

On the Single-Sideband Transform for MVDR Beamformers

Vitor Probst Curtarelli^{1,*} , Israel Cohen¹ 

¹ Andrew and Erna Viterbi Faculty of Electrical and Computer Engineering, Technion–Israel Institute of Technology, Technion City, Haifa 3200003, Israel

* Correspondence: vitor.c@campus.technion.ac.il

Abstract: In order to explore different beamforming applications, this paper investigates the application of the Single-Sideband Transform (SSBT) for constructing a Minimum-Variance Distortionless-Response (MVDR) beamformer in the context of the convolutive transfer function (CTF) model for short-window time-frequency transforms by making use of filter-banks and their properties. Our study aims to optimize the appropriate utilization of SSBT in this endeavor, by examining its characteristics and traits. We address a reverberant scenario with multiple noise sources, aiming to minimize both undesired interference and reverberation in the output. Through simulations reflecting real-life scenarios, we show that employing the SSBT correctly leads to a beamformer that outperforms the one obtained when via the Short-Time Fourier Transform (STFT), while exploiting the SSBT's property of it being real-valued. Two approaches were developed with the SSBT, one naive and one refined, with the later being able to ensure the desired distortionless behavior, which is not achieved by the former.

Keywords: Single-sideband transform; MVDR beamformer; Filter-banks; Array signal processing; Signal enhancement.

1. Introduction

Beamformers are an important tool for signal enhancement, being employed in a plethora of applications from hearing aids [1] to source localization [2] to imaging [3,4]. Among the possible ways to use such devices is to implement them the time-frequency domain [5], which allows the exploitation of frequency-related information while also dynamically adapting to signal changes over time. The most widely used instruments for time-frequency analysis are transforms, from which the Short-Time Fourier Transform (STFT) [6,7] stands out in terms of spread and commonness. However, alternative transforms can also be employed implemented [8–10], each offering unique perspective and information regarding the signal, possibly leading to different outputs.

Among these alternatives, the Single-Sideband Transform (SSBT) [11,12] is of great interest, given its real-valued frequency spectrum. It has been shown that the SSBT works particularly well with short analysis windows [11]. Therefore, if we use the convolutive transfer function (CTF) model [13] to study the desired signal model, the SSBT can lend itself to be useful, if we think about the beamforming process through the lenses of filter-banks [14,15]. Thus, by applying this transform within this context it is possible to pull off

Citation: Curtarelli, V. P.; Cohen, I. On the Single-Sideband Transform for MVDR Beamformers. *Algorithms* **2023**, *1*, 0. <https://doi.org/>

Received:
Revised:
Accepted:
Published:

Copyright: © 2024 by the authors. Submitted to *Algorithms* for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

superior performances than only with the STFT. However, it is important to be aware of the limitations of the transform, in order to properly utilize it to try and achieve better outputs.

Two of the most important goals in beamforming are the minimization of noise in the output signal, and the distortionless-ness of the desired signal, both being achieved by the Minimum-Variance Distortionless-Response (MVDR) beamformer [16,17]. As the MVDR beamformer can be used on the time-frequency domain without restrictions on the transform chosen, it is possible to explore and compare the performance of this filter, when designing it through different time-frequency transforms.

Motivated by this, our paper explores the SSB transform and its application on the subject of beamforming within the context of the CTF model. We propose an approach for the CTF that allows the separation of desired and undesired speech components for reverberant environments, and employ this approach for designing the MVDR beamformer. We also explore the traits and limitations of the SSBT, and how to properly adapt the MVDR beamformer to this new transform's constraints. We show that a beamformer designed using the SSBT can surpass the STFT one, while also conforming to the distortionless constraint.

2. Frequency and Time-Frequency Transforms

When studying signals and systems, often frequency and time-frequency transforms are used in order to change the signal domain [18], allowing the exploitation of different patterns and informations inherent to the signal. We from now on assume that all time-domain signals are real-valued.

For continuous time and frequency domains, the Fourier Transform (FT) is defined as

$$\begin{aligned} X_{\mathcal{F}}(f) &\equiv \mathcal{F}\{x(t)\}(f) \\ &= \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt \end{aligned} \quad (1)$$

We define the Real Fourier Transform (RFT) similarly, being cleverly constructed such that its frequency spectrum is real-valued without loss of information, as

$$\begin{aligned} X_{\mathcal{R}}(f) &\equiv \mathcal{R}\{x(t)\}(f) \\ &= \sqrt{2}\mathcal{R}\left\{\int_{-\infty}^{\infty} x(t) e^{-j2\pi ft + j\frac{3\pi}{4}} dt\right\} \\ &= \int_{-\infty}^{\infty} x(t) [-\cos(2\pi ft) + \sin(2\pi ft)] dt \end{aligned} \quad (2)$$

and the Inverse Real Fourier Transform (IRFT) as (see Property A-II)

$$\begin{aligned} x(t) &\equiv \mathcal{R}^{-1}\{X_{\mathcal{R}}(f)\}(t) \\ &= \sqrt{2}\mathcal{R}\left\{\int_{-\infty}^{\infty} X_{\mathcal{R}}(f) e^{j2\pi ft - j\frac{3\pi}{4}} df\right\} \end{aligned} \quad (3)$$

We can also define the RFT in terms of the FT through a simple substitution of Eq. (1) in Eq. (2) (see Property A-I), such that

$$\begin{aligned} X_{\mathcal{R}}(f) &= \sqrt{2}\mathbb{R}\left\{X_{\mathcal{F}}(f)e^{j\frac{3\pi}{4}}\right\} \\ &= -X_{\mathcal{F}}^{\mathbb{R}}(f) - X_{\mathcal{F}}^{\mathbb{I}}(f) \end{aligned} \quad (4)$$

One can also write the RFT in terms of the FT as

$$X_{\mathcal{R}}(f) = \frac{1}{\sqrt{2}}\left(e^{j\frac{3\pi}{4}}X_{\mathcal{F}}(f) + e^{-j\frac{3\pi}{4}}X_{\mathcal{F}}(-f)\right) \quad (5)$$

from which we deduce that

$$X_{\mathcal{F}}(f) = \frac{1}{\sqrt{2}}\left(e^{-j\frac{3\pi}{4}}X_{\mathcal{R}}(f) + e^{j\frac{3\pi}{4}}X_{\mathcal{R}}(-f)\right) \quad (6)$$

2.1. Convolution

Given an impulse response $h(t)$ for a LIT system, the FT the convolution theorem states that

$$h(t) * x(t) \stackrel{\mathcal{F}}{=} H_{\mathcal{F}}(f)X_{\mathcal{F}}(f) \quad (7)$$

where $\stackrel{\mathcal{F}}{=}$ indicates a Fourier transform pair. It can be shown that this theorem isn't strictly valid for the RFT (see Property A-III). That is, if $H_{\mathcal{R}}(f)$ and $X_{\mathcal{R}}(f)$ are the RFT's of $h(t)$ and $x(t)$ respectively, then

$$h(t) * x(t) \not\stackrel{\mathcal{R}}{=} H_{\mathcal{R}}(f)X_{\mathcal{R}}(f) \quad (8)$$

However, it is possible to prove that there is an equivalent of the convolution theorem for the RFT (see Property A-IV), where

$$h(t) * x(t) \stackrel{\mathcal{R}}{=} X_{\mathcal{R}}(f)H_{\mathcal{F}}^{\mathbb{R}}(f) + X_{\mathcal{R}}(-f)H_{\mathcal{F}}^{\mathbb{I}}(f) \quad (9)$$

where, for a given frequency f , the convolution's output on the RFT domain depends on both it and its dual frequency $-f$.

2.2. Discrete time-frequency transforms

Given a time-domain signal $x[n]$, its Short-time Fourier Transform (STFT) [6,7] is

$$X_{\mathcal{F}}[l, k] = \sum_{n=0}^{K-1} w[n]x[n + l \cdot O]e^{-j2\pi k \frac{(n+l \cdot O)}{K}} \quad (10)$$

where $w[n]$ is an analysis window of length K ; and O is the overlap between windows of the transform, usually $O = \lfloor K/2 \rfloor$. The STFT can be seen as a discretization of the FT, while also applying it over different "snippets" of time.

The Single-Sideband Transform (SSBT) [11] is similarly defined, being the RFT's windowed discrete-time adaptation. The SSB transform of $x[n]$ is defined as

$$X_{\mathcal{S}}[l, k] = \sqrt{2}\mathbb{R}\left\{\sum_{n=0}^{K-1} w[n]x[n + l \cdot O]e^{-j2\pi k \frac{(n+l \cdot O)}{K} + j\frac{3\pi}{4}}\right\} \quad (11)$$

One advantage of using the STFT is that we only need to work with $\lfloor (K+1)/2 \rfloor + 1$ frequency bins, given its complex-conjugate behavior. Meanwhile, the SSBT requires all K bins to correctly capture all information of $x[n]$, however it is real-valued.

Assuming that all K bins of the STFT are available, like with Eqs. (4) to (6) we have¹

$$\begin{aligned} X_S[l, k] &= \sqrt{2} \Re \{ X_{\mathcal{F}}[l, k] e^{j\frac{3\pi}{4}} \} \\ &= -\Re \{ X_{\mathcal{F}}[l, k] \} + \Im \{ X_{\mathcal{F}}[l, k] \} \end{aligned} \quad (12)$$

$$X_S[l, k] = \frac{1}{\sqrt{2}} \left(e^{-j\frac{3\pi}{4}} X_{\mathcal{F}}[l, k] + e^{j\frac{3\pi}{4}} X_{\mathcal{F}}[l, K-k] \right) \quad (13)$$

$$X_{\mathcal{F}}[l, k] = \frac{1}{\sqrt{2}} \left(e^{j\frac{3\pi}{4}} X_S[l, k] + e^{-j\frac{3\pi}{4}} X_S[l, K-k] \right) \quad (14)$$

As was the case for the RFT, the SSBT also doesn't hold the convolution theorem the same way as the STFT does. However, similarly to what was shown in Eq. (9), we can write the convolution on the SSBT domain as

$$h[n] * x[n] \stackrel{\mathcal{S}}{\rightleftharpoons} X_S[l, k] H_{\mathcal{F}}^{\mathbf{R}}[k] - X_S[l, K-k] H_{\mathcal{F}}^{\mathbf{I}}[k] \quad (15)$$

or, with the convolutive transfer function (CTF) model [13],

$$h[n] * x[n] \stackrel{\mathcal{S}}{\rightleftharpoons} X_S[l, k] * H_{\mathcal{F}}^{\mathbf{R}}[l, k] - X_S[l, K-k] * H_{\mathcal{F}}^{\mathbf{I}}[l, k] \quad (16)$$

in which this convolution is done over the frames l .

2.3. Relative transfer functions

Given two systems that share an input $x[n]$ each with an impulse response $h_1[n]$ and $h_2[n]$, on the STFT domain $H_1[l, k]$ and $H_2[l, k]$, we can calculate their relative transfer functions (RTF's), respective to a common input. We denote these RTF's $A_1[l, k]$ and $A_2[l, k]$, respective for each system.

Let us denote $Y_1[l, k] = H_1[k] X[l, k]$, under the MTF model, and similarly for $Y_2[l, k]$. With the STFT, we write $X_1[l, k] = H_1[k] X[l, k]$, and thus $Y_1[l, k] = A_1[k] X_1[l, k]$ with $A_1[k] = 1$. We can obtain $A_2[k]$ as

$$A_2[k] = \frac{H_2[k]}{H_1[k]} \quad (17)$$

which trivially satisfies that $A_2[k] X_1[l, k] = H_2[k] X[l, k]$. These RTF's can be calculated as

$$A_m[k] = \frac{\mathbb{E} \{ X_m[l, k] X_1^*[l, k] \}}{\mathbb{E} \{ X_1[l, k] X_1^*[l, k] \}} \quad (18)$$

where $\mathbb{E}\{\cdot\}$ is the expectation operator.

This isn't as straight-forward with the SSBT, since after the convolution each frequency depends on its conjugate as well. However, by considering each system to have two inputs

¹ For the abuse of notation, we let $X_S[l, K] \equiv X_S[l, 0]$, and equally for $X_{\mathcal{F}}[l, K]$.

$X'[l, k] = X[l, k]$ and $X''[l, k] = X[l, K - k]$ and two transfer functions $H'_m[k]$ and $H''_m[k]$ (where m represents the system's index), then our outputs can be described as

$$Y_1[l, k] = H'_1[k]X'[l, k] + H''_1[k]X''[l, k] \quad (19a)$$

$$Y_2[l, k] = H'_2[k]X'[l, k] + H''_2[k]X''[l, k] \quad (19b)$$

From Eq. (15), we easily see that

$$H'_m[k] = H'_m[K - k] = H^{\mathbf{R}}_{\mathcal{F},m}[k] \quad (20a)$$

$$H''_m[k] = -H''_m[K - k] = H^{\mathbf{I}}_{\mathcal{F},m}[k] \quad (20b)$$

We let

$$X'_1[l, k] = H'_1[k]X'[l, k] \quad (21a)$$

$$X''_1[l, k] = H''_1[k]X''[l, k] \quad (21b)$$

and

$$\begin{aligned} X_1[l, k] &= H'_1[k]X'[l, k] + H''_1[k]X''[l, k] \\ &= X'_1[l, k] + X''_1[l, k] \end{aligned} \quad (22)$$

in which $X'_1[l, k]$ and $X''_1[l, k]$ are the inputs processed by the first system, and $X_1[l, k]$ is the observable input signal. Through these, we get that

$$Y_1[l, k] = A'_1[k]X'_1[l, k] + A''_1[k]X''_1[l, k] \quad (23a)$$

$$Y_2[l, k] = A'_2[k]X'_1[l, k] + A''_2[k]X''_1[l, k] \quad (23b)$$

where

$$A^n_m[k] = \frac{H^n_m[k]}{H^n_1[k]} \quad (24)$$

are the RTF's for the n -th input, between the m -th system and the reference (assumed to be $m = 1$). Trivially, $A'_1[k] = A''_1[k] = 1$. Note that, for this, we must be able to estimate each $G^n_m[k]$ separately, which may not be easy. For example, the technique in Eq. (18) used for the STFT isn't applicable here. Using it directly on the observable input $X_m[l, k]$ would yield

$$\frac{\mathbb{E}\{X_m[l, k]X_1^*[l, k]\}}{\mathbb{E}\{X_1[l, k]X_1^*[l, k]\}} = \frac{H'_mH'_1 + H''_mH''_1}{H'^2_1 + H''^2_1} \quad (25)$$

where we used that different frequency bins are independent and all are zero-mean. The exact same would be obtained for $K - k$, thus this tool isn't useful to give us any insight into the desired RTF's of the form from Eq. (24), opposite to with the STFT, where this is possible. Also, since we don't have access to each input $X'_1[l, k]$ and $X''_1[l, k]$ separately, we can't do this practically on each of them, following Eq. (18).

This same formulation can be used with the CTF model. That is, if we consider $A^n_m[l, k]$ as a convolutive gain, then

$$A^n_m[l, k] \approx \frac{H^n_m[l, k]}{H^n_1[0, k]} \quad (26)$$

this being valid for both the STFT (where $n = 1$ is the only option), and for the SSBT. This is an approximation, since different windows in a time-frequency transform aren't independent and thus can contribute to the gain of one-another.

We will from now on assume that we have the RTF's according to Eqs. (23) and (24), allowing the mathematical continuation of the problem.

3. Signal Model and Beamforming

Let there be a device that consists of M sensors and a loudspeaker (LS) in a reverberant environment, in which there also is a desired source, both traveling with a speed c . We also assume the presence of undesired noise at each sensor. For simplicity we assume that all sources are spatially stationary, although this condition can be easily removed.

We denote $Y_{m,k}[l]$ as the signal at the m -th sensor on the time-frequency domain, being represented by

$$Y_{m,k}[l] = X_{m,k}[l] + S_{m,k}[l] + R_{m,k}[l] \quad (27)$$

where $X_{m,k}[l]$ is a desired signal component, $S_{m,k}[l]$ is the undesired loudspeaker signal that is captured by the sensors, and $R_{m,k}[l]$ is uncorrelated white noise present in the sensors. m is the sensor index ($1 \leq m \leq M$), k is the frequency bin index ($0 \leq k < K$), and l is the decimated-time index. We will use a different notation to the one previously used, for ease of reading.

Treating the path and reverberations it takes between the desired source and the m -th sensor as a system, and with the CTF model, in the STFT domain $X_{m,k}[l]$ can be represented as

$$X_{m,k}[l] = A_{m,k}[l] * X'_{1,k}[l] \quad (28)$$

where $X'_{1,k}[l] = H_{1,k}[l] * X_k[l]$ is the desired signal at the reference sensor ($m = 1$), $A_{m,k}[l]$ is the RTF for each sensor, and $H_{1,k}[l]$ is the desired signal's transfer function between the source and the reference sensor. Meanwhile, in the SSBT domain we can represent it as

$$X_{m,k}[l] = A'_{m,k}[l] * X'_{1,k}[l] + A''_{m,k}[l] * X''_{1,k}[l] \quad (29)$$

$$X'_{1,k}[l] = H'_{1,k}[l] * X_k[l] \quad (30a)$$

$$X''_{1,k}[l] = H''_{1,k}[l] * X_{K-k}[l] \quad (30b)$$

where we now have two desired inputs for each sensor, and two RTF's as well. Using Eqs. (20) to (22), we have that $X'_{1,k}[l] \neq X''_{1,k-k}[l]$, even though both originate from $X_k[l]$. Likewise, $X''_{1,k}[l] \neq X'_{1,K-k}[l]$, for the same reason.

Also, since the STFT formulation can also be represented by the SSBT one, with $A'_{m,k}[l] = A_{m,k}[l]$ and $A''_{m,k}[l] = 0$. Therefore, the mathematical developments will be done the SSBT formulation in mind, as it is a less restricting model.

Note that $A_{m,k}[l]$ isn't strictly a causal response, depending on the direction of arrival and features of the reverberant environment, as well as relative delays between the sources at each sensor. We will assume that there are Δ non-causal samples in $A_{m,k}[l]$. It is trivial to see that, for Eq. (29) to be respected, $A'_{1,k}[l] = A''_{1,k}[l] = \delta_{0,l}$, a Kronecker delta at $l = 0$.

We consider the delayed signal $X_{m,k}[l + \lambda]$, such that, when expanding the convolutions, we have

$$X_{m,k}[l + \lambda] = \sum_{\tau} \left(A'_{m,k}[\tau] X'_{1,k}[l + \lambda - \tau] + A''_{m,k}[\tau] X''_{1,k}[l + \lambda - \tau] \right) \quad (31)$$

Now we explicit the contribution of $X''_{1,k}[l]$ on the summation, leading us to

$$\begin{aligned} X_{m,k}[l + \lambda] &= A'_{m,k}[\lambda] X'_{1,k}[l] + A''_{m,k}[\lambda] X''_{1,k}[l] \\ &\quad + \sum_{\tau \neq \lambda} \left(A'_{m,k}[\tau] X'_{1,k}[l + \lambda - \tau] + A''_{m,k}[\tau] X''_{1,k}[l + \lambda - \tau] \right) \\ &= A'_{m,k}[\lambda] X'_{1,k}[l] + A''_{m,k}[\lambda] X''_{1,k}[l] + Q_{m,k}[l + \lambda] \end{aligned} \quad (32)$$

where the first two terms are the contributions of the desired signals at the time of interest l , and $Q_{m,k}[l + \lambda]$ are the remaining terms of the convolution, which can be regarded as only reverberation. Using this on Eq. (27) with Eq. (29), we have that

$$Y_{m,k}[l + \lambda] = A'_{m,k}[\lambda]X'_{1,k}[l] + A''_{m,k}[\lambda]X''_{1,k}[l] + Q_{m,k}[l + \lambda] + S_{m,k}[l + \lambda] + R_{m,k}[l + \lambda] \quad (33)$$

We consider L_Y previous and \bar{L}_Y future samples of $Y_{m,k}[l]$ that are influenced by $X'_{1,k}[l]$ (that is, $A'_{m,k}[\lambda] \neq 0$ or $A''_{m,k}[\lambda] \neq 0$), and define $L = L_Y + \bar{L}_Y + 1$. With this, we get

$$\mathbf{Y}_{m,k}[l] = \mathbf{A}'_{m,k}X'_{1,k}[l] + \mathbf{A}''_{m,k}X''_{1,k}[l] + \mathbf{Q}_{m,k}[l] + \mathbf{S}_{m,k}[l] + \mathbf{R}_{m,k}[l] \quad (34)$$

in which

$$\mathbf{Y}_{m,k}[l] = \left[Y_{m,k}[l + \bar{L}_Y], \dots, Y_{m,k}[l], \dots, Y_{m,k}[l - L_Y] \right]^T \quad (35)$$

and similarly for all other signal vectors, and

$$\mathbf{A}_{m,k}^n = \left[A_{m,k}[L_Y], \dots, A_{m,k}[0], \dots, A_{m,k}[-L_Y] \right]^T \quad (36)$$

Now stacking all sensors into a vector, we get

$$\mathbf{y}_k[l] = \mathbf{a}'_k X'_{1,k}[l] + \mathbf{a}''_k X''_{1,k}[l] + \mathbf{q}_k[l] + \mathbf{s}_k[l] + \mathbf{r}_k[l] \quad (37)$$

with

$$\mathbf{y}_k[l] = \left[\mathbf{Y}_{1,k}^T[l], \dots, \mathbf{Y}_{M,k}^T[l] \right]^T \quad (38)$$

and the same for \mathbf{a}_k^n , $\mathbf{q}_k[l]$, $\mathbf{s}_k[l]$ and $\mathbf{r}_k[l]$, with them all being $ML \times 1$ vectors.

3.1. Filtering and the MPDR beamformer

We want to recover the desired signal at the reference sensor, $X_{1,k}[l] = X'_{1,k}[l] + X''_{1,k}[l]$ (see Eq. (29) with $m = 1$), without any distortion. For this, a linear filter $\mathbf{f}_k[l]$ will be employed, producing an estimate $Z_k[l]$ of our desired signal, such that

$$\begin{aligned} Z_k^n[l] &\approx X_{1,k}[l] \\ &= \mathbf{f}_k^H[l] \mathbf{y}_k[l] \end{aligned} \quad (39)$$

with $(\cdot)^H$ being the transposed-complex-conjugate operator. This process can also be interpreted as

$$\begin{aligned} Z_k[l] &= \sum_m \bar{\mathbf{f}}_{m,k}^H[l] \bar{\mathbf{y}}_{m,k}[l] \\ &= \sum_m F_{m,k}^*[l] * Y_{m,k}[l] \end{aligned} \quad (40)$$

where $\bar{\mathbf{f}}_{m,k}[l]$ is the $L_Y \times 1$ part of $\mathbf{f}_k[l]$ that filters the m -th sensor, and $F_{m,k}[l]$ is its signal-form counterpart. In this sense, the filtering process can be interpreted as the sum across all sensors of the convolution between the signal and the observations.

Going back to Eq. (39), with Eq. (37) we can write

$$Z_k[l] = \mathbf{f}_k^H[l] \mathbf{a}'_k X'_{1,k}[l] + \mathbf{f}_k^H[l] \mathbf{a}''_k X''_{1,k}[l] + \mathbf{f}_k^H[l] \mathbf{q}_k[l] + \mathbf{f}_k^H[l] \mathbf{s}_k[l] + \mathbf{f}_k^H[l] \mathbf{r}_k[l] \quad (41)$$

From this, we easily see that to achieve a distortionless response from the desired signal, we must have that

$$\mathbf{f}_k^H[l] \mathbf{a}'_k = 1 \quad (42a)$$

$$\mathbf{f}_k^H[l] \mathbf{a}''_k = 1 \quad (42b)$$

which will ensure that both components of the desired signal are undistorted. For the STFT, only the first constraint is considered, since in this case we have that $\mathbf{a}''_k = \mathbf{0}$ and thus the second condition is impossible. With this, we write our constraint matrix as

$$\mathbf{f}_k^H[l] \mathbf{C}_k = \mathbf{i}^T \quad (43)$$

where, for the STFT, $\mathbf{C}_k = \mathbf{a}_k$ and $\mathbf{i} = 1$; and, for the SSBT, we have

$$\mathbf{C}_k = \begin{bmatrix} \mathbf{a}'_k & \mathbf{a}''_k \end{bmatrix}_{ML \times 2} \quad (44a)$$

$$\mathbf{i} = \begin{bmatrix} 1 & 1 \end{bmatrix} \quad (44b)$$

To minimize the variance of the output signal while obeying the distortionless constraint, a Minimum-Power Distortionless Response (MPDR) beamformer will be used, it being defined as

$$\mathbf{f}_{\text{mpdr};k}[l] = \min_{\mathbf{f}_k[l]} \mathbf{f}_k^H[l] \mathbf{\Phi}_{y_k}[l] \mathbf{f}_k[l] \text{ s.t. } \mathbf{f}_k^H[l] \mathbf{C}_k = \mathbf{i}^T \quad (45)$$

where $\mathbf{\Phi}_{y_k}[l]$ is the correlation matrix of the observed signal $\mathbf{y}_k[l]$. The solution to this minimization problem

$$\mathbf{f}_{\text{mpdr};k}[l] = \mathbf{\Phi}_{y_k}^{-1}[l] \mathbf{C}_k \left[\mathbf{C}_k^H \mathbf{\Phi}_{y_k}^{-1}[l] \mathbf{A}_k \right]^{-1} \mathbf{i} \quad (46)$$

Author Contributions: Conceptualization, I. Cohen and V. Curtarelli; Methodology, V. Curtarelli; Software, V. Curtarelli; Writing—original draft: V. Curtarelli; Writing—review and editing, I. Cohen and V. Curtarelli; Supervision, V. Curtarelli. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Pazy Research Foundation, and the Israel Science Foundation (grant no. 1449/23).

Data Availability Statement: The source-code for the simulations developed here is available at <https://github.com/VCurtarelli/py-ssb-ctf-bf>.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CTF	Convulsive Transfer Function
DSRF	Desired Signal Reduction Factor
MPDR	Minimum-Power Distortionless-Response
MTF	Multiplicative Transfer Function
SNR	Signal-to-Noise Ratio
SSBT	Single-Sideband Transform
STFT	Short-Time Fourier Transform

Appendix A. Properties of the Real Fourier Transform

200

$$X_{\mathcal{F}}(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt \quad (\text{A.1})$$

$$X_{\mathcal{R}}(f) = \sqrt{2}\mathbb{R}\left\{\int_{-\infty}^{\infty} x(t) e^{-j2\pi ft + j\frac{3\pi}{4}} dt\right\} \quad (\text{A.2})$$

$$x(t) = \sqrt{2}\mathbb{R}\left\{\int_{-\infty}^{\infty} X_{\mathcal{R}}(f) e^{j2\pi ft - j\frac{3\pi}{4}} df\right\} \quad (\text{A.3})$$

Property A-I: The FT and RFT can be obtained from one another.

201

Proof: By manipulating Eq. (A.2) we can get

202

$$X_{\mathcal{R}}(f) = \sqrt{2}\mathbb{R}\left\{\left(\int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt\right) e^{j\frac{3\pi}{4}}\right\} \quad (\text{i.1})$$

where the term in parenthesis is trivially the FT of $x(t)$ from Eq. (A.1). Thus, we get that

203

$$\begin{aligned} X_{\mathcal{R}}(f) &= \mathbb{R}\left\{\left(X_{\mathcal{F}}^{\mathcal{R}}(f) + jX_{\mathcal{F}}^{\mathcal{I}}(f)\right) \cdot (-1 + j)\right\} \\ &= -X_{\mathcal{F}}^{\mathcal{R}}(f) - X_{\mathcal{F}}^{\mathcal{I}}(f) \end{aligned} \quad (\text{i.2})$$

Likewise, using that the FT of a real signal is complex-conjugate, such that $X_{\mathcal{F}}(f) = X_{\mathcal{F}}^*(-f)$, it is easy to see that

204

$$X_{\mathcal{R}}(-f) = -X_{\mathcal{F}}^{\mathcal{R}}(f) + X_{\mathcal{F}}^{\mathcal{I}}(f) \quad (\text{i.3})$$

205

With this, we have that

206

$$\begin{aligned} \sqrt{2}X_{\mathcal{R}}(f) e^{-j\frac{3\pi}{4}} &= X_{\mathcal{F}}^{\mathcal{R}}(f) + X_{\mathcal{F}}^{\mathcal{I}}(f) + jX_{\mathcal{F}}^{\mathcal{R}}(f) + jX_{\mathcal{F}}^{\mathcal{I}}(f) \\ \sqrt{2}X_{\mathcal{R}}(-f) e^{j\frac{3\pi}{4}} &= X_{\mathcal{F}}^{\mathcal{R}}(f) - X_{\mathcal{F}}^{\mathcal{I}}(f) - jX_{\mathcal{F}}^{\mathcal{R}}(f) + jX_{\mathcal{F}}^{\mathcal{I}}(f) \end{aligned} \quad (\text{i.4})$$

and therefore

207

$$\begin{aligned} \frac{X_{\mathcal{R}}(f) e^{-j\frac{3\pi}{4}} + X_{\mathcal{R}}(-f) e^{j\frac{3\pi}{4}}}{\sqrt{2}} &= X_{\mathcal{F}}^{\mathcal{R}}(f) + jX_{\mathcal{F}}^{\mathcal{I}}(f) \\ &= X_{\mathcal{F}}(f) \end{aligned} \quad (\text{i.5})$$

Similarly, using Eq. (i.1) we have that

208

$$X_{\mathcal{R}}(f) = \sqrt{2}\mathbb{R}\left\{X_{\mathcal{F}}(f) e^{j\frac{3\pi}{4}}\right\} \quad (\text{i.6})$$

Using the property of complex numbers that $\mathbb{R}\{a\} = \frac{a+a^*}{2}$, then

209

$$X_{\mathcal{R}}(f) = \frac{X_{\mathcal{F}}(f) e^{j\frac{3\pi}{4}} + X_{\mathcal{F}}(-f) e^{-j\frac{3\pi}{4}}}{\sqrt{2}} \quad (\text{i.7})$$

210

Property A-II: The IFFT is the inverse of the RFT.

211

Proof: From Property A-I, we have that

$$\mathcal{R}^{-1}\{X_{\mathcal{R}}(f)\}(t) = \frac{X_{\mathcal{F}}(f)e^{j\frac{3\pi}{4}} + X_{\mathcal{F}}(-f)e^{-j\frac{3\pi}{4}}}{\sqrt{2}} \quad (\text{ii.1})$$

Substituting this in Eq. (A.3), we have

$$\begin{aligned} \mathcal{R}^{-1}\{X_{\mathcal{R}}(f)\}(t) &= \sqrt{2}\mathbb{R}\left\{\int_{-\infty}^{\infty}\left(\frac{X_{\mathcal{F}}(f)e^{j\frac{3\pi}{4}} + X_{\mathcal{F}}(-f)e^{-j\frac{3\pi}{4}}}{\sqrt{2}}\right)e^{j2\pi ft - j\frac{3\pi}{4}}df\right\} \\ &= \mathbb{R}\left\{\int_{-\infty}^{\infty}\left(X_{\mathcal{F}}(f)e^{j2\pi ft} + X_{\mathcal{F}}(-f)e^{j2\pi ft - j\frac{3\pi}{2}}\right)df\right\} \end{aligned} \quad (\text{ii.2})$$

The first term expands to the inverse Fourier transform of $X_{\mathcal{F}}(f)$, which is trivially $x(t)$; the second term is the inverse Fourier transform of $X_{\mathcal{F}}(-f)$, which is from the time reversal property is $x(-t)$. Therefore,

$$\begin{aligned} \mathcal{R}^{-1}\{X_{\mathcal{R}}(f)\}(t) &= \mathbb{R}\left\{x(t) + x(-t)e^{-j\frac{3\pi}{2}}\right\} \\ &= \mathbb{R}\{x(t) + jx(-t)\} \\ &= x(t) \end{aligned} \quad (\text{ii.3})$$

thus concluding the proof.

Property A-III: The convolution theorem doesn't apply for the RFT.

Proof: Let $h(t)$ be the impulse response of an LIT system, with input $x(t)$. It is trivial that the system's output, $y(t)$, is given by

$$y(t) = h(t) * x(t) \quad (\text{iii.1})$$

with $*$ being the convolution operator. For the Fourier transform, through the convolution theorem it is trivial that

$$Y_{\mathcal{F}}(f) = H_{\mathcal{F}}(f)X_{\mathcal{F}}(f) \quad (\text{iii.2})$$

Expanding these in terms of real and imaginary parts (omitting the frequency index for clarity),

$$Y_{\mathcal{F}} = H_{\mathcal{F}}^{\mathcal{R}}X_{\mathcal{F}}^{\mathcal{R}} + jH_{\mathcal{F}}^{\mathcal{R}}X_{\mathcal{F}}^{\mathcal{I}} + jH_{\mathcal{F}}^{\mathcal{I}}X_{\mathcal{F}}^{\mathcal{R}} - H_{\mathcal{F}}^{\mathcal{I}}X_{\mathcal{F}}^{\mathcal{I}} \quad (\text{iii.3})$$

Now in the RFT domain, with Eq. (i.2) we have that

$$\begin{aligned} X_{\mathcal{R}}(f) &= -X_{\mathcal{F}}^{\mathcal{R}}(f) - X_{\mathcal{F}}^{\mathcal{I}}(f) \\ H_{\mathcal{R}}(f) &= -H_{\mathcal{F}}^{\mathcal{R}}(f) - H_{\mathcal{F}}^{\mathcal{I}}(f) \end{aligned} \quad (\text{iii.4})$$

Assuming the convolution theorem true for the RFT,

$$Y_{\mathcal{R}} = H_{\mathcal{F}}^{\mathcal{R}}X_{\mathcal{F}}^{\mathcal{R}} + H_{\mathcal{F}}^{\mathcal{R}}X_{\mathcal{F}}^{\mathcal{I}} + H_{\mathcal{F}}^{\mathcal{I}}X_{\mathcal{F}}^{\mathcal{R}} + H_{\mathcal{F}}^{\mathcal{I}}X_{\mathcal{F}}^{\mathcal{I}} \quad (\text{iii.5})$$

Now, by applying Eq. (i.2) on Eq. (iii.3), we have

$$\tilde{Y}_{\mathcal{R}} = -H_{\mathcal{F}}^{\mathcal{R}}X_{\mathcal{F}}^{\mathcal{R}} - H_{\mathcal{F}}^{\mathcal{R}}X_{\mathcal{F}}^{\mathcal{I}} - H_{\mathcal{F}}^{\mathcal{I}}X_{\mathcal{F}}^{\mathcal{R}} + H_{\mathcal{F}}^{\mathcal{I}}X_{\mathcal{F}}^{\mathcal{I}} \quad (\text{iii.6})$$

where it is explicit that $Y_{\mathcal{R}}(f) \neq \tilde{Y}_{\mathcal{R}}(f)$. Therefore, the RFT of the convolution (Eq. (iii.6)) is not the product of the RFT's of the signals (Eq. (iii.5)), and thus the convolution theorem doesn't hold for the RFT.

Property A-IV: *There is an equivalent of the convolution theorem for the RFT.*

Proof: From Eq. (iii.6), we have our objective for the “convolution theorem”-equivalent for the RFT. From both Eqs. (i.2) and (i.3), we have

$$\begin{aligned} X_{\mathcal{R}}(f) &= -X_{\mathcal{F}}^{\mathcal{R}}(f) - X_{\mathcal{F}}^{\mathcal{I}}(f) \\ X_{\mathcal{R}}(-f) &= -X_{\mathcal{F}}^{\mathcal{R}}(f) + X_{\mathcal{F}}^{\mathcal{I}}(f) \\ H_{\mathcal{R}}(f) &= -H_{\mathcal{F}}^{\mathcal{R}}(f) - H_{\mathcal{F}}^{\mathcal{I}}(f) \\ H_{\mathcal{R}}(-f) &= -H_{\mathcal{F}}^{\mathcal{R}}(f) + H_{\mathcal{F}}^{\mathcal{I}}(f) \end{aligned} \quad (\text{iv.1})$$

We will omit the frequency dependency in the FT values. Taking the possible combinations, we have

$$X_{\mathcal{R}}(f)H_{\mathcal{R}}(f) = H_{\mathcal{F}}^{\mathcal{R}}X_{\mathcal{F}}^{\mathcal{R}} + H_{\mathcal{F}}^{\mathcal{R}}X_{\mathcal{F}}^{\mathcal{I}} + H_{\mathcal{F}}^{\mathcal{I}}X_{\mathcal{F}}^{\mathcal{R}} + H_{\mathcal{F}}^{\mathcal{I}}X_{\mathcal{F}}^{\mathcal{I}} \quad (\text{iv.2a})$$

$$X_{\mathcal{R}}(f)H_{\mathcal{R}}(-f) = H_{\mathcal{F}}^{\mathcal{R}}X_{\mathcal{F}}^{\mathcal{R}} + H_{\mathcal{F}}^{\mathcal{R}}X_{\mathcal{F}}^{\mathcal{I}} - H_{\mathcal{F}}^{\mathcal{I}}X_{\mathcal{F}}^{\mathcal{R}} - H_{\mathcal{F}}^{\mathcal{I}}X_{\mathcal{F}}^{\mathcal{I}} \quad (\text{iv.2b})$$

$$X_{\mathcal{R}}(-f)H_{\mathcal{R}}(f) = H_{\mathcal{F}}^{\mathcal{R}}X_{\mathcal{F}}^{\mathcal{R}} - H_{\mathcal{F}}^{\mathcal{R}}X_{\mathcal{F}}^{\mathcal{I}} + H_{\mathcal{F}}^{\mathcal{I}}X_{\mathcal{F}}^{\mathcal{R}} - H_{\mathcal{F}}^{\mathcal{I}}X_{\mathcal{F}}^{\mathcal{I}} \quad (\text{iv.2c})$$

$$X_{\mathcal{R}}(-f)H_{\mathcal{R}}(-f) = H_{\mathcal{F}}^{\mathcal{R}}X_{\mathcal{F}}^{\mathcal{R}} - H_{\mathcal{F}}^{\mathcal{R}}X_{\mathcal{F}}^{\mathcal{I}} - H_{\mathcal{F}}^{\mathcal{I}}X_{\mathcal{F}}^{\mathcal{R}} + H_{\mathcal{F}}^{\mathcal{I}}X_{\mathcal{F}}^{\mathcal{I}} \quad (\text{iv.2d})$$

Taking the difference between Eq. (iv.2a) and Eq. (iv.2d), and the sum of Eq. (iv.2b) and Eq. (iv.2c), we have

$$X_{\mathcal{R}}(f)H_{\mathcal{R}}(f) - X_{\mathcal{R}}(-f)H_{\mathcal{R}}(-f) = 2\left(H_{\mathcal{F}}^{\mathcal{R}}X_{\mathcal{F}}^{\mathcal{I}} + H_{\mathcal{F}}^{\mathcal{I}}X_{\mathcal{F}}^{\mathcal{R}}\right) \quad (\text{iv.3a})$$

$$X_{\mathcal{R}}(f)H_{\mathcal{R}}(-f) + X_{\mathcal{R}}(-f)H_{\mathcal{R}}(f) = 2\left(H_{\mathcal{F}}^{\mathcal{R}}X_{\mathcal{F}}^{\mathcal{R}} - H_{\mathcal{F}}^{\mathcal{I}}X_{\mathcal{F}}^{\mathcal{I}}\right) \quad (\text{iv.3b})$$

and therefore, to achieve Eq. (iii.6), we let

$$\begin{aligned} Y_{\mathcal{R}}(f) &= \frac{-X_{\mathcal{R}}(f)H_{\mathcal{R}}(f) + X_{\mathcal{R}}(-f)H_{\mathcal{R}}(-f) - X_{\mathcal{R}}(f)H_{\mathcal{R}}(-f) - X_{\mathcal{R}}(-f)H_{\mathcal{R}}(f)}{2} \\ &= X_{\mathcal{R}}(f)\frac{-H_{\mathcal{R}}(f) - H_{\mathcal{R}}(-f)}{2} + X_{\mathcal{R}}(-f)\frac{-H_{\mathcal{R}}(f) + H_{\mathcal{R}}(-f)}{2} \end{aligned} \quad (\text{iv.4})$$

Finally, from Eq. (iv.1), we achieve

$$Y_{\mathcal{R}}(f) = X_{\mathcal{R}}(f)H_{\mathcal{F}}^{\mathcal{R}}(f) + X_{\mathcal{R}}(-f)H_{\mathcal{F}}^{\mathcal{I}}(f) \quad (\text{iv.5})$$

and, for its conjugate frequency (that is, for $-f$), we have

$$\begin{aligned} Y_{\mathcal{R}}(-f) &= X_{\mathcal{R}}(-f)H_{\mathcal{F}}^{\mathcal{R}}(-f) + X_{\mathcal{R}}(f)H_{\mathcal{F}}^{\mathcal{I}}(-f) \\ &= X_{\mathcal{R}}(-f)H_{\mathcal{F}}^{\mathcal{R}}(f) - X_{\mathcal{R}}(f)H_{\mathcal{F}}^{\mathcal{I}}(f) \end{aligned} \quad (\text{iv.6})$$

Property A-V: *Frequencies in the RFT have the same variance as their FT counterpart.*

Proof: We now assume that $X_{\mathcal{F}}(f)$ is the transform of a random process, such that its real and imaginary parts are independent and identically distributed with zero mean. Taking the complex correlation of a given frequency in the FT domain,

$$\mathbb{E}\{X_{\mathcal{F}}(f)X_{\mathcal{F}}^*(f)\} = \mathbb{E}\{X_{\mathcal{F}}^{\mathcal{R}}(f)^2 + X_{\mathcal{F}}^{\mathcal{I}}(f)^2\} \quad (\text{v.1})$$

Since they are identically distributed, we denote

$$\mathbb{E}\{X_{\mathcal{F}}^{\mathcal{R}}(f)^2\} = \mathbb{E}\{X_{\mathcal{F}}^{\mathcal{I}}(f)^2\} = \sigma_f^2 \quad (\text{v.2})$$

and therefore

$$\mathbb{E}\{X_{\mathcal{F}}(f)X_{\mathcal{F}}^*(f)\} = 2\sigma_f^2 \quad (\text{v.3})$$

Now in the RFT domain, we take the correlation of a given frequency Eq. (i.2),

$$\mathbb{E}\{X_{\mathcal{R}}(f)^2\} = \mathbb{E}\{X_{\mathcal{F}}^{\mathcal{R}}(f)^2 + 2X_{\mathcal{F}}^{\mathcal{R}}(f)X_{\mathcal{F}}^{\mathcal{I}}(f) + X_{\mathcal{F}}^{\mathcal{I}}(f)^2\} \quad (\text{v.4})$$

Using that $X_{\mathcal{F}}^{\mathcal{R}}(f)$ and $X_{\mathcal{F}}^{\mathcal{I}}(f)$ are i.i.d. and zero-mean, the cross terms are zero, and thus

$$\begin{aligned} \mathbb{E}\{X_{\mathcal{R}}(f)^2\} &= 2\sigma_f^2 \\ &= \mathbb{E}\{X_{\mathcal{F}}(f)X_{\mathcal{F}}^*(f)\} \end{aligned} \quad (\text{v.5})$$

It is trivial to see that the same applies for $X_{\mathcal{R}}(-f)$.

Property A-VI: Conjugate frequencies in the RFT domain are independent.

Proof: We take the same assumptions as those in Property A-V. Taking the complex correlation between the two conjugate frequencies,

$$\mathbb{E}\{X_{\mathcal{F}}(f)X_{\mathcal{F}}^*(-f)\} = \mathbb{E}\{X_{\mathcal{F}}^{\mathcal{R}}(f)^2 + 2jX_{\mathcal{F}}^{\mathcal{R}}(f)X_{\mathcal{F}}^{\mathcal{I}}(f) - X_{\mathcal{F}}^{\mathcal{I}}(f)^2\} \quad (\text{vi.1})$$

Using that the $X_{\mathcal{F}}^{\mathcal{R}}(f)$ and $X_{\mathcal{F}}^{\mathcal{I}}(f)$ are independent and zero-mean, the cross terms are zero, and with Eq. (v.2) then

$$\mathbb{E}\{X_{\mathcal{F}}(f)X_{\mathcal{F}}^*(-f)\} = 0 \quad (\text{vi.2})$$

This result is known, but nonetheless it is useful to show it, since this same procedure will be used for the RFT.

We now consider Eqs. (i.2) and (i.3). Taking the correlation between the two conjugate frequencies yields

$$\mathbb{E}\{X_{\mathcal{R}}(f)X_{\mathcal{R}}(-f)\} = \mathbb{E}\{X_{\mathcal{F}}^{\mathcal{R}}(f)^2 - X_{\mathcal{F}}^{\mathcal{I}}(f)^2\} \quad (\text{vi.3})$$

Under the same assumptions that the real and imaginary parts of $X_{\mathcal{F}}(f)$ are identically distributed, we get the same result as before, where

$$\mathbb{E}\{X_{\mathcal{R}}(f)X_{\mathcal{R}}(-f)\} = 0 \quad (\text{vi.4})$$

Note that, with the RFT, we didn't use the complex correlation, since it is real-valued.

Lastly, we take the correlation between conjugate frequencies of the output of a system according to Eq. (iv.5),

$$\begin{aligned} \mathbb{E}\{Y_{\mathcal{R}}(f)Y_{\mathcal{R}}(-f)\} &= H_{\mathcal{F}}^{\mathcal{R}}(f)^2\mathbb{E}\{X_{\mathcal{R}}(f)X_{\mathcal{R}}(-f)\} - H_{\mathcal{F}}^{\mathcal{R}}(f)H_{\mathcal{F}}^{\mathcal{I}}(f)\mathbb{E}\{X_{\mathcal{R}}(f)^2\} \\ &\quad + H_{\mathcal{F}}^{\mathcal{R}}(f)H_{\mathcal{F}}^{\mathcal{I}}(f)\mathbb{E}\{X_{\mathcal{R}}(-f)^2\} + H_{\mathcal{F}}^{\mathcal{I}}(f)^2\mathbb{E}\{X_{\mathcal{R}}(f)X_{\mathcal{R}}(-f)\} \end{aligned} \quad (\text{vi.5})$$

Since $X_{\mathcal{R}}(f)$ and $X_{\mathcal{R}}(-f)$ are independent and zero-mean, the first and last terms are zero, and also the other two cancel each other out since $\mathbb{E}\{X_{\mathcal{R}}(f)^2\} = \mathbb{E}\{X_{\mathcal{R}}(-f)^2\} = 2\sigma_f^2$. Therefore,

$$\mathbb{E}\{Y_{\mathcal{R}}(f)Y_{\mathcal{R}}(-f)\} = 0 \quad (\text{vi.6})$$

References

1. Lobato, W.; Costa, M.H. Worst-Case-Optimization Robust-MVDR Beamformer for Stereo Noise Reduction in Hearing Aids. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* **2020**, *28*, 2224–2237. <https://doi.org/10.1109/TASLP.2020.3009831>.
2. Chen, J.; Kung Yao,.; Hudson, R. Source localization and beamforming. *IEEE Signal Processing Magazine* **2002**, *19*, 30–39. <https://doi.org/10.1109/79.985676>.
3. Lu, J.Y.; Zou, H.; Greenleaf, J.F. Biomedical ultrasound beam forming. *Ultrasound in Medicine & Biology* **1994**, *20*, 403–428. [https://doi.org/10.1016/0301-5629\(94\)90097-3](https://doi.org/10.1016/0301-5629(94)90097-3).
4. Nguyen, N.Q.; Prager, R.W. Minimum Variance Approaches to Ultrasound Pixel-Based Beamforming. *IEEE Transactions on Medical Imaging* **2017**, *36*, 374–384. <https://doi.org/10.1109/TMI.2016.2609889>.
5. Benesty, J.; Cohen, I.; Chen, J. *Fundamentals of signal enhancement and array signal processing*; John Wiley & Sons: Hoboken, NJ, 2017.
6. Kıymık, M.; Güler, İ.; Dizibüyük, A.; Akın, M. Comparison of STFT and wavelet transform methods in determining epileptic seizure activity in EEG signals for real-time application. *Computers in Biology and Medicine* **2005**, *35*, 603–616. <https://doi.org/10.1016/j.combiomed.2004.05.001>.
7. Pan, C.; Chen, J.; Shi, G.; Benesty, J. On microphone array beamforming and insights into the underlying signal models in the short-time-Fourier-transform domain. *The Journal of the Acoustical Society of America* **2021**, *149*, 660–672. <https://doi.org/10.1121/10.0003335>.
8. Chen, W.; Huang, X. Wavelet-Based Beamforming for High-Speed Rotating Acoustic Source. *IEEE Access* **2018**, *6*, 10231–10239. <https://doi.org/10.1109/ACCESS.2018.2795538>.
9. Yang, Y.; Peng, Z.K.; Dong, X.J.; Zhang, W.M.; Meng, G. General Parameterized Time-Frequency Transform. *IEEE Transactions on Signal Processing* **2014**, *62*, 2751–2764. <https://doi.org/10.1109/TSP.2014.2314061>.
10. Almeida, L. The fractional Fourier transform and time-frequency representations. *IEEE Transactions on Signal Processing* **1994**, *42*, 3084–3091. <https://doi.org/10.1109/78.330368>.
11. Crochiere, R.E.; Rabiner, L.R. *Multirate digital signal processing*; Prentice-Hall signal processing series, Prentice-Hall: Englewood Cliffs, N.J, 1983.
12. Oyzerman, A. Speech Dereverberation in the Time-Frequency Domain. Master's thesis, Technion - Israel Institute of Technology, Haifa, Israel, 2012.
13. Talmon, R.; Cohen, I.; Gannot, S. Relative Transfer Function Identification Using Convolutional Transfer Function Approximation. *IEEE Transactions on Audio, Speech, and Language Processing* **2009**, *17*, 546–555. <https://doi.org/10.1109/TASL.2008.2009576>.
14. Kumatani, K.; McDonough, J.; Schacht, S.; Klakow, D.; Garner, P.N.; Li, W. Filter bank design based on minimization of individual aliasing terms for minimum mutual information subband adaptive beamforming. In Proceedings of the 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, Las Vegas, NV, March 2008; pp. 1609–1612. <https://doi.org/10.1109/ICASSP.2008.4517933>.
15. Gopinath, R.; Burrus, C. A tutorial overview of filter banks, wavelets and interrelations. In Proceedings of the 1993 IEEE International Symposium on Circuits and Systems, Chicago, IL, USA, 1993; pp. 104–107. <https://doi.org/10.1109/ISCAS.1993.393668>.
16. Capon, J. High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE* **1969**, *57*, 1408–1418. <https://doi.org/10.1109/PROC.1969.7278>.
17. Erdogan, H.; Hershey, J.R.; Watanabe, S.; Mandel, M.I.; Roux, J.L. Improved MVDR Beamforming Using Single-Channel Mask Prediction Networks. In Proceedings of the Interspeech 2016. ISCA, September 2016, pp. 1981–1985. <https://doi.org/10.21437/Interspeech.2016-552>.
18. DeMuth, G. Frequency domain beamforming techniques. In Proceedings of the ICASSP '77. IEEE International Conference on Acoustics, Speech, and Signal Processing, Hartford, CT, USA, 1977; Vol. 2, pp. 713–715. <https://doi.org/10.1109/ICASSP.1977.1170316>.

19. Bai, M.R.; Ih, J.G.; Benesty, J. *Acoustic Array Systems: Theory, Implementation, and Application*, 1 ed.; Wiley, 2013. <https://doi.org/10.1002/9780470827253>. 322
20. Habets, E. RIR Generator, 2020. 323
21. Nielsen, J.K.; Jensen, J.R.; Jensen, S.H.; Christensen, M.G. The Single- and Multichannel Audio Recordings Database (SMARD). In Proceedings of the Int. Workshop Acoustic Signal Enhancement, Sep. 2014. 324

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content. 325