On the average directivity factor attainable with a beamformer incorporating null constraints

Dovid Y. Levin, Emanuël A.P. Habets, *Senior member, IEEE,* and Sharon Gannot, *Senior member, IEEE*

***Abstract*—The directivity factor (DF) of a beamformer describes its spatial selectivity and ability to suppress diffuse noise which arrives from all directions. For a given array configuration, it is possible to design beamforming weights which maximize the DF for a particular look-direction, while enforcing nulls for a set of undesired directions. In general, the resulting DF is dependent upon the specific look- and null directions. Using the same array, one may apply a different set of weights designed for any other feasible set of look- and null directions. In this contribution we show that when the optimal DF is averaged over all look directions the result equals the number of sensors minus the number of null constraints. This result holds, regardless of the positions and spatial responses of the individual sensors, and of the null directions. The result generalizes to more complex wave-propagation domains (e.g., reverberation).**

I. Introduction

A beamformer produces a weighted combination of signals corresponding to different elements. The weights may be designed so as to aid the reception (or transmission) of signals from a certain direction (the *look direction*), and to block certain other undesired directions. Applications of beamforming are widespread [1] and encompass such areas as radar, sonar, communications, seismology, astronomy, oceanography, medical tomography, and acoustic signal processing.

The directivity factor (DF) of a beamformer describes its spatial selectivity and ability to suppress diffuse noise which arrives from all directions. For a given array constellation, it is possible to select weights which maximize the DF for a particular look direction. The criteria for selecting weights may also incorporate null constraints to block undesired directions. In general, the addition of constraints will reduce the DF.

Although the array constellation is usually fixed and cannot be easily changed, the array weights may be readily modified. Thus, a single array can be used for multiple scenarios with different look- and null directions. The DF attained when using optimal weights is dependent on these directions.

Previous work [2]–[4] has shown that when the optimal DF is averaged over all look directions, the result equals the number of array elements. In this contribution, we extend this result to incorporate null constraints and prove that the ensuing average equals the number of array elements minus the number of null constraints. This property holds regardless of the selection of array constellation and is generalizable to more complex wave propagation domains, e.g. sound propagation in reverberant enclosures.

D.Y. Levin and S. Gannot are with the Faculty of Engineering, Bar-Ilan University, Ramat-Gan, Israel. (e-mail: david.levin@live.biu.ac.il ; sharon.gannot@biu.ac.il).

E.A.P. Habets is with the International Audio Laboratories Erlangen (a joint institution between the University of Erlangen-Nuremberg and Fraunhofer IIS), Erlangen, Germany (e-mail: emanuel.habets@audiolabs-erlangen.de).

II. Background and notation

The theory of beamforming relates to both receiving and transmitting arrays. Without loss of generality, we adopt the vantage point of a receiving array.

A. Steering vectors and beamforming

Let us consider an array consisting of N sensors receiving a signal from a source. A *steering vector* (of dimensions $N \times 1$) describes the relationship between a source signal and the signals received at the sensors. Different factors affect the amplitude and phase values of the steering vector. For example when a plane wave impinges on an array, the positions of the sensors, the direction of arrival (DOA), and the beampatterns of the individual sensors all have an impact. The resulting steering vector is

$$\mathbf{v}(\mathbf{u},\omega) = \mathbf{b}(\mathbf{u},\omega) \odot \exp\{j\,(\omega/c)\cdot \mathbf{P}^T\mathbf{u}\}\,, \qquad (1)$$

where the DOA is denoted by the unit-vector $\mathbf{u}$ and ω is the angular frequency. The beampatterns of the individual sensors are contained in the $N \times 1$ vector $\mathbf{b}(\mathbf{u},\omega)$. $\mathbf{P}$ is a $3 \times N$ matrix denoting the positions of the sensors (in Cartesian coordinates), and c is the velocity of wave propagation; hence, $\exp\{j\,(\omega/c)\cdot \mathbf{P}^T\mathbf{u}\}$ describes the phase change due to wave propagation [1]. The operator $\odot$ represents the Hadamard (element-wise) product. We note that (1) portrays a fairly straightforward scenario; more complex cases may incorporate near-field effects and multi-path, e.g. reverberation.

A practical scenario may incorporate multiple sources as well as noise. Each source has a distinct corresponding steering vector. From superposition, the received signals are

$$\mathbf{x}(\omega) = \sum_{k=1}^{K} \mathbf{v}_k s_k(\omega) + \mathbf{n}(\omega)\,, \qquad (2)$$

where $\mathbf{x}(\omega) = [x_1(\omega)\ldots x_N(\omega)]^T$ contains the received signals, $s_k(\omega)$ is the signal produced by the k-th source, $\mathbf{v}_k$ is the corresponding steering vector, and $\mathbf{n}(\omega)$ is noise.

The beamformer produces an output signal by performing a weighted sum of the input channels

$$y(\omega) = \mathbf{w}^H(\omega)\mathbf{x}(\omega)\,, \qquad (3)$$

where $\mathbf{w}(\omega) = [w_1(\omega)\ldots w_N(\omega)]^T$ contains the weights corresponding to each sensor. Substitution of (2) into (3) shows that the respective signals are scaled by a factor of $\mathbf{w}^H(\omega)\mathbf{v}_k$. The response of the beamformer to a signal from an arbitrary direction is described by the beampattern

$$\mathrm{BP}(\mathbf{u},\omega) = \mathbf{w}(\omega)^H\mathbf{v}(\mathbf{u},\omega)\,. \qquad (4)$$

Similarly, the beam-power is defined as $|\mathrm{BP}(\mathbf{u},\omega)|^2$. The level of the noise at the output is given by[1]

$$E\{|\mathbf{w}^H\mathbf{n}|^2\} = \mathbf{w}^H E\{\mathbf{n}\mathbf{n}^H\}\mathbf{w} = \mathbf{w}^H\mathbf{R}\mathbf{w}\,, \tag{5}$$

where $E\{\cdot\}$ represents statistical expectation and $\mathbf{R} = E\{\mathbf{n}\mathbf{n}^H\}$ is the noise covariance matrix.

B. LCMV beamforming

It may be desirable to constrain the responses towards the K source signals to certain prespecified values, while minimizing the noise level. This goal can be attained by the linearly constrained minimum variance (LCMV) beamformer which solves

$$\mathbf{w}_{\text{LCMV}} = \arg\min_{\mathbf{w}} \mathbf{w}^H\mathbf{R}\mathbf{w} \quad \text{s.t.} \quad \mathbf{w}^H\mathbf{V} = \mathbf{c}^T\,, \tag{6}$$

where $\mathbf{V}$ is a matrix whose columns consist of the K steering vectors $\mathbf{V} = [\mathbf{v}_1 \cdots \mathbf{v}_K]$, and $\mathbf{c}$ is a column vector whose elements represent the desired responses. It is assumed that $\mathbf{V}$ is formulated to contain $K \le N$ columns which are linearly independent. This precludes overdetermined systems and rules out conflicting or redundant constraints. The solution to (6) is [5]

$$\mathbf{w}_{\text{LCMV}} = \mathbf{R}^{-1}\mathbf{V}(\mathbf{V}^H\mathbf{R}^{-1}\mathbf{V})^{-1}\mathbf{c}^*\,. \tag{7}$$

Substituting (7) into (5), the beamformer's noise level is

$$\mathbf{w}_{\text{LCMV}}^H\mathbf{R}\,\mathbf{w}_{\text{LCMV}} = \mathbf{c}^T(\mathbf{V}^H\mathbf{R}^{-1}\mathbf{V})^{-1}\mathbf{c}^*\,. \tag{8}$$

C. The directivity factor

The DF of a beamformer describes the ratio of beam-power in the look direction relative to the average beam-power over all directions. A high DF corresponds to sharper and more focused beams. For a beamformer with a set of weights $\mathbf{w}$, the spherical DF is defined [1], [6] as:

$$\mathrm{DF}_{\text{sph}}(\mathbf{w},\mathbf{u}_0) = \frac{|\mathbf{w}^H\mathbf{v}(\mathbf{u}_0)|^2}{\frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi}|\mathbf{w}^H\mathbf{v}(\mathbf{u})|^2\,\sin(\theta)d\theta\,d\phi}\,, \tag{9}$$

where the DOA $\mathbf{u} = [\cos(\phi)\sin(\theta),\,\sin(\phi)\sin(\theta),\,\cos(\theta)]^T$ is specified by azimuth and inclination parameters. The spherical DF of (9) is the standard definition. On occasion, only DOAs within the *xy*-plane (i.e. $\theta = \frac{\pi}{2}$) are of interest. In this case, an alternative version – cylindrical DF – is defined [6]:

$$\mathrm{DF}_{\text{cyl}}(\mathbf{w},\mathbf{u}_0) = \frac{|\mathbf{w}^H\mathbf{v}(\mathbf{u}_0)|^2}{\frac{1}{2\pi}\int_0^{2\pi}|\mathbf{w}^H\mathbf{v}(\mathbf{u})|^2\,d\phi}\,. \tag{10}$$

Both (9) and (10) subscribe to the general form (proposed by the authors in [4]):

$$\mathrm{DF}_{\text{gen}}(\mathbf{w},\mathbf{q}_0) = \frac{|\mathbf{w}^H\mathbf{v}(\mathbf{q}_0)|^2}{\kappa\int_{\mathbf{q}\in\mathcal{Q}}|\mathbf{w}^H\mathbf{v}(\mathbf{q})|^2\,A(\mathbf{q})d\mathbf{q}}\,, \tag{11}$$

where the steering-vectors are specified by the parameter vector $\mathbf{q}$, integration is performed over some set of interest $\mathcal{Q}$ in the parameter space, and $A(\mathbf{q})$ is a function denoting the weight for each instance of $\mathbf{q}$ [such as the Jacobian $\sin(\theta)$ of (9), or some measure of likelihood]. The normalizing constant κ is defined as $1/\int_{\mathbf{q}\in\mathcal{Q}}A(\mathbf{q})d\mathbf{q}$. Equation (11) is used in Sec. V to formulate generalizations of the DF.

The denominator of (11) can be expressed as

$$\mathbf{w}^H\left(\kappa\int_{\mathbf{q}\in\mathcal{Q}}\mathbf{v}(\mathbf{q})\mathbf{v}^H(\mathbf{q})\,A(\mathbf{q})d\mathbf{q}\right)\mathbf{w} = \mathbf{w}^H\mathbf{\Phi}\mathbf{w}\,, \tag{12}$$

where $\mathbf{\Phi}$ is defined as

$$\mathbf{\Phi} = \kappa\int_{\mathbf{q}\in\mathcal{Q}}\mathbf{v}(\mathbf{q})\mathbf{v}^H(\mathbf{q})\,A(\mathbf{q})d\mathbf{q} \tag{13}$$

and is assumed to be nonsingular (since otherwise there is a redundancy of sensors). Substituting (12) into (11) yields

$$\mathrm{DF}_{\text{gen}}(\mathbf{w},\mathbf{q}_0) = \frac{|\mathbf{w}^H\mathbf{v}(\mathbf{q}_0)|^2}{\mathbf{w}^H\mathbf{\Phi}\mathbf{w}}\,. \tag{14}$$

The matrix $\mathbf{\Phi}$ which consists of components corresponding to all $\mathbf{q}\in\mathcal{Q}$ can also be seen as describing an isotropic diffuse noise field. Specifically, if the noise component $\mathbf{n}$ is formed by uncorrelated signals arriving from all directions with identical power, the noise covariance in (5) has the form $\mathbf{R}_{\text{diff}} = \sigma^2_{\text{diff}}\mathbf{\Phi}$ where σ^2_{diff} is proportional to the field's strength [4, fn. 5].

III. Maximization of the DF with constraints

In this Section, we derive the optimal value of DF with null constraints. This optimization problem is shown to be analogous to the LCMV beamforming discussed in II-B. We then provide a geometric interpretation of the result.

A. Optimization

We wish to maximize the DF expressed in (14). It should be noted that $\mathbf{w}$ and $a\mathbf{w}$ (for $a \neq 0$) have the same DF. Consequently, we may constrain the numerator to have unity value without affecting the DF. The problem of maximizing (14) is equivalent to minimizing $\mathbf{w}^H\mathbf{\Phi}\mathbf{w}$ subject to the constraint $\mathbf{w}^H\mathbf{v}(\mathbf{q}_0) = 1$. Additional constraints may also be incorporated. Specifically, it may be desirable to place nulls in certain directions corresponding to known interference.

The problem of DF maximization may be formulated in the mold of (6), with the noise matrix $\mathbf{R}$ replaced by $\mathbf{\Phi}$. Additionally, the matrix $\mathbf{V}$ is set such that the first column vector is the desired look direction $\mathbf{v}_1 = \mathbf{v}(\mathbf{q}_0)$ and all subsequent columns consist of the $M = K-1$ steering vectors for directions corresponding to null constraints. For subsequent derivations, it is convenient to partition $\mathbf{V}$ into block form as $\mathbf{V} = [\mathbf{v}_1\ \mathbf{T}]$. The constraint vector becomes $\mathbf{c} = [1\ \mathbf{0}_{1\times M}]^T$ where $\mathbf{0}_{1\times M}$ is a vector of zeros with the specified dimensions.

For the considered scenario, the optimal DF can be obtained from (14) using (8) and the fact that $\mathbf{w}^H\mathbf{v}_1 = 1$:

$$\mathrm{DF}_{\text{opt}} = \frac{1}{[1\ \mathbf{0}_{1\times M}]([\mathbf{v}_1\ \mathbf{T}]^H\mathbf{\Phi}^{-1}[\mathbf{v}_1\ \mathbf{T}])^{-1}[1\ \mathbf{0}_{1\times M}]^T}\,. \tag{15}$$

After expanding the terms in parentheses, (15) can be written as

$$\mathrm{DF}_{\text{opt}} = \frac{1}{\begin{bmatrix}1 & \mathbf{0}_{1\times M}\end{bmatrix}\begin{bmatrix}\mathbf{v}_1^H\mathbf{\Phi}^{-1}\mathbf{v}_1 & \mathbf{v}_1^H\mathbf{\Phi}^{-1}\mathbf{T}\\ \mathbf{T}^H\mathbf{\Phi}^{-1}\mathbf{v}_1 & \mathbf{T}^H\mathbf{\Phi}^{-1}\mathbf{T}\end{bmatrix}^{-1}\begin{bmatrix}1\\ \mathbf{0}_{M\times 1}\end{bmatrix}}\,. \tag{16}$$

[1] In the interest of conciseness, explicit dependence on frequency is dropped from this point onwards.

The net effect of the triple product in the denominator of (16) is to select the first element of the matrix inverse. The matrix inversion in (16) can be expressed using the block inversion rule [7] as

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & \cdot \\ \cdot & \cdot \end{bmatrix}, \tag{17}$$

yielding

$$\mathrm{DF}_{\mathrm{opt}} = \mathbf{v}_1^H \mathbf{\Phi}^{-1} \mathbf{v}_1 - \mathbf{v}_1^H \mathbf{\Phi}^{-1} \mathbf{T} (\mathbf{T}^H \mathbf{\Phi}^{-1} \mathbf{T})^{-1} \mathbf{T}^H \mathbf{\Phi}^{-1} \mathbf{v}_1 \,. \tag{18}$$

We note that when no null constraints are specified, the final term is absent and $\mathrm{DF}_{\mathrm{opt}} = \mathbf{v}_1^H \mathbf{\Phi}^{-1} \mathbf{v}_1$ [4]. A result somewhat similar to (18) was presented in [8] for the simpler case of a single null constraint. The expression for $\mathrm{DF}_{\mathrm{opt}}$ in (18) proves instrumental for the calculation of the average of $\mathrm{DF}_{\mathrm{opt}}$ over all look directions in Sec. IV, and leads to a geometric interpretation.

B. Interpretation

The terms in (18) lend themselves to an elegant geometric interpretation. Let $\tilde{\mathbf{v}}_1 = \mathbf{\Phi}^{-\frac{1}{2}} \mathbf{v}_1$ and $\tilde{\mathbf{T}} = \mathbf{\Phi}^{-\frac{1}{2}} \mathbf{T}$ (i.e., *whitened* versions of $\mathbf{v}$ and $\mathbf{T}$). The first term of (18) $\mathbf{v}_1^H \mathbf{\Phi}^{-1} \mathbf{v}_1$ is then the squared Euclidean norm: $\|\tilde{\mathbf{v}}_1\|^2$. This describes the maximum DF attainable *without* null constraints.

Now let us define $\tilde{\mathbf{v}}_{\parallel} = \tilde{\mathbf{T}}(\tilde{\mathbf{T}}^H \tilde{\mathbf{T}})^{-1} \tilde{\mathbf{T}}^H \tilde{\mathbf{v}}_1$; this is the projection of $\tilde{\mathbf{v}}_1$ onto the column space of $\tilde{\mathbf{T}}$. The second term of (18) corresponds to $-\|\tilde{\mathbf{v}}_{\parallel}\|^2$ and describes the reduction of the DF due to the null constraints. The net result is

$$\mathrm{DF}_{\mathrm{opt}} = \|\tilde{\mathbf{v}}_{\perp}\|^2 \,, \tag{19}$$

where $\tilde{\mathbf{v}}_{\perp} = \tilde{\mathbf{v}}_1 - \tilde{\mathbf{v}}_{\parallel}$, i.e the component of $\tilde{\mathbf{v}}_1$ which is orthogonal to $\tilde{\mathbf{T}}$.

In a more general sense, (19) applies to LCMV beamforming for suppressing arbitrary noise (with $\mathbf{R}$ replacing $\mathbf{\Phi}$). The output signal to noise ratio (SNR) attainable without null constraints is $\sigma_s^2 \|\tilde{\mathbf{v}}_1\|^2 = \sigma_s^2 \mathbf{v}_1^H \mathbf{R}^{-1} \mathbf{v}_1$, with $\sigma_s^2 = E\{|s|^2\}$ denoting the signal power. When null constraints are employed, the SNR is reduced to $\sigma_s^2 \|\tilde{\mathbf{v}}_{\perp}\|^2$.

IV. Average of Optimal DF over all Directions

The weights which optimize the DF depend on the look direction and on the null constraint directions. Therefore for a given array constellation, $\mathrm{DF}_{\mathrm{opt}}$ is subject to change for scenarios with different specifications. We wish to calculate the value of $\mathrm{DF}_{\mathrm{opt}}$ averaged over all directions.

For the moment, we assume that the null constraints ($\mathbf{T}$) are constant. For these constraints, we select the weights which optimize the DF for each look direction $\mathbf{v}_1(\mathbf{q})$. This yields a look direction dependent $\mathrm{DF}_{\mathrm{opt}}(\mathbf{q})$. The average of $\mathrm{DF}_{\mathrm{opt}}(\mathbf{q})$ over all look directions is[2]

$$\begin{aligned} \overline{\mathrm{DF}}_{\mathrm{opt}} &= \kappa \int_{\mathbf{q} \in \mathcal{Q}} \mathrm{DF}_{\mathrm{opt}}(\mathbf{q})\, A(\mathbf{q}) d\mathbf{q} \\ &= \left\langle \mathrm{DF}_{\mathrm{opt}}(\mathbf{q}) \right\rangle , \end{aligned} \tag{20}$$

[2] *All directions* does include the null-constraint directions, leading to an internal contradiction in the constraints (unity vs. null) for which the beamformer is not defined. Formally, (18) yields $\mathrm{DF}_{\mathrm{opt}} = 0$ for these cases. At any rate, the constraints are presumably consistent (i.e., do not lead to contradiction) almost everywhere (a.e.).

where we have denoted the averaging integral as $\langle \cdot \rangle \triangleq \kappa \int_{\mathbf{q} \in \mathcal{Q}} (\cdot)\, A(\mathbf{q}) d\mathbf{q}$ in order to streamline the notation.

Since the integrand of (20) is a scalar, it is equal to its own trace. Hence,

$$\begin{aligned} \overline{\mathrm{DF}}_{\mathrm{opt}} &= \Big\langle \mathrm{tr}\{\mathbf{v}_1^H \mathbf{\Phi}^{-1} \mathbf{v}_1 \\ &\qquad - \mathbf{v}_1^H \mathbf{\Phi}^{-1} \mathbf{T} (\mathbf{T}^H \mathbf{\Phi}^{-1} \mathbf{T})^{-1} \mathbf{T}^H \mathbf{\Phi}^{-1} \mathbf{v}_1\} \Big\rangle \\ &= \Big\langle \mathrm{tr}\{\mathbf{v}_1^H \mathbf{\Phi}^{-1} \mathbf{v}_1\} \\ &\qquad - \mathrm{tr}\{\mathbf{v}_1^H \mathbf{\Phi}^{-1} \mathbf{T} (\mathbf{T}^H \mathbf{\Phi}^{-1} \mathbf{T})^{-1} \mathbf{T}^H \mathbf{\Phi}^{-1} \mathbf{v}_1\} \Big\rangle \\ &= \Big\langle \mathrm{tr}\{\mathbf{\Phi}^{-1} \mathbf{v}_1 \mathbf{v}_1^H\} \\ &\qquad - \mathrm{tr}\{\mathbf{\Phi}^{-1} \mathbf{T} (\mathbf{T}^H \mathbf{\Phi}^{-1} \mathbf{T})^{-1} \mathbf{T}^H \mathbf{\Phi}^{-1} \mathbf{v}_1 \mathbf{v}_1^H\} \Big\rangle \\ &= \mathrm{tr}\left\{ \mathbf{\Phi}^{-1} \left\langle \mathbf{v}_1 \mathbf{v}_1^H \right\rangle \right\} \\ &\qquad - \mathrm{tr}\left\{ \mathbf{\Phi}^{-1} \mathbf{T} (\mathbf{T}^H \mathbf{\Phi}^{-1} \mathbf{T})^{-1} \mathbf{T}^H \mathbf{\Phi}^{-1} \left\langle \mathbf{v}_1 \mathbf{v}_1^H \right\rangle \right\} \end{aligned} \tag{21}$$

in which the following stages are applied. First, $\overline{\mathrm{DF}}_{\mathrm{opt}}$ is expressed as averaging over the trace of (18). In the second stage, the linearity of the trace operator is used to reformulate the integrand as the sum of two traces. In the third stage, we reorder terms using the property that the trace of a product is invariant to a cyclic permutation of the order of multiplicands. Finally, the order of the trace and the integration is reversed and constant terms are moved outside the integration.

The term $\left\langle \mathbf{v}_1 \mathbf{v}_1^H \right\rangle$ which appears twice in the final stage is $\mathbf{\Phi}$ by definition [see (13)]. Therefore, $\mathbf{\Phi}^{-1} \left\langle \mathbf{v}_1 \mathbf{v}_1^H \right\rangle = \mathbf{\Phi}^{-1} \mathbf{\Phi} = \mathbf{I}_N$, where $\mathbf{I}_N$ is an $N \times N$ identity matrix. Substituting back into (21) and again applying cyclic reordering yields

$$\begin{aligned} \overline{\mathrm{DF}}_{\mathrm{opt}} &= \mathrm{tr}\left\{\mathbf{I}_N\right\} - \mathrm{tr}\left\{ \mathbf{\Phi}^{-1} \mathbf{T} (\mathbf{T}^H \mathbf{\Phi}^{-1} \mathbf{T})^{-1} \mathbf{T}^H \right\} \\ &= \mathrm{tr}\left\{\mathbf{I}_N\right\} - \mathrm{tr}\left\{ \mathbf{T}^H \mathbf{\Phi}^{-1} \mathbf{T} (\mathbf{T}^H \mathbf{\Phi}^{-1} \mathbf{T})^{-1} \right\} \\ &= \mathrm{tr}\left\{\mathbf{I}_N\right\} - \mathrm{tr}\left\{\mathbf{I}_M\right\} \,. \end{aligned} \tag{22}$$

Since the trace of an identity matrix equals its rank, it follows that

$$\overline{\mathrm{DF}}_{\mathrm{opt}} = N - M \,. \tag{23}$$

We have thus proved that for an arbitrary choice of null directions, $\overline{\mathrm{DF}}_{\mathrm{opt}}$ equals the number of sensors minus the number of linearly independent null constraints. Since this is true for any specific choice, it is certainly true when averaging over any possible set of null directions (i.e., a set of feasible $\mathbf{T}$ matrices).

V. Discussion

A. Generalized propagation

Although the classical definition of DF corresponds to plane waves [as in (9) and (10)], the formulation presented in (11) does not specify the form of propagation or the parameter space. Instead of the plane wave propagation (1), more complex models may be employed leading to a broader perspective of DF. The validity of our main result (23) is not limited to a particular propagation model and can hence be applied to generalized versions of DF.

For example, we may define the DF for a microphone array operating in an echoic room. In this case, $\mathbf{q}$ would be the

Cartesian coordinates $[x\ y\ z]^T$, $\mathcal{Q}$ of (11) and (20) would be the region of interest, and $\mathbf{v}(\mathbf{q})$ would be an acoustic transfer function incorporating the direct path and reverberation. When averaging $\mathrm{DF}_{\mathrm{opt}}$ over the region of interest, (23) dictates that the result is $N - M$.

B. Regions of integration

The derivation in Sec. IV relies on the fact that $\langle \mathbf{v}\mathbf{v}^H \rangle = \mathbf{\Phi}$. This results from the integrations in (13) and in (20) both being over the same region $\mathcal{Q}$. However, one may be interested in averaging $\mathrm{DF}_{\mathrm{opt}}$ over a different region. For instance, the designer may have prior knowledge that the desired source is present only in a certain region $\mathcal{Q}'$. Replacing $\mathcal{Q}$ of (20) with $\mathcal{Q}'$ will result in[3] $\langle \mathbf{v}\mathbf{v}^H \rangle = \mathbf{\Phi}'$ which is generally different than $\mathbf{\Phi}$. The resulting average,

$$\overline{\mathrm{DF}}'_{\mathrm{opt}} = \mathrm{tr}\left\{\mathbf{\Phi}^{-1}\mathbf{\Phi}'\right\} - \mathrm{tr}\left\{\mathbf{\Phi}^{-1}\mathbf{T}(\mathbf{T}^H\mathbf{\Phi}^{-1}\mathbf{T})^{-1}\mathbf{T}^H\mathbf{\Phi}^{-1}\mathbf{\Phi}'\right\}, \tag{24}$$

does not necessarily equal $N - M$. In some cases, judicious design of the array constellation may provide higher DF values.

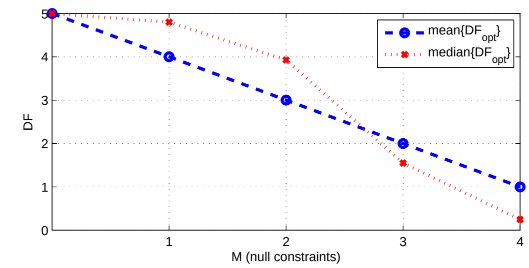
C. Average DF and distribution of DFs

Knowledge of the *average* DF does not guarantee that this value will be attained for a particular scenario. In fact, when the look direction is close to a null direction, the ensuing DF will be close to zero. A more complete portrayal is provided by the distribution of DFs.

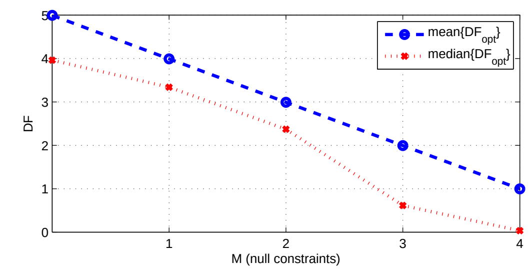
To illustrate this point, we now examine the well-known case of a uniform linear array (ULA). We conducted Monte Carlo simulations for a ULA of $N = 5$ omnidirectional elements. For each trial, 5 directions were selected at random (from a uniform distribution on the unit sphere); the fist direction was designated as the look direction and the other four reserved for potential null constraints. The optimal DF was calculated using (18) for a scenario with no null constraints and for scenarios with 1 to 4 null constraints. This was done for one million sets of randomly selected directions. The Monte Carlo procedure was repeated for two different interelement spacings: (i) 0.5 wavelengths, and (ii) 0.1 wavelengths.

Fig. 1 shows the sample means and medians for these two array configurations. The sample means are in agreement with (23); the sample medians are not given by (23) and depend on the array configuration. Fig. 2 shows the sample cumulative distribution functions (cdfs) for the two configurations. Here too, the results depend on the configuration. For example, with half-wavelength spacing the maximal DF a ULA can attain is N. With no null constraints, this is attained in all look directions [1] (leading to the "brick wall" form of the cdf for $M = 0$ in Fig. 2(a)). As null constraints are added, the probability of lower DFs increases. For a 0.1 wavelength spacing, higher DFs are attainable. Specifically, in the endfire direction the DF of an array with vanishingly small spacing approaches N^2 [6], [9] (e.g., in Fig. 2(b), the DF approaches 25). For other directions, the DF can be significantly smaller.

[3]Similarly, replacing $A(\mathbf{q})$ with $A'(\mathbf{q})$ would have the same effect.

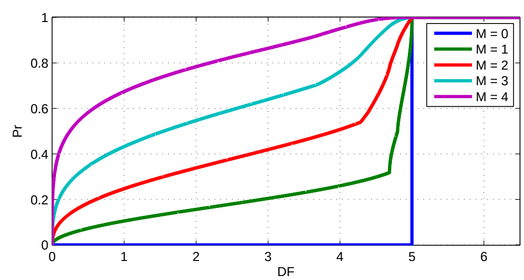


(a) interelement spacing: 0.5 wavelengths.

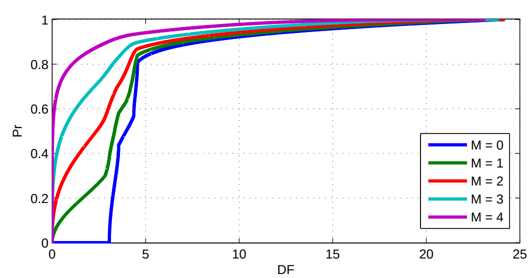


(b) interelement spacing: 0.1 wavelengths.

Fig. 1. Sample mean and median of DF for uniform linear array (ULA) with $N = 5$ elements and M null constraints.



(a) interelement spacing: 0.5 wavelengths.



(b) interelement spacing: 0.1 wavelengths.

Fig. 2. Sample cumulative distribution function (cdf) DF for uniform linear array (ULA) with $N = 5$ elements and M null constraints.
Note that the range of the x-axis differs in the two subplots.

VI. Conclusion

In this letter, we proved that when using optimal weights the average DF of a beamformer (over *all* directions) equals the number of array elements minus the number of null constraints. This result does not depend on the array constellation, and holds for generalized propagation regimes. When the DF is averaged over a particular region of interest, then the resulting average DF is given by (24), which may depend on the array constellation.

REFERENCES

[1] H. Van Trees, *Detection, Estimation, and Modulation Theory*. New York, USA: Wiley, Apr. 2002, vol. IV, Optimum Array Processing.

[2] E. Gilbert and S. Morgan, "Optimum design of directive antenna arrays subject to random variations," *Bell Syst. Tech. J*, vol. 34, pp. 637–663, 1955.

[3] H. Cox, "Sensitivity considerations in adaptive beamforming," NATO advanced study institute on signal processing with particular reference to underwater acoustics, Tech. Rep., 1972.

[4] D. Y. Levin, E. A. P. Habets, and S. Gannot, "A generalized theorem on the average array directivity factor," *IEEE Signal Process. Lett.*, vol. 20, no. 9, pp. 877–880, Sept 2013.

[5] H. Cox, R. Zeskind, and M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 10, pp. 1365–1376, Oct 1987.

[6] G. W. Elko, "Superdirectional microphone arrays," in *Acoustic signal processing for telecommunication*, S. L. Gay and J. Benesty, Eds. Springer, 2000, pp. 181–237.

[7] R. A. Horn and C. R. Johnson, *Matrix analysis*, 2nd ed. Cambridge University Press, 2013.

[8] A. Fenwick, "An array processor optimum subject to a constraint," *J. Acoust. Soc. Am.*, vol. 49, no. 4A, pp. 1108–1110, 1971.

[9] A. Uzkov, "An approach to the problem of optimum directive antenna design," *Comptes Rendus (Doklady) de l'Academie des Sciences de l'URSS*, vol. 53, no. 1, pp. 35–38, 1946.