Published in final edited form as:

J Acoust Soc Am. 2014 March ; 135(3): 1278–1293. doi:10.1121/1.4863196.

Modeling the direction-continuous time-of-arrival in head-related transfer functions

Harald Ziegelwanger^{a)} and Piotr Majdak

Acoustics Research Institute, Austrian Academy of Sciences, Wohllebengasse 12-14, A-1040 Vienna, Austria

Abstract

Head-related transfer functions (HRTFs) describe the filtering of the incoming sound by the torso, head, and pinna. As a consequence of the propagation path from the source to the ear, each HRTF contains a direction-dependent, broadband time-of-arrival (TOA). TOAs are usually estimated independently for each direction from HRTFs, a method prone to artifacts and limited by the spatial sampling. In this study, a continuous-direction TOA model combined with an outlier-removal algorithm is proposed. The model is based on a simplified geometric representation of the listener, and his/her arbitrary position within the HRTF measurement. The outlier-removal procedure uses the extreme studentized deviation test to remove implausible TOAs. The model was evaluated for numerically calculated HRTFs of sphere, torso, and pinna under various conditions. The accuracy of estimated parameters was within the resolution given by the sampling rate. Applied to acoustically measured HRTFs of 172 listeners, the estimated parameters were consistent with realistic listener geometry. The outlier removal further improved the goodness-of-fit, particularly for some problematic fits. The comparison with a simpler model that fixed the listener position to the center of the measurement geometry showed a clear advantage of listener position as an additional free model parameter.

I. INTRODUCTION

Head-related transfer functions (HRTFs) are direction-dependent filters describing the acoustic filtering of the head, torso, and pinna (Blauert, 1997; Mehrgardt and Mellert, 1977; Møller *et al.*, 1995). They provide cues for the localization and externalization of virtual sound sources presented via headphones, a procedure known as binaural reproduction technique (Wightman and Kistler, 1989). With a complete set of room-specific HRTFs, arbitrary combinations of virtual sound sources can be presented via so called virtual auditory displays (Xie, 2013).

Binaural pairs of HRTFs provide interaural time differences (ITDs) which are important for the lateralization of sound-sources in the horizontal plane (Strutt, 1894; Macpherson and Middlebrooks, 2002). While the ITD is a relative quantity and thus sufficient for the description of an acoustic scene with a single sound source, the inter-*source* time difference,

^{a)}Author to whom correspondence should be addressed. harald.ziegelwanger@oeaw.ac.at.

PACS number(s): 43.60.Uv, 43.66.Yw, 43.60.Jn, 43.66.Pn [ZHM]

i.e., the delay between individual sound sources is required for an acoustically correct representation of multiple sound sources in a scene. This inter-source time difference can only be derived when the absolute delay arising from the acoustic wave propagation path from each source to a listener's ear is available. Those delays can be described by the directional time-of-arrival (TOA), i.e., a delay in an HRTF as a function of the sound direction. Note that for a given direction, the difference between the left-ear TOA and the right-ear TOA corresponds to the broadband¹ ITD.

The TOA can be derived from the spectral phase of an HRTF by separating the spectral phase in three components: Minimum phase, frequency-independent TOA, and excess phase. While the minimum phase can be calculated with the Hilbert transform (Oppenheim *et al.*, 1999), a robust estimation of the TOA from an impulse response is not trivial (Defrance *et al.*, 2008). In the research field of binaural hearing, most of the TOA estimation algorithms consider single directions only, mainly focusing on *ITD* estimation. Because of the close relation between the TOA and the ITD, some of the ITD estimation methods also implicitly estimate the TOA for both ears. These methods can be classified in frequency-domain (Huopaniemi and Smith, 1999; Jot *et al.*, 1995), time-domain (Jot *et al.*, 1995; Møller *et al.*, 1995), and other methods (Nam *et al.*, 2008; Wightman and Kistler, 2005). All these methods are susceptible to measurement noise, and even though low-pass or high-pass filtering have been used to improve the estimation stability (Algazi *et al.*, 2001a), the methods produce artifacts, which correction is not trivial (Busson *et al.*, 2005). The artifacts can be found by following the TOA along the direction but can not be considered by an estimator processing each sound direction separately. The detection and correction is thus, usually done *post hoc* in a manual procedure.

In other research fields, sophisticated methods like particle filtering considering spatial evolution of TOA (Michalopoulou and Jain, 2012) have been developed to overcome these problems. Applied to the field of HRTFs, in such a *joint-directional* timing model, the timing information (TOA or ITD) of the sound propagation paths around the human head is modeled as a continuous function of direction. A simple example for such a model is the ITD model for the horizontal plane (Kuhn, 1977), which describes ITD by two model parameters: The radius of the circle and the orientation of the interaural axis. This model assumes that ITD can be correctly described by a function of the lateral angle in the horizontal plane and has been extended to a two-dimensional model (Savioja *et al.*, 1999). In order to estimate the parameters of such a model for a specific listener, the timing estimations from HRTFs measured for different directions are used to fit a continuous function of direction. In such a model, potentially incorrect timing estimations do not *define* the timing for the corresponding direction, but they, together with the correct estimations, *contribute* to a model as a continuous function of direction. Wang *et al.* (2009) proposed to fit a polynomial function with truncated order to model ITD as a two-dimensional function of direction. Even though this model could be a starting point for a TOA model, the parameters of the polynomial function have a very weak relation to the geometry of the

¹While the ITD is a frequency-dependent quantity with larger values at low frequencies (Kuhn, 1977), it can be sufficiently approximated by a frequency-independent quantity as indicated for a clothed manikin (Kuhn, 1977) or human listeners (Kulkarni *et al.*, 1999). Thus, in this study, both TOA and ITD are considered being frequency independent.

measurement setup and physical interpretation of such model parameters seems to be rather difficult. Thus, in this study, we propose and investigate the use of a continuous-direction model of the TOA with a close relation to the geometry of the HRTF measurement setup.

II. GENERAL METHODS

A. Coordinate systems

Two coordinate systems are used for the directional representation of HRTF data. First, we use the geodesic coordinate system (Fig. 1, left panel), in which the direction of a sound source is described by the azimuth φ and elevation θ angle. A horizontal plane is selected by an elevation angle. The horizontal plane selected by $\theta = 0^\circ$ is called the interaural horizontal plane. Second, we use the interaural-polar coordinate system, which is the geodesic coordinate system rotated such that the polar axis becomes equivalent with the interaural axis (Fig. 1, right panel). In the interaural-polar coordinate system, a sagittal plane is selected by the lateral angle Φ and the direction of a source within a sagittal plane is described by the polar angle Θ . This coordinate system more distinctly addresses the human sound-localization mechanisms because the lateral direction is mostly associated with binaural cues and polar direction is mostly associated with spectral cues (Morimoto and Aokata, 1984).

B. HRTF material

In order to evaluate the proposed models, parameters resulting from the models were compared to the parameters actually used in the HRTF measurements. In order to provide well-defined actual parameters, we used numerically calculated HRTFs, for which the simulated measurement setup can be perfectly constructed. Further, the model was applied to acoustically measured HRTFs of human listeners. The exact parameters actually used in HRTF measurements, such as position, rotation, and size of the listener's head, can not be perfectly reconstructed and thus are not used for the evaluation.

1. Numerical calculation—HRTFs of three different objects were numerically calculated. The first object, Sphere, was a rigid sphere with two virtual microphones arbitrarily positioned on the surface. The second object, sphere-and-torso (SAT), was the object Sphere combined with a simplified torso placed below the sphere. The third object was sphere-torso-pinna (STP), which was the SAT with a stitched pinna on the sphere, see Fig. 2. Even though the contribution of the pinna was expected to be rather small, the STP is considered as a relevant link to the actual morphology of a human listener.

The sphere and the torso were constructed based on icosphere objects in Blender.² The torso was an ellipsoid with axes of 200 mm $\times$ 520 mm $\times$ 180 mm (width $\times$ depth $\times$ height) with its center positioned 190 mm below the center of the sphere. The torso and sphere were linked with a neck, which was modeled as a cylinder with a radius of 50 mm. The pinna was created in MakeHuman³ with the application's default parameters. All objects were meshed to an average edge length of 3 mm. Top and side view on the STP object is shown in Fig. 2.

²Blender 2.64, available from <http://www.blender.org> (date last viewed July 8, 2013).

³Available from <http://www.makehuman.org> (date last viewed July 8, 2013).

For all objects, the sound pressure in the simulated free-field was calculated with the Burton-Miller collocation boundary-element method coupled with the multi-level fast-multipole method (for more details see Kreuzer *et al.*, 2009). Complex spectra were calculated in the range of 0.2 to 20 kHz with a frequency resolution of 0.2 kHz. The principle of reciprocity was used to decrease the computational costs in the calculation. HRTFs were calculated for 1550 directions at a 3-m distance. The directions were in the elevation between -30° to 80° in steps of 5° . The azimuthal range was 360° in steps of 2.5° within the interaural horizontal plane and larger steps within other horizontal planes in order to obtain an approximately constant directional resolution. Finally, the head-related impulse responses (HRIRs) were obtained by applying the inverse Fourier transform on the complex spectra and resampling to 48 kHz. The final duration of the HRIRs was 5 ms.

The actual parameters, for which the HRTFs were calculated are described in the corresponding sections. In total, HRTFs of 98 objects were calculated.⁴

2. Acoustic measurements—Acoustically measured HRTFs of in total 172 human listeners from three HRTF databases were considered.

The database ARI⁵ contains HRTFs measured in a semi-anechoic sound chamber with the blocked-ear-canal method (Møller *et al.*, 1995) for 1550 directions at a distance of 1.2 m. The horizontal and vertical range was 360° and -30° to 80° , respectively, with the horizontal and vertical resolution of 2.5° and 5° , respectively (for more details see Majdak *et al.*, 2010). In our study, the HRTFs of 77 listeners from NH2 to NH132 were used (not all IDs are populated between NH2 and NH132). In text, this database is referred to as DB1.

The database CIPIC⁶ contains HRTFs measured in an acoustic facility with the blocked-ear-canal method for 45 listeners and 1250 directions at a distance of 1 m. The horizontal and vertical range of directions was 360° and -80° to 80° , respectively, with the horizontal and vertical resolution of 5.625° and 5° , respectively (for more details see Algazi *et al.*, 2001b). In text, this database is referred to as DB2.

The database LISTEN⁷ contains 51 HRTF sets measured in an anechoic sound chamber with the blocked-ear-canal method for 187 directions at a distance of 1.95 m. The horizontal and vertical range was 360° and -45° to 90° , respectively, with the horizontal and vertical resolution of 15° and 15° , respectively. HRTFs of subject #34 have been excluded from our study because of some directions containing invalid HRIRs in terms of the total energy of 30 dB below that for the other directions. In text, this database is referred to as DB3.

⁴Available from <http://www.kfs.oeaw.ac.at/hrtf/database/amt/ziegelwanger2013.zip> (date last viewed August 4, 2012).

⁵Retrieved from <http://sofacooustics.org/data/database/ari/> (date last viewed January 12, 2012), originating from the “(short)” HRTF set from <http://www.kfs.oeaw.ac.at/hrtf> (date last viewed January 12, 2012).

⁶Retrieved from <http://sofacooustics.org/data/database/cipic/> (date last viewed January 11, 2012), originating from <http://interface.cipic.ucdavis.edu/sound/hrtf.html> (date last viewed January 11, 2012).

⁷Retrieved from <http://sofacooustics.org/data/database/listen/> (date last viewed January 2, 2012), originating from the “compensated” HRTF set from <http://recherche.ircam.fr/equipes/salles/listen/download.html> (date last viewed January 11, 2012) with the sampling rate set to 48 kHz.

III. TOA ESTIMATORS

In this section, we review the relevant TOA estimators and investigate their results when applied to HRTFs.

A. General

The estimation of TOA relies on the separation of the minimum phase, frequency-independent TOA, and excess phase in an HRTF. While most of the existing methods focus on the estimation of ITD, some of them implicitly estimate the TOA. The methods can be classified as time-domain, frequency-domain, and mixed methods.

Frequency-domain methods rely on the phase spectrum. In order to estimate the TOA, a line is fit to the phase spectrum and the slope of the line is defined as the TOA (Huopaniemi and Smith, 1999; Jot *et al.*, 1995). The TOA has also been derived from the group delay, i.e., the frequency derivative of the phase, by averaging the group delay over the frequency range of 1 to 5 kHz (Jot *et al.*, 1995), or the frequency range of 0.5 to 2 kHz (Huopaniemi and Smith, 1999). These frequency-domain methods suffer from the phase wrap around 2π and the corresponding difficulties in the unwrap algorithm when a change of $-\pi$ or π cannot be distinguished (Tribolet, 1977). The unwrap errors directly affect the TOA estimation. Even though the group delay can be directly calculated from the complex spectrum, it still might yield noisy results (Raykar *et al.*, 2005).

Time-domain methods do not suffer from the phase wrap. The simplest time-domain method for the TOA estimation is the amplitude threshold, in which TOA is the time where a (pre-processed) HRIR reaches a threshold for the first time. In particular, the following combinations of preprocessing and thresholds have been used: The maximum (or 20 dB below the maximum) of the absolute-valued HRIR (Møller *et al.*, 1995), its envelope (Busson *et al.*, 2005), its low-passed (Algazi *et al.*, 2001a), or its high-passed versions. In a similar time-domain method, the centroid of the HRIR has been used to calculate the ITD and therefore might be interpreted as TOA (Jeppesen and Møller, 2005; Jot *et al.*, 1995). The correspondence of those methods to the frequency-domain methods is not simple. As a combination of time- and frequency method, zero-crossing method has been used to estimate the group delay in the time domain (Lee *et al.*, 2009), but it yielded noisy results for realistic signals.

Based on the generalized cross-correlation method (Knapp and Carter, 1976), an interesting TOA estimator has been proposed (Wightman and Kistler, 2005; Nam *et al.*, 2008), where the HRIR h and its minimum-phase version $h_{\min}$ were cross-correlated and the resulting maximum was evaluated,

$$\hat{\tau} = \underset{k}{\operatorname{argmax}} \left| \sum_{n=-\infty}^{\infty} h[n] h_{\min}[n+k] \right|, \quad (1)$$

with $\hat{\tau}$ as the estimated TOA and n as the sample index corresponding to the discrete time. Note that $h_{\min}$ can be easily calculated from the causal part of an HRTF's complex cepstrum (Oppenheim *et al.*, 1999).

This method has some interesting properties. First, it might be less vulnerable to phase artifacts in spectral notches because the amplitude spectrum of the cross-correlation is the squared amplitude spectrum of the HRTF, and thus, frequency regions with little energy do not contribute much to the TOA estimation. Second, if an HRTF is a “minimum-phase plus delay” system, which seems to be approximately valid for at least some directions of the incoming sound (Kulkarni *et al.*, 1999; Mehrgardt and Mellert, 1977) then the excess phase is approximately zero. Consequently, the cross-correlation in Eq. (1) reduces to a delayed pulse in the time domain, whose delay can be easily estimated and corresponds *exactly* to the TOA. Thus, this method works well for directions for which HRTFs are approximately “minimum-phase plus delay” systems. Interestingly, the transfer function of the pressure at the surface of a sphere is minimum phase (Constan and Hartmann, 2003, p. 1007 and Fig. 6). However, HRTFs of a human can not be minimum-phase systems because of the reflections at the torso and the pinna. Pinna contributes in higher frequencies only and torso contributes in lower frequencies, with stronger contributions from the ipsilateral side (Blauert, 1997). While the impact of these contributions on ITDs is frequency dependent and rather small (Kuhn, 1977, Figs. 11, 12, and 13), their impact on the excess phase in HRTFs seems to be not clarified yet.

Thus, in our study, we do not rely on the assumption that HRTFs are minimum-phase systems. We rather apply different TOA estimators on HRTFs and investigate their compatibility with our model.

B. Application

In order to estimate the impact of the TOA estimators, results from various TOA estimators were investigated when applied on different HRTFs, with a particular focus on the impact of torso, pinna, and the acoustic measurement noise. To this end, numerically calculated HRTFs represented the noise-free condition in which HRTFs of SAT and STP were used to demonstrate the impact of torso and pinna in comparison to HRTFs of a sphere. Acoustically measured HRTFs of an exemplary listener (NH89, DB1) were used to investigate the usefulness of TOA estimators under more realistic conditions.

The following TOA estimators were considered: (1) TOA at the maximum of the absolute HRIR (MAX; Møller *et al.*, 1995); (2) TOA at the centroid of the HRIR (CTD; Jeppesen and Møller, 2005); (3) TOA as the average group delay of the HRTF, calculated from the negative frequency derivative of the phase, and averaged between 1 and 5 kHz (AGD; Jot *et al.*, 1995); and (4) TOA at the maximum of the cross-correlation between the HRIR and its minimum phase version (MCM; Nam *et al.*, 2008) according to Eq. (1).

Figure 3 shows estimated TOAs for the interaural horizontal plane with one TOA estimator for each panel. The torso does not seem to have affected the estimated TOAs much. For example, for the MCM estimator, the difference between the TOAs of SAT and that of the Sphere was $0.4 \mu\text{s}$ (average over all directions). Compared to torso, the pinna seems to have an effect on the estimated TOA. For the MCM estimator, the average difference between the TOAs of SAT and that of STP was $-12.0 \mu\text{s}$. It seems like the combination of the pinna and torso affected the TOA estimation substantially, but the effect of the torso alone was negligible.

TOAs obtained for the acoustically measured HRTFs seem to follow a different pattern with artifacts for particular directions. These findings support others' observations that TOA estimation from acoustically measured HRTFs is a difficult task. Especially the artifacts might yield larger errors in modeling and might require a special handling.

Even though the comparison between the different TOA estimators was only used to illustrate their results, the differences are evident. Especially, the MAX estimator seems to produce many artifacts. Also, the CTD estimator seems to result in systematically larger TOAs than the MCM estimator. Nevertheless, at this stage, we are not able to reliably judge the compatibility of the TOA estimators with the proposed model, and thus, in the following section, we consider all of them in the evaluation of the model.

IV. ON-AXIS TOA MODEL

In this section, the directional TOA model is derived. The model aims at describing simplified geometric properties of the listener and the HRTF measurement setup. The model is fit to TOAs estimated by one of the previously described TOA estimators, and the results are evaluated.

A. Model derivation

The listener's head is modeled as a rigid sphere that is placed exactly in the center of the measurement system. Thus, the sphere center and the measurement center are assumed to be coincident. The model further assumes a stationary position of listener and the loudspeakers. Figure 4 illustrates the model with the arbitrary position $\vec{e} = [\phi_e, \theta_e]$ of the ear on the sphere. The sound propagation path between a sound source and the listener's ear is modeled corresponding to the sound propagation path from a loudspeaker to the in-ear-canal microphone in the HRTF measurement. A plane-wave propagation is assumed, thus, the curvature of the incident sound waves generated by the sound sources are assumed to be small in relation to the entire sound propagation paths. TOA is modeled as the time the sound requires to travel along the propagation path for a given direction.

For a given direction $\vec{h} = [\phi_h, \theta_h]$ of a sound source, the TOA is split in a direction-independent and direction-dependent part. The direction-independent TOA τ is the propagation path between a loudspeaker and the closest point on the sphere. For the direction-dependent TOA and its corresponding sound propagation path, the cases for the ipsilateral and the contralateral sound sources are considered separately. The distinction is based on the angular distance α , which is

$$\alpha = \arccos(\vec{e} \cdot \vec{h}). \quad (2)$$

For the ipsilateral sound sources $\alpha < \pi/2$, the direction-dependent propagation path s_1 is

$$s_1 = r [1 - \sin \theta_e \sin \theta_h - \cos \theta_e \cos \theta_h \cos(\phi_e - \phi_h)] \quad \text{for } \alpha \leq \frac{\pi}{2}. \quad (3)$$

For the contralateral sound sources $\alpha > \pi/2$, the direction-dependent propagation path additionally considers the scattering path around the sphere and is

$$s_1 = r + r \arccos [\sin \theta_e \sin \theta_h + \cos \theta_e \cos \theta_h \cos (\phi_e - \phi_h)] - \frac{r\pi}{2} \text{ for } \alpha > \frac{\pi}{2}.$$

With the speed of sound c , the TOA is then

$$\tilde{\tau} = \frac{s_1}{c} + \bar{\tau}. \quad (5)$$

In order to obtain the model parameters $\bar{\tau}$, r , φ_e , and θ_e , Eq. (5) is fit to the set of estimated TOAs $\hat{\tau}$. The fit is done by minimizing the squared error between $\tilde{\tau}$ and $\hat{\tau}$ using an iterative trust-region-reflective algorithm implemented as *lsqcurvefit*.⁸ This algorithm is a subspace trust-region method and is based on the interior-reflective Newton method (Coleman and Li, 1996). It implements a robust and efficient globalization strategy for computing a local minimum of a possibly non-convex and once continuously differentiable function. In each iteration, the algorithm approximately solves a linear system using the method of preconditioned conjugate gradients, until the termination tolerance is reached. In our study, the termination tolerance was set to 10^{-6} . For a general review of nonlinear least-squares methods, see Dennis (1977).

The algorithm converges at a local minimum, which distance to the global minimum depends on the initial values. Thus, providing reasonable initial values is important. For example when fitting the model to HRTFs of human, the initial values should correspond to average human parameters. Further, trivial values like $r = 0$ must be avoided. In this study, the initial values were chosen in each evaluation and/or application separately depending on the origin of the data.

Finally, the algorithm requires constrains. For the on-axis model, the constrains were always the same, namely, ear position constrained to $\varphi_e = \pm 180^\circ$ and $\theta_e = \pm 90^\circ$ (effectively unconstrained) and the radius constrained to ± 6 cm of the radius initial value.

Note that the model is fit to estimated TOAs for each ear separately.

B. Evaluation

In order to evaluate the goodness of the model fit, we used the adjusted norm of residuals (ANR), i.e., the square root of the adjusted sum of squares of residuals (Bajorski, 2011). The adjustment happens by dividing the sum by the number of residuals, which makes the norm less dependent on the number of HRTF directions. ANR varies from 0 to infinity—the smaller the norm the better the fit. It provides an idea about the average (in terms of root-mean-square) size of the residuals per fit.

⁸MATLAB R2011a, available from <http://www.mathworks.de/products/matlab/> (date last viewed July 8, 2013).

1. Effect of parameter variation—The on-axis TOA model was evaluated addressing the following: (1) general ability of the model to describe TOA; (2) ability to describe TOAs estimated from HRTFs measured under conditions represented by spheres of different sizes and different microphone positions; (3) effect of initial values on the goodness of fit. To this end, HRTFs were calculated for the object Sphere under conditions resulting from the following parameter combinations: r was 77.5, 87.5, and 97.5 mm, φ_e was -40° to $+40^\circ$ in steps of 10° , θ_e was -10° to $+10^\circ$ in steps of 5° . Each combination of the actual parameters is referred to as a condition. The position of the left and right ear microphones was varied independently, however, only the most relevant combinations of parameters were tested. Note that $r = 87.5$ mm addresses the average head radius of human population (Hartley and Fry, 1921). Four TOA estimators (MAX, CTD, AGD, and MCM) were used to estimate TOAs from each HRTF set. For the fits, the initial values were $r = 87.5$ mm, $\varphi_e = \pm 90^\circ$, and $\theta_e = 0^\circ$. Initial values were identical in all conditions.

Table I shows the ANRs and the parameter errors, i.e., differences between the actual and estimated parameters, averaged over all conditions for each of the TOA estimators. Recall that the smaller ANR, the smaller the difference between the modeled and estimated TOAs, and thus, the better the model represents the estimated TOAs. Also, the smaller the parameter error, the more accurate the model estimated the actual parameter of the geometry.

For the MAX, CTD, and MCM estimators, the ANRs were well below the sampling interval of $20.8 \mu\text{s}$. The sphere radius errors were within the spatial resolution of the propagation paths (7.1 mm at $c = 340$ m/s and sampling rate of 48 kHz). The ear position errors were within the range of 1.2° . For the MCM estimator, Fig. 5 shows the estimated model parameters for each condition.

For the AGD estimator, the ANR was three times larger than for the other estimators. Also the parameter errors were larger than for the other estimators. The estimated radii were on average 15.5 mm larger than the actual radii used in the HRTF simulations.

The fits seem to be robust to the initial values: Identical initial values were used in all conditions, and still the estimated model parameters converged near the actual values.

2. Effect of sphere, torso, and pinna—The proposed model approximates the listener's head as a sphere. When applied to HRTFs of human listeners, torso and pinna might affect the model results. In order to estimate the impact of torso and pinna, the on-axis model was fit to TOAs estimated for calculated HRTFs of the objects Sphere, SAT, and STP. For the HRTF calculations, the sphere radius was 87.5 mm and the virtual microphones were placed at typical ear positions, i.e., $\varphi_e = \pm 85^\circ$ and $\theta_e = -10^\circ$. Four TOA estimators (MAX, CTD, AGD, and MCM) were used to estimate TOAs, then the on-axis model was fit to the estimated TOA sets. The initial values were $r = 90$ mm, $\varphi_e = \pm 90^\circ$, and $\theta_e = 0^\circ$, i.e., slightly different than the actual parameters of the HRTF calculations.

Table II shows ANRs and model parameters for the four TOA estimators. Generally, the results seem to depend on both TOA estimator and evaluated object. For MAX, CTD, and

MCM estimators, the ANRs were below $10.4 \mu\text{s}$, which is only a fraction of the sampling interval of $20.8 \mu\text{s}$. The estimated ear positions had errors of 3° or smaller, the estimated radii had errors of 7.7 mm or smaller. For the AGD estimator, the ANR increased up to $54.1 \mu\text{s}$, the estimated ear position and radius had errors as large as 4.3° and 17.5 mm , respectively. Such large errors indicate that the AGD estimator is not compatible with the proposed model.

For the remaining TOA estimators, the ANRs were fractions of the sampling interval, even when the torso and pinna were considered in HRTF simulations. This indicates that the proposed model was generally able to describe the TOAs even in the presence of torso and pinna. The estimated radii increased, however, when the torso and even more when the pinna were considered in HRTF simulations. For MAX, CTD, and MCM estimators, the radius error increased from 3.9 , 4.5 , and 2.7 mm (Sphere), respectively, to 7.9 , 7.9 , and 5.7 mm (STP), respectively. It seems like the pinna had the most effect on the estimated radius—from Sphere to SAT, the increase in the radius error averaged over the three estimators was 0.6 mm , from SAT to STP, it was 2.9 mm . This indicates that pinna, not the torso, had the most prominent effect on the estimated radii. This effect might be related to the longer propagation paths around the pinna, an aspect not considered in the proposed model.

The combination of the on-axis model with the MCM estimator showed the most accurate parameter estimations, even in the presence of pinna and torso. For STP, the radius error was 5.7 mm which is within the spatial resolution of the propagation paths of 7.1 mm . The estimated ear position had an error of 1.8° . We thus, consider the MCM estimator as an estimator best compatible with the proposed model and focus on this estimator in the following sections.

C. Application

The model is intended to be applied on acoustically measured HRTFs of human listeners. Thus, in this section, we analyze the results of fitting the model to TOAs obtained from HRTFs of human listeners (described in Sec. II B 2). For each listener and ear, TOAs were estimated using the MCM estimator and the model was fit. The initial values were $r = 87.5 \text{ mm}$, $\varphi_e = \pm 90^\circ$, and $\theta_e = 0^\circ$.

Figure 6 shows the estimated model parameters as histograms over all listeners. Each histogram was fit to a Gaussian function in order to estimate the average and the standard deviation of the distributions. The statistics of the model parameters and ANRs are shown in Table III. The ANRs are positive values and can be approximated by the normal distribution only when their distribution has a small variance. Thus, for ANRs, medians were calculated. Further, the interaural radius differences (IRDs), i.e., the listener-specific differences between the radii estimated for the left and the right ears were calculated. The average IRD was $12.5 \pm 22.3 \text{ mm}$ (average $\pm$ standard deviation), with individual IRDs ranging from -57.2 to 80.8 mm , see Fig. 7 (top panel). For the three databases DB1, DB2, and DB3, the average IRD was 11.0 , 4.8 , and 21.6 mm , respectively. The *unsigned* IRD was $19.8 \pm 16.2 \text{ mm}$. For the three databases, the average of unsigned IRDs was 21.9 , 7.2 , and 27.7 mm , respectively.

D. Discussion

The on-axis model describes an HRTF measurement with a listener placed in the center of the coordinate system. The evaluation performed on the calculated HRTFs showed a good correspondence between the actual parameters used for HRTF simulations and the parameters estimated by the on-axis model. The model relies on a TOA estimator, and the model using the MCM estimator showed the best correspondence between the actual and estimated parameters. For the MCM estimator, the ANRs were in the range of few microseconds, which is a small fraction of the sampling interval, indicating a good fit of the model to the data. The differences between the estimated and the actual ear positions were smaller than 1.6° , which is smaller than the minimum audible angle being in the range of 2° (Mills, 1958). The differences between the estimated and the actual radius were smaller than 5.6 mm, which is within the resolution of the propagation paths given by the sampling interval and the speed of sound.

The model was then applied to acoustically measured HRTFs of 172 listeners. Note that for acoustically measured HRTFs, an evaluation in terms of comparing the estimated and actual parameters can not take place. The estimated parameters can be, however, checked for plausibility. For example, the ear position or the sphere radius should correspond to those usually found in humans. Further, the model independently fit to TOAs obtained for the two ears of the same listener should yield in similar estimated radii. In other words, an IRD substantially deviating from zero should raise evidence for implausible results.

Interestingly, the check of plausibility showed mixed results. The binaural average of the estimated radius was with 86.95 mm close to the average head radius of adult humans of 87.5 mm (Hartley and Fry, 1921). Also the estimated average ear positions, nearly most lateral (φ_e of $\pm 88^\circ$) and a little below the head center (θ_e of -7°) seem to be close to those usually found in humans (Algazi *et al.*, 2001b; Burkhard and Sachs, 1975). There are, however, two serious concerns about the results.

First, the ANRs (binaural average of $23.6 \mu\text{s}$) were much larger than those obtained for calculated HRTFs (binaural average for the MCM estimator of $6.1 \mu\text{s}$). The larger ANRs might result from artifacts in the TOA estimations, which underlying reason might be noise and inaccuracies involved in the acoustic measurements. In order to reduce the larger ANRs, in the next section, we propose to handle these artifacts before fitting the model.

Second, the estimated radii substantially differed between the two ears of each listener. For one listener, the IRD was even as large as 80.8 mm. Such large discrepancy raises evidence that the on-axis model might be a too simple representation of HRTF measurement geometry. This issue seems to be independent from the larger ANRs, as supported by the rather low correlation coefficients for the correlation between the IRD and the ANR of 0.22 and -0.05 for the left and right ears, respectively. Thus, in the two next sections, we propose enhancements of the on-axis model aiming at reducing the large IRDs and ANRs in particular when applied to acoustically measured HRTFs of human listeners.

V. DIRECTIONAL OUTLIER REMOVAL

In this section, we address the large ANRs which resulted when the on-axis model was applied to acoustically measured HRTFs. The working hypothesis was that a few artifacts caused the large ANRs and that removing the corresponding directions before fitting the model will result in smaller ANRs.

Generally, artifacts can be defined as outliers in the sense of estimations being numerically distant from the rest of the estimations (Grubbs, 1969). Figure 3 shows the estimated TOAs for an exemplary listener and from that visual inspection, some of the data points might be manually classified as outliers.

There are many methods for detection and removal of outliers from data (Barnett and Lewis, 1994). Most unsupervised methods rely on the assumption of normally distributed data and test individual data points for the probability of deviating from the normality (NIST/SEMATECH, 2002). Grubbs' test is such a test, it is, however, only recommended when testing for a single outlier. When the number of outliers is not exactly known, the generalization of Grubbs' test known as the extreme studentized deviate (ESD) test is recommended (Rosner, 1983). Since we can not specify the exact number of the outliers, the ESD test seems to be the appropriate test for our outlier detection.

The ESD test requires many data points in the pool, thus, TOAs from all available directions of an HRTF set are pooled. Unfortunately, TOAs are correlated with Φ , and thus, if collapsed over Φ and Θ , the TOAs were not normally distributed. This situation is illustrated in Fig. 8, in which the top panels show the TOAs as a function of lateral angle Φ , i.e., collapsed over the polar angles Θ , for HRTFs of STP (left column) and of NH89 from DB1 (right column). The TOAs were obtained from the MCM estimator and black points represent outliers found later by the ESD test (see later). While in that representation some data can be visually identified as outliers, most of them would become undetectable when further collapsing over Φ . Basically, by further collapsing over Φ , the variance of the pool would be large and the outliers would remain hidden among the majority of the other data.

In order to better fulfill the assumption of normality and to obtain a data pool with a smaller variance, the *residuals* resulting from the on-axis model fit are considered. The bottom panels of Fig. 8 show the residuals corresponding to the TOAs shown in the top panels. Note that in that representation, the outliers will remain far from the majority of the data points even when further collapsing over Φ . Thus, the residuals show a better approximation of a normal distribution with a smaller variance, which makes them a more suitable basis for the unsupervised outlier detection with the ESD test.

In the ESD test, the hypothesis of no outliers in the data set is tested at a given significance level up to an upper bound of outliers (compare Sec. 1.3.5.17.3 in NIST/SEMATECH, 2002). The significance level is usually chosen as 0.05 and was also used in our study. The upper bound can be chosen as a compromise of finding only a few outliers when tested for clean data and but still removing most of the outliers when tested for noisy data. In our study, TOAs obtained for calculated HRTFs of STP represented clean data and TOAs obtained for acoustically measured HRTFs represented noisy data. This comparison is

illustrated in Fig. 8. The black points represent outliers resulting from the ESD tested with an upper bound of outlier rate of 1%, the gray points represent the remaining data. From this visual inspection, it seems like the ESD test was able to detect outliers that also in a supervised procedure would have been removed.

In order to estimate an appropriate upper bound of outliers, the outlier rate and ANRs were calculated for the objects Sphere (all conditions from Sec. IV B 1), STP and SAT (from Sec. IV B 2), and human listeners (from Sec. IV C). The MCM estimator and the on-axis model were used as described in Sec. IV C. The on-axis model was fit to TOAs and ANRs were calculated. Then, ESD tests were performed with upper bounds of the outlier rate of 0.01, 0.1, 1, and 10%. The outliers were removed and the on-axis models were fit again, this time based on the outlier-adjusted set of TOAs. Outlier rates and the final ANRs were calculated.

The statistics of the outlier rates and ANRs are summarized in Table IV. For the object Sphere, no outliers were detected in all conditions, and thus, these conditions are not shown. ANRs and outlier rates are positive values and can be approximated by the normal distribution only when their distribution has a small variance. Thus, in addition to the average and standard deviation, the median, and maximum values are shown.

The outlier rate increased with increasing upper bound. For calculated HRTFs (STP and SAT), the outlier removal did not affect the ANRs at all, and it did not affect the outlier rates up to the upper bound of 1%. For measured HRTFs, the outlier removal had an effect on the ANR and outlier rate even at the smallest upper bound tested, but an increase of the upper bound from 1% to 10% did not substantially change the ANR.

Thus, the upper bound of outlier rate of 1% seems to be a good compromise. It did not have much effect on the calculated HRTFs (clean TOAs) but a significant effect on the measured HRTFs (noisy TOAs). For the latter, the outlier removal substantially improved poor fits (from largest ANR of $407 \mu\text{s}$ to largest ANR of $89.5 \mu\text{s}$), resulting in reduced variability across the listeners (from standard deviation of 46.5 to that of $13.5 \mu\text{s}$). The outlier rate was at a maximum of 6.4%, with a median over all listeners of 0.16%, i.e., statistically below the upper bound of 1%.

VI. OFF-AXIS TOA MODEL

The on-axis model, evaluated for calculated HRTFs did not raise concerns even for HRTFs of a complex object like the STP. However, when applied to acoustically measured HRTFs of human listeners, the estimated IRDs were larger than expected. In this section, we extend the on-axis model in order to address those large IRDs. We analyze the problem, build a hypothesis for the reason of the large IRDs, and test it. Then, a model referred to as off-axis model is proposed. It is supposed to result in smaller IRDs even when applied to acoustically measured HRTFs of human listeners. The model is evaluated and finally applied to acoustically measured HRTFs.

A. Problem analysis

Figure 7 (top panel) shows the histograms of the IRD resulting from the on-axis model fit to HRTFs of human listeners. This histogram represents the statistics of IRDs provided in Sec. IV C. Note particularly the large variability indicating that for some listeners, the IRD was in the range of a few centimeters. For a comparison, the IRDs obtained for the calculated HRTFs of the Sphere (all conditions from Sec. IV B 1), SAT, and STP were much smaller, with an average of 0.0 ± 0.1 mm.

Seeking for an explanation of this observation, we compared the TOAs estimated by the MCM estimator and the TOAs from the on-axis model of an exemplary listener (NH89, DB1). The corresponding estimated model parameters are provided in Table III. Figure 9 shows the corresponding TOAs as functions of the azimuth angle in the interaural horizontal plane. For the left ear, the difference between the smallest and the largest TOA was $784 \mu\text{s}$. This difference would correspond to a maximum ITD of $784 \mu\text{s}$, if the HRTFs measured for the left ear were mirrored to the right side along the median plane. Note that this procedure is sometimes used in order to save time during the measurements (e.g., Gardner and Martin, 1995). For the *right* ear, the difference between the smallest and the largest TOA was $532 \mu\text{s}$, and thus, the maximum ITD would be $532 \mu\text{s}$. In the light of this finding, it is not surprising that our model converges at interaurally different sphere radii (107.3 mm and 68.5 mm for the left and the right ear, respectively) yielding an interaural asymmetry. Interestingly, the fits shown in Fig. 9 do not raise evidence for a fail of the model for any of the two ears. The asymmetry appears to be an intrinsic property of the measured TOAs and not an artifact of the modeling itself.

Such an asymmetry might result from a laterally non-centered position of the listener's head in the HRTF measurement. Interestingly, the descriptions of the three HRTF databases claim that the listener was placed at the center of the measurement, i.e., on-axis. In order to investigate how a lateral offset of an object affects the asymmetry of the TOA, the on-axis model was applied to HRTFs calculated for the objects Sphere with radius of 77.5, 87.5, and 97.5 mm, each of them placed 0, 10, and 20 mm to the left of the center. The other parameters were $\varphi_e = \pm 90^\circ$, and $\theta_e = 0^\circ$. The MCM estimator was used to obtain the TOAs. For the fits, the initial values were $r = 90$ mm, $\varphi_e = \pm 80^\circ$, and $\theta_e = 10^\circ$. Figure 10 shows the actual and estimated parameters. While the resulting ANR was below $7.3 \mu\text{s}$ in all conditions, the IRDs increased with increasing off-axis position of the object.

This finding suggests that the large IRDs obtained from the on-axis model when applied to acoustically measured HRTFs originate from a potential listener's lateral offset in the measurement. The on-axis model does not consider such an offset and thus results in a large IRD. Thus, in the following, we derive an off-axis model, which allows arbitrary positions of the listener's head. Note that a procedure jointly fitting the on-axis model to TOAs for both ears and forcing the IRD to zero would be an option as well. In such a procedure, however, the ANR would increase as a compromise to fit an average radius to TOAs for both ears. A model allowing arbitrary positions might have the potential to preserve the ANR and still reduce the IRD.

B. Model derivation

We consider the spherical head model to be arbitrarily positioned in the measurement system (Fig. 11). In order to consider the arbitrary position, $\vec{M} = [x_M y_M z_M]$ represents the offset of the sphere center relative to the measurement center. Note that while the head model might be arbitrarily positioned, its position still must be stationary, i.e., must not vary over the time. The direction-independent part of the sound propagation path ends on the surface of a centered sphere with radius $r + |\vec{M}|$ (see Fig. 11, bottom panel). The direction-dependent part of the sound propagation path is represented by the sum of three paths s_1 , s_2 , and s_3 , with s_1 from Eqs. (3) and (4). In order to derive s_2 , the angular distance β between the HRTF direction $\vec{h}$ and the position vector $\vec{M}$ is calculated as follows:

$$\beta = \arccos [-\cos \theta_h (x_M \cos \phi_h + y_M \sin \phi_h) - z_M \sin \phi_h]. \quad (6)$$

Then, the sound-propagation path s_2 is given by

$$s_2 = -r + |\vec{M}| \cos \beta + \sqrt{r^2 + 2|\vec{M}|^2 \cos^2 \beta + r|\vec{M}|}. \quad (7)$$

Further, the sound-propagation path s_3 is

$$s_3 = \frac{|\vec{M}| \sin \beta}{2 \left(|\vec{M}| + r \right) \tan \frac{\gamma}{2}}, \quad (8)$$

with

$$\gamma = \pi - \beta - \arccos \left(1 - \frac{2r s_2 + s_2^2}{2|\vec{M}|^2 + 2r|\vec{M}|} \right). \quad (9)$$

Finally, the modeled TOA $\tilde{\tau}$ is

$$\tilde{\tau} = \frac{(s_1 + s_2 + s_3)}{c} + \bar{\tau}. \quad (10)$$

In order to obtain the model parameters $\vec{M}$, $\bar{\tau}$, r , ϕ_e , and θ_e , Eq. (10) is fit to the set of estimated TOAs $\hat{\tau}$ using the same algorithm as that used for the on-axis model.

The off-axis model has many degrees of freedom and there are many local minima at which the fit might converge. In order to account for the increase in degrees of freedom, the termination tolerance was reduced to 10^{-8} . The initial values, however, require a special handling. If the initial values used for the on-axis model ($r = 87.5$ mm and $\vec{M} = 0$) were directly used for the off-axis model, the fit might converge at a local minimum at $\vec{M} = 0$ and a radius corresponding to that obtained from the on-axis model, which, again then, would yield an IRD comparable to that obtained from the on-axis model. In order to help the algorithm to converge at a global minimum, we used the estimated parameters from the on-

axis model as the initial values for the off-axis model. This was done for φ_e , θ_e , and $\bar{\tau}$, but not for $\vec{M}$ (in the on-axis model, $\vec{M}$ was fixed to zero per definition) and not for r (resulted from the on-axis model in a large IRD). In order to still provide reasonable initial values for r and $\vec{M}$, the initial value for r was the binaural average of the radii obtained from the on-axis model for both ears. The initial value for $\vec{M}$ was set to $\vec{M} = [1 \quad 0.5 \cdot \text{IRD} \quad 1] \text{ mm}$, because $\vec{M}$ resulting from the off-axis model was supposed to show a lateral offset toward the direction of the ear with the larger radius estimated by the on-axis model.

Even with those initial values, the fitting algorithm might still converge at $\vec{M} = 0$ and at r of the on-axis model. In order to prevent that, constraints were used with r at $\pm 0.25 \text{ IRD}$, $|\vec{M}|$ of $\pm 10 \text{ cm}$, φ_e of $\pm 45^\circ$, θ_e of $\pm 45^\circ$, and $\bar{\tau}$ of $\pm 1 \text{ ms}$. Note that with that strategy, the initial values for the off-axis model depend on the results from the on-axis model, for which the initial values must be provided.

C. Evaluation

1. General ability—The general ability of the off-axis model to describe the estimated TOAs was evaluated. To this end, the evaluation from Sec. IV B 2 was repeated for the off-axis model. The MCM estimator was applied on HRTFs of the objects Sphere, SAT, and STP. Then, for each object, the on-axis model was fit (initial values: $r = 90 \text{ mm}$, $\varphi_e = \pm 80^\circ$, and $\theta_e = 10^\circ$) and subsequently, the off-axis model was fit (initial values based on the results from the on-axis model). The full sets of TOAs were used, i.e., outliers were not removed.

Table V shows ANRs and estimated parameters. The results are similar to those from the on-axis model, demonstrating the general ability of the off-axis model to describe the estimated TOAs.

2. Effect of parameter variation—The off-axis model was evaluated for TOAs estimated from HRTFs obtained under different conditions, represented by spheres of different sizes and different microphone positions. This evaluation corresponds to that from Sec. IV B 1 for the MCM estimator. Correspondingly, the evaluation was performed under similar conditions, with the modification that after fitting the on-axis model, the off-axis model was fit to the full sets of TOAs.

The resulting errors, i.e., differences between the actual and estimated parameters are shown in rows labeled as Centered in the Table VI. The off-axis model described the different conditions with the radius errors of approximately 1.6 mm, the off-axis position errors smaller than 0.4 mm, and the ear position errors of approximately 1.2° . The ANRs ($6.2 \pm 0.1 \mu\text{s}$) were similar to those for the on-axis model ($6.5 \pm 0.1 \mu\text{s}$).

3. Susceptibility to the lateral offset—The susceptibility of the off-axis model to HRTFs of non-centered object is an important evaluation. The on-axis model resulted in large IRDs when applied to such HRTFs and the off-axis model was supposed to result in much smaller IRDs. To this end, the off-axis model was applied to the HRTFs calculated for the object Sphere placed at a combination of following offsets: x_M of 0, -10 , and -20 mm ;

y_M of 0, 10, and 20 mm; and z_M of 0 and -10 mm (compare Sec. VI A). The other parameters were $r = 77.5, 87.5$, and 97.5 mm, $\varphi_e = \pm 90^\circ$, and $\theta_e = 0^\circ$. The MCM estimator was used to obtain the TOAs and initial values from Sec. IV D were used. The off-axis models were fit to the full sets of TOAs.

The estimated parameters are shown in Fig. 12. The parameter errors are shown in rows labeled as Non-centered in the Table VI. The IRD was 0.0 ± 0.1 mm on average over all conditions tested.

4. Effect of outlier removal—In this evaluation, the effect of the outlier removal on the off-axis model was evaluated on the calculated HRTFs of SAT, STP, and centered and non-centered object Sphere. While these conditions represent clean data, in this evaluation, the potential negative impact of the outlier removal on the off-axis model is investigated. The evaluation corresponds to those from the previous three sections with a single modification: With the results from the on-axis model, the outlier detection was performed with the upper bound of outlier rate of 1% and then the off-axis models were fit to the outlier-adjusted sets of TOAs.

For SAT and STP, the outlier rate was 0.13% and 0.06%, respectively, and the estimated parameters and ANRs are shown in Table V (rows labeled as O-A). An effect of the outlier removal cannot be observed.

For the object Sphere, the parameter errors and ANRs are shown in Table VI. For the centered objects (variation of r and φ_e ; rows labeled as Centered), the outlier rate was 0%, and thus, the effect of the outlier removal is not given. For the non-centered objects (variation of $\vec{M}$, rows labeled as Non-centered), the average outlier rate was 0.7%. Compared to the full TOA sets, the estimated parameter error change was in the range of 1 mm and 1° , and the ANRs decreased by $0.1 \mu\text{s}$.

In summary, for calculated HRTFs, the outlier removal did not raise any evidence for a significant negative effect on the results of the off-axis model.

D. Application

The off-axis model was applied to HRTF sets of human listeners. This application corresponds to that from Sec. IV C, with the modification of using the off-axis model. The MCM estimator was used to obtain TOAs from acoustically measured HRTFs of human listeners described in Sec. II B 2. Then, for each listener, the on-axis model was fit with the initial values of $r = 87.5$ mm, $\varphi_e = \pm 90^\circ$, and $\theta_e = 0^\circ$. Further, off-axis models were fit to both full and outlier-adjusted TOA sets.

The rows labeled as All in Table V show the estimated parameters for the full (rows labeled as Full) and outlier-adjusted (rows labeled as O-A) TOA sets. The outlier removal reduced the variability in the goodness-of-fit, as indicated by the decrease of the ANR standard deviation from $46.3 \mu\text{s}$ (Full) to $12.5 \mu\text{s}$ (O-A). Also the estimated parameters show a slightly reduced variability when the outlier removal was applied.

Figure 7(bottom panel) shows the histogram of the IRDs obtained with the off-axis model fit to the outlier-adjusted TOA sets. The average IRD was 0.05 ± 1.54 mm and -0.14 ± 0.97 mm, with individual IRDs ranging from -7.1 to 9.9 mm and from -4.2 to 3.8 mm for the full and outlier-adjusted TOA sets, respectively.

E. Discussion

The off-axis model, applied to HRTF calculated for centered objects like SAT, STP, and Sphere, reconstructed the actual geometry parameters with an error up to a small fraction of the sampling interval (of $20.8 \mu\text{s}$ or 7.1 mm). Also for non-centered objects, the model seems to be able to yield correct parameters, as shown in Fig. 12 for various HRTF sets of the non-centered object Sphere. With an average IRD of 0.0 ± 0.1 mm, the model did not fail for the non-centered objects. Recall that this issue was the main problem of the on-axis model. The average ANR was in the range of a few microseconds, indicating a good fit of the model. The outlier removal did not show any negative effects on the results for calculated HRTFs.

The off-axis model and outlier removal were applied to acoustically measured HRTFs of human listeners. The average and largest outlier rates were 0.7% and 6.4% , respectively. Note that at that point, the estimated parameters can not be compared to the actual parameters of the measured objects, but, they can be checked for plausibility. For example, the estimated radius was 89.2 ± 5.4 mm when averaged across both ears and all listeners from the three HRTF databases. Assuming an adequate human sub-population in those databases, the estimated average radius is close to the average head radius of adult humans of 87.5 mm (Hartley and Fry, 1921). Also the estimated average ear positions, nearly most lateral (φ_e of $\pm 85^\circ$) and a little below the head center (θ_e of -12°), correspond to those usually found in human listeners (Algazi *et al.*, 2001b; Burkhard and Sachs, 1975).

The main motivation for the off-axis model were the large IRDs obtained with the on-axis model. For acoustically measured HRTFs, the IRDs obtained with the off-axis model (0.03 ± 1.12 mm) were much smaller and showed smaller variance than that obtained with the on-axis model (12.3 ± 22.9 mm). Thus, the off-axis model seems to fulfill the expectations of decreasing the estimated IRD without increasing the ANR. Together with the outlier removal, the off-axis model was even able to further decrease the ANR (binaural average of $28.6 \pm 12.7 \mu\text{s}$).

Two further issues were observed. First, the application of the on-axis model resulted in a positive average IRD (12.5 mm), which implies a consistent lateral offset in the measurement position. Indeed, the application of the off-axis model resulted in a positive average lateral offset y_M (on average 6.6 mm), indicating that the listeners were, on average, placed to the left of the center in the measurements in all three HRTF databases.

Second, the radius was consistently overestimated, in both on-axis and off-axis models. On average, the approximate bias was 2.5 mm for the SAT, and 5.7 mm for the STP conditions. The origin of the consistent overestimation of the radius is unclear. One speculative explanation might be the pinna, which might increase the effective propagation delay captured in the HRTFs, but is missing in the TOA model. Actually, an anthropometric

model of the ITD showed a significant statistical importance of the pinna protruding height as a parameter of the model (Xie, 2013), showing evidence for the relevance of the longer propagation path with the pinna. In our model, the underestimation of the propagation paths yields an overestimation of the head radius, an issue which might be considered in further improvements of the proposed model.

VII. CONCLUSION

A direction-continuous model for the monaural broadband TOA describing the HRTF measurement geometry was proposed. The model is fit to direction-discrete TOAs resulting from a TOA estimator applied on a set of HRTFs. The first model version, the on-axis model, assumes an exact placement of the listener in the center of the HRTF measurement setup. The on-axis model has been evaluated with calculated HRTFs of objects involving a sphere, torso, and pinna. While it was able to accurately reconstruct the geometric parameters used in calculated HRTFs, it failed when applied to acoustically measured HRTFs of human listeners in terms of large IRDs and ANRs. In the evaluation, four TOA estimators were used. The MCM estimator appeared to be the most-compatible estimator for the proposed model.

Basing on the observation that the large IRDs are linked with an off-axis placement of the object in HRTF simulations, a model able to represent measurement geometry with an arbitrary placement of the listener was proposed. This off-axis model was evaluated with calculated HRTFs of on-axis and off-axis placed objects, resulting in parameter errors smaller than the spatial resolution of the propagation paths. When applied to acoustically measured HRTFs of human listeners, the off-axis model resulted in parameters consistent with realistic listener geometry, particularly, yielding IRDs close to zero.

In order to improve the goodness of fit, an unsupervised outlier removal based on the ESD test was proposed. Together with the off-axis model, the outlier removal resulted in substantially reduced ANRs, particularly improving the problematic fits without any interaction of the operator.

The results from the application of the off-axis model show the relevance of an arbitrary position when modeling TOAs obtained from realistic HRTF measurements in human listeners. Interestingly, distinct lateral offsets were found in HRTFs of most of the listeners in all three HRTF databases tested, indicating that in those measurements lateral shifts have been involved.

The parameters of the off-axis TOA model describe the geometric setup used for the HRTF measurement. The model output is the TOA as a continuous directional function. The implementations of the TOA models with outlier removal are available in the Auditory Modeling Toolbox (Søndergaard and Majdak, 2013). These implementations allow to further analyze and modify HRTFs. For example, TOAs can be removed from each HRTF, and only the model parameters have to be stored. Such an HRTF set can be stored using shorter impulse responses. Also, the continuous representation of the TOA enables a spatially smooth interpolation of the timing between HRTFs. This is especially interesting

for the interpolation along a horizontal plane where broadband timing cues are most salient. For a binaural reproduction of virtual sound sources, the original HRTFs can be resynthesized by calculating the TOA function according to the model parameters and shifting the shortened HRIRs to the correct position in time. By considering the TOAs for each ear separately, such a representation is also suitable for binaural reproduction of multiple sound sources.⁹ There are potentially further applications for the proposed off-axis model, e.g., reconstruction of listener's position in a HRTF measurement,¹⁰ reduction of measured positions in HRTF measurements,¹¹ and pre-processing for spatial decomposition techniques with spherical harmonics (Pollow *et al.*, 2012), where a smooth spatial phase is essential.

Finally, it is important to note that the proposed model relies on a TOA estimator. The perceptual relevance of the MCM estimator and the off-axis model is, however, not clear yet. First, the model should be able to represent TOAs from HRTF sets even not originating from human beings. For example, the model might be applied to HRTFs of apes, humanoid robots, or other systems, which fulfill the requirements on the geometry representation. Second, the perceptual relevance of the TOA estimator should be defined, e.g., the hypothesis that an HRTF with the original TOA, replaced by the modeled TOA can not be discriminated from the HRTF with the original TOA. While in listener-specific HRTFs, the original perceptually relevant TOA is unknown, it still somehow must be removed. This can be done, for example, by using the minimum-phase version of the HRTF, which results in a modification of the excess phase, and unfortunately, human listeners are sensitive to phase (Moore and Glasberg, 1989). A preliminary study controlling for the effects of both the TOA model and the minimum-phase approximation did not show any major problems of the TOA model (Ziegelwanger, 2012). A more rigorous perceptual evaluation of the off-axis TOA model when applied on human HRTFs is required.

ACKNOWLEDGMENTS

We thank Christian Kasess for fruitful discussions on statistics, and Bernhard Laback and Wolfgang Kreuzer for their comments on a draft of this paper. This study was supported by the Austrian Science Fund (FWF, P 24124-N13) and the University of Music and Performing Arts, Graz.

NOMENCLATURE

α	Angular distance between a sound source and an ear
β	Angular distance between a sound source and $\vec{M}$

⁹In some applications of virtual acoustics, in order to increase the rendering efficiency, a binaural pair of HRTFs is modeled as a pair of minimum-phase impulse responses and an ITD. While the ITD is a relative quantity and thus sufficient for the description of an acoustic scene with a single sound source, the inter-source time difference is important for the acoustically correct rendering of multiple sound sources in a scene. The inter-source time difference can only be derived when the absolute delay arising from the acoustic wave propagation path from each source to a listener's ear is available, i.e., the TOA. The inter-source time difference can not be derived from the ITD. The proposed model describes the TOA as a function of direction, showing potential for an efficient rendering of multiple sound sources.

¹⁰If the actual listener position can be reconstructed after the HRTF measurements, the actually measured HRTF directions can be calculated and corrected. Then, the exact placement of the listener in the center of the measurement setup will be not required.

¹¹For a reduction of the measurement positions, a direction-continuous model of HRTFs is required. The proposed model addresses the direction-continuous timing only, and the direction-continuous spectral magnitude is still to be solved.

c	Speed of sound
γ	Auxiliary angle for the off-axis model
$\vec{e}$	Position of the ear on the sphere, $\vec{e} = [\phi_e, \theta_e]$
φ	Azimuth angle (general)
φ_e	Azimuth angle of the ear on the sphere
φ_h	Azimuth angle of the incoming sound
Φ	Lateral angle (general)
h	Head-related impulse response (HRIR)
h_{min}	Minimum-phase version of h
$\vec{h}$	Direction of the incoming sound, $\vec{h} = [\phi_h, \theta_h]$
θ	Elevation angle (general)
θ_e	Elevation angle of the ear on the sphere
θ_h	Elevation angle of the incoming sound
Θ	Polar angle
$\vec{M}$	Center of the sphere, $\vec{M} = [x_M, y_M, z_M]$
r	Radius of the sphere
s_1, s_2, s_3	Sound propagation paths
τ	Time of arrival (TOA)
$\hat{\tau}$	Direction-discrete τ resulting from an estimator
$\tilde{\tau}$	Direction-continuous τ resulting from a model
$\bar{\tau}$	Direction-independent τ
x_M, y_M, z_M	Cartesian coordinates of $\vec{M}$

References

- Algazi VR, Avendano C, Duda RO. Estimation of a spherical-head model from anthropometry. J. Audio Eng. Soc. 2001a; 49:472–479.
- Algazi, VR.; Duda, RO.; Thompson, DM.; Avendano, C. The CIPIC HRTF database. Proceedings of 2001 IEEE Workshop on Applications of Signal Processing to Audio and Electroacoustics; New York, NY. 2001b. p. 99-102.
- Bajorski, P. Statistics for Imaging, Optics, and Photonics. John Wiley and Sons; Hoboken, NJ: 2011. p. 220-227.
- Barnett, V.; Lewis, T. Outliers in Statistical Data. John Wiley and Sons; Hoboken, NJ: 1994. p. 1-616.
- Blauert, J. Spatial hearing: The Psychophysics of Human Sound Localization. MIT Press; Cambridge, MA: 1997. p. 1-494.
- Burkhard MD, Sachs RM. Anthropometric manikin for acoustic research. J. Acoust. Soc. Am. 1975; 58:214–222. [PubMed: 1150969]

- Busson, S.; Nicol, R.; Katz, BFG. Subjective investigations of the interaural time difference in the horizontal plane. 118th Convention of the Audio Engineering Society; Barcelona, Spain. 2005. Convention Paper 6324
- Coleman TF, Li Y. An interior trust region approach for nonlinear minimization subject to bounds. *SIAM J. Optimiz.* 1996; 6:418–445.
- Constan ZA, Hartmann WM. On the detection of dispersion in the head-related transfer function. *J. Acoust. Soc. Am.* 2003; 114:998–1008. [PubMed: 12942979]
- Defrance G, Daudet L, Polack J-D. Finding the onset of a room impulse response: Straightforward? *J. Acoust. Soc. Am.* 2008; 124:248–254.
- Dennis, JE, Jr.. Nonlinear least-squares. In: Jacobs, D., editor. *State of the Art in Numerical Analysis*. Academic Press; Waltham, MA: 1977. p. 269–312.
- Gardner WG, Martin KD. HRTF measurements of a KEMAR. *J. Acoust. Soc. Am.* 1995; 97:3907–3908.
- Grubbs FE. Procedures for detecting outlying observations in samples. *Technometrics.* 1969; 11:1–21.
- Hartley RVL, Fry TC. The binaural location of pure tones. *Phys. Rev.* 1921; 18:431–442.
- Huopaniemi, J.; Smith, JO. Spectral and time-domain preprocessing and the choice of modeling error criteria for binaural digital filters. *Proceedings of the 16th International Conference on Spatial Sound Reproduction of the Audio Engineering Society; Rovaniemi, Finland.* 1999. p. 301–312.
- Jeppesen, J.; Møller, H. Cues for localization in the horizontal plane. 118th Convention of the Audio Engineering Society; Barcelona, Spain. 2005. Convention Paper 6323
- Jot, JM.; Larcher, V.; Warusfel, O. Digital signal processing issues in the context of binaural and transaural stereophony. 98th Convention of the Audio Engineering Society; Paris, France. 1995. Convention Paper 3980
- Knapp, C.; Carter, GC. The generalized correlation method for estimation of time delay. *IEEE Trans. Acoust. Speech Sig. Proc., ASSP-24;* 1976. p. 320–327.
- Kreuzer W, Majdak P, Chen Z. Fast multipole boundary element method to calculate head-related transfer functions for a wide frequency range. *J. Acoust. Soc. Am.* 2009; 126:1280–1290. [PubMed: 19739742]
- Kuhn GF. Model for the interaural time differences in the azimuthal plane. *J. Acoust. Soc. Am.* 1977; 62:157–167.
- Kulkarni A, Isabelle SK, Colburn HS. Sensitivity of human subjects to head-related transfer-function phase spectra. *J. Acoust. Soc. Am.* 1999; 105:2821–2840. [PubMed: 10335633]
- Lee, B.; Choi, J.; Kim, M. Interaural time difference estimation using cross-power spectrum phase with zerocrossings. 4th International Conference on Autonomous Robots and Agents (ICARA); 2009. p. 222–226.
- Macpherson EA, Middlebrooks JC. Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited. *J. Acoust. Soc. Am.* 2002; 111:2219–2236. [PubMed: 12051442]
- Majdak P, Goupell MJ, Laback B. 3-D localization of virtual sound sources: effects of visual environment, pointing method, and training. *Atten. Percept. Psychophys.* 2010; 72:454–469. [PubMed: 20139459]
- Mehrgardt S, Mellert V. Transformation characteristics of the external human ear. *J. Acoust. Soc. Am.* 1977; 61:1567–1576. [PubMed: 893803]
- Michalopoulou Z-H, Jain R. Particle filtering for arrival time tracking in space and source localization. *J. Acoust. Soc. Am.* 2012; 132:3041–3052. [PubMed: 23145590]
- Mills AW. On the minimum audible angle. *J. Acoust. Soc. Am.* 1958; 30:237–246.
- Møller H, Sørensen MF, Hammershøj D, Jensen CB. Head-related transfer functions of human subjects. *J. Audio Eng. Soc.* 1995; 43:300–321.
- Moore BCJ, Glasberg BR. Difference limens for phase in normal and hearing-impaired subjects. *J. Acoust. Soc. Am.* 1989; 86:1351–1365. [PubMed: 2808909]
- Morimoto M, Aokata H. Localization cues in the upper hemisphere. *J. Acoust. Soc. Jap.* 1984; 5:165–173.

- Nam, J.; Abel, J.; Smith, JO, III. A method for estimating interaural time difference for binaural synthesis. 125th Convention of the Audio Engineering Society; San Francisco, CA. 2008. Convention Paper 7612
- NIST/SEMATECH. [date last viewed September 27, 2013] e-Handbook of Statistical Methods. 2002. <http://www.itl.nist.gov/div898/handbook/>
- Oppenheim, AV.; Schafer, RW.; Buck, JR. Discrete-Time Signal Processing. Prentice Hall; Upper Saddle River, NJ: 1999. 904 pages
- Pollow M, Nguyen K-V, Warusfel O, Carpentier T, Müller-Trapet M, Vorländer M, Noisternig M. Calculation of head-related transfer functions for arbitrary field Points using spherical harmonics decomposition. *Acta Acust. Acust.* 2012; 98:72–82.
- Raykar VC, Duraiswami R, Yegnanarayana B. Extracting the frequencies of the pinna spectral notches in measured head related impulse responses. *J. Acoust. Soc. Am.* 2005; 118:364–374. [PubMed: 16119357]
- Rosner B. Percentage points for a generalized ESD many-outlier procedure. *Technometrics.* 1983; 25:165–172.
- Savioja L, Houpaniemi J, Lokki T, Väänäen R. Creating interactive virtual acoustic environments. *J. Audio Eng. Soc.* 1999; 47:675–705.
- Søndergaard, P.; Majdak, P. The auditory-modeling toolbox. In: Blauert, J., editor. *The Technology of Binaural Listening*. Springer; Berlin, Germany: 2013. p. 19-44.
- Strutt, JW. *The Theory of Sound*. Macmillan; London, UK: 1894. 510 pages
- Tribolet J. A new phase unwrapping algorithm. *IEEE Trans. Acoust. Speech Sig. Proc.* 1977; 25:170–177.
- Wang L, Yin F, Chen Z. A hybrid compression method for head-related transfer functions. *Appl. Acoust.* 2009; 70:1212–1218.
- Wightman FL, Kistler DJ. Headphone simulation of free-field listening I: Stimulus synthesis. *J. Acoust. Soc. Am.* 1989; 85:858–867. [PubMed: 2926000]
- Wightman FL, Kistler DJ. Measurement and validation of human HRTFs for use in hearing research. *Acta Acust. Acust.* 2005; 91:429–439.
- Xie, B. *Head-Related Transfer Function and Virtual Auditory Display*. J. Ross Publishing; Plantation, FL: 2013. 482 pages
- Ziegelwanger, H. *Modell zur effizienten Kodierung von Signallaufzeiten für die binaurale Wiedergabe virtueller Schallquellen*. University of Music and Performing Arts; Graz, Austria: 2012. p. 1-112.

FIG. 1.

(Color online) Coordinate systems. Geodesic system (left panel) described by the azimuth ϕ and the elevation θ angles. Horizontal-polar system (right panel) described by the lateral Φ and the polar Θ angles.

FIG. 2.
Top and side view of the object STP.

FIG. 3. TOAs resulting from TOA estimators in the horizontal interaural plane applied on calculated HRTFs of the objects Sphere, SAT, and STP, as well as on measured HRTFs of an exemplary listener (NH89, DB1).

FIG. 4.

(Color online) On-axis model. Left panel: Sphere representing the head placed in the origin of the coordinate system. The circle represents the great circle defined by the direction of the incoming wave ($\vec{h}$), the center of the sphere (0), and the position of the ear ($\vec{e}$). Right panel: Projected view on the great circle showing the propagation path s_1 defining the TOA.

FIG. 5.

Estimated on-axis model parameters from HRTFs calculated for the centered object Sphere. Condition: combination of actual parameters used in HRTF simulations. Circle and cross: Parameters estimated for the left and right ears, respectively. Gray lines: Actual parameters.

FIG. 6.

Relative frequencies (%) of estimated on-axis model parameters from left-ear (top row) and right-ear (bottom row) HRTFs of human listeners. Lines: normal distribution fitted to the data. ϕ_e : Positive and negative values correspond to the left and the right ears, respectively.

FIG. 7. Estimated on-axis model IRDs (top) and off-axis model IRDs (bottom) from acoustically measured HRTFs of listeners. Line: normal distribution fitted to the data.

FIG. 8.

Relative TOAs (top row) and on-axis model fit residuals (bottom row) from HRTFs of STP (left column) and of an exemplary listener (NH89, DB1; right column). Black points: data classified as outliers by the ESD test with an upper bound of outlier rate of 1% (see text). The reference for the relative TOAs is the smallest TOA in each HRTF set. Horizontal lines: ± 1 sampling interval.

FIG. 9.

Relative TOAs of an exemplary listener (NH89, DB1) in the interaural horizontal plane for the left (black) and right (gray) ear as results from the MCM estimator (symbols) and the on-axis model (lines). The reference for the relative TOAs is the smallest TOA in HRTF sets of both ears.

FIG. 10.
Estimated on-axis model parameters from HRTFs calculated for the non-centered object Sphere. Other details as in Fig. 5.

FIG. 11.

(Color online) Off-axis model. Top-left panel: Sphere representing the head placed off-axis of the coordinate system. The large circle is the sphere's tangent circumscribed circle with the center in the origin of the coordinate system. Bottom-left panel: Projected view on that tangent circle showing the off-axis sphere and the first two components of the propagation paths. Right panel: Projected view on the great circle of the sphere showing all components of the propagation path. Note that the two projections are not necessarily in the same plane.

FIG. 12.

Estimated off-axis model parameters from HRTFs calculated for the non-centered object Sphere. Other details as in Fig. 5. ϕ_e : Positive and negative values correspond to the left and the right ears, respectively.

TABLE I

ANRs and parameter errors (average ± 1 standard deviation) resulting from fitting the on-axis model to TOAs estimated from HRTFs of the object Sphere (135 conditions, compare Sec. IV B 1). EST: TOA estimator. Parameter errors: Differences between the estimated and actual parameters. The averages and standard deviations are across the tested conditions.

EST	r error (mm)	ϕ_e error ($^\circ$)	θ_e error ($^\circ$)	ANR (μ s)
MAX	2.8 ± 0.2	1.1 ± 0.0	0.8 ± 0.0	6.5 ± 0.1
CTD	3.2 ± 0.2	1.2 ± 0.0	0.9 ± 0.1	6.9 ± 0.1
AGD	15.4 ± 0.9	1.9 ± 0.0	2.5 ± 0.0	23.6 ± 0.6
MCM	1.5 ± 0.2	1.1 ± 0.0	0.6 ± 0.0	6.1 ± 0.0

TABLE II

Parameters and ANRs resulting from fitting the on-axis model to TOAs estimated from HRTFs of the objects Sphere, SAT, and STP. Actual parameters: $r = 87.5$ mm, $\varphi_e = 85^\circ$, and $\theta_e = -10^\circ$.

	r (mm)					φ_e ($^\circ$)					θ_e ($^\circ$)					ANR (μ s)				
	MAX	CTD	AGD	MCM	MAX	CTD	AGD	MCM	MAX	CTD	AGD	MCM	MAX	CTD	AGD	MCM	MAX	CTD	AGD	MCM
Sphere	91.4	91.9	105.0	90.2	84.9	84.9	84.3	84.9	84.9	-9.6	-9.7	-8.3	-9.9	6.1	6.5	19.8	6.3			
SAT	92.1	92.9	98.8	90.4	84.9	84.9	85.1	84.9	-9.0	-8.9	-7.2	-9.6	8.0	9.8	54.1	6.6				
STP	95.4	95.4	100.7	93.2	82.9	82.8	81.5	83.2	-9.8	-9.6	-5.7	-9.4	9.4	10.4	52.8	8.8				

TABLE III

Parameters and ANRs (average $\pm$ 1 standard deviation) resulting from fitting the on-axis model to TOAs estimated from acoustically measured HRTFs of human listeners. L: Left ear. R: Right ear. All: Results for all listeners. NH89: Results for a single listener (NH89, DB1).

	Ear	r (mm)	φ_e ($^\circ$)	θ_e ($^\circ$)	ANR (μ s)
All	L	93.2 ± 12.1	88.1 ± 4.6	-5.7 ± 5.3	40.9 ± 47.1
	R	80.7 ± 11.5	-87.7 ± 8.0	-8.0 ± 5.6	38.4 ± 41.4
DB1	L	88.9 ± 6.2	86.4 ± 3.4	-6.7 ± 3.0	49.4 ± 17.4
	R	84.1 ± 5.1	-86.2 ± 4.1	-8.1 ± 3.4	47.9 ± 16.3
DB2	L	92.9 ± 12.5	87.6 ± 4.3	-6.7 ± 3.1	27.1 ± 8.3
	R	81.9 ± 12.4	-87.4 ± 7.6	-8.6 ± 5.1	29.2 ± 7.8
DB3	L	97.7 ± 14.0	90.4 ± 5.2	-3.6 ± 8.2	54.6 ± 82.5
	R	76.1 ± 12.9	-89.5 ± 10.5	-7.5 ± 7.6	44.0 ± 73.4
NH89	L	107.3	82.3	-7.8	29.9
	R	68.5	-72.6	-12.6	38.7

TABLE IV

Outlier rates obtained from the outlier detector and ANRs obtained from fitting the on-axis model to outlier-adjusted TOA sets. UB: upper bound of outlier rates used in the outlier detector (hyphen: without outlier detector). Avg $\pm$ SD: Average $\pm$ 1 standard deviation. Q₂: Median (second quartile). All: Results for measured HRTFs of all listeners.

	UB (%)	ANR (μ s)			Outlier rate (%)	
		Avg $\pm$ SD	Q ₂	Max	Q ₂	Max
SAT	-	6.7	-	-	-	-
	0.01	6.7	-	-	0	-
	0.1	6.6	-	-	0.29	-
	1	6.5	-	-	0.38	-
	10	6.4	-	-	0.81	-
STP	-	8.9	-	-	-	-
	0.01	8.9	-	-	0	-
	0.1	8.9	-	-	0.06	-
	1	8.8	-	-	0.10	-
	10	8.6	-	-	0.68	-
All	-	40.4 $\pm$ 46.5	24.0	407	-	-
	0.01	30.9 $\pm$ 13.6	22.5	89.5	0.00	2.67
	0.1	30.5 $\pm$ 13.5	22.2	89.5	0.06	4.28
	1	29.8 $\pm$ 13.5	21.6	89.5	0.16	6.42
	10	28.2 $\pm$ 13.4	20.4	89.5	0.84	10.7
DB1	-	27.6 $\pm$ 8.7	22.3	53.9	-	-
	1	25.0 $\pm$ 8.9	20.3	53.0	0.32	4.65
DB2	-	56.2 $\pm$ 43.9	38.9	322	-	-
	1	42.8 $\pm$ 18.2	29.0	89.5	0.08	6.42
DB3	-	51.8 $\pm$ 79.1	25.5	407	-	-
	1	25.8 $\pm$ 3.7	23.3	40.8	0.00	6.42

TABLE V

Parameters and ANRs (average ± 1 standard deviation) resulting from fitting the off-axis model to TOAs estimated from HRTFs of SAT, STP, and all listeners (All). Full: Fits to full TOA sets, i.e., without outlier removal. O-A: Fits to the outlier-adjusted TOA sets. L: Left ear. R: Right ear.

	TOA set	Ear	r (mm)	ϕ_e (°)	θ_e (°)	x_M (mm)	y_M (mm)	z_M (mm)	ANR (μ s)
SAT	Full	L	90.3	85.0	-8.5	0.0	-0.2	-2.2	6.5
		R	90.3	-84.9	-8.5	-0.1	0.2	-2.1	6.8
	O-A	L	90.3	84.9	-8.8	0.0	-0.2	-1.8	6.3
		R	90.3	-84.9	-8.8	-0.1	0.2	-1.7	6.6
STP	Full	L	93.2	82.9	-9.3	0.3	0.1	0.6	9.1
		R	93.2	-83.1	-9.3	0.8	0.0	0.9	9.4
	O-A	L	93.2	82.8	-9.3	0.2	0.1	0.6	9.0
		R	93.2	-83.2	-9.4	1.0	0.0	1.0	9.4
All	Full	L	89.6 $\pm$ 5.5	85.2 $\pm$ 8.4	-9.7 $\pm$ 10.7	-4.6 $\pm$ 12.2	7.1 $\pm$ 13.9	7.1 $\pm$ 17.0	39.2 $\pm$ 46.3
		R	89.5 $\pm$ 5.9	-83.3 $\pm$ 9.9	-12.7 $\pm$ 7.9	-6.5 $\pm$ 12.7	6.6 $\pm$ 17.9	9.9 $\pm$ 14.0	36.3 $\pm$ 40.9
	O-A	L	88.6 $\pm$ 5.3	85.2 $\pm$ 9.2	-12.3 $\pm$ 6.9	-4.3 $\pm$ 12.6	7.0 $\pm$ 13.3	10.5 $\pm$ 12.1	28.4 $\pm$ 13.0
		R	88.5 $\pm$ 5.5	-84.6 $\pm$ 9.3	-13.9 $\pm$ 7.4	-4.7 $\pm$ 12.3	6.2 $\pm$ 17.2	11.0 $\pm$ 12.0	28.9 $\pm$ 12.5

TABLE VI

ANRs and parameter errors (average ± 1 standard deviation) resulting from fitting the off-axis model to TOAs estimated from HRTFs of the object Sphere. Centered: Conditions in which r and $\vec{e}$ varied for $\vec{M}=0$. Non-centered: Conditions in which $\vec{M}$ varied. Other details as in Table V.

Condition	TOA set	Ear	r error (mm)	φ_e error ($^\circ$)	θ_e error ($^\circ$)	x_M error (mm)	y_M error (mm)	z_M error (mm)	ANR (μ s)
Centered	Full	L	1.6 ± 0.2	-1.3 ± 0.3	0.9 ± 0.5	-0.2 ± 0.5	0.1 ± 0.2	-0.4 ± 0.9	6.1 ± 0.1
		R	1.6 ± 0.2	1.2 ± 0.4	0.8 ± 0.6	-0.1 ± 0.5	-0.1 ± 0.2	-0.3 ± 0.9	6.1 ± 0.1
	O-A	L	1.6 ± 0.2	-1.3 ± 0.3	0.9 ± 0.5	-0.2 ± 0.5	0.1 ± 0.2	-0.4 ± 0.9	6.1 ± 0.1
		R	1.6 ± 0.2	1.2 ± 0.4	0.8 ± 0.6	-0.1 ± 0.5	-0.1 ± 0.2	-0.3 ± 0.9	6.1 ± 0.1
Non-centered	Full	L	2.2 ± 0.6	-1.3 ± 0.3	0.7 ± 0.3	-0.1 ± 0.4	-0.8 ± 0.6	-0.0 ± 0.5	6.2 ± 0.1
		R	2.6 ± 0.8	1.3 ± 0.3	0.6 ± 0.3	-0.0 ± 0.5	0.8 ± 0.8	0.2 ± 0.6	6.2 ± 0.1
	O-A	L	2.3 ± 0.5	-1.2 ± 0.3	0.7 ± 0.3	0.0 ± 0.4	-0.9 ± 0.5	0.0 ± 0.5	6.2 ± 0.1
		R	2.3 ± 0.8	1.1 ± 0.3	0.5 ± 0.4	0.2 ± 0.5	0.6 ± 0.8	0.3 ± 0.7	6.2 ± 0.1