ON MULTICAST BEAMFORMING FOR MINIMUM OUTAGE

Diploma Thesis

By

VASILEIOS G. NTRANOS

Submitted to the Department of Electronic Eng & Computer Engineering (ECE)

in partial fulfillment of the requirements for the ECE Diploma Degree.

TECHNICAL UNIVERSITY OF CRETE

Advisor: Professor Nicholaos Sidiropoulos

Co-advisor: Associate Professor Athanasios Liavas

Co-advisor: Associate Professor Georgios Karistinos

August 2008

Dedication

To my father George, my mother Joanne and my brother Achilleas,

who offered me unconditional love and support

throughout the course of this thesis.

ACKNOWLEDGMENTS

I would like to express my gratitude to my supervisor, Prof. Nicholas Sidiropoulos, whose expertise and research insight added considerably to my undergraduate experience. His inspiration, enthusiasm and incredible ability of explaining things, have been the ingredients to the successful completion of this thesis. I would also like to thank Prof. Leandros Tassioulas whose ideas were the cornerstone of this work. Finally, I would like to thank all of my friends who believed in me from the very beginning and especially Evita for her love, patience and understanding.

CONTENTS

LIST OF FIGURES

ABSTRACT

The multicast beamforming problem is considered from the viewpoint of minimizing outage probability subject to a transmit power constraint. The main difference with the point-to-point transmit beamforming problem is that in multicast beamforming the channel is naturally modeled as a Gaussian mixture, as opposed to a single Gaussian distribution. The Gaussian components in the mixture model user clusters of different means (locations) and variances (spreads). It is shown that minimizing outage probability subject to a transmit power constraint is an NP-hard problem when the number of Gaussian kernels, J, is greater than or equal to the number of transmit antennas, N. Through dimensionality reduction, it is also shown that the problem is practically tractable for $2 - 3$ Gaussian kernels. An approximate solution based on the Markov inequality is also proposed. This is simple to compute for any J and N, and often works well in practice.

1. INTRODUCTION

Consider a base station or wireless access point that uses an antenna array to transmit *common* information to a pool of users, each equipped with a single receive antenna. When the channel vectors of all users are known at the transmitter, it is possible to beamform in a way that directs power towards the users and limits wasteful radiation in other directions. This is a physical layer multicasting approach that has been recently investigated in a series of papers [7, 5, 4, 3]. The design formulations in [7, 5, 4, 3] target signal to (interference plus) noise ratio (SNR) guarantees: they either minimize total transmitted power subject to guaranteed SNR for each receiver, or maximize the minimum SNR subject to an overall transmitted power constraint.

Exact channel state information (CSI) will never be available in practice, in which case it is impossible to guarantee instantaneous SNR. An alternative is to offer average (expected) SNR guarantees. The channel correlation matrices (which vary far slower than the actual channel realizations) are then sufficient for transmit optimization, and the solutions in [7, 5, 4, 3] carry over almost verbatim.

The drawback is that persistent deep fading can occur in this case, which is unacceptable for delay-sensitive applications. An alternative is to start with a set of nominal channel vectors, allow limited perturbation, and aim for a conservative design that guarantees a certain SNR for every allowable perturbation; see [2, 4] for related results in the context of multicasting, and references therein for earlier work on robust unicast beamforming.

A different approach is pursued here. The channel vector is modeled as random, with a known distribution. The objective is to design the weight vector of the transmit beamformer to minimize the outage probability, i.e., the probability that the useful received signal power falls below a certain threshold. In a multicast context, this has the following interpretation: If one draws a large number of channel vectors, then (under ergodic mixing conditions) the fraction of terminals served will be approximately one minus the outage probability. Minimizing the outage probability thus approximately maximizes the number of users served. In a single-user (unicast) context, averaging is with respect to the temporal channel variation, and minimizing outage approximately maximizes the fraction of time that the channel meets the quality of service demand.

Minimum outage probability beamforming has been considered in the literature, in the context of point-to-point multiple-input single-output systems [8] and receive beamforming for the cellular uplink [1]. The main difference with our setup is that in [8, 1] the channel is modeled using a (single) Gaussian distribution, whereas we adopt a Gaussian mixture. This is natural for wireless multicasting, where subscribers are spatially dispersed in a non-uniform fashion (e.g., clustered in malls, squares, campuses, etc). Also, given enough kernels, it is possible to approximate almost any density by a Gaussian mixture. Finally, we note that unicast beamforming under an outage probability constraint has been considered in [10].

2. RESULTS

Let the channel vectors be drawn from a Gaussian mixture distribution

$$f(\mathbf{h}) = \sum_{j=1}^{J} p_j \mathcal{N}(\mathbf{h}; \mathbf{m}_j, \sigma_j^2 \mathbf{I}),$$

where $\mathcal{N}((\cdot); \mathbf{m}, \mathbf{C})$ denotes a multivariate Gaussian distribution of mean vector $\mathbf{m}$ and covariance matrix $\mathbf{C}$, assumed diagonal for simplicity. Let $y := \mathbf{w}^T \mathbf{h}$. Then,

$$f(y; \mathbf{w}) = \sum_{j=1}^{J} p_j \mathcal{N}(y; \mathbf{w}^T \mathbf{m}_j, \sigma_j^2 ||\mathbf{w}||^2),$$

i.e., a mixture of univariate Gaussians (throughout, $||\cdot||$ denotes the Euclidean norm). Consider

$$\min_{||\mathbf{w}||^2 = P} Pr[|y| < \gamma] \iff$$

$$\min_{||\mathbf{w}||^2 = P} \sum_{j=1}^{J} p_j \int_{-\gamma}^{\gamma} \mathcal{N}(y; \mathbf{w}^T \mathbf{m}_j, \sigma_j^2 ||\mathbf{w}||^2) \iff$$

$$\min_{||\mathbf{w}||^2 = P} \sum_{j=1}^{J} p_j \int_{-\gamma}^{\gamma} \mathcal{N}(y; \mathbf{w}^T \mathbf{m}_j, \sigma_j^2 P).$$

2.1 NP-HARDNESS

Claim 1: Computing $\min_{||\mathbf{w}||^2=P} Pr[|y| < \gamma]$ is NP-hard.

Proof 1: Consider the special case where $\sigma_j = \sigma$, $p_j = 1/J$, $\forall j$. Using the Cauchy-Schwartz inequality, it can be shown (cf. [9]) that

$$|\mathbf{w}^T\mathbf{m}_j| - \epsilon||\mathbf{w}|| \leq |\mathbf{w}^T\mathbf{h}_j| \leq$$

$$|\mathbf{w}^T\mathbf{m}_j| + \epsilon||\mathbf{w}||, \forall \mathbf{h}_j \in B_\epsilon(\mathbf{m}_j),$$

where $B_\epsilon(\mathbf{m}_j)$ denotes a ball of radius ϵ centered at $\mathbf{m}_j$. Let $\mathbf{h}_j$ be drawn from the j-th component pdf $\mathcal{N}(\mathbf{h}; \mathbf{m}_j, \sigma^2\mathbf{I})$. Given ϵ and $\delta > 0$, we can pick $\sigma = \sigma(\epsilon, \delta)$ such that $Pr[\mathbf{h}_j \in B_\epsilon(\mathbf{m}_j)] \geq 1 - \delta$. Let

$$p_{out}(\mathbf{w}) := \frac{1}{J}\sum_{j=1}^{J} p_{out|j}(\mathbf{w}),\ p^*_{out} := \min_{||\mathbf{w}||^2=P} p_{out}(\mathbf{w}),$$

where $p_{out|j}(\mathbf{w}) := Pr[|\mathbf{w}^T\mathbf{h}_j| < \gamma]$. With $1(\text{condition})$ denoting the indicator function, and $E_{\mathbf{h}_j}[\cdot]$ the expectation conditioned on the j-th component,

$$p_{out|j}(\mathbf{w}) = E_{\mathbf{h}_j}[1(|\mathbf{w}^T\mathbf{h}_j| < \gamma)].$$

For $\mathbf{h}_j \in B_\epsilon(\mathbf{m}_j)$, it holds

$$|\mathbf{w}^T\mathbf{m}_j| + \epsilon||\mathbf{w}|| < \gamma \Rightarrow |\mathbf{w}^T\mathbf{h}_j| < \gamma,$$

and therefore

$$1(|\mathbf{w}^T\mathbf{m}_j| < \gamma - \epsilon||\mathbf{w}||) \leq 1(|\mathbf{w}^T\mathbf{h}_j| < \gamma).$$

It follows that, for all $\mathbf{h}_j$,

$$1(|\mathbf{w}^T\mathbf{m}_j| < \gamma - \epsilon||\mathbf{w}||)1(\mathbf{h}_j \in B_\epsilon(\mathbf{m}_j)) \leq 1(|\mathbf{w}^T\mathbf{h}_j| < \gamma).$$

Taking $E_{\mathbf{h}_j}[\cdot]$ we obtain

$$(1-\delta)1(|\mathbf{w}^T\mathbf{m}_j| < \gamma - \epsilon||\mathbf{w}||) \leq p_{out|j}(\mathbf{w}).$$

In a similar vain, for $\mathbf{h}_j \in B_\epsilon(\mathbf{m}_j)$, it holds

$$|\mathbf{w}^T\mathbf{h}_j| < \gamma \Rightarrow |\mathbf{w}^T\mathbf{m}_j| - \epsilon||\mathbf{w}|| < \gamma,$$

and therefore

$$1(|\mathbf{w}^T\mathbf{h}_j| < \gamma) \leq 1(|\mathbf{w}^T\mathbf{m}_j| < \gamma + \epsilon||\mathbf{w}||).$$

It follows that, for all $\mathbf{h}_j$,

$$1(|\mathbf{w}^T\mathbf{h}_j| < \gamma) \leq 1(|\mathbf{w}^T\mathbf{m}_j| < \gamma + \epsilon||\mathbf{w}||)1(\mathbf{h}_j \in B_\epsilon(\mathbf{m}_j)) +$$

$$1 - 1(\mathbf{h}_j \in B_\epsilon(\mathbf{m}_j)),$$

where the last term is a trivial upper bound that applies to the complement of $B_\epsilon(\mathbf{m}_j)$. Again taking $E_{\mathbf{h}_j}[\cdot]$ we obtain

$$p_{out|j}(\mathbf{w}) \leq (1-\delta)1(|\mathbf{w}^T\mathbf{m}_j| < \gamma + \epsilon||\mathbf{w}||) + \delta.$$

Combining the two inequalities, we have

$$(1-\delta)1(|\mathbf{w}^T\mathbf{m}_j| < \gamma - \epsilon||\mathbf{w}||) \leq p_{out|j}(\mathbf{w}) \leq$$

$$(1-\delta)1(|\mathbf{w}^T\mathbf{m}_j| < \gamma + \epsilon||\mathbf{w}||) + \delta,$$

Averaging out over j and taking the minimum over $\mathbf{w}$ with $||\mathbf{w}|| = \sqrt{P}$ yields

$$\frac{1-\delta}{J} \min_{||\mathbf{w}||=\sqrt{P}} \sum_{j=1}^{J} 1(|\mathbf{w}^T\mathbf{m}_j| < \gamma - \epsilon\sqrt{P}) \leq p^*_{out}(\gamma) \leq$$

$$\frac{1-\delta}{J} \min_{||\mathbf{w}||=\sqrt{P}} \sum_{j=1}^{J} 1(|\mathbf{w}^T\mathbf{m}_j| < \gamma + \epsilon\sqrt{P}) + \delta,$$

where we have also made explicit that p^*_{out} depends on γ. It follows that

$$p^*_{out}(t - \epsilon\sqrt{P}) - \delta \leq \frac{1-\delta}{J} \min_{||\mathbf{w}||=\sqrt{P}} \sum_{j=1}^{J} 1(|\mathbf{w}^T\mathbf{m}_j| < t) \leq$$

$$p^*_{out}(t + \epsilon\sqrt{P}),$$

for all $t \in (\epsilon\sqrt{P}, 1 - \epsilon\sqrt{P})$. Notice now that $p^*_{out}(\cdot)$ is a continuous function, whereas

$$\frac{1-\delta}{J} \min_{||\mathbf{w}||=\sqrt{P}} \sum_{j=1}^{J} 1(|\mathbf{w}^T\mathbf{m}_j| < t)$$

only takes discrete values, separated by $\frac{1-\delta}{J}$. Recall that $\epsilon > 0$, $\delta > 0$, but otherwise up to our control. Pick $0 < \delta < \frac{1}{J+1}$ (which implies $\delta < \frac{1-\delta}{J}$) and ϵ sufficiently small to sandwich $\frac{1-\delta}{J} \min_{||\mathbf{w}||=\sqrt{P}} \sum_{j=1}^{J} 1(|\mathbf{w}^T\mathbf{m}_j| < t)$ within an interval strictly less than $\frac{1-\delta}{J}$. This leaves no ambiguity - computing $p^*_{out}(t \pm \epsilon\sqrt{P})$

pin-points the exact value of $\min_{||\mathbf{w}||=\sqrt{P}} \sum_{j=1}^{J} 1(|\mathbf{w}^T \mathbf{m}_j| < t)$. In particular, this answers the question of whether or not it is possible to find a $\mathbf{w}$ of norm $\sqrt{P}$ such that $|\mathbf{w}^T \mathbf{m}_j| \geq t$, $\forall j \in \{1, \cdots, J\}$. The latter is the decidability version of a problem shown to be NP-hard in [7] for $J \geq N$.

2.2 Special case: $J = 1$

When there is only one Gaussian kernel ($J = 1$), minimizing $Pr[|y| < \gamma]$ under $||\mathbf{w}||^2 = P$ reduces to maximizing $|\mathbf{w}^T \mathbf{m}_1|$ under the same constraint. From the Cauchy-Schwartz inequality, the optimum $\mathbf{w}$ is simply $\mathbf{m}_1$ scaled to power P (note there is freedom to choose the sign; if $\mathbf{m}_1 = \mathbf{0}$, then any $\mathbf{w}$ on the sphere of radius $\sqrt{P}$ is equally good). The solution is trivial in this case - but also interesting in the following way: If one draws a large number of channel vectors, then $1 - Pr[|y| < \gamma]$ is an estimate of the fraction of terminals that will be served, thus picking $\mathbf{w}$ to minimize $Pr[|y| < \gamma]$ approximately maximizes the number of terminals served in the "large sample" regime (the fraction converges to $1 - Pr[|y| < \gamma]$ under quite general ergodic mixing conditions, and notably when the channel vectors are drawn from a product distribution with $\mathcal{N}(\mathbf{h}; \mathbf{m}_1, \sigma_1^2 \mathbf{I})$ as marginal). This is interesting, because even if the channel vectors are exactly known at the transmitter, exactly maximizing the number of terminals served is NP-hard in this case, and even approximate solutions are non-trivial, see [6]. Thus, when the number of terminals is large, we can approximately maximize the number of terminals served by matching the weight vector to the mean vector; but exactly maximizing the number of terminals served is prohibitive, even when all channel vectors are known exactly at the transmitter.

Remark 1: The above two results generalize to the complex case (channel elements and weights are both complex-valued), using results from [7] and [8].

2.3 CASE: $J = 2$, GENERAL N

By adding a second Gaussian kernel in the channel vector model, the problem of minimizing the outage probability becomes non-trivial:

$$\min_{||\mathbf{w}||^2=P} \sum_{j=1}^{2} p_j \int_{-\gamma}^{\gamma} \mathcal{N}(y; \mathbf{w}^T\mathbf{m}_j, \sigma_j^2 P). \iff$$

$$\min_{||\mathbf{w}||^2=P} \sum_{j=1}^{2} p_j \left[\mathcal{Q}\left(\frac{-\gamma - \mathbf{w}^T\mathbf{m}_j}{\sigma_j}\right) - \mathcal{Q}\left(\frac{\gamma - \mathbf{w}^T\mathbf{m}_j}{\sigma_j}\right)\right].$$

Let

$$\mathcal{M}(\mathbf{w}) = \sum_{j=1}^{2} p_j \left[\mathcal{Q}\left(\frac{-\gamma - \mathbf{w}^T\mathbf{m}_j}{\sigma_j}\right) - \mathcal{Q}\left(\frac{\gamma - \mathbf{w}^T\mathbf{m}_j}{\sigma_j}\right)\right],$$

Then, $\mathbf{w}_o = arg\min_{||\mathbf{w}||^2=P} \mathcal{M}\left(\mathbf{w}\right)$ will lie on the subspace $\mathcal{V}$, spanned by the mean vectors $\mathbf{m}_1, \mathbf{m}_2$ (otherwise, power allocated in a direction out of $\mathcal{V}$ would be wasted). Then $\mathbf{w}^T\mathbf{m}_j$ can be parameterized as $||\mathbf{w}||||\mathbf{m}_j|| \cos(\angle w - \angle m_j)$, where $\angle \mathbf{x} := \arccos(\mathbf{x}^T\mathbf{v}_r/||\mathbf{x}||||\mathbf{v}_r||)$ is the angle between a vector $\mathbf{x}$ and a reference vector $\mathbf{v}_r$ in the two dimentional space $\mathcal{V}$. For simplicity we may take $\mathbf{v}_r = \mathbf{m}_1/||\mathbf{m}_1||$, and the objective function becomes

$$\mathcal{M}(\angle w) = \sum_{j=1}^{2} p_j \Bigg[\mathcal{Q}\Bigg(\frac{-\gamma - \sqrt{P}||\mathbf{m}_j|| \cos(\angle w - \angle m_j)}{\sigma_j}\Bigg) \\ - \mathcal{Q}\Bigg(\frac{\gamma - \sqrt{P}||\mathbf{m}_j|| \cos(\angle w - \angle m_j)}{\sigma_j}\Bigg)\Bigg],$$

and $\angle w_o = arg\min_{||\mathbf{w}||^2=P} \mathcal{M}(\angle w)$ determines $\mathbf{w}_o = arg\min_{||\mathbf{w}||^2=P} \mathcal{M}\left(\mathbf{w}\right)$ as follows:

$$\mathbf{w}_o = \sqrt{P}\mathbf{V}\mathbf{Q}\mathbf{V}^T\mathbf{v}_r,$$

where

$$\mathbf{Q} := \begin{bmatrix} \cos(\angle w) & -\sin(\angle w) \\ \sin(\angle w) & cos(\angle w) \end{bmatrix}$$

is a rotation matrix, $\mathbf{V} := [\mathbf{v}_1, \mathbf{v}_2]$ is the orthonormal basis of $\mathcal{V}$ and $\mathbf{v}_r$ is the reference vector used above.

We can therefore find the optimal $\angle w_o$ and $\mathbf{w}_o$ (up to desired accuracy) by performing one-dimensional fine grid search of $\mathcal{M}(\angle w)$. Alternatively, we can use a relatively coarser grid search, followed by a steepest descent iteration. The objective function

$$\mathcal{M}(\mathbf{w}) = \sum_{j=1}^{J} p_j \Bigg[\mathcal{Q}\Bigg(\frac{-\gamma - \mathbf{w}^T\mathbf{m}_j}{\sigma_j}\Bigg) - \mathcal{Q}\Bigg(\frac{\gamma - \mathbf{w}^T\mathbf{m}_j}{\sigma_j}\Bigg)\Bigg],$$

is differentiable with

$$\nabla\mathcal{M}(\mathbf{w}) = \sum_{j=1}^{J} \frac{p_j\mathbf{m}_j}{\sigma_j^2\sqrt{2\pi}} \Bigg[e^{\frac{(-\gamma - \mathbf{w}^T\mathbf{m}_j)^2}{2\sigma_j^2}} - e^{\frac{(\gamma - \mathbf{w}^T\mathbf{m}_j)^2}{2\sigma_j^2}} \Bigg].$$

In our experiments, the results of coarser grid search followed by steepest descent were comparable to those obtained using fine grid search.

2.4 CASE: $J = 3$, GENERAL N

The case of three Gaussian kernels in the channel model is a practical upper bound on computing the globally optimal beamforming weight vector through grid search of the objective function

$$\mathcal{M}(\mathbf{w}) = \sum_{j=1}^{3} p_j \left[\mathcal{Q}\left(\frac{-\gamma - \mathbf{w}^T \mathbf{m}_j}{\sigma_j} \right) - \mathcal{Q}\left(\frac{\gamma - \mathbf{w}^T \mathbf{m}_j}{\sigma_j} \right) \right].$$

Again, $\mathbf{w}_o = arg\min_{||\mathbf{w}||^2=P} \mathcal{M}(\mathbf{w})$ will lie on the subspace $\mathcal{V}$, spanned by the mean vectors $\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3$. Let $\mathbf{V} := [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$ be an orthonormal basis of subspace $\mathcal{V}$. We can use spherical coordinates (r, θ, ϕ) to write the weight and mean vectors as:

$$\mathbf{w} = \sqrt{P}\mathbf{V}\hat{\mathbf{r}}_{\theta_w,\phi_w}, \ \mathbf{m}_j = ||\mathbf{m}_j||\mathbf{V}\hat{\mathbf{r}}_{\theta_j,\phi_j},$$

where

$$\hat{\mathbf{r}}_{\theta_j,\phi_j} = \begin{bmatrix} \cos(\theta_j)\sin(\phi_j) \\ \sin(\theta_j)\sin(\phi_j) \\ \cos(\phi_j) \end{bmatrix}$$

is the spherical position vector, θ_j is the angle between the projection of the j^{th} mean vector on $\mathbf{v}_1\mathbf{v}_2$ plane and $\mathbf{v}_1$, and ϕ_j is the angle between the j^{th} mean vector and $\mathbf{v}_3$.

Now, exactly as in case of $J = 2$, we may write $\mathcal{M}(\mathbf{w})$ as a function of θ_w and ϕ_w and perform a two-dimentional fine grid search in $\mathcal{M}(\theta_w, \phi_w)$ to efficiently compute $\{\theta_{w_o}, \phi_{w_o}\} = arg \min_{||\mathbf{w}||^2=P} \mathcal{M}\left(\theta_w, \phi_w\right)$. The optimal beamformng vector is then given by $\mathbf{w}_o = \sqrt{P}\mathbf{V}\hat{\mathbf{r}}_{\theta_{w_o},\phi_{w_o}}$.

2.5 COMPLEX CASE: $J = 2$, GENERAL N

We now generalize to the complex case.

$$f(\mathbf{h}) = \sum_{j=1}^{J} p_j \mathcal{CN}(\mathbf{h}; \mathbf{m}_j, \sigma_j^2 \mathbf{I}),$$

where $\mathcal{CN}((\cdot); \mathbf{m}, \mathbf{C})$ denotes a complex multivariate Gaussian distribution of mean vector $\mathbf{m}$ and covariance matrix $\mathbf{C}$, assumed diagonal for simplicity. Let $z := \mathbf{w}^H \mathbf{h}$, where ^{H} denotes Hermitian (conjugate) transpose. Then

$$f(z; \mathbf{w}) = \sum_{j=1}^{J} p_j \mathcal{CN}(z; \mathbf{w}^H \mathbf{m}_j, \sigma_j^2 ||\mathbf{w}||^2).$$

The optimal beamforming vector can be found by,

$$\min_{||\mathbf{w}||^2=P} Pr[|z| < \gamma] \Longleftrightarrow$$

$$\min_{||\mathbf{w}||^2=P} \sum_{j=1}^{J} p_j \iint_{\mathbf{A}} \mathcal{CN}(z; \mathbf{w}^H \mathbf{m}_j, \sigma_j^2 ||\mathbf{w}||^2),$$

where $\mathbf{A}$ is a disc of radius γ in the complex plane. The above integral is given by [11]

$$\iint_{\mathbf{A}} \mathcal{CN}(z; \mathbf{w}^H \mathbf{m}_j, \sigma_j^2 ||\mathbf{w}||^2) = \mathcal{P}\left[\left(\frac{\gamma}{\sigma_j ||\mathbf{w}||} \right)^2 \Big|_2, \left(\frac{|\mathbf{w}^H \mathbf{m}_j|}{\sigma_j ||\mathbf{w}||} \right)^2 \right]$$

where $\mathcal{P}[\chi^2|_2, \lambda]$ is the cdf of the non-central χ^2 distribution with two degrees of freedom and non-centrality parameter λ. Let

$$\mathcal{C}(\mathbf{w}) := \sum_{j=1}^{J} p_j \mathcal{P}\left[\left(\frac{\gamma}{\sigma_j||\mathbf{w}||}\right)^2\Big|_2, \left(\frac{|\mathbf{w}^H\mathbf{m}_j|}{\sigma_j||\mathbf{w}||}\right)^2\right].$$

Again, $\mathbf{w}_o = arg\min_{||\mathbf{w}||^2=P} \mathcal{C}(\mathbf{w})$ will lie on the subspace spanned by the complex mean vectors $\mathbf{m}_1, \mathbf{m}_2$. Thus, all candidate beamforming vectors are $\mathbf{w} = c_1\mathbf{m}_1 + c_2\mathbf{m}_2$, with c_1, c_2 complex numbers such that $||\mathbf{w}||^2 = P$. We can use this constraint to find a relationship between c_1 and c_2:

$$||\mathbf{w}||^2 = P \Longleftrightarrow$$

$$(c_1^*\mathbf{m}_1^H + c_2^*\mathbf{m}_2^H)(c_1\mathbf{m}_1 + c_2\mathbf{m}_2) = P \Longleftrightarrow$$

$$|c_1|^2||\mathbf{m}_1||^2 + |c_2|^2||\mathbf{m}_2||^2 + 2\Re\{c_1^*c_2\mathbf{m}_1^H\mathbf{m}_2)\} = P.$$

Using the rotational invariance of the outage probability in the complex plane $Pr[|\mathbf{w}^H\mathbf{h}| < \gamma] = Pr[|e^{j\omega}\mathbf{w}^H\mathbf{h}| < \gamma]$ we can take c_2 to be real without loss of generality. Thus:

$$|c_1|^2||\mathbf{m}_1||^2 + c_2^2||\mathbf{m}_2||^2 + 2c_2\Re\{c_1^*\mathbf{m}_1^H\mathbf{m}_2)\} = P.$$

We can now compute the optimal beamforming vector by performing a two-dimensional grid search in $\mathcal{C}(\mathbf{w})$ -one dimension for $\angle c_1 \in [0, 2\pi)$ and one for $|c_1| \in (0, \sqrt{P}/||\mathbf{m}_1||)$. For every c_1, we can compute c_2 through the constraint equation

$$c_{21,2} = \frac{-\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha},$$

where

$$\alpha = ||\mathbf{m}_2||^2,$$

$$\beta = 2\Re\{c_1^* \mathbf{m}_1^H \mathbf{m}_2\},$$

$$\gamma = |c_1|^2 ||\mathbf{m}_1||^2 - P.$$

Note that we need to check only one root for c_2: For every $c_1, c_1' = -c_1 \in \mathbb{C}$ it is easy to see that $c_{21} = -c_{22}' \Rightarrow c_1\mathbf{m}_1 + c_{21}\mathbf{m}_2 = -(c_1'\mathbf{m}_1 + c_{22}'\mathbf{m}_2) \Rightarrow \mathbf{w} = -\mathbf{w}'$. These two beamforming vectors are equivalent in terms of minimizing the outage probability because $\mathcal{C}(\mathbf{w}) = \mathcal{C}(-\mathbf{w}) = \mathcal{C}(\mathbf{w}')$.

2.6 General Covariance matrix case

Up to this point we have made the assumption that channel vectors are drawn from a Gaussian mixture distribution with diagonal covariance matrix. In general, one may have to drop that assumption to best fit his channel model. In that case, $f(\mathbf{h}) = \sum_{j=1}^{J} p_j \mathcal{CN}(\mathbf{h}; \mathbf{m}_j, \mathbf{C}_j)$ and the minimization problem becomes:

$$\min_{||\mathbf{w}||^2 = P} \sum_{j=1}^{J} p_j \iint_{\mathbf{A}} \mathcal{CN}(z; \mathbf{w}^H \mathbf{m}_j, \mathbf{w}^H \mathbf{C}_j \mathbf{w}).$$

Note that the optimal beamforming vector no longer lies on the subspace spanned by the J mean vectors since $\mathbf{w}$ affects not only the mean of each univariate Gaussian in the mixture, but its variance as well. We will only consider a special but important case in the sequel.

2.6.1 Special case: real-valued, J=1

The problem of minimizing the outage probability when $J = 1$ becomes non-trivial in the general covariance matrix case. The outage probability is given by

$$M(\mathbf{w}) = \int_{-\gamma}^{\gamma} \mathcal{N}(y; \mathbf{w}^T\mathbf{m}, \mathbf{w}^T\mathbf{C}\mathbf{w}).$$

Consider problem

$$Q : \min_{||\mathbf{w}||^2=P} M(\mathbf{w}).$$

Claim 2: The optimal beamforming vector, $\mathbf{w}_o = arg\min_{||\mathbf{w}||^2=P} M(\mathbf{w})$, will lie on the subspace spanned by the mean vector $\mathbf{m}$ and the principal eigenvector of the covariance matrix $\mathbf{C}$.

Proof 2: The minimization problem Q can equivalently be written as:

$$Q' : \quad \min_{|c|\leq\sqrt{P}||\mathbf{m}||} M \begin{pmatrix} arg\max \mathbf{w}^T\mathbf{C}\mathbf{w} \\ s.t : ||\mathbf{w}||^2 = P \\ \mathbf{w}^T\mathbf{m} = c \end{pmatrix}.$$

Let

$$\mathbf{w}_o(c) := arg\max \mathbf{w}^T\mathbf{C}\mathbf{w}$$

$$s.t : ||\mathbf{w}||^2 = P$$

$$\mathbf{w}^T\mathbf{m} = c$$

Every $\mathbf{w}$ that satisfies the above constrains is equidistant from $\frac{c\cdot\mathbf{m}}{||\mathbf{m}||^2}$ and can be written as:

$$\mathbf{w} = \frac{c\cdot\mathbf{m}}{||\mathbf{m}||^2} + \mathbf{v},$$

with $\mathbf{v}^T\mathbf{m} = 0$ and $||\mathbf{v}||^2 = P - c^2/||\mathbf{m}||^2$. The maximization problem can be reformulated in terms of $\mathbf{v}$ as:

$$\mathbf{v}_o(c) := arg\max[\mathbf{m}^T\mathbf{C}\mathbf{m}(c^2/||\mathbf{m}||^2)^2 + \mathbf{v}^T\mathbf{C}\mathbf{v}]$$

$$s.t : ||\mathbf{v}||^2 = P - c^2/||\mathbf{m}||^2$$

$$\mathbf{v}^T\mathbf{m} = 0$$

which is equivalent to

$$\mathbf{v}_o(c) := arg\max \mathbf{v}^T\mathbf{C}\mathbf{v}$$

$$s.t : ||\mathbf{v}||^2 = P - c^2/||\mathbf{m}||^2$$

$$\mathbf{v}^T\mathbf{m} = 0,$$

since $\mathbf{m}^T\mathbf{C}\mathbf{m}(c^2/||\mathbf{m}||^2)^2$ is a constant. Furthermore, $\mathbf{C}$ can be written as:

$$\begin{aligned}\mathbf{C} =& \mathbf{U}^T\mathbf{D}\mathbf{U} \\ &= (\mathbf{U_m} + \mathbf{U}')^T\mathbf{D}(\mathbf{U_m} + \mathbf{U}') \\ &= \mathbf{U_m}^T\mathbf{D}\mathbf{U_m} + \mathbf{U}'^T\mathbf{D}\mathbf{U}',\end{aligned}$$

where $\mathbf{U_m} = \frac{\mathbf{m}\mathbf{m}^T}{||\mathbf{m}||^2}\mathbf{U}$ and $\mathbf{U}' = \mathbf{U} - \frac{\mathbf{m}\mathbf{m}^T}{||\mathbf{m}||^2}\mathbf{U}$, are the parallel and perpendicular to $\mathbf{m}$ components of $\mathbf{U}$ respectively. Substituting, we have:

$$\begin{aligned}\mathbf{v}_o(c) := arg\max \mathbf{v}^T\mathbf{U_m}^T\mathbf{D}\mathbf{U_m}\mathbf{v} + \mathbf{v}^T\mathbf{U}'^T\mathbf{D}\mathbf{U}'\mathbf{v} \\ s.t: ||\mathbf{v}||^2 = P - c^2/||\mathbf{m}||^2 \\ \mathbf{v}^T\mathbf{m} = 0,\end{aligned}$$

which yields

$$\begin{aligned}\mathbf{v}_o(c) := arg\max \mathbf{v}^T\mathbf{U}'^T\mathbf{D}\mathbf{U}'\mathbf{v} \\ s.t: ||\mathbf{v}||^2 = P - c^2/||\mathbf{m}||^2 \\ \mathbf{v}^T\mathbf{m} = 0,\end{aligned}$$

since $\mathbf{v}^T\mathbf{U_m}^T\mathbf{D}\mathbf{U_m}\mathbf{v} = 0$. The solution to the latter problem is:

$$\begin{aligned}\mathbf{v}_o(c) &= \frac{\sqrt{P - \frac{c^2}{||\mathbf{m}||^2}}}{||\mathbf{u_1'}||} \cdot \mathbf{u_1'} \\ &= \frac{\sqrt{P - \frac{c^2}{||\mathbf{m}||^2}}}{||\mathbf{u_1} - \frac{\mathbf{u_1}^T\mathbf{m}}{||\mathbf{m}||^2} \cdot \mathbf{m}||} \cdot (\mathbf{u_1} - \frac{\mathbf{u_1}^T\mathbf{m}}{||\mathbf{m}||^2} \cdot \mathbf{m}),\end{aligned}$$

where $\mathbf{u_1}'$ and $\mathbf{u_1}$ are the principal eigenvectors of $\mathbf{U}'^T\mathbf{D}\mathbf{U}'$ and $\mathbf{C}$ respectively. Consequently, $\mathbf{w}_o(c) = \mathbf{v}_o(c) + \frac{c \cdot \mathbf{m}}{||\mathbf{m}||^2}$ will lie in the subspace spanned by $[\mathbf{m}, \mathbf{u}_1]$.

The above analysis holds for every c in $[-\sqrt{P}||\mathbf{m}||, \sqrt{P}||\mathbf{m}||]$, and that completes the proof.

In order to find the optimal beamforming vector in the general covariance matrix case we only need to perform one-dimensional fine grid search in $M(\sqrt{P}\mathbf{V}\mathbf{Q}(\theta)\mathbf{V}^T\mathbf{v}_r)$, where

$$\mathbf{Q}(\theta) := \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & cos(\theta) \end{bmatrix}$$

is a rotation matrix, $\mathbf{V}$ is the orthonormal basis of the subspace spanned by the mean vector and the principal eigenvector of the covariance matrix, and $\mathbf{v}_r$ is a reference vector in that space.

2.7 Markov Approximation

From Markov's inequality we have that $Pr[x \geq t] \leq t^{-1}E[x]$, for any non-negative random variable. We can thus consider approximating (the real case is considered for simplicity, but the approach generalizes to the complex case)

$$\min_{||\mathbf{w}||^2=P} Pr[|y| < \gamma] \Longleftrightarrow \max_{||\mathbf{w}||^2=P} Pr[|y| \geq \gamma] \Longleftrightarrow$$

$$\max_{||\mathbf{w}||^2=P} Pr[y^2 \geq \gamma^2]$$

by

$$\max_{||\mathbf{w}||^2=P} E[y^2],$$

thus maximizing an upper bound on the actual objective function (when put in maximization form). Now,

$$E[y^2] = \sum_{j=1}^{J} p_j \left((\mathbf{w}^T \mathbf{m}_j)^2 + \sigma_j^2 P \right),$$

thus we may

$$\max_{||\mathbf{w}||^2 = P} \sum_{j=1}^{J} p_j (\mathbf{w}^T \mathbf{m}_j)^2.$$

Solution of the latter problem is easy. Let

$$\mathbf{D} := diag([\sqrt{p_1}, \cdots, \sqrt{p_J}\,]),\ \mathbf{M} := [\mathbf{m}_1, \cdots, \mathbf{m}_J]^T,$$

then $\mathbf{w}_{app} = arg \max_{||\mathbf{w}||^2 = P} E[y^2]$ is given by the principal right singular vector of the matrix $\mathbf{DM}$ scaled to power P. Of course, $\mathbf{w}_o$ does not in general solve the original problem of minimizing outage (maximizing service) probability; but it is interesting to note that in the special case of $J = 1$ (single Gaussian Kernel) it does. Also note that $\mathbf{w}_o$ *is not* $\sum_{j=1}^{J} p_j \mathbf{m}_j$ normalized to power P, as quick intuition would perhaps suggest. To appreciate this, consider for example what happens when $J = 2$, $p_1 = p_2 = 1/2$, and $\mathbf{m}_2 = -\mathbf{m}_1$.

3. NUMERICAL RESULTS

3.1 REAL CASE

In the cases when $J = 2$ and $J = 3$ we can efficiently compute the optimal beamforming vector by performing low-dimensional fine grid search, as already shown. In this case, we can evaluate how far is the solution based on Markov approximation from the optimal one. In four different scenarios for each case ($J = 2, J = 3$), we computed $\mathbf{w}_{opt} = arg\min_{||\mathbf{w}||^2=P} \mathcal{M}(\mathbf{w})$ and $\mathbf{w}_{app} = arg\max_{||\mathbf{w}||^2=P} E[y^2]$ through a fine grid search algorithm and the Markov approximation respectively. The parameters for the different scenarios are given below. The results are summarized in Figs. 3.1 - 3.8, where curves are parameterized by the number of transmit antennas, N.

3.1.1 ANGLE BETWEEN MEAN VECTORS (J=2)

Fig. 3.1 plots outage probability results for $p_1 = p_2 = 1/2$, $\sigma_1^2 = \sigma_2^2 = 1$, $||\mathbf{m}_1||^2 = ||\mathbf{m}_2||^2 = N$, $||\mathbf{w}||^2 = P = 4$, as $\hat{\phi} := \angle m_2 - \angle m_1$ varies in $[0, \pi)$.

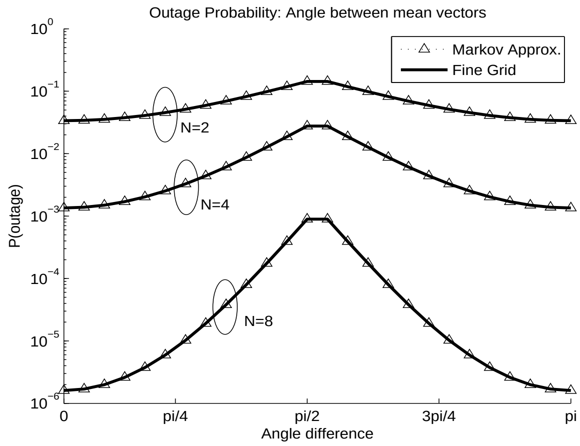

Fig. 3.1: Angle between mean vectors($J = 2$): Outage Probability as a function of the angle: $\mathcal{M}(\mathbf{w}_{opt})$ and $\mathcal{M}(\mathbf{w}_{apr})$ versus $\hat{\phi}$.

3.1.2 MAGNITUDE OF MEAN VECTORS (J=2)

Fig. 3.2 plots outage probability results for $p_1 = p_2 = 1/2$, $\sigma_1^2 = \sigma_2^2 = 1$, $\hat{\phi} = \pi/3$, $||\mathbf{m}_1||^2 = N$, $||\mathbf{w}||^2 = P = 4$, as $||\mathbf{m}_2||$ varies in $(0, 10]$.

3.1.3 VARIANCE OF THE GAUSSIAN KERNELS (J=2)

Fig. 3.3 plots outage probability results for $p_1 = p_2 = 1/2$, $\hat{\phi} = \pi/3$, $||\mathbf{m}_1||^2 = ||\mathbf{m}_2||^2 = N$, $\sigma_1^2 = 1$, $||\mathbf{w}||^2 = P = 4$, as σ_2^2 varies in $(0, 10]$.

3.1.4 MIXTURE PROBABILITY (J=2)

Fig. 3.4 plots outage probability results for $\sigma_1^2 = \sigma_2^2 = 1$, $||\mathbf{m}_1||^2 = ||\mathbf{m}_2||^2 = N$ $\hat{\phi} = \pi/3$, $||\mathbf{w}||^2 = P = 4$, as p_1 varies in $[0.1, 0.9]$.

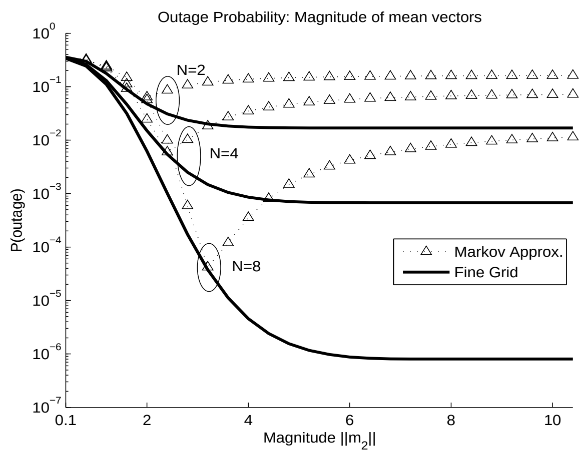

Fig. 3.2: Magnitude of mean vectors($J = 2$): Outage Probability as a function of the magnitude: $\mathcal{M}(\mathbf{w}_{opt})$ and $\mathcal{M}(\mathbf{w}_{apr})$ versus $||\mathbf{m}_2||$.

3.1.5 ANGLE BETWEEN MEAN VECTORS (J=3)

Fig. 3.5 plots outage probability results for $p_j = 1/3$, $\sigma_j^2 = 1$, $||\mathbf{m}_j||^2 = N,\ \forall j$, $(\theta_1, \phi_1) = (\frac{\pi}{4}, \frac{\pi}{4})$, $(\theta_2, \phi_2) = (\frac{3\pi}{4}, \frac{\pi}{4})$, $\theta_3 = \frac{\pi}{4}$, $||\mathbf{w}||^2 = P = 4$, as ϕ_3 varies in $[0, \pi)$.

3.1.6 MAGNITUDE OF MEAN VECTORS (J=3)

Fig. 3.6 plots outage probability results for $p_j = 1/3$, $\sigma_j^2 = 1,\ \forall j$, $(\theta_1, \phi_1) = (\frac{\pi}{4}, \frac{\pi}{4})$, $(\theta_2, \phi_2) = (\frac{3\pi}{4}, \frac{\pi}{4})$, $(\theta_3, \phi_3) = (\frac{\pi}{2}, \frac{\pi}{2})$, $||\mathbf{m}_1||^2 = ||\mathbf{m}_2||^2 = N$, $||\mathbf{w}||^2 = P = 4$, as $||\mathbf{m}_3||$ varies in $(0, 5]$.

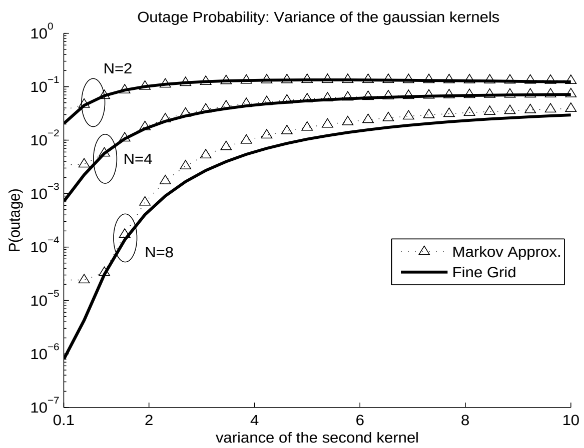

Fig. 3.3: Variance of the Gaussian Kernels($J = 2$): Outage Probability as a function of the variance: $\mathcal{M}(\mathbf{w}_{opt})$ and $\mathcal{M}(\mathbf{w}_{apr})$ versus σ_2^2.

3.1.7 VARIANCE OF THE GAUSSIAN KERNELS (J=3)

Fig. 3.7 plots outage probability results for $p_j = 1/3$, $||\mathbf{m}_j||^2 = N,\ \forall j$, $(\theta_1, \phi_1) = (\frac{\pi}{4}, \frac{\pi}{4})$, $(\theta_2, \phi_2) = (\frac{3\pi}{4}, \frac{\pi}{4})$, $(\theta_3, \phi_3) = (\frac{\pi}{2}, \frac{\pi}{2})$, $\sigma_1^2 = 2$, $\sigma_2^2 = 1$, $||\mathbf{w}||^2 = P = 4$, as σ_3^2 varies in $(0, 5]$.

3.1.8 MIXTURE PROBABILITY (J=3)

Fig. 3.8 plots outage probability results for $\sigma_j^2 = 1$, $||\mathbf{m}_j||^2 = N,\ \forall j$, $(\theta_1, \phi_1) = (\frac{\pi}{4}, \frac{\pi}{4})$, $(\theta_2, \phi_2) = (\frac{3\pi}{4}, \frac{\pi}{4})$, $(\theta_3, \phi_3) = (\frac{\pi}{2}, \frac{\pi}{2})$, $p_1 = p_2 = (1 - p_3)/2$,,$||\mathbf{w}||^2 = P = 4$, as p_3 varies in $[0.1, 0.9]$.

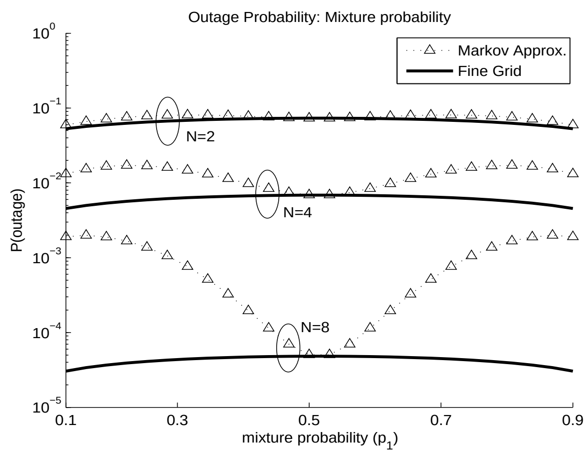

Fig. 3.4: Mixture Probability($J = 2$): Outage Probability as a function of the mixture probability: $\mathcal{M}(\mathbf{w}_{opt})$ and $\mathcal{M}(\mathbf{w}_{apr})$ versus p_1.

3.2 COMPLEX CASE

We now turn to the complex case. Only $J = 2$ kernels are considered. The results are summarized in Figs. 3.9 - 3.12, where curves are parameterized by the number of transmit antennas, N.

3.2.1 ANGLE BETWEEN MEAN VECTORS

Fig. 3.9 plots outage probability results for $p_1 = p_2 = 1/2$, $\sigma_1^2 = \sigma_2^2 = 1$, $||\mathbf{m}_1||^2 = ||\mathbf{m}_2||^2 = N$, $||\mathbf{w}||^2 = P = 1$, as the angle between the mean vectors, θ, varies in $[0, \pi)$.

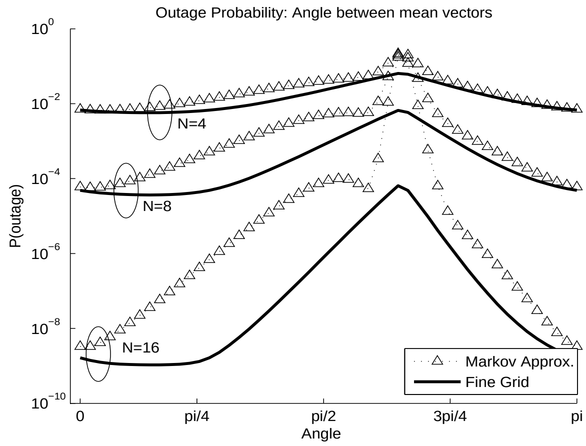

Fig. 3.5: Angle between mean vectors($J = 3$): Outage Probability as a function of the angle: $\mathcal{M}(\mathbf{w}_{opt})$ and $\mathcal{M}(\mathbf{w}_{apr})$ versus ϕ_3.

3.2.2 MAGNITUDE OF MEAN VECTORS

Fig. 3.10 plots outage probability results for $p_1 = p_2 = 1/2, \;\; \sigma_1^2 = \sigma_2^2 = 1$, $\theta = \pi/3$, $||\mathbf{m}_1||^2 = N$, $||\mathbf{w}||^2 = P = 1$, as $||\mathbf{m}_2||$ varies in $(0, 5]$.

3.2.3 VARIANCE OF THE GAUSSIAN KERNELS

Fig. 3.11 plots outage probability results for $p_1 = p_2 = 1/2, \;\; \theta = \pi/3$, $||\mathbf{m}_1||^2 = ||\mathbf{m}_2||^2 = N$, $\sigma_1^2 = 1$, $||\mathbf{w}||^2 = P = 1$, as σ_2^2 varies in $(0, 5]$.

3.2.4 MIXTURE PROBABILITY

Fig. 3.12 plots outage probability results for $\sigma_1^2 = \sigma_2^2 = 1, \;\; ||\mathbf{m}_1||^2 = ||\mathbf{m}_2||^2 = N \; \theta = \pi/3$, $||\mathbf{w}||^2 = P = 1$, as p_1 varies in$[0.1, 0.9]$.

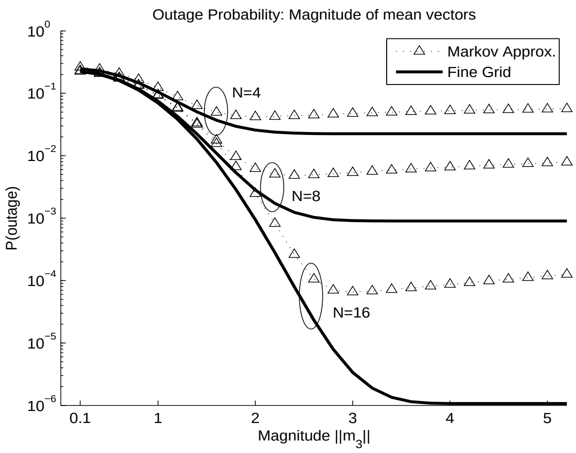

Fig. 3.6: Magnitude of mean vectors($J = 3$): Outage Probability as a function of the magnitude: $\mathcal{M}(\mathbf{w}_{opt})$ and $\mathcal{M}(\mathbf{w}_{apr})$ versus $||\mathbf{m}_3||$.

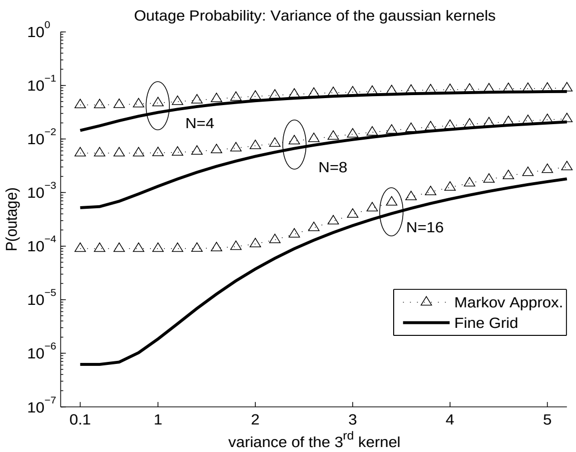

Fig. 3.7: Variance of the Gaussian Kernels($J = 3$): Outage Probability as a function of the variance: $\mathcal{M}(\mathbf{w}_{opt})$ and $\mathcal{M}(\mathbf{w}_{apr})$ versus σ_3^2.

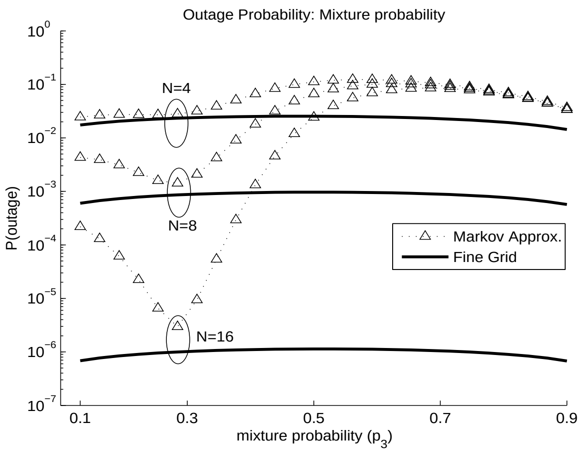

Fig. 3.8: Mixture Probability($J = 3$): Outage Probability as a function of the mixture probability: $\mathcal{M}(\mathbf{w}_{opt})$ and $\mathcal{M}(\mathbf{w}_{apr})$ versus p_3.

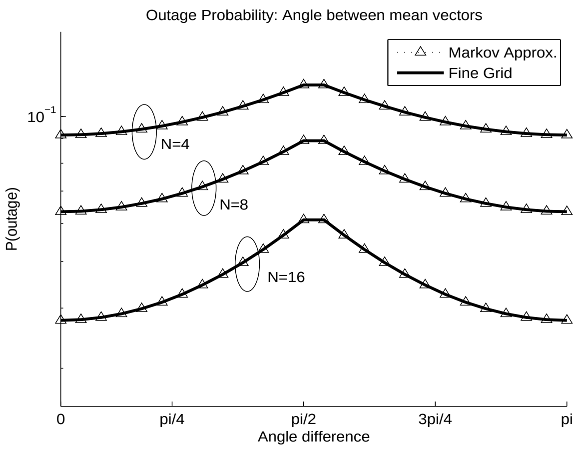

Fig. 3.9: Angle between mean vectors: Outage Probability as a function of the angle: $\mathcal{C}(\mathbf{w}_{opt})$ and $\mathcal{C}(\mathbf{w}_{apr})$ versus θ.

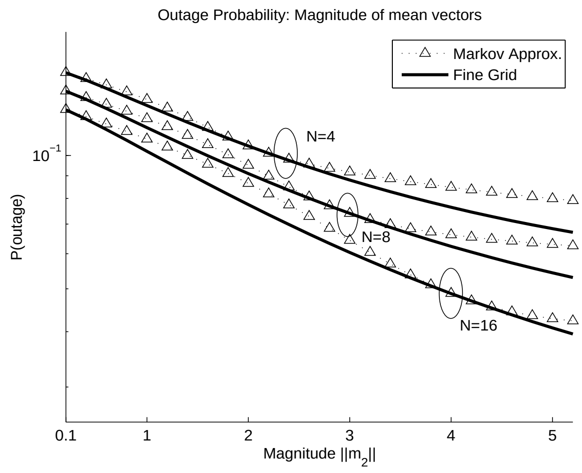

Fig. 3.10: Magnitude of mean vectors: Outage Probability as a function of the magnitude: $\mathcal{C}(\mathbf{w}_{opt})$ and $\mathcal{C}(\mathbf{w}_{apr})$ versus $||\mathbf{m}_2||$.

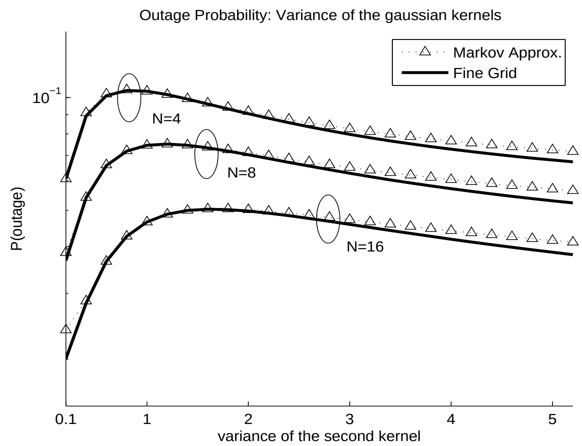

Fig. 3.11: Variance of the Gaussian Kernels: Outage Probability as a function of the variance: $\mathcal{C}(\mathbf{w}_{opt})$ and $\mathcal{C}(\mathbf{w}_{apr})$ versus σ_2^2.

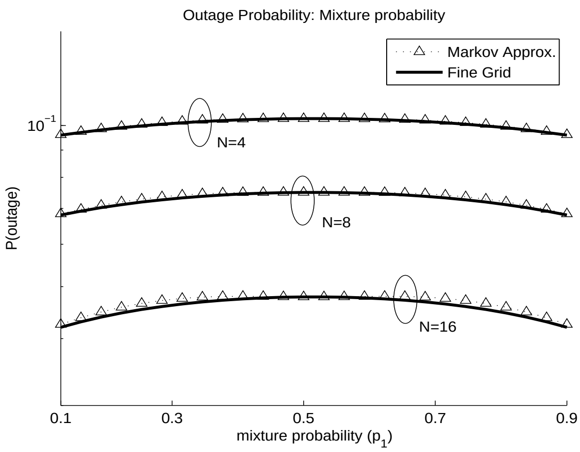

Fig. 3.12: Mixture Probability: Outage Probability as a function of the mixture probability: $\mathcal{C}(\mathbf{w}_{opt})$ and $\mathcal{C}(\mathbf{w}_{apr})$ versus p_1.

4. CONCLUSIONS

The multicast beamforming problem was considered from the viewpoint of minimizing outage probability subject to a transmit power constraint. In a multicast context, the channel is naturally modeled as a Gaussian mixture, as opposed to a single Gaussian distribution. The different Gaussian kernels model user clusters of different means (locations) and variances (spreads). It was shown that minimizing outage probability subject to a transmit power constraint is an NP-hard problem when the number of Gaussian kernels, J, is greater than or equal to the number of transmit antennas, N. Through dimensionality reduction, it was also shown that the problem is practically tractable for $2-3$ Gaussian kernels. An approximate solution based on the Markov inequality was also proposed.

In the real case, the Markov approximation can be very accurate, but appears sensitive to near-far and mixture probability imbalance effects. For a large number of transmit antennas, N, the Markov approximation brakes down in the presence of such imbalances - the gap from the optimal solution is significant. The reason is that the principal right singular vector of the matrix $\mathbf{DM}$ then tends to align with the dominant component(s), effectively ignoring weaker ones.

Interestingly, the Markov approximation seems to be far more accurate in the complex case. This corroborates findings in [5], which showed that related approximation problems are easier in the complex case.

BIBLIOGRAPHY

[1] B.K. Chalise, A. Czylwik, A., "Robust uplink beamforming based upon minimum outage probability criterion," in *Proc. Global Telecommunications Conference (GLOBECOM)* 29 Nov.-3 Dec. 2004, vol.6, pp. 3974-3978.

[2] E. Karipidis, N.D. Sidiropoulos, Z.-Q. Luo, "Convex Transmit Beamforming for Downlink Multicasting to Multiple Co-channel Groups," *in Proc. IEEE ICASSP 2006*, May 14-19, 2006, Toulouse, France.

[3] E. Karipidis, N.D. Sidiropoulos, Z.-Q. Luo, "Quality of Service and Max-min-fair Transmit Beamforming to Multiple Co-channel Multicast Groups," *IEEE Trans. on Signal Proc.*, vol. 56, no. 3, pp. 1268-1279, Mar. 2008.

[4] E. Karipidis, N.D. Sidiropoulos, Z.-Q. Luo, "Far-field Multicast Beamforming for Uniform Linear Antenna Arrays," *IEEE Trans. on Signal Processing*, vol. 55, no. 10, pp. 4916-4927, Oct. 2007.

[5] Z.-Q. Luo, N.D. Sidiropoulos, P. Tseng, P., and S. Zhang, "Approxima-

tion Bounds for Quadratic Optimization with Homogeneous Quadratic Constraints," *SIAM Journal on Optimization*, vol. 18, no. 1, pp. 1-28, Feb. 2007.

[6] E. Matskani, N.D. Sidiropoulos, and L. Tassiulas, "On Multicast Beamforming and Admission Control for UMTS-LTE," *Proc. ICASSP 2008*, Mar. 30 - Apr. 4, 2008, Las Vegas, Nevada.

[7] N.D. Sidiropoulos, T.N. Davidson, and Z-Q. Luo, "Transmit Beamforming for Physical Layer Multicasting," *IEEE Transactions on Signal Processing*, vol. 54, no. 6, part 1, pp. 2239-2251, June 2006.

[8] Y. Xie, C.N. Georghiades, A. Arapostathis, "Minimum Outage Probability Transmission With Imperfect Feedback for MISO Fading Channels," *IEEE Trans. on Wireless Communications*, vol. 4, no. 3, pp. 1084-1091, May 2005.

[9] S.A. Vorobyov, A.B. Gershman, Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: a solution to the signal mismatch problem,", *IEEE Trans. on signal Processing*, vol. 51, no. 2, pp. 313-324, Feb. 2003.

[10] S.A. Vorobyov, Y. Rong, A.B. Gershman, "Robust adaptive beamforming using probability-constrained optimization," in *Proc. IEEE Statistical Signal Processing Workshop*, 17-20 July 2005, pp. 934-939.

[11] M. Abramowitz and I. A. Stegun. Handbook of Mathematical Functions and Formulas, Graphs, and Mathematical Tables. Chapter 26.3.24. Government Printing Office, Washington, D.C., 1972.