocked auditory canal recording and blocked auditory canal HPTF based equalization for both HPs considered.

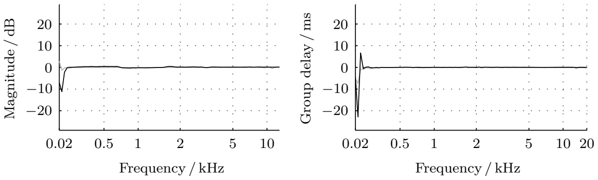

Fig. 5: Spectrally smoothed relation of the blocked auditory canal headphone selection criterion to the BSQC for binaural synthesis with blocked auditory canal recording and blocked auditory canal headphone transfer function based equalization, implemented and evaluated using an artificial head. Sennheiser HD 800 (black) and Stax λ pro NEW headphones (gray).

The spurious artifacts visible in Figure 5 are likely to be caused by implementation constraints regarding the equalization filter design [2, 21].

7. CONCLUSIONS AND OUTLOOK

In this paper, a blocked auditory canal headphone selection criterion (HPSC) is proposed and evaluated. The criterion can be used to predict, based on the four artificial head transfer function measurements given by equation 17, whether a specific pair of headphones is suitable for binaural synthesis with blocked auditory canal recording and headphone transfer function measurement.

Speaking descriptively, the HPSC checks whether the transfer functions connecting the sound pressure spectra at the miniature microphones in the blocked auditory canals and the sound pressure spectra at the eardrums in the open auditory canals (the complex valued blocking factors) are identical for loudspeaker and headphone playback. In case the HPSC predicts deviations for a specific headphone specimen, the applicability of the specimen can be decided in view of the intended application based on the magnitude spectrum and group delay errors predicted by the HPSC.

If an artificial head is employed that allows for repeatably positioning the microphones at the eardrum positions and at the entrances to the blocked auditory canals, the procedure provides valid results in the whole audible frequency range. Furthermore, a simple relation to the overall binaural synthesis transfer functions defined as the binaural synthesis quality criterion is given inherently.

The method proposed for the same purpose by [1] allows for valid measurements only in the frequency range below 7 kHz, primarily due to being founded on the transmission line theory and procedural reasons [4]. Further, it requires using probe microphones, which provide systematically less signal to noise ratio than the miniature or artificial head microphones used here.

Based on the theory presented in this paper, a large scale study is currently being set up at our laboratory, addressing the HPSC compliance of headphones. To be representative for the respective headphone models, a sample of at least five specimens per model will be included in the test. Headphone manufacturers interested in having one or more circum- or supra-aural headphone models included in the study are welcome to contact the author.

8. ACKNOWLEDGMENTS

The author is indebted to Prof. Hugo Fastl for providing valuable insight on the subject and the possibility for this work, supported by grant FA 140/4 of Deutsche Forschungsgemeinschaft (DFG).

9. REFERENCES

[1] Henrik Møller. Fundamentals of binaural technology. *Applied Acoustics*, 36:171–218, 1992.

[2] Florian Völk. System theory of binaural synthesis. In *131st AES Convention*, Convention Paper 8568, 2011.

[3] Don H. Johnson. Equivalent circuit concept: the voltage-source equivalent. *Proceedings of the IEEE*, 91(4):636–640, 2003.

[4] Henrik Møller, Dorte Hammershøi, Clemen Boje Jensen, and Michael Friis Sørensen. Transfer characteristics of headphones measured on human ears. *J. Audio Eng. Soc.*, 43(4):203–217, April 1995.

[5] Allen R. Groth. High-fidelity sound system equalization by analysis of standing waves. *J. Audio Eng. Soc.*, 22(10):795–799, 1974.

[6] Trevor J. Cox and Peter D'Antonio. Room optimizer: A computer program to optimize the placement of listener, loudspeakers, acoustical surface treatment and room dimensions in critical listening rooms. In *103rd AES Convention*, Preprint 4555, 1997.

[7] Henrik Møller, Michale Friis Sørensen, Dorte Hammershøi, and Clemen Boje Jensen. Head-related transfer functions of human subjects. *J. Audio Eng. Soc.*, 43(5):300–321, May 1995.

[8] Dorte Hammershøi and Henrik Møller. Sound transmission to and within the human ear canal. *J. Acoust. Soc. Am.*, 100(1):408–427, July 1996.

[9] D. Hammershøi and H. Møller. Sound transmission from the free field to the eardrum. *Acustica – acta acustica*, 82(Suppl. 1):S 87, 1996.

[10] Futoshi Asano, Yoiti Suzuki, and Toshio Sone. Role of spectral cues in median plane localization. *J. Acoust. Soc. Am.*, 88(1):159–168, July 1990.

[11] John C. Middlebrooks and David M. Green. Sound localization by human listeners. *Annu. Rev. Psychol.*, 42:135–159, 1991.

[12] Frederic L. Wightman and Doris J. Kistler. Monaural sound localization revisited. *J. Acoust. Soc. Am.*, 101(2):1050–1063, February 1997.

[13] G. v. Bismarck. Timbre of steady sounds: Scaling of sharpness. In *Proceedings of 7th International Congress on Acoustics, ICA 1971*, volume 3, pages 353–356, 1971.

[14] Hugo Fastl and Eberhard Zwicker. *Psychoacoustics – Facts and Models*. Springer, Berlin Heidelberg, 3rd edition, 2007.

[15] W. Aures. Der Wohlklang: Eine Funktion von Schärfe, Rauhigkeit, Klanghaftgkeit und Lautheit (Sensory pleasantness: a function of sharpness, roughness, tonality, and loudness). In *Fortschritte der Akustik, DAGA '84*, pages 735–738. DPG, Bad Honnef, 1984.

[16] G. v. Bismarck. Sharpness as an attribute of the timbre of steady sounds. *Acustica*, 30:159–172, 1974.

[17] Florian Völk. Messtechnische Verifizierung eines datenbasierten binauralen Synthesesystems (Metrological verification of a data based binaural synthesis system). In *Fortschritte der Akustik, DAGA 2010*, pages 1049–1050. Dt. Gesell. für Akustik e.V., Berlin, 2010.

[18] Florian Völk and Hugo Fastl. Locating the missing 6 db by loudness calibration of binaural synthesis. In *131st AES Convention*, Convention Paper 8488, 2011.

[19] H. Hudde and S. Schmidt. Sound fields in generally shaped curved ear canals. *J. Acoust. Soc. Am.*, 125(5):3146–3157, May 2009.

[20] Florian Völk, Andreas Dunstmair, Tobias Riesenweber, and Hugo Fastl. Bedingungen für gleichlaute Schalle aus Kopfhörern und Lautsprechern (Requirements for equally loud sounds from headphones and loudspeakers). In *Fortschritte der Akustik, DAGA 2011*, pages 825–826. Dt. Gesell. für Akustik e.V., Berlin, 2011.

[21] Florian Völk. Inter- and intra-individual variability in blocked auditory canal transfer functions of three circum-aural headphones. In *131st AES Convention*, Convention Paper 8465, 2011.

[22] Brian F. G. Katz. Acoustic absorption measurement of human hair and skin within the audible frequency range. *J. Acoust. Soc. Am.*, 108(5 Pt. 1):2238–2242, November 2000.

[23] Bradley E. Treeby, Jie Pan, and Roshun M. Paurobally. The effect of hair on auditory localization cues. *J. Acoust. Soc. Am.*, 122(6):3586–3597, December 2007.

[24] Bradley E. Treeby, Jie Pan, and Roshun M. Paurobally. An experimental study of the acoustic impedance characteristics of human hair. *J. Acoust. Soc. Am.*, 122(4):2107–2117, October 2007.