Abstract—Fixed beamformers maintain response that is independent of the signal and interference statistics. Frequency-invariant fixed beamformers can achieve constant beamwidth across all frequencies which results in lower signal distortion and they have lower computation complexity compared to its adaptive counterpart. However, unlike data-dependent beamformers, their sidelobe attenuation is poor with respect to the direction of the interferences. In this paper we propose a method to improve the sidelobe attenuation while retaining the advantages of lower computational complexity and frequency-invariance. This is achieved by introducing frequency-invariant nulls in the interference directions. We also show how the weights for each null direction can be combined with the fixed beamformer to form the effective weights of the proposed beamformer.

I. INTRODUCTION

Microphone arrays are widely used in teleconferencing systems and hands-free telephony for source localization, speech enhancement and speaker recognition applications [1]-[5]. Beamforming algorithms employing multi-channel microphone arrays have been widely used for speech enhancement. These algorithms can broadly be divided into adaptive and fixed beamforming methods. Among the adaptive methods, the minimum variance distortionless response (MVDR) (also known as the Capon beamformer [6]), linearly constrained minimum variance (LCMV) [7] and generalized sidelobe canceler (GSC) [8] are the most popular. Since adaptive beamformers are data dependent, the beamformer weights have to be continuously updated based on the received signals.

Fixed beamformers, on the other hand, represent a class of beamformers which maintain a constant response for a particular look direction independent of the signal and noise/interference statistics. Compared to adaptive beamformers, although they achieve a lower signal-to-noise ratio (SNR), fixed beamformers are attractive in terms of computational complexity and ease of real-time implementation since the filter weights can be pre-calculated and saved in a lookup table. For a fixed array aperture, the bandwidth of a beamformer is inversely proportional to the signal frequency. As such, direct implementation of a fixed beamformer can introduce low-pass filter distortion due to non-uniform gain, particularly for large bandwidth signals such as speech.

Frequency-invariant fixed beamformers (FIBs) provide constant frequency response with uniform beamwidth across all frequencies. The design of FIB using different techniques for different microphone array configuration have been studied in great detail. In [9], the frequency-invariance property was obtained using the technique of harmonic nesting. While this approach ensures octave-independent beampattern, the spatial resolution within each octave is frequency dependent. To achieve frequency-invariance within each octave, a combination of harmonic nesting and filter-and-sum beamforming has been proposed in [10]. Both these methods focus on linear arrays. In [11] and [12], FIB design was studied for circular arrays and further extended to concentric circular and spherical arrays. A general FIB design for multi-dimensional arrays using separate primary filters for each array element and a common secondary filter has been proposed in [13]. The design is complex when applied for higher-dimensional arrays since the dilation property of the primary filters is not guaranteed. A relatively simpler design using multi-dimensional inverse Fourier transforms has been discussed in [14], [15], [16]. Recently, a direct optimization approach for the FIB design using the technique of convex optimization has been proposed in [17], [18]. To achieve higher computational efficiency, closed-form solutions for the FIB design using the least squares and eigenfilter methods have been discussed in [19], [20] and [21].

The inherent disadvantage of FIB is the inadequate sidelobe attenuation, which limits its applications for interference cancellation. In this paper, we achieve good sidelobe attenuation by combining multiple frequency-invariant null (FIN) beamformers along with the fixed beamformer. We show how the filter coefficients corresponding to the frequency-invariant fixed and null beamformers are combined to obtain the effective filter coefficients.

The outline of this paper is as follows: Section II gives a brief review of the general wideband beamformer structure. Section III discusses the least square solution for the design of FIB, while the proposed method is discussed in Section IV. Section V presents the simulation results and conclusions are drawn in Section VI.

II. REVIEW OF WIDEBAND BEAMFORMING

The general structure of a wideband beamformer using tapped-delay lines (TDL) is shown Fig. 1, where \( J \) is the number of filter taps associated with each of the \( M \) microphones. Such a beamformer samples the signals both spatially and temporally and its response can be expressed as a function
of the frequency \( f \) and the azimuth angle \( \theta \) given by

\[
P(f, \theta) = \sum_{m=0}^{M-1} \sum_{j=0}^{J-1} w_{m,j} e^{-j2\pi f (\tau_m(\theta) + jT_s)},
\]

where \( * \) denotes the conjugate operation, \( \tau_m(\theta) \) is the spatial propagation delay between the \( m \)th microphone and the reference microphone and \( T_s \) is the time delay between the adjacent taps of the TDL. The vector form of (1) can be written as

\[
P(f, \theta) = \mathbf{w}^H \mathbf{s}(f, \theta),
\]

where

\[
\mathbf{w} = [w_{0,0}, \ldots, w_{M-1,0}, \ldots, w_{0,J-1}, \ldots, w_{M-1,J-1}]^T
\]

is the (complex) filter coefficient vector. The variable \( \mathbf{s}(f, \theta) \) is the \( MJ \times 1 \) steering vector defined as

\[
\mathbf{s}(f, \theta) = \mathbf{s}_{T_r}(f) \otimes \mathbf{s}_{r}(f, \theta),
\]

such that \( \otimes \) denotes the Kronecker product,

\[
\mathbf{s}_{T_r}(f) = \begin{bmatrix} e^{-j2\pi f T_s}, \ldots, e^{-j(J-1)2\pi f T_s} \end{bmatrix}^T,
\]

and

\[
\mathbf{s}_{r}(f, \theta) = \begin{bmatrix} e^{-j2\pi f \tau_m(\theta)}, e^{-j2\pi f \tau_1(\theta)}, \ldots, e^{-j2\pi f \tau_{M-1}(\theta)} \end{bmatrix}^T.
\]

For the case of a uniformly spaced linear microphone array with spacing \( d \) and wave propagation velocity \( c \), the time delay is given by \( \tau_m(\theta) = \frac{d}{2c} \cos \theta \).

III. FIB BASED ON LEAST SQUARES

The general weighted least squares cost function for the design of a wideband beamformer is given by [22],

\[
J_{LS} = \int \int F(f, \theta)|P(f, \theta) - P_d(f, \theta)|^2 df d\theta,
\]

where \( F \) and \( \Theta \) denote the frequency range of interest and azimuth angles, respectively. The variable \( F(f, \theta) \) is a positive real-valued weighting function corresponding to the mainlobe and sidelobes while \( P(f, \theta) \) and \( P_d(f, \theta) \) define the actual and desired directional responses, respectively. By selecting a uniform grid of frequencies and azimuth angles, we assume \( P_d(f, \theta) = 1 \) and \( F(f, \theta) = 1 \) for the mainlobe and \( P_d(f, \theta) = 0 \) and \( F(f, \theta) = \alpha \), where \( 0 < \alpha < 1 \) for the sidelobes. Defining \( f_i \) as frequency for the \( i \)th frequency index and \( \Theta_k \) as the azimuth for the \( k \)th azimuth index, the cost function in (7) can then be rewritten, using (2), as

\[
J_{LS} = \sum_{f_i \in F} \sum_{\theta_k \in \Theta_m} |\mathbf{w}^H \mathbf{s}(f_i, \Theta_k) - 1|^2 + \alpha \sum_{f_i \in F, \theta_k \in \Theta_s} |\mathbf{w}^H \mathbf{s}(f_i, \Theta_k)|^2,
\]

where \( \Theta_m \) and \( \Theta_s \) denote the azimuth range corresponding to mainlobe and sidelobes, respectively.

To achieve frequency-invariant property, the ‘response variation’ criterion has been proposed [17]. The criterion

\[
J_{rv} = \sum_{f_i \in F} \sum_{\theta_k \in \Theta_m} |\mathbf{w}^H \mathbf{s}(f_i, \Theta_k) - \mathbf{w}^H \mathbf{s}(f_i, \Theta_s)|^2,
\]

defines the Euclidean distance between the response at the reference frequency \( f_i \) and that of all frequencies over the range of azimuth angles \( \Theta_F \). This cost function along with (8) offers a tradeoff between sidelobe attenuation and the frequency-invariance property. This tradeoff can be varied through the use of a control parameter \( \beta \) in the following FIB cost function

\[
J_{FIB} = J_{LS} + \beta J_{rv}.
\]

It is useful to note that the frequency-invariant property can be obtained either in the mainlobe direction or the full azimuth range. If it is considered over the whole azimuth range, i.e., including the sidelobe region, the spectrum energy of the beamformer must be minimized only at the reference frequency \( f_i \) instead of over the entire frequency range.

Substituting \( J_{LS} \) from (8) to (10), we obtain the FIB cost function

\[
J_{FIB} = \sum_{\theta_k \in \Theta_m} |\mathbf{w}^H \mathbf{s}(f_i, \Theta_k) - 1|^2 + \alpha \sum_{\theta_k \in \Theta_s} |\mathbf{w}^H \mathbf{s}(f_i, \Theta_k)|^2 + \beta J_{rv}.
\]

The above cost function is then solved using the Lagrange multiplier method to obtain the optimum solution [20]

\[
\mathbf{w} = \mathbf{Q}_{FIB}^{-1} \mathbf{a}_{FIB},
\]
FIB is that the sidelobe attenuation may not be sufficient as shown in Fig. 2. These nulls can be achieved by minimizing the spectrum estimation of the effective filter coefficients corresponding to FINs and that of FIB.

where

$$Q_{\text{FIB}} = \sum_{\theta_k} s(f_s, \theta_k) s^H(f_s, \theta_k) + \alpha \sum_{\theta_k} s(f_s, \theta_k) s^H(f_s, \theta_k)$$

$$+ \beta \sum_{f_s, \theta_k} (s(f_s, \theta_k) - s(f_s, \theta_k))(s(f_s, \theta_k) - s(f_s, \theta_k))^H$$

and

$$a_{\text{FIB}} = \sum_{\theta_k} s(f_s, \theta_k).$$

The drawback of direct implementation of FIB for speech signals lies in the fact that it is challenging to achieve a frequency-invariant response at low frequencies where most of the speech energy is present. This drawback is particularly pronounced when the array aperture is small [23]. When the frequency-invariant property is considered for the whole speech band, the mainlobe beamwidth will be large and the frequency-invariant response at low frequencies where most of the speech energy is present. This drawback is particularly pronounced when the array aperture is small [23]. When the frequency-invariant property is considered for the whole speech band, the mainlobe beamwidth will be large and the frequency-invariant nulls (FIN) are combined with the FIB so that the effective beamformer can achieve unity gain in the direction of the desired source and nulls in the direction of the interferences.

For the estimation of interfering source directions one may use any of the existing methods [25], [26]. For example, by applying a complex-valued independent component analysis (ICA) algorithm such as the FastICA [27] in any one of the frequency bins, we can estimate the unmixing matrix based on which the rows can then be used to estimate the direction(s) of the source(s) [28], [29]. In this work however, we assume that the directions of the interferers are known. In the following sections, we describe the design of the broadband nulls and estimation of the effective filter coefficients by combining the filter coefficients corresponding to FINs and that of FIB.

IV. THE PROPOSED CFIBN APPROACH

As described in Section III, the main disadvantage of the FIB is that the sidelobe attenuation may not be sufficient to suppress the interferences. On the other hand, data-dependent beamformers such as the MVDR beamformer will form nulls in the direction of the interferences and this results in the suppression of interfering signals. To overcome the above disadvantage of the FIB, we propose to design a combined frequency-invariant fixed beamformer with null (CFIBN) in the direction of interferences. In our proposed approach, the cost function for the null design is then given by minimizing

$$J_{\text{NULL}} = \sum_{\theta_k} s(f_s, \theta_k)^2 + \beta J_{\text{TV}},$$

subject to

$$\sum_{\theta_k} w^H s(f_s, \theta_k) = 1, \quad \Theta_{\text{null}} \not\subset \Theta,$$

where $\Theta_{\text{null}}$ is the desired null direction which corresponds to the direction of the interference. The above cost function can be solved using the Lagrange multiplier method which results in

$$w = R^{-1} p / p^H R^{-1} p.$$
where

\[
R = \sum_{\theta_k \in \Theta_{null}} s(f_r, \theta_k) s^H(f_r, \theta_k) + \beta \sum_{f_r \in \mathcal{F}} \sum_{\theta_k \epsilon \Theta} (s(f_r, \theta_k) - s(f_r, \theta_k)) (s(f_r, \theta_k) - s(f_r, \theta_k))^H
\]

and

\[
p = \sum_{\theta_k \epsilon \Theta} s(f_r, \theta_k), \quad \Theta_{null} \not\subset \Theta.
\]

Fig. 4 shows an example for a frequency-invariant broadband null for an azimuth angle of 30° based on the above design, using a uniform linear microphone array of twenty microphones with an inter-microphone spacing of 2.14 cm.

### B. Effective filter coefficients

After obtaining the beamformer filter coefficients corresponding to the look direction and null directions, they have to be combined to obtain the effective filter coefficients. Denoting the weight vectors corresponding to the FIB by \( \mathbf{w}_{\text{FIB}} \) and that of the \( N \) FINs by \( \mathbf{w}_{\text{FIN}_1}, \mathbf{w}_{\text{FIN}_2}, \ldots, \mathbf{w}_{\text{FIN}_N} \), the response of the equivalent beamformer having unity response in the desired direction and nulls in the directions of the \( N \) interferences can be written as

\[
\mathbf{w}_{\text{CPFIN}}^H \mathbf{s}(f, \theta) = \mathbf{w}_{\text{FIB}}^H \mathbf{s}(f, \theta) \circ \mathbf{w}_{\text{FIN}_1}^H \mathbf{s}(f, \theta) \circ \mathbf{w}_{\text{FIN}_2}^H \mathbf{s}(f, \theta) \circ \cdots \circ \mathbf{w}_{\text{FIN}_N}^H \mathbf{s}(f, \theta),
\]

where \( \circ \) denotes the Hadamard product and \( \mathbf{w}_{\text{CPFIN}} \) is the effective filter coefficients of the equivalent beamformer. It is important to note that before computing the equivalent weights, beamformer responses of FIB and FINs have to be normalized using corresponding filter coefficients. This normalization ensures that there is equal weightage for individual beamformers.

Following that of (20), the effective filter coefficients can then be calculated as

\[
\mathbf{w}_{\text{CPFIN}} = \left[ \mathbf{w}_{\text{FIB}}^H \mathbf{s}(f, \theta) \circ \mathbf{w}_{\text{FIN}_1}^H \mathbf{s}(f, \theta) \circ \mathbf{w}_{\text{FIN}_2}^H \mathbf{s}(f, \theta) \circ \cdots \circ \mathbf{w}_{\text{FIN}_N}^H \mathbf{s}(f, \theta) \right] \left[ \mathbf{s}(f, \theta) \mathbf{s}^H(f, \theta) \right]^{-1}.
\]

Applying the above constraint to all the frequencies and directions, (21) can be expressed as

\[
\mathbf{w}_{\text{CPFIN}} = \left[ \mathbf{w}_{\text{FIB}}^H \mathbf{S}(f, \theta) \circ \mathbf{w}_{\text{FIN}_1}^H \mathbf{S}(f, \theta) \circ \mathbf{w}_{\text{FIN}_2}^H \mathbf{S}(f, \theta) \circ \cdots \circ \mathbf{w}_{\text{FIN}_N}^H \mathbf{S}(f, \theta) \right] \left[ \mathbf{S}(f, \theta) \mathbf{S}^H(f, \theta) \right]^{-1},
\]

where

\[
\mathbf{S}(f, \theta) = [s(f_1, \theta_1), s(f_1, \theta_2), \cdots, s(f_1, \theta_K), s(f_2, \theta_1), s(f_2, \theta_2), \cdots, s(f_2, \theta_K), \cdots, s(f_I, \theta_K)]^T
\]

given that \( I \) is the number of frequency bins and \( K \) is the number of directions.

The advantage of the above design compared to data dependent beamformers such as the MVDR is in its practical implementation. We can pre-calculate the filter coefficients \( \mathbf{w}_{\text{FIB}} \) and \( \mathbf{w}_{\text{FIN}} \) for all the possible directions and the matrix \( \left( \mathbf{S}(f, \theta) \mathbf{S}^H(f, \theta) \right)^{-1} \) so that they can be combined together for any combination of source and interference directions without the need for a significant increase in computational complexity. In addition, unlike data-dependent methods, the proposed method does not require the estimation of noise covariance matrix, which can often be challenging under low SNR environment.

### V. Simulation Results

To verify the performance of the proposed approach, a uniform linear microphone array with \( M = 20 \), each employing a \( J = 300 \) tap finite impulse response (FIR) filter is considered for the design of both FIB and FIN in our simulations. We have also used a sampling frequency of 16 kHz. The spacing between microphones \( d \) is set at half the wavelength corresponding to maximum signal frequency and this corresponds to 2.14 cm. The look direction is kept at 90° and the direction of the interferers are set to 30° and 150°. The sidelobe weighting parameter \( \alpha \) and the tradeoff parameter \( \beta \) is set at 0.01 and 0.95, respectively. The frequency of interest is assumed to range from 100 Hz to 3200 Hz, with the reference frequency \( f_r \) set at 1700 Hz.

The beampattern and the polar pattern obtained using the proposed method is shown in Fig. 5 and 6, respectively. Compared with the beampattern and polar pattern of the conventional FIB shown in Fig. 2 and 3, it can be seen that the interferences from 30° and 150° are attenuated significantly, while maintaining sufficient frequency-invariant response in the desired direction. It is worth noting that, unlike for the case of MVDR in Fig. 7 and 8, for FIB in Fig. 2, even though there are nulls in the sidelobes, they are not in the direction of
the interferences. This is due to the fact that the interference direction information were not used for the design of FIB. By incorporating this information, one could significantly improve the attenuation in the direction of interferences. In addition, the beampattern of the MVDR is non-uniform across frequencies and this results in distortion when applied for speech enhancement applications. It is also important to note that the MVDR assumes that the interference covariance matrix is known. Knowledge of this covariance matrix is challenging in practice particularly in an environment where the SNR is low.

**VI. CONCLUSIONS**

We proposed a method for improving the sidelobe attenuation of a frequency-invariant fixed beamformer by incorporating nulls in the interference directions. The proposed method involves the design of frequency-invariant fixed beamformers and broadband nulls as well as the pre-calculation of the beamformer filter coefficients for all the possible directions. Frequency-invariant fixed beamformers with nulls in the direction of interferences and unity response in the direction of the desired source can then be formed by combining the above pre-calculated beamformer filter coefficients. This approach does not require significant computational complexity and hence the method is suitable for real-time applications.

**REFERENCES**


