Meas. Sci. Technol. **XX** (XXXX) XXXXXX https://doi.org/XXXX/XXXX

Method for Time-of-Flight Estimation of Low Frequency Acoustic Signals in Reverberant and Noisy Environment with Sparse Impulse Response

Michael Elfering[1,2], Sven Annas[1], Hans-Arno Jantzen[1] and Uwe Janoske[2]

[1] Faculty of Mechanical Engineering, FH Münster University of Applied Sciences, Steinfurt, Germany
[2] Faculty of Mechanical Engineering and Safety Engineering, Bergische Universität Wuppertal, Wuppertal, Germany

E-mail: michael.elfering@fh-muenster.de

Received xxxxxx
Accepted for publication xxxxxx
Published xxxxxx

Abstract

For acoustic procedures which rely on the speed of sound to derive process parameters, the determination of the acoustic time of flight is essential. In this work, a method for the determination of the time of flight (TOF) is presented. It is intended for reverberant and noisy environments and can be applied in the gas holdup determination in bubbly liquids via acoustic transmission tomography (GHATT) for example. The method includes the selection and design of the transmitted signal to optimize the disambiguate of the autocorrelation, the narrowing of the time window based on the Fractional Fourier Transform (FrFT) to accelerate the TOF estimation. Furthermore, it includes the consideration of the system-induced signal distortion through prior quasi-anechoic measurements and the sparse reconstruction of the spatial impulse response for TOF estimation using non-negative sparse deconvolution algorithms. The method is tested analytically on numerically generated signals and various sparse deconvolution algorithms are investigated with respect to their applicability and limitations.

Keywords: GHATT, NLARS, reverberation, Fractional Fourier Transform

1. Introduction

Time of Arrival (TOA) and Time of Flight (TOF) estimation is a widely used and widely studied part of Time Delay Estimation (TDE) [1]. Among others, TOF is applied in localization problems with active radar and sonar systems. In the last two decades, the knowledge gained has been increasingly applied to the determination of transit times in acoustic transmission tomography systems [2, 3]. In these systems, unlike in localization, the distance between the sound sources and receivers is known in advance, and the speed of sound is the desired measurand. The knowledge of the speed of sound allows the characterization of further process variables. The dependence of process parameters on the speed of sound is already utilized in various acoustic systems and applications and is, furthermore, the subject of current

research in other areas. For example, high frequency ultrasonic signals in the form of ultrasonic tomography are applied in areas of non-destructive testing (NDT) [4] and medical diagnostics [5]. Signals with audible frequencies are used to determine temperatures and flow velocities in industrial furnaces [6], the atmosphere as well as contained environments [7–9], and in oceanography [3]. Furthermore, gas holdup determination in bubbly liquids via acoustic transmission tomography (GHATT) is the subject of current research. Elfering et. al [10, 11] utilize the dependency of sound velocity on volumetric gas content in GHATT. This dependency, first described mathematically by Wood [12], is valid for the so-called subcritical frequency range (i.e., below the resonance frequency for radial pulsations of the bubbles) [13]. Thus, for the time-of-flight determination in GHATT, the resonance frequency of the bubbles represents a natural upper frequency bound for the signals to be used.

Depending on the application and the environmental conditions, noise as well as reflections from surfaces occur in transmission tomography, thereby complicating the accurate determination of the transit times using established methods.

In acoustic transmission tomography where reflections are superimposed on the direct sound, various methods have been established, they include the following methods [14–16]. Simple cross-correlation-based methods are widely used in TOF estimation but prove to be particularly vulnerable when direct sound is superimposed with reflections, since here the receiver signal is not a simple time-delayed copy of the transmitted signal. More advanced cross correlation methods that are more robust to reflections are susceptible to noise and signal distortion [15, 17]. TOF estimation based on the Akaike Information Criterion (AIC), can be applied directly to the receiver signal without a reference signal, similar to threshold methods and other characteristic functions. However, the estimated TOF is highly dependent on the noise component. For cross-talk interference and reflection echoes, they are only applicable under certain conditions [15].

In ultrasonic pulse echo methods, which determine distances based on TOF of reflected or scattered high-frequency signals, blind deconvolution approaches or deconvolution approaches with modeled building blocks are used for the TOF estimation [14, 16]. These modeled building blocks of the dictionary describe theoretical signal components that can be expected under defined reflection and scattering conditions of the high-frequency signals. In contrast, these models are not useful for transmission tomography with low-frequency signals, since in these cases the building blocks can better be derived directly from the defined transmitted signal.

The method presented in this paper is intended to counteract the limitations of previous methods for use in GHATT and other acoustic transmission tomography applications. For this purpose, the proposed method exploits the non-negative sparse structure of the room impulse response of reverberant rooms by applying a non-negative sparse deconvolution for the first time in TOF estimation. It focuses on the sparse early reflections (echoes), which promptly follow the direct sound, and are thus of particular relevance for TOF estimation. A signal design is described, which on the one hand is unambiguously correlated. On the other hand, the short signal durations mean that later dense reflections (reverb) do not superimpose with the direct sound and can therefore be ignored in the deconvolution. The proposed method offers the possibility to consider distortions caused by the system, e.g. by the transducers. Fractional Fourier transform (FrFT), which works particularly effectively with linear frequency modeled signals, is used for truncation of the receiver signal before deconvolution.

The remainder of the paper is organized as follows. First, in Section 2, an overview of TOF estimation approaches is given, and, based on this, a new approach is derived in Section 3. Finally, in Section 4, this new approach is tested using simulated signals.

2. Methods for Time-of-Flight Estimation

For TOF estimation in transmission tomography, the transmitted signal and received signal must be compared. In the following, both the source signal and the recorded received signal are assumed to be known, which can generally be taken for granted in acoustic transmission tomography like GHATT.

2.1 Signal Model

Under ideal environmental conditions, the received signal $x(t)$ of a measurement path is a time-delayed and attenuated copy of the source signal with additive noise.

$$x(t) = \alpha \cdot s(t - \tau_{TOF}) + w(t) \tag{1}$$

Here, α describes the attenuation factor, $s(t)$ denotes the transmitted source signal, $w(t)$ denotes the additive noise and τ_{TOF} denotes the TOF. Therefore, the ideal propagation model only considers the direct path. This model can be applied to an ideal free field environment without any substantial acoustic reflections.

In many situations, considering only the direct path is insufficient because reflecting walls and surfaces affect the recorded signal. In such cases, the transmitted signal arrives at the receiver multiple times with individual attenuation and time delay. This multipath model, as it is commonly called, can be described by an additive superposition of several signals from individual paths [18].

$$x(t) = \sum_{j=1}^{k} \alpha_j \cdot s\big(t - \tau_j\big) + w(t) \tag{2}$$

Here, k represents the number of signal paths; α_j and τ_j denote the individual attenuation factors and propagation times of these paths. The TOF is the propagation time τ_j, which is associated with the direct path. Hence, it has the minimum value of all propagation times. In the case of discrete time sampling with a sampling period T, equation (2) leads to (3).

$$x(nT) = \sum_{j=1}^{k} \alpha_j \cdot s\left(nT - \tau_j\right) + w(nT), n \in \mathbb{Z} \qquad (3)$$

Although this multipath model is not valid for every environment [1], it is valid and has been established for most applications with hard reflecting boundaries and non-dispersive media such as room acoustic environments [19].

The discrete multipath model in (3) represents a special case of a more general convolutional reverberant model. At this point, it must be noted that the times τ_j are non-discrete in general. Hereafter, they are discretized with the sampling period T, in the same way as the recorded signals are. From now on, they are treated as integer multiples of T.

$$\mathbf{x}[n] = \sum_{l=1}^{L} \mathbf{h}[l] \cdot \mathbf{s}[n-l] + \mathbf{w}[n] \qquad (4)$$

Here, $\mathbf{h} \in \mathbb{R}^L$ is the impulse response of the system. In the special case given in (3), the impulse response is sparse and consists of only k non-zero entries with amplitudes α_j at the points $l = \tau_j T^{-1}$ for $j = 1, \dots, k$.

$$\mathbf{h}[l] = \begin{cases} \alpha_j & \text{if } l = \tau_j T^{-1} \text{ for } j = 1, \dots, k \\ 0 & \text{otherwise} \end{cases} \qquad (5)$$

The sparse convolution model described in (4) can be rewritten as a discrete convolution in vector notation with $\mathbf{x} \in \mathbb{R}^N$, $\mathbf{s} \in \mathbb{R}^M$, $\mathbf{h} \in \mathbb{R}^L$ and additive noise $\mathbf{w} \in \mathbb{R}^N$:

$$\mathbf{x} = \mathbf{s} * \mathbf{h} + \mathbf{w} \qquad (6)$$

With $*$ the discrete convolution operator is denoted. Alternatively, this convolution can be expressed as a linear matrix product of $\mathbf{S}$ and $\mathbf{h}$, where $\mathbf{S} \in \mathbb{R}^{N \times L}$ with $L = N - M + 1$ denotes a Toeplitz matrix with possible time-delayed copies of the transmitted signal $\mathbf{s} = [s_1, \dots, s_M]^{\mathrm{T}}$ [20]:

$$\mathbf{x} = \mathbf{S}\,\mathbf{h} + \mathbf{w} \qquad (7)$$

with

$$\mathbf{S} = [\mathbf{s_1} \quad \mathbf{s_2} \quad \cdots \quad \mathbf{s_L}] = \begin{bmatrix} s_1 & 0 & \cdots & 0 \\ \vdots & s_1 & \ddots & \vdots \\ s_M & \vdots & \ddots & 0 \\ 0 & s_M & \ddots & s_1 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & s_M \end{bmatrix}$$

In enclosed spaces with acoustically reflective walls, such as an industrial plant, the transmission model can be approximated by (7). The transmitted signal of a GHATT measurement thus arrives at the receiver temporally overlapping with the reflections.

2.2 Sparse Deconvolution Approach

An obvious approach to determine the TOF of signals according to (6) and (7) is to invert the convolution (deconvolution). The goal is to invert the problem and to reconstruct the impulse response $\hat{\mathbf{h}} \in \mathbb{R}^L$. For a TOF estimation, the first peak of this reconstructed impulse response has to be identified. This peak has the shortest time of flight and thus represents the direct sound. Therefore, its time delay corresponds to the sought estimated TOF $\hat{\tau}_{TOF}$. A direct linear deconvolution without constraints on the impulse response is sensitive to additive noise and corresponds to the ordinary least-square optimization without constraints or penalization term. For real-life signals with noise, a sparse deconvolution is more suitable and established for determining the impulse response of multipath systems [1]. The following underdetermined optimization problem arises:

$$\underset{\hat{\mathbf{h}}}{\operatorname{argmin}} \left\|\mathbf{x} - \mathbf{S}\,\hat{\mathbf{h}}\right\|_2 \text{ s.t. } \left\|\hat{\mathbf{h}}\right\|_0 \leq \hat{k} \qquad (8)$$

Here, the ℓ_0-pseudo-norm $\|\cdot\|_0$ denotes the number of non-zero components with an upper bound $\hat{k}$. For solving this problem, the two families of Greedy algorithms and convex optimization have been established for reconstructing the sparse signal from the measured values.

2.2.1 Greedy Algorithms. Orthogonal Matching Pursuit (OMP) and Orthogonal Least Squares (OLS) are two important representatives of Greedy algorithms [21]. $\boldsymbol{r}_j$ represents the residual of the j-th iteration.

$$\mathbf{r}_j = \mathbf{x} - \mathbf{S}\,\hat{\mathbf{h}}\big|_j$$

As Greedy algorithms, they both start from an all zero solution of $\hat{\mathbf{h}}$, and initialize the residual $\mathbf{r}_0 \in \mathbb{R}^N$with the recorded measurement $\mathbf{r}_0 = \mathbf{x}$ (i.e. at the beginning no atom $\mathbf{s}_i$ of the dictionary $\mathbf{S}$ is selected). From there, they both recursively increase the number of non-zero entries in $\hat{\mathbf{h}}$ one-by-one at each iteration, but they differ in the way of choosing new entries. In OMP [22] at the j-th iteration, the atom $\mathbf{s}_i$ that is maximally correlated with the residual $\mathbf{r}_{j-1}$ of the previous iteration is chosen:

$$\underset{i}{\operatorname{argmax}} \left|\langle \mathbf{r}_{j-1}, \mathbf{s}_i \rangle\right|,$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product. Then, the values of the non-zero identified entries of $\hat{\mathbf{h}}$ are chosen such that $\mathbf{r}_j$ is orthogonal to $\mathbf{S}\,\hat{\mathbf{h}}$:

$$\hat{\mathbf{h}} = \underset{\hat{\mathbf{h}}}{\operatorname{argmin}} \left\| \mathbf{x} - \mathbf{S}\,\hat{\mathbf{h}} \right\|_2$$

In OLS [23], on the other hand, the ℓ_2-norm $\left\|\mathbf{r}_j\right\|_2$ is minimized by adjusting the values among all previously identified supports including one new component simultaneously:

$$\underset{\hat{\mathbf{h}}_{j-1},h_i,i}{\operatorname{argmin}} \left\| \mathbf{x} - \mathbf{S}\,\hat{\mathbf{h}}_{j-1} - \mathbf{s}_i\, h_i \right\|_2$$

The recursive Greedy algorithms OLS and OMP are repeated until $j = \hat{k}$.

2.2.2 Convex Optimization. In convex optimization, the sparse deconvolution problem is addressed not by bounding the ℓ_0-pseudo-norm, as in (8), but by an ℓ_1-penalization [21]. Thus, the sparse deconvolution problem is transformed into the following optimization problem by the least absolute shrinkage and selection operator (LASSO) approach:

$$\underset{\hat{\mathbf{h}}}{\operatorname{argmin}} \frac{1}{2} \left\| \mathbf{x} - \mathbf{S}\,\hat{\mathbf{h}} \right\|_2^2 + \lambda \left\| \hat{\mathbf{h}} \right\|_1 \qquad (9)$$

The advantage of switching to this convex problem is, that efficient algorithms like the modified least-angle regression (LARS), the so-called homotopy algorithm, exist to approach their solution. Here, starting from $\lambda_0 = \max|\mathbf{S}^{\mathbf{T}}\,\mathbf{x}|$, iteratively reducing the value of λ until a total of $\hat{k}$ distinct entries in $\hat{\mathbf{h}}$ have been identified as non-zero in at least one of the iteration steps. [24]

2.2.3 Non-negative sparse Deconvolution. Since the direct sound as well as the reverberant wall reflections are recorded non-inverted by the receiver, the reliability of the sparse deconvolution can be improved by restricting the impulse response to non-negative entries only. Lin and Lee exploited this non-negative, sparse structure of acoustic room impulse response first [25]. Here, the sparse impulse response was reconstructed from the received signal using least-square optimization with ℓ_1-penalization. The problem from (8) thus becomes (10).

$$\underset{\hat{\mathbf{h}}}{\operatorname{argmin}} \left\| \mathbf{x} - \mathbf{S}\,\hat{\mathbf{h}} \right\|_2 \text{ s.t. } \left\| \hat{\mathbf{h}} \right\|_0 \leq \hat{k}, \qquad \hat{h}_i \geq 0 \qquad (10)$$

The problem described in (10) is similar to the non-negative matrix factorization (NMF), which already provides sparse solutions. However, sparsity in regular NMF results as a by-product, since only non-negative components are allowed [26]. However, there are several approaches of NMF which enforce sparseness. Moreover, there are extensions of the Greedy algorithms [20] and convex optimization approaches [27] that enforce non-negativity. Bruckstein et al. first introduced a non-negative implementation of the OMP algorithm (NNOMP) [28]. Where the non-negativity of $\hat{\mathbf{h}}$, besides selecting only non-negative atoms, is ensured by determining the values using the non-negative least square (NNLS). For a faster implementation, Yaghoobi et al. [29] developed the fast implementation as Fast NNOMP (FNNOMP), which dispensed with the time-intensive computation of the pseudo-inverse of the dictionary. The FNNOMP algorithm has a slightly different atom selection procedure which also does not always guarantee identical results. Peharz and Pernkopf proposed a sparse approach to NMF (sNMF, Sparse NNLS), pointing out parallels to NNOMP, even though identical outputs are not always guaranteed [26]. Yaghoobi and Davies first proposed a non-negative OLS (NNOLS) and a faster approximated suboptimal NNOLS (sNNOLS) [30]. A non-negative implementation of the LARS algorithm (NLARS) was proposed by Mørup et al. [27].

2.3 Akaike Information Criterion

The use of the Akaike Information Criterion (AIC) for TOF estimation of seismic signals was described by Sleeman and van Eck [31]. It was further developed and used for time of flight estimation of ultrasonic signals for the first time by Li et al. [32]. Here the AIC is calculated as:

$$\mathrm{AIC}(n) = n \cdot \ln\big(\mathrm{Var}(x_1 \dots x_n)\big) + \big(N - (n+1)\big) \cdot \ln\big(\mathrm{Var}(x_{n+1} \dots x_N)\big)$$

The TOF $\tau_{\mathrm{TOF,AIC}}$ is calculated as a weighted average:

$$\tau_{\mathrm{TOF,AIC}} \coloneqq \sum_{n=1}^{N} W_n \cdot (n-1)T$$

with

$$W_n \coloneqq \frac{\exp(-\Delta_n/2)}{\sum_{j=1}^{N} \exp\big(-\Delta_j/2\big)}$$

$$\Delta_n \coloneqq \mathrm{AIC}(n) - \min(\mathrm{AIC})$$

Since this method is susceptible to interference and cross-talk, it was adapted by Bao and Jia for labscale use of an acoustic transmission tomography system to determine temperature distributions [15]. They added an adaptive analysis window, which limits the analysis to the relevant region and ignores later reflections. In addition, they introduced a phase correction, which improves the robustness of the AIC against signal noise. The phase of the signal is reconstructed from $\tau_{\mathrm{TOF,AIC}}$ to the maximum of the signal's envelope. The phase shift of the reconstructed phase with respect to a reference signal is estimated and considered in the TOF estimation. This has significantly improved the results; however, a reliable correction assumes that the phase within the evaluated range corresponds to that of the direct sound and has not been superimposed and influenced by reflections

significantly within this range. This step is therefore of limited use for reflections that closely follow the direct sound.

2.4 Cross-correlation-based Methods

Another more intuitive way to determine the TOF is the generalized cross correlation (GCC). Here, the transmitted and the recorded signal are correlated in a frequency-dependent weighted manner [33].

$$R_{sx}(\tau) = \mathcal{F}^{-1}\{\Psi\, \overline{\mathcal{F}\{s\}}\, \mathcal{F}\{x\}\} \tag{11}$$

$\mathcal{F}\{\cdot\}$ describes the Fourier transformation; $\bar{\cdot}$ describes the complex conjugate; $\mathcal{F}^{-1}\{\cdot\}$ describes the inverse Fourier transform function and Ψ denotes a weighting function. In the case of non-reverberant sound propagation, as in (1), the estimated TOF $\hat{\tau}_{TOF,GCC}$ can be obtained by the maximum of the correlation and a positive constant weighting function.

$$\hat{\tau}_{\mathrm{TOF,GCC}} \coloneqq \underset{\tau}{\operatorname{argmax}}\big(R_{sx}(\tau)\big) \tag{12}$$

However, this method is shown to be unsuitable for multipath problems. Therefore, the weighting function according to the generalized cross correlation with phase transform (GCC-PHAT) has been established for these applications [33]. It is defined as

$$\Psi_{PHAT} \coloneqq \left|\overline{\mathcal{F}\{s\}}\, \mathcal{F}\{x\}\right|^{-1}.$$

This results in sharp peaks in the GCC-PHAT function for ideal noise-free signals, similar to those peaks in h in (4). These peaks can be assigned to the propagation times τ_j. The estimated TOF $\hat{\tau}_{TOF,GCC}$ results from the first peak of the correlation function, as in (3). However, the GCC-PHAT is sensitive to additive noise and the improvement of this issue is the subject of current research [17].

2.5 Fractional Fourier Transform based Methods

As an alternative to the methods mentioned above, there is an established analytical approach for pulse compression and TOF estimation. In this method, linear frequency modeled signals (sweeps) are generated and transmitted by a transducer. The recorded signals can be analyzed using the Fractional Fourier Transform (FrFT) [34] defined as

$$\mathcal{F}^{a}\{x(t)\}(u) \coloneqq \int_{-\infty}^{\infty} B_{\varphi}(u,t)\, x(t)\, \mathrm{d}t \tag{13}$$

with

$$B_{\varphi}(u,t) \coloneqq A_{\varphi} \exp\left[i\pi\left(u^2\cot(\varphi) - \frac{2ut}{\sin(\varphi)} + t^2\cot(\varphi)\right)\right],$$

$$A_{\varphi} \coloneqq \frac{\exp\left(-\frac{i\pi}{4}\operatorname{sgn}(\sin(\varphi)) + \frac{i\varphi}{2}\right)}{|\sin(\varphi)|^{\frac{1}{2}}}$$

and

$$\varphi \coloneqq a\frac{\pi}{2}.$$

With this method, the signals are not examined in the time or Fourier domain but in an intermediate domain, the fractional Fourier domain u. If the choice of the order a and the angle φ of the fractional domain is tuned to the slope of the chirp-rate (i.e. the rate of frequency change) of the signal, the energy of the signals of every path projected into the fractional domain will be compressed [35]:

$$a_{opt} = -\frac{2}{\pi}\cdot\arctan\left(\frac{T_s}{2\,T^2\,N\,\Delta f}\right)$$

This helps to identify the signals of individual paths (s. Figure 1) and has been successfully implemented in NDT [35] [36] and medical diagnostics [37]. However, there are limitations to the TOF estimation; the cause is analogous to the classical Fourier transform, where signals with finite signal duration yield main lobes with significant width in Fourier domain. Signals with multiple closely spaced frequency components superimpose in the Fourier domain. This also applies to FrFT with swept signals and closely spaced reflections, as in GHATT. Despite these limitations, there is an application of this method for these signals (s. Section 4).

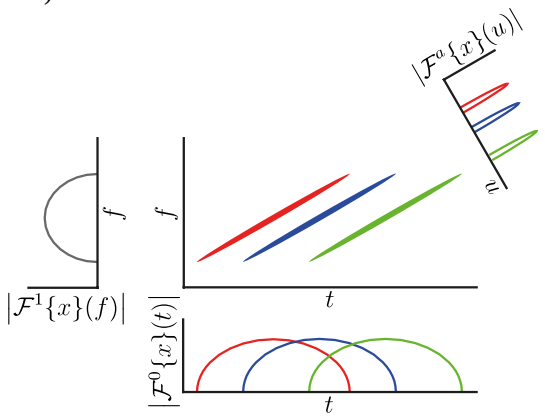

Figure 1. Schematic representation of the Fractional Fourier Transform of a signal with three time-delayed linear sweeps

3 Proposed Method for TOF estimation

Building on previous work, a new method for determining the TOF and the effective sound slowness of low-frequency acoustic signals between transmitters and receivers is presented below. This method can be used in GHATT to determine the gas content distribution and other acoustic transmission tomography application. Figure 2 schematically shows the steps of the proposed method.

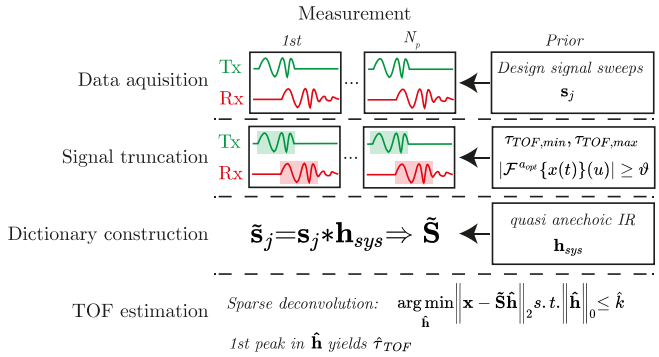

Figure 2. Scheme of the individual steps of the proposed method

The proposed method uses a non-negative sparse deconvolution and exploits the fact of the non-negative sparse structure of the room impulse response of sound-hard rooms for TOF estimation for the first time. To implement this, the method concentrates on the sparse early reflections (echo), which follow closely in time to the direct sound. A signal design is described, which consists of one or more short signal sweeps, so that the direct sound does not overlap with the later dense reflections (reverb). This short signal duration has the advantage that the dense reverb does not have to be reconstructed during deconvolution, since its influence on the TOF estimation is negligible, as these signal components do not arrive simultaneously with the direct sound. A consideration of system-conditioned distortions is realized by prior quasi-anechoic measurements during the deconvolution. FrFT is used for the first time for truncation of the receiver signal before deconvolution, which is particularly robust for linear frequency-modulated signals.

3.1 Signal Design

Two aspects are particularly relevant for the selection of the frequency spectrum of the transmitted signal. On the one hand, the signal should only contain frequency components that can be reliably reproduced by the transducers. On the other hand, when determining the signal propagation time in bubbly liquids, the frequency spectrum has to be selected in such a way that all signal components are completely below the resonance frequency of the bubbles [13]. Only in this range, the Wood correlation is valid for the GHATT measurements and the dispersion of the phase velocity becomes negligible.

For pulse compression and to reduce temporal overlap of the direct sound with the reflections, the proposed method uses a short signal with sufficient length and energy for a good signal-to-noise ratio (SNR) and a linearly modulated frequency according to (14).

$$s(t) = W(t) \cdot \mathcal{R}\mathrm{e}\left[A \cdot \exp\left(2\pi i \cdot \left[\left(f_0 - \frac{\Delta f}{2}\right)t + \frac{\Delta f}{2T_s}t^2\right]\right)\right], \quad \text{for } 0 \leq t \leq T_s \tag{14}$$

Here, $W(t)$ denotes a window function, $\mathcal{R}e(\cdot)$ denotes the real part of the signal, A denotes the maximum amplitude of the signal, f_0 is the center frequency, Δf is the bandwidth, and $T_s = (M-1)T$ is the duration of the sweep. The linear frequency modulation provides a pulse compression after a cross correlation. This is beneficial for TOF estimation with Greedy algorithms or convex optimization, since the interferences between the different columns of the dictionary are reduced (lower correlation).

In principle, the SNR can be improved by increasing the signal duration. However, increasing the signal duration leads to a greater temporal overlap of the direct and reflected sound. Therefore, an alternative is chosen in this work. Instead of one long signal, several short signals with intermediate pauses are used. That means, first a primary signal is emitted. At the same time, the received signals are recorded. After the time window in which the direct sound and the first reflections occur, the recording is stopped. Now, the sound energy can decay in the measuring volume. After that, another signal is emitted, so that the received signal cannot be influenced by the reverberation of the first signal. The signals used and their frequencies are different in each instance. The signals are selected in such a way that the improvement of the pulse compression is achieved in the best possible way.

The goal of pulse compression is to make the autocorrelation as close as possible to an ideal Dirac function, so that the correlation of the atoms in the dictionary is minimized. Sufficient pulse compression thus directly improves the disambiguation of the correlation-based atom selection of the OMP. But also the LS term of the OLS and the LARS algorithms benefit from the improved autocorrelation, since there is a simple and direct relation between the cross-correlation and the mean-squared distance of two signal vectors [38].

A single signal sweep with limited bandwidth always shows side lobes in the autocorrelation. This can lead to ambiguities in peak detection, especially with short observation times. The reduction of side lobes is therefore important and can be achieved by a suitable window function. In the following, a von-Hann window is used to reduce unwanted frequency components and the unwanted side lobes of the autocorrelation. In order to optimize the autocorrelation by additional signals beyond this, the individual signals must be appropriately tuned to each other. For this purpose, it is desirable that the negative side lobes of one signal coincide with the positive ones of another signal in order to cancel them by destructive interference. This is achieved, for example, by scaling the center frequency $f_{0,j}$ and the bandwidth Δf_j of the total of N_p individual signals by the factor $a_{f,j}$:

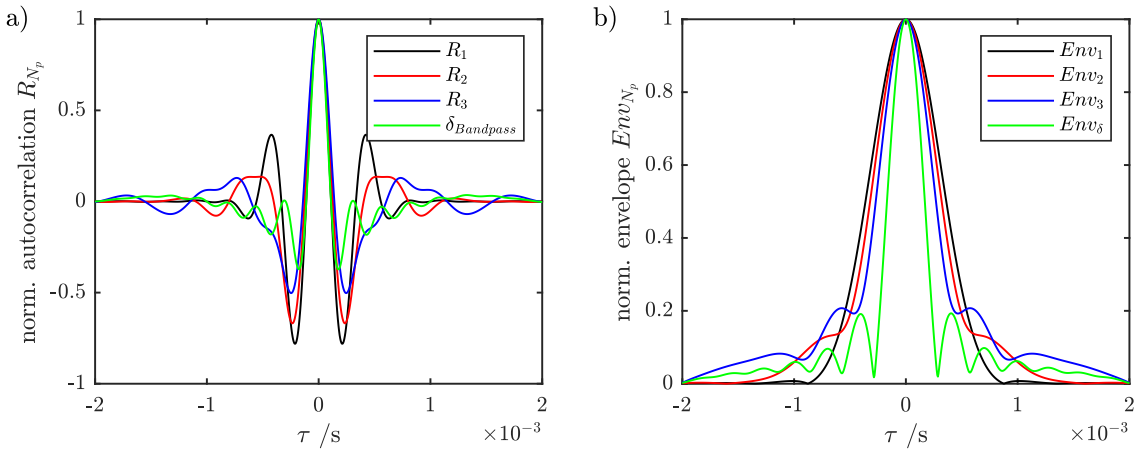

Figure 3. a) Normalized autocorrelation R_{N_p} with N_p signal sweeps and, for comparison, a frequency-limited Dirac function $\delta_{Bandpass}$; b) Envelopes Env_{N_p} of the autocorrelations; the use of multiple successive signal sweeps allows a significant improvement of the unambiguity of the autocorrelations by reducing the side lobes.

$$\Delta f_j = a_{f,j} \cdot \Delta f$$
$$f_{0,j} = a_{f,j} \cdot f_0 \qquad (15)$$

with

$$a_{f,j} = \left(\frac{3}{2}\right)^{N_p - j}$$

Figure 3 a) shows the normalized autocorrelation of a single sweep and the normalized autocorrelation of two and three consecutive sweeps with individually scaled frequencies according to (15). To ensure comparability, the total bandwidth Δf = 3500 Hz around the total center frequency f_0 = 2250 Hz is kept. This means that only frequencies between 500…4000 Hz are used in each case. It can be seen that, by using multiple signals, the unambiguity of the autocorrelations R_2 and R_3 can be significantly improved compared to the autocorrelation R_1 with only a single sweep; Hence, the side lobes have significantly lower amplitudes. The shown Dirac function $\delta_{Bandpass}$, which is frequency-limited by a bandpass, illustrates the optimum which can be realized with signals using the limited frequency range; it only serves as a comparison. Also, the main lobe of the envelope Env_{N_p} of the autocorrelations in Figure 3 b), which considers not only the real part of the autocorrelation but also the imaginary part of its analytic signal, is determined according to

$$Env_{N_p} := \left| R_{N_p} + i\mathcal{H}\left\{R_{N_p}\right\} \right|$$

with the Hilbert transform $\mathcal{H}\{\cdot\}$.

It is shown that by combining multiple sweeps, the amplitudes of the first negative as well as the first positive side lobes can be significantly reduced without changing the total bandwidth or the duration T_s of the single sweep. A requirement for the efficient application of this multiple sweep method is that the decay time is significantly smaller than the temporal changes of the impulse responses.

3.2 Signal Recording

First, the transmitter and receiver are placed at precisely known positions in the medium to be examined. The measurement signal is transmitted by one acoustic transmitter and received by one or more receivers. For the precise determination of the transit time, a parallel recording of the receiving signals simultaneously with the transmitted signal via a multi-channel acquisition unit is required.

3.3 Truncation of Received Signal

For the TOF estimation only the direct sound is relevant. The first reflections, which occur in superposition with the direct sound, are additionally relevant for its estimation. Therefore, only the time window in which the direct sound and the first reflections occur is of interest. For the determination of this time window, two methods are suitable. The first method requires the knowledge of a maximum and minimum possible speed of sound. This is always the case with GHATT, since the influence of the gas content on the speed of sound c varies only to a known extent [13]. Therefore, also the TOF sought can only be within a limited time window $(\tau_{TOF,min}, \tau_{TOF,max})$. If, in addition, prior information about the minimum and maximum expected gas hold up $(\varepsilon_{min}, \varepsilon_{max})$ and the distance d are known, this time window can be further restricted.

$$\tau_{TOF,min} = \frac{d}{c(\varepsilon_{min})}$$

$$\tau_{TOF,max} = \frac{d}{c(\varepsilon_{max})}$$

In addition, a second method can be used to further limit the time window. Compared to the Short Time Fourier

Transform (STFT) [16], the FrFT is especially suitable for swept signals, since the FrFT is used to identify signal regions with high energy content in the fractional frequency domain matching the chirp-rate of the transmitted sweep. Although the FrFT has its limitation in the identification of individual sweeps in close succession, it can be used to identify the relevant time window (s. Section 4). This is done by selecting the signal part which exceeds a defined threshold ϑ in the fractional frequency domain:

$$|\mathcal{F}^{a_{opt}}\{x(t)\}(u)| \geq \vartheta$$

Subsequently, the region in u identified as relevant can be transferred back to the time domain:

$$u^* = \frac{u}{a_u} \qquad a_u = \cos(\varphi_{opt}) \tag{16}$$

The intersection of the determined time windows results in the relevant range for the TOF estimation. The original signal $\mathbf{x}'$ can be limited to this time range. The result is the truncated signal $\mathbf{x}$.

3.4 Dictionary Construction

The dictionary required for deconvolution is constructed from the transmitted signal recorded simultaneously. First, the sweep of the transmitted signals is extracted. If the time of transmission is not known exactly, it can be determined by the theoretical signal with the help of the GCC-PHAT. The transmitted signal is then cropped to the known signal length M.

Depending on the transducers used, the distortion of the transmission system has to be considered. That is, if the signal is considerably distorted by the transmitters and/or the receivers, the recorded signal results from the convolution of the transmitted signal $\mathbf{s}$ with the anechoic impulse response of the measurement system $\mathbf{h}_{sys}$ and the sought impulse response of the room $\mathbf{h}$. Thus, (6) becomes

$$\mathbf{x} = \tilde{\mathbf{s}} * \mathbf{h} + \mathbf{w} \tag{17}$$

with

$$\tilde{\mathbf{s}} = \mathbf{s} * \mathbf{h}_{\mathrm{sys}}.$$

Thus, $\tilde{\mathbf{s}}$ corresponds to the transmitted signal $\mathbf{s}$ after considering the distortion caused by the system. In NDT using ultrasound, the distortion of the system is approximated by convolution models [39]. To account for this system-induced bias in GHATT, $\mathbf{h}_{\mathrm{sys}}$ has to be determined in advance under anechoic or quasi-anechoic conditions (see [40] for more information). From $\tilde{\mathbf{s}}$, the dictionary $\tilde{\mathbf{S}}$ is constructed for the deconvolution according to (7) as a Toeplitz matrix.

3.5 TOF Estimation

For the TOF estimation, the deconvolution of (17) is performed using one of the algorithms presented in Section 3 to estimate $\hat{\mathbf{h}}$. The use of a $\hat{k}$ -sparse non-negative algorithm is feasible if it is guaranteed that the direct sound is among the most $\hat{k}$-dominant signal components in the evaluated time window. The first peak of the reconstructed impulse response $\hat{\mathbf{h}}$ provides the transit time of the direct sound when the truncation of the original signal $\mathbf{x}'$ is considered.

4 Analytical Testing of Proposed Method

For the analytical evaluation and testing of the proposed method for the determination of the TOF, their individual steps from section 4 are then tested for signals with pseudo-random impulse responses. On the one hand, this investigation serves to evaluate the applicability of the steps and, on the other hand, it allows a qualitative comparison of the sparse deconvolution algorithms. In the end the Proposed Method is compared to the established methods (GCC, GCCPHAT, AIC) for TOF estimation. Although this analytical investigation does not provide direct insight into the absolute reliability of real GHATT system it does shows which algorithms are suitable for use given certain influencing parameters and boundary conditions.

4.1 Signal Design

For the analytical investigation of the proposed method, the evaluation of signals with known underlying impulse response is useful, so that a direct comparison of the reconstructed results with the true impulse response and the true TOF can be made. For this purpose, the simulated received signals $\mathbf{x}$ are generated according to (6) in such a way that it corresponds approximately to real GHATT signals. Therefore, the transmitted signal $\mathbf{s}$ is designed as linear sweeps according to (14) using a von-Hann window and a frequency spectrum of 500…4000 Hz (Δf = 3500 Hz; f_0 = 2250 Hz), the signal duration T_s = 4 ms and the sampling rate T^{-1} = 192 kHz.

The true impulse response $\mathbf{h}$ has a duration of 4 ms, resulting in a signal duration of the received signal of 8 ms. The impulse response $\mathbf{h}$ is sparse and contains k = 6 positive components (peaks), which are randomly distributed over the duration of the impulse response. This results in a median time interval of 667 µs between two successive peaks. This interval corresponds to only 1.5 times the period of the mean frequency f_0 and therefore leads to a significant temporal overlap of the sweeps in the received signal. To investigate the influence of the amplitudes α_j (5), they are varied according to a unity distribution resembling the inverse law of sound pressure:

$$\alpha_j \sim \mathcal{U}(0, 1/j) \text{ for } j = 1, \dots, k$$

For studies with multiple sweeps, the frequencies of the individual sweeps are selected according to (15). Additive white noise **w** is added, so that corresponding SNR is reached:

$$SNR = 10\log_{10}\left(\frac{P_{\mathbf{S}\,\mathbf{h}}}{P_{\mathbf{w}}}\right)\ \text{dB}$$

with the power of the noise-free signal $P_{\mathbf{S}\,\mathbf{h}} = \|\mathbf{S}\,\mathbf{h}\|_2^2/N$ and the noise power (i.e. noise variance) $P_{\mathbf{w}} = \|\mathbf{w}\|_2^2/N$.

4.2 Time Window Identification via FrFT

For a reliable determination of the time range relevant for the TOF estimation via FrFT, it is necessary that the sweeps clearly raise the magnitude of the FrFT above its noise floor. Figure 4 shows the FrFT for an example signal with SNR = 0 dB. The abscissa is scaled according to (16) in order to project u onto the time t. For comparison, the true underlying impulse response **h** is shown.

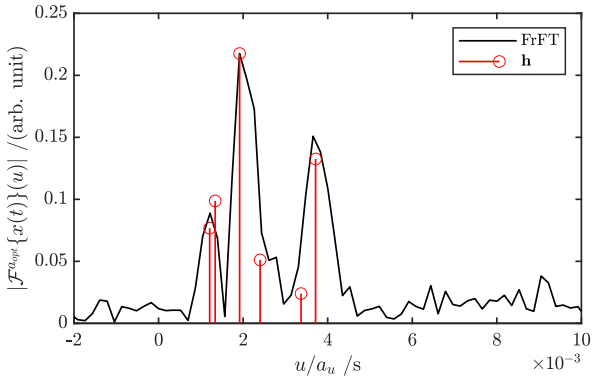

Figure 4. Fractional Fourier Transform of an exemplary signal with SNR = 0 dB, the underlying impulse response of the signal is shown as stems, time has been projected onto the abscissa by the factor a_u

The testing showed that even at a low SNR of 0 dB, the peaks in the FrFT domain can still be clearly distinguished from the noise. Although it is not possible to determine the time precisely, especially for sweeps in close succession, there is nevertheless a high degree of congruence between the impulse response and the results of the FrFT. Thus, the FrFT seems to be very suitable for the determination of the time window for the subsequent steps of the TOF estimation. For this purpose, a threshold $\vartheta = \frac{1}{4}\cdot\max(|\mathcal{F}^{a_{opt}}\{x(t)\}|)$ provided a reliable indicator of the signal reception and thus of the relevant range for the TOF estimation.

4.3 Comparison of Deconvolution Methods

In order to statistically evaluate the various sparse deconvolution algorithms from section 3 for TOF estimation application, 300 pseudorandom signals are generated for each algorithm and for each SNR, and then TOF estimation is performed according to section 4. In the following, the Greedy algorithms NNOMP [41, 42], FNNOMP [43] and sNMF [26] of the family of OMP as well as NNOLS and sNNOLS [41, 42] of the family of OLS and the convex approach NLARS [44] are tested. In particular, the influence of the $\hat{k}$ reconstructed peaks in $\hat{\mathbf{h}}$ is investigated. For this purpose, $\hat{k}$ was varied around the true value $k = 6$.

4.3.1 Single Sweep Signals. First, receiver signals with a single sweep over the entire defined frequency range of 500...4000 Hz were generated and evaluated. Figure 5 shows the computation time of the investigated algorithms as a function of the $\hat{k}$ reconstructed peaks in $\hat{\boldsymbol{h}}$ for receiver signals with a single sweep. Here, it can be seen that the computation time increases with the number of $\hat{k}$ components to be reconstructed for all algorithms. NLARS and NNOLS show the strongest dependence; whereas sNMF suffers the least from the increase of $\hat{k}$. For the signals examined, FNNOMP always provided the fastest calculation for all $\hat{k}$.

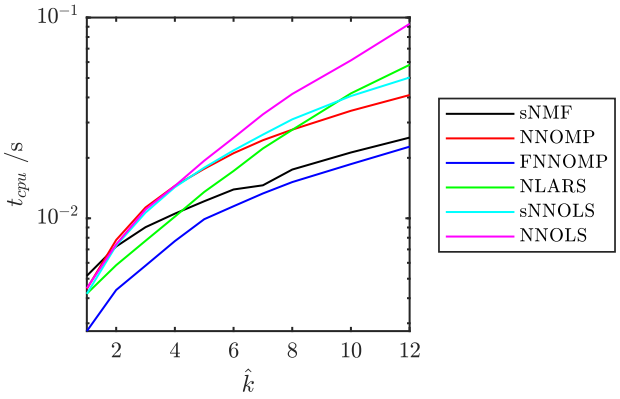

Figure 5. Computation time of the various algorithms as a function of the $\hat{k}$ reconstructed peaks in $\hat{\mathbf{h}}$ for receiver signals with one sweep.

Figure 6 shows the percentage of correctly determined TOF of the various algorithms at various SNRs as a function of the $\hat{k}$ reconstructed peaks in $\hat{\mathbf{h}}$ for receiver signals with a sweep. All estimated TOF are considered to be correct, which deviate at most 2 samples (10.4 µs) from the true TOF. A deviation of an amount this small can be considered uncritical in most cases with GHATT. Besides, it must be considered that the achieved values of the TOF estimation with pseudo-random impulse responses primarily serve the qualitative evaluation of the algorithms among themselves and cannot be transferred directly as absolute values to a real GHATT measurement. Hence, the estimated TOF declared as false mostly occur for receiver signals with very low pseudo-random amplitude α_1 of the first peak. Such low first peaks are not to be expected in real GHATT measurements to this extent. For an in-depth view of the deviation distribution, Figure 7 shows the 50 %, 75 % and 90 % percentiles of the absolute deviations $E_\tau := |\tau - \hat{\tau}|$ of the estimated TOF $\hat{\tau}$ from the true TOF τ.

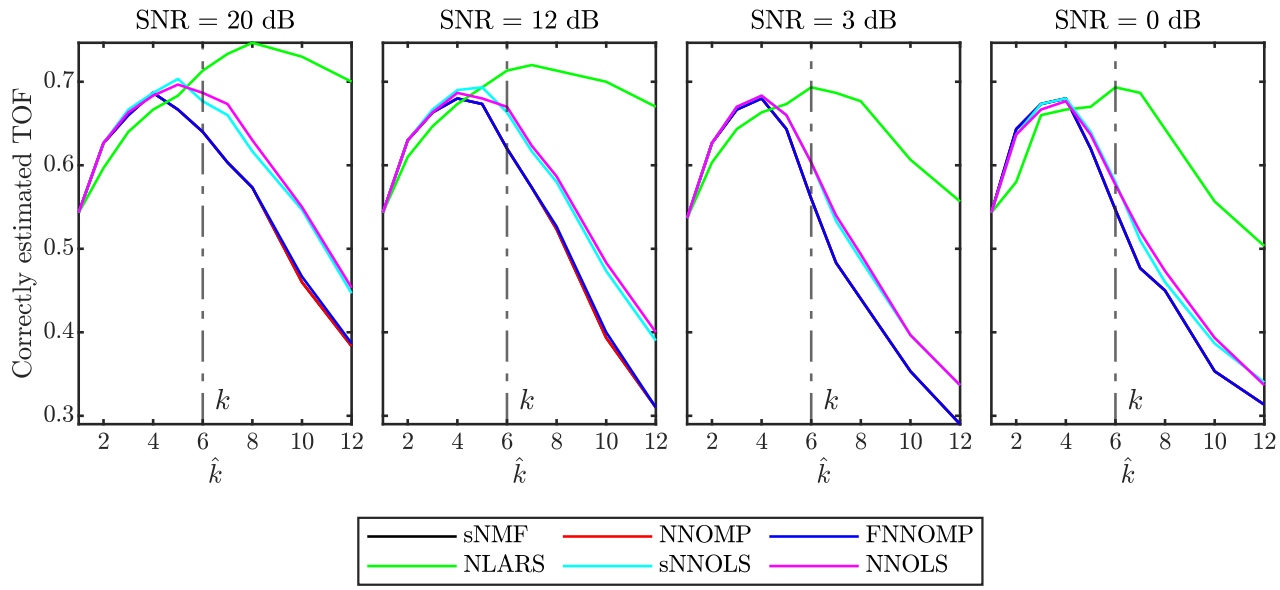

Figure 6. Percentage of correctly determined TOF of the various algorithms at various SNRs as a function of the $\hat{k}$ reconstructed peaks in $\hat{\mathbf{h}}$ for receiver signals with one sweep. k is dash-dotted for reference.

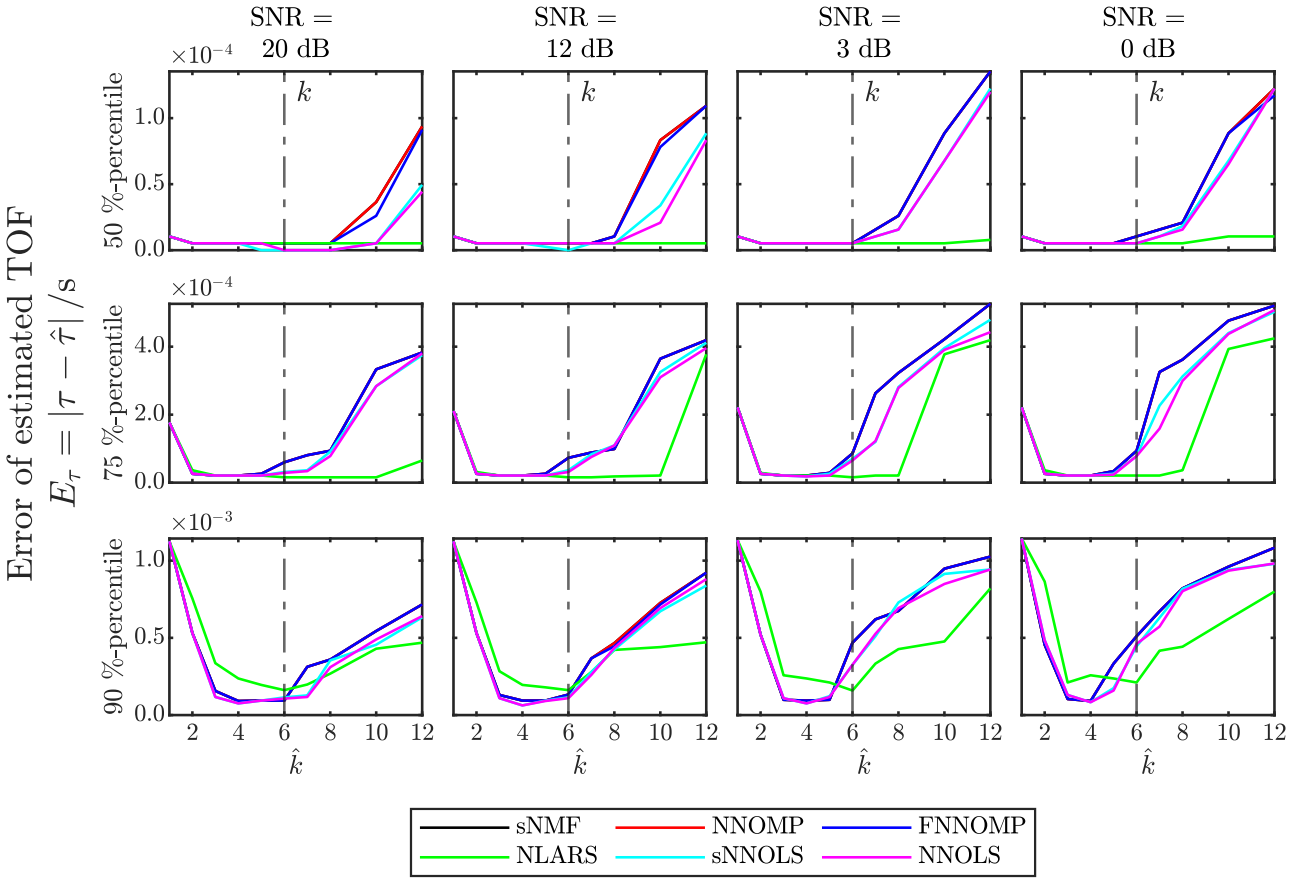

Figure 7. Percentiles of the deviations of the estimated TOF from the true TOF for the various algorithms at various SNRs as a function of the $\hat{k}$ reconstructed peaks in $\hat{\mathbf{h}}$ for receiver signals with one sweep. k is dash-dotted for reference.

The investigation shows that all the OMP approaches (sNMF, NNOMP, FNNOMP) deliver almost identical results, only the calculation times differ. Here, the FNNOMP shows the fastest computation of all OMP algorithms for the short signals with only one sweep. The OLS approaches always deliver slightly better results compared to the OMP

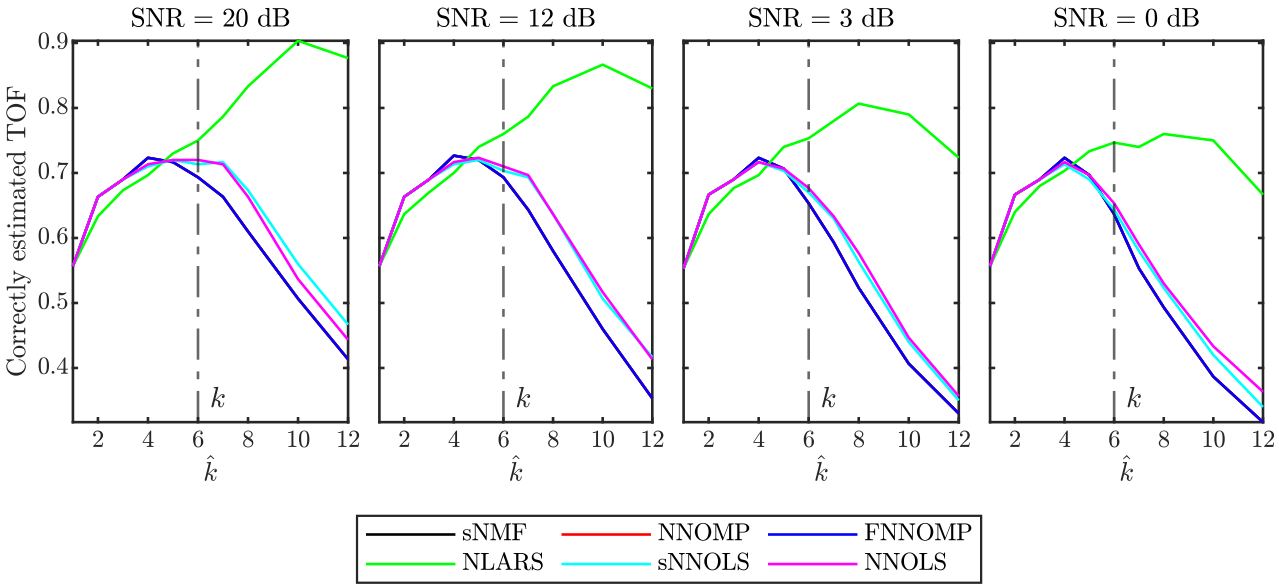

Figure 8. Percentage of correctly determined TOF of the various algorithms at various SNRs as a function of the $\hat{k}$ reconstructed peaks in $\hat{\mathbf{h}}$ for receiver signals with three sweeps. k is dash-dotted for reference.

approaches. However, this comes at the price of a significantly increased computation time. The results of the NLARS algorithm are clearly superior, showing significantly less dependence of the results on $\hat{k}$. This means that with this approach the suboptimal choice of $\hat{k}$ is less critical for a satisfying result. It is noticeable that with the Greedy algorithms the optimum is achieved when $\hat{k}$ slightly underestimates the true value k. In the case of the convex approach, however, the opposite is true. This can be explained by the fact that in the NLARS, similar to the LARS, selected entries $\hat{\mathbf{h}}$ can become zero again in a later iteration step. Here $\hat{k}$ only represents an upper limit, which can also be undershot if sufficient. With the Greedy algorithms, an increased $\hat{k}$ increasingly leads to a false estimation of the TOF. The reason for this is an overfitting of the signal, which results in non-zero entries within the reconstructed impulse response $\hat{\mathbf{h}}$ ahead of the first true peak. The fact that LARS usually requires more iteration steps than OMP algorithms for comparable solution accuracy is known [21, 45] and can also be assumed to apply to the non-negative modifications based on the results. The fact that Greedy algorithms produce less accurate results for highly correlated atoms in the dictionary [45] appears to be also true for its non-negative approaches.

4.3.2 Multiple Sweep Signals. In section 4 it is already shown that the use of multiple sweeps with tuned frequencies according to (15) can disambiguate the autocorrelation of the transmitted signals. Hereafter, it will be examined to what extent this optimized correlation behavior can be transferred to the deconvolution algorithms. For this purpose the receiver signals were generated and evaluated with three sweeps according to (15) with the identical 300 pseudo random impulse responses from above.

The use of three consecutive sweeps triples the total signal duration and thus also increases the computing time for sparse deconvolution. However, the time does not increase equally for all algorithms. Figure 9 shows the computation time of the various algorithms as a function of the $\hat{k}$ reconstructed peaks in $\hat{\mathbf{h}}$ for receiver signals with three sweeps. Compared to the computation time for signals with one sweep, the computation time for signals with three sweeps results in three times for FNNOMP and 2.5 times for NNOLS, sNNOL, NNOMP. For sNMF and NLARS however, the time increases by only 25% and 20%, respectively.

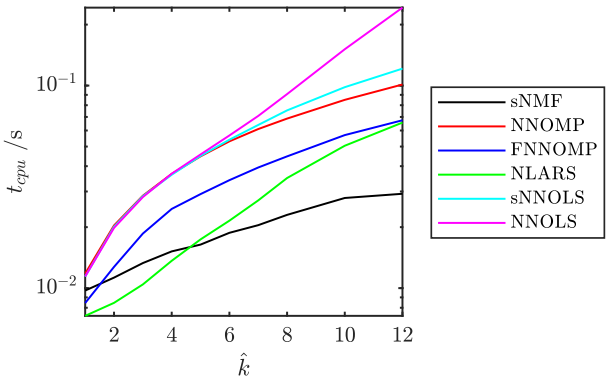

Figure 9. Computation time of the various algorithms as a function of the $\hat{k}$ reconstructed peaks in $\hat{\mathbf{h}}$ for receiver signals with three sweeps.

Figure 8 shows the percentage of correctly determined TOF of the various algorithms at various SNRs as a function of the $\hat{k}$ reconstructed peaks in $\hat{\mathbf{h}}$ for receiver signals with three sweeps. For the Greedy algorithms, a small improvement in

the results of about 5% is shown. Also, the independence of $\hat{k}$ is improved to a minor extent here. For the NLARS, on the other hand, there is a significant improvement in the correctly determined TOF. The maximum value for SNR = 20 dB increases from 75% to over 90%. The independence of $\hat{k}$ is also significantly improved at all SNRs for NLARS.

4.3.3 Summary. In summary, it can be seen that the OMP algorithms sNMF and FNNOMP provide the fastest computation. However, the Greedy algorithms, especially the OMP algorithms, require a precise prior estimation of $\hat{k}$. A misestimation of $\hat{k}$ usually leads to a significant reduction in the quality of the results. The NLARS approach demands significantly longer computation times for signals with one sweep compared to the Greedy algorithms, but also delivers satisfying results over a broader range of $\hat{k}$. Another advantage of the NLARS approach is the possibility to further improve the results by using signals with more than one sweep. It has been shown that the use of three sweeps significantly improves the results for the entire range of $\hat{k}$ and SNR. It is significantly better than those of the Greedy algorithms, without the computation time suffering from the longer signals.

4.4 Comparison to established Time of Flight Methods

In the following, the proposed method is compared to the established methods (GCC, GCCPHAT, AIC) for TOF estimation. For these tests, 300 signals are generated and subsequently evaluated in the same way as in Section 4.1. For the proposed method (with NLARS and $\hat{k}$ = 8), GCC and GCCPHAT, signals with three sweeps according to (15) are used. The AIC is implemented including the adaptive window as introduced by Bao and Jia [15] to increase the robustness to interference from reflections and crosstalk. A sinusoidal burst with 2 µs duration and steep amplitude slope is used as the transmit signal for AIC, as this gave the best results (cf. [46]).

Figure 10 shows the correctly estimated TOF of the investigated methods at different SNR. The proposed method is particularly robust against noise and achieves the highest reliability in TOF estimation. Only with almost noise-free signals (SNR = 96 dB) the estimates of the proposed method could be surpassed by the AIC, which however, like the GCCPHAT, suffers strongly from the influence of the noise. Although the GCC is also very robust against noise, only about two-thirds of the TOF are correctly determined by the GCC compared to the proposed method.

Comparable results show up for the percentiles of the error of estimated TOF in Fig. 11. Here, it can be observed that for GCCPHAT and AIC the error of the estimated TOF increases significantly with increasing noise. GCC and the proposed method work robustly against noise. Only at low noise (SNR > 24 dB) the 90 %-percentile error of the AIC was lower compared to the proposed method.

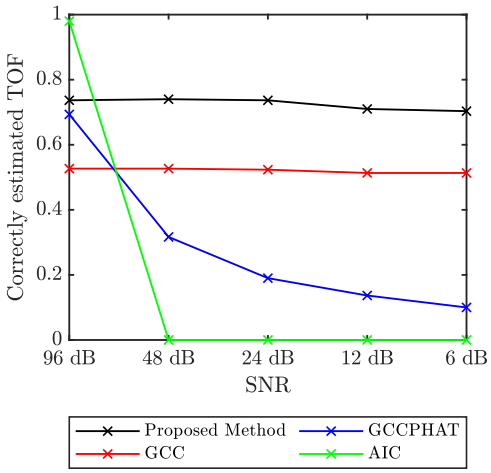

Figure 10. Percentage of correctly determined TOF of proposed method compared to GCC, GCCPHAT and AIC at various SNRs

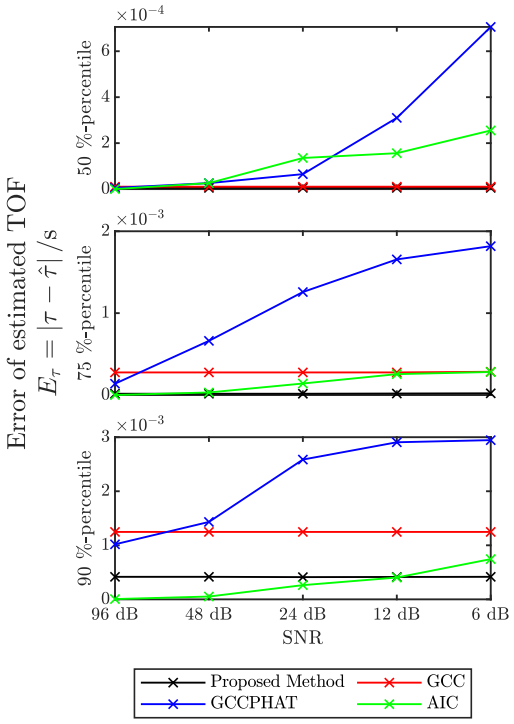

Figure 11. Percentiles of the deviations of the estimated TOF from the true TOF for proposed method compared to GCC, GCCPHAT and AIC at various SNRs

The proposed method appears to be a suitable method for TOF estimation, especially in the case of highly noisy signals with strong reflections. For the simulated GHATT signals, the quality of the estimated TOF by the proposed method clearly exceeds that of established methods.

5 Conclusion

A new method for TOF estimation was presented, which consists of the preparatory steps for data acquisition, signal truncation and dictionary construction, as well as the final step of TOF estimation by means of sparse non-negative deconvolution. The method shows promising possibilities especially for low frequency applications such as GHATT systems. Depending on the room conditions, wall reflections occur, which lead to ambiguous results with conventional methods. The method exploits the non-negativity of reverberant reflections. Furthermore, it allows to consider the system-related distortions of the transducer and the measurement equipment on the basis of prior quasi-anechoic measurements.

Based on the analytical signals, various sparse non-negative algorithms were investigated comparatively. The NLARS algorithm in particular was found to be suitable for practical application. It provided the highest quality of results and is only weakly dependent on the sparsity $\hat{k}$ estimated in advance.

The unambiguity of the autocorrelation of the signals and thus also the quality of the results, in particular of the NLARS, could be improved by the use of multiple sweeps with tuned frequencies, so that a reliable TOF estimation can be performed in moderately transient processes.

For the simulated GHATT signals with superimposed reflections and high noise level, the proposed methods shows a superior quality of estimated TOF compared to established methods.

Furthermore, the findings of this work are to be numerically as well as experimentally investigated for practical GHATT systems, depending on various influencing factors, so that a tool for active investigations with (low-frequency) temporally overlapping signals is created.

Acknowledgements

This work was supported by the German Federal Ministry of Education and Research (16ES0984).

References

[1] Huang Y and Benesty J 2004 *Audio Signal Processing for Next-Generation Multimedia Communication Systems* (Boston, MA: Kluwer Academic Publishers) (eng)

[2] Bleisteiner M, Barth M and Raabe A 2016 Tomographic reconstruction of indoor spatial temperature distributions using room impulse responses *Meas. Sci. Technol.* **27** 35306

[3] Dushaw B 2014 Acoustic Tomography, Ocean *Encyclopedia of Remote Sensing* ed E G Njoku (New York, NY: Springer New York) pp 4–11

[4] Ida N and Meyendorf N eds 2019 *Handbook of Advanced Non-Destructive Evaluation*

[5] Ashfaq M 2007 Measuring and signal processing techniques for ultrasound computed tomography *Dissertation* Ruhr-Universität Bochum, Bochum

[6] Green S F 1985 An acoustic technique for rapid temperature distribution measurement *The Journal of the Acoustical Society of America* **77** 759–63

[7] Arnold K, Ziemann A and Raabe A 1999 Acoustic tomography inside the atmospheric boundary layer *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere* **24** 133–7

[8] Barth M and Raabe A 2011 Acoustic tomographic imaging of temperature and flow fields in air *Meas. Sci. Technol.* **22** 35102

[9] Bao Y, Jia J and Polydorides N 2017 Real-time temperature field measurement based on acoustic tomography *Meas. Sci. Technol.* **28** 74002

[10] Elfering M, Borgmann D, Czajka H, Jantzen H-A, Annas S and Janoske U 2019 Introduction to an acoustic tomography system for the determination of the gas holdup *MikroSystemTechnik Kongress 2019* 1st edn (Berlin: VDE VERLAG) pp 219–21

[11] Elfering M, Czajka H, Annas S and Jantzen H-A 2019 Messanordnung und Verfahren zum Bestimmen einer Gasgehaltsverteilung DE 10 2019 128 230 A1 (Elfering M, Czajka H, Annas S and Jantzen H-A)

[12] Wood A B 1930 *A textbook of sound: Being an account of the physics of vibrations with special reference to recent theoretical and technical developments* 1st edn (The Macmillan company)

[13] Temkin S 2005 *Suspension Acoustics* (Cambridge: Cambridge University Press)

[14] Lu Z, Ma F, Yang C and Chang M 2020 A novel method for Estimating Time of Flight of ultrasonic echoes through short-time Fourier transforms *Ultrasonics* **103** 106104 (eng)

[15] Bao Y and Jia J 2019 Improved time-of-flight estimation method for acoustic tomography system *IEEE Trans. Instrum. Meas.* 1

[16] Lian M, Liu H, Bo Q, Zhang T, Li T and Wang Y 2019 An improved matching pursuit method for overlapping echo separation in ultrasonic thickness measurement *Meas. Sci. Technol.* **30** 65001

[17] Lee R, Kang M-S, Kim B-H, Park K-H, Lee S Q and Park H-M 2020 Sound Source Localization Based on GCC-PHAT With Diffuseness Mask in Noisy and Reverberant Environments *IEEE Access* **8** 7373–82

[18] Fuchs J-J 1999 Multipath time-delay detection and estimation *IEEE Trans. Signal Process.* **47** 237–43

[19] Chen J, Benesty J and Huang Y 2006 Time Delay Estimation in Room Acoustic Environments: An Overview *EURASIP J. Adv. Signal Process.* **2006** 1861

[20] Nguyen T T 2019 Orthogonal greedy algorithms for non-negative sparse reconstruction *PhD Thesis* Université de Lorraine, Nancy

[21] Crespo Marques E, Maciel N, Naviner L, Cai H and Yang J 2019 A Review of Sparse Recovery Algorithms *IEEE Access* **7** 1300–22

[22] Pati Y C, Rezaiifar R and Krishnaprasad P S 1993 Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition *Proceedings of 27th*

Asilomar Conference on Signals, Systems and Computers (27th Asilomar Conference on Signals, Systems and Computers) (Pacific Grove, CA, USA, 1-3 Nov. 1993) (IEEE Comput. Soc. Press) pp 40–4

[23] Chen S, Billings S A and Luo W 1989 Orthogonal least squares methods and their application to non-linear system identification *International Journal of Control* **50** 1873–96

[24] Donoho D L and Tsaig Y 2008 Fast Solution of l1-Norm Minimization Problems When the Solution May Be Sparse *IEEE Trans. Inform. Theory* **54** 4789–812

[25] Lin Y and Lee D D 2006 Bayesian regularization and nonnegative deconvolution for room impulse response estimation *IEEE Trans. Signal Process.* **54** 839–47

[26] Peharz R and Pernkopf F 2012 Sparse nonnegative matrix factorization with ℓ0-constraints *Neurocomputing* **80** 38–46

[27] Mørup M, Madsen K H and Hansen L K 2008 - 2008 Approximate L0 constrained non-negative matrix and tensor factorization *2008 IEEE International Symposium on Circuits and Systems (2008 IEEE International Symposium on Circuits and Systems - ISCAS 2008) (Seattle, WA, USA, 18.05.2008 - 21.05.2008)* (IEEE) pp 1328–31

[28] Bruckstein A M, Elad M and Zibulevsky M 2008 On the Uniqueness of Nonnegative Sparse Solutions to Underdetermined Systems of Equations *IEEE Trans. Inform. Theory* **54** 4813–20

[29] Yaghoobi M, Di Wu and Davies M E 2015 Fast Non-Negative Orthogonal Matching Pursuit *IEEE Signal Process. Lett.* **22** 1229–33

[30] Yaghoobi M and Davies M E 2015 - 2015 Fast non-negative orthogonal least squares *2015 23rd European Signal Processing Conference (EUSIPCO) (2015 23rd European Signal Processing Conference (EUSIPCO)) (Nice, 31.08.2015 - 04.09.2015)* (IEEE) pp 479–83

[31] Sleeman R and van Eck T 1999 Robust automatic P-phase picking: an on-line implementation in the analysis of broadband seismogram recordings *Physics of the Earth and Planetary Interiors* **113** 265–75

[32] Li C, Huang L, Duric N, Zhang H and Rowe C 2009 An improved automatic time-of-flight picker for medical ultrasound tomography *Ultrasonics* **49** 61–72 (eng)

[33] Brandstein M S and Silverman H F 1997 A robust method for speech signal time-delay estimation in reverberant rooms *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing (1997 IEEE International Conference on Acoustics, Speech, and Signal Processing) (Munich, Germany, 21-24 April 1997)* (IEEE Comput. Soc. Press) pp 375–8

[34] Ozaktas H M, Arikan O, Kutay M A and Bozdagt G 1996 Digital computation of the fractional Fourier transform *IEEE Trans. Signal Process.* **44** 2141–50

[35] Lu Y, Kasaeifard A, Oruklu E and Saniie J 2010 - 2010 Performance evaluation of fractional Fourier transform(FrFT) for time-frequency analysis of ultrasonic signals in NDE applications *2010 IEEE International Ultrasonics Symposium (2010 IEEE Ultrasonics Symposium (IUS)) (San Diego, CA, USA, 11.10.2010 - 14.10.2010)* (IEEE) pp 2028–31

[36] Cowell D M J and Freear S 2010 Separation of overlapping linear frequency modulated (LFM) signals using the fractional fourier transform *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* **57** 2324–33 (eng)

[37] Arif M, Cowell D M J and Freear S 2010 Pulse compression of harmonic chirp signals using the fractional fourier transform *Ultrasound in Medicine & Biology* **36** 949–56 (eng)

[38] Carrier A and Got J-L 2014 A maximum a posteriori probability time-delay estimation for seismic signals *Geophysical Journal International* **198** 1543–53

[39] Chang Y, Zi Y, Zhao J, Yang Z, He W and Sun H 2017 An adaptive sparse deconvolution method for distinguishing the overlapping echoes of ultrasonic guided waves for pipeline crack inspection *Meas. Sci. Technol.* **28** 35002

[40] Havelock D, Kuwano S and Vorländer M eds 2009 *Handbook of Signal Processing in Acoustics* (New York, NY: Springer New York) (eng)

[41] Nguyen T T, Idier J, Soussen C and Eh-Hadi Djermoune 2019 *Non-Negative Greedy Algorithms* (Code Ocean) (en)

[42] Nguyen T T, Idier J, Soussen C and Djermoune E-H 2019 Non-Negative Orthogonal Greedy Algorithms *IEEE Trans. Signal Process.* **67** 5643–58

[43] Chatterjee A 2019 *KMSCD Endmember Learning with FNNOMP representation for Hyperspectral Scene Simulators* (Code Ocean) (en)

[44] Mørup M and Madsen K H 2007 *NLARS algorithm*

[45] Hameed M A 2012 Comparative analysis of orthogonal matching pursuit and least angle regression *Master Theses* Michigan State University, Michigan

[46] Espinosa L, Bacca J, Prieto F, Lasaygues P and Brancheriau L 2018 Accuracy on the Time-of-Flight Estimation for Ultrasonic Waves Applied to Non-Destructive Evaluation of Standing Trees: A Comparative Experimental Study *Acta Acustica united with Acustica* **104** 429–39