COLOCATED ANTENNA ARRAYS: DESIGN DESIDERATA FOR WIRELESS COMMUNICATIONS

Samuli Visuri and Dirk T.M. Slock

Institut Eurécom, 2229 route des Crêtes, B.P. 193
06904 Sophia Antipolis Cedex, FRANCE
Tel.: +33 4 93002626, Fax: +33 4 93002627
email:{visuri,slock}@eurecom.fr

ABSTRACT

The research we report on in this paper is motivated by the drive towards antenna array miniaturization: going from an array of identical antennas that get spread out in space towards an array of colocalized antennas with differing responses. We address the problem of designing the colocated antenna array response for two applications: diversity reception using maximum ratio combining and MIMO communication capacity. The design criteria are derived under specific models regarding the distribution of the parameters of the frequency flat channel. We are particularly interested in evaluating the performance of two opposite extremes in terms of design: very selective beamspace type of operation versus spatially allpass antennas. We also show that the pathwise channel model converges for large number of paths to the stochastic channel model with separable correlation between TX and RX sides.

1. INTRODUCTION

In future wireless communications, antenna arrays will play a central role. In general, the main motivation to use multiple-element arrays at the receiver and/or at the transmitter is the added spatial dimension that can be used to enhance the overall network performance. Roughly speaking the benefit obtained from the multiple antennas grows as a function of the number of antenna elements.

The physical size of an array limits the number of the antenna elements that can be used, especially at the mobile terminal side. For example, it is a well known fact that to fully exploit the added spatial dimension with an array that consists of identical elements, the elements should have half-wavelength spacing.

Recently there have been studies showing that colocalized antenna arrays consisting of only one physical antenna element are useful in obtaining spatial diversity. An example of such a colocalized antenna is a biconical horn antenna where the different array "elements" are obtained by exciting different modes of the antenna [1, 2]. The benefit obtained by using colocated antenna arrays is the reduced size of the overall antenna. Spatial diversity means that a given antenna has a response that varies with the position of the antenna in space. Spatial diversity at a small spatial scale is in fact due to angular diversity, which means that a given point in space multiple wavefronts arrive from different angles. A given antenna response will combine these wavefronts into a superposition that varies with position. At a larger spatial scale, other sources of spatial diversity arise to variations in the wavefront scenario. Hence, small arrays of identical antennas pick up mostly angular diversity. This diversity can equally well be picked up by colocated antennas with differing angular responses, leading to different superpositions of the same set of wavefronts.

In this paper we study the design of the colocated antenna array response for two applications: diversity reception using maximum ratio combining and MIMO communication capacity. We take full freedom to design the arrays' steering vector as a function of the angle of an incoming/departing signal. This is the major difference to the previously published antenna array design results, that have mainly been focusing on designing placements of (identical) array elements in order to enhance DOA estimation performance. See, for example, [3, 4, 5] and references therein.

We obtain analytical results showing that in most of the cases the array elements should be uncorrelated and have equal variances (after averaging over angle distributions also). We show that this kind of behavior is possible to obtain by two extremes in design, spatially allpass elements where the phase response depends on the angle distribution, or very selective totally sectorized antennas.

2. DIVERSITY RECEPTION

2.1. Array design

We first consider antenna array design for diversity reception using MRC combining. Assume that a single narrowband signal undergoes multipath propagation and arrives to an N element antenna array from L directions $\theta_1, \theta_2, \dots, \theta_L$. The time delays associated to the multipaths are assumed to be much smaller than the inverse bandwidth of the signal, so there is no delay spread. We consider the baseband equivalent signal model

$$\boldsymbol{x}(t) = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{c}s(t) + \boldsymbol{v}(t),$$

where $\boldsymbol{x}(t)$ is an $N \times 1$ received signal vector, $s(t)$ is the transmitted signal, $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_L]^T$,

$$\boldsymbol{A}(\boldsymbol{\theta}) = [\boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_2), \dots, \boldsymbol{a}(\theta_L)]$$

is an $N \times L$ matrix of $N \times 1$ steering vectors $\boldsymbol{a}(\theta)$ and $\boldsymbol{v}(t)$ is an $N \times 1$ noise vector. For notational convenience we simply write $\boldsymbol{A}$ instead of $\boldsymbol{A}(\boldsymbol{\theta})$ when there is no possibility of confusion. The vector $\boldsymbol{c} = [c_1, \dots, c_L]^T$ contains the random multipath complex

Eurécom's research is partially supported by its industrial partners: Ascom, Swisscom, Thales Communications, ST Microelectronics, GEGETEL, Motorola, France Télécom, Bouygues Télécom, Hitachi Europe Ltd. and Texas Instruments. The work reported herein was also partially supported by the French RNRT project ANTIPODE.

amplitudes, which are assumed to be of zero mean, circular complex Gaussian with different variances, and mutually independent, i.e. $c_i \sim \mathcal{CN}(0, \sigma_i^2)$. We also assume that the DOAs θ_i are random, i.i.d., and independent of the fading coefficients. The overall channel from the transmitter to the array elements is given by a vector $\boldsymbol{h} = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{c}$.

After maximum ratio combining, the channel part becomes

$$\xi = ||\boldsymbol{h}||^2,$$

where $||\cdot||$ denotes the Euclidean vector norm.

From a diversity point of view we would like ξ to have a high average value and vary as little as possible around the average value. A good way to describe this behavior is via the normalized variance

$$\frac{\sigma_\xi^2}{(E\{\xi\})^2} = \frac{E\{|\xi|^2\}}{(E\{\xi\})^2} - 1, \tag{1}$$

where σ_ξ^2 denotes the variance of ξ and E denotes the expectation operator. The variance and the expectation are calculated with respect to all random variables, i.e. with respect to the complex amplitudes and the angles. Our goal is to design a steering vector $\boldsymbol{a}(\theta)$ that minimizes the above characteristic of variation.

The properties that an optimal steering vector should fulfill are given by the following theorem.

Theorem 1 *A steering vector $\boldsymbol{a}(\theta)$ minimizes the criterion (1) if and only if the following conditions are fulfilled.*

(i)

$$E_\theta\{a_i(\theta)a_k(\theta)^*\} = 0, \quad i \neq k \tag{2}$$

(ii)

$$||\boldsymbol{a}(\theta)||^2 = c > 0 \quad \textit{w.p.1} \tag{3}$$

(iii)

$$E_\theta|a_i(\theta)|^2 = \frac{c}{N} \quad \forall i \tag{4}$$

Proof: See the appendix.

To summarize, the elements of an optimal array should be uncorrelated, the array should have constant response to all the directions where the density function of the DOAs is positive, and the array elements should have equal variances. Furthermore, since $E(\xi) = \sum_{i=1}^{L} \sigma_i^2 E\{||\boldsymbol{a}(\theta)||^2\}$ (see the appendix), $E(\xi)$ gets maximized by maximizing c in (3).

2.2. Solution Examples

How can we then choose a steering vector that satisfies the constraints (2) – (4). In what follows we give two different type of solutions: array of allpass elements and array of completely sectorized elements. We assume that $||\boldsymbol{a}(\theta)|| = 1 \; \forall\theta$.

2.2.1. Arrays of spatial allpass elements

If the modulus of the elements of the steering vector does not depend on θ, we know that generally

$$a_i(\theta) = \frac{1}{\sqrt{N}} e^{jf_i(\theta)},$$

where $f_i(\theta)$ is some function of θ.

One solution satisfying (2) is, for example,

$$a_i(\theta) = \frac{1}{\sqrt{N}} e^{j2\pi(\alpha - i)F_\theta(\theta)},$$

where F_θ is the cumulative distribution function of the DOAs and $\alpha \in \mathbb{R}$. Note that in this case the random variable $F_\theta(\theta)$ is uniformly distributed on the interval $[0, 1)$ and hence for $i \neq k$

$$E_\theta\{a_i(\theta)a_k^*(\theta)\} = \frac{1}{N}\int_0^1 e^{j2\pi(k-i)u} du = 0.$$

A more general solution is given by

$$a_i(\theta) = \frac{1}{\sqrt{N}} e^{j(2\pi(\alpha \pm i)rF_\theta(\theta) + \phi_i)}, \tag{5}$$

where $\alpha \in \mathbb{R}$, r is a positive integer and ϕ_i is a constant phase for element i.

It is interesting to note that if the p.d.f. of the angles is $f_\theta(\theta) = \frac{1}{2}\cos(\theta)I_{[-\pi/2,\pi/2]}(\theta)$, the c.d.f. is $\frac{1}{2}(\sin(\theta) + 1)$ and by letting $\alpha = 1$, $r = 1$ $\phi_i = -i\pi$ and using the "plus sign" in (5) we get

$$a_i(\theta) = \frac{1}{\sqrt{N}} e^{-j2\pi(i-1)\frac{1}{2}\sin(\theta)}$$

In other words, if the distribution of the angles is the cosine distribution over the interval $[-\pi/2, \pi/2]$, the resulting array maximizing the diversity is a ULA with $\lambda/2$ spacing (λ denotes the wavelength of the signal). The parameter r in this case describes the interelement spacing of the ULA. In general, the array in case of cosine angle distribution could be a ULA with $(r\lambda)/2$ spacing.

2.2.2. Arrays of spatially sectorized elements

If we allow the modulus of the elements of the steering vector to depend on θ, a possible solution is to use completely spatially sectorized elements. Let $f(\theta)$ be the density function of the angles and let $\Theta = \{\theta : \; f(\theta) > 0\}$. Define $\Theta_1, \ldots, \Theta_N$ such that

$$\bigcup_{i=1}^{N} \Theta_i = \Theta$$

and

$$\Theta_i \cap \Theta_j = \emptyset, \quad i \neq j$$

Finally let

$$a_i(\theta) = I_{\Theta_i}(\theta), \quad i = 1, \ldots, N. \tag{6}$$

Here

$$I_{\Theta_i}(\theta) = \begin{cases} 1 & \theta \in \Theta_i \\ 0 & \theta \neq \Theta_i \end{cases}$$

If Θ_i:s are chosen such that $P[\theta \in \Theta_i] = \frac{1}{N}$, the conditions are fulfilled. An array corresponding to the steering vector defined in (6) is consist from completely sectorized elements. Note that the sectorization implies automatically the condition (2). The condition (3) implies that an particular element must have a constant modulus in the sector. Finally condition (4) implies that the size of the element's sector must be such that $P[\theta \in \Theta_i] = \frac{1}{N}$

3. MIMO COMMUNICATION CAPACITY

3.1. Pathwise channel model versus separable spatial correlation channel model

We now turn our attention to design steering vectors maximizing the ergodic capacity of MIMO (Multi Input Multi Output) channel with N receive and M transmit antennas.

The maximization is done by assuming a specific flat fading channel model, where the elements of the channel matrix are correlated and the correlation is separable between the transmit and receive array [6, 7, 8]. We call this channel model separable spatial correlation channel model. To make the discussion more general, we first show that this channel model is obtained asymptotically from a more general pathwise channel model [9].

Let $\boldsymbol{y}$ be the received $N \times 1$ signal vector. In this part we consider the following single user linear channel model

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{v}, \tag{7}$$

where $\boldsymbol{x}$ is the $M \times 1$ transmitted signal vector and $\boldsymbol{v}$ is the $N \times 1$ noise vector which is assumed to be complex circular Gaussian with covariance matrix $\sigma_v^2 I$.

In the pathwise channel model the $N \times M$ MIMO channel matrix is given by

$$\boldsymbol{H} = s_L^{-1} \sum_{l=1}^{L} c_l \boldsymbol{a}(\theta_l) \boldsymbol{b}^T(\phi_l), \tag{8}$$

where L is the number of multipaths and c_l, $i = 1, \ldots, L$ denote the complex multipath amplitudes. Vector $\boldsymbol{a}(\theta)$ is the steering vector of the receive array and $\theta_1, \ldots, \theta_L$ are the DOAs of the multipath components. Similarly $\boldsymbol{b}(\phi)$ is the $M \times 1$ steering vector of the transmit antenna array and $\phi_1, \ldots, \phi_L$ denote the angle of departure of the multipath components. For further discussion of the above channel model see, for example, [9]. We assume that the $c_l \sim \mathcal{CN}(0, \sigma_l^2)$. The normalizing constant s_L in the model is the square root of the sum of the path variances, i.e. $s_L^2 = \sum_{l=1}^{L} \sigma_l^2$. The angles $\theta_i \sim \theta$ are assumed to be i.i.d. and angles $\phi_i \sim \phi$ are assumed to be i.i.d. Finally all random variables are assumed to be mutually independent.

The following theorem states the connection between the pathwise and separable spatial channel models.

Theorem 2 *Let $\boldsymbol{H}$ be as in (8). Write $\boldsymbol{\Sigma}_a = E\{\boldsymbol{a}(\theta)\boldsymbol{a}^H(\theta)\}$ and $\boldsymbol{\Sigma}_b = E\{\boldsymbol{b}(\phi)\boldsymbol{b}^H(\phi)\}$. Assume that for a given $\epsilon > 0$ and L sufficiently large*

$$\sigma_k < \epsilon s_L, \quad k = 1, \ldots, L.$$

Then, as $L \to \infty$, $\boldsymbol{H}$ converges in distribution to that of the random matrix

$$\boldsymbol{\Sigma}_a^{1/2} \boldsymbol{W} \boldsymbol{\Sigma}_b^{T/2}, \tag{9}$$

where $\boldsymbol{W}$ is an $N \times M$ random matrix with i.i.d. elements that are complex circular Gaussian with mean 0 and variance 1.

The channel matrix in (9) is the channel matrix that is used in the separable spatial correlation channel model.

3.2. Antenna array design

In what follows we assume that the channel matrix is distributed as given in (9). Because of Theorem 2, the obtained results are also approximately valid for the channel model given in (8), when the number of multipaths is large. The analysis is performed under the constraints $E\{||\boldsymbol{a}(\theta)||^2\} = 1$ and $E\{||\boldsymbol{b}(\phi)||^2\} = 1$.

Assume that the realization of the channel $\boldsymbol{H}$ is known at the receiver, but not at the transmitter. Assume also that the input vector $\boldsymbol{x}$ is complex circularly symmetric Gaussian with covariance matrix $\boldsymbol{Q}$. In this case the mutual information between the channel input and output is given by

$$I(\boldsymbol{Q}) = E\{\log\det\left[I + \frac{1}{\sigma_v^2}\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^H\right]\},$$

and the ergodic capacity under the power constraint $E\{\boldsymbol{x}^H\boldsymbol{x}\} \le 1$ is [10]

$$C = \max_{\boldsymbol{Q}:\mathrm{tr}\{\boldsymbol{Q}\}\le 1} I(\boldsymbol{Q}).$$

Our goal is to design array steering vectors $\boldsymbol{a}(\theta)$ and $\boldsymbol{b}(\phi)$ such that the ergodic capacity is maximized.

Under the channel model given in (9), the mutual information is

$$I(\boldsymbol{Q}) = E\{\log\det(\boldsymbol{I} + \rho\boldsymbol{\Sigma}_a^{1/2}\boldsymbol{W}\boldsymbol{\Sigma}_b^{T/2}\boldsymbol{Q}\boldsymbol{\Sigma}_b^{*/2}\boldsymbol{W}^H\boldsymbol{\Sigma}_a^{H/2}\}, \tag{10}$$

where $\rho = \frac{1}{\sigma_v^2}$.

Write $\boldsymbol{\Sigma}_a = \boldsymbol{U}_a\Lambda_a\boldsymbol{U}_a^H$ and $\boldsymbol{\Sigma}_b = \boldsymbol{U}_b\Lambda_b\boldsymbol{U}_b^H$ according to the standard eigenvalue decomposition. If the matrix square roots are considered to be hermitian symmetric, $\boldsymbol{\Sigma}_a^{1/2} = \boldsymbol{U}_a\Lambda_a^{1/2}\boldsymbol{U}_a^H$ and $\boldsymbol{\Sigma}_b^{1/2} = \boldsymbol{U}_b\Lambda_b^{1/2}\boldsymbol{U}_b^H$. It is straightforward to show that

$$I(\boldsymbol{Q}) = E\{\log\det(\boldsymbol{I} + \rho\boldsymbol{\Lambda}_a^{1/2}\boldsymbol{W}\boldsymbol{\Lambda}_b^{1/2}\boldsymbol{Q}'\boldsymbol{\Lambda}_b^{1/2}\boldsymbol{W}^H\boldsymbol{\Lambda}_a^{1/2})\},$$

where $\boldsymbol{Q}' = \boldsymbol{U}_b^T\boldsymbol{Q}\boldsymbol{U}_b^*$.

3.2.1. Receiver side

Let $\boldsymbol{\Gamma}$ be the hermitian symmetric square root matrix of $\Lambda_b^{1/2}\boldsymbol{Q}'\boldsymbol{\Lambda}_b^{1/2}$. Note that

$$I(\boldsymbol{Q}) = E\log\det(\boldsymbol{I} + \rho\boldsymbol{\Gamma}\boldsymbol{W}^H\boldsymbol{\Lambda}_a\boldsymbol{W}\boldsymbol{\Gamma})\}. \tag{11}$$

Since $\mathrm{tr}\{\boldsymbol{\Lambda}_a\} = 1$ and for any $N \times N$ unitary matrix $\boldsymbol{U}$ the distribution of $\boldsymbol{\Gamma}\boldsymbol{W}^H\boldsymbol{U}$ is the same as the distribution of $\boldsymbol{\Gamma}\boldsymbol{W}^H$, the results presented in [10] now imply that for *any* $\boldsymbol{\Gamma}$, the matrix $\boldsymbol{\Lambda}_a$ that maximizes the expression (11) is given by

$$\boldsymbol{\Lambda}_a = \frac{1}{N}\boldsymbol{I}.$$

Hence the optimal array in the receive side should have uncorrelated elements with equal variances.

3.2.2. Transmission side

The remaining task is to maximize

$$C' = \max_{\boldsymbol{Q}:\mathrm{tr}\{\boldsymbol{Q}\}\le 1} I'(\boldsymbol{Q}),$$

where

$$I'(\boldsymbol{Q}) = E\{\log\det(\boldsymbol{I} + \rho/N\boldsymbol{W}\boldsymbol{\Lambda}_b^{1/2}\boldsymbol{Q}'\boldsymbol{\Lambda}_b^{1/2}\boldsymbol{W}^H)\},$$

with respect to $\boldsymbol{\Sigma}_b$. Unfortunately, giving a general solution is not possible, because the optimal solution depends on the value of ρ/N. To see this, let $\rho/N << 1$ (low SNR). In this case we may use a first order approximation

$$\begin{aligned} I'(\boldsymbol{Q}) &\approx E\mathrm{tr}\{\rho/N\boldsymbol{W}\boldsymbol{\Lambda}_b^{1/2}\boldsymbol{Q}'\boldsymbol{\Lambda}_b^{1/2}\boldsymbol{W}^H\} \\ &= \rho\frac{M}{N}\mathrm{tr}\{\boldsymbol{\Lambda}_b\boldsymbol{Q}'\}. \end{aligned}$$

For any $\mathbf{\Lambda}_b$, the optimal matrix $\boldsymbol{Q}'$ is given by

$$\boldsymbol{Q}' = \text{diag}\{0, \ldots, 0, 1, \ldots, 0\},$$

where the only nonzero diagonal element is in the position corresponding to the largest diagonal element of $\mathbf{\Lambda}_b$ (if there is no unique maximum, we may choose a position of any of the "maximum" elements). Hence in this case $C' \approx \rho M/N\lambda_i$, where λ_i is the maximum eigenvalue of $\mathbf{\Lambda}_b$. Because $\text{tr}\Lambda_b = 1$, the optimal covariance matrix should be of the form $\Sigma_b = \boldsymbol{b}\boldsymbol{b}^H$, where $||\boldsymbol{b}||^2 = 1$.

On the other hand, when $N \geq M$ and $\rho/N >> 1$ (high SNR), we may use an approximation

$$\begin{aligned} I'(\boldsymbol{Q}) &\approx E \log\det\{\rho/N \boldsymbol{W}\mathbf{\Lambda}_b^{1/2}\boldsymbol{Q}'\mathbf{\Lambda}_b^{1/2}\boldsymbol{W}^H\} \\ &= \log\det\{\mathbf{\Lambda}_b\} + \log\det\{\boldsymbol{Q}\} + \text{constant} \end{aligned}$$

Therefore, in this case the optimal matrix is given by $\mathbf{\Sigma}_b = \frac{1}{M}\boldsymbol{I}$.

When $N < M$ and $\rho/N >> 1$, the maximization of

$$E\{\det(\boldsymbol{W}\mathbf{\Lambda}_b^{1/2}\boldsymbol{Q}'\mathbf{\Lambda}_b^{1/2}\boldsymbol{W}^H)\},$$

leads to $\mathbf{\Lambda}_b$ of the form

$$\frac{1}{N}\begin{bmatrix} \boldsymbol{I}_N & \mathbf{0}_{N\times(M-N)} \\ \mathbf{0}_{(M-N)\times N} & \mathbf{0}_{(M-N)\times(M-N)} \end{bmatrix}.$$

In this case the eigenvectors of $\mathbf{\Sigma}_b$ are arbitrary.

3.3. Solution examples

Since for the optimal receive array the array covariance matrix should be a constant times the identity matrix, the two examples given in the case of diversity reception fulfill the design criterion. For the receive array for high SNR case these two examples are also (approximately) optimal, when $N \geq M$. For low SNR case the covariance matrix of an optimal array should have rank one. Hence for optimal transmitter array the steering vector should be constant as a function of the angle, or that the transmit array should have only one antenna. The latter case is of course more realizable.

4. CONCLUSION

In the paper we showed that for most of the cases, an optimal array should have uncorrelated elements and the elements should have equal variances. The examples then show that this behavior may be obtained, for example, by two extremes: spatially allpass elements or completely sectorized elements. In case of spatially allpass elements the elements' phase responses depend on the distribution of the angles whereas for sectorized antennas the widths of the sectors depend on the angle distribution.

In practice it is hard to state which solution is more robust to deviations from the assumed angle distribution. In the sectorized solution the elements are always uncorrelated but the requirement of equal variances may be easily violated. On the other hand, in the spatial allpass solution the equal variance assumption is always satisfied, but a small deviation from the assumed angle distribution may cause the elements to be correlated.

In theory, robustifying (hardly realizable) solutions exist. For example, a robustifying sectorized solution consists of choosing the array to have spatial comb filter like sectors, i.e. one element has many sectors of small width and the sectors of the differing elements are beside each other. In this case the equal variance condition holds approximately, if the distribution of the angles is smooth enough inside the sectors.

Further work might involve consideration of outage capacity, and channel knowledge at the transmitter.

A. PROOFS OF THE RESULTS

A.1. Proof of theorem 1

In what follows we use the notation $E_{\boldsymbol{c}}$ ($E_{\boldsymbol{\theta}}$) to denote expectation with respect to the complex amplitudes (DOAs). We start by calculating the expected value of ξ. It is given by

$$\begin{aligned} E\{\xi\} &= E_{\boldsymbol{\theta}} E_{\boldsymbol{c}}\{\boldsymbol{c}^H\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{c}\} = E_{\boldsymbol{\theta}} \text{tr}\{\boldsymbol{A}^H\boldsymbol{A}E_{\boldsymbol{c}}\{\boldsymbol{c}\boldsymbol{c}^H\}\} \\ &= \sum_{i=1}^{L} \sigma_i^2 E_{\theta_i}||\boldsymbol{a}(\theta_i)||^2 \\ &= E||\boldsymbol{a}(\theta)||^2 \sum_{i=1}^{L}\sigma_i^2, \end{aligned}$$

where the last equality follows from the i.i.d. assumption of the angles.

We now proceed by calculating the expectation $E\{|\xi|^2\}$. Using the result of the expectation of the product of Gaussian random vectors [11] we get

$$\begin{aligned} E_{\boldsymbol{\theta}}\{E_{\boldsymbol{c}}|\xi|^2\} &= E_{\boldsymbol{\theta}} E_{\boldsymbol{c}}\{\boldsymbol{c}^H\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{c}\boldsymbol{c}^H\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{c}\} \\ &= E_{\boldsymbol{\theta}}\left\{E_{\boldsymbol{c}}\{\boldsymbol{c}^H\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{c}\}E_{\boldsymbol{c}}\{\boldsymbol{c}^H\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{c}\}\right\} \\ &+ E_{\boldsymbol{\theta}}\{\text{tr}\{\boldsymbol{A}^H\boldsymbol{A}E\{\boldsymbol{c}\boldsymbol{c}^H\}\boldsymbol{A}^H\boldsymbol{A}E\{\boldsymbol{c}\boldsymbol{c}^H\}\}\} \\ &= E_{\boldsymbol{\theta}}\left(\text{tr}\{\boldsymbol{A}^H\boldsymbol{A}\mathbf{\Sigma}_{\boldsymbol{c}}\}\right)^2 + E_{\boldsymbol{\theta}}\text{tr}\left\{\boldsymbol{A}^H\boldsymbol{A}\mathbf{\Sigma}_{\boldsymbol{c}}\boldsymbol{A}^H\boldsymbol{A}\mathbf{\Sigma}_{\boldsymbol{c}}\right\}, \end{aligned}$$

where $\Sigma_{\boldsymbol{c}} = E\{\boldsymbol{c}\boldsymbol{c}^H\}$. Now

$$E_{\boldsymbol{\theta}}\left(\text{tr}\{\boldsymbol{A}^H\boldsymbol{A}\mathbf{\Sigma}_{\boldsymbol{b}}\}\right)^2 = E\{\left(\sum_{i=1}^{L}\sigma_i^2||\boldsymbol{a}(\theta_i)||^2\right)^2\}$$

$$= (\sum_{i=1}^{L}\sigma_i^4)E\{||\boldsymbol{a}(\theta)||^4\} + (\sum_{i\neq j}\sigma_i^2\sigma_j^2)\left[E\{||\boldsymbol{a}(\theta)||^2\}\right]^2,$$

and

$$E_{\boldsymbol{\theta}}\text{tr}\left\{\boldsymbol{A}^H\boldsymbol{A}\mathbf{\Sigma}_{\boldsymbol{c}}\boldsymbol{A}^H\boldsymbol{A}\mathbf{\Sigma}_{\boldsymbol{c}}\right\} = E_{\boldsymbol{\theta}}||\boldsymbol{D}\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{D}||_F^2,$$

where $||\cdot||_F$ denotes the Frobenius norm and $\boldsymbol{D}$ is the matrix square root of $\mathbf{\Sigma}_{\boldsymbol{c}}$: $\boldsymbol{D} = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_L)$. By noting that

$$||\boldsymbol{D}\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{D}||_F^2 = \sum_{i=1}^{L}\sum_{j=1}^{L}|\sigma_i\sigma_j\boldsymbol{a}^H(\theta_i)\boldsymbol{a}(\theta_j)|^2$$

we get

$$\begin{aligned} E_{\boldsymbol{\theta}}||\boldsymbol{D}\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{D}||_F^2 &= \sum_{i=1}^{L}\sigma_i^4 E||\boldsymbol{a}(\theta_i)||^4 \\ &+ \sum_{i\neq j}\sigma_i^2\sigma_j^2 E\{|\boldsymbol{a}^H(\theta_i)\boldsymbol{a}(\theta_j)|^2\} \\ &= \sum_{i=1}^{L}\sigma_i^4 E||\boldsymbol{a}(\theta)||^4 + \sum_{i\neq j}\sigma_i^2\sigma_j^2||\mathbf{\Sigma}_{\boldsymbol{a}}||_F^2 \end{aligned}$$

where

$$
\begin{aligned}
||\boldsymbol{\Sigma}_{\boldsymbol{a}}||_F^2 &= ||E\{\boldsymbol{a}(\theta)\boldsymbol{a}^H(\theta)\}||_F^2 \\
&= \sum_{i=1}^{N}(E|a_i(\theta)|^2)^2 + \sum_{i\neq j}|E\{a_i(\theta)a_j^*(\theta)\}|^2
\end{aligned}
$$

The criterion (1) may hence be written as

$$
\begin{aligned}
\frac{\sigma_\xi^2}{(E\{\xi\})^2} &= \frac{(2\sum_{i=1}^{L}\sigma_i^4)E\{||\boldsymbol{a}(\theta)||^4\}}{\left(E||\boldsymbol{a}(\theta)||^2\sum_{i=1}^{L}\sigma_i^2\right)^2} + \frac{\sum_i\left(E\{|a_i(\boldsymbol{\theta})|^2\}\right)^2}{(E||\boldsymbol{a}(\theta)||^2)^2} \\
&+ \frac{\sum_{i\neq j}|E\{a_i(\theta)a_j^*(\theta)\}|^2}{(E||\boldsymbol{a}(\theta)||^2)^2} + \frac{\sum_{i\neq j}\sigma_i^2\sigma_j^2}{(\sum_i\sigma_i^2)^2} - 1
\end{aligned}
$$

By using Jensen's inequality we get that

$$
\frac{E||\boldsymbol{a}(\theta)||^4}{(E||\boldsymbol{a}(\theta)||^2)^2} \geq 1.
$$

The equality is achieved if and only if $||\boldsymbol{a}(\theta)||^2 = E||\boldsymbol{a}(\theta)||^2$ with probability one. Similarly Jensen's inequality gives that

$$
\frac{\sum_{i=1}^{N}(E|a_i(\theta)|^2)^2}{[E\{||\boldsymbol{a}(\theta)||^2\}]^2} \geq \frac{1}{N}
$$

and the equality is achieved if and only if $E|a_i(\theta)|^2 = \frac{E\{||\boldsymbol{a}(\theta)||^2}{N}$ $\forall i$. Clearly

$$
\sum_{i\neq j}|E\{a_i(\theta)a_j^*(\theta)\}|^2 \geq 0.
$$

Thus the theorem is proved.

A.2. Proof of theorem 2

Define $\boldsymbol{h}_l = c_l\boldsymbol{b}(\phi_l)\otimes\boldsymbol{a}(\theta_l)$. Note that $\text{vec}(\boldsymbol{H}) = s_L^{-1}\sum_{l=1}^{L}\boldsymbol{h}_l$,

$$
E\{\boldsymbol{h}_l\boldsymbol{h}_l^T\} = \boldsymbol{0}_{NM\times NM},
$$

$$
\begin{aligned}
E\{\boldsymbol{h}_l\boldsymbol{h}_l^H\} &= E\{|c_l|^2(\boldsymbol{b}_l\otimes\boldsymbol{a}_l)(\boldsymbol{b}_l^H\otimes\boldsymbol{a}_l^H)\} \\
&= \sigma_l^2E\{(\boldsymbol{b}_l\boldsymbol{b}_l^H)\otimes(\boldsymbol{a}_l\boldsymbol{a}_l^H)\} \\
&= \sigma_l^2\boldsymbol{\Sigma}_b\otimes\boldsymbol{\Sigma}_a \\
&= \boldsymbol{\Sigma}_l.
\end{aligned}
$$

Write $\tilde{\boldsymbol{h}} = [\Re\{\boldsymbol{h}\}^T\Im\{\boldsymbol{h}\}^T]^T$. Note that

$$
\begin{aligned}
\tilde{\boldsymbol{\Sigma}}_l &= E\{\tilde{\boldsymbol{h}}_l\tilde{\boldsymbol{h}}_l^T\} \\
&= \frac{\sigma_l^2}{2}\begin{bmatrix}\Re\{\boldsymbol{\Sigma}_b\otimes\boldsymbol{\Sigma}_a\} & -\Im\{\boldsymbol{\Sigma}_b\otimes\boldsymbol{\Sigma}_a\} \\ \Im\{\boldsymbol{\Sigma}_b\otimes\boldsymbol{\Sigma}_a\} & \Re\{\boldsymbol{\Sigma}_b\otimes\boldsymbol{\Sigma}_a\}\end{bmatrix} \\
&= \frac{\sigma_l^2}{2}\widetilde{\boldsymbol{\Sigma}_b\otimes\boldsymbol{\Sigma}}_a.
\end{aligned}
$$

Let $\boldsymbol{q}$ be any fixed $2MN$-vector. Then under the condition on the path variances stated in the theorem, the central limit theorem states that $s_L^{-1}\sum_{l=1}^{L}\boldsymbol{q}^T\tilde{\boldsymbol{h}}_l\boldsymbol{q}$ converges in distribution to $N(0, \frac{1}{2}\boldsymbol{q}^T\widetilde{\boldsymbol{\Sigma}_b\otimes\boldsymbol{\Sigma}}_a\boldsymbol{q})$ distribution (note that the case $\boldsymbol{q}^T\widetilde{\boldsymbol{\Sigma}_b\otimes\boldsymbol{\Sigma}}_a\boldsymbol{q} = 0$ corresponds to degenerate Gaussian distribution). Therefore we may state that asymptotically

$$
s_L^{-1}\sum_{l=1}^{L}\tilde{\boldsymbol{h}}_l \sim \frac{1}{\sqrt{2}}(\widetilde{\boldsymbol{\Sigma}_b\otimes\boldsymbol{\Sigma}}_a)^{1/2}\tilde{\boldsymbol{w}}
$$

where the elements of the $2NM\times 1$ random vector $\tilde{\boldsymbol{w}}$ are i.i.d. Gaussian with mean 0 and variance 1. Because of the structure of the matrix $\widetilde{\boldsymbol{\Sigma}_b\otimes\boldsymbol{\Sigma}}_a$, asymptotically

$$
s_L^{-1}\sum_{l=1}^{L}\boldsymbol{h}_l \sim (\boldsymbol{\Sigma}_b^{1/2}\otimes\boldsymbol{\Sigma}_a^{1/2})\boldsymbol{w},
$$

where the elements of the $MN\times 1$ random vector $\boldsymbol{w}$ are i.i.d. complex circular Gaussian with mean 0 and variance 1. The result follows because

$$
\text{devec}\{(\boldsymbol{\Sigma}_b^{1/2}\otimes\boldsymbol{\Sigma}_a^{1/2})\boldsymbol{w}\} = \boldsymbol{\Sigma}_a^{1/2}\text{devec}\{\boldsymbol{w}\}\boldsymbol{\Sigma}_b^{T/2}.
$$

B. REFERENCES

[1] F. Demmerle and W. Wiesbeck, "A biconical multibeam antenna for space-division multiple access," *IEEE Trans. Antennas and Propagation*, vol. 46, no. 6, pp. 782–787, 1998.

[2] T. Svantesson, "An antenna solution for MIMO channels: The multimode antenna," in *Proc. 34th Asilomar Conf. Signals Systems and Computers*, vol. 2, pp. 1617–1621, 2000.

[3] U. Baysal and R. Moses, "On the geometry of isotropic wideband arrays," in *Proc. of Int'l Conf. on Acoust. Speech and Signal Proc.*, vol. 2, pp. 3045–3048, 2002.

[4] F. Athley, "Optimization of element positions for direction finding with sparce arrays," in *Proc. of the 11th IEEE Signal Procesing Workshop on Statistical Signal Processing*, pp. 516–519, 2001.

[5] Y. Erhel and L. Bertel, "A method of direction finding operating on an heterogeneous array," in *Proc. EUSIPCO*, 1998.

[6] C.-N. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 637–650, 2002.

[7] C. Martin and B. Ottersten, "Analytic approximations of eigenvalue moments and mean channel capacity for MIMO channels," in *Proc. of Int'l Conf. on Acoust. Speech and Signal Proc.*, pp. 2839–2392, 2002.

[8] D. Gespert, H. Bolcskei, H. Gore, and A. Paulraj, "MIMO wireless channels: Capacity and performance prediction," in *Proc. of 2000 IEEE Global Telecommunications Conference*, vol. 2, pp. 1083–1088, 2000.

[9] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Transactions on Communications*, vol. 46, no. 3, pp. 357–366, 1998.

[10] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[11] P. Janssen and P. Stoica, "On the expectation of the product of four matrix-valued Gaussian random variables," *IEEE Trans. Automatic Control*, vol. 33, no. 9, pp. 867–870, 1988.