Quadratic Beamforming for Magnitude Estimation

Gal Itzhak
Faculty of Electrical Engineering
Technion –Israel Institute of Technology
Haifa, Israel
galitz@campus.technion.ac.il

Jacob Benesty
INRS-EMT, University of Quebec
Montreal, Canada
benesty@emt.inrs.ca

Israel Cohen
Faculty of Electrical Engineering
Technion –Israel Institute of Technology
Haifa, Israel
icohen@ee.technion.ac.il

***Abstract*—In this paper, we introduce an optimal quadratic Wiener beamformer for magnitude estimation of a desired signal. For simplicity, we focus on a two-microphone array and develop an iterative algorithm for magnitude estimation based on a quadratic multichannel noise reduction approach. We analyze two test cases, with uncorrelated and correlated noises. In each, we derive the appropriate versions of the Wiener beamformer, as well as their corresponding unbiased magnitude estimators. We compare the root-mean-squared errors (RMSEs) for the linear and quadratic Wiener beamformers and show that for low input signal-to-noise ratios (SNRs), the RMSE obtained with the proposed approach is either lower than or equal to the RMSE obtained with the linear Wiener beamformer, depending on the type of noise and its distribution.**

I. Introduction

Magnitude estimation of a desired signal of interest is a common task in a wide area of fields, including communications, target detection, and speech enhancement. The desired signal is typically only observable through noisy samples, that is, it is corrupted by noise, which may critically damage the performance of the application at hand. Consequently, a large number of studies have addressed this issue by exploiting data either from a single microphone or a sensor array.

Most commonly with communications and speech signals, processing is done in the frequency domain. That is, a frame of consecutive time-domain samples is transformed into the frequency domain by applying the fast Fourier transform (FFT), yielding a set of analysis coefficients, which can be processed more efficiently than the time-domain samples. This is particularly significant with multichannel methods, in which time-domain noisy observations are sampled simultaneously in multiple sensors. These methods typically seek for a linear optimal solution with respect to some criterion, looking to estimate both the desired signal phase and magnitude [1]–[4]. On the contrary, single-channel approaches may either attempt to estimate the complex desired signal [5]–[8] or may directly attempt to estimate its magnitude [9]–[17], which is known to be more prominent than its phase for some applications.

Recently, a quadratic noise reduction approach was suggested [17], [18]. The idea behind the quadratic approach is to estimate the spectral power of the desired signal by applying a complex-valued beamformer, which takes into consideration data from higher-order moments. The quadratic beamformer is applied to a modified version of the noisy observations vector and may be seen as a generalization of the traditional linear beamformer.

In this paper, we introduce a quadratic beamformer for a desired signal magnitude estimation, which is optimal in terms of the RMSE. For simplicity, we focus on a two-microphone array and develop an iterative algorithm for magnitude estimation based on the quadratic multichannel noise reduction approach. We analyze two test cases, with uncorrelated and correlated noises. In each case, we derive the appropriate version of the Wiener beamformer, as well as the corresponding unbiased magnitude estimator. We compare the RMSEs with the linear and quadratic Wiener beamformers and show that for low input SNRs, the RMSE obtained with the proposed approach is either lower than or equal to the RMSE obtained with the linear Wiener beamformer, depending on the type of noise and its distribution.

The rest of the paper is organized as follows. In Section II, we present the signal model. In Section III, we introduce the quadratic beamforming approach. In Section IV, we analyze two test cases, with both uncorrelated and correlated noises. We derive the quadratic optimal beamformers in terms of minimum RMSE and use them to derive unbiased magnitude estimators. Then, in Section V, we demonstrate the advantage of the quadratic approach over the linear one through simulations. Finally, we summarize this work in Section VI.

II. Signal Model

Consider an array consisting of M omnidirectional microphones. The received signals at the frequency index f are expressed as [3], [19]

$$Y_m(f) = X_m(f) + V_m(f), \ m = 1, 2, \ldots, M, \tag{1}$$

where $Y_m(f)$ is the mth microphone signal, $X_m(f)$ is the zero-mean desired speech signal, and $V_m(f)$ is the zero-mean additive noise. It is assumed that the desired signal and noise are uncorrelated.

Considering the first microphone as the reference, we may express (1) in a vector notation:

$$\mathbf{y}(f) = \mathbf{d}_{\theta_\mathrm{d}}(f) X_1(f) + \mathbf{v}(f), \tag{2}$$

where

$$\mathbf{y}(f) = \begin{bmatrix} Y_1(f) & Y_2(f) & \cdots & Y_M(f) \end{bmatrix}^T,$$

This research was supported by the ISF-NSFC joint research program (grant No. 2514/17), and the Pazy Research Foundation.

$\mathbf{v}(f)$ is defined in a similar manner to $\mathbf{y}(f)$, the superscript ^{T} is the transpose operator, and

$$\mathbf{d}_{\theta_\mathrm{d}}(f) = \begin{bmatrix} 1 & e^{-\jmath 2\pi f \delta \cos\theta_\mathrm{d}/c} & \cdots & e^{-\jmath 2\pi f \delta (M-1)\cos\theta_\mathrm{d}/c} \end{bmatrix}^T \tag{3}$$

is the frequency-domain steering vector, considering the farfield planar wave model [1], [2]. In addition, θ_d is the desired speech signal incident angle, δ is the inter-element spacing, $c = 340$ m/s is the speed of sound, and $\jmath = \sqrt{-1}$ is the imaginary unit.

Since $\mathbf{y}(f)$ is the sum of two uncorrelated components, its correlation matrix is

$$\begin{aligned} \mathbf{\Phi}_\mathbf{y}(f) &= E\left[\mathbf{y}(f)\mathbf{y}^H(f)\right] \\ &= \phi_{X_1}(f)\mathbf{d}_{\theta_\mathrm{d}}(f)\mathbf{d}_{\theta_\mathrm{d}}^H(f) + \mathbf{\Phi}_\mathbf{v}(f), \end{aligned} \tag{4}$$

where $E[\cdot]$ denotes mathematical expectation, the superscript ^{H} is the conjugate-transpose operator, $\phi_{X_1}(f) = E\left[|X_1(f)|^2\right]$ is the variance of $X_1(f)$, and $\mathbf{\Phi}_\mathbf{v}(f) = E\left[\mathbf{v}(f)\mathbf{v}^H(f)\right]$ is the 2nd-order correlation matrix of $\mathbf{v}(f)$ whose top-left element is $\phi_{V_1}(f) = E\left[|V_1(f)|^2\right]$.

III. Quadratic Beamforming

Conventionally, with an array of M sensors, beamforming is performed by applying a complex-valued linear filter, $\mathbf{h}(f)$ of length M, to the observation signal vector, $\mathbf{y}(f)$, i.e., [3], [19]

$$\begin{aligned} \widehat{X}(f) &= \mathbf{h}^H(f)\mathbf{y}(f) \\ &= X_1(f)\mathbf{h}^H(f)\mathbf{d}_{\theta_\mathrm{d}}(f) + \mathbf{h}^H(f)\mathbf{v}(f), \end{aligned} \tag{5}$$

where the filter output, $\widehat{X}(f)$, is an estimate of $X_1(f)$. We note that $\widehat{X}(f)$ is complex, that is, it carries information on both the magnitude and phase of the desired signal.

Recently, a quadratic noise reduction approach was suggested in [18]. According to this technique, we can estimate the spectral power of $\widehat{X}(f)$ defined in (5) for a given complex-valued beamformer $\widetilde{\mathbf{h}}(f)$ of length M^2 by

$$\left|\widehat{X}(f)\right|^2 = \widetilde{\mathbf{h}}^H(f)\widetilde{\mathbf{y}}(f), \tag{6}$$

where $\widetilde{\mathbf{y}}(f) = \mathbf{y}^*(f) \otimes \mathbf{y}(f)$, with the superscript * being the complex-conjugate operator and $\otimes$ the Kronecker product. Additionally, it was shown that

$$\begin{aligned} \phi_{\widehat{X}}(f) &= E\left[\left|\widetilde{\mathbf{h}}^H(f)\widetilde{\mathbf{y}}(f)\right|\right] \\ &\approx \left|\phi_{X_1}(f)\widetilde{\mathbf{h}}^H(f)\widetilde{\mathbf{d}}_{\theta_\mathrm{d}}(f) + \widetilde{\mathbf{h}}^H(f)\mathrm{vec}\left[\mathbf{\Phi}_\mathbf{v}(f)\right]\right|, \end{aligned} \tag{7}$$

where $\widetilde{\mathbf{d}}_{\theta_\mathrm{d}}(f) = \mathbf{d}_{\theta_\mathrm{d}}^*(f) \otimes \mathbf{d}_{\theta_\mathrm{d}}(f)$ is the quadratic steering vector and $\mathrm{vec}[\cdot]$ is the vectorization operator.

IV. Analysis of Two Test Cases

As of this point, for the sake of simplicity, let us assume that $M = 2$ and $\theta_\mathrm{d} = 90^o$ (it should be noted, though, that this approach is indeed general and not limited to certain array sizes or incident angles). Hence, (2) reduces to

$$\begin{aligned} \mathbf{y} &= \begin{bmatrix} X \\ X \end{bmatrix} + \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \\ &= \begin{bmatrix} ae^{\jmath\phi_a} \\ ae^{\jmath\phi_a} \end{bmatrix} + \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}, \end{aligned} \tag{8}$$

where a and ϕ_a are the magnitude and phase of the desired signal, respectively, and the explicit dependence on frequency is dropped to lighten the notation. In this model, we assume the desired signal is a deterministic and unknown variable. With a known to be more prominent than ϕ_a for speech enhancement purposes [10], we further assume that $\phi_a = 0$. Consequently, our objective is to derive optimal estimators for the real and positive variable a in two key cases: uncorrelated and correlated noises.

A. Uncorrelated Noise

Let us assume that V_1 and V_2 are independent, real, zero-mean, identically-distributed random variables whose variances is σ^2. Adopting the RMSE as the performance criterion, the optimal linear beamformer is given by the linear Wiener beamformer [20], [21]:

$$\begin{aligned} \mathbf{h}_\mathrm{W} &= \phi_X \mathbf{\Phi}_\mathbf{y}^{-1}\mathbf{d}_{\theta_\mathrm{d}} \\ &= a^2\mathbf{\Phi}_\mathbf{y}^{-1}\mathbf{1}_2 \\ &\doteq \left[H_\mathrm{W}(1) \quad H_\mathrm{W}(2)\right]^T, \end{aligned} \tag{9}$$

where $\mathbf{1}_2$ is an "all-ones" vector of length 2, and we assume that the correlation matrix $\mathbf{\Phi}_\mathbf{y}$ is either given or can be estimated from the noisy observations. Then, considering (8), an unbiased estimator for a is given by

$$\hat{a}\left[\mathbf{h}_\mathrm{W}\right] = \sqrt{\max\left\{\frac{|\mathbf{h}_\mathrm{W}^T\mathbf{y}|^2 - \sigma^2\|\mathbf{h}_\mathrm{W}\|^2}{|\mathbf{h}_\mathrm{W}^T\mathbf{1}_2|^2}, 0\right\}}, \tag{10}$$

where $\|\cdot\|$ is the Euclidean norm.

We follow a similar protocol with $\widetilde{\mathbf{h}}$ where (6) is employed. That is, we are interested in solving the following optimization criterion:

$$\min_{\widetilde{\mathbf{h}}} E\left|\widetilde{\mathbf{h}}^T\widetilde{\mathbf{y}} - a^2\right|^2, \tag{11}$$

whose solution is given by

$$\begin{aligned} \widetilde{\mathbf{h}}_\mathrm{W} &= a^2\mathbf{\Phi}_{\widetilde{\mathbf{y}}}^{-1}E\left(\widetilde{\mathbf{y}}\right) \\ &= a^2\mathbf{\Phi}_{\widetilde{\mathbf{y}}}^{-1}\begin{bmatrix} a^2+\sigma^2 \\ a^2 \\ a^2 \\ a^2+\sigma^2 \end{bmatrix} \\ &\doteq \left[\widetilde{H}_\mathrm{W}(1) \quad \widetilde{H}_\mathrm{W}(2) \quad \widetilde{H}_\mathrm{W}(3) \quad \widetilde{H}_\mathrm{W}(4)\right]^T, \end{aligned} \tag{12}$$

where the 4th-order correlation matrix $\mathbf{\Phi}_{\widetilde{\mathbf{y}}} = E\left[\widetilde{\mathbf{y}}\widetilde{\mathbf{y}}^H\right]$ is assumed to be known or can be estimated from the noisy observations. We refer to $\widetilde{\mathbf{h}}_{\mathrm{W}}$ as the quadratic Wiener filter. Therefore, in a similar manner to (10), we obtain an unbiased estimator for a based on $\widetilde{\mathbf{h}}_{\mathrm{W}}$:

$$\hat{a}\left[\widetilde{\mathbf{h}}_{\mathrm{W}}\right] = \sqrt{\max\left\{\frac{\widetilde{\mathbf{h}}_{\mathrm{W}}^T\widetilde{\mathbf{y}} - \sigma^2\left[\widetilde{H}_{\mathrm{W}}(1) + \widetilde{H}_{\mathrm{W}}(4)\right]}{\widetilde{\mathbf{h}}_{\mathrm{W}}^T\mathbf{1}_4}, 0\right\}}, \tag{13}$$

where $\mathbf{1}_4$ is an "all-ones" vector of length 4.

While their structures exhibit some level of similarity, some key differences between $\mathbf{h}_{\mathrm{W}}$ and $\widetilde{\mathbf{h}}_{\mathrm{W}}$ should be addressed. For example, $\mathbf{h}_{\mathrm{W}}$ is, in general, designed to estimate a complex-valued variable, whereas $\widetilde{\mathbf{h}}_{\mathrm{W}}$ is designed to estimate a real and positive variable. In addition, $\mathbf{h}_{\mathrm{W}}$ only requires the second order-statistics of the noisy observations, but $\widetilde{\mathbf{h}}_{\mathrm{W}}$ takes advantage of their 4th-order statistics. As a result, the RMSE with $\hat{a}\left[\widetilde{\mathbf{h}}_{\mathrm{W}}\right]$ is expected to be potentially lower than with $\hat{a}\left[\mathbf{h}_{\mathrm{W}}\right]$. Note that the inversion of $\mathbf{\Phi}_{\widetilde{\mathbf{y}}}$ requires more multiplication operations than $\mathbf{\Phi}_{\mathbf{y}}$, but when M is small, the additional complexity is insignificant.

We end this part by pointing out that both versions of the Wiener beamformers require the estimate of a to be known in advance. Since this is the value we wish to estimate, we will employ an iterative procedure in which every iteration consists of two steps: (a) deriving the appropriate beamformer for a given value of a and (b) using that beamformer and its corresponding estimator to generate a new estimate for a. It can be verified that due to the convex nature of the problem, the convergence of the beamformers, and thereby the estimate of a, is guaranteed. Summary of the magnitude estimation algorithm with the quadratic Wiener beamformer, given multiple noisy observations, is elaborated in Algorithm 1. We note that the estimation process with the linear Wiener beamformer is similar, but requires the following modifications: (a) equations (12) and (13) are replaced by (9) and (10), respectively, (b) lines 5 and 6 are omitted as $\mathbf{\Phi}_{\mathbf{y}}$ is computed directly from $\{\mathbf{y}_n\}_{n=1}^N$ in line 7, and (c) the expression $\mathbf{h}_{\mathrm{W}} \leftarrow [1\ 0]^T$ replaces line 9.

B. Correlated Noise

The correlated noise case corresponds, for example, to directional interferences. That is, the same noise signal is received in both microphones but with a frequency-dependent phase difference. Hence, with two uncorrelated real directional interferences V_1 and V_2, (8) reduces to

$$\mathbf{y} = \begin{bmatrix} a \\ a \end{bmatrix} + \begin{bmatrix} V_1 \\ V_1 e^{-\jmath 2\pi f \delta \cos\theta_{\mathrm{i},1}/c} \end{bmatrix} + \begin{bmatrix} V_2 \\ V_2 e^{-\jmath 2\pi f \delta \cos\theta_{\mathrm{i},2}/c} \end{bmatrix}, \tag{14}$$

where $\theta_{\mathrm{i},1}$ and $\theta_{\mathrm{i},2}$ are the respective incident angles of V_1 and V_2. We assume $V_1, V_2 \sim \mathcal{N}\left(0, \sigma^2/2\right)$, and note that the aforementioned phase differences turn the problem from real to complex.

Algorithm 1 Magnitude Estimation with the Quadratic Wiener Beamformer

```
1:  Input: {y_n}_{n=1}^N,                 ▷ set of N noisy vectors
2:         N_0,                  ▷ number of samples to estimate Φ_ỹ
3:         I_0,                            ▷ number of iterations
4:         a_0                             ▷ initial guess for a
5:  for n=1:N do
6:      ỹ_n ← y_n^* ⊗ y_n                  ▷ modify the observations
7:  Φ_ỹ ← (1/N_0) Σ_{n=1}^{N_0} ỹ_n ỹ_n^H
8:  â ← a_0                                 ▷ initialize â
9:  h̃_W ← [1 0 0 0]^T                      ▷ initialize h̃_W
10: for i=1:I_0 do
11:     obtain h̃_W using (12)              ▷ update h̃_W
12:     for n=1:N do
13:         obtain â_n using (13)
14:     â ← (1/#[â_n>0]) Σ_{[â_n>0]} â_n    ▷ update â
15: Output: â                 ▷ desired signal magnitude estimate
```

We turn our attention to the well-known beampattern, which exhibits the ULA response to a plane wave impinging from the direction θ. With a linear beamformer $\mathbf{h}$ of length M, the beampattern is defined by

$$\mathcal{B}_\theta\left[\mathbf{h}\right] = \mathbf{h}^H\mathbf{d}_\theta, \tag{15}$$

where the steering vector $\mathbf{d}_\theta$ is defined as in (3). It is well known that a linear beamformer of length $M = 2$ is only capable of placing a single zero in its beampattern (in addition to the distortionless constraint) [3]. Therefore, we cannot completely eliminate the two directional interferences simultaneously. Instead, we will use the linear Wiener beamformer, $\mathbf{h}_{\mathrm{W}}$, from the previous part and derive a corresponding magnitude estimator.

Recalling (7), we may define an analogous power beampattern with a quadratic beamformer $\widetilde{\mathbf{h}}$ of length $M^2 = 4$ by

$$\begin{aligned} \widetilde{\mathcal{B}}_\theta\left[\widetilde{\mathbf{h}}\right] &= \widetilde{\mathbf{h}}^H\widetilde{\mathbf{d}}_\theta \\ &= \begin{bmatrix} \widetilde{H}(3) \\ \widetilde{H}(1) + \widetilde{H}(4) \\ \widetilde{H}(2) \end{bmatrix}^H \begin{bmatrix} e^{\jmath 2\pi f\delta\cos\theta/c} \\ 1 \\ e^{-\jmath 2\pi f\delta\cos\theta/c} \end{bmatrix} \\ &= \widetilde{\mathbf{g}}^H \begin{bmatrix} 1 \\ e^{-\jmath 2\pi f\delta\cos\theta/c} \\ e^{-\jmath 4\pi f\delta\cos\theta/c} \end{bmatrix}, \end{aligned} \tag{16}$$

where

$$\widetilde{\mathbf{h}} = \left[\widetilde{H}(1)\ \ \widetilde{H}(2)\ \ \widetilde{H}(3)\ \ \widetilde{H}(4)\right]^T, \tag{17}$$

$$\widetilde{\mathbf{g}} = e^{-\jmath 2\pi f\delta\cos\theta/c}\left[\widetilde{H}(3)\ \ \widetilde{H}(1) + \widetilde{H}(4)\ \ \widetilde{H}(2)\right]^T. \tag{18}$$

We observe that the power beampattern of a beamformer $\widetilde{\mathbf{h}}$ of length 2 is mathematically equal to a linear beampattern of an alternative beamformer $\widetilde{\mathbf{g}}$ of length 3 whose elements are formed by linear combinations of the elements of $\widetilde{\mathbf{h}}$. We deduce that $\widetilde{\mathbf{h}}$ is capable of placing two distinct nulls in its power beampattern.

Next, we will adapt the two versions of the Wiener beamformer and derive appropriate estimators. As $\mathbf{h}_{\mathrm{W}}$ depends merely on $\mathbf{\Phi}_{\mathbf{y}}$, whereas $\widetilde{\mathbf{h}}_{\mathrm{W}}$ depends on both $\mathbf{\Phi}_{\widetilde{\mathbf{y}}}$ and $E\left(\widetilde{\mathbf{y}}\right)$, the linear beamformer remains the same as in (9), but the quadratic beamformer changes to

$$\begin{aligned} \widetilde{\mathbf{h}}_{\mathrm{W}} &= a^2 \mathbf{\Phi}_{\widetilde{\mathbf{y}}}^{-1} E\left(\widetilde{\mathbf{y}}\right) \\ &= a^4 \mathbf{\Phi}_{\widetilde{\mathbf{y}}}^{-1} \mathbf{1}_4 + \sigma^2 \ \beta_{\zeta}, \end{aligned} \tag{19}$$

where

$$\beta_{\zeta} = \left[\begin{array}{cccc} 1 & \zeta/2 & \zeta^{*}/2 & 1 \end{array}\right]^T, \tag{20}$$

$$\zeta = e^{-\jmath 2\pi f \delta \cos\theta_{\mathrm{i},1}/c} + e^{-\jmath 2\pi f \delta \cos\theta_{\mathrm{i},2}/c}. \tag{21}$$

In addition, the magnitude estimators are adapted accordingly. We immediately have

$$\hat{a}\left[\mathbf{h}_{\mathrm{W}}\right] = \sqrt{\max\left\{\frac{|\mathbf{h}_{\mathrm{W}}^H \mathbf{y}|^2 - \sigma^2\left[\Re\{H_{\mathrm{W}}^{*}(1) H_{\mathrm{W}}(2)\zeta^{*}\} - \|\mathbf{h}_{\mathrm{W}}\|^2\right]}{|\mathbf{h}_{\mathrm{W}}^H \mathbf{1}_2|^2}, 0\right\}} \tag{22}$$

and

$$\hat{a}\left[\widetilde{\mathbf{h}}_{\mathrm{W}}\right] = \sqrt{\max\left\{\frac{\widetilde{\mathbf{h}}_{\mathrm{W}}^H\left[\widetilde{\mathbf{y}} - \sigma^2\beta_{\zeta}\right]}{\widetilde{\mathbf{h}}_{\mathrm{W}}^H \mathbf{1}_4}, 0\right\}}, \tag{23}$$

which can be both verified to be real and non-negative. We note that Algorithm 1 applies for the correlated noise case as well by appropriately modifying the expressions for the Wiener beamformers and the estimates of a.

V. Simulations

Let us begin with the uncorrelated noise case. We set $a = 1$, $\phi_a = 0$, and generate $N = 10,000$ independent realizations of V_1 and V_2 drawn from two distinct probability distributions: normal, that is, $V_1, V_2 \sim \mathcal{N}\left(0, \sigma^2\right)$, and exponential, that is, $V_1, V_2 \sim \exp\left(1/\sigma\right)$. We note that the mean value is subtracted from each noise sample of the exponential distribution to form zero-mean samples. We use $N_0 = 500$ realizations to generate estimates of $\mathbf{\Phi}_{\mathbf{y}}$ and $\mathbf{\Phi}_{\widetilde{\mathbf{y}}}$, respectively. Next, we perform the following iterative procedure for each of the beamformers which consists of $I_0 = 5$ iterations (although the simulations clearly indicated that $I_0 = 3$ iterations are enough). First, the beamformer is derived with the latest estimate of a fixed. Then, it is used to generate $10,000$ new estimates for a, out of which the positive estimates are averaged to acquire a single valid estimate. We note that both filters are initialized as identity filters and that the initial magnitude is $a_0 = 5$.

We repeat this experiment for varying values of the broadband input SNR from -20 dB to 20 dB, where it is defined by

$$\begin{aligned} \mathrm{iSNR} &= \frac{\int_f \phi_X(f) df}{\int_f \phi_{V_1}(f) df} \\ &= \frac{a^2}{\sigma^2} \end{aligned} \tag{24}$$

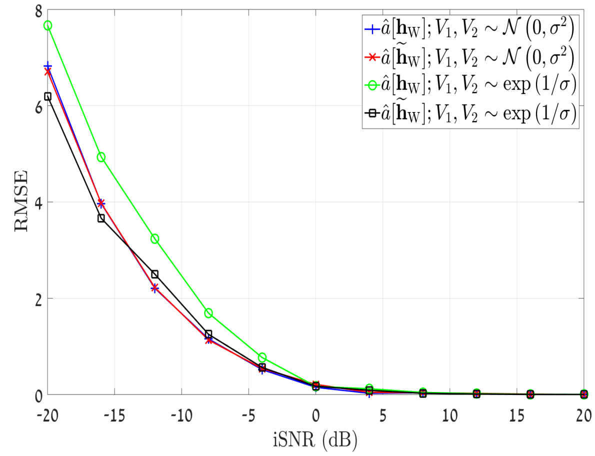


Fig. 1: Magnitude RMSE with the linear and quadratic Wiener beamformers, $\mathbf{h}_{\mathrm{W}}$ and $\widetilde{\mathbf{h}}_{\mathrm{W}}$, with two types of uncorrelated additive noise: normally- and (zero-mean) exponentially-distributed.

and employ the aforementioned RMSE defined by

$$\mathrm{RMSE}\left[\mathbf{h}_{\mathrm{W}}\right] = \sqrt{E\left|a - \hat{a}\left[\mathbf{h}_{\mathrm{W}}\right]\right|^2}, \tag{25}$$

$$\mathrm{RMSE}\left[\widetilde{\mathbf{h}}_{\mathrm{W}}\right] = \sqrt{E\left|a - \hat{a}\left[\widetilde{\mathbf{h}}_{\mathrm{W}}\right]\right|^2}, \tag{26}$$

as the performance measure. The RMSE values as a function for the input SNR are shown in Fig. 1. We observe that for high input SNRs, the RMSE converges to zero, with both estimators and noise distributions. For low input SNRs and normally-distributed noise, both estimators perform the same. This results from the fact that with normally-distributed noise the latent information in higher-order moments is limited. For example, the 3rd-order moment is strictly zero. On the contrary, with the exponentially-distributed noise, the RMSE with the quadratic Wiener beamformer is lower than with the linear Wiener beamformer, with the performance gap reducing as the input SNR increases.

We now turn to the correlated noise case. We maintain the same simulation settings of the uncorrelated noise case and generate samples according to the model in (14). We set $f = 4$ kHz, $\delta = 5$ mm, $\theta_{\mathrm{i},1} = 0^o$, and $\theta_{\mathrm{i},2} = 180^o$. We examine the RMSEs with the two Wiener beamformers and their respective beampatterns (power beampattern with $\widetilde{\mathbf{h}}_{\mathrm{W}}$). The results are depicted in Figs 2 and 3, respectively. We observe that the RMSE with $\widetilde{\mathbf{h}}_{\mathrm{W}}$ is significantly lower than with $\mathbf{h}_{\mathrm{W}}$, with the former achieving a practically zero RMSE for input SNRs higher than 0 dB. As before, for high input SNRs, the RMSE converges to zero with both beamformers. Addressing the beampatterns, we note that while $\mathbf{h}_{\mathrm{W}}$ exhibits a constant "all-pass" pattern, $\widetilde{\mathbf{h}}_{\mathrm{W}}$ exhibits an over-40 dB attenuation in both $\theta_{\mathrm{i},1}$ and $\theta_{\mathrm{i},2}$, whereas θ_{d} remains distortionless. Clearly, such a performance cannot be obtained using a linear beamformer of length $M = 2$.

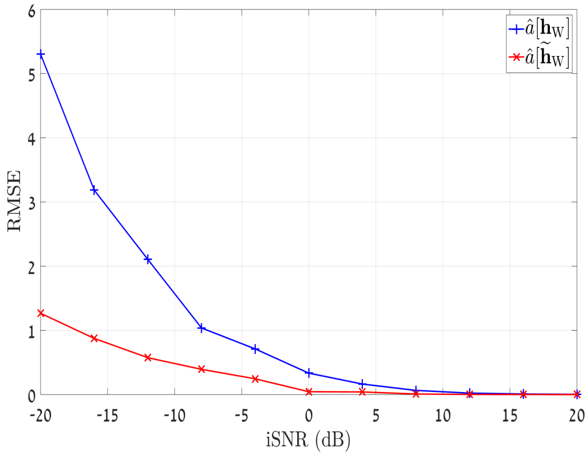
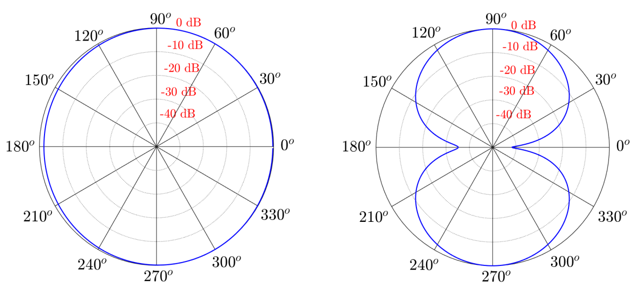


Fig. 2: Magnitude RMSE with the linear and quadratic Wiener beamformers, $\mathbf{h}_{\mathrm{W}}$ and $\widetilde{\mathbf{h}}_{\mathrm{W}}$, with two normally-distributed directional interferences.



(a) (b)

Fig. 3: Beampattern of the linear Wiener beamformer $\mathbf{h}_{\mathrm{W}}$ and a power beampattern of the quadratic Wiener beamformer $\widetilde{\mathbf{h}}_{\mathrm{W}}$, with two normally-distributed directional interferences at $\theta_{\mathrm{i},1} = 0^o$ and $\theta_{\mathrm{i},2} = 180^o$. (a) $\mathbf{h}_{\mathrm{W}}$ and (b) $\widetilde{\mathbf{h}}_{\mathrm{W}}$.

VI. Conclusions

We have introduced an optimal quadratic Wiener beamformer for a desired signal magnitude estimation which utilizes information from higher-order moments. To simplify the formulation, we assumed a two-microphone array, but the generalization to any array is straightforward. We developed an iterative algorithm for magnitude estimation based on a quadratic multichannel noise reduction approach, and addressed two types of additive noise: uncorrelated and correlated. For each noise type, we derived a quadratic version of the Wiener beamformer and a respective unbiased magnitude estimator, and compared their performances to the linear versions of the Wiener beamformer. With uncorrelated noise, we have shown that the quadratic magnitude estimator yields a lower RMSE with respect to the linear estimator in low input SNRs, in case the noise is exponentially distributed. When the noise is drawn from a normal distribution, both estimators perform equally. With correlated noise, we have shown that the quadratic beamformer eliminates two spatial directions, as opposed to a single direction with any linear beamformer. For low input SNRs, this resulted in a significantly lower RMSE using the quadratic beamformer.

References

[1] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*, Simon and Schuster, Inc., USA, 1992.

[2] H.L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*, Detection, Estimation, and Modulation Theory. Wiley, 2004.

[3] J. Benesty, I. Cohen, and J. Chen, *Fundamentals of Signal Enhancement and Array Signal Processing*, Wiley-IEEE Press, 1st edition, 2018.

[4] J. Benesty, I. Cohen, and J. Chen, *Array Processing - Kronecker Product Beamforming*, Springer-Verlag, Switzerland, 2019.

[5] K. Paliwal and A. Basu, "A speech enhancement method based on kalman filtering," in *ICASSP '87. IEEE International Conference on Acoustics, Speech, and Signal Processing*, April 1987, vol. 12, pp. 177–180.

[6] I. Cohen and B. Berdugo, "Speech enhancement for non-stationary noise environments," *Signal Processing*, vol. 81, no. 11, pp. 2403 – 2418, 2001.

[7] I. Cohen and S. Gannot, *Spectral Enhancement Methods*, pp. 873–901, Springer Berlin Heidelberg, 2008.

[8] J. Benesty, J. Chen, and E. Habets, *Speech Enhancement in the STFT Domain*, Springer-Verlag Berlin Heidelberg, Berlin, 2012.

[9] R. McAulay and M. Malpass, "Speech enhancement using a soft-decision noise suppression filter," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 28, no. 2, pp. 137–145, Apr 1980.

[10] Y. Ephraim and D. Malah, "Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 32, no. 6, pp. 1109–1121, Dec 1984.

[11] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error log-spectral amplitude estimator," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 2, pp. 443–445, Apr 1985.

[12] P.J. Wolfe and J.S. Godsill, "Efficient Alternatives to the Ephraim and Malah Suppression Rule for Audio Signal Enhancement," *EURASIP Journal on Advances in Signal Processing*, vol. 2003, no. 10, Sep 2003.

[13] J. S. Erkelens, R. C. Hendriks, R. Heusdens, and J. Jensen, "Minimum mean-square error estimation of discrete fourier coefficients with generalized gamma priors," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 6, pp. 1741–1752, Aug 2007.

[14] Richard C. Hendriks, Heusdens Richard, and Jesper Jensen, "Log-spectral magnitude mmse estimators under super-gaussian densities.," 01 2009, pp. 1319–1322.

[15] Kuldip Paliwal, Kamil Wójcicki, and Belinda Schwerin, "Single-channel speech enhancement using spectral subtraction in the short-time modulation domain," *Speech Communication*, pp. 450–475, 2010.

[16] Y. Wang and M. Brookes, "Model-based speech enhancement in the modulation domain," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 3, pp. 580–594, March 2018.

[17] G. Itzhak, J. Benesty, and I. Cohen, "Quadratic approach for single-channel noise reduction," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2020, 2020.

[18] G. Itzhak, J. Benesty, and I. Cohen, "Nonlinear kronecker product filtering for multichannel noise reduction," *Speech Communication*, vol. 114, pp. 49 – 59, 2019.

[19] J. Benesty, J. Chen, and Y. Huang, *Microphone Array Signal Processing*, Springer-Verlag Berlin Heidelberg, 2008.

[20] J. Chen, J. Benesty, Y. Huang, and S. Doclo, "New insights into the noise reduction wiener filter," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, no. 4, pp. 1218–1234, July 2006.

[21] P. C. Loizou, *Speech Enhancement: Theory and Practice*, CRC Press, Inc., Boca Raton, FL, USA, 2nd edition, 2013.