# THE BIAS OF THE MVDR BEAMFORMER OUTPUTS UNDER DIAGONAL LOADING

*Raj Rao Nadakuditi* *

Massachusetts Institute of Technology
Department of Electrical Engineering
and Computer Science
MIT/WHOI Joint Program
Cambridge, MA 02139

*Alan Edelman*

Massachusetts Institute of Technology
Department of Mathematics
Cambridge, MA 02139.

## ABSTRACT

The MVDR beamformer is the most extensively used array processing algorithm and involves inverting the sample covariance matrix. In the snapshot deficient scenario, when the number of sensors is greater than or approximately equal to the number of snapshots, the eigenvalues of the resulting sample covariance matrix are poorly conditioned. Diagonal loading is then applied to the sample covariance matrix. Expressions for the bias of the resulting MVDR beamformer outputs in the sidelobe region are presented that are exact for asymptotically large arrays. Numerical simulations confirm the accuracy of these asymptotic expressions when predicting the bias of the outputs of moderately large arrays.

## 1. INTRODUCTION

Analytical expressions for the bias and variance of the minimum variance distortionless response (MVDR) beamformer outputs under diagonal loading have been considered as open problems by researchers in array processing [1]. A closed-form expression for the bias of the diagonally loaded MVDR beamformer in the sidelobe region is presented that has been derived using infinite random matrix techniques [2–5].

A part of our work is similar to that of Mestre and Lagunas [6] in the sense that we too use infinite random matrix techniques to characterize the bias of the MVDR beamformer under diagonal loading . Our work is different from theirs because we explicitly derive closed-form analytical expressions for the bias even in the regime when the number of snapshots is less than the number of sensors for *both* forms of the MVDR beamformer. These expressions are useful because they help analytically predict the extent to which the sidelobe levels in MVDR beamformers are likely to rise when diagonal loading is applied to compensate for finite sample size constraints in large arrays.

Sections 2 and 3, respectively, introduce the classical and the diagonally loaded MVDR beamformer. Section 4 presents some results using infinite random matrix theory that will be used in Section 5 to derive the analytical expressions for the bias in asymptotically large arrays. Simulations in Section 6 demonstrate how these asymptotic expressions are accurate for predicting the bias even when the arrays are finite dimensional. The conclusions are summarized in Section 7.

*This work was supported by ONR Grant N00014-00-1-0048

## 2. THE MVDR BEAMFORMER

The MVDR algorithm, like many other adaptive algorithms, forms the sample covariance matrix $\hat{K}$ from the $M$ received snapshots of time samples and frequency bins near the center frequency of interest. This sample covariance matrix is given by

$$\hat{K} = \frac{1}{M}\sum_{i=1}^{M} \underline{x}_i \underline{x}_i^H \tag{1}$$

where $\underline{x}_i$ are the $N \times 1$ snapshot vectors. The MVDR beamformer is then given by

$$\hat{P}(\underline{u}) = \frac{1}{\underline{V}^H(\underline{u})\hat{K}^{-1}\underline{V}(\underline{u})} \tag{2}$$

where $\underline{V}(\underline{u})$ is the manifold vector. If the elements of $\underline{x}_i$ are assumed to be independent identically distributed (i.i.d.) complex Gaussian random variables with mean zero and variance one, then $\hat{K}$ is simply the extensively studied [7,8] ("pure") Wishart matrix . Using this knowledge, Capon and Goodman demonstrated [9] that $\hat{P}(\underline{u})$ has a complex, chi-squared distribution with $M - N + 1$ degrees of freedom. This leads to their well-known and extensively used expressions for the bias and variance of the MVDR beamformer outputs. These are, respectively,

$$\text{Bias} = N\, E[\hat{P}(\underline{u})] = \frac{M-N+1}{M} \tag{3}$$

$$\text{Variance} = N\, \sigma^2_{\hat{P}(\underline{u})} = \frac{M-N+1}{M^2}. \tag{4}$$

Note the bias and variance expressions in (3) and (4) assume that the manifold vector is normalized such that $\underline{V}^H(\underline{u})\underline{V}(\underline{u}) = N$.

Equations (3) and (4) are important for array processing because they help analytically characterize how a limited number of snapshots $M$ affect the performance of the MVDR beamformer. In particular, since $\hat{K}$ was modelled as a Wishart matrix, the expressions allow us to characterize the performance of the MVDR beamformer, in an admittedly ad-hoc manner, under broad sector nulling. In a practical setting these expressions are important and widely used, because they predict the sidelobe levels due to snapshot constraints when forming the sample covariance matrix.

## 3. DIAGONAL LOADING

The MVDR beamformer involves explicitly computing the sample covariance matrix in (1). In the snapshot deficient case, $\hat{K}$ is rank deficient when $M < N$ or poorly conditioned with low eigenvalues for $M \approx N$. In many realistic operating environments, this snapshot deficiency cannot be overcome because of physical or stationarity constraints. Hence, diagonal loading is applied to the sample covariance matrix which is replaced by

$$\hat{K}(\delta) = \hat{K} + \delta I. \tag{5}$$

With this, the weight vector of the MVDR processor is given by

$$\underline{w}(\underline{u} \mid \delta) = \frac{\hat{K}^{-1}(\delta)\underline{V}(\underline{u})}{\underline{V}^H(\underline{u})\hat{K}^{-1}(\delta)\underline{V}(\underline{u})}. \tag{6}$$

The MVDR beam output can be calculated in two ways. These are given by

$$\hat{P}_1(\underline{u}) = \underline{w}^H(\underline{u} \mid \delta)\hat{K}\,\underline{w}(\underline{u} \mid \delta) \tag{7}$$

$$\hat{P}_2(\underline{u}) = \frac{1}{\underline{V}^H(\underline{u})\hat{K}^{-1}(\delta)\underline{V}(\underline{u})}. \tag{8}$$

Of the two ways above, (7) is the most extensively used method since (8) has been observed to have a detrimental impact on the detectability of low level signals as a consequence of the increased sidelobe levels. Except for the very special, and practically unrealistic situation, of a single snapshot, there are no published analytic results [1] for the bias of the MVDR beam outputs in (7) and (8) for the snapshot deficient case.

Though the MVDR beamformer is amongst the oldest and most extensively used adaptive array processing algorithm, analytically characterizing its performance due to finite snapshots constraints has remained an open problem, as Baggeroer and Cox state in [1]. Before we analyze its performance in a specific, though useful, scenario we will derive some expressions using infinite random matrix theory that will turn out to be useful later on.

## 4. INFINITE RANDOM MATRIX THEORY

For an $N \times N$ matrix $A_N$, the eigenvalue distribution function (e.d.f.) $F^{A_N}(x)$ is defined as

$$F^{A_N}(x) = \frac{\text{Number of eigenvalues of } A_N \leq x}{N}. \tag{9}$$

As defined, the e.d.f. is right continuous and possibly atomic i.e. with step discontinuities at discrete points. In practical terms, the derivative of (9), referred to as the (eigenvalue) probability distribution function (p.d.f.), is simply the appropriately normalized histogram of the eigenvalues of $A_N$.

The moments of the matrix $A_N$ are defined as

$$M_k^{A_N} \equiv \varphi(A_N^k) = \frac{1}{N}\operatorname{tr}(A_N^k) = \int x^k dF^{A_N}(x). \tag{10}$$

**Theorem 1 (The Marčenko-Pastur distribution [3–5]).** *Let $G_N$ be an $N \times M$ matrix with i.i.d. zero mean unit variance elements whose higher order moments are bounded as well. Let $N/M \to c > 0$ as $M, N \to \infty$. Let $W_N(c) = 1/M\, GG^*$ denote an $M \times N$ "generalized" [1] Wishart matrix. Then, the e.d.f. $F^{W_N}$ strongly converges as $M, N \to \infty$ to the (non-random) distribution function $F^W(x)$ given by*

$$\frac{dF^W(x)}{dx} = \max\left(0, 1 - \frac{1}{c}\right)\delta(x) + \frac{\sqrt{(x - b_-)(b_+ - x)}}{2\pi x c} I_{[b_-, b_+]} \tag{11}$$

*where, $b_{\pm} = (1 \pm \sqrt{c})^2$ and $I_{[b_-, b_+]}$ is the indicator function that is equal to $1$ for $b_- < x < b_+$ and $0$ elsewhere.*

Furthermore, when $W_N$ is a "pure" Wishart matrix, its eigenvectors will be Haar distributed for *all* $N$ [8]. When $W_N$ is a "generalized" Wishart matrix, its eigenvectors will be Haar distributed [10] when $N \to \infty$.

Consider, now, the $N \times N$ matrix $D_N = (W_N(c) + \delta\, I)^{-1}$. The eigenvalues of the matrix $D_N$ can be deterministically inferred from the eigenvalues of the matrix $W_N$. Hence, the convergence of the e.d.f. $F^{W_N}(x)$ implies convergence of the e.d.f $F^{D_N}(x)$ as $M, N \to \infty$. Additionally the eigenvectors of $D_N$ will also be Haar distributed. We are, however, interested in the limiting moments of the matrix $D_N$. We can choose to derive the limiting e.d.f $F^D(x)$ from $F^W(x)$ in (11) and then compute its moments using (10). Alternately, the limiting moments of $D_N$ can be directly expressed in terms of the limiting e.d.f. $dF^W(x)$ as simply

$$\varphi(D^k) = \lim_{N\to\infty} \varphi(D_N^k) = \int \frac{1}{(x+\delta)^k} dF^W(x). \tag{12}$$

The second term on the right hand side of (11) can, with an appropriate change of variables, be rewritten as a Beta distribution whose properties, including the moments, are well known [11,12]. The first and second limiting moments of $D_N$ can hence, with some tedious but relatively straightforward algebra, be expressed as

$$\varphi(D) = \frac{-1 + c - \delta + \sqrt{1 - 2\,c + 2\,\delta + c^2 + 2\,c\,\delta + \delta^2}}{2c\,\delta} \tag{13}$$

$$\varphi(D^2) = \frac{1 + c\,\delta + c\sqrt{c^2 - 2\,c + 2\,c\,\delta + 1 + 2\,\delta + \delta^2}}{2\sqrt{c^2 - 2\,c + 2\,c\,\delta + 1 + 2\,\delta + \delta^2}c\delta^2} + \frac{c^2 - \sqrt{c^2 - 2\,c + 2\,c\,\delta + 1 + 2\,\delta + \delta^2} + \delta - 2\,c}{2\sqrt{c^2 - 2\,c + 2\,c\,\delta + 1 + 2\,\delta + \delta^2}c\delta^2}. \tag{14}$$

We note that a more elegant way of obtaining moments of random matrices, such as $D_N$ in this case, without having to resort to cumbersome integration is to use the techniques developed in [2, 13] instead. As we shall shortly see, these moments and a random matrix lemma that we will derive next, will prove to be important in our analysis of the bias of the MVDR beamformer.

**Lemma 1.** *Let $\underline{V}(\underline{u})$ be an $N \times 1$ complex manifold vector with $\underline{V}^H(\underline{u})\underline{V}(\underline{u}) = N$. Let $D_N$ be an random matrix independent of $\underline{V}(\underline{u})$ whose e.d.f. converges to $F^D(x)$ as $N \to \infty$ and whose eigenvectors are Haar distributed. Then,*

$$\lim_{N\to\infty} E[\underline{V}^H(\underline{u})D_N^k\underline{V}(\underline{u})] = \operatorname{tr}(D^k) = N\,\varphi(D^k). \tag{15}$$

*where $\varphi(D^k) = \int x^k\, dF^D(x)$ is the limiting moment of $D_N$.*

[1] It is a "pure" Wishart matrix when the i.i.d. elements are normally distributed.

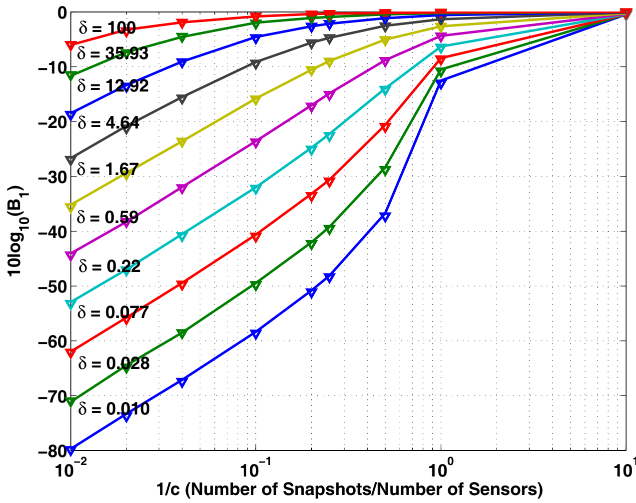


**Fig. 1**: Bias, $B_1$, of the diagonally loaded MVDR beamformer output obtained using (7) for $N = 100$ and $c = N/M$ compared with experimental values averaged over 1000 trials.

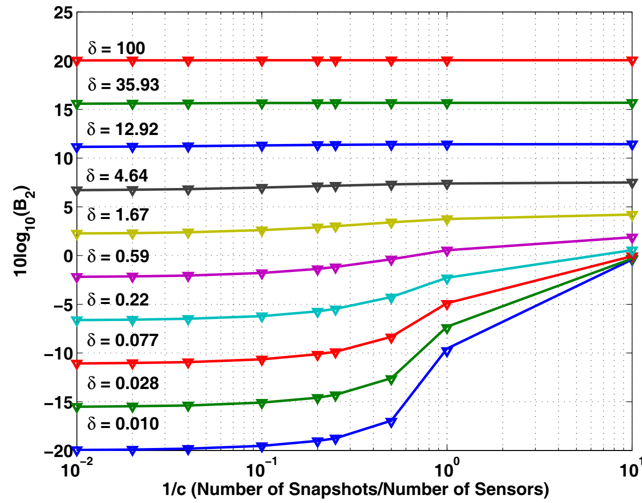


**Fig. 2**: Bias $B_2$ of the diagonally loaded MVDR beamformer outputs obtained using (8) for $N = 100$ and $c = N/M$ compared with experimental values averaged over 1000 trials.

*Proof.* Let $D_N$ be written as $D_N = Q_N\, L_N\, Q_N^*$ where $Q_N$ is a Haar unitary matrix [8] and $L_N$ is a diagonal matrix. Since the e.d.f. of $D_N$ converges as $N \to \infty$ to the non-random distribution function $F^D(x)$, the e.d.f. of the diagonal matrix $L_N$ will also converge to $F^L(x) = F^D(x)$ as $N \to \infty$. Hence, $\lim_{N\to\infty} \varphi(L_N^k) = \varphi(L^k) = \varphi(D^k)$. Furthermore,

$$
\begin{aligned}
E[\underline{V}^H(\underline{u})D_N^k\underline{V}(\underline{u})] &= E[\underline{V}^H(\underline{u})Q_N\, L_N^k\, Q_N^H\underline{V}(\underline{u}) && (16)\\
&= E[\mathrm{tr}(L_N^k Q_N^H \underline{V}(\underline{u})\underline{V}^H(\underline{u})Q_N))] && (17)\\
&= E[\mathrm{tr}(L_N^k\, \underline{q}\,\underline{q}^H)] && (18)
\end{aligned}
$$

where $\underline{q} = Q_N^H\underline{V}(\underline{u})$ is a $N \times 1$ vector with $E[\underline{q}\underline{q}^H] = I$. Since the trace and expectation commute, and $L_N$ and $\underline{q}\underline{q}^H$ are independent of each other, (18) can be rewritten as

$$
\begin{aligned}
\lim_{N\to\infty} E[\underline{V}^H(\underline{u})D_N^k\underline{V}(\underline{u})] &= \lim_{N\to\infty} \mathrm{tr}(E[L_N^k]\, E[\underline{q}\underline{q}^H]) && (19)\\
&= \lim_{N\to\infty} \mathrm{tr}(E[L_N^k]\, I) && (20)\\
&= \lim_{N\to\infty} E[\mathrm{tr}(L_N^k)] && (21)\\
&= N\,\varphi(L^k) && (22)
\end{aligned}
$$

□

## 5. BIAS FOR ASYMPTOTICALLY LARGE ARRAYS

Let us assume,, for the rest of this analysis that $M, N \to \infty$ so that the e.d.f. and the moments of $\hat{K}$, and $\hat{K}(\delta)$ are non-random. When $M < N$ or $M \approx N$, the matrix $\hat{K}$ given by (1) is replaced by $\hat{K}(\delta) = \hat{K} + \delta\, I$. This is simply the matrix $D_N$ we discussed earlier. The resulting MVDR beamformer may then be computed using either (7) or (8).

### 5.1. Case 1

The MVDR output in the sidelobe region given by (7) can be written after substituting the expression for $\underline{w}(\underline{u} \mid \delta)$ as

$$
\hat{P}_1(\underline{u}) = \frac{\underline{V}^H(\underline{u})\hat{K}^{-1}(\delta)\hat{K}\hat{K}^{-1}(\delta)\underline{V}(\underline{u})}{\left(\underline{V}^H(\underline{u})\hat{K}^{-1}(\delta)\underline{V}(\underline{u})\right)^2}. \qquad (23)
$$

As $M, N \to \infty$, both the numerator and the denominator of (23) are non-random . This means that the numerator and the denominator can be evaluated separately and then substituted into (23) to obtain the required expression for $\hat{P}_1(\underline{u})$. By Lemma 1, the denominator of (23) can be expressed as

$$
\left(\underline{V}^H(\underline{u})\hat{K}^{-1}(\delta)\underline{V}(\underline{u})\right)^2 = \left(N\,\varphi(\hat{K}^{-1}(\delta))\right)^2. \qquad (24)
$$

Since $\hat{K}^{-1}(\delta) = D_N$, $\varphi(\hat{K}^{-1}(\delta))$ is given by (13). Hence to obtain the expression for $\hat{P}_1(\underline{u})$ it is necessary to evaluate the numerator of (23). Since $\hat{K}(\delta) = \hat{K} + \delta\, I$, so that $\hat{K} = \hat{K}(\delta) - \delta\, I$, the numerator of (23) can be written as

$$
\begin{aligned}
&\underline{V}^H(\underline{u})\hat{K}^{-1}(\delta)\hat{K}\hat{K}^{-1}(\delta)\underline{V}(\underline{u}) = \\
&\qquad \underline{V}^H(\underline{u})\left(\hat{K}^{-1}(\delta) - \delta\,\hat{K}^{-2}(\delta)\right)\underline{V}(\underline{u}) \qquad (25)
\end{aligned}
$$

which can be rewritten using Lemma 1 as simply.

$$
\begin{aligned}
&\underline{V}^H(\underline{u})\hat{K}^{-1}(\delta)\hat{K}\hat{K}^{-1}(\delta)\underline{V}(\underline{u}) = \\
&\qquad N\left(\varphi(\hat{K}^{-1}(\delta)) - \delta\,\varphi(\hat{K}^{-2}(\delta))\right). \qquad (26)
\end{aligned}
$$

There are two terms on the right hand side of (26). The expression for $\varphi(\hat{K}^{-1}(\delta))$ is given by (13) while the expression for $\varphi(\hat{K}^{-2}(\delta))$ is given by (14). Hence, the relationships in (26),(14),(24) and (13) can be used to obtain the denominator and numerator respectively of the expression in (23). These values can then be substituted in (23) to obtain an expression for $\hat{P}_1(\underline{u})$ using some tedious yet straightforward algebra. If the bias of (7) were defined as $B_1 = N\,\hat{P}_1(\underline{u})$ then it can easily shown that the required expression is given by

$$B_1 = \frac{\left(-\sqrt{c^2-2\,c+2\,c\delta+1+2\,\delta+\delta^2}+c+1+\delta\right)}{\left(-c+1+\delta-\sqrt{c^2-2\,c+2\,c\delta+1+2\,\delta+\delta^2}\right)^2} \times \frac{2\,c\,\delta^2}{\sqrt{c^2-2\,c+2\,c\delta+1+2\,\delta+\delta^2}}. \quad (27)$$

### 5.2. Case 2

Here the MVDR output under diagonal loading in the sidelobe region given by (8) can be written using Lemma 1 as

$$\begin{aligned}\hat{P}_2(\underline{u}) &= \frac{1}{\underline{V}^H(\underline{u})\hat{K}^{-1}(\delta)\underline{V}(\underline{u})} \\ &= \frac{1}{N\,\varphi(\hat{K}^{-1}(\delta))}.\end{aligned} \quad (28)$$

Substituting the expression for $\varphi(\hat{K}^{-1}(\delta)) = \varphi(D_N)$ given by (13) in (28) will give us an expression for $\hat{P}_2(\underline{u})$. If the bias were to be defined as $B_2 = N\hat{P}_2(\underline{u})$ then the bias of the MVDR output given by (8) is simply

$$B_2 = \frac{2c\,\delta}{(-1+c-\delta+\sqrt{1-2\,c+2\,\delta+c^2+2\,c\,\delta+\delta^2})}. \quad (29)$$

When $\delta = 0$, it can be shown, upon using L'Hospital's Rule, that (29) reduces to $B_2 = 1 - c = 1 - N/M$ which is simply the Capon-Goodman result in (3) for large $N$.

Note that as $M, N \to \infty$, $\hat{P}_1(\underline{u})$, and $\hat{P}_2(\underline{u})$ become deterministic with variance $zero$. It is possible to analyze the variance of the MVDR outputs in the sidelobe region formed using (7) and (8) for finite $M, N$. The techniques involved are described in [2].

## 6. SIMULATION RESULTS

Equations (27) and (29) are expressions for the bias of the MVDR beamformer outputs in the sidelobe region as $M, N \to \infty$. Figures 1 and 2 compare the analytical expressions in (27) and (29) (solid line) with 1000 realizations (triangles) of the diagonally loaded beamformer with $N = 100$ for a different range of values for $\delta$ and $c = N/M$. Given the excellent agreement exhibited in these figures, it is clear that though these expressions were derived using infinite random matrix techniques, they accurately predict the bias of the MVDR beamformer outputs under diagonal loading even when $M$ and $N$ are finite.

## 7. CONCLUSIONS

Closed form analytical expressions for the bias of the diagonally loaded MVDR beamformer outputs in the sidelobe region have been derived that are exact for asymptotically large arrays. Numerical simulations have been used to justify the validity of these asymptotic expressions for finite dimensional arrays as well. While the problem of determining the bias under the presence of an arbitrary number of interferers at arbitrary positions (relative to the steering direction) remains open, our expressions help provide an answer, unknown previously, in a scenario that is of relevance.

# ACKNOWLEDGEMENTS

We would like to thank Professor Art Baggeroer and Dr. Jim Preisig for their feedback and encouragement.

---

## 8. REFERENCES

[1] A. B. Baggeroer and H. Cox, "Passive sonar limits upon nulling multiple moving ships with large aperture arrays," in *Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, 1999, vol. 1, pp. 103–108.

[2] R. Rao Nadakuditi and Alan Edelman, "The polynomial method for random matrices.," Preprint.

[3] V.A. Marčenko and L.A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Math USSR Sbornik*, vol. 1, pp. 457–483, 1967.

[4] Sang Il Choi and Jack W. Silverstein, "Analysis of the limiting spectral distribution of large dimensional random matrices," vol. 54(2), pp. 295–309. 1995.

[5] Jack W. Silverstein, "Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 55(2), pp. 331–339, 1995.

[6] X. Mestre and M. A. Lagunas, "On diagonal loading for minimum variance beamformers," in *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT 2003)*, Darmstadt (Germany), December 14–17.

[7] J. Wishart, "The generalized product moment distribution in samples from a normal multivariate population," *Biometrika*, vol. 20 A, pp. 32–52, 1928.

[8] Alan Edelman, *Eigenvalues and Condition Numbers of Random Matrices*, Ph.D. thesis, Massachusetts Institute of Technology, 1989.

[9] J. Capon and N. R. Goodman, "Probability distributions for estimator of the frequency wavenumber spectrum," *Proceedings of the IEEE*, vol. 58, pp. 1785–1786, October 1970.

[10] Jack W. Silverstein, "On the eigenvectors of large-dimensional sample covariance matrices," *J. Multivariate Anal.*, vol. 30, no. 1, pp. 1–16, 1989.

[11] M. V. Jambunathan, "Some properties of Beta and Gamma distributions," *Ann. Math. Stat.*, vol. 25, pp. 401–405, 1954.

[12] M. Abramowitz and I.A. Stegun, Eds., *Handbook of Mathematical Functions*, Dover Publications, New York, 1970.

[13] Alan Edelman, "Course notes for 18.338, M.I.T.," Fall 2004.