*Article*

# On the Single-Sideband Transform for MVDR Beamformers

**Vitor Probst Curtarelli[1,*], Israel Cohen[1]**

[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 applicaitons, 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.

**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:



## 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 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

$$\begin{aligned} X_{\mathcal{F}}(f) &\equiv \mathcal{F}\{x(t)\}(f) \\ &= \int_{-\infty}^{\infty} x(t)e^{-\mathrm{j}2\pi f t}\mathrm{d}t \end{aligned} \tag{1}$$

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

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

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

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

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

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

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

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

from which we deduce that

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

### *2.1. Convolution*

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

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

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

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

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

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

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

$$
X_{\mathcal{F}}[l,k] = \sum_{n=0}^{K-1} w[n]x[n+l\cdot O]e^{-\mathrm{j}2\pi k\frac{(n+l\cdot O)}{K}} \tag{10}
$$

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

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

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 have[1]

$$\begin{aligned} X_{\mathcal{S}}[l,k] &= \sqrt{2}\mathbb{R}\left\{X_{\mathcal{F}}[l,k]e^{\mathrm{j}\frac{3\pi}{4}}\right\} \\ &= -\mathbb{R}\{X_{\mathcal{F}}[l,k]\} + \mathbb{I}\{X_{\mathcal{F}}[l,k]\} \end{aligned} \tag{12}$$

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

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

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

$$h[n] * x[n] \overset{\mathcal{S}}{\rightleftharpoons} X_{\mathcal{S}}[l,k]H^{\mathbb{R}}_{\mathcal{F}}[k] + X_{\mathcal{S}}[l,K-k]H^{\mathbb{I}}_{\mathcal{F}}[k] \tag{15}$$

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

$$h[n] * x[n] \overset{\mathcal{S}}{\rightleftharpoons} X_{\mathcal{S}}[l,k] * H^{\mathbb{R}}_{\mathcal{F}}[l,k] + X_{\mathcal{S}}[l,K-k] * H^{\mathbb{I}}_{\mathcal{F}}[l,k] \tag{16}$$

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.

Under the MTF model, let us denote $Y_1[l,k]$ as the first system's output, given by

$$Y_1[l,k] = H_1[k]X[l,k] \tag{17}$$

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

$$A_2[k] = \frac{H_2[k]}{H_1[k]} \tag{18}$$

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

$$A_m[k] = \frac{\mathbb{E}\left\{X_m[l,k]X_1^*[l,k]\right\}}{\mathbb{E}\left\{X_1[l,k]X_1^*[l,k]\right\}} \tag{19}$$

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 $X_{\mathcal{S}}[l,K] \equiv X_{\mathcal{S}}[l,0]$, 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

$$Y_1[l,k] = H'_1[k]X'[l,k] + H''_1[k]X''[l,k] \tag{20}$$

and in the same way for $Y_2[l,k]$ under the SSBT. From Eq. (15), we easily see that

$$H'_m[k] = H'_m[K-k] = H^{\mathbb{R}}_{\mathcal{F};m}[k] \tag{21a}$$
$$H''_m[k] = -H''_m[K-k] = H^{\mathbb{I}}_{\mathcal{F};m}[k] \tag{21b}$$

We let

$$X'_1[l,k] = H'_1[k]X'[l,k] \tag{22a}$$
$$X''_1[l,k] = H''_1[k]X''[l,k] \tag{22b}$$

and

$$\begin{aligned} X_1[l,k] &= H'_1[k]X'[l,k] + H''_1[k]X''[l,k] \\ &= X'_1[l,k] + X''_1[l,k] \end{aligned} \tag{23}$$

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

$$Y_1[l,k] = A'_1[k]X'_1[l,k] + A''_1[k]X''_1[l,k] \tag{24a}$$
$$Y_2[l,k] = A'_2[k]X'_1[l,k] + A''_2[k]X''_1[l,k] \tag{24b}$$

where

$$A^n_m[k] = \frac{H^n_m[k]}{H^n_1[k]} \tag{25}$$

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. (19) used for the STFT isn't applicable here. Using it directly on the observable input $X_m[l,k]$ would yield

$$\frac{\mathbb{E}\{X_m[l,k]X_1[l,k]\}}{\mathbb{E}\{X_1[l,k]X_1[l,k]\}} = \frac{H'_m[k]H'_1[k] + H''_m[k]H''_1[k]}{H_1'^2[k] + H_1''^2[k]} \tag{26}$$

where we used that different frequency bins are i.i.d. and zero-mean (see Property VI in Appendix A). Noticeably this isn't the same as the desired result from Eq. (25), however we can work around it.

First, it is easy to see that

$$\begin{aligned} \mathbb{E}\{X_m[l,k]X_1[l,k]\} &= \left(H'_m[k]H'_1[k] + H''_m[k]H''_1[k]\right)\sigma^2_{X,k} \\ \mathbb{E}\{X_m[l,k]X_1[l,K-k]\} &= \left(H'_m[k]H'_1[k] - H''_m[k]H''_1[k]\right)\sigma^2_{X,k} \end{aligned} \tag{27}$$

where $\sigma^2_{X,k} = \mathbb{E}\left\{(X'[l,k])^2\right\} = \mathbb{E}\left\{(X''[l,k])^2\right\}$. From Eq. (27), we can see that we can calculate $H'_m[k]H'_1[k]\sigma^2_{X,k}$ and $H''_m[k]H''_1[k]\sigma^2_{X,k}$ as

$$H'_m[k]H'_1[k]\sigma^2_{X,k} = \frac{\mathbb{E}\{X_m[l,k]X_1[l,k]\} + \mathbb{E}\{X_m[l,k]X_1[l,K-k]\}}{2} \tag{28a}$$
$$H''_m[k]H''_1[k]\sigma^2_{X,k} = \frac{\mathbb{E}\{X_m[l,k]X_1[l,k]\} - \mathbb{E}\{X_m[l,k]X_1[l,K-k]\}}{2} \tag{28b}$$

and from this we can easily calculate $(H_1'[k])^2\sigma_{X,k}^2$ and $(H_1''[k])^2\sigma_{X,k}^2$ as well. Therefore, it is trivial that the gains $A_m^n[k]$ from Eq. (25) can be calculated as

$$\begin{aligned} A_m'[k] &= \frac{H_m'[k]}{H_1'[k]} \\ &= \frac{\mathbb{E}\{X_m[l,k]X_1[l,k]\} + \mathbb{E}\{X_m[l,k]X_1[l,K-k]\}}{\mathbb{E}\{X_1[l,k]X_1[l,k]\} + \mathbb{E}\{X_1[l,k]X_1[l,K-k]\}} \end{aligned} \tag{29a}$$

$$\begin{aligned} A_m''[k] &= \frac{H_m''[k]}{H_1''[k]} \\ &= \frac{\mathbb{E}\{X_m[l,k]X_1[l,k]\} - \mathbb{E}\{X_m[l,k]X_1[l,K-k]\}}{\mathbb{E}\{X_1[l,k]X_1[l,k]\} - \mathbb{E}\{X_1[l,k]X_1[l,K-k]\}} \end{aligned} \tag{29b}$$

These same derivations can be used with the CTF model. That is, if we consider $A_m^n[l,k]$ as a convolutive gain, then

$$A_m^n[l,k] \approx \frac{H_m^n[l,k]}{H_1^n[0,k]} \tag{30}$$

This is 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. Under this model, our systems become

$$Y_m[l,k] = A_m[l,k] * X_1[l,k] \tag{31}$$

for the STFT, where $X_1[l,k] = H_1[l,k] * X[l,k]$, and

$$Y_m[l,k] = A_m'[l,k] * X_1'[l,k] + A_m''[l,k] * X_1''[l,k] \tag{32}$$

for the SSBT, in which $X_1^n[l,k] = H_1^n[l,k] * X^n[l,k]$. We raise attention, since using Eqs. (21) to (23) we have that $X_1'[l,k] \neq X_1''[l,K-k]$, even though both originate from $X_k[l]$. Likewise, for the same reason $X_1''[l,k] \neq X_1'[l,K-k]$.

Notice that the STFT formulation can be seen as a particular case of the SSBT one (comparing Eq. (17) with Eq. (20), or Eq. (31) with Eq. (32)), where $A_m'[l,k] = A_m[l,k]$ and $A_m''[l,k] = 0$. Therefore, from now on we will use the SSBT formulation from Eq. (32), as it is a less restricting model.

## 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

$$Y_{m,k}[l] = X_{m,k}[l] + S_{m,k}[l] + R_{m,k}[l] \tag{33}$$

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 multiple paths between the desired source and the $m$-th sensor as a system, with the CTF model $X_{m,k}[l]$ can be represented by Eq. (32), as

$$X_{m,k}[l] = A'_{m,k}[l] * X'_{1,k}[l] + A''_{m,k}[l] * X''_{1,k}[l] \tag{34}$$

$$X'_{1,k}[l] = H'_{1,k}[l] * X_k[l] \tag{35a}$$

$$X''_{1,k}[l] = H''_{1,k}[l] * X_{K-k}[l] \tag{35b}$$

where $X'_{1,k}[l]$ is the portion of the desired signal at the reference sensor of the direct frequency, $X''_{1,k}[l]$ is the portion of the conjugate frequency $K-k$, $H^n_{1,k}[l]$ are the desired signal's transfer functions between source and reference, and $A^n_{m,k}[l]$ are the RTF's between each sensor and the reference. This formulation models both the STFT and the SSBT, as with the STFT we just take $A''_{m,k}[l] = H''_{m,k}[l] = 0$, as there is no cross-over between conjugate frequencies.

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 $\Delta$ non-causal samples in $A_{m,k}[l]$. It is trivial to see that, for Eq. (34) 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

$$X_{m,k}[l+\lambda] = \sum_{\tau}\Big(A'_{m,k}[\tau]X'_{1,k}[l+\lambda-\tau] + A''_{m,k}[\tau]X''_{1,k}[l+\lambda-\tau]\Big) \tag{36}$$

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

$$\begin{aligned} X_{m,k}[l+\lambda] &= A'_{m,k}[\lambda]X'_{1,k}[l] + A''_{m,k}[\lambda]X''_{1,k}[l] \\ &\quad + \sum_{\tau\neq\lambda}\Big(A'_{m,k}[\tau]X'_{1,k}[l+\lambda-\tau] + A''_{m,k}[\tau]X''_{1,k}[l+\lambda-\tau]\Big) \\ &= A'_{m,k}[\lambda]X'_{1,k}[l] + A''_{m,k}[\lambda]X''_{1,k}[l] + Q_{m,k}[l+\lambda] \end{aligned} \tag{37}$$

where the first two terms are the contributions of the desired signals $X'_{1,k}[l]$ and $X''_{1,k}[l]$ 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. (33) with Eq. (34), we have that

$$Y_{m,k}[l+\lambda] = A'_{m,k}[\lambda]X'_{1,k}[l] + A''_{m,k}[\lambda]X''_{1,k}[l] + Q_{m,k}[l+\lambda] + S_{m,k}[l+\lambda] + R_{m,k}[l+\lambda] \tag{38}$$

We consider $L_Y$ previous and $\bar{L}_Y$ future frames of $Y_{m,k}[l]$, and let $L = L_Y + \bar{L}_Y + 1$ the total number of frames. With this, we define $\mathbf{Y}_{m,k}[l]$ as a vector containing those $L$ samples,

$$\mathbf{Y}_{m,k}[l] = \mathbf{A}'_{m,k}X'_{1,k}[l] + \mathbf{A}''_{m,k}X''_{1,k}[l] + \mathbf{Q}_{m,k}[l] + \mathbf{S}_{m,k}[l] + \mathbf{R}_{m,k}[l] \tag{39}$$

in which

$$\mathbf{Y}_{m,k}[l] = \Big[\, Y_{m,k}[l+\bar{L}_Y]\,,\ \cdots\,,\ Y_{m,k}[l]\,,\ \cdots\,,\ Y_{m,k}[l-L_Y] \,\Big]^{\mathsf{T}} \tag{40}$$

and similarly for all other signals, and

$$\mathbf{A}^n_{m,k} = \Big[\, A_{m,k}[\bar{L}_Y]\,,\ \cdots\,,\ A_{m,k}[0]\,,\ \cdots\,,\ A_{m,k}[-L_Y] \,\Big]^{\mathsf{T}} \tag{41}$$

with all of them being $L \times 1$ vectors.

Now stacking these signals in a vector sensor-wise, we get

$$\mathbf{y}_k[l] = \mathbf{a}'_k X'_{1,k}[l] + \mathbf{a}''_k X''_{1,k}[l] + \mathbf{q}_k[l] + \mathbf{s}_k[l] + \mathbf{r}_k[l] \tag{42}$$

with

$$\mathbf{y}_k[l] = \left[ \mathbf{Y}^{\mathsf{T}}_{1,k}[l]\,,\ \cdots\,,\ \mathbf{Y}^{\mathsf{T}}_{M,k}[l] \right]^{\mathsf{T}} \tag{43}$$

and the same for $\mathbf{a}^n_k$, $\mathbf{q}_k[l]$, $\mathbf{s}_k[l]$ and $\mathbf{r}_k[l]$, where them all are $ML \times 1$ 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. (34) 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

$$\begin{aligned} Z^n_k[l] &\approx X_{1,k}[l] \\ &= \mathbf{f}^{\mathsf{H}}_k[l]\mathbf{y}_k[l] \end{aligned} \tag{44}$$

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

$$\begin{aligned} Z_k[l] &= \sum_m \bar{\mathbf{f}}^{\mathsf{H}}_{m,k}[l]\bar{\mathbf{y}}_{m,k}[l] \\ &= \sum_m F^*_{m,k}[l] * Y_{m,k}[l] \end{aligned} \tag{45}$$

where $\bar{\mathbf{f}}_{m,k}[l]$ is the $L \times 1$ 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. (44), with Eq. (42) we can write

$$Z_k[l] = \mathbf{f}^{\mathsf{H}}_k[l]\mathbf{a}'_k X'_{1,k}[l] + \mathbf{f}^{\mathsf{H}}_k[l]\mathbf{a}''_k X''_{1,k}[l] + \mathbf{f}^{\mathsf{H}}_k[l]\mathbf{q}_k[l] + \mathbf{f}^{\mathsf{H}}_k[l]\mathbf{s}_k[l] + \mathbf{f}^{\mathsf{H}}_k[l]\mathbf{r}_k[l] \tag{46}$$

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

$$\mathbf{f}^{\mathsf{H}}_k[l]\mathbf{a}'_k = 1 \tag{47a}$$

$$\mathbf{f}^{\mathsf{H}}_k[l]\mathbf{a}''_k = 1 \tag{47b}$$

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

$$\mathbf{f}^{\mathsf{H}}_k[l]\mathbf{C}_k = \mathbf{i}^{\mathsf{T}} \tag{48}$$

where, for the STFT, $\mathbf{C}_k = \mathbf{a}_k$ and $\mathbf{i} = 1$; and, for the SSBT, $\mathbf{C}_k = \left[\mathbf{a}'_k\,,\ \mathbf{a}''_k\right]$, and $\mathbf{i} = [1\,,\ 1]$.

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

$$\mathbf{f}_{\mathrm{mpdr};k}[l] = \min_{\mathbf{f}_k[l]} \mathbf{f}^{\mathsf{H}}_k[l]\mathbf{\Phi}_{\mathbf{y}_k}[l]\mathbf{f}_k[l] \text{ s.t. } \mathbf{f}^{\mathsf{H}}_k[l]\mathbf{C}_k = \mathbf{i}^{\mathsf{T}} \tag{49}$$

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

$$\mathbf{f}_{\mathrm{mpdr};k}[l] = \boldsymbol{\Phi}_{\mathbf{y}_k}^{-1}[l]\mathbf{C}_k\left[\mathbf{C}_k^{\mathrm{H}}\boldsymbol{\Phi}_{\mathbf{y}_k}^{-1}[l]\mathbf{A}_k\right]^{-1}\mathbf{i} \tag{50}$$

Obviously, with the SSBT all conjugate-transpose operations are replaced with simple transposes, as in this transform all signals and matrices are real-valued.

**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 | Convolutive 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

The Fourier Transform (FT) is defined as

$$X_{\mathcal{F}}(f) = \int_{-\infty}^{\infty} x(t)e^{-\mathrm{j}2\pi f t}\mathrm{d}t \tag{A.1}$$

with an inverse

$$x(t) = \int_{-\infty}^{\infty} X_{\mathcal{F}}(f)e^{\mathrm{j}2\pi f t}\mathrm{d}t \tag{A.2}$$

The Real Fourier Transform (RFT) is defined as

$$X_{\mathcal{R}}(f) = \sqrt{2}\mathbb{R}\left\{\int_{-\infty}^{\infty} x(t)e^{-\mathrm{j}2\pi f t+\mathrm{j}\frac{3\pi}{4}}\mathrm{d}t\right\} \tag{A.3}$$

with an inverse

$$x(t) = \sqrt{2}\mathbb{R}\left\{\int_{-\infty}^{\infty} X_{\mathcal{R}}(f)e^{\mathrm{j}2\pi ft-\mathrm{j}\frac{3\pi}{4}}\,\mathrm{d}f\right\} \tag{A.4}$$

**Property I:** *The FT and RFT are bijective transformations of one-another.*

**Proof:** By manipulating Eq. (A.3) we can get

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

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

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

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

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

With this, we have that

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

and therefore

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

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

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

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

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

Since for each $X_{\mathcal{F}}(f)$ there exists one and only one $X_{\mathcal{R}}(f)$, and vice-versa, it is possible to define a bijective transformation $T$ such that

$$X_{\mathcal{F}}(f) \overset{T}{\rightleftharpoons} X_{\mathcal{R}}(f) \tag{i.8}$$

---

**Property II:** *The IRFT is the inverse of the RFT.*

**Proof:** From Property I, we have that

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

Substituting this in Eq. (A.4), we have

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

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,

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

thus concluding the proof.

---

**Property 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

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

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

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

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

$$Y_{\mathcal{F}} = H_{\mathcal{F}}^{\mathbb{R}}X_{\mathcal{F}}^{\mathbb{R}} + \mathrm{j}H_{\mathcal{F}}^{\mathbb{R}}X_{\mathcal{F}}^{\mathbb{I}} + \mathrm{j}H_{\mathcal{F}}^{\mathbb{I}}X_{\mathcal{F}}^{\mathbb{R}} - H_{\mathcal{F}}^{\mathbb{I}}X_{\mathcal{F}}^{\mathbb{I}} \tag{iii.3}$$

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

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

Assuming the convolution theorem true for the RFT,

$$Y_{\mathcal{R}} = H_{\mathcal{F}}^{\mathbb{R}}X_{\mathcal{F}}^{\mathbb{R}} + H_{\mathcal{F}}^{\mathbb{R}}X_{\mathcal{F}}^{\mathbb{I}} + H_{\mathcal{F}}^{\mathbb{I}}X_{\mathcal{F}}^{\mathbb{R}} + H_{\mathcal{F}}^{\mathbb{I}}X_{\mathcal{F}}^{\mathbb{I}} \tag{iii.5}$$

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

$$\tilde{Y}_{\mathcal{R}} = -H_{\mathcal{F}}^{\mathbb{R}}X_{\mathcal{F}}^{\mathbb{R}} - H_{\mathcal{F}}^{\mathbb{R}}X_{\mathcal{F}}^{\mathbb{I}} - H_{\mathcal{F}}^{\mathbb{I}}X_{\mathcal{F}}^{\mathbb{R}} + H_{\mathcal{F}}^{\mathbb{I}}X_{\mathcal{F}}^{\mathbb{I}} \tag{iii.6}$$

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 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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

---

**Property 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,

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

Since they are identically distributed, we denote

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

and therefore

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

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

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

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

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

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

---

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

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

$$\mathbb{E}\left\{X_{\mathcal{F}}(f)X_{\mathcal{F}}^*(\text{-}f)\right\} = \mathbb{E}\left\{X_{\mathcal{F}}^{\mathbb{R}}(f)^2 + 2\mathrm{j}X_{\mathcal{F}}^{\mathbb{R}}(f)X_{\mathcal{F}}^{\mathbb{I}}(f) - X_{\mathcal{F}}^{\mathbb{I}}(f)^2\right\} \tag{vi.1}$$

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

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

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

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

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

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

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 Eqs. (iv.5) and (iv.6),

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

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

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

---

**Proposition A1.** *AAA BBBB*

## 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.
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.compbiomed.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 Convolutive 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.
20. Habets, E. RIR Generator, 2020.
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.