SPATIAL-FOURIER RETRIEVAL OF HEAD-RELATED IMPULSE RESPONSES FROM FAST CONTINUOUS-AZIMUTH RECORDINGS IN THE TIME-DOMAIN

Christoph Urbanietz and Gerald Enzner

Ruhr-Universität Bochum, Department of Electrical Engineering and Information Technology,
Institute of Communication Acoustics, 44801 Bochum, Germany
christoph.urbanietz@rub.de, gerald.enzner@rub.de

ABSTRACT

Fast and comprehensive acquisition of head-related impulse responses (HRIRs) continues in the interest of rich application scenarios. Various HRIR resolutions have been presented based on discrete stop-and-go measurement, or comprehensive measurement equipment, or continuous-azimuth acquisition with moving apparatus. Some researchers have further mapped the recorded HRIR table to spatial Fourier format. In the interest of both fast acquisition and high accuracy, in this paper, we propose direct retrieval of the spatial Fourier format from time-domain recordings of the measurement signal. Specifically, we revert to the moving apparatus for HRIR measurement and express a generative model of the recorded signal based on a hypothetical spatial Fourier basis. In this approach, the spatial-Fourier model is meant to entirely capture the spatial variation of the HRIR. By least-squares inference we can then fit that model to the entire recording. Due to our time-domain treatment the model and the processing are entirely clean from block artifacts, while the huge dimension of the problem requires attention in the implementation. It turns out that the proposed algorithm significantly outperforms the previous NLMS-based continuous-azimuth acquisition of HRIR with moving apparatus.

***Index Terms*—** HRIR Acquisition, Continuous Measurement, Least Squares, Spatial Fourier Analysis

1. INTRODUCTION

Nowadays binaural rendering has many applications including hearing aids [1], voice communication, music performance, games [2], sonic detection and orientation [3], and virtual acoustic environments [4–11]. For binaural rendering head-related transfer functions (HRTFs) or their time domain representation, the head-related impulse responses (HRIRs), are needed as they describe the acoustic transfer system for a spatialized sound arriving at the ear. With the knowledge of the HRIR the sound from a virtual localized sound source can be simulated at the ear.

The demand for appropriate HRTFs can be covered either by simulated or measured HRTFs. In the case of simulated HRTFs a standard model of a head shape can be assumed or it can be individualized by taking data from an individual head, e.g., by measuring the ear shape [12, 13]. In the case of measured HRTFs there are many standard databases of non-individualized HRTFs but in principle the measurement can also be done for individuals or an individual can be best fitted to a database [14]. Many databases of HRTFs exist, but most of them exhibit a coarse discrete spatial resolution, not fulfilling the demand of high-resolution rendering approaches [15]. Conventionally the table of HRTFs is measured by a stop-and-go procedure [16–20]. An interpolation can be done using spatial, i.e., mainly angular basis functions that are fitted to the measurement points [21, 22].

There are also approaches that calculate the HRTFs from continuous measurements with unconstrained [23–26] or constrained [27–29] head movements, where some are able to deliver high-resolution at least in the azimuth direction. A further benefit of these quasi-continuous approaches is that a whole measurement can be done in a small amount of time, e.g., in only one minute, whereas the stop-and-go approach needs more time.

In this paper we introduce an improvement of these continuous algorithms in order to get a better precision of the measured HRIRs under given circumstances. Therefore we introduce a signal model of the measuring process that includes the spatial variation of the transfer function during measurement inherently. Within this model we calculate the optimal HRIR in the sense of a least-square error w.r.t. the observed measurements. The result is a continuous representation of the HRIR in spatial basis functions, coming directly from the evaluation of the measured signals and not indirectly from interpolation of spatially sparse measurements as it is usually done. As the computational dimension of this problem is high we show different approaches to overcome the computational cost. The proposed algorithm is compared with a state-of-the-art quasi-continuous azimuth measuring approach [27].

The remainder of the paper is organized as follows. Sec. 2 briefly revisits the state-of-the-art continuous-azimuth HRIR measurement. Sec. 3 then describes the improved reconstruction algorithm and Sec. 4 shows options to lower the computational costs of the new algorithm. Sec. 5 experimentally confirms the improvement of the new concept before conclusions are drawn in Sec. 6.

2. RELATION TO PRIOR ART

In the following we concentrate on the HRIR of one ear, the other follows equivalently. A continuous-azimuth measurement of the HRIR as described in [27] and visualized in Fig. 1 relies on a measurement of the noisy ear-signals $y(k)$ during a full 360° revolution of the subject of interest with a known acoustic excitation signal $x(k)$ at discrete time k. The HRIR of the length N in vector format

$$\boldsymbol{h}(\phi_k) = (h_0(\phi_k), h_1(\phi_k), \dots, h_{N-1}(\phi_k))^T \tag{1}$$

depends on the relative angle ϕ_k between head and sound source at time k. With some observation noise $n(k)$ we get

$$y(k) = \boldsymbol{h}^T(\phi_k)\boldsymbol{x}(k) + n(k) \tag{2}$$

This work is supported by Nr. EFRE-0800372 NRW grant, LS-1-1-044d.

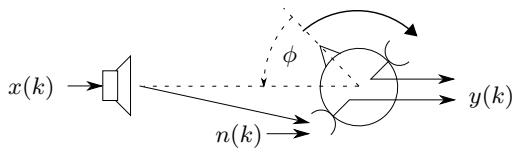

Fig. 1. Dynamic measurement of an HRIR.

with

$$\boldsymbol{x}(k) = (x(k), x(k-1), \dots x(k-N+1))^T. \quad (3)$$

The task now is to reconstruct $\boldsymbol{h}(\phi_k)$ from the measurement $y(k)$ and the excitation signal $x(k)$ under the knowledge of ϕ_k.

In [27] an algorithm was presented that identifies the time varying system from $x(k)$ to $y(k)$ with an adaptive filter and stores the actual state of the adaptive filter as estimate of the HRIR at the corresponding angle ϕ_k. This algorithm is able to retrieve an estimation of $\boldsymbol{h}$ for each ϕ_k from the continuous-azimuth measurement. However there is a trade-off in this algorithm since the normalized least mean-square (NLMS) algorithm used for system identification in [27] can either rapidly track the variation due to the spatial characteristics of the HRIR (large adaptation constant μ_0) or reject observation noise in the measurement by averaging (small μ_0).

In this paper we introduce a model of the HRIR that represents the spatial characteristics of the HRIR inherently, specifically a spatial basis representation of the HRIR, and an algorithm for averaging with respect to this model to overcome the observation noise.

3. CONCEPT OF HRIR RECONSTRUCTION

3.1. HRIR Representation in Spatial Basis Functions

We express the HRIR coefficients $h_n(\phi)$ in the horizontal plane as a weighted superposition of angular basis functions $c_q(\phi)$ by

$$h_n(\phi) = \sum_{q=-Q}^{Q} a_{q,n}^* c_q(\phi), \quad (4)$$

where $a_{p,n}$ are the coefficients of this representation, $.^*$ denotes the complex conjugate and n is ranging from 0 to $N-1$. Q is referred to as the spatial order and $P = 2Q+1$ is the number of spatial coefficients. In this work we use the complex Fourier basis given by

$$c_q(\phi) = \exp(jq\phi), \quad (5)$$

where j is the imaginary unit. The corresponding representation of the HRIR is

$$h_n(\phi) = \sum_{q=-Q}^{Q} a_{q,n}^* \exp(jq\phi). \quad (6)$$

A benefit of this choice of spatial basis that we will later make use of is that

$$c_m^*(\phi) c_n(\phi) = \exp(j(n-m)\phi) \quad (7)$$

only depends on the difference between n and m and, on the other hand, we have an orthogonal basis, so that

$$\int_0^{2\pi} c_m^*(\phi) c_n(\phi) \mathrm{d}\phi = 2\pi\delta_{m,n}, \quad (8)$$

where $\delta_{m,n}$ is the Kronecker delta. The vector representation (1) of the HRIR can then be written as

$$\boldsymbol{h}(\phi) = \sum_{q=-Q}^{Q} \boldsymbol{a}_q^* c_q(\phi), \quad (9)$$

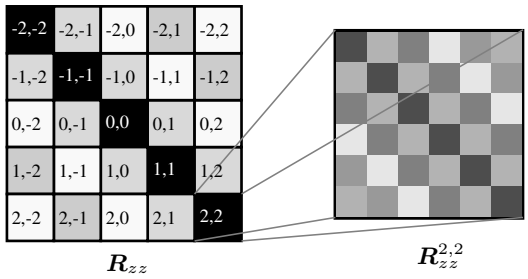

Fig. 2. Example structure of $\boldsymbol{R}_{zz}$ with $Q=2$, $P=2Q+1=5$, $N=6$, with exact block-Toeplitz structure of $\boldsymbol{R}_{zz}$, nearly block-diagonal $\boldsymbol{R}_{zz}$, and nearly Toeplitz structure of the block $\boldsymbol{R}_{zz}^{m,n}$.

where $\boldsymbol{a}_q = (a_{q,0}, a_{q,1}, \dots, a_{q,N-1})^T$. With (9) we can describe the measurement (2) as

$$y(k) = \left(\sum_{q=-Q}^{Q} \boldsymbol{a}_q^H c_q(\phi_k) \boldsymbol{x}(k) \right) + n(k), \quad (10)$$

where $\boldsymbol{a}_q^H$ is the Hermitian transpose of $\boldsymbol{a}_q$ and with stacked vectors of size $N \cdot P \times 1$ we further get

$$y(k) = \boldsymbol{a}^H \boldsymbol{z}(k) + n(k) \quad (11)$$

with

$$\boldsymbol{a} = \begin{pmatrix} \boldsymbol{a}_{-Q} \\ \boldsymbol{a}_{-Q+1} \\ \vdots \\ \boldsymbol{a}_Q \end{pmatrix} \quad \text{and} \quad \boldsymbol{z}(k) = \begin{pmatrix} c_{-Q}(\phi_k)\boldsymbol{x}(k) \\ c_{-Q+1}(\phi_k)\boldsymbol{x}(k) \\ \vdots \\ c_Q(\phi_k)\boldsymbol{x}(k) \end{pmatrix}. \quad (12)$$

3.2. Least-Squares Solution

We achieve the optimal $\boldsymbol{a}$ by minimizing the least-squares error

$$J(\boldsymbol{a}) = \sum_{k=0}^{L-1} \left| \Big(y(k) - \boldsymbol{a}^H \boldsymbol{z}(k) \Big) \right|^2 \quad (13)$$

with respect to $\boldsymbol{a}$. Here L is the number of measurement samples of the recorded signal $y(k)$. In all further considerations this is the duration of a full 360° revolution in ϕ. As the solution we get

$$\boldsymbol{a} = \boldsymbol{R}_{zz}^{-1} \boldsymbol{r}_{zy} \quad (14)$$

with

$$\boldsymbol{R}_{zz} = \sum_{k=0}^{L-1} \boldsymbol{z}(k)\boldsymbol{z}^H(k), \; \boldsymbol{r}_{zy} = \sum_{k=0}^{L-1} \boldsymbol{z}(k) y(k). \quad (15)$$

4. EFFICIENT IMPLEMENTATION

With realistic sizes of N and Q, e.g., $N=256$ and $Q=39$, the calculation of the matrix $\boldsymbol{R}_{zz}$ of size $N \cdot P \times N \cdot P$ becomes very large (i.e., more than 400 million elements) and therefore requires too much memory and computation time. We thus have to exploit the structure of $\boldsymbol{R}_{zz}$ in order to make the implementation of the algorithm more practical. Specifically $\boldsymbol{R}_{zz}$ has a block structure with $P \times P$ blocks $\boldsymbol{R}_{zz}^{m,n}$ where each block is of size $N \times N$ as visualized in Fig. 2. The block in row m and column n is given by

$$\boldsymbol{R}_{zz}^{m,n} = \sum_{k=0}^{L-1} c_m^*(\phi_k)\boldsymbol{x}(k) c_n(\phi_k) \boldsymbol{x}^H(k). \quad (16)$$

4.1. Block-Toeplitz and Sub-Toeplitz Structure

Due to (7) each block $\boldsymbol{R}_{zz}^{m,n}$ with the same difference between row m and column n is identical and therefore $\boldsymbol{R}_{\boldsymbol{zz}}$ has a block-Toeplitz structure. Thus we can reconstruct the whole matrix $\boldsymbol{R}_{\boldsymbol{zz}}$ from its first block-row or block-column. This saves us factor P in time for acquiring $\boldsymbol{R}_{zz}$. This simplification preserves exact results compared to calculating $\boldsymbol{R}_{\boldsymbol{zz}}$ in a straight-forward manner. Nevertheless the calculation effort is still quite high. Thus we rely on the independence of $x(k)$ and ϕ_k, respectively $x(k)$ and $c_m(\phi_k)$, to express $\boldsymbol{R}_{zz}^{m,n}$ for many observations, i.e., for large L, as

$$\boldsymbol{R}_{zz}^{m,n} \approx \mathcal{E}\{c_m^* \boldsymbol{x} c_n \boldsymbol{x}^H\} = \mathcal{E}\{c_m^*(\phi) c_n(\phi)\}\mathcal{E}\{\boldsymbol{x}\boldsymbol{x}^H\}, \quad (17)$$

where $\mathcal{E}$ is the expectation operator, such that $\boldsymbol{R}_{zz}^{m,n}$ becomes nearly Toeplitz structure itself. An approximation of $\boldsymbol{R}_{zz}^{m,n}$ can hence be obtained from its first row or column and an approximation of the whole $\boldsymbol{R}_{zz}$ can be obtained from merely the first row or column of $\boldsymbol{R}_{zz}$. This saves an overall acquisition time factor of $P \cdot N$.

4.2. Taylor Approximation for Matrix Inversion

The former simplifications save time in acquiring $\boldsymbol{R}_{zz}$, but this very large matrix still has to be stored and inverted. Since we do a full equidistant 360° revolution for ϕ the expectation $\mathcal{E}\{c_m^*(\phi)c_n(\phi)\}$ is due to (8) non-zero only for $m = n$. Therefore for large L the matrix $\boldsymbol{R}_{zz}$ is nearly block-diagonal. Inspired by [30] we separate $\boldsymbol{R}_{zz}$ into a matrix $\boldsymbol{D}$ containing the diagonal blocks and $\boldsymbol{\Delta}$ containing the minor off-diagonal blocks with $\boldsymbol{R}_{zz} = \boldsymbol{D} + \boldsymbol{\Delta}$ and express $\boldsymbol{R}_{zz}^{-1}$ as

$$\boldsymbol{R}_{zz}^{-1} = ((\boldsymbol{I} + \boldsymbol{\Delta}\boldsymbol{D}^{-1})\boldsymbol{D})^{-1} = \boldsymbol{D}^{-1}(\boldsymbol{I} + \boldsymbol{\Delta}\boldsymbol{D}^{-1})^{-1}, \quad (18)$$

where $\boldsymbol{I}$ is the identity matrix. By Taylor expansion of $(\boldsymbol{I} + \boldsymbol{\Delta}\boldsymbol{D}^{-1})^{-1}$ at $\boldsymbol{\Delta}\boldsymbol{D}^{-1} = \boldsymbol{0}$ we then get a good approximation

$$\boldsymbol{R}_{zz}^{-1} \approx \boldsymbol{D}^{-1}\left(\boldsymbol{I} + \sum_{b=1}^{N_b}(-1)^b(\boldsymbol{\Delta}\boldsymbol{D}^{-1})^b\right) \quad (19)$$

by truncating the expansion at order N_b. Then $\boldsymbol{a}$ can be calculated in a block-wise manner since each block of the inverse is given by

$$(\boldsymbol{R}_{zz}^{-1})^{m,n} = \boldsymbol{S}_0^{m,n} + \boldsymbol{S}_1^{m,n} + \ldots + \boldsymbol{S}_R^{m,n} + \ldots \quad (20)$$

$$\boldsymbol{S}_0^{m,n} = (\boldsymbol{D}^{0,0})^{-1}\delta_{m,n} \quad (21)$$

$$\boldsymbol{S}_1^{m,n} = -(\boldsymbol{D}^{0,0})^{-1}\boldsymbol{\Delta}^{m,n}(\boldsymbol{D}^{0,0})^{-1} \quad (22)$$

$$\boldsymbol{S}_R^{m,n} = (-1)^R \sum_{l_1 \ldots l_{R-1}=1}^{P}\left(\prod_{r=0}^{R-1}(\boldsymbol{D}^{0,0})^{-1}\boldsymbol{\Delta}^{l_r,l_{r+1}}\right)(\boldsymbol{D}^{0,0})^{-1} \quad (23)$$

where $\boldsymbol{\Delta}^{m,n}$ and $(\boldsymbol{D}^{0,0})^{-1} = (\boldsymbol{D}^{n,n})^{-1}$ can always be reconstructed from the first row of $\boldsymbol{R}_{zz}$ according to its structure.

4.3. Steepest Descent Solution

Instead of calculating (14) we can also find a solution to vector $\boldsymbol{a}$ in

$$\boldsymbol{r}_{zy} = \boldsymbol{R}_{zz}\boldsymbol{a} \quad (24)$$

with an iterative algorithm, the steepest descent (SD) algorithm [31]. We get $\boldsymbol{a}$ by

$$\boldsymbol{\rho}_d = \boldsymbol{r}_{zy} - \boldsymbol{R}_{zz}\boldsymbol{a}_d \quad (25)$$

$$\beta_d = \left(\boldsymbol{\rho}_{d-1}^H \boldsymbol{\rho}_{d-1}\right) / \left(\boldsymbol{\rho}_{d-1}^H \boldsymbol{R}_{zz}\boldsymbol{\rho}_{d-1}\right) \quad (26)$$

$$\boldsymbol{a}_d = \boldsymbol{a}_{d-1} + \beta_d \boldsymbol{\rho}_{d-1} \quad (27)$$

where we call d the iteration index and we choose $\boldsymbol{a}_0$ as zero-vector.

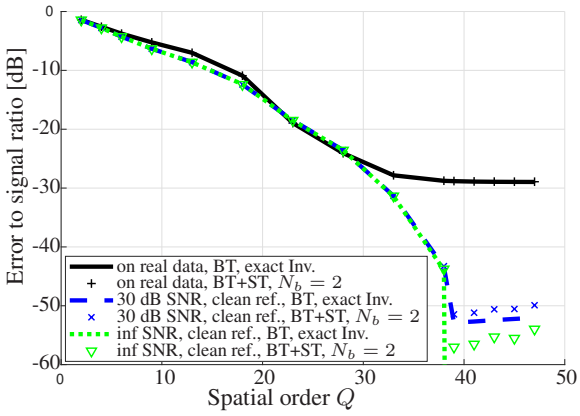

Fig. 3. ESR over spatial order Q for different experimental setups.

5. EXPERIMENTAL EVALUATION

As a measure for the accuracy of the HRIR reconstruction we define the error-to-signal ratio (ESR) between a reconstruction $\hat{y}(k)$ of the measured signal based on the estimated coefficients $\boldsymbol{a}$ and the original measured $y(k)$. For real measurements we rely on the noisy $y(k)$ as reference, but in simulations we also have a noise-free reference $y_{\mathrm{ref}}(k)$, with $y(k) = y_{\mathrm{ref}}(k) + n(k)$. Therefore we have two ESR-definitions:

$$\mathrm{ESR}_{\mathrm{noisy}} = \sum_{k=0}^{L-1}(y(k) - \hat{y}(k))^2 \Big/ \sum_{k=0}^{L-1} y(k)^2 \quad (28)$$

$$\mathrm{ESR}_{\mathrm{clean}} = \sum_{k=0}^{L-1}(y_{\mathrm{ref}}(k) - \hat{y}(k))^2 \Big/ \sum_{k=0}^{L-1} y_{\mathrm{ref}}(k)^2 . \quad (29)$$

For the noisy ESR, the SNR of $y(k)$ is a lower bound even if the linear system (HRIR) is perfectly represented by $\boldsymbol{a}$.

As a first step, in Fig. 3 we evaluate the proposed algorithm for various Q on real data to find an appropriate spatial order. The recording was done with a KMS3 dummy head of Head Acoustics GmbH with DPA d:screet SMK-SC 4060 microphone placed in the ears. The revolution time T_{360} was 60 s at $f_s = 44100$Hz. In all evaluations we use $N = 256$. We applied the exact algorithm, using the block Toeplitz (BT) structure and the exact inverse and, on the other hand, the algorithm using also the sub-Toeplitz (ST) structure and an approximation of the inverse with $N_b = 2$. Both deliver nearly the same results. Further we see that a larger number of spatial coefficients gives better results up to a limit around $Q = 39$. This corresponds to the maximum spatial frequency found in [32].

For further evaluations, in order to have a known noise-free ground-truth $y_{\mathrm{ref}}(k)$, we do simulations using a synthetic random HRIR with a spatial band-limit according to the former observations $Q = 39$. From this synthetic HRIR we generate $y_{\mathrm{ref}}(k)$ according to (10) and $y(k)$ by adding noise with known SNR.

The results for $\mathrm{SNR} = 30\mathrm{dB}$ and $\mathrm{SNR} = \mathrm{inf}$ are additionally shown in Fig. 3. With a sufficient large Q the clean ESR becomes better than the simulated measuring SNR. If Q is larger or equal to the simulated order of $Q = 39$ we get $\mathrm{ESR}_{\mathrm{clean}}$ between -60dB and -50dB, or in the case of infinite SNR and exact BT algorithm $\mathrm{ESR}_{\mathrm{clean}}$ is limited only by machine precision. The more efficient Taylor inverse introduces an limitation at about -55dB for a second order approximation.

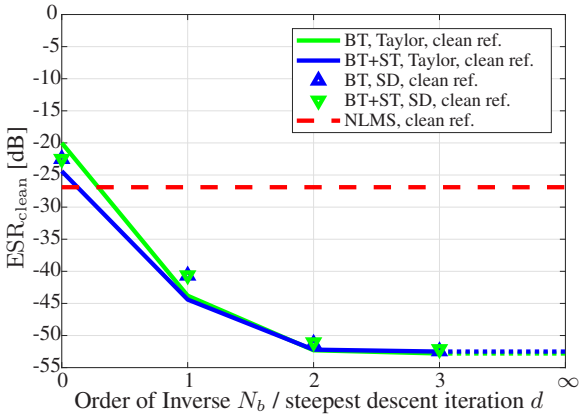

Fig. 4. ESR at $\mathrm{SNR}_{\mathrm{in}} = 30$ dB, $Q = 39$ for various N_b.

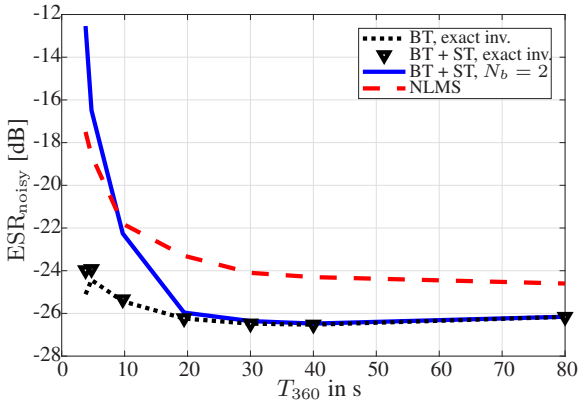

Fig. 5. ESR on real data, $Q = 39$.

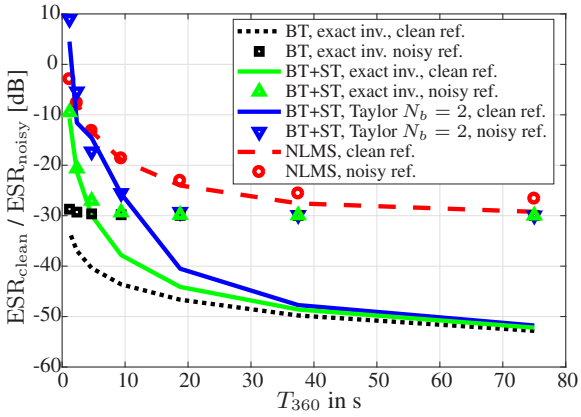

Fig. 6. ESR, simulated data at $\mathrm{SNR}_{\mathrm{in}} = 30$ dB, $Q = 39$.

As reference for further evaluations we take a state-of-the-art NLMS algorithm for continuous-azimuth acquisition on the horizontal plane as described in [27] where we store the HRIR for every processing sample, i.e., every 10^{-4} degree. The reference algorithm is tested for $\mu_0 = 1.0, 0.3$ and 0.1 and the best result $\mu_0 = 1.0$ is taken for the comparison. For the proposed algorithm we choose $Q = 39$. We test the BT and BT+ST implementation with Taylor approximation for different truncation orders N_b and the SD approximation for different iteration numbers d. $N_b = 0$ is equivalent to use only the diagonal blocks of $\boldsymbol{R}_{zz}$ while $N_b = \infty$ performs an exact inverse of $\boldsymbol{R}_{zz}$. We choose $\mathrm{SNR} = 30$ dB in Fig. 4 as a realistic scenario for HRIR measurement. In this case the approximation from utilizing a sub-Toeplitz assumption is not a limiting factor and we get almost same results for the exact BT algorithm and the BT+ST approximation. Nearly full performance is already reached for $N_b = 2$, or $d = 2$, respectively. Therefore we choose the BT+ST approximation with Taylor approximation at $N_b = 2$ as an appropriate algorithm for further considerations. It gives a good trade-off between computational effort and performance. The proposed algorithm outperforms the NLMS reference algorithm already with $N_b = 1$. The benefit of the proposed algorithm is that the model itself contains the spatial characteristics of the HRIR inherently, so that the whole amount of samples can be used to overcome the observation noise by averaging. For the NLMS algorithm instead there is a trade-off between rejecting the additive noise by averaging (small μ_0) and the ability to track the spatial characteristics of the HRIR (large μ_0).

In Fig. 5 (real data) and Fig. 6 (simulated data) we evaluate the performance of the proposed algorithm as a function of the acquisition time T_{360} at a sampling rate of 44.1 kHz. The proposed algorithm and the NLMS reference algorithm both benefit from longer acquisition time while the proposed algorithm with exact inverse outperforms the reference in all cases. The approximation of the inverse suffers heavily from short acquisition times, i.e., small L. For very short times it falls behind the NLMS algorithm but for anything beyond 10s it improves over it and reaches the level of the proposed algorithm with exact inverse for $T_{360} \geq 20$s. As suspected in the theory, the approximation of the inverse only holds for large L.

Since the $\mathrm{ESR}_{\mathrm{noisy}}$ is limited by the input SNR we cannot directly see the true performance of the algorithm under consideration with respect to the sought HRIR. Therefore we did the same evaluation with simulated $\mathrm{ESR}_{\mathrm{clean}}$ in Fig. 6. We see that the NLMS reference is again limited to the input SNR but the proposed algorithm can benefit from larger measurement time and clearly outperforms the reference. The SD approximation gives results very close to the Taylor approximation and is therefore not shown here. We see that both, the approximation in the inverse and the ST approximation only hold for long acquisition times. For any $T_{360} > 20$s these approximations are however appropriate and therefore the ESR yields nearly the level of the exact algorithm. For the sake of completeness we depict also the $\mathrm{ESR}_{\mathrm{noisy}}$ in Fig. 6 and see a similar saturation as in Fig. 5, which confirms that the results are consistent.

6. CONCLUSIONS

We introduced a new algorithm for determining HRIRs from continuous-azimuth measurements. The proposed algorithm was compared to a state-of-the-art NLMS algorithm and we showed that our algorithm is superior in all tested configurations with an appropriate implementation. The remarkable benefit over the reference algorithm is founded upon the inherent spatial HRIR model, so that the whole amount of measurement samples can be used to reject the noise floor by averaging. Although the algorithm is computationally expensive compared to the state-of-the-art algorithm we showed variations of the proposed algorithm that reduce the computational costs, either without losing precision, or with only negligible performance loss.

7. REFERENCES

[1] S. Doclo, W. Kellermann, S. Makino, and S. E. Nordholm, "Multichannel signal enhancement algorithms for assisted listening devices: Exploiting spatial diversity using multiple microphones," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 18–30, Mar. 2015.

[2] V. R. Algazi and R. O. Duda, "Headphone-based spatial sound," *IEEE Signal Process. Mag.*, vol. 28, no. 1, pp. 33–42, Jan. 2011.

[3] T. Hermann, A. Hunt, and J.G. Neuhoff, *The Sonification Handbook*, Logos Publishing House, 2011.

[4] J. Blauert, *Spatial Hearing - The Psychophysics of Human Sound Localization*, MIT Press, Cambridge, rev. edition, 1997.

[5] D. R. Begault, *3D Sound for Virtual Reality and Multimedia*, Academic Press Professional, Inc., San Diego, CA, USA, 1994.

[6] M. Geier, J. Ahrens, and S. Spors, "The soundscape renderer: A unified spatial audio reproduction framework for arbitrary rendering methods," in *124th AES Conv.*, 2008.

[7] M. Vorländer, *Auralization (RWTHedition)*, Springer Berlin Heidelberg, 2007.

[8] S. Malik, J. Fligge, and G. Enzner, "Continuous HRTF acquisition vs. HRTF interpolation for binaural rendering of dynamical auditory virtual environments," in *ITG Congr. Speech Commun.* Oct. 2010, VDE VERLAG GmbH.

[9] K. Sunder, J. He, E. L. Tan, and W. S. Gan, "Natural sound rendering for headphones: Integration of signal processing techniques," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 100–113, Mar. 2015.

[10] F. Rumsey, *Spatial Audio*, Music Technology Series. Focal Press, 2001.

[11] S. Carlile, *Virtual Auditory Space: Generation and Applications*, Neuroscience Intelligence Unit. Landes Bioscience, 1996.

[12] H. Ziegelwanger, P. Majdak, and W. Kreuzer, "Numerical calculation of listener-specific head-related transfer functions and sound localization: Microphone model and mesh discretization," *J. Acoust. Soc. Am.*, vol. 138, no. 1, pp. 208–222, 2015.

[13] M. Pollow, K. Nguyen, O. Warusfel, T. Carpentier, M. Müller-Trapet, M. Vorländer, and M. Noisternig, "Calculation of head-related transfer functions for arbitrary field points using spherical harmonics decomposition," *Acta Acustica united with Acustica*, vol. 98, no. 1, pp. 72–82, 2012.

[14] H. Gamper, A. Politis, M. Thomas, and I. Tashev, "Applications of 3D spherical transforms to personalization of head-related transfer functions," in *IEEE Int. Conf. on Acoust., Speech, and Signal Process.*, Mar. 2016.

[15] C. Urbanietz and G. Enzner, "Binaural rendering of dynamic head and sound source orientation using high-resolution HRTF and retarded time," in *IEEE Int. Conf. on Acoust., Speech and Signal Process.*, Apr. 2018, pp. 566–570.

[16] A. Andreopoulou, D. R. Begault, and B. F. G. Katz, "Inter-laboratory round robin HRTF measurement comparison," *IEEE J. of Select. Topics in Signal Process.*, vol. 9, no. 5, pp. 895–906, Aug. 2015.

[17] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, "The CIPIC HRTF database," in *Proc. of IEEE Workshop on the Appl. of Signal Process. to Audio and Acoustics*, Oct. 2001.

[18] W. G. Gardner and K. D. Martin, "HRTF measurements of a KEMAR," *J. Acoust. Soc. Am.*, vol. 97, pp. 3907–3908, 1995.

[19] J. Blauert, M. Brueggen, K. Hartung, A. W. Bronkhorst, R. Drullmann, G. Reynaud, L. Pellieux, W. Krebber, and R. Sottek, "The AUDIS catalog of human HRTFs," in *16th Int. Congr. of Acoustics*, May 1988.

[20] N. Gupta, A. Barreto, M. Joshi, and J. C. Agudelo, "HRTF database at FIU DSP Lab," in *IEEE Int. Conf. on Acoustics, Speech and Signal Proc.*, Mar. 2010, pp. 169–172.

[21] T. Ajdler, C. Faller, L. Sbaiz, and M. Vetterli, "Sound field analysis along a circle and its applications to HRTF interpolation," *J. Audio Eng. Soc*, vol. 56, no. 3, pp. 156–175, 2008.

[22] W. Zhang, T. D. Abhayapala, R. A. Kennedy, and R. Duraiswami, "Insights into head-related transfer function: Spatial dimensionality and continuous representation," *J. Acoust. Soc. Am.*, vol. 127, no. 4, pp. 2347–2357, 2010.

[23] S. Li and J. Peissig, "Fast estimation of 2d individual HRTFs with arbitrary head movements," in *Proc. of 22nd Int. Conf. on Digital Signal Process. (DSP)*, Aug 2017, pp. 1–5.

[24] N. Hahn, W. Hahne, and S. Spors, "Dynamic Measurement of Binaural Room Impulse Responses Using an Optical Tracking System," in *Int. Conf. on Spatial Audio*, Graz, Austria, Sept. 2017.

[25] J. He, R. Ranjan, W. Gan, N. Kumar Chaudhary, N. Duy, and R. Gupta, "Fast continuous measurement of HRTFs with unconstrained head movements for 3D audio," *J. of the Audio Eng. Soc.*, 2018.

[26] Sebastian Nagel, Tobias Kabzinski, Stefan Kühl, Christiane Antweiler, and Peter Jax, "Acoustic head-tracking for acquisition of head-related transfer functions with unconstrained subject movement," in *AES Int. Conf. on Audio for Virtual and Augmented Reality*, Aug. 2018.

[27] G. Enzner, "Analysis and optimal control of LMS-type adaptive filtering for continuous-azimuth acquisition of head related impulse responses," in *IEEE Int. Conf. on Acoust., Speech and Signal Process.*, Mar. 2008, pp. 393–396.

[28] G. Enzner, "3D-continuous-azimuth acquisition of head-related impulse responses using multi-channel adaptive filtering," in *IEEE Workshop on Appl. of Signal Process. to Audio and Acoust.*, Oct. 2009.

[29] G. Enzner, Chr. Antweiler, and S. Spors, "Trends in acquisition of individual head-related transfer functions," in *The Technology of Binaural Listening*, Jens Blauert, Ed., pp. 57–92. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[30] P. Sprent, "Inversion of nearly diagonal matrices," *The Mathematical Gazette*, vol. 49, no. 368, pp. 184–189, 1965.

[31] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, third edition, 1996.

[32] G. Enzner, M. Krawczyk, F. Hoffmann, and M. Weinert, "3D reconstruction of HRTF-fields from 1D continuous measurements," in *IEEE Workshop on Appl. of Signal Process. to Audio and Acoust.*, Oct 2011, pp. 157–160.