On the Diversity-Multiplexing Tradeoff of Unconstrained Multiple-Access Channels

Yair Yona, and Meir Feder, Fellow, IEEE

Abstract

In this work the optimal diversity-multiplexing tradeoff (DMT) is investigated for the multiple-input multiple-output (MIMO) fading multiple-access channels with no power constraints (infinite constellations). For $K$ users ($K > 1$), $M$ transmit antennas for each user, and $N$ receive antennas, infinite constellations in general and lattices in particular are shown to attain the optimal DMT of finite constellations for the case $N \geq (K + 1)M - 1$, i.e. user limited regime. On the other hand for the case $N < (K + 1)M - 1$ it is shown that infinite constellations cannot attain the optimal DMT. This is in contrast to the point-to-point case where infinite constellations are DMT optimal for any $M$ and $N$. In general, this work shows that when the network is heavily loaded, i.e. $K > \max(1, \frac{N-M+1}{M})$, taking into account the shaping region in the decoding process plays a crucial role in pursuing the optimal DMT. By investigating the cases where infinite constellations are optimal and suboptimal, this work also gives a geometrical interpretation to the DMT of IC’s in multiple-access channels.

I. INTRODUCTION

Employing multiple antennas in a point-to-point wireless channel increases the number of degrees of freedom available for transmission. This is illustrated for the ergodic case in [1,2], where $M$ transmit and $N$ receive antennas increase the capacity by a factor of $\min(M, N)$. The number of degrees of freedom utilized by the transmission scheme is referred to as \textit{multiplexing gain}. Another advantage of employing multiple antennas is the potential increase in the transmitted signal reliability. The fact that multiple antennas increase the number of independent links between antenna pairs, enables the error probability to decrease, i.e. add diversity. If for high signal to noise ratio (SNR) the error probability is proportional to $\text{SNR}^{-d}$, then we state that the \textit{diversity order} is $d$.

For the point-to-point setting, Zheng and Tse [3] characterized the optimal diversity-multiplexing tradeoff (DMT) of the quasi-static Rayleigh flat-fading channel, i.e. for each multiplexing gain they found the best attainable diversity order. The optimal DMT is a piecewise linear function connecting the points $(M - l)(N - l), l = 0, \ldots, \min(M, N)$. The transmission scheme in [3] uses random codes. Subsequent works presented more structured schemes that attain the optimal DMT. El Gamal et al. [4] showed by using probabilistic methods that lattice space-time (LAST) codes attain the optimal DMT by using minimum-mean square error (MMSE) estimation followed by lattice decoding. Later, explicit coding schemes based on lattices and cyclic-division algebra [5], [6] were shown to attain the optimal DMT by using maximum-likelihood (ML) decoding, and also by using MMSE estimation followed by lattice decoding [7]. These coding schemes take into consideration the finiteness of the codebook in the decoder. A question that remained open was whether lattices can achieve the optimal DMT by using \textit{regular} lattice decoding, i.e. decoder that takes into account the infinite lattice without considering the shaping region or the power constraint. In order to answer this question, the work in [8] presented an analysis of the performance of infinite constellations (IC’s) in multiple-input multiple-output (MIMO) fading channels. A new tradeoff was presented between the IC’s average number of dimensions per channel use, i.e. the IC dimensionality divided by the number of channel uses, and the best attainable DMT. By choosing the right average number of dimensions per channel use, it was shown [8] that IC’s in general and more specifically lattices using regular lattice decoding, attain the optimal DMT of finite constellations.

For multiple-access channels, where a number of users transmit to a single receiver, the number of users in the network affects the multiplexing gain and diversity order. For instance, for a network with $K$ users transmitting at the same rate, the number of available degrees of freedom for each user is $\min(M, N)$. Tse et al. [9] characterized the optimal DMT of a network with $K$ users, where each user has $M$ transmit antennas and the receiver has $N$ antennas. For the symmetric case, where the users transmit at the same multiplexing gain $r$, i.e. $r_1 = \cdots = r_K = r$, the optimal DMT takes the following elegant form [9]:

- For $r \in \left[0, \min \left(\frac{N}{K+1}, M \right) \right]$ the optimal symmetric DMT equals to the optimal DMT of a point-to-point channel with $M$ transmit and $N$ receive antennas $d_{M,N}^{c,(FC)} (r)$.
- For $r \in \left[\min \left(\frac{N}{K+1}, M \right), \min \left(M, \frac{N}{K} \right) \right]$ the optimal symmetric DMT equals to the optimal DMT of a point-to-point channel with all $K$ users pulled together $d_{K,M,N}^{c,(FC)} (Kr)$.

This research was supported by THE ISRAEL SCIENCE FOUNDATION grant No. 634/09. The authors are with the Department of Electrical Engineering – Systems, Tel-Aviv University, Ramat-Aviv 69978, Israel (e-mails: {yairyo,meir}@eng.tau.ac.il).
Similar to the evolution in the point-to-point case, random codes were used in [9]. Later Nam and El Gamal [10] showed that a random ensemble of LAST codes attains the optimal DMT for multiple-access channels using MMSE estimation followed by lattice decoding over the lattice induced by the $K$ users. An explicit coding scheme based on lattices and cyclic division algebra that attains the optimal DMT using ML decoding was presented in [11].

In this paper we investigate the optimal DMT of lattices using regular lattice decoding, i.e. decoding without taking into consideration the power constraint, in MIMO Rayleigh fading multiple-access channels. The result is rather surprising; unlike the point-to-point case where the tradeoff between dimensions and diversity enables to attain the optimal DMT, we show that for the multiple-access channels the optimal DMT is attained only when $N \geq (K + 1) M - 1$, i.e. user limited regime. On the other hand when the network is heavily loaded we show that IC’s or lattices using regular lattice decoding, can not attain the optimal DMT.

In the first part of this paper an upper bound on the optimal symmetric DMT IC’s can achieve is derived. The upper bound is attained by finding for each multiplexing gain $r$, the average number of dimensions per channel use for each user, that maximizes the diversity order. For the case $N < (K + 1) M - 1$ it is shown that the optimal DMT of IC’s does not coincide with the optimal DMT of finite constellations. Moreover, for $N < (K - 1) M + 1$ it is shown that the optimal DMT of IC’s in the symmetric case is inferior compared to the optimal DMT of finite constellations, for any value of $r$ except for the edges $r = 0, \frac{1}{N}$. On the other hand for the case $N \geq (K + 1) M - 1$, by choosing the correct average number of dimensions per channel use for each user, it is shown that the upper bound on the optimal DMT of IC’s coincides with the optimal DMT of finite constellations $d_{M,N}^{*(PC)}(\max (r_1, \ldots, r_K))$.

In the second part of this paper, a transmission scheme that attains the optimal DMT for $N \geq (K + 1) M - 1$ is presented. Each user in this scheme transmits according to the DMT optimal scheme for the point-to-point channel, presented in [8]. By analyzing the receiver joint ML decoding performance, it is shown that this transmission scheme attains the optimal DMT of finite constellations. We wish to emphasize that the proposed transmission scheme is more involved than simply using orthogonalization between users, which in general is shown to be suboptimal for IC’s. The proposed transmission scheme requires $N + M - 1$ channel uses to attain the optimal DMT, which is smaller than $N + K M - 1$, the number of channel uses required in [9] (the dependence in the number of users lies in the fact that $N \geq (K + 1) M - 1$). Finally, the algebraic analysis of the transmission scheme geometrically explains why for $N \geq (K + 1) M - 1$ the optimal DMT equals to the optimal DMT of the point-to-point channel of each user, i.e. why the optimal DMT equals $d_{M,N}^{*(PC)} (\max (r_1, \ldots, r_K))$.

As a basic illustrative example of the results we consider the following two cases. In the first case assume a network with two users ($K = 2$), where each user has a single transmit antenna ($M = 1$), and a receiver with a single receive antenna ($N = 1$). In this case the optimal DMT of finite constellations in the symmetric case [2] equals $1 - r$ for $r \in [0, \frac{1}{4}]$, and $2 - 4r$ for $r \in [\frac{1}{3}, \frac{1}{2}]$. For IC’s it is shown in this setting that the optimal DMT for the symmetric case equals $1 - 2r$ for $r \in [0, \frac{1}{2}]$, which is strictly inferior except for $r = 0, \frac{1}{2}$. In the second case, by merely adding another receive antenna, i.e. $M = 1, N = K = 2$, the optimal DMT of IC’s coincides with finite constellations optimal DMT $d_{1,2}^{*(PC)} (\max (r_1, r_2))$.

It is important to note that for the case $N < (K + 1) M - 1$ this paper shows the sub-optimality of IC’s compared to the optimal DMT of finite constellations. However, in this case an explicit analytical expression for the upper bound on the optimal DMT of IC’s is given only for the symmetric case, where for the general case the upper bound is presented in the form of optimization problem. Indeed, for the case $N < (K + 1) M - 1$ finding an explicit expression to the general upper bound (the non-symmetric case) on the optimal DMT of IC’s, together with a transmission scheme that attains it, still remains an open problem. On the other hand, for the case $N \geq (K + 1) M - 1$ this paper provides both analytical upper bound to the optimal DMT of IC’s, and also a transmission scheme that attains it.

The outline of the paper is as follows. In section III basic definitions for the fading multiple-access channels and IC’s are given. Section IV presents an upper bound on the optimal DMT of IC’s, and shows the sub-optimality of IC’s for the case $N < (K + 1) M - 1$. Transmission scheme that attains the optimal DMT of finite constellations for the case $N \geq (K + 1) M - 1$ is presented in section V. Finally, in section VI we discuss the results in this paper and present a geometrical interpretation to the DMT of IC’s in multiple-access channels.

II. BASIC DEFINITIONS

A. Channel Model

We consider $K$-user multiple access channels where each user has $M$ transmit antennas, and the receiver has $N$ antennas. We assume perfect knowledge of all channels at the receiver, and no channel knowledge at the transmitters. We also assume quasi static flat-fading channel for each user. The channel model is as follows:

$$y_t = \sum_{i=1}^{K} H^{(i)} \cdot x^{(i)} + \rho^{-\frac{1}{2}} w_t \quad t = 1, \ldots, T$$

where $x^{(i)}$, $t = 1, \ldots, T$ is user $i$ transmitted signal, $w_t \sim \mathcal{CN}(0, \frac{1}{\rho^{2}} I_N)$ is the additive noise where $\mathcal{CN}$ denotes complex-normal, $I_N$ is the $N$-dimensional unit matrix, and $y_t \in \mathbb{C}^N$. $H^{(i)}$ is the fading matrix of user $i$. It consists of $N$ rows and $M$
Similarly to the definitions in [8] the multiplexing gain of user’s 
channel uses. Let us consider user 
block on the diagonal equals
The VNR in the transmitter of user 
there exists rotated 
Let \( \sqrt{\lambda_i} \geq \cdots \geq \sqrt{\lambda_1} > 0 \). For large values of \( \rho \), we state that \( f(\rho) \gtrsim g(\rho) \) when \( \lim_{\rho \to \infty} \frac{\ln(f(\rho))}{\ln(\rho)} \geq \frac{\ln(g(\rho))}{\ln(\rho)} \), and also define \( \lesssim \), \( \doteq \) in a similar manner by substituting \( \gtrsim \) with \( \lesssim \), respectively.

**B. Infinite Constellations**

Infinite constellation (IC) is a countable set \( S = \{ s_1, s_2, \ldots \} \subset \mathbb{C}^n \). Let cube\( \{a\} \subset \mathbb{C}^n \) be a (probably rotated) \( l \)-complex dimensional cube \( \{ a \} \subset \mathbb{C}^n \) with edge of length \( a \) centered around zero. We define an IC \( S_i \) to be \( l \)-complex dimensional if there exists rotated \( l \)-complex dimensional cube \( \text{cube}(a) \) such that \( S_i \subset \text{cube}(a) \) and \( l \) is minimal. \( M(S_i, a) = \#(S_i \bigcap \text{cube}(a)) \) is the number of points of the IC \( S_i \) inside \( \text{cube}(a) \). In [12], the \( n \)-complex dimensional IC density was defined as

\[
\gamma_G = \limsup_{a \to \infty} \frac{M(S_n, a)}{a^{2n}}
\]

and the volume to noise ratio (VNR) for the additive white Gaussian noise (AWGN) channel was given as

\[
\mu_G = \frac{\gamma_G}{2\pi e\sigma^2}
\]

where \( \sigma^2 \) is the noise variance of each component.

We now turn to the IC definitions in the transmitters. We define the average number of dimensions per channel use as the IC dimension divided by the number of channel uses. Let us consider user \( i \), where \( 1 \leq i \leq K \). We denote the average number of dimensions per channel use by \( D_i \). Let us consider a \( D_i T \)-complex dimensional sequence of IC’s - \( S_{D_i T}^{(i)}(\rho) \), where \( D_i \leq M \), \( T \) is the number of channel uses, and \( \sum_{i=1}^{K} D_i \leq L \). First we define \( \gamma_{tr}^{(i)} = \rho^{\frac{1}{1-T}} \) as the density of \( S_{K T}^{(i)}(\rho) \) in transmitter \( i \). Similarly to the definitions in [8] the multiplexing gain of user’s IC is defined as

\[
r_i = \lim_{\rho \to \infty} \frac{1}{T} \log_\rho (\gamma_{tr}^{(i)} + 1) = \lim_{\rho \to \infty} \frac{1}{T} \log_\rho (\rho^{\frac{1}{1-T}} + 1), \quad 0 \leq r_i \leq D_i.
\]

The VNR in the transmitter of user \( i \) is

\[
\mu_{tr}^{(i)} = \frac{\gamma_{tr}^{(i)} - \frac{1}{1-T}}{2\pi e\sigma^2}
\]

where \( \sigma^2 = \frac{1}{2\pi e} \) is each component’s additive noise variance. Now let us concatenate the users IC’s in accordance with [9]. We denote \( D = \sum_{i=1}^{K} D_i \). The concatenation yields an equivalent \( DT \)-complex dimensional IC, \( S_{D T}^{(i)}(\rho) \), that has multiplexing gain \( \sum_{i=1}^{K} r_i \), density \( \gamma_{tr} = \rho^{\frac{1}{1-T}} \) and VNR \( \mu_{tr} = \rho^{\frac{1}{1-T}} \). In this case we get in [3] that the transmitted signal \( x \in S_{D T}^{(i)}(\rho) \subset \mathbb{C}^{K M T} \).

In the receiver we first define the set \( H_{ex} \cdot \text{cube}_{D T}(a) \) as the multiplication of each point in \( \text{cube}_{D T}(a) \) with the matrix \( H_{ex} \). In a similar manner, the IC induced by the channel in the receiver is \( S_{D T}^{(i)} = H_{ex} \cdot S_{D T}^{(i)} \). The set \( H_{ex} \cdot \text{cube}_{D T}(a) \) is almost surely \( D \cdot T \)-complex dimensional (where \( D \leq L \)). In this case

\[
M(S_{D T}^{(i)}, a) = |S_{D T}^{(i)} \bigcap \text{cube}_{D T}(a)| = |S_{D T}^{(i)} \bigcap [H_{ex} \cdot \text{cube}_{D T}(a)]|.\]

We define the receiver density as

\[
\gamma_{rc} = \limsup_{a \to \infty} \frac{M(S_{D T}^{(i)}, a)}{\text{Vol}(H_{ex} \cdot \text{cube}_{D T}(a))}
\]

i.e., the upper limit on the ratio of the number of IC points in \( H_{ex} \cdot \text{cube}_{D T}(a) \), and the volume of \( H_{ex} \cdot \text{cube}_{D T}(a) \). Note that for \( N \geq K M \) and \( D = K M \) we get \( \gamma_{rc} \approx \rho^{\frac{1}{1-T}} \cdot \prod_{i=1}^{K M} \lambda_i^{-T} \) and \( \mu_{rc} = \rho^{1 - \frac{1}{1-T}} \cdot \prod_{i=1}^{K M} \lambda_i^{1/T} \). The joint decoder
average decoding error probability, over the points of the effective IC \( S_{D,T}(\rho) \), for a certain channel realization \( H \), is defined as

\[
\overline{P_e}(H, \rho) = \limsup_{a \to \infty} \frac{\sum_{x' \in S_{D,T}(H, e_{\text{cube}, D}(a))} P_e(x', H, \rho)}{M(S_{D,T}, a)}
\]

(6)

where \( P_e(x', H, \rho) \) is the error probability associated with \( x' \). The average decoding error probability of \( S_{D,T}(\rho) \) over all channel realizations is \( \overline{P_e}(\rho) = E_H (\overline{P_e}(H, \rho)) \). The diversity order is defined as

\[
d = -\lim_{\rho \to \infty} \rho \log(\overline{P_e}(\rho)).
\]

(7)

In practice finite constellations are transmitted even when performing regular lattice decoding in the receiver. Based on the results in [13] it was shown in [8] that finite constellation with multiplexing gain \( r \) can be carved from a lattice with multiplexing gain \( r \), while maintaining the same performance when regular lattice decoder is employed in the receiver. In our case it also applies for each of the users, i.e. carving finite constellations with multiplexing gains tuple \((r_1, \ldots, r_K)\) that satisfy the power constraint, from lattices with multiplexing gains tuple \((r_1, \ldots, r_K)\). In the receiver the performance is maintained by performing regular lattice decoding on the effective lattice.

C. Additional Notations

We further denote by \( d^*_{M,N} (r) \) the optimal DMT of finite constellations, and by \( d^*_{M,N} (r) \) the upper bound on the optimal DMT of any IC with average number of dimensions per channel use \( D \), both in a point to point channel with \( M \) transmit and \( N \) receive antennas. For the multiple access channels with \( K \) users, \( M \) transmit antennas for each user, and \( N \) receive antennas, we denote by \( d^{*,(FC)}_{K,M,N} (r) \) the optimal DMT of finite constellations in the symmetric case, and by \( d^{*,(IC)}_{K,M,N} (r) \), \( d_{K,M,N} (r_1, \ldots, r_K) \) the upper bounds on the optimal DMT of the unconstrained multiple-access channels for the symmetric case, and for multiplexing gains tuple \((r_1, \ldots, r_K)\) respectively.

We denote \( r_{\max} = \max(r_1, \ldots, r_K) \), i.e. the maximal multiplexing gain in the multiplexing gains tuple. In addition for any \( A \subseteq \{1, \ldots, K\} \) we define \( R_A = \sum_{a \in A} r_a \) and \( D_A = \sum_{a \in A} D_a \).

III. UPPER BOUND ON THE BEST DIVERSITY-MULTIPLEXING TRADEOFF

In this section we show that for \( N < (K + 1)M - 1 \) the DMT of the unconstrained multiple-access channels is suboptimal compared to the optimal DMT of finite constellations. On the other hand for \( N \geq (K + 1)M - 1 \), we derive an upper bound on the optimal DMT that coincides with the optimal DMT of finite constellations.

In subsection III-A we lower bound the error probability of any IC in the multiple-access channels, by using lower bounds from point-to-point channels. We use these lower bounds to formulate an upper bound on the optimal DMT of IC’s in multiple-access channels, in the form of an optimization problem. In subsection III-B we solve this optimization problem for the symmetric case. We compare the optimal DMT of IC’s to the optimal DMT of finite constellations, and find the cases where IC’s are suboptimal in subsection III-C. Finally in subsection III-D we give a convexity argument that shows for the symmetric case that whenever the optimal DMT is not a convex function IC’s are suboptimal.

A. Upper Bound on the Diversity-Multiplexing-Tradeoff

We lower bound the error probability of the unconstrained multiple-access channels in Lemma 1. Based on this lower bound we present in Theorem 2 an upper bound on the optimal DMT of IC’s.

Assume user \( i \) transmits \( D_iT \)-complex dimensional IC, with average number of dimensions per channel use \( D_i \) and \( T \) channel uses. The following lemma lower bounds the average decoding error probability of the \( K \)-users IC \( \overline{P_e}(D_1, \ldots, D_K; T) (\rho, r_1, \ldots, r_K) \), where \((D_1, \ldots, D_K)\) is the tuple of average number of dimensions per channel use, \( T \) is the number of channel uses and \((r_1, \ldots, r_K)\) is the tuple of multiplexing gains.

Lemma 1.

\[
\overline{P_e}(D_1, \ldots, D_K; T) (\rho, r_1, \ldots, r_K) \geq \max_{A \subseteq \{1, \ldots, K\}} \left( \overline{P_e}(D_A; T) (\rho, R_A) \right)
\]

where \( \overline{P_e}(D_A; T) (\rho, R_A) \) is the lower bound derived in [8] on the error probability of any IC with \( T \) channel uses, \( D_A = \sum_{a \in A} D_a \) average number of dimensions per channel use, and multiplexing gain \( R_A = \sum_{a \in A} r_a \), in a point-to-point channel with \(|A| \cdot M \) transmit and \( N \) receive antennas.

Proof: By considering the extended channel model [3], we get that the \( K \) distributed transmitters transmit an effective \( \left( \sum_{i=1}^K D_i \right) T \)-complex dimensional IC, over \( T \) channel uses, with multiplexing gain \( \sum_{i=1}^K r_i \). The error probability of this IC is lower bounded by the lower bound on the error probability of any IC with average number of dimensions per channel use \( \sum_{i=1}^K D_i \), \( T \) channel uses, and multiplexing gain \( \sum_{i=1}^K r_i \), in a point-to-point channel with \( KM \) transmit and \( N \) receive antennas. Such a lower bound on the error probability was derived in [3] for each channel realization ([8] Theorem 1), and then
for the average over all channel realizations when $\rho$ is large (\textcolor{red}{Theorem 2}). Now consider the set $A \subseteq \{1, \ldots, K\}$. In case a genie tells the receiver the transmitted messages of users $\{1, \ldots, K\} \setminus A$, the optimal receiver attains an error probability that lower bounds the $K$-user optimal receiver error probability. Without loss of optimality, the optimal receiver can subtract them from the received signal, and get a new $|A|$-users unconstrained multiple-access channels with average number of dimensions per channel use $\{D_0\}_{a \in A}$, $T$ channel uses, and multiplexing gain $\sum_{a \in A} r_a$. In a similar manner, the error probability of this $|A|$-users channel is lower bounded by the lower bound on the error probability of any IC with $\sum_{a \in A} D_a$ average number of dimensions per channel use, $T$ channel uses, and multiplexing gain $\sum_{a \in A} r_a$, derived in \textcolor{red}{[8]}. Hence, the maximal lower bound on the error probability between all $A \subseteq \{1, \ldots, K\}$ also sets a lower bound on the error probability. This concludes the proof. \hfill $\blacksquare$

Next we wish to formulate an upper bound on the DMT of IC’s in the $K$-user unconstrained multiple-access channels. We derive this bound based on the lower bound on the error probability presented in \textcolor{red}{Lemma 1} and on an upper bound on the DMT of IC’s for the point-to-point channel, presented in \textcolor{red}{[8]}. Let us begin by presenting the upper bound on the DMT for the point-to-point channel.

\textbf{Theorem 1 (\textcolor{red}{Theorem 2}).} For any sequence of IC’s $S_{D,T}(\rho)$ with $D$ average number of dimensions per channel use, in a point-to-point channel with $M$ transmit and $N$ receive antennas, the DMT $d_{M,N}(r)$ is upper bounded by

$$d_{M,N}(r) \leq d_{M,N}^{D,T}(r) = \frac{M \cdot N}{D} (D - r)$$

for $0 \leq D \leq \frac{M \cdot N}{M + N - 1}$, and

$$d_{M,N}^{D,T}(r) \leq d_{M,N}^{D,T}(r) = \frac{(M - l) (N - l)}{D - l} \cdot (D - r)$$

for $\frac{M \cdot N - (l-1)l}{N + M - 1 - 2(l-1)} \leq D \leq \frac{M \cdot N - (l+1)l}{N + M - 1 - 2l}$, and $l = 1, \ldots, \min(M, N) - 1$. In all cases $0 \leq r \leq D$.

Based on Lemma 1 and Theorem 1 we formulate the following upper bound on the optimal DMT of the multiple-access channels.

\textbf{Theorem 2.} The optimal DMT of any IC with multiplexing gains tuple $(r_1, \ldots, r_K)$ is upper bounded by

$$d_{K,M,N}^{*,(IC)}(r_1, \ldots, r_K) = \max_{(D_1, \ldots, D_K) \in D} \min_{A \subseteq \{1, \ldots, K\}} \left( d_{|A|,M,N}^{*,DA}(R_A) \right)$$

where $D = \{D_1, \ldots, D_K \mid 0 \leq D_i \leq M, \sum_{i=1}^{K} D_i \leq L\}$.

\textbf{Proof:} From \textcolor{red}{Lemma 1} we get a lower bound on the error probability of any sequence of effective IC’s $S_{\sum_{i=1}^{K} D_i}(\rho)$, transmitted by the $K$ users. The lower bound on the error probability can be translated to an upper bound on the diversity order. In addition the lower bound on the error probability depends on lower bounds on the error probabilities in point-to-point channels. Hence, we can use the upper bound on the DMT in point-to-point channels, presented in \textcolor{red}{Theorem 1} to get the following upper bound on the DMT of a tuple of average number of dimensions per channel use $(D_1, \ldots, D_K)$

$$\min_{A \subseteq \{1, \ldots, K\}} \left( d_{|A|,M,N}^{*,DA}(R_A) \right)$$

Maximizing over $(D_1, \ldots, D_K) \in D$ yields the upper bound on the optimal DMT. \hfill $\blacksquare$

\textbf{B. Characterizing the Optimal Symmetric DMT}

We wish to characterize an upper bound on the optimal DMT of IC’s in the symmetric case, i.e. $r_1 = \cdots = r_K = r$. Later we will use this upper bound in order to show the sub-optimality of the unconstrained multiple-access channels in the case $N < (K + 1)M - 1$, and also to show that the DMT upper bound coincides with the optimal DMT of finite constellations in the case $N \geq (K + 1)M - 1$.

\textcolor{red}{Lemmas 2, 3, 4, 5} present the relations between $d_{i,M,N}^{*,D}(i \cdot r), i = 1, \ldots, K$ for different values of $N$. We use these lemmas in order to upper bound the optimal DMT in the symmetric case in \textcolor{red}{Theorem 4}.

Based on \textcolor{red}{Theorem 2} we can state that the optimal DMT for the symmetric case for $K$ users is upper bounded by

$$d_{K,M,N}^{*,(IC)}(r) = \max_{(D_1, \ldots, D_K) \in D} \min_{A \subseteq \{1, \ldots, K\}} \left( d_{|A|,M,N}^{*,DA}(|A| \cdot r) \right)$$

i.e. we wish solve the aforementioned optimization problem for each $0 \leq r \leq \frac{L}{K}$. In order to solve this optimization problem we first solve a simpler optimization problem for the case $D_1 = \cdots = D_K = D$, i.e. each user transmits $D$ average number of dimensions per channel use. In this case the upper bound in \textcolor{red}{[8]} takes a simpler form

$$\max_{1 \leq i \leq K} \min_{D} \left( d_{i,M,N}^{*,D}(i \cdot r) \right)$$

\hfill $\textcolor{red}{(9)}$
where \( 0 \leq D \leq \frac{K}{M} \). After solving this optimization problem, we will show that choosing \( D_1 = \cdots = D_K = D \) also yields the optimal solution for \( \mathfrak{F} \).

In order to solve the optimization problem in (9), we first need to present some properties on the relations between \( d^*_{M,N} (i \cdot r) \), \( 1 \leq i \leq K \). We begin by presenting a property on the behavior of \( d^*_{M,N} (\cdot) \) as a function of \( D \).

**Corollary 1** (\cite{8} Corollary 1). For \( 0 \leq D \leq \frac{M-N}{N+M-1} \), we get

\[
d^*_{M,N} (0) = MN.
\]

For \( \frac{M-N-(l-1)l}{N+M-1-2l(l-1)} \leq D \leq \frac{M-N-(l+1)(l)}{N+M-1-2l} \), and \( l = 1, \ldots, \min\{M,N\} - 1 \) we get

\[
d^*_{M,N} (l) = (M-l) \cdot (N-l).
\]

A simple interpretation of Corollary 1 is that for the case \( 0 \leq D \leq \frac{M-N}{N+M-1} \) the straight lines \( d^*_{M,N} (\cdot) \) that represent the upper bounds on the DMT, all have the same “anchor” point at multiplexing gain \( r = 0 \), i.e. they all have diversity order \( MN \) for \( r = 0 \), and each line equals to zero for \( r = D \). On the other hand, for \( \frac{M-N-(l-1)l}{N+M-1-2l} \leq D \leq \frac{M-N-(l+1)(l)}{N+M-1-2l} \), and \( l = 1, \ldots, \min\{M,N\} - 1 \), the straight lines equal to \( (M-l) \cdot (N-l) \) for multiplexing gain \( r = l \), and again each line equals to zero for \( r = D \). Figure 1 illustrates this property for the case \( M = N = 2 \). Another property relates to the optimal DMT of finite constellations in the symmetric case.

**Theorem 3** (\cite{9} Theorem 3). The optimal DMT of finite constellations in the symmetric case equals

\[
d^*_{K,M,N} (r) = \begin{cases} 
  d^*_{M,N} (\cdot) & 0 \leq r \leq \min \left( \frac{N}{K+1}, M \right) \\
  d^*_{K,M,N} (K \cdot r) & \min \left( \frac{N}{K+1}, M \right) \leq r \leq \min \left( \frac{N}{K}, M \right)
\end{cases}
\]

In order to solve the optimization problem in (9) we present several lemmas related to the inequalities between \( d^*_{i,M,N} (i \cdot r) \) for \( 1 \leq i \leq K \). The proofs of these lemmas rely mainly on Corollary 1 and Theorem 3.

**Lemma 2.** For \( N \geq (K+1)M - 1 \) we get

\[
d^*_{M,N} (r) \leq d^*_{i,M,N} (i \cdot r) \quad 2 \leq i \leq K
\]

for any \( 0 \leq r \leq D \) and \( 0 \leq D \leq M \).

*Proof:* The proof is in appendix A.

An example of Lemma 2 for the case \( M = K = 2 \) and \( N = 4 \) is illustrated in Figure 2.

**Lemma 3.** For \( N < (K+1)M - 1 \) we get

\[
d^*_{M,N} (r) \leq d^*_{i,M,N} (i \cdot r) \quad 2 \leq i \leq K - 1
\]

for any \( 0 \leq D \leq \frac{K}{M} \) and \( 0 \leq r \leq D \).

*Proof:* The proof is in appendix B.
For Lemma 4, prove two more properties that will enable us to find the optimal DMT of IC’s in the symmetric case.

It can be seen that for this setting $d_{2M,N}^{*,2D}(2r) > d_{M,N}^{*,D}(r)$.

From Lemmas 2, 3 we can see that the optimization problem in (9) involves only $d_{M,N}^{*,D}(r)$ and $d_{K,M,N}^{*,K,D}(K \cdot r)$. Now we prove two more properties that will enable us to find the optimal DMT of IC’s in the symmetric case.

**Lemma 4.** For $N < (K - 1) M + 1$ we get

$$\max_{0 \leq D \leq K} \min_{1 \leq l \leq K} d_{l,M,N}^{*,D}(i \cdot r) = d_{M,N}^{*,K}(r) = M \cdot N - M \cdot K \cdot r \quad 0 \leq r \leq \frac{N}{K}$$

**Proof:** The proof is in appendix C.

From Lemma 4 we can see that for the multiple-access channels, when $N < (M - 1) K + 1$ the optimal DMT of IC’s is smaller than finite constellations optimal DMT for any value of $r$ except for $r = 0$ and $r = \frac{N}{K}$. Figure 3 illustrates Lemma 4 for the case $M = N = K = 2$. Next we wish to show the cases where $d_{M,N}^{*,D}(r)$ and $d_{K,M,N}^{*,K}(K \cdot r)$ coincide.

From Lemmas 2, 3 we can see that the optimization problem in (9) involves only $d_{M,N}^{*,D}(r)$ and $d_{K,M,N}^{*,K,D}(K \cdot r)$. Now we prove two more properties that will enable us to find the optimal DMT of IC’s in the symmetric case.

**Lemma 4.** For $N < (K - 1) M + 1$ we get

$$\max_{0 \leq D \leq K} \min_{1 \leq l \leq K} d_{l,M,N}^{*,D}(i \cdot r) = d_{M,N}^{*,K}(r) = M \cdot N - M \cdot K \cdot r \quad 0 \leq r \leq \frac{N}{K}$$

**Proof:** The proof is in appendix C.

From Lemma 4 we can see that for the multiple-access channels, when $N < (M - 1) K + 1$ the optimal DMT of IC’s is smaller than finite constellations optimal DMT for any value of $r$ except for $r = 0$ and $r = \frac{N}{K}$. Figure 3 illustrates Lemma 4 for the case $M = N = K = 2$. Next we wish to show the cases where $d_{M,N}^{*,D}(r)$ and $d_{K,M,N}^{*,K}(K \cdot r)$ coincide.

The following lemma sets another building block in upper bounding the optimal DMT in the symmetric case when $N = (K - 1) M + 1 + l$, $l = 0, \ldots, 2M - 3$. It finds the average number of dimensions per channel use that leads to the equality $d_{M,N}^{*,D}(r) = d_{K,M,N}^{*,K,D}(r)$ for any value of $r$, and also shows for which values of $r$ these straight lines are equal to the optimal DMT of finite constellations in a point-to-point channel.

**Lemma 5.** For $N = (K - 1) M + 1 + l < (K + 1) M - 1$, where $l = 0, \ldots, 2M - 3$, we get for average number of dimensions per channel use per user $D_l = \frac{MN - \frac{l}{2}((\frac{l}{2}) + 1) - 2((\frac{1}{2}) + 1)\left(\frac{1}{2} - \frac{l}{2}\right)}{N + M - 1 - l}$ that

$$d_{M,N}^{*,D_l}(r) = d_{K,M,N}^{*,K,D_l}(K \cdot r) = d^{*}(r) = MN - \frac{l}{2} \cdot \left(\frac{l}{2} + 1\right) - 2 \cdot \left(\frac{1}{2} + 1\right) \cdot \left(\frac{1}{2} - \frac{l}{2}\right) - (N + M - 1 - l) r$$

where $0 \leq r \leq D_l$. In addition

$$d_{M,N}^{*(FC)}\left(\frac{l}{2} + 1\right) = d^{*}\left(\frac{1}{2} + 1\right)$$
and also

\[ d^{*(FC)}_{K,M,N}(K-1)M + \left\lfloor \frac{l+1}{2} \right\rfloor = d^*(\frac{(K-1)M + \left\lfloor \frac{l+1}{2} \right\rfloor}{K}) \]

Proof: The proof is in appendix \[D\]

An example that illustrates Lemma \[5\] for the case \( M = K = 2 \) and \( N = 4 \) is given in Figure \[4\].

![Figure 4](image)

**Fig. 4.** \( d^* (r) \) for the case \( M = K = 2 \) and \( N = 4 \), i.e. \( l = 1 \). Note that \( d^* (1) = d^{*8/5}_{2,4}(1) = d^{*8/5,F C}_{2,4}(1) \) and \( d^* (\frac{3}{2}) = d^*_{4,4}(3) = d^{*4,F C}_{4,4}(3) \).

Now we are ready to characterize the upper bound on the optimal DMT of IC’s in the symmetric case. Recall that for \( N = (K-1)M + 1 + l < (K+1)M - 1 \), \( l = 0, \ldots, 2M-3 \)

\[ d^* (r) = MN - \left\lfloor \frac{l}{2} \right\rfloor \cdot \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right) - 2 \cdot \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right) \cdot \left( \frac{l}{2} - \left\lfloor \frac{l}{2} \right\rfloor \right) - (N + M - 1 - l)r. \]

**Theorem 4.** The optimal DMT of IC’s in the symmetric case is upper bounded by:

For \( N \geq (K+1)M - 1 \)

\[ d^{*(IC)}_{K,M,N} (r) = d^{*(FC)}_{M,N} (r). \]

For \( N < (K-1)M + 1 \)

\[ d^{*(IC)}_{K,M,N} (r) = M \cdot N - K \cdot M \cdot r. \]

For \( N = (K-1)M + 1 + l < (K+1)M - 1 \), where \( l = 0, \ldots, 2M-3 \)

\[ d^{*(IC)}_{K,M,N} (r) = \begin{cases} 
 d^{*(FC)}_{M,N} (r) & 0 \leq r \leq \left\lfloor \frac{l}{2} \right\rfloor + 1 \\
 d^* (r) & \left\lfloor \frac{l}{2} \right\rfloor + 1 \leq r \leq \frac{(K-1)M + \left\lfloor \frac{l+1}{2} \right\rfloor}{K} \\
 d^{*(FC)}_{K,M,N} (Kr) & \frac{(K-1)M + \left\lfloor \frac{l+1}{2} \right\rfloor}{K} \leq r \leq \frac{L}{K}.
\end{cases} \]

Proof: The proof is in appendix \[E\]

Figure \[4\] also presents \( d^{*(IC)}_{K,M,N} (r) \) for the case \( M = K = 2 \) and \( N = 4 \) (which leads to \( l = 1 \)).

**C. Comparison to Finite Constellations**

In this subsection we compare the optimal DMT of finite constellations to the upper bound on the optimal DMT of IC’s (in general, not only for the symmetric case). This comparison enables us to show that for \( N \geq (K+1)M - 1 \) the upper bound on the optimal DMT of IC’s coincides with the optimal DMT of finite constellations. On the other hand for \( N < (K+1)M - 1 \) we show that the upper bound on the optimal DMT of IC’s is inferior compared to the optimal DMT of finite constellations. This leads to the conclusion that in the range \( N < (K+1)M - 1 \), the best DMT IC’s can attain is suboptimal compared to the optimal DMT of finite constellations.

In Lemma \[6\] we compare the upper bound on the optimal DMT of IC’s in the symmetric case, to the optimal DMT of finite constellations. Then we use this result to prove in Theorem \[5\] that the optimal DMT of IC’s is suboptimal when \( N < (K+1)M - 1 \).

We begin by showing when the upper bound on the optimal DMT of IC’s in the symmetric case, \( d^{*(IC)}_{K,M,N} (r) \), is suboptimal compared to the optimal DMT of finite constellations.
Lemma 6. For either $N \geq (K + 1) M - 1$ or $K = 2$, $M = s + 1$, $N = 3 \cdot s$, where $s \geq 1$ and $s \in \mathbb{Z}$ we get
\[ d_{K,M,N}^{\text{IC}} (r) = d_{K,M,N}^{\text{DFC}} (r). \]

For $N < (K - 1) M + 1$
\[ d_{K,M,N}^{\text{IC}} (r) < d_{K,M,N}^{\text{DFC}} (r) \quad 0 < r < \frac{N}{K}. \]

For $N = (K - 1) M + 1 + l < (K + 1) M - 1$ where $l = 0, \ldots, 2M - 3$
\[ d_{K,M,N}^{\text{IC}} (r) < d_{K,M,N}^{\text{DFC}} (r) \quad \left\lfloor \frac{l}{2} \right\rfloor + 1 < r < \frac{(K - 1) M + \left\lfloor \frac{l+1}{2} \right\rfloor}{K}. \]

Proof: The full proof is in appendix E. In a nutshell the proof is based on the properties of $d_{M,N}^{A,D} (r)$ derived in Corollary 1 and also on the results in Theorem 4. It is important to note that for $K = 2$, $M = s + 1$ and $N = 3s$ we get that $d_{K,M,N}^{\text{IC}} (r) = d_{K,M,N}^{\text{DFC}} (r)$ because in this case $\left\lfloor \frac{l}{2} \right\rfloor + 1 = \frac{(K - 1) M + \left\lfloor \frac{l+1}{2} \right\rfloor}{K}$. 

The sub-optimality of $d_{K,M,N}^{\text{IC}} (r)$ for the case $N < (K - 1) M + 1$ is illustrated in Figure 3 where the sub-optimality for the case $N = (K - 1) M + 1 + l$, where $l = 0, \ldots, 2M - 3$ is illustrated in Figure 4.

Now we are ready to present the cases where the upper bound on the optimal DMT of the unconstrained multiple-access channels coincides with the optimal DMT of finite constellations, and the cases where the optimal DMT of the unconstrained multiple-access channels is suboptimal compared to the optimal DMT of finite constellations.

Theorem 5. For $N \geq (K + 1) M - 1$ the optimal DMT of the unconstrained multiple-access channels is upper bounded by $d_{M,N}^{\text{DFC}} (\max (r_1, \ldots, r_K))$ the optimal DMT of finite constellations. For $N < (K + 1) M - 1$ the best DMT that can be attained by the unconstrained multiple-access channels is inferior compared to the optimal DMT of finite constellations.

Proof: The full proof is in appendix G. The proof outline is as follows. Recall that in Theorem 2 we have shown that the optimal DMT of IC’s is upper bounded by
\[ d_{K,M,N}^{\text{IC}} (r_1, \ldots, r_K) = \max_{(D_1, \ldots, D_K) \in \mathcal{D}} \min_{A \subseteq \{1, \ldots, K\}} \left( d_{|A|,M,N}^{A,D} (R_A) \right). \]

For $N \geq (K + 1) M - 1$ we show that this term is upper and lower bounded by $d_{M,N}^{\text{DFC}} (\max (r_1, \ldots, r_K))$, which is the optimal DMT of finite constellations in this case.

For the case $N < (K + 1) M - 1$ we show that the optimal DMT is not attained by finding a set of multiplexing gain tuples $(r_1, \ldots, r_K) \in \mathcal{B}$ for which $d_{K,M,N}^{\text{IC}} (r_1, \ldots, r_K) < d_{M,N}^{\text{DFC}} (r_1, \ldots, r_K)$. Based on Lemma 6 we get for $r_1 = \cdots = r_K = r$ that there exist a set of multiplexing gains for which $d_{K,M,N}^{\text{IC}} (r) < d_{K,M,N}^{\text{DFC}} (r)$, except for the case $K = 2$, $M = s + 1$ and $N = 3s$, where $s \geq 1$ is an integer. For this case showing that $d_{2,s+1,3s}^{\text{IC}} (r_1, r_2) < d_{2,s+1,3s}^{\text{DFC}} (r_1, r_2)$ is more involved and requires considering the case $r_1 \neq r_2$ (see appendix C for the full proof). An illustrative example of the method of proof for this case is presented in Figures 5 and 6.

![Figure 5](image-url)

Fig. 5. The upper bound on the optimal DMT of IC’s in the symmetric case for $K = 2$, $M = 3$, $N = 6$. Note that for this case we get $\left\lfloor \frac{l}{2} \right\rfloor + 1 = \frac{N}{K+1} = \frac{(K-1)M+1+\left\lfloor \frac{l+1}{2} \right\rfloor}{K}$. In addition this upper bound coincides with the optimal DMT of finite constellations in the symmetric case. Finally, for this case we get $d_{3,6}^{\text{IC}} (r) = d_{6,6}^{\text{DFC}} (2r)$. 


shown rather easily by the following arguments. It is based on the fact that a function that equals to the maximum between where the optimal DMT of finite constellations is not necessarily so. The convexity of the optimal DMT of IC’s can be

\[ \text{D. Discussion: Convexity Vs. Non-Convexity of the Optimal DMT} \]

It is interesting to note that the upper bound on the optimal DMT of IC’s in the symmetric case is a convex function, where the optimal DMT of finite constellations is not necessarily so. The convexity of the optimal DMT of IC’s can be shown rather easily by the following arguments. It is based on the fact that a function that equals to the maximum between straight lines is a convex function. For \( N \geq (K + 1) M - 1 \) the optimal DMT of IC’s in the symmetric case is simply upper bounded by \( d_{M,N}^{(FC)} (r) \) which is a maximization between straight lines, and therefore is a convex function. For the case \( N < (K - 1) M + 1 \) the upper bound on the optimal DMT of IC’s in the symmetric case is a straight line. Finally, for \( N = (K - 1) M + 1 + l < (K + 1) M - 1 \), where \( l = 0, \ldots, 2M - 3 \), the upper bound on the optimal symmetric DMT of IC’s equals to the maximization between the first \( \lfloor \frac{l}{2} \rfloor + 1 \) straight lines constituting \( d_{K,M,N}^{(FC)} (r) \) and \( d^* (r) \), and the last \( M - \lfloor \frac{l+1}{2} \rfloor \) straight lines constituting \( d_{K,M,N}^{(IC)} (K \cdot r) \). This maximization also yields a convex function.

On the other hand the optimal DMT of finite constellations in the symmetric case is not necessarily a convex function. See Figure 4 for illustration. In fact the optimal DMT is not a convex function whenever \( N < (K - 1) M + 1 \), or when \( N = (K - 1) M + 1 + l < (K + 1) M - 1 \) and \( \frac{l}{2} + 1 \neq \frac{(K - 1) M + l}{K} \), where \( l = 0, \ldots, 2M - 3 \). It results from the following arguments. For \( N < (K - 1) M + 1 \) we get \( \frac{M}{N + M - 1} > \frac{N}{K} \), and so \( d_{M,N}^{(FC)} (r) = d_{M,N}^{(FC)} (1) > 0 \). In addition \( d_{K,M,N}^{(FC)} (r) = d_{K,M,N}^{(FC)} (\frac{1}{K}, r) \) for \( 0 \leq r < \min \left( 1, \frac{N}{K+1} \right) \). Based on these facts and on the facts that \( d_{K,M,N}^{(FC)} (r) \) is a piecewise linear function and \( d_{K,M,N}^{(FC)} (\frac{N}{K}) = 0 \), we get that \( d_{K,M,N}^{(IC)} (r) \) is not a convex function. For the case \( N = (K - 1) M + 1 + l < (K + 1) M - 1 \), where \( l = 0, \ldots, 2M - 3 \), we know that

\[ d_{K,M,N}^{(IC)} (r) = d^* (r) < d_{K,M,N}^{(FC)} (r) \]

Since \( d^* (r) \) is a straight line it necessarily means that \( d_{K,M,N}^{(FC)} (r) \) is not a convex function whenever \( \frac{l}{2} + 1 \neq \frac{(K - 1) M + l}{K} \). For the case \( \frac{l}{2} + 1 = \frac{(K - 1) M + l}{K} \) we get \( d_{K,M,N}^{(FC)} (r) = d_{K,M,N}^{(IC)} (r) \), and so in this case the optimal DMT of finite constellations in the symmetric case is also a convex function. Finally, for the case \( N \geq (K + 1) M - 1 \) the optimal DMT in the symmetric case equals \( d_{M,N}^{(FC)} \) and as aforementioned it is a convex function. Therefore, we can state that whenever the
optimal DMT of finite constellations in the symmetric case is not a convex function, the optimal DMT of IC’s is suboptimal.

Finally, a question that may arise is whether it is possible to find an extension of orthogonal designs [14] to the multiple-access channels, i.e. a transmission scheme that enables to separate the space-time code from the symbols required for transmission. The most notable example of such a transmission scheme is the Alamouti scheme [15] for the case of two transmit antennas and a single receive antenna. For example, in this case transmitting the information itself over the space-time code enables to separate the space-time code from the symbols required for transmission.

The most notable example of such a transmission scheme is the Alamouti scheme [15] for the case of two transmit antennas and a single receive antenna. For example, in this case transmitting the information itself over the space-time code enables to separate the space-time code from the symbols required for transmission. This is due to the fact that when the constellation size is infinite, the performance is sub-optimal. Hence, in this sense there is no extension of orthogonal designs to the multiple-access channels.

In this section we show that the upper bound on the DMT of the unconstrained multiple-access channels, derived in section III is achievable for the case $N \geq (K + 1) M - 1$ by a sequence of IC’s in general and lattices in particular. Essentially, we show for $N \geq (K + 1) M - 1$ that IC’s attain DMT that equals to $d_{K,N}^\star$($r_1, \ldots, r_K$) = $d_{M,N}^\star$($\max(r_1, \ldots, r_K$)).

We begin by showing in subsection IV-A that simple orthogonal transmission approaches such as time-division multiple-access (TDMA) or code-division multiple-access (CDMA) will result in sub-optimal performance for $N \geq (K + 1) M - 1$. Then, we introduce in subsection IV-B the transmission scheme for each user, followed by presentation of the effective channel induced by the transmission scheme in subsection IV-C. We derive in subsection IV-D for each channel realization an upper bound on the error probability of the ML decoder of an ensemble of K IC’s. Finally, in subsection IV-E we average this upper bound over the channel realizations, and show that the optimal DMT is attained for $N \geq (K + 1) M - 1$.

### A. Orthogonal Transmission is Sub-optimal

In this subsection we show the sub-optimality of transmission methods that create in the receiver orthogonalization between different independent streams, for any channel realization. The advantage of these transmission schemes is their simplicity. By assigning the IC’s or lattices correctly in the space, they enable to consider each stream independently and reduce the decoding problem to the point-to-point scenario. Such an approach is very natural when considering IC’s in general and lattices in particular, as it involves assigning the streams with dimensions or subspaces that remain orthogonal in the receiver for each channel realization. The IC related to a certain stream lies within the assigned subspace. We show for $N \geq (K + 1) M - 1$ that such transmission method is sub-optimal as it requires each user to give up too many dimensions to create the orthogonalization.
In the receiver, orthogonal transmission scheme enables each independent stream to lie within a subspace orthogonal to the other streams, for each channel realization. In order for a transmission scheme to fulfill this property, the streams must be assigned with orthogonal subspaces already in the transmitter, i.e., must be assigned with orthogonal subspaces in $\mathbb{C}^{MT}$ assuming there are $T$ channel uses. Hence, orthogonal transmission schemes require the partition of at most $M$ number of dimensions per channel use between all users. On the other hand, $N \geq (K+1)M-1$ leads to $N \geq K \cdot M$, and so potentially the $K$ users could transmit together up to $KM$ dimensions per channel use, but not orthogonally. The optimal DMT for the symmetric case for $N \geq (K+1)M-1$ is $d^*_{M,N}(FC) (r)$ from Corollary 3 and Theorem 4. We know that in the range $M-1 \leq r \leq M$ the optimal DMT is obtained only when each user transmits $M$ average number of dimensions per channel use, i.e., the $K$ users must transmit together $KM$ dimensions per channel use. Hence, orthogonal transmission is not provided with enough dimensions per channel use to obtain the last line of the optimal DMT. This leads to its sub-optimality.

As a first example we consider an orthogonal transmission scheme that takes the natural partition to $K$ streams induced by the multiple-access channels. In order to obtain orthogonalization for this case, a different user is transmitting in each channel use, while the others wait for their turn to transmit. This transmission method is referred to as TDMA. Let us consider the symmetric case where each user transmits with multiplexing gain $r$. Assuming there are $T$ channel uses, in the symmetric case each user transmits only on $\frac{T}{K}$ channel uses. In this case each user can obtain the point-to-point performance of a channel with $M$ transmit and $N$ receive antennas, using $\frac{T}{K}$ channel uses. However, in order for each user to obtain multiplexing gain $r$ per channel use, it must transmit at multiplexing gain $Kr$ in the $\frac{T}{K}$ channel uses, which leads to DMT performance of $d^*_{M,N}(FC) (Kr)$. This shows the sub-optimality of TDMA.

Another transmission approach is assigning an independent stream for each transmit antenna. This is equivalent to considering a multiple-access channel with $KM$ users, each with a single transmit antenna. Let us consider for example a multiple-access channel with $M = 1$, $K$ users and $N \geq K$. In this case the optimal DMT for the symmetric case equals $d^*_{1,N}(FC) (r)$. On the other hand for CDMA each user is assigned with an orthogonal subspace in $\mathbb{C}^T$, assuming there are $T$ channel uses. In this way each stream can obtain the performance of a point-to-point channel with a single transmit antenna and $N$ receive antennas. However, for the orthogonalization to hold each user is assigned with $\frac{T}{K}$ dimensional subspace, which must be orthogonal to the other users subspaces. Hence, in order for each user to obtain multiplexing gain $r$ per channel use, he must transmit at multiplexing gain $Kr$ over the $\frac{T}{K}$ dimensional subspace. This leads to suboptimal DMT performance of $d^*_{1,N}(FC) (Kr)$.

**B. The Transmission Scheme**

From subsection [V-A] we get that an optimal transmission scheme must allow different users to lie in overlapping subspaces in the receiver, i.e. in the receiver the users can not reside in orthogonal subspaces. Essentially, in the proposed transmission scheme each user transmits as if the channel was a point-to-point channel with $M$ transmit and $N$ receive antennas. Hence, each user transmission matrix is identical to the transmission matrix presented in [8].

We denote the transmission matrix of user $i$ by $G^{(i)}_{l}$, where $l = 0, \ldots, M-1$ and $i = 1, \ldots, K$. $G^{(i)}_{l}$ has $M$ rows that represent the transmission antennas, and $T_l = N + M - 1 - 2 \cdot l$ columns that represent the number of channel uses. $G^{(i)}_{l}$ transmits $D_l = \frac{NM-(l+1)}{N+M-1-2l}$ average number of dimensions per channel use in the following manner.

Consider a channel with $M$ transmit and $N$ receive antennas.

1) For $D_{M-1} = \frac{M(N-M+1)}{N+M-1} = M$: the matrix $G^{(i)}_{M-1}$ has $N - M + 1$ columns (channel uses). In the first column transmit symbols $x_{1}, \ldots, x_{M}$ on the $M$ antennas, and in the $N-M+1$ column transmit symbols $x_{M(N-M)+1}, \ldots, x_{M(N-M+1)}$ on the $M$ antennas.

2) For $D_{l}, l = 0, \ldots, L-2$: the matrix $G^{(i)}_{l}$ has $M + N - 1 - 2 \cdot l$ columns. We add to $G^{(i)}_{l+1}$ the transmission scheme for $D_{l+1}$, two columns in order to get $G^{(i)}_{l+1}$. In the first added column transmit $l+1$ symbols on antennas $1, \ldots, l+1$. In the second added column transmit different $l+1$ symbols on antennas $M-l, \ldots, M$.

According to the definition of the transmission scheme we can see that the different users transmit the same average number of dimensions per channel use. Let us denote the transmission scheme of the first $k$ users by

$$G^{(1,\ldots,k)}_{l} = \left( G^{(1)}_{l}, \ldots, G^{(k)}_{l} \right)$$

(10)

$G^{(1,\ldots,k)}_{l}$ is a $k \cdot M \times T_{l}$ matrix. Note that $G^{(1,\ldots,k)}_{l}$ transmits $k \cdot D_{l} T_{l}$ average number of dimensions per channel use. Later in this section we show that $G^{(1,\ldots,K)}_{l}$ attains the optimal DMT in the range $l \leq r_{max} \leq l+1$.

**Example:** $M = 2$, $N = 5$ and $K = 2$. In this case the transmission scheme for $D_{0} = \frac{10}{6}$, $D_{1} = \frac{8}{4}$ ($G^{(1,2)}_{0}$, $G^{(1,2)}_{1}$ respectively)
is as follows:

\[
G_{t}^{1(1,2)} = \begin{pmatrix} G_{t}^{1(1)} \\ G_{t}^{2(2)} \end{pmatrix} = \begin{pmatrix}
G_{t}^{(1)} \\
G_{t}^{(2)}
\end{pmatrix} = \begin{pmatrix}
x_1 & x_3 & x_5 & x_7 & 0 & x_{17} \\
x_2 & x_4 & x_6 & x_8 & 0 & x_{18} \\
- & - & - & - & - & - \\
x_9 & x_{11} & x_{13} & x_{15} & x_{19} & 0 \\
x_{10} & x_{12} & x_{14} & x_{16} & 0 & x_{20}
\end{pmatrix}.
\]

(11)

\[
D_{1} = \frac{4}{v_{0}} G_{1}^{(1,2)} \\
D_{0} = \frac{4}{v_{0}} G_{0}^{(1,2)}
\]

C. The Effective Channel

Next we define the effective channel matrix induced by the transmission scheme of the first \(k\) users \(G_{t}^{(1,...,k)}\), where \(k = 1, \ldots, K\). Let us denote the first \(k\) users transmission at time instance \(t\) by

\[
\tilde{z}_{t} = (\tilde{z}_{t}^{(1)}, \ldots, \tilde{z}_{t}^{(k)})^T \quad t = 1, \ldots, T_{t}.
\]

In accordance with the channel model from (1) we get

\[
y_{t} = H_{t}^{(1,...,k)} \cdot \tilde{z}_{t} \quad t = 1, \ldots, T_{t},
\]

where \(H_{t}^{(1,...,k)} = (H_{t}^{(1)}, \ldots, H_{t}^{(k)})\), is an \(N \times k \cdot M\) matrix. The multiplication \(H_{t}^{(1,...,k)} \cdot G_{t}^{(1,...,k)}\) yields a matrix with \(N\) rows and \(T_{t}\) columns, where each column equals to \(H_{t}^{(1,...,k)} \cdot \tilde{z}_{t}\), \(t = 1, \ldots, T_{t}\). Each user is transmitting \(D_{t} T_{t}\)-complex dimensional IC with \(D_{t} T_{t}\)-complex symbols, i.e. \(G_{t}^{(1)}\) has exactly \(D_{t} T_{t}\) non-zero values representing the \(D_{t} T_{t}\) complex-dimensional IC within \(\mathbb{C}^{M T}\). Together, the first \(k\) users transmit an effective \(k \cdot D_{t} T_{t}\)-dimensional complex IC within \(\mathbb{C}^{k \cdot M T}\). For each column of \(G_{t}^{(1,...,k)}\), denoted by \(g_{m}^{(k)}\), \(m = 1, \ldots, T_{t}\), we define the effective channel that \(g_{m}^{(k)}\) sees as \(\tilde{H}_{m}\). It consists of the columns of \(H_{t}^{(1,...,k)}\) that correspond to the non-zero entries of \(g_{m}^{(k)}\), i.e. \(H_{t}^{(1,...,k)} \cdot \tilde{z}_{m}^{(k)} = \tilde{H}_{m} \cdot g_{m}^{(k)}\), where \(\tilde{H}_{m}\) equals to the non-zero entries of \(g_{m}^{(k)}\). As an example assume without loss of generality that only the first \(l_{m}\) entries of \(g_{m}^{(k)}\) are not zero. In this case \(\tilde{H}_{m}\) is an \(N \times l_{m}\) matrix that equals to the first \(l_{m}\) columns of \(H_{t}^{(1,...,k)}\). In accordance with (3), \(H_{t}^{(1,k)}\) is an \(N T_{t} \times k D_{t} T_{t}\) block diagonal matrix consisting of \(T_{t}\) blocks. Since each block in \(H_{t}^{(1,k)}\) corresponds to the multiplication of \(H_{t}^{(1,...,k)}\) with different column in \(G_{t}^{(1,...,k)}\), the blocks of \(H_{t}^{(1,k)}\) equal \(\tilde{H}_{m}\), \(m = 1, \ldots, T_{t}\). Note that in the effective matrix \(N T_{t} \geq k \cdot D_{t} T_{t}\).

Next we elaborate on the structure of the blocks of \(H_{t}^{(1,k)}\). For this reason we denote the \(m^\text{th}\) column of \(H_{t}^{(1,...,k)}\) by \(\tilde{h}_{m}^{(k)}\), \(m = 1, \ldots, k \cdot M\). The transmission scheme has \(N + M - 1 - 2 \cdot l\) columns. The entries of the first \(N + M + 1\) columns of \(G_{t}^{(1,...,k)}\), \(\tilde{g}_{1}^{(k)}, \ldots, \tilde{g}_{N+M+1}^{(k)}\) are all different from zero. Hence, the first \(N + M + 1\) blocks of \(H_{t}^{(1,k)}\) are

\[
\tilde{H}_{m} = H_{t}^{(1,...,k)} \quad m = 1, \ldots, N + M + 1.
\]

(12)

After the first \(N + M + 1\) columns we have \(M - 1 - l\) pairs of columns. For each pair we have

\[
\tilde{H}_{N+M+2v} = \tilde{H}_{N+M+2(v-1)} \setminus \{ \tilde{h}_{M-(v-1)}^{(k)}, \tilde{h}_{2M-(v-1)}^{(k)}, \ldots, \tilde{h}_{kM-(v-1)}^{(k)} \} = \{ \tilde{h}_{1}, \ldots, \tilde{h}_{M}, \tilde{h}_{M+1}, \ldots, \tilde{h}_{2M}, \ldots, \tilde{h}_{(k-1)M+1}, \ldots, \tilde{h}_{kM-v} \}
\]

and

\[
\tilde{H}_{N+M+2v+1} = \tilde{H}_{N+M+2(v-1)+1} \setminus \{ \tilde{h}_{v+1}, \tilde{h}_{v+2}, \ldots, \tilde{h}_{v+kM} \} = \{ \tilde{h}_{v+1}, \ldots, \tilde{h}_{M}, \tilde{h}_{M+v+1}, \ldots, \tilde{h}_{2M}, \ldots, \tilde{h}_{(k-1)M+v+1}, \ldots, \tilde{h}_{kM} \}
\]

(13)

(14)

where \(v = 1, \ldots, M - 1 - l\).

Example: consider \(M = 2\), \(N = 5\) and \(K = 2\) as presented in (11). In this case \(l = 0, 1\) and we have \(D_{0} = \frac{10}{v_{0}}\) and \(D_{1} = \frac{8}{v_{0}}\) respectively. In addition \(H_{t}^{(1,2)} = H_{t}^{(1)} \odot H_{t}^{(2)} = (\tilde{h}_{1}, \tilde{h}_{2}, \tilde{h}_{3}, \tilde{h}_{4})\). We begin with \(k = 1\). In this case we get a point-to-point channel with 2 transmit and 5 receive antennas \(H^{(1)} = (\tilde{h}_{1}, \tilde{h}_{2}, \tilde{h}_{3}, \tilde{h}_{4})\), which leads to the following effective channels

1) \(D_{1} = 2: H_{t}^{(1,1,k=1)}\) is generated from the multiplication of the 5 \(\times 2\) matrix \(H^{(1)}\) with the four columns of the transmission matrix \(G_{1}^{(1)}\). In this case \(H_{t}^{(1,1)}\) is a \(20 \times 8\) block diagonal matrix, consisting of four blocks, each block equals to \(H^{(1)}\).

2) \(D_{0} = \frac{10}{v_{0}}: H_{t}^{(0,1,k=1)}\) is a \(30 \times 10\) block diagonal matrix consisting of six blocks. The first four blocks are equal to \(H^{(1)}\). The additional two blocks (induced by columns 5-6 of \(G_{1}^{(1)}\)) are vectors. We get that \(\tilde{H}_{5} = \tilde{h}_{1}\) and \(\tilde{H}_{6} = \tilde{h}_{2}\).

For the case \(k=2\) the effective channel induced by \(G_{t}^{(1,2)}\) is as follows.

1) \(D_{1} = 2: \) In this case the effective channel \(H_{t}^{(1,2)}\) is a \(20 \times 16\) matrix consisting of four blocks, where each block equals \(H^{(1,2)} = (H^{(1)}, H^{(2)})\).
For any \( k \) we show that the upper bound on the error probability for the point-to-point channel derived in [8] can be used to upper bound the probability for each of these events. The full proof is in appendix I.

We present \( H_{eff}^{(0), k} \) of our example in equation (15).

Now we consider the rows of \( G^{(1, \ldots, k)}_l \). Each row of the transmission matrix is related to the column of \( H^{(1, \ldots, k)}_l \) that multiplies it, i.e. row \( j \) in \( G^{(1, \ldots, k)}_l \) corresponds to column \( h_j \). In case there is a non-zero entry of row \( j \) in column \( m \) of \( G^{(1, \ldots, k)}_l \), it means that \( h_j \) occurs in \( H_m \). In the next lemma we examine the number of occurrences of a certain column of \( H^{(1, \ldots, k)}_l \) in the blocks of \( H_{eff}^{(l), k} \).

**Lemma 7.** For any \( k = 1, \ldots, K \) consider column \( h_{a - M + b} \) in \( H^{(1, \ldots, k)}_l \), where \( a = 0, \ldots, k - 1 \) and \( b = 1, \ldots, M \). In this case \( h_{a - M + b} \) occurs only in the first \( m = N - M + 1 + \min(M - l - 1, M - b) + \min(M - l - 1, b - 1) \) blocks of \( H_{eff}^{(l), k} \).

**Proof:** Straight forward from the definition of the blocks of \( H_{eff}^{(l), k} \) in (12), (13) and (14).

**D. Upper Bound on the Error Probability**

In this subsection we derive for each channel realization an upper bound on the error probability of the joint ML decoder of \( K \) ensembles of IC’s transmitted on the unconstrained multiple-access channels, assuming each IC is \( D_l T_l \)-complex dimensional.

In accordance with the definitions in [V-C] we denote the effective channel of any set of users pulled together by \( H_{eff}^{(l), s} \), where \( s \subseteq \{1, \ldots, K\} \). We define \( |H_{eff}^{(l), s}_l| \cdot H_{eff}^{(l), s} = \rho^{-\sum_{i=1}^{s} D_l T_l} \eta_i^{(s)} \), where \( \rho \eta_i^{(s)} \) is the \( i \)th singular value of \( H_{eff}^{(l), s} \), \( 1 \leq i \leq |s| \cdot D_l T_l \). We also define \( \eta_i^{(s)} = (\eta_1^{(s)}, \ldots, \eta_{|s| \cdot D_l T_l})^T \). Note that in our setting \( N T_l \geq K \cdot D_l T_l \).

**Theorem 6.** Consider \( K \) ensembles of \( D_l T_l \)-complex dimensional IC’s transmitted on the unconstrained multiple-access channels with effective channel \( H_{eff}^{(l), s} \) and densities \( \gamma_{tr} \) and \( \eta_i^{(s)} \) for any \( s \subseteq \{1, \ldots, K\} \) and any \( 1 \leq i \leq |s| \cdot D_l T_l \). The average decoding error probability of the joint ML decoder is upper bounded by

\[
P_e(H_{eff}^{(l), K}, \rho) \leq \sum_{s \subseteq \{1, \ldots, K\}} D(|s| \cdot D_l T_l) \rho^{-\sum_{i=1}^{s} D_l T_l} \eta_i^{(s)}
\]

\[
= \sum_{s \subseteq \{1, \ldots, K\}} D(|s| \cdot D_l T_l) \rho^{-\sum_{i=1}^{s} D_l T_l \gamma_{tr} \sum_{i=1}^{s} D_l T_l} |H_{eff}^{(l), s}_l|^T \cdot H_{eff}^{(l), s}\n\]

where \( D(|s| \cdot D_l T_l) \) is a constant independent of \( \rho \), and \( \eta_i^{(s)} \geq 0 \) for any \( s \subseteq \{1, \ldots, K\} \) and any \( 1 \leq i \leq |s| \cdot D_l T_l \).

**Proof:** The proof is based on dividing the error event into events of error for different sets of users (disjoint events). Then we show that the upper bound on the error probability for the point-to-point channel derived in [8] can be used to upper bound the probability for each of these events. The full proof is in appendix I.

We wish to emphasize that the constraint of \( \eta_i^{(s)} \geq 0 \), for \( i = 1, \ldots, |s| \cdot D_l T_l \) and for any \( s \subseteq \{1, \ldots, K\} \) results from the fact that the same ensemble is upper bounded for any channel realization. In cases where it is possible to fit an ensemble to each channel realization, i.e. the case where the transmitter knows the channel, the upper bound applies also without this restriction.

**E. Achieving the Optimal DMT**

In this subsection we show that the transmission scheme proposed in [V-B] attains the optimal DMT for \( N \geq (K + 1) M - 1 \), \( d_r^*(FC) (\max(r_1, \ldots, r_K)) \). We base the proof on the upper bound on the error probability derived in Theorem 6 This upper bound consists of the sum of several terms, one for each \( s \subseteq \{1, \ldots, K\} \). Each term depends on the determinant corresponding to its effective channel \( |H_{eff}^{(l), s}_l|^T \cdot H_{eff}^{(l), s}\n\) of each ensemble lower bound the determinant of the orthogonal elements of its columns (when the number of rows is larger

\footnote{Note that in [V-C] we considered the case of the first \( k \) users where \( k = 1, \ldots, K \). The extension to any \( s \subseteq \{1, \ldots, K\} \) is immediate.}
than the number of columns). We average the upper bound over the channel realizations and show it attains the optimal DMT in Theorem 4 and also prove that the results apply to lattices when regular lattice decoder is employed in the receiver, in Theorem 8.

Each term in the upper bound in Theorem 6 can be viewed as the error probability of a point-to-point channel with \(|s| \cdot M\) transmit antennas and \(N\) receive antennas, while transmitting an \(|s| \cdot D_1 \cdot T\)-complex dimensional IC in the method described in IV-B. We wish to emphasize that in this subsection we show that the terms corresponding to \(|s| = 1\) attain the required optimal DMT since each user uses an optimal transmission scheme for the point-to-point channel with \(M\) transmit and \(N\) receive antennas. However, for the terms corresponding to \(1 < |s| \leq K\) the effective transmission scheme is no longer optimal and does not necessarily attain the optimal DMT for a point-to-point channel with \(|s| \cdot M\) transmit and \(N\) receive antennas. In fact it does not even necessarily attain \(d^*_{|s| \cdot M} (\max (r_1, \ldots, r_K))\). Hence, the challenge in this subsection is to upper bound the DMT of these terms and show that, although not optimal for the corresponding point-to-point channel, they attain the optimal DMT of the multiple-access channels for the case \(N \geq (K + 1) M - 1\).

The average decoding error probability equals to the average over all channel realizations, i.e.

\[
\overline{\text{Pe}}(\rho) = E_H \left( \overline{\text{Pe}} \left( H^{(l)}_{\text{eff}}, K, \rho \right) \right).
\]  

(16)

Based on Theorem 5 we get the following upper bound on the average decoding error probability

\[
\overline{\text{Pe}}(\rho) \leq \sum_{s \subseteq \{1, \ldots, K\}} E_H \left( D(|s| \cdot D_1 \cdot T) \rho^{-T_l(|s| D_1 - \sum_i \rho_i)} \cdot |H^{(l)}_{\text{eff}}(s)\|^\dagger \cdot H^{(l)}_{\text{eff}}(s)\|^\dagger \right) \cdot E_H \left( H^{(l)}_{\text{eff}}(s)\|^\dagger \cdot H^{(l)}_{\text{eff}}(s)\|^\dagger \right)^{-1}.
\]  

(17)

Note that \(E_H \left( |H^{(l)}_{\text{eff}}(s)\|^\dagger \cdot H^{(l)}_{\text{eff}}(s)\|^\dagger \right) = E_H \left( |H^{(l)}_{\text{eff}}(s)\|^\dagger \right)\) for any \(|s| = k\), where \(k = 1, \ldots, K\), i.e. the mean value for any the users equals to the mean value for the first \(k\) users. Therefore, by replacing \(H^{(l)}_{\text{eff}}(s)\) with \(H^{(l)}_{\text{eff}}(s)\) we can write (17) as follows

\[
\overline{\text{Pe}}(\rho) \leq \sum_{s \subseteq \{1, \ldots, K\}} D(|s| \cdot D_1 \cdot T) \rho^{-T_l(|s| D_1 - \sum_i \rho_i)} \cdot E_H \left( |H^{(l)}_{\text{eff}}(s)\|^\dagger \cdot H^{(l)}_{\text{eff}}(s)\|^\dagger \right)^{-1}.
\]  

(18)

where \(H^{(l)}_{\text{eff}}(s)\) is the effective channel of the first \(|s|\) users, as defined in subsection IV-C.

The channel matrix \(H\) consists of \(N \cdot K \cdot M\) i.i.d entries, where each entry has distribution \(h_{i,j} \sim \mathcal{CN}(0,1), 1 \leq i \leq N\) and \(1 \leq j \leq K \cdot M\). Without loss of generality we consider the case where the columns of \(H\) are drawn sequentially from left to right, i.e. \(h_{1,j}\) is drawn first, then \(h_{2,j}\) is drawn et cetera. Column \(h_{j,1}\) is an \(N\)-dimensional vector. Given \(h_{1,1}, \ldots, h_{1,N-1}\), let us denote by \(\tilde{h}_j \in \mathbb{C}^N\) the elements of the projection of \(h_{j,1}\) on an orthonormal basis that depends on \(h_{1,1}, \ldots, h_{j-1,1}\). We can write

\[
\tilde{h}_j = \Theta(h_{1,1}, \ldots, h_{j-1,1}) \cdot \tilde{h}_j
\]  

(19)

where \(\Theta(\cdot)\) is an \(N \times N\) unitary matrix. \(\Theta(\cdot)\) is chosen such that:

1) The first element of \(\tilde{h}_j\), \(h_{j,1}\), is in the direction of \(\tilde{h}_{j-1}\).

2) The second element, \(h_{2,j}\), is in the direction orthogonal to \(\tilde{h}_{j-1}\), in the hyperplane spanned by \(\{\tilde{h}_{1,1}, \tilde{h}_{2,1}\}\).

3) Element \(h_{3,j-1}\) is in the direction orthogonal to the hyperplane spanned by \(\{\tilde{h}_{2,1}, \ldots, \tilde{h}_{j-1,1}\}\) inside the hyperplane spanned by \(\{\tilde{h}_{1,1}, \ldots, \tilde{h}_{j-1,1}\}\).

4) The rest of the \(N - j + 1\) elements are in directions orthogonal to the hyperplane \(\{h_{1,1}, \ldots, h_{j-1,1}\}\).

Note that \(h_{i,j}, 1 \leq i \leq N, 1 \leq j \leq K \cdot M\) are i.i.d random variables with distribution \(\mathcal{CN}(0,1)\). Let us denote by \(h_{j,1}, \ldots, h_{j-k,1}\) the component of \(\tilde{h}_j\) which resides in the \(N - k\) subspace which is perpendicular to the space spanned by \(\{\tilde{h}_{j-1,1}, \ldots, \tilde{h}_{j-k,1}\}\).

In this case we get

\[
\|\tilde{h}_{j,1}, \ldots, h_{j-k,1}\|^2 = \sum_{i=k+1}^N |\tilde{h}_{i,j}|^2, 1 \leq k \leq j - 1.
\]  

(20)

If we assign \(|\tilde{h}_{i,j}|^2 = \rho^{-\xi_{i,j}}\), we get that the probability density function (PDF) of \(\xi_{i,j}\) is

\[
f(\xi_{i,j}) = C \cdot \log \rho \cdot \rho^{-\xi_{i,j}} \cdot e^{-\rho^{-\xi_{i,j}}}
\]  

(21)

where \(C\) is a normalization factor. In our analysis we assume a very large value for \(\rho\). Hence we can neglect events where \(\xi_{i,j} < 0\) since in this case the PDF (21) decreases exponentially as a function of \(\rho\). For a very large \(\rho\), \(\xi_{i,j} \geq 0, 1 \leq i \leq N\) and \(1 \leq j \leq K \cdot M\), the PDF takes the following form

\[
f(\xi_{i,j}) \propto \rho^{-\xi_{i,j}} \quad \xi_{i,j} \geq 0.
\]  

(22)

In this case by assigning in (20) the vector \(\xi_j = (\xi_{1,j}, \ldots, \xi_{N,j})^T\) with PDF which is proportional to \(\rho^{-\sum_{i=1}^N \xi_{i,j}}\), we get
where $1 \leq k \leq j - 1$. In addition
\begin{equation}
\|h_{aM,b}\|^2 \geq \rho - \min_{z \in \{1, \ldots , N\}} \xi_z, j
\end{equation}

As presented in (18), in order to calculate the upper bound on the error probability we need to consider only the effective channel of the first $|s|$ users, $1 \leq |s| \leq K$. Hence, in order to obtain an upper bound on the error probability we wish to lower bound the determinant $|H_{\text{eff}, s}^{(1), s}, H_{\text{eff}, s}^{(1), s}|$ by lower bounding the contribution of each column in the channel matrix $H$ to the determinant. The following lemma presents a lower bound on the determinant.

**Lemma 8.**
\begin{equation}
|H_{\text{eff}, s}^{(1), s}\| H_{\text{eff}, s}^{(1), s}| \geq \prod_{i=0}^{s-1} \prod_{b=1}^{M} \rho^{-(N-M+1+\min(M-1,M-b)) \cdot \min_{z \in \{aM+b, \ldots , N\}} \xi_z, aM+b} \prod_{b=2}^{M} \sum_{i=1}^{\min(M-l-1, b'-1)} \prod_{z,aM} \min_{z \in \{aM+b', \ldots , N\}} \xi_z, aM+b'.
\end{equation}

**Proof:** The proof is in appendix J. Essentially, the term $(N-M+1+\min(M-1,M-b)) \cdot \min_{z \in \{aM+b, \ldots , N\}} \xi_z, aM+b$ indicates that in the lower bound column $h_{aM+b}$ occurs $N-M+1+\min(M-1,M-b)$ times with $h_{aM+b}$ to its left. Therefore, only the elements of $h_{aM+b}$ which are orthogonal to this set of columns, $\xi_z, aM+b$, where $aM+b \leq z \leq N$ contribute to the lower bound.

The term $\prod_{i=1}^{\min(M-l-1, b'-1)} \cdot \min_{z \in \{aM+b'-i, \ldots , N\}} \xi_z, aM+b'$ indicates that column $h_{aM+b'}$ occurs $\min(M-l-1, b'-1)$ times. However, this time we handle the contribution of the orthogonal elements more carefully. For $1 \leq i \leq \min(M-l-1, b'-1)$ we consider the elements in $h_{aM+b'}$ which are orthogonal to the set of columns $h_{aM+b'}$, $h_{aM+b'-i}$, $\ldots$, $h_{aM+b'-i-1}$.

Now we are ready to lower bound the transmission scheme DMT, based on the lower bound on the determinant in Lemma 8. Let us denote the maximal multiplexing gain by $r_{\text{max}} = \max(1, \ldots , K)$, and also assume $l = \lceil r_{\text{max}} \rceil$.

**Theorem 7.** Consider $K$ sequences of ensembles of $D(T_1)$-complex dimensional IC’s transmitted over the unconstrained multiple-access channels, where each user transmits multiplexing-gain $r_i$ using $G_{\text{opt}}^{(i)} |r_{\text{max}}|$ for each segment of the optimal DMT there exists a sequence of IC’s that attain DMT which is lower bounded by $d_{M,N}^{(\text{opt})} (r_{\text{max}})$.

**Proof:** We use the upper bound on the error probability derived in Theorem 6 and the lower bound on the determinant in Theorem (157) in order to give a new upper bound on the error probability. We average this upper bound over the channel realization, and show that for large $\rho$ the diversity order of the most dominant error event is lower bounded by $d_{M,N}^{(\text{opt})} (r_{\text{max}})$. The full proof is in appendix K.

In Theorem 8 we have shown that for $N \geq (K+1) M - 1$ the DMT of any IC is upper bounded by $d_{M,N}^{(\text{opt})} (r_{\text{max}})$. On the other hand in Theorem 7 we have shown that there exists a sequences of IC’s that attain DMT which is lower bounded by $d_{M,N}^{(\text{opt})} (r_{\text{max}})$. Hence, the transmission scheme must attain the optimal DMT.

In the next theorem we prove the existence of a sequence of lattices that attains the optimal DMT as in Theorem 2.

**Theorem 8.** For each tuple of multiplexing gains $(r_1, \ldots , r_K)$ there exist $K$ sequences of $2D(T_1)$-real dimensional lattices transmitted over the unconstrained multiple access channels that attain diversity order of $d_{M,N}^{(\text{opt})} (r_{\text{max}})$, when regular lattice decoder is employed, where $l = \lceil r_{\text{max}} \rceil$.

**Proof:** See appendix K.

Now we show that for each segment of the optimal DMT there exists a sequence of $K$ lattices that attains it, i.e. the optimal DMT consists of $M$ segments, each in the range $l \leq r_{\text{max}} \leq l + 1$ where $l = 0, \ldots , M-1$, and there are $M$ sequences of lattices that attain it.

**Corollary 2.** For the case $N \geq (K+1) M - 1$ each segment of the optimal DMT of the unconstrained multiple-access channels, $d_{M,N}^{(\text{opt})} (r_{\text{max}})$, is attained by a sequence of $K$, $2D_{r_{\text{max}}} T_{r_{\text{max}}}$-real dimensional lattices.

**Proof:** See appendix K.

V. DISCUSSION

In this section we discuss the results presented in the paper. As an illustrative example we consider the case where there are two users, each with two transmit antennas, i.e. $K = M = 2$. We consider the symmetric case where $r_1 = r_2 = r$, and explain based on Theorem 4 why for $N = 2, 4$ IC’s are suboptimal. On the other hand based on Theorem 5 and Theorem 6
we explain why the optimal DMT is attained for $N \geq 5$. The analysis in this section is somewhat loosed and we refer the reader to sections III, IV for the full analysis.

We begin by giving a short reminder to the behavior of lattices in a point-to-point channel when $M = N = 2$, as presented in [8]. We consider in this discussion lattices although the results apply to general IC’s. The optimal DMT equals $d_{2,2}^{s,\text{(FC)}}(r) = 4 - 3r$ in the range $0 \leq r \leq 1$, and in order to attain it the average number of dimensions per channel use, $D$, must be equal to $\frac{4}{3}$. We wish to explain why when $D \neq \frac{4}{3}$ the optimal DMT is not attained in the range $0 \leq r \leq 1$. For lattices, obtaining multiplexing gain $r > 0$ requires scaling each dimension of the lattice by $\rho = \frac{r}{D}$. When $D < \frac{4}{3}$ diversity order of 4 may be attained for $r = 0$. However, the scaling is too strong and does not enable to attain the optimal DMT for any $r > 0$ (there are not enough degrees of freedom to attain the straight line $4 - 3r$). On the other hand when $D > \frac{4}{3}$, the lattice “fills” too much of the space and the channel induces error probability that does not enable to attain diversity order of 4 for $r = 0$, and therefore does not allow attaining the optimal DMT in the range $0 \leq r \leq 1$. Hence, choosing $D = \frac{4}{3}$ balances the effect of the scaling and the channel for the lattice and allows to attain the optimal DMT in the range $0 \leq r \leq 1$. We now follow this intuition to discuss the multiple-access channels.

A. Why for $N < (K + 1) M - 1$ IC’s are Suboptimal

The error event in multiple-access channels can be divided into the disjoint error events of any subset of the users, as described in Theorem [6] Consider a certain subset of users $s \subseteq \{1, \ldots, K\}$. Due to the distributed nature of the multiple-access channels, the error probability for this subset is upper bounded by the error probability of a point-to-point channel with $|s| \cdot M$ transmit and $N$ receive antennas, i.e. corresponding to a point-to-point channel where the users in $s$ are pulled together. Hence, the DMT in the multiple-access channels is determined by the most probable error event. For the unconstrained multiple-access channels the problem is more involved as each IC has a certain average number of dimensions per channel use. Assume user $i$ has $D_i$ average number of dimensions per channel use, where $1 \leq i \leq K$. When considering the error event of users in $s$, we consider an IC with $\sum_{i \in s} D_i$ average number of dimensions per channel use. The DMT in this error event is upper bounded by $d_{\sum_{i \in s} D_i}^{s,\text{(FC)}}(|s| \cdot r)$, i.e. the bounds derived in [8] for the point-to-point channel. In case the dimensions of any subset of the users do not “align”, i.e. in case a certain subset of the users has average number of dimensions per channel use that is too large or too small to attain the optimal DMT, we get sub-optimality. In this subsection we take as example the case $M = K = 2$ and explain why for $N = 2, 4$ the dimensions do not align, and therefore the optimal DMT is not attained.

Let us begin with the case $M = K = 2$. In this case the optimal DMT in the symmetric case equals

$$d_{2,2}^{s,\text{(FC)}}(r) = \max_{0 \leq r \leq 1} \left\{ \begin{array}{ll} d_{2,2}^{s,\text{(IC)}}(r) & \text{if } 0 \leq r \leq \frac{2}{3} \\ d_{4,2}^{s,\text{(FC)}}(2r) & \text{if } \frac{2}{3} < r \leq 1 \end{array} \right\}$$

On the other hand the optimal DMT of IC’s in this case is upper bounded by $d_{2,2}^{s,\text{(IC)}}(r) = 4(1 - r)$, which is smaller than the optimal DMT for any $0 < r < 1$. First, note that in the symmetric case we must choose $D_1 = D_2$ to maximize the IC’s DMT, i.e. the users have the same average number of dimensions per channel use. The sub-optimality results from the fact that $N = 2$ and so each user can not transmit more than one average number of dimensions per channel use, where in [8] it was shown that each user needs to transmit $\frac{4}{3}$ average number of dimensions per channel use in order to attain $d_{2,2}^{s,\text{(FC)}}(r)$ in the range $0 \leq r \leq \frac{2}{3}$. In addition, the maximal diversity order each user may attain is 4 since $M = N = 2$, and also $d_{2,2}^{s,\text{(IC)}}(r)$ is a straight line. Hence, even when transmitting one dimension per channel use the DMT must be smaller than $6 - 6r$. Therefore, in this case the dimension mismatch manifest itself in the fact that $N$ is too small even to attain the first line of $d_{2,2}^{s,\text{(FC)}}(r)$.

For $K = M = 2$ and $N = 4$ it was shown in Theorem 4 for the symmetric case that IC’s are suboptimal in the range $1 < r < \frac{4}{3}$. In this range the DMT of IC’s is upper bounded by $7 - 4r$, attained when $D_1 = D_2 = \frac{7}{4}$. The dimension mismatch manifests itself in this example both in error events of a single user, and the error event of both users. For error events of a single user the optimal DMT is $d_{2,4}^{s,\text{(FC)}}(r)$ which is also the optimal DMT of the multiple-access channels in the range $1 < r \leq \frac{N}{K+1} = \frac{4}{5}$. The average number of dimensions per channel use required to attain $d_{2,4}^{s,\text{(FC)}}(r)$ for $1 \leq r \leq 2$ is 2 which is larger than $D_1 = D_2 = \frac{7}{4}$. Therefore, for the single user error events the scaling of the IC of each user is too strong and does not enable to attain the optimal DMT. On the other hand, for the two users error event the optimal DMT is $d_{4,4}^{s,\text{(FC)}}(2r)$ which is also the optimal DMT in the range $\frac{4}{3} \leq r \leq 2$. The effective IC of the two users pulled together has average number of dimensions per channel use $D_1 + D_2 = \frac{7}{2}$, which is too large compared to what is required to attain $d_{2,2}^{s,\text{(FC)}}(2r)$ in the range $1 < r < \frac{3}{2}$. Hence, for this error event we get that the effective IC fills too much of the space and so the channel does not enable to attain the optimal DMT.

B. Why for $N \geq (K + 1) M - 1$ IC’s Attain the Optimal DMT

For the case where $N \geq (K + 1) M - 1$ there is no longer a dimension mismatch. However, the condition that there is no dimension mismatch is merely a necessary condition in order to attain the optimal DMT. Hence, in this subsection we will
explain why the optimal DMT is attained based on the transmission scheme presented in subsection IV-B and on the effective channel presented in IV-C.

We consider as an example the case $M = K = 2$ and $N = 5$. We show why in this case the single user performance $d_{2,5}^{(FC)}(r_{\text{max}})$ is attained. For simplicity we will focus on the symmetric case. Essentially, we show in this example that IC’s attain the first DMT line, $10 - 6r$, which coincides with the optimal DMT $d_{2,5}^{(FC)}(r)$ in the range $0 \leq r \leq 1$. The transmission scheme is $G_{0}^{(1,2)}$ presented in [11]. Note that each user transmits an optimal transmission scheme for the point-to-point channel with 2 transmit and 5 receive antennas. Hence, the DMT of the error events of each of the users, is upper bounded by $10 - 6r$ the optimal DMT in the range $0 \leq r \leq 1$. What is left to show is that the DMT of the error event of the two users is also upper bounded by $10 - 6r$. In this case we consider the effective lattice of the two users pulled together, i.e. an error event of a lattice transmitted in a point-to-point channel with 2 transmit and 5 receive antennas. For this lattice the average number of dimensions per channel use equals $D_1 + D_2 = \frac{10}{2}$. We will show that for $r = 0$ this lattice attains diversity order 10. This will lead to DMT $10 - 6r$ since the lattice DMT is a straight line and $D_1 + D_2 = \frac{10}{2}$.

In the receiver, the effective radius of the lattice of the two users pulled together, when $r = 0$, is

$$r_{\text{eff}}^2 \triangleq |V|^{-\frac{1}{2}} = \gamma_{\text{rc}}^{-\frac{1}{2}}(D_1 + D_2)^{\frac{1}{2}} \triangleq |H_{\text{eff}}^{(l=0),K}|H_{\text{eff}}^{(l=0),K}^{-\frac{1}{2}}$$

(26)

where $|V| = \gamma_{\text{rc}}^{-\frac{1}{2}}$ is the volume of the Voronoi region of the effective lattice in the receiver. Recall that for lattices $r_{\text{eff}} \geq r_{\text{packing}} = d_{\text{min}}^{(\text{lattice})}$, where $r_{\text{packing}}$ and $d_{\text{min}}^{(\text{lattice})}$ are the packing radius and the minimal distance of the lattice respectively. We are interested in the event where $r_{\text{eff}}^2$ is in the order of the additive noise variance $\rho^{-1}$. In this case $(d_{\text{min}}^{(\text{lattice})})^2$ is in the order of the noise variance or worse, and so the error probability does not reduce with $r$. In subsection IV-E it is shown that this event is the dominant error event in determining the DMT of the transmission scheme. From (26) we get that $H_{\text{eff}}^{(l=0),K}$ determines the effective radius in the receiver. From [11] and the description of the effective channel in subsection IV-C we get that $H_{\text{eff}}^{(l=0),K}$ is a block diagonal matrix, where 4 of its blocks equal $H \in \mathbb{C}^{4 \times 5}$. For large $r$, the most probable error event $(r_{\text{eff}}^2 = \rho^{-1})$ occurs when the determinant of $H$ reduces with $r$, and the determinants of the rest of the blocks in $H_{\text{eff}}^{(l=0),K}$ remain constant with $r$. Note that if $|H^\dagger H| = \rho^{-\alpha}$, then most likely that the smallest singular value of $H$ equals $\rho^{-\alpha}$ and the rest of the singular values remain constant [3]. In this case we get $|H^\dagger H| = \rho^{-\alpha}$ with a PDF proportional to $\rho^{-2\alpha}$. By assigning $(D_1 + D_2) T = 20$ and $|H_{\text{eff}}^{(l=0),K}|H_{\text{eff}}^{(l=0),K}^{-\frac{1}{2}} = |H^\dagger H| = \rho^{-\alpha}$ in (26) we get that

$$r_{\text{eff}}^2 = |H^\dagger H|^{-\frac{1}{2}} = \rho^{-\alpha}$$

(27)

with a PDF proportional to $\rho^{-2\alpha}$. Hence, $r_{\text{eff}}^2 = \rho^{-1}$ when $\alpha = -5$. Based on subsection IV-E we get for large $r$ that this is the most dominant error event, and by assigning $\alpha = 5$ we get that it happens with probability $\rho^{-10}$. Therefore, in this case diversity order of 10 is attained.

For general $N = (K + 1) M - 1$ each user transmits an optimal transmission scheme for a point-to-point channel with $M$ transmit and $N$ receive antennas. Since the users do not cooperate, in the worst case we get that $H_{\text{eff}}^{(l=0),K}$ has $N - M + 1$ blocks that equal $H \in \mathbb{C}^{N \times M}$. For large $r$ we get that $|H^\dagger H| = \rho^{-\alpha}$ with a PDF proportional to $\rho^{-\alpha(N - K \cdot M + 1)}$. In this case $(D_1 + D_2) T = K \cdot M \cdot M$ and so we get

$$r_{\text{eff}}^2 \triangleq \rho^{-\frac{(N - M + 1)\alpha}{K \cdot M}}$$

(28)

Since $N = (K + 1) M - 1$ we get that $N - M + 1 = K \cdot M$ and $N - K \cdot M + 1 = M$. Hence, by substituting in (28) we get

$$r_{\text{eff}}^2 \triangleq \rho^{-\frac{\alpha}{K \cdot M}}$$

(29)

with a PDF proportional to $\rho^{-M \cdot \alpha}$. Hence we get for large $r$ that $r_{\text{eff}}^2 = \rho^{-1}$ with probability $\rho^{-MN}$, which leads to a diversity order $MN$ when $r = 0$. For any error event of the users in $s \subseteq \{1, \ldots, K\}$, the diversity order will be larger or equal to $MN$ when $r = 0$.

In summary, since the users do not cooperate we get in the worst case $N - M + 1$ occurrences of $H$ in the blocks of $H_{\text{eff}}^{(l=0),K}$. However, when $N \geq (K + 1) M - 1$ there are sufficient amount of receive antennas such that the probability that $H$ has small determinant, is small enough to compensate for the impact of $H$ on $r_{\text{eff}}^2$.

VI. SUMMARY AND FURTHER RESEARCH

This work investigates the DMT of the unconstrained multiple-access channels. For the case $N \geq (K + 1) M - 1$ an explicit upper bound on the optimal DMT of IC’s for any multiplexing-gain tuple is presented. The upper bound coincides with the optimal DMT of multiple-access channels for finite constellations. A transmission scheme that attains this upper bound is also introduced and analyzed.

For the case $N < (K + 1) M - 1$ an upper bound on the optimal DMT of IC’s is derived. For the general case this upper bound remains in the form of a maximization problem. This maximization problem depends on $|s|$, the number of IC’s pulled together, where $1 \leq |s| \leq K$, and on the average number of dimensions per channel use for each user. On the other hand for
finite constellations the maximization depends only on the number of users pulled together. Hence, finding the upper bound on the optimal DMT of IC’s is more involved. For the symmetric case, where all users transmit with the same multiplexing gain, an explicit upper bound on the optimal DMT of IC’s is presented for the case $N < (K + 1)M - 1$. By using this upper bound, it is shown that IC’s are suboptimal compared to finite constellations in this case.

While this work presents a transmission scheme that attains the optimal DMT for the case $N \geq (K + 1)M - 1$, for the case $N < (K + 1)M - 1$ the upper bound on the optimal DMT of IC’s is attained only for some cases. For instance whenever $N = 1$, orthogonalization attains the optimal DMT of IC’s in the symmetric case. Also for $K = 2$, $M = 2$ and $N = 3$, the transmission scheme presented in this paper attains the upper bound on the optimal DMT of IC’s in the symmetric case. However, finding a transmission scheme that attains the upper bound on the optimal DMT for all $N < (K + 1)M - 1$, remains an open problem even in the symmetric case.

APPENDIX A

PROOF OF LEMMA 1

The proof outline is as follows. First we show that for finite constellations, the single user DMT is smaller than the contracted optimal DMT of any number of users (up to $K$) pulled together. Then we use this relation, together with the anchor points presented in Corollary 1 for the upper bound on IC’s DMT, in order to prove the lemma.

Since $K > 1$ and $M$ are positive integers, we get for $N \geq (K + 1)M - 1$ that $M = \frac{N}{K + 1}$, where $1 \leq i \leq K$. Hence for any $d^i_{i,M,N}(i \cdot r)$, the range of average number of dimensions per channel use per user is $0 \leq D \leq \min(M, \frac{N}{K + 1}) = M$, where $1 \leq i \leq K$.

We begin by showing that $d^i_{i,M,N}(FC)(r)$ is smaller or equal to $d^i_{i,M,N}(i \cdot r)$ for $2 \leq i \leq K$, where $d^i_{i,M,N}(i \cdot r)$ is the optimal DMT of finite constellations contracted by $i$, in a point-to-point channel with $i \cdot M$ transmit and $N$ receive antennas. For the case $N > (K + 1)M - 1$ we get that $\frac{N}{K + 1} \geq M$. Hence we also get that $\frac{N}{K + 1} \geq M$ for $1 \leq i \leq K$. Hence from Theorem 3 we can see that

$$d^i_{i,M,N}(FC)(r) \leq d^i_{i,M,N}(i \cdot r) \quad 2 \leq i \leq K$$

by replacing $K$ with $i$.

For $N = (K + 1)M - 1$ we still get that $\frac{N}{K + 1} \geq M$ for $1 \leq i \leq K - 1$, and again based on Theorem 3

$$d^i_{i,M,N}(FC)(r) \leq d^i_{i,M,N}(i \cdot r) \quad 2 \leq i \leq K - 1.$$  \hspace{1cm} (31)

There remains a case of $i = K$. We can see that for $N = (K + 1)M - 1$ we get $M - \frac{1}{K} \leq \frac{N}{K + 1} \leq M$. Hence we get from Theorem 3

$$d^i_{i,M,N}(FC)(r) \leq d^i_{i,M,N}(K \cdot r) \quad 0 \leq r \leq M - \frac{1}{K}.$$  \hspace{1cm} (32)

For $M - \frac{1}{K} \leq r \leq M$ both $d^i_{i,M,N}(FC)(r)$ and $d^i_{i,M,N}(K \cdot r)$ are on the last straight line of the piecewise linear functions. By simply assigning $N = (K + 1)M - 1$ we get for $M - \frac{1}{K} \leq r \leq M$

$$d^i_{i,M,N}(FC)(r) = d^i_{i,M,N}(K \cdot r) = KM (M - r).$$  \hspace{1cm} (33)

From (30)-(33) we get for $N \geq (K + 1)M - 1$ and $0 \leq r \leq M$ that

$$d^*_{i,M,N}(FC)(r) \leq d^*_{i,M,N}(i \cdot r) \quad 2 \leq i \leq K.$$  \hspace{1cm} (34)

So far we have proved the relation between the contracted optimal DMT of finite constellations with different number of users pulled together. Next we wish to use it in order to prove the relation between $d^i_{i,M,N}(i \cdot r)$ for $1 \leq i \leq K$. In [8, Corollary 2] it was shown that for $0 < D \leq \min(M, N)$

$$d^i_{i,M,N}(FC)(r) \leq d^*_{i,M,N}(r) \quad 0 \leq r \leq D.$$  \hspace{1cm} (35)

On the other hand from Corollary 1 we can see that

$$d^i_{i,M,N}(l) = d^i_{i,M,N}(FC)(l) = (i \cdot M - l)(N - l) \quad 1 \leq i \leq K$$

for $l = 0$ when $0 \leq i \cdot D \leq \frac{i \cdot MN}{i \cdot M + N - 1}$, and also for $l = 1, \ldots, i \cdot M - 1$ when $\frac{i \cdot MN - (i - 1)l}{i \cdot M + N - 1 - 2l} \leq i \cdot D \leq \frac{i \cdot MN - (i - 1)l}{i \cdot M + N - 1 - 2l}$. Hence based on (34)-(36), and the fact that $d^i_{i,M,N}(i \cdot r)$ is a contraction of $d^i_{i,M,N}(r)$ for $2 \leq i \leq K$ we get

$$d^i_{i,M,N}(0) \geq d^i_{i,M,N}(0) \quad 2 \leq i \leq K.$$  \hspace{1cm} (37)

for $0 \leq D \leq \frac{MN}{i \cdot M + N - 1}$, and

$$d^i_{i,M,N}(l) \geq d^i_{i,M,N}(\frac{M}{2}) \quad 2 \leq i \leq K.$$  \hspace{1cm} (38)
for \(i = 1, \ldots, i \cdot M - 1\) and \(\frac{MN - \frac{i}{i+1} \cdot (i+1)}{i+M + N - 1 - 2(l-1)} \leq D \leq \frac{MN - \frac{i}{i+1} \cdot (i+1)}{i+M + N - 1 - 2l}.\) Since \(d^*_{i:M,N} (i \cdot r), 1 \leq i \leq K,\) are straight lines as a function of \(r,\) and also all of these straight lines are equal zero for \(r = D,\) i.e. \(d^*_{i:M,N} (i \cdot D) = 0\) for \(1 \leq i \leq K,\) the inequalities in (37), (38) leads to

\[
d^*_{i:M,N} (r) \leq d^*_{i:M,N} (i \cdot r) \quad 2 \leq i \leq K
\]

for any \(0 \leq D \leq M\) and \(0 \leq r \leq D.\) This concludes the proof.

**APPENDIX B**

**PROOF OF LEMMA 3**

First note that \(\frac{N}{i+1} \leq \frac{K}{i} \) for \(1 \leq i \leq K - 1.\) Hence from Theorem 3 we get that

\[
d^*_{i:M,N} (r) \leq d^*_{i:M,N} (i \cdot r) \quad 2 \leq i \leq K - 1
\]

for \(0 \leq r \leq \frac{K}{i}.\) Based on (35), (36), (39) and Corollary 1 we get that

\[
d^*_{i:M,N} (0) \geq d^*_{i:M,N} (0) \quad 2 \leq i \leq K - 1
\]

for \(0 \leq D \leq \frac{MN}{i+M + N - 1},\) and

\[
d^*_{i:M,N} (l) \geq d^*_{i:M,N} \left( \frac{l}{i} \right) \quad 2 \leq i \leq K - 1
\]

for \(l = 1, \ldots, i \cdot M - 1\) and \(\frac{MN - \frac{i}{i+1} \cdot (i+1)}{i+M + N - 1 - 2(l-1)} \leq D \leq \frac{MN - \frac{i}{i+1} \cdot (i+1)}{i+M + N - 1 - 2l}.\) Again, since \(d^*_{i:M,N} (i \cdot r), 1 \leq i \leq K,\) are straight lines as a function of \(r,\) and also all of these straight lines are equal to zero for \(r = D,\) the inequalities in (40), (41) lead to

\[
d^*_{i:M,N} (r) \leq d^*_{i:M,N} (i \cdot r) \quad 2 \leq i \leq K - 1
\]

for any \(0 \leq D \leq \frac{K}{i} \) and \(0 \leq r \leq D.\)

**APPENDIX C**

**PROOF OF LEMMA 4**

Since \(M \geq 1\) we get for \(N < (K-1)M + 1\) that \(L = \frac{M}{N}.\) Hence we can consider the range \(0 \leq r \leq \frac{N}{K}.\) We begin the proof by showing that for \(N < (K-1)M + 1,\) \(d^*_{i:M,N} (r)\) is inferior compared to \(d^*_{K:M,N} (K \cdot r),\) for any \(0 \leq D \leq \frac{N}{K}.\) Then we show that the maximization over \(d^*_{i:M,N} (r)\) yields \(M \cdot N - M \cdot K \cdot r.\)

We begin by showing that

\[
d^*_{i:M,N} (r) \leq d^*_{K:M,N} (K \cdot r) \quad 0 \leq D \leq \frac{N}{K}
\]

for \(0 \leq r \leq D.\) By assigning \(D = \frac{N}{K}\) in \(d^*_{K:M,N} (K \cdot r)\) we get

\[
d^*_{K:M,N} (K \cdot r) = (K \cdot M - N + 1) \cdot (N - Kr).
\]

Since \(N < (K-1)M + 1\) we get

\[
d^*_{K:M,N} (0) = (K \cdot M - N + 1) \cdot N > M \cdot N.
\]

From Corollary 1 we get that

\[
d^*_{K:M,N} (0) \leq d^*_{K:M,N} (0) \quad 0 \leq D \leq \frac{N}{K}
\]

and also

\[
d^*_{M,N} (0) \leq M \cdot N \quad 0 \leq D \leq \frac{N}{K}.
\]

Since \(d^*_{i:M,N} (i \cdot r), 1 \leq i \leq K\) are straight lines as a function of \(r,\) that equal to zero for \(r = D,\) and also based on (42), (43), (44) and Lemma 3 we get

\[
d^*_{M,N} (r) \leq d^*_{i:M,N} (i \cdot r) \quad 1 \leq i \leq K
\]

for any \(0 \leq D \leq \frac{N}{K}\) and \(0 \leq r \leq D.\) Hence the optimization problem takes the following form

\[
\max_D \min_{1 \leq i \leq K} d^*_{i:M,N} (i \cdot r) = \max_D d^*_{M,N} (r) \quad 0 \leq r \leq \frac{N}{K}.
\]

For \(N < (K-1)M + 1\) we get that \(\frac{N}{K} < \frac{MN}{N+M-1}.\) Also, from Corollary 1 we get that \(d^*_{M,N} (0) = M \cdot N\) for \(0 \leq D \leq \frac{MN}{N+M-1}.\) Hence, in the range \(0 \leq D \leq \frac{N}{K}\) we get a set of straight lines as a function of \(r, d^*_{M,N} (r),\) where \(d^*_{M,N} (0) = MN.\)
and \( d^*_{M,N}(D) = 0 \). As a result the maximal value for each \( r \) is attained for \( D = \frac{N}{K} \), and equals
\[
\max_D d^*_{M,N}(r) = d^*_{M,N}(r) = MN - KMr \quad 0 \leq r \leq \frac{N}{K}.
\] (47)

APPENDIX D

PROOF OF LEMMA

The outline of the proof is as follows. We begin by finding the straight line that equals \( d^{(FC)}_{M,N}(\lfloor \frac{l}{2} \rfloor + 1) \) when \( r = \lfloor \frac{l}{2} \rfloor + 1 \), and also equals \( d^{(FC)}_{K,M,N}(K - 1)M + \lfloor \frac{l}{2} \rfloor \) for \( r = \frac{(K-1)M+\lfloor \frac{l}{2} \rfloor}{K} \). Then we show that the average number of dimensions per channel use that rotate around the anchor point \( KM \), and also for \( d^{(FC)}_{K,M,N}(K \cdot r) \), \( D_t \) is in the range of average number of dimensions per channel use that rotate around the anchor point \( D_t^{(FC)}(K \cdot \frac{K}{M} + \lfloor \frac{l}{2} \rfloor) \).

By showing that the straight line fulfills Corollary 1 for both cases, we get that the straight line equals \( d^{(FC)}_{M,N}(r) \) and also \( d^{(FC)}_{K,M,N}(K \cdot r) \).

Let us denote the straight line by
\[
d^* (r) = MN - \lfloor \frac{l}{2} \rfloor \cdot \left( \lfloor \frac{l}{2} \rfloor + 1 \right) - 2 \cdot \left( \lfloor \frac{l}{2} \rfloor + 1 \right) \cdot \left( \frac{l}{2} - \lfloor \frac{l}{2} \rfloor \right) - (N + M - 1 - l) r.
\]
First we wish to show that \( d^* (\lfloor \frac{l}{2} \rfloor + 1) = d^{(FC)}_{M,N}(\lfloor \frac{l}{2} \rfloor + 1) \), and also that \( d^* \left( \frac{(K-1)M+\lfloor \frac{l}{2} \rfloor}{K} \right) = d^{(FC)}_{K,M,N}(K - 1)M + \lfloor \frac{l}{2} \rfloor \).

By simply assigning \( r = \lfloor \frac{l}{2} \rfloor + 1 \) we get
\[
d^* \left( \lfloor \frac{l}{2} \rfloor + 1 \right) = \left( N - \lfloor \frac{l}{2} \rfloor - 1 \right) \cdot \left( M - \lfloor \frac{l}{2} \rfloor - 1 \right) = d^{(FC)}_{M,N} \left( \lfloor \frac{l}{2} \rfloor + 1 \right).
\] (48)

For \( r = \frac{(K-1)M+\lfloor \frac{l}{2} \rfloor}{K} \) we consider two cases. In the first case assume \( l = 2b \), i.e. \( l \) is even. Under this assumption \( \lfloor \frac{l}{2} \rfloor = b \), and so \( r = \frac{(K-1)M+2b}{K} \). By assigning \( KM = N + M - 2b \) in \( d^* (r) \) we get
\[
d^* \left( \frac{(K-1)M+b}{K} \right) = MN - b \cdot (b + M + 1) - (K - 1)M^2 = (N - (K - 1)M - b) \cdot (M - b) = d^{(FC)}_{K,M,N}(K - 1)M + b + 1).
\]

In the second case \( l = 2b + 1 \), i.e. \( l \) is odd. In this case we get \( \lfloor \frac{l+1}{2} \rfloor = b + 1 \), \( \lfloor \frac{l}{2} \rfloor = b \) and \( r = \frac{(K-1)M+2b+1}{K} \). By assigning \( KM = N + M - 2b \) in \( d^* (r) \) we get
\[
d^* \left( \frac{(K-1)M+b+1}{K} \right) = MN - (b + 1) \cdot (b + M + 1) - (K - 1)M^2 = d^{(FC)}_{K,M,N}(K - 1)M + b + 1).
\]

Hence from both cases we get
\[
d^* \left( \frac{(K-1)M+\lfloor \frac{l+1}{2} \rfloor}{K} \right) = d^{(FC)}_{K,M,N} \left( (K - 1)M + \lfloor \frac{l+1}{2} \rfloor \right).
\] (49)

Now we wish to show that \( d^* (r) = d^{(FC)}_{K,M,N}(r) = d^{(FC)}_{K,M,N}(K \cdot r) \). We begin by showing that \( d^* (r) = d^{(FC)}_{M,N}(r) \). First note that
\[
d^* (D_t) = d^{(FC)}_{M,N}(D_t) = d^{(FC)}_{K,M,N} (K \cdot D_t) = 0.
\] (50)

Now let us denote \( D^*_{\lfloor \frac{l}{2} \rfloor} = \frac{MN-\lfloor \frac{l}{2} \rfloor \cdot (\lfloor \frac{l}{2} \rfloor + 1)}{N+M-1-2\lfloor \frac{l}{2} \rfloor} \) and \( D^*_{\lfloor \frac{l}{2} \rfloor+1} = \frac{MN-\lfloor \frac{l}{2} \rfloor \cdot (\lfloor \frac{l}{2} \rfloor + 2)}{N+M-1-2\lfloor \frac{l}{2} \rfloor} \). We wish to show that
\[
d^{(FC)}_{M,N}(0) = M \cdot N - \left( \lfloor \frac{l}{2} \rfloor + 1 \right) \cdot \left( \lfloor \frac{l}{2} \rfloor + 2 \right) < d^* (0) \leq M \cdot N - \lfloor \frac{l}{2} \rfloor \cdot \left( \lfloor \frac{l}{2} \rfloor + 1 \right) = d^{(FC)}_{M,N}(0) \cdot \left( \lfloor \frac{l}{2} \rfloor + 1 \right).
\] (51)

In the first case we take \( l = 2b \). In this case
\[
d^* (0) = M \cdot N - b \cdot (b + 1).
\]

On the other hand we also get
\[
M \cdot N - \lfloor \frac{l}{2} \rfloor \cdot \left( \lfloor \frac{l}{2} \rfloor + 1 \right) = M \cdot N - b \cdot (b + 1) = d^* (0)
\]
which proves (51) for the first case. In the second case we consider \( l = 2b + 1 \). In this case
\[
d^* (0) = M \cdot N - (b + 1)^2.
\]
For this case we also get $M \cdot N - \left( \frac{b+1}{2} \right) \left( \left( \frac{b+1}{2} \right) + 1 \right) = M \cdot N - b \cdot (b + 1)$ and $M \cdot N - \left( \frac{b+1}{2} \right) \left( \left( \frac{b+1}{2} \right) + 2 \right) = M \cdot N - (b + 1) \cdot (b + 2)$. It can be easily shown that for $b \geq 0$

$$M \cdot N - (b + 1) \cdot (b + 2) < d^*(0) = M \cdot N - (b + 1)^2 \leq M \cdot N - b \cdot (b + 1)$$

which proves (51) for the second case. From Corollary 1 and (48) we know that

$$d^* \left( \frac{l}{2} + 1 \right) = d_{M,N}^* \left( \frac{l}{2} + 1 \right) = d_{M,N}^* \left( \frac{l}{2} + 1 \right) = d_{M,N}^* \left( \frac{l}{2} + 1 \right).$$

(52)

Since $d^* (r)$, $d_{M,N}^* (r)$ and $d_{M,N}^* (r)$ are all straight lines that fulfill (51), (52) we get

$$d_{M,N}^* \left( \frac{l}{2} + 1 \right) \leq D_1 < D_{M,N}^* \left( \frac{l}{2} + 1 \right).$$

(53)

As a result, from Corollary 1 and (55) we get

$$d_{M,N}^* \left( \frac{l}{2} + 1 \right) = d_{M,N}^* \left( \frac{l}{2} + 1 \right).$$

(54)

Since $d^* (r)$ and $d_{M,N}^* (r)$ are straight lines and based on the equalities in (48), (50) and (54) we get

$$d^* (r) = d_{K,M,N}^* (r).$$

(55)

Next we prove $d^* (r) = d_{K,M,N}^* (K \cdot r)$. Let us denote $r = \frac{(K-1)M+b+1}{K}$ and $D_{r_1} = \frac{MN-(K(r_1-1)r_1)}{K}$. We wish to show

$$d_{K,M,N}^* (0) \leq d^* (0) < d_{K,M,N}^* (0).$$

(56)

We consider two cases. For the first case we take $l = 2 \cdot b$. In this case we get $r_{2b} = \frac{(K-1)M+b}{K}$, $d^* (0) = M \cdot N - b \cdot (b + 1)$ and $N = (K - 1) M + 1 + 2b$. Hence we get

$$d_{K,M,N}^* (0) = KMN - ((K - 1) M + b) (N - b) = MN + b (K - 1) M + b^2.$$ 

(57)

Since $N = (K - 1) M + 1 + 2b$ we get

$$MN - b (N - (K - 1) M) + b^2 = MN - (b + 1)^2.$$ 

(58)

From (57) and (58) we get $d^* (0) = d_{K,M,N}^* (0)$, which proves (56) for the first case. For the second case we take $l = 2b+1$. In this case $r_{2b+1} = \frac{(K-1)M+b+1}{K}$, $d^* (0) = M \cdot N - (b + 1)^2$ and $N = (K - 1) M + 2b + 2$. For this case we get

$$d_{K,M,N}^* (0) = KMN - ((K - 1) M + b) (N - b - 1) = MN + b + 1 (K - 1) M - bN + b (b + 1).$$ 

(59)

Hence according to (56) we need to show

$$MN + b + 1 (K - 1) M - bN + b (b + 1) > MN - (b + 1)^2.$$ 

(60)

By assigning $(K - 1) M = N - 2b - 2$ we get from (60) $N > b + 1$. Since $0 \leq l = 2b + 1 \leq 2M - 3$, the maximal value of $b$ is $b = M - 2$, which gives for $N = (K - 1) M + 2b + l$

$$N > M > M - 1 \geq b + 1.$$ 

Hence we get

$$d^* (0) < d_{K,M,N}^* (0) = d_{K,M,N}^* (0).$$ 

(61)

On the other hand we get

$$d_{K,M,N}^* (0) = KMN - ((K - 1) M + 1 + b) (N - b).$$ 

(62)

Hence according to (56), (62) we need to show that

$$MN + b (K - 1) M - N (b + 1) + b (b + 1) \leq MN - (b + 1)^2$$ 

(63)

which again leads to $N > b + 1$. Hence we get

$$d_{K,M,N}^* (0) = d_{K,M,N}^* (0).$$ 

(64)
From (61) and (64) we get (56) for the second case. Hence we have proved (56). From Corollary 1 and (49) we know that
\[
d^* \left( \frac{(K - 1) M + \left[ \frac{l + 1}{2} \right]}{K} \right) = d^*_{K,M,N} \left( \frac{(K - 1) M + \left[ \frac{l + 1}{2} \right]}{K} \right)
\]
\[
= d^*_{K,M,N} \left( (K - 1) M + \left[ \frac{l + 1}{2} \right] \right) = d^*_{K,M,N} \left( (K - 1) M + \left[ \frac{l + 1}{2} \right] \right). \tag{65}
\]
Since \( d^* \), \( d^*_{K,M,N} \), and \( d^*_{K,M,N} \) are straight lines that fulfill (56), (65) we get
\[
D^*_r < D_l \leq D^*_r + 2.
\tag{66}
\]
As a result, from Corollary 1 and (66) we get
\[
d^*_{K,D_l} \left( (K - 1) M + \left[ \frac{l + 1}{2} \right] \right) = d^*_{K,M,N} \left( (K - 1) M + \left[ \frac{l + 1}{2} \right] \right). \tag{67}
\]
Since \( d^* \) and \( d^*_{K,D_l} \) are straight lines, and based on the equalities in (49), (50) and (67) we get
\[
d^* \left( K \cdot r \right) = d^*_{K,M,N} \left( K \cdot r \right). \tag{68}
\]
From (55), (68) we get the first part of the Lemma, where from (54), (67) we get the second part of the Lemma.

APPENDIX E

PROOF OF THEOREM 4

We begin by showing that \( d^*_{K,M,N} (r) \) is the solution of the optimization problem in (5), i.e. the case where all users have the same average number of dimensions per channel use, \( D \). Then we show that this is also the solution for (8).

First we find \( \max_D \min_{1 \leq i \leq K} \left( d^*_{i,M,N} (i \cdot r) \right) \), where \( 0 \leq r \leq \frac{L}{K} \). For the case \( N \geq (K + 1) M - 1 \), we can see from Lemma 2 that
\[
\max_D \min_{1 \leq i \leq K} \left( d^*_{i,M,N} (i \cdot r) \right) = \max_D d^*_{M,N} (r) = d^*_{M,N} (r). \tag{69}
\]

For the case \( N < (K - 1) M + 1 \) it was shown in Lemma 4 that \( d^*_{K,M,N} (r) \) is the optimization problem solution. For \( N = (K - 1) M + 1 + l \), where \( l = 0, \ldots, 2M - 3 \) we know from Lemma 5 that \( d^*_{M,N} (r) \) is smaller than \( d^*_{i,M,N} (i \cdot r) \) for \( 2 \leq i \leq K - 1 \) and any \( 0 \leq D \leq \frac{L}{K} \), \( 0 \leq r \leq D \). Hence the optimization problem for this case boils down to
\[
\max_D \min \left\{ d^*_{M,N} (r), d^*_{K,M,N} (K \cdot r) \right\} \tag{69}
\]
for \( 0 \leq D \leq \frac{L}{K} \) and \( 0 \leq r \leq D \). From Lemma 5 we know that \( d^*_{M,N} (\lfloor \frac{L}{2} \rfloor + 1) = d^*_{M,N} (\lfloor \frac{L}{2} \rfloor + 1) \). As a result, based on Corollary 1 we get that for \( 0 < D \leq D_l \)
\[
d^*_{M,N} (\lfloor \frac{L}{2} \rfloor + 1) \leq d^*_{M,N} (\lfloor \frac{L}{2} \rfloor + 1) = d^*_{M,N} (\lfloor \frac{L}{2} \rfloor + 1)
\]
and also
\[
d^*_{M,N} (r) = 0 \leq d^*_{M,N} (r) \quad r \geq D.
\]

Hence we get for \( 0 < D \leq D_l \)
\[
d^*_{M,N} (r) \leq d^*_{M,N} (r) \quad \lfloor \frac{L}{2} \rfloor + 1 \leq r \leq \frac{L}{K}. \tag{70}
\]

In a similar manner we also know from Lemma 5 that \( d^*_{K,M,N} ((K - 1) M + \left[ \frac{l + 1}{2} \right]) = d^*_{K,M,N} ((K - 1) M + \left[ \frac{l + 1}{2} \right]) \). As a result, based on Corollary 1 we get that for \( D_l \leq D \leq \frac{L}{K} \)
\[
d^*_{K,M,N} ((K - 1) M + \left[ \frac{l + 1}{2} \right]) \leq d^*_{K,M,N} ((K - 1) M + \left[ \frac{l + 1}{2} \right]) = d^*_{K,M,N} ((K - 1) M + \left[ \frac{l + 1}{2} \right])
\]
and also
\[
d^*_{K,M,N} (K \cdot r) = 0 \leq d^*_{K,M,N} (K \cdot r) \quad r \geq D_l.
\]

Since \( D_l \geq \frac{(K - 1) M + \left[ \frac{l + 1}{2} \right]}{K} \), and these are straight lines, we also get for \( D_l \leq D \leq \frac{L}{K} \)
\[
d^*_{K,M,N} (K \cdot r) \leq d^*_{K,M,N} (K \cdot r) \quad 0 \leq r \leq \frac{(K - 1) M + \left[ \frac{l + 1}{2} \right]}{K}. \tag{71}
\]
Hence, based on (70), (71) and the fact that \( d_{M,N}^{*,D_k} (r) = d_{K,M,N}^{*,K,D_k} (K \cdot r) = d^*(r) \) (Lemma 5), we get that
\[
\max_D \min \left\{ d_{M,N}^{*,D_k} (r), d_{K,M,N}^{*,K,D_k} (K \cdot r) \right\} = d^*(r) = d_{M,N}^{*(IC)} (r) \quad \left\lfloor \frac{l}{2} \right\rfloor + 1 \leq r \leq \frac{(K-1)M + \left\lfloor \frac{l+1}{2} \right\rfloor}{K}.
\]

(72)

Next we find the solution for \( 0 \leq r \leq \left\lfloor \frac{l}{2} \right\rfloor + 1 \). Our starting point is \( D = D_l \) for which \( d_{M,N}^{*,D_k} (r) = d_{K,M,N}^{*,K,D_k} (K \cdot r) \). Since \( d^* \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right) = d_{M,N}^{*(FC)} \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right) \) we get from Corollary 1 and (53) that
\[
\frac{MN - \left\lfloor \frac{l}{2} \right\rfloor \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right)}{M + N - 1 - 2 \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right)} \leq D_l < \frac{MN - \left\lfloor \frac{l}{2} \right\rfloor \left( \left\lfloor \frac{l}{2} \right\rfloor + 2 \right)}{M + N - 1 - 2 \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right)}.
\]

(73)

We also get from Corollary 1 that for \( D_l \leq D \leq \frac{K}{K} \) \( d_{M,N}^{*,D} (r) \leq d_{K,M,N}^{*(FC)} (r) \).

(74)

In addition it can be easily shown that for \( N = (K-1)M + 1 \), where \( l = 0, \ldots, 2M - 3 \)
\[
\left\lfloor \frac{l}{2} \right\rfloor + 1 \leq \frac{N}{K+1} \leq \frac{(K-1)M + \left\lfloor \frac{l+1}{2} \right\rfloor}{K}
\]

(75)

by considering the cases where \( l \) is even and odd, i.e. the cases where \( l = 2b \) and \( l = 2b + 1 \). For the case \( \frac{MN - \left\lfloor \frac{l}{2} \right\rfloor \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right)}{M + N - 1 - 2 \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right)} \leq D \leq D_l \) assume \( d_{K,M,N}^{*,K,D} (K \cdot r) \) rotates around anchor point with multiplexing gain \( m \). In this case there are two possibilities. The first possibility \( \left\lfloor \frac{l}{2} \right\rfloor + 2 \leq m \leq \frac{K}{K} \), where \( m \in \mathbb{Z} \). When this is the case, we get from Corollary 1 that in the range \( \frac{MN - \left\lfloor \frac{l}{2} \right\rfloor \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right)}{M + N - 1 - 2 \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right)} \leq D \leq D_l \)
\[
d_{M,N}^{*,D} \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right) = d_{K,M,N}^{*,K,D} \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right) \leq d_{K,M,N}^{*,K,D} \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right).
\]

(76)

For the second possibility \( 0 \leq m \leq \left\lfloor \frac{l}{2} \right\rfloor + 1 \) we get from (75) and Theorem 1 that
\[
d_{K,M,N}^{*,K,D} (K \cdot m) = d_{K,M,N}^{*,K,D} (K \cdot m) \geq d_{M,N}^{*(FC)} (m) \geq d_{M,N}^{*,D} (m).
\]

(77)

In addition \( d_{M,N}^{*,D} (D) = d_{K,M,N}^{*,K,D} (K \cdot D) = 0 \). Since these are straight lines we get in the range \( \frac{MN - \left\lfloor \frac{l}{2} \right\rfloor \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right)}{M + N - 1 - 2 \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right)} \leq D \leq D_l \)
\[
d_{M,N}^{*,D} (r) \leq d_{K,M,N}^{*,K,D} (K \cdot r).
\]

(78)

By induction, for \( \frac{MN - (s-1) \left\lfloor \frac{l}{2} \right\rfloor \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right)}{M + N - 1 - 2(s-1) \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right)} \leq D \leq \frac{MN - s \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right)}{M + N - 1 - 2s \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right)} \), \( s = \left\lfloor \frac{l}{2} \right\rfloor + 1, \ldots, 1 \), assuming \( d_{K,M,N}^{*,K,D} (K \cdot r) \geq d_{M,N}^{*,D} (r) \), where \( D(s) = \frac{MN - (s-1) \left\lfloor \frac{l}{2} \right\rfloor \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right)}{M + N - 1 - 2(s-1) \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right)} \) we get from similar arguments to (75)-(78) that
\[
d_{M,N}^{*,D} (r) \leq d_{K,M,N}^{*,K,D} (K \cdot r).
\]

(79)

Finally for \( 0 < D \leq \frac{MN}{N+1} \), from the same arguments as in (79) we also get
\[
d_{M,N}^{*,D} (r) \leq d_{K,M,N}^{*,K,D} (K \cdot r).
\]

(80)

Hence, from (78), (79) and (80) we get that in the range \( 0 < D \leq D_l \)
\[
\max_D \min \left\{ d_{M,N}^{*,D} (r), d_{K,M,N}^{*,K,D} (K \cdot r) \right\} = \max_D d_{M,N}^{*,D} (r).
\]

(81)

Since \( D_l \geq \frac{MN - \left\lfloor \frac{l}{2} \right\rfloor \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right)}{M + N - 1 - 2 \left( \left\lfloor \frac{l}{2} \right\rfloor + 1 \right)} \) (79), and also from (74), (81) we get based on Corollary 1
\[
\max_D \min \left\{ d_{M,N}^{*,D} (r), d_{K,M,N}^{*,K,D} (K \cdot r) \right\} = d_{M,N}^{*,D} (r) = d_{K,M,N}^{*,K,D} (K \cdot r) \quad 0 \leq r \leq \left\lfloor \frac{l}{2} \right\rfloor + 1.
\]

(82)

Now we wish to find \( d_{K,M,N}^{*,K,D} (K \cdot r) \) for \( \frac{(K-1)M + \left\lfloor \frac{l}{2} \right\rfloor + 1}{K} \leq r \leq \frac{L}{K} \). Let us denote \( r_l = \frac{(K-1)M + \left\lfloor \frac{l}{2} \right\rfloor + 1}{K} \). Since
\[
d_{K,M,N}^{*,K,D} \left( \left( K-1 \right) M + \left\lfloor \frac{l}{2} \right\rfloor + 1 \right) = d_{K,M,N}^{*,K,D} \left( \left( K-1 \right) M + \left\lfloor \frac{l}{2} \right\rfloor + 1 \right)
\]
we get (66)
\[
\frac{NM - (K \cdot r_l - 1) r_l}{KM + N - 1 - 2(K \cdot r_l - 1)} \leq D_l \leq \frac{NM - r_l (K \cdot r_l + 1)}{KM + N - 1 - 2(K \cdot r_l)}.
\]

(83)
Based on Corollary 1 we get in the range \(0 < D \leq D_l\)
\[
d^{s,K,D}_{K,M,N} (K \cdot r) \leq d^{s,(FC)}_{K,M,N} (K \cdot r).
\] (84)

For \(D_l < D \leq \frac{MN - 2(r_1 + 1)}{K + N - 1 - 2r_1}\) assume \(d^{s,D}_{M,N} (r)\) rotates around anchor point with multiplexing gain \(\frac{K}{r}\), where \(m \in \mathbb{Z}\). In case \(0 \leq m < (K - 1)M + \left\lfloor \frac{r + 1}{2} \right\rfloor\), based on Corollary 1 and Lemma 5 we get
\[
d^{s,D}_{M,N} \left(\frac{(K - 1)M + \left\lfloor \frac{r + 1}{2} \right\rfloor}{K}\right) = d^{s,(FC)}_{K,M,N} \left(\frac{(K - 1)M + \left\lfloor \frac{r + 1}{2} \right\rfloor}{K}\right) \geq d^{s,K,D}_{K,M,N} \left(\frac{(K - 1)M + \left\lfloor \frac{r + 1}{2} \right\rfloor}{K}\right).
\] (85)

In case \((K - 1)M + \left\lfloor \frac{r + 1}{2} \right\rfloor \leq m \leq L\) we get from (75) and Theorem 3 that
\[
d^{s,D}_{M,N} (m) = d^{s,(FC)}_{M,N} (m) \geq d^{s,K,D}_{K,M,N} (K \cdot m) \geq d^{s,K,D}_{K,M,N} (K \cdot m).
\] (86)

We also get \(d^{s,D}_{M,N} (D) = d^{s,K,D}_{K,M,N} (K \cdot D) = 0\). Since these are straight lines, we get for \(D_l < D \leq \frac{MN - 2(r_1 + 1)}{K + N - 1 - 2r_1}\)
\[
d^{s,D}_{M,N} (r) \geq d^{s,K,D}_{K,M,N} (K \cdot r).
\] (87)

Similarly to (79) it can be shown by induction for \(\frac{M + \left\lfloor \frac{r + 1}{2} \right\rfloor}{K + N - 1 - 2s} \leq D \leq \frac{MN - 2(r_1 + 1)}{K + N - 1 - 2s}\), \(s = (K - 1)M + \left\lfloor \frac{r + 1}{2} \right\rfloor + 1, \ldots, L - 1\), that
\[
d^{s,D}_{M,N} (r) \geq d^{s,K,D}_{K,M,N} (K \cdot r).
\] (88)

Hence from (84), (87) and (88) we get
\[
\max_D \min \left\{d^{s,D}_{M,N} (r), d^{s,K,D}_{K,M,N} (K \cdot r)\right\} = d^{s,(FC)}_{K,M,N} (K \cdot r) = d^{s,(IC)}_{K,M,N} (r) \geq \frac{(K - 1)M + \left\lfloor \frac{r + 1}{2} \right\rfloor}{K} \leq r \leq \frac{L}{K}.
\] (89)

The remaining open point for \(N = (K - 1)M + 1 + l, l = 0, \ldots, 2M - 3\) is the case
\[
\left\lfloor \frac{l}{2} \right\rfloor + 1 = \frac{(K - 1)M + \left\lfloor \frac{r + 1}{2} \right\rfloor}{K}.
\] (90)

First we would like to find when this equality takes place. For this we consider two cases. First let us consider \(l = 2b\). For this case (90) takes the following form
\[
K \cdot (b + 1) = (K - 1)M + b
\]
which leads to
\[
b = M - \frac{K}{K - 1}.
\]

Since \(b \geq 0\), \(M \geq 1\) and \(K \geq 2\) are integers, we get that this equality can only hold when \(K = 2\). In this case we get \(M = b + 2\) and \(N = 3(b + 1)\). Since both \(M \geq 1\) and \(N \geq 1\), we get that \(b \geq 2\). Hence by assigning \(s = b + 1\) we get (90) for \(K = 2, M = s + 1\) and \(N = 3s\), where \(s \geq 1\) is an integer. For the second case we consider \(l = 2b + 1\). In this case by assigning in (90) we get \(b = M - 1\). However we know that \(l = 2b + 1 \leq 2M - 3\), and so \(b \leq M - 2\). Hence for \(l = 2b + 1\) (90) can not take place. From (75), (90) we get
\[
\left\lfloor \frac{l}{2} \right\rfloor + 1 = \frac{N}{K + 1} = \frac{(K - 1)M + \left\lfloor \frac{r + 1}{2} \right\rfloor}{K}.
\] (91)

In addition, (90) holds only for \(l = 2b\). For this case simply by assigning \(l = 2b\) we get
\[
D^{s,(FC)}_{\frac{l}{2}} = D_l = D^{s}_{r_1}.
\] (92)

Hence, we are interested in finding \(d^{s,(IC)}_{K,M,N} (r)\) for the case \(K = 2, M = s + 1\) and \(N = 3s\), where \(s \geq 1\) is an integer. For \(D > D_l\) we get \(d^{s,D}_{s+1,3s} (r) \leq d^{s,(FC)}_{s+1,3s} (r)\). From the same arguments we get for \(\frac{N}{K + 1} \leq r \leq \frac{K}{K} = 1\) that the optimal
solution is $d^{*,(FC)}_{2,(s+1),3s}(2 \cdot r)$. Hence we get
\[
d^{*,(IC)}_{K,M,N}(r) = d^{*,(IC)}_{2,3s+1,2s}(r) = \begin{cases} 
    d^{*,(FC)}_{s+1,3s}(r) & 0 \leq r \leq \frac{N}{K+1} = s \\
    d^{*,(FC)}_{2s+1,3s}(2 \cdot r) & s \leq r \leq 3s.
\end{cases} \tag{93}
\]

So far we have shown that
\[
\max_{A \subseteq \{1, \ldots, K\}, D_{i} \neq D_{r}} \min_{D} \left\{ d^{*,D}_{M,N}(r), d^{*,K-D}_{K-M,N}(K \cdot r) \right\} = d^{*,(IC)}_{K,M,N}(r). \tag{94}
\]

Now we wish to show that this is also the solution of (8). We begin with the case for which $d^{*,(IC)}_{K,M,N}(r) = d^{*,(FC)}_{M,N}(r)$. This is the case for $N \geq (K + 1) M - 1$, and also for $N = (K - 1) M - l$, $l = 0, \ldots, 2M - 3$ when $0 \leq r \leq \frac{K}{r} + 1$. As a base line we consider the case $D_{1} = \ldots = D_{K} = D_{r}$, where $D_{r}$ is the optimal average number of dimensions per channel use per user in (94). Without loss of generality assume user $i$ has $D_{i} \neq D_{r}$. In this case based on Corollary [1] we get
\[
\min_{A \subseteq \{1, \ldots, K\}, D_{i} \neq D_{r}} \left( d^{*,\sum_{a \in A} D_{a} (|A| \cdot r) \leq d^{*,(FC)}_{M,N}(r) = \max_{D} \min \left\{ d^{*,D}_{M,N}(r), d^{*,K-D}_{K-M,N}(K \cdot r) \right\} \right). \tag{95}
\]

Hence the optimal solution must be $d^{*,(IC)}_{K,M,N}(r)$, attained for $D_{1} = \ldots = D_{K} = D_{r}$. Now consider the case where $d^{*,(IC)}_{K,M,N}(r) = d^{*,(FC)}_{M,N}(r)$, for which $N = (K - 1) M + 1 + l$, where $l = 0, \ldots, 2M - 3$ and $\frac{(K-1)M+\frac{l}{r}+\frac{1}{r}}{K} \leq r \leq \frac{L}{K}$. In this case the optimal solution in (94) for the $K$ users pulled together is attained for $K \cdot D_{r}$. Let us assume that $\sum_{i=1}^{K} D_{i} \neq K \cdot D_{r}$. In this case we get
\[
\min_{A \subseteq \{1, \ldots, K\}, D_{i} < \frac{N}{K}} \left( d^{*,\sum_{a \in A} D_{a} (|A| \cdot r) \leq d^{*,(FC)}_{M,N}(r) = \max_{D} \min \left\{ d^{*,D}_{M,N}(r), d^{*,K-D}_{K-M,N}(K \cdot r) \right\} \right). \tag{96}
\]

Hence the optimal solution must be $d^{*,(IC)}_{K,M,N}(r)$. Now let us consider the case $N < (K - 1) M + 1$. In this case the optimal solution in (94) is attained for $D_{r} = \frac{N}{K}$. Without loss of generality assume $D_{i} < \frac{N}{K}$. In this case we get from Corollary [1] that
\[
\min_{A \subseteq \{1, \ldots, K\}, D_{i} < \frac{N}{K}} \left( d^{*,\sum_{a \in A} D_{a} (|A| \cdot r) \leq M - K \cdot r \leq \max_{D} \min \left\{ d^{*,D}_{M,N}(r), d^{*,K-D}_{K-M,N}(K \cdot r) \right\} \right). \tag{97}
\]

which shows again that $d^{*,(IC)}_{K,M,N}(r)$ is the solution. Finally we consider the case where $d^{*,(IC)}_{K,M,N}(r) = d^{*}(r)$, i.e. the case where $N = (K - 1) M + 1 + l$, $l = 0, \ldots, 2M - 3$ and $\frac{(K-1)M+\frac{l}{r}+\frac{1}{r}}{K} \leq r \leq \frac{(K-1)M+\frac{l}{r}+\frac{1}{r}}{K}$. Without loss of generality assume $D_{i} \neq D_{i}$. In this case we get from Corollary [1]
\[
\min_{A \subseteq \{1, \ldots, K\}, D_{i} \neq D_{r}} \left( d^{*,\sum_{a \in A} D_{a} (|A| \cdot r) \leq d^{*}(r) = \max_{D} \min \left\{ d^{*,D}_{M,N}(r), d^{*,K-D}_{K-M,N}(K \cdot r) \right\} \right) \tag{98}
\]

which shows that $d^{*,(IC)}_{K,M,N}(r)$ is the optimal solution. This concludes the proof.

\section*{Appendix F}
\section*{Proof of Lemma 6}

For $N \geq (K + 1) M - 1$ it can be easily shown based on Lemma 2 and Corollary [1] that
\[
d^{*,(FC)}_{K,M,N}(r) = d^{*,(FC)}_{M,N}(r) = d^{*,(IC)}_{K,M,N}(r). \tag{99}
\]

For $N < (K - 1) M + 1$ we get $\frac{K}{r} = \frac{N}{K}$. In this case from (112), (113) and (114) we get
\[
d^{*,D}_{M,N}(r) < d^{*,K-D}_{K-M,N}(K \cdot r) \leq d^{*,(FC)}_{M,N}(K \cdot r) \quad 0 < D \leq \frac{N}{K} \tag{100}
\]
for $0 < r < D$. In addition since $\frac{N}{K} < \frac{M+N}{N+M-1}$ and $0 < D \leq \frac{N}{K}$ we get from Corollary [1] that
\[
d^{*,(IC)}_{K,M,N}(r) = M - K \cdot r < d^{*,(FC)}_{M,N}(r) \quad 0 < r \leq \frac{N}{K}. \tag{101}
\]

Since $d^{*,(FC)}_{K,M,N}(r)$ consists of $d^{*,(FC)}_{M,N}(r)$ and $d^{*,(FC)}_{K,M,N}(K \cdot r)$ we get from (100), (101) that
\[
d^{*,(IC)}_{K,M,N}(r) < d^{*,(FC)}_{K,M,N}(r) \quad 0 < r < \frac{N}{K}. \tag{100}
\]

For $N = (K - 1) M + 1 + l$ where $l = 0, \ldots, 2M - 3$, recall that we denoted $D_{l} = \frac{MN - 2}{N+M-1-l}$.
and also \( r_l = \frac{(K-1)M+\lfloor t+1\rfloor}{K} \). In [53], it was shown that \( D_l < \frac{MN-(\lfloor t+1\rfloor)(\lfloor t+2\rfloor)}{M+N-1-2(\lfloor t+1\rfloor)} \). Hence, from Corollary [1] we get

\[
d^*(r) = d_{M,N}^*(D_2) (r) < d_{M,N}^* (FC) (r) \quad \lceil \frac{t}{2} \rceil + 1 < r \leq \frac{L}{K}.
\]

On the other hand from [66] we get \( D_l > \frac{MN-r_l(K-r_l-1)}{K-M(N+1-2(K-r_l-1))} \). Hence based on Corollary [1] we get

\[
d^*(r) = d_{K,M,N}^*(K \cdot r) < d_{K,M,N}^* (FC) (K \cdot r) \quad 0 \leq r < \frac{(K-1)M + \lfloor t+1/2 \rfloor}{K}.
\]

Since \( d_{K,M,N}^* (FC) (r) \) consists of \( d_{M,N}^* (FC) (r) \) and \( d_{K,M,N}^* (K \cdot r) \), we get from (102), (103)

\[
d^*(r) < d_{K,M,N}^* (FC) (r) \quad \lceil \frac{t}{2} \rceil + 1 < r < \frac{(K-1)M + \lfloor t+1/2 \rfloor}{K}.
\]

The remaining open point for \( N = (K-1)M + 1 + t \), where \( l = 0, \ldots, 2M - 3 \) is the case

\[
\lceil \frac{t}{2} \rceil + 1 = \frac{(K-1)M + \lfloor t+1/2 \rfloor}{K}.
\]

In Theorem [4] it was shown (see equation [91]) appendix [3] that we get equality for \( K = 2, M = s + 1 \) and \( N = 3s \), where \( s \geq 1 \) is an integer. According to Theorem [3] for this case the optimal DMT of finite constellations equals

\[
d_{2,s+1,3s}^* (FC) (r) = \begin{cases} 
d_{2,s+1,3s}^* (FC) (r), & 0 \leq r \leq \frac{N}{K+1} = s \\
d_{2,s+1,3s}^* (2 \cdot r), & s \leq r \leq 3s.
\end{cases}
\]

Hence, from [53] we get \( d_{2,s+1,3s}^* (FC) (r) = d_{2,s+1,3s}^* (IC) (r) \). By simply assigning we get that in this case \( N < (K+1)M - 1 \). This concludes the proof.

**APPENDIX G**

**PROOF OF THEOREM [5]**

We begin by finding for \( N \geq (K+1)M - 1 \) an upper bound on the DMT of the unconstrained multiple-access channels, that equals to the optimal DMT of finite constellations \( d_{M,N}^* (\max (r_1, \ldots, r_K)) \). The proof relies on the upper bound on the optimal DMT in the symmetric case \( d_{K,M,N}^* (r) \). For \( N \geq (K+1)M - 1 \) it was shown in Lemma [6] that

\[
d_{K,M,N}^* (r) = d_{M,N}^* (r).
\]

From Theorem [2] we get that the optimal DMT is upper bounded by

\[
\max_{(D_1, \ldots, D_K) \in \Delta} \min_{A \subseteq \{1, \ldots, K\}} d_{|A|,M,N}^* (D_A) (R_A).
\]

We wish to solve (106). We solve it by finding upper and lower bounds on (106) that coincide. For the rate tuple \((r_1, \ldots, r_K)\) recall the definition \( r_{\text{max}} = \max (r_1, \ldots, r_K) \). We begin by lower bounding the optimization problem terms. Based on Lemma [2] and the fact that \( d_{M,N}^* (i \cdot r), i = 1, \ldots, K \) are straight lines as a function of \( r \) we get

\[
d_{|A|,M,N}^* \left( \sum_{a \in A} D_a r_a \right) \geq d_{|A|,M,N}^* (\sum_{a \in A} D_a (|A| \cdot r_{\text{max}})) \geq d_{M,N}^* (\sum_{a \in A} D_a (|A| \cdot r_{\text{max}})) \forall A \subseteq \{1, \ldots, K\}.
\]

Hence, we get

\[
\min_{A \subseteq \{1, \ldots, K\}} d_{|A|,M,N}^* \left( \sum_{a \in A} D_a r_a \right) \geq \min_{A \subseteq \{1, \ldots, K\}} d_{M,N}^* (r_{\text{max}}).
\]

From Corollary [1] we know that

\[
\max_{D} d_{M,N}^* (r_{\text{max}}) = d_{M,N}^* (FC) (r_{\text{max}})
\]

obtained for \( D_{\text{max}} = \frac{MN-1(r_{\text{max}}+1)}{M+N-1-2(r_{\text{max}})} \). Hence, from (108), (109) we get

\[
\max_{(D_1, \ldots, D_K) \in \Delta} \min_{A \subseteq \{1, \ldots, K\}} d_{|A|,M,N}^* (D_A) \geq \max_{D} \min_{A \subseteq \{1, \ldots, K\}} d_{M,N}^* (r_{\text{max}}) = d_{M,N}^* (FC) (r_{\text{max}}).
\]
obtained for \( D_1 = \cdots = D_K = D_{\max} \). Next we upper bound the optimization problem and show it coincides with the lower bound. Without loss of generality assume \( r_1 = r_{\max} \). In this case we get

\[
\min_{A \subseteq \{1, \ldots, K\}} d^*_{|A|} D_A \left( \sum_{a \in A} r_a \right) \leq d^*_M (r_{\max}).
\]

From (109), (111) we can write

\[
\max_{(D_1, \ldots, D_K) \in D} \min_{A \subseteq \{1, \ldots, K\}} d^*_{|A|} D_A (R_A) \leq \max_{D_i} d^*_M (r_{\max}) = d^*((FC) (r_{\max})
\]

obtained for \( D_i = D_{\max} \). Hence, from (110), (112) we get

\[
\max_{(D_1, \ldots, D_K) \in D} \min_{A \subseteq \{1, \ldots, K\}} d^*_{|A|} D_A (R_A) = d^*((FC) (r_{\max})
\]

which is the optimal DMT of finite constellations.

Now we show for \( N < (K + 1) M - 1 \) that the optimal DMT of the unconstrained multiple-access channels is suboptimal compared to the optimal DMT of finite constellations. We do that by showing that there exists a set \( B \) of multiplexing gain tuples \((r_1, \ldots, r_K)\) for which

\[
\max_{(D_1, \ldots, D_K) \in D} \min_{A \subseteq \{1, \ldots, K\}} d^*_{|A|} D_A (R_A) < d^*_{K,M,N} (r_1, \ldots, r_K) \quad \forall (r_1, \ldots, r_K) \in B,
\]

where \( d^*_{K,M,N} (r_1, \ldots, r_K) \) is the optimal DMT of finite constellations. We divide the sub-optimality proof of \( N < (K + 1) M - 1 \) to several cases. We begin with the case \( N < (K - 1) M + 1 \). For this case we show the sub-optimality by considering symmetric multiplexing gain tuples, i.e. \( r_1 = \cdots = r_K = r \). In this case the optimization problem (106) solution equals \( d^*_{K,M,N} (r) \). From Lemma 4 we get that

\[
d^*_{K,M,N} (r) < d^*_{K,M,N} (r) = d^*_{K,M,N} (r, \ldots, r) \quad 0 < r < \frac{N}{K}.
\]

Hence, in this case we have proved the sub-optimality based on the optimal DMT in the symmetric case. Now we prove the sub-optimality for the case \( N = (K - 1) M + l \), where \( l = 0, \ldots, 2M - 3 \). In Lemma 5 we have showed for \( r_1 = \cdots = r_K = r \) that

\[
d^*_{K,M,N} (r) < d^*_{K,M,N} (r) \quad \left( \frac{l}{2} + 1 < r < \frac{(K - 1) M + \lfloor \frac{l+1}{2} \rfloor}{K} \right).
\]

Hence, for the case where \( \left( \frac{l}{2} + 1 = \frac{(K - 1) M + \lfloor \frac{l+1}{2} \rfloor}{K} \right) \) this shows the sub-optimality of any IC’s DMT. Therefore, in order to complete the sub-optimality proof we are left only with the case \( \left( \frac{l}{2} + 1 = \frac{(K - 1) M + \lfloor \frac{l+1}{2} \rfloor}{K} \right) \) only when \( K = 2, M = s + 1 \) and \( N = 3 s \), where \( s \geq 1 \) is an integer. Note that in this case the upper bound on the optimal DMT of IC’s in the symmetric case equals to the optimal DMT of finite constellations. Hence, in this case we can not obtain the sub-optimality from the symmetric case and we need to find a set of multiplexing gain tuples \( B \) for which

\[
\max_{(D_1, D_2)} \min_{(r_1, r_2)} \left( d^*_{2,s+1,3s} (r_1), d^*_{2,s+1,3s} (r_1 + r_2), d^*_{2,s+1,3s} (r_2) \right) < d^*_{2,s+1,3s} (r_1, r_2) \quad (r_1, r_2) \in B.
\]

We defer the proof of (115) to appendix H. In a nutshell we are interested in finding a set of the optimal DMT of finite constellations equals to the two user optimal DMT, i.e. \( d^*_{2,s+1,3s} (r_1 + r_2) = d^*_{2,s+1,3s} (r_1, r_2) \), where the IC’s single user expressions \( d^*_{2,s+1,3s} (r_1) \) or \( d^*_{2,s+1,3s} (r_2) \) will be smaller than \( d^*_{2,s+1,3s} (r_1, r_2) \) for any \( D_1, D_2 \) for which \( d^*_{2,s+1,3s} (r_1 + r_2) = d^*_{2,s+1,3s} (r_1, r_2) \). Figure 5 shows the optimal DMT of finite constellations for the case \( K = 2, M = 3 \) and \( N = 6 \), and Figure 6 illustrates the aforementioned description of the proof method for the same setting.

APPENDIX H

FINAL PART OF THE PROOF OF THEOREM 5

In order to find the set \( B \) we first present several properties of \( d^*_{2,s+1,3s} (r) \), i.e. the optimal DMT of IC’s in the symmetric case, for this case. First note that from Theorem 4 we get

\[
d^*_{2,s+1,3s} (r) = \left\{ \begin{array}{ll}
d^*_{2,s+1,3s} (r) & 0 \leq r \leq \frac{N}{K+1} = s \\
d^*_{2,s+1,3s} (2 \cdot r) & s \leq r \leq \min \left( s + 1, \frac{3}{2} s \right) \end{array} \right.
\]

An example of \( d^*_{2,s+1,3s} (r) \) for the case \( M = 3, N = 6 \) and \( K = 2 \), i.e. \( s = 2 \), is given in Figure 5.
From simple assignment of the values of $M$, $N$ and $K$ we get that $l = 2(s - 1)$. We know from Lemma 5, Theorem 3 and (91) that
\[
d_{s+1,3s}^* \left( \frac{N}{K+1} \right) = d_{s+1,3s}^{P,*(FC)} \left( \frac{N}{K+1} \right) = d_{2s+1,3s}^{P,FC} \left( \frac{K \cdot N}{K+1} \right) = d_{2(s+1),3s}^{2D_1} \left( \frac{K \cdot N}{K+1} \right).
\]
Hence, from (92) and (116) we get
\[
d_{s+1,3s}^{*,D_1^*} (r) = d_{2(s+1),3s}^{*,2D_1^*} (2 \cdot r).
\]
Finally, from Corollary 1 we get
\[
d_{s+1,3s}^{*,D_1^*} (r) = d_{s+1,3s}^{*,2D_1^*} (2 \cdot r) = d_{s+1,3s}^{*,(FC)} (r) \quad s - 1 \leq r \leq \frac{N}{K+1} = s
\]
and
\[
d_{s+1,3s}^{*,D_1^*} (r) = d_{s+1,3s}^{*,2D_1^*} (2 \cdot r) = d_{s+1,3s}^{*,(FC)} (2 \cdot r) \quad s \leq r \leq s + \frac{1}{2}
\]
i.e., The last line of $d_{s+1,3s}^{*,(FC)} (r)$ before $\frac{N}{K+1} = s$, and the first line of $d_{2(s+1),3s}^{*,(FC)} (2r)$ after $s$ are equal. To sum up, for $\left\lfloor \frac{1}{2} \right\rfloor + 1 = \frac{(K-1)M+1}{K}$ the optimal DMT of IC’s in the symmetric case is upper bounded by a piecewise linear function as expected, and we have found the straight line coincide with it for $s - 1 \leq r \leq s + \frac{1}{2}$. We are interested in finding a set of multiplexing gain tuples $B$, for which (115) is fulfilled. In a nutshell we are interested in finding a set such that the optimal DMT of finite constellations equals to the two user optimal DMT, where IC’s single user expressions will be smaller than the optimal DMT of finite constellations for any $D_1, D_2$ for which the IC’s two users expression equals to the optimal DMT of finite constellations. Figure 5 illustrates the aforementioned description of the proof method.

From (118) and Corollary 1 we know that
\[
d_{s+1,3s}^{*,(FC)} (r) = d_{s+1,3s}^{*,D_1^*} + 1 (r) \quad s \leq r \leq s + 1.
\]
Hence, for certain $s < r_0 < s + \frac{1}{2}$, we are interested in the set for which $r_1 = r_0 + \epsilon$, $r_2 = r_0 - \epsilon$ such that $s < r_0 + \epsilon < s + \frac{1}{2}$ and also
\[
d_{s+1,3s}^{*,D_1^*} (r_0) = d_{s+1,3s}^{*,(FC)} (2r_0) < d_{s+1,3s}^{*,(FC)} (r_0 + \epsilon) = d_{s+1,3s}^{*,D_1^*} + 1 (r_0 + \epsilon)
\]
where the first equality results from (119). Note that (121) can be fulfilled because from Corollary 1 we get $d_{s+1,3s}^{*,D_1^*} (r) < d_{s+1,3s}^{*,D_1^*} + 1 (r)$ for $r > s$. In order to translate this condition to $\epsilon$ we write the following inequality
\[
d_{s+1,3s}^{*,D_1^*} + 1 (r_0 + \epsilon) = MN - \left( \left\lfloor \frac{1}{2} \right\rfloor + 1 \right) \cdot \left( \left\lfloor \frac{1}{2} \right\rfloor + 2 \right) - \left( N + M - 1 - 2 \cdot \left\lfloor \frac{1}{2} \right\rfloor + 1 \right) (r_0 + \epsilon) >
\]
\[
MN - \left( \left\lfloor \frac{1}{2} \right\rfloor + 1 \right) - \left( N + M - 1 - 2 \cdot \left\lfloor \frac{1}{2} \right\rfloor \right) r_0 = d_{s+1,3s}^{*,D_1^*} (r_0)
\]
When assigning $K = 2$, $M = s + 1$ and $N = 3s$ we get
\[
\epsilon < \frac{r_0}{s} - 1.
\]
Hence, the set of multiplexing gain tuples we are considering is
\[
B_{r_0} = \left\{ r_1, r_2 \mid r_1 = r_0 + \epsilon, r_2 = r_0 - \epsilon, 0 < \epsilon < \min \left( r_0 + \frac{r_0}{s} - 1, s + \frac{1}{2} \right) - r_0 \right\}
\]
where $s < r_0 < s + \frac{1}{2}$ is a parameter determining the set. From (92) Lemma 7] we get that the optimal DMT of finite constellations equals
\[
d_{s+1,3s}^{*,(FC)} (r_1, r_2) = \min \left( d_{s+1,3s}^{*,(FC)} (r_1), d_{s+1,3s}^{*,(FC)} (r_2), d_{s+1,3s}^{*,(FC)} (r_1 + r_2) \right).
\]
Considering $(r_1, r_2) \in B_{r_0}$, based on (121), (124) and the fact that $d_{s+1,3s}^{*,(FC)} (r)$ is a straight line, we get
\[
d_{s+1,3s}^{*,(FC)} (r_1, r_2) = \min \left( d_{s+1,3s}^{*,(FC)} (r_1 + r_2) \right) \quad 0 < \epsilon < \min \left( r_0 + \frac{r_0}{s} - 1, s + \frac{1}{2} \right) - r_0.
\]
Hence, in order to prove (115) we need to show for certain $0 < r_0 < s + \frac{1}{2}$ that
\[
\max_{D_1, D_2} \min \left( d_{s+1,3s}^{*} (r_0 + \epsilon), d_{2(s+1),3s}^{*} (2r_0), d_{s+1,3s}^{*} (r_0 - \epsilon) \right) < d_{2(s+1),3s}^{*} (2r_0)
\]  
(127)
where $0 < \epsilon < \min \left( r_0 + \frac{r_0}{s}, 1, s + \frac{1}{2} \right) - r_0$. We begin the proof by taking the symmetric case, i.e. $D_1 = D_2$, as a baseline. We assign $D_1 = D_2 = D_0^* = D^*_{\frac{1}{2}}$. From (119) we get that $d_{2(s+1),3s}^{*} (2r_0) = d_{s+1,3s}^{*} (r_0) = d_{2(s+1),3s}^{*} (2r_0)$. Hence for the symmetric case we get
\[
\min \left( d_{s+1,3s}^{*} (r_0 + \epsilon), d_{s+1,3s}^{*} (r_0 - \epsilon) \right) = d_{s+1,3s}^{*} (r_0 + \epsilon) < d_{2(s+1),3s}^{*} (2r_0).
\]  
(128)
Since $s < r_0 < s + \frac{1}{2}$ is not an anchor point, we get from Corollary 11 and (119) that $d_{2(s+1),3s}^{*} (2r_0) = d_{2(s+1),3s}^{*} (2r_0)$ if and only if $D_1 + D_2 = 2D_0^* = 2D^*_{\frac{1}{2}}$. Hence, in order for $d_{2(s+1),3s}^{*} (2r_0)$ (127) to attain the optimal DMT of finite constellations, we must choose
\[
D_1 + D_2 = 2D^*_{\frac{1}{2}}.
\]  
(129)
From (121), (128) we know that
\[
d_{s+1,3s}^{*} (r_0 + \epsilon) < d_{2(s+1),3s}^{*} (2r_0) < d_{s+1,3s}^{*} (r_0 + \epsilon).
\]  
(130)
Hence from Corollary 11 we can deduce that there must exist $D' = D_{\frac{1}{2}} + \epsilon'$, where $0 < \epsilon' < D_{\frac{1}{2}} + 1 - D_{\frac{1}{2}}$, such that
\[
d_{s+1,3s}^{*} (r_0 + \epsilon) = d_{2(s+1),3s}^{*} (2r_0).
\]  
(131)
We divide the assignment of $D_1$ into several cases. For the case $0 < D_1 < D^*$ we get from Corollary 11 and the fact that $s < r_0 + \epsilon < s + \frac{1}{2}$ is not an anchor point
\[
d_{s+1,3s}^{*} (r_0 + \epsilon) < d_{s+1,3s}^{*} (r_0 + \epsilon) = d_{2(s+1),3s}^{*} (2r_0).
\]  
(132)
Hence in this range the optimal DMT of finite constellations is not obtained. For the case $D_1 = D' = D_{\frac{1}{2}} + \epsilon'$, we have shown (131) that $d_{s+1,3s}^{*} (r_0 + \epsilon)$ equals to the optimal DMT of finite constellations. According to (129) we need to assign
\[
D_2 = D'' = D_{\frac{1}{2}} - \epsilon' \quad \text{in order to get}
\]
\[
d_{s+1,3s}^{*} (r_0 + \epsilon) = d_{2(s+1),3s}^{*} (2r_0). \]
So far we have shown that the first two terms in the left side of (127) can attain the optimal DMT of finite constellations for $D_1 = D'$. We are left with the third term that equals to the straight line $d_{s+1,3s}^{*} (r)$. We consider two cases. For the first case we assume $D'' \leq r_0 - \epsilon$. In this case we get
\[
d_{s+1,3s}^{*} (r_0 - \epsilon) = 0 < d_{2(s+1),3s}^{*} (2r_0).
\]  
(133)
For the second case we assume $D'' > r_0 - \epsilon$. From symmetry considerations it can be easily shown that the straight line $d' (r)$ that fulfills $d' (s) = d_{s+1,3s}^{*} (s) = d_{s+1,3s}^{*} (s)$ and $d' \left( D'' \right) = 0$, also fulfills
\[
d' (r_0 - \epsilon) = d_{s+1,3s}^{*} (r_0 + \epsilon) = d_{2(s+1),3s}^{*} (2r_0).
\]  
(134)
Since $D'' < D_{\frac{1}{2}}$, according to Corollary 11 we get
\[
d_{s+1,3s}^{*} (s) < d_{s+1,3s}^{*} (s) = d' (s).
\]  
(135)
Since $d_{s+1,3s}^{*} (D'') = d' \left( D'' \right) = 0$ and these are straight lines we get
\[
d_{s+1,3s}^{*} (r) < d' (r) \quad 0 < r < D''
\]  
(136)
and so according to (135)
\[
d_{s+1,3s}^{*} (r_0 - \epsilon) < d' (r_0 - \epsilon) = d_{2(s+1),3s}^{*} (2r_0).
\]  
(137)
Thus, the third term in the left side of (127) $d_{s+1,3s}^{*} (r_0 - \epsilon)$ is smaller than the optimal DMT of finite constellations. Finally,
we consider the case $D_1 > D'$. For this case we get $D_2 < D'' < D'_1$, which according to Corollary 1 leads to

$$d_{s+1,3s}^e(r_0 - \epsilon) < d_{s+1,3s}^e(r_0 + \epsilon) < d_{2(s+1),3s}^e(2r_0).$$  \hfill (138)

From (132), (133), (137) and (138) we have proved that

$$\max_{(D_1, D_2)} \min d_{s+1,3s}^e(r_0 + \epsilon), d_{2(s+1),3s}^e(2r_0), d_{s+1,3s}^e(r_0 - \epsilon) < d_{2(s+1),3s}^e(2r_0).$$  \hfill (139)

This concludes the proof.

### APPENDIX I

#### PROOF OF THEOREM 3

We base our proof on the techniques developed by Poltyrev [12] for the AWGN channel and extended in [8] to colored channels in the point-to-point case. We begin by partitioning the error event into several disjoint events of errors for subsets of the users. We relate each of these error events to the point-to-point channel of the relevant users pulled together. Then we use the bounds derived in [8] to upper bound each of the error events probabilities.

When the ML decoder makes an error it means that the decoded word is different from the transmitted signal for at least one of the users. Hence, we can break the error probability into the following sum of disjoint events

$$\overline{P_e}(H_{\text{eff}}^{(1),K}, \rho) = \sum_{s \subseteq \{1, \ldots, K\}} \overline{P_e}(H_{\text{eff}}^{(1),s}, \rho)$$  \hfill (140)

where $\overline{P_e}(H_{\text{eff}}^{(1),s}, \rho)$ is the probability of error to words that induce error on the users in $s$. Note that the event of error to users in $s$ depends only on $H_{\text{eff}}^{(1),s}$ and not on $H_{\text{eff}}^{(1),\{1, \ldots, K\}}$. We wish to upper bound $\overline{P_e}(H_{\text{eff}}^{(1),s}, \rho)$ for any $s \subseteq \{1, \ldots, K\}$.

Based on (12) we have the following upper bound on the error probability of the joint ML decoder when transmitting $\mathbf{x} \in S_{K, D_1, T_1}$

$$P_e(\mathbf{x}') \leq P_r(\|\hat{\mathbf{n}}_{ex}\| \geq R) + \sum_{l \in \text{Ball}(\mathbf{x}', 2R) \cap S_{K, D_1, T_1}, l \neq \mathbf{x}'} P_r(\|\mathbf{l} - \mathbf{x}' - \hat{\mathbf{n}}_{ex}\| < \|\hat{\mathbf{n}}_{ex}\|)$$  \hfill (141)

where $S_{K, D_1, T_1}$ is the $K \times D_1 \times T_1$-complex dimensional effective IC of the $K$ users. Ball$(\mathbf{x}', 2R)$ is a $K \times D_1 \times T_1$-complex dimensional ball of radius $2R$ centered around $\mathbf{x}'$, and $\mathbf{n}_{ex}$ is the effective noise in the $K \times D_1 \times T_1$-complex dimensional hyperplane where the effective IC resides. Instead of calculating (141), we focus on upper bounding the probability of decoding words that lead to an error only for the users in $s \subseteq \{1, \ldots, K\}$ (140). This will lead to an upper bound on the error probability. Hence, we begin by considering the error probability of $\mathbf{x}'$ to words that are different from $\mathbf{x}'$ only in the entries of the users in $s$. Based on our ensemble, this is the error event of users in $s$ almost surely (with probability 1). This error event is equivalent to the error event of a word $\mathbf{z}'$, which is a vector of length $|s| \times D_1 \times T_1$ that resides within an $|s| \times D_1 \times T_1$-complex dimensional IC $S_{|s|, D_1, T_1}$, when $\mathbf{z}''$ equals to $\mathbf{z}'$ in the entries of the users in $s$, and the other words in $S_{|s|, D_1, T_1}$ are equal, in the entries of the users in $s$, to words in $S_{K, D_1, T_1}$, that lead to an error for the users in $s$. Hence, we wish to upper bound the error probability of $\mathbf{z}'' \in S_{|s|, D_1, T_1}$. Based on the expressions in (141) we get that this upper bound can be written as

$$P_r(\|\hat{\mathbf{n}}_{ex}\| \geq R') + \sum_{l \in \text{Ball}(\mathbf{z}'', 2R') \cap S_{|s|, D_1, T_1}, l \neq \mathbf{z}'',} P_r(\|\mathbf{l} - \mathbf{z}'' - \hat{\mathbf{n}}_{ex}\| < \|\hat{\mathbf{n}}_{ex}\|)$$  \hfill (142)

where Ball$(\mathbf{z}'', 2R')$ is a $|s| \times D_1 \times T_1$-complex dimensional ball of radius $2R'$ centered around $\mathbf{z}''$, and $\mathbf{n}_{ex}$ is the effective noise in the $|s| \times D_1 \times T_1$-complex dimensional hyperplane $S_{|s|, D_1, T_1}$ resides.

Next we upper bound the average decoding error probability of an ensemble of finite constellations, which later will extend to ensemble of IC’s. Note that the upper bounds on the error probability of IC’s in (140), (141) also apply to finite constellations. Assume user $j$ code-book contains $\gamma_{tr}^{(j)|B_{2D_1T_1}}$ words, where each word is drawn independently and uniformly within $\text{cube}_{D_1T_1}(\mathbf{b})$, $j = 1, \ldots, K$. Recall from (11) that $\gamma_{tr}^{(j)} = \rho^{T_1}$. The $K$ users constitute together an ensemble of $\prod_{j=1}^{K} \gamma_{tr}^{(j)|B_{2D_1T_1}}$ words, where a word in the ensemble is sampled from a uniform distribution in $\text{cube}_{|s|, D_1, T_1}(\mathbf{b})$ (not all words are drawn independently). In fact any subset of the users $s \subseteq \{1, \ldots, K\}$ corresponds to an ensemble of $\prod_{i \in s} \gamma_{tr}^{(i)|B_{2D_1T_1}}$ words, where again a word in the ensemble is sampled from a uniform distribution, this time in $\text{cube}_{|s|, D_1, T_1}(\mathbf{b})$. Hence, the number of codewords that are different in the entries of the users in $s$ is upper bounded by $\prod_{i \in s} \gamma_{tr}^{(i)|B_{2D_1T_1}}$. These words are in fact drawn independently in the entries of the users in $s$. Based on these arguments and since the ML decoder decides on the word with minimal Euclidean distance from the observation, we get for each word in the ensemble that the probability of error for users in $s \subseteq \{1, \ldots, K\}$ is upper bounded by the average decoding error probability of an ensemble consisting of $\prod_{i \in s} \gamma_{tr}^{(i)|B_{2D_1T_1}}$ words drawn independently and uniformly within $\text{cube}_{|s|, D_1, T_1}(\mathbf{b})$, with effective channel $H_{\text{eff}}^{(1),s}$. In [8] Theorem 3 an upper bound on the average decoding error probability of this ensemble was derived. By choosing for any
\[ R^2_s = R_{\text{eff}}^2 = \frac{2|s| \cdot D_t \cdot T_t}{2 \pi c} \rho - T_t |s| D_t - \sum_{i=1}^{s} D_i \cdot T_i \frac{s(s)}{|s| \cdot D_t \cdot T_t} \]

to get for the ensemble the following upper bound on the probability of error for users in \( s \)

\[ \Pr_c^{\text{FC}}(s, \eta(s)) \leq D'(|s| \cdot D_t \cdot T_t) \rho - T_t |s| D_t - \sum_{i=1}^{s} D_i \cdot T_i + \sum_{i=1}^{s} D_i \cdot T_i \eta_i(s) \quad \forall s \subseteq \{1, \ldots, K\} \]  

(143)

where \( D'(|s| \cdot D_t \cdot T_t) \geq 1 \) and \( \eta_i(s) \geq 0, i = 1, \ldots, |s| \cdot D_t \cdot T_t. \)

So far we have upper bounded the probability of error of users in \( s, \) in an ensemble of finite constellations, for any \( s \subseteq \{1, \ldots, K\}. \) Now we extend this ensemble of finite constellations into an ensemble of IC’s with density \( \gamma_{trr}^{(j)} \) for user \( j, \) where \( j = 1, \ldots, K. \) We show that extending the ensemble of finite constellations to ensemble of IC’s does not change the upper bound on the error probability. Let us consider for user \( j \) a certain finite constellation from the ensemble \( C_0^j(\rho, b) \subseteq \text{cube}_{D_t \cdot T_t}(b). \) In accordance, for the ensemble of users relates to \( s \) let us denote a certain finite constellation from the effective ensemble by \( C_0^s(\rho, b) \subseteq \text{cube}_{|s| \cdot D_t \cdot T_t}(b). \) We extend each finite constellation into IC by extending each user finite constellation in the following manner

\[ IC^j(\rho, D_t T_t) = C_0^j(\rho, b) + (b + b') \cdot \mathbb{Z}^{2 |s| \cdot D_t \cdot T_t} \]  

(144)

where without loss of generality \( \rho = 1 \) we assumed that \( \text{cube}_{D_t \cdot T_t}(b) \subseteq \mathbb{C}^{D_t T_t}. \) Therefore for the users in \( s \subseteq \{1, \ldots, K\} \) we get an effective IC

\[ IC^s(\rho, |s| \cdot D_t \cdot T_t) = C_0^s(\rho, b) + (b + b') H_{\text{eff}}^{(l),s} \cdot \mathbb{Z}^{2 |s| \cdot D_t \cdot T_t} \]  

(145)

In the receiver we get

\[ IC^s(\rho, |s| \cdot D_t \cdot T_t, H_{\text{eff}}^{(l),s}) = H_{\text{eff}}^{(l),s} \cdot C_0(\rho, b) + (b + b') H_{\text{eff}}^{(l),s} \cdot \mathbb{Z}^{2 |s| \cdot D_t \cdot T_t} \]  

(146)

By extending each finite constellation in the ensemble into an IC according to the method presented in (145), (146) we get a new ensemble of IC’s. We would like to set \( b \) and \( b' \) to be large enough such that the ensemble average decoding error probability has the same upper bound as in (143), and the users densities are equal to \( \gamma_{trr}^{(j)} \) up to a coefficient, where \( j = 1, \ldots, K. \) First we like to set a value for \( b' \). For a word within the set \( \{ H_{\text{eff}}^{(l),s} \cdot C_0(\rho, b) \} \), increasing \( b' \) decreases the error probability inflicted by the codewords outside the set \( \{ H_{\text{eff}}^{(l),s} \cdot C_0(\rho, b) \} \), for any \( s \subseteq \{1, \ldots, K\} \). In [8, Theorem 3] we have shown that for any \( \eta_i(s) \geq 0, \) by choosing \( b' = \sqrt{\frac{|s| \cdot D_t \cdot T_t}{\pi c}} \rho^{- T_t |s| D_t - \sum_{i=1}^{s} D_i \cdot T_i} + \epsilon, \) where \( \epsilon > 0, \) we get for \( \rho \geq 1 \)

\[ \Pr_c(H_{\text{eff}}^{(l),s}, \rho) = E_{C_0}(P_{e_c}^{IC}(H_{\text{eff}}^{(l),s} \cdot C_0)) \leq D(|s| \cdot D_t \cdot T_t) \rho - T_t |s| D_t - \sum_{i=1}^{s} D_i \cdot T_i + \sum_{i=1}^{s} D_i \cdot T_i \eta_i(s) \]  

(147)

where \( E_{C_0}(P_{e_c}^{IC}(H_{\text{eff}}^{(l),s} \cdot C_0)) \) is the average decoding error probability of the ensemble of IC’s defined in (146), and \( D(|s| \cdot D_t \cdot T_t) \geq D'(|s| \cdot D_t \cdot T_t). \) Hence, choosing \( b' \) to be the maximal value between \( \sqrt{\frac{|s| \cdot D_t \cdot T_t}{\pi c}} \rho^{- T_t |s| D_t - \sum_{i=1}^{s} D_i \cdot T_i} + \epsilon, \) where \( s \subseteq \{1, \ldots, K\} \) will enable to satisfy (147) for any \( s. \)

Next, we set the value of \( b \) to be large enough such that for each user, each IC density from the ensemble in (146), \( \gamma_{trc}^{(j)}, \) equals \( \gamma_{trr}^{(j)} \) up to a factor of 2, where \( j = 1, \ldots, K. \) By choosing \( b = b' \cdot \rho' \) we get

\[ \gamma_{trc}^{(j)} = \gamma_{trr}^{(j)} \cdot \left( \frac{b}{b + b'} \right)^{2 D_t T_t} = \gamma_{trr}^{(j)} \cdot \frac{1}{1 + \rho'^{\epsilon}}. \]

Hence, for \( \rho \geq 1 \) we get

\[ \frac{1}{2} \gamma_{trr}^{(j)} \leq \gamma_{trc}^{(j)} \leq \gamma_{trr}^{(j)}. \]  

(148)

As a result we also get

\[ \mu_{trc}^{(j)} \leq \frac{\gamma_{trc}^{(j)} - \gamma_{trr}^{(j)}}{\gamma_{trr}^{(j)}} \leq 2 \mu_{trr}^{(j)}. \]

(149)

Hence, from (146) and (147) we get that

\[ \Pr_c(H_{\text{eff}}^{(l),K}, \rho) \leq \sum_{s \subseteq \{1, \ldots, K\}} D(|s| \cdot D_t \cdot T_t) \rho^{- T_t |s| D_t - \sum_{i=1}^{s} D_i \cdot T_i} \cdot |H_{\text{eff}}^{(l),s} \cdot H_{\text{eff}}^{(l),s}|^{-1} \]

and from (148) we get that user \( j \) has multiplexing gain \( r_j \) as required, where \( j = 1, \ldots, K. \) This concludes the proof.

\(^2\) In case \( \text{cube}_{D_t \cdot T_t}(b) \) is a rotated cube within \( \mathbb{C}^{M \cdot T_t} \), then the replication is done according the corresponding \( M \cdot T_t \times D_t \cdot T_t \) matrix with orthonormal columns.
**APPENDIX J**

**PROOF OF LEMMA 3**

$H_{\text{eff}}^{(t),|s|}$ is a block diagonal matrix. Hence the determinant of $|H_{\text{eff}}^{(t),|s|}^\dagger \cdot H_{\text{eff}}^{(t),|s|}|$ can be expressed as

$$|H_{\text{eff}}^{(t),|s|}^\dagger \cdot H_{\text{eff}}^{(t),|s|}| = \prod_{i=1}^{T_t} |\tilde{H}_i^\dagger \cdot \tilde{H}_i|.$$  \hspace{1cm} (150)

Assume $\tilde{H}_i = (\tilde{h}_{1i}, \ldots, \tilde{h}_{mi})$, i.e. $\tilde{H}_i$ has $m$ columns. In this case we can state that the determinant

$$|\tilde{H}_i^\dagger \cdot \tilde{H}_i| = \|\tilde{h}_i\|^2 \|\tilde{h}_{i+1}\|^2 \cdots \|\tilde{h}_{m-1,m}\|^2.$$  

Note that $\tilde{H}_i$ has more rows than columns. The columns of $\tilde{H}_i$ are subset of the columns of the channel matrix $H$. Hence, in order to quantify the contribution of a certain column of $H$, $\tilde{h}_j$, $j = 1, \ldots, K \cdot M$, to the determinant we need to consider the blocks where it occurs. We know that the contribution of $\tilde{h}_j$ to these determinants can be quantified by taking into account the columns to its left in each block, i.e. by taking into account $\{\tilde{h}_1, \ldots, \tilde{h}_{j-1}\}$.

Based on (23) and (24) we can quantify the contribution of $\tilde{h}_j$ to $|H_{\text{eff}}^{(t),|s|}^\dagger \cdot H_{\text{eff}}^{(t),|s|}|$ by

$$\|\tilde{h}_j\|^{2b_j(0)} \prod_{k=1}^{j-1} \|\tilde{h}_{j-k}\|^{2b_j(k)} = \rho - \sum_{k=0}^{j-1} b_j(k) \cdot \min_{x \in \{k+1, \ldots, N\}} \xi_x.$$ 

where $b_j(k)$ is the number of occurrences of $\tilde{h}_j$ in the blocks of $H_{\text{eff}}^{(t),|s|}$, with only $\{\tilde{h}_{j-1}, \ldots, \tilde{h}_{j-k}\}$ to its left, $b_j(0)$ is the number of occurrences of $\tilde{h}_j$ with no columns to its left. Hence, the determinant is obtained by multiplying the contribution of each column in $H_{\text{eff}}^{(t),|s|}$

$$|H_{\text{eff}}^{(t),|s|}^\dagger \cdot H_{\text{eff}}^{(t),|s|}| = \prod_{j=1}^{[s]M} \|\tilde{h}_j\|^{2b_j(0)} \prod_{k=1}^{j-1} \|\tilde{h}_{j-k}\|^{2b_j(k)} = \rho - \sum_{k=0}^{j-1} b_j(k) \cdot \min_{x \in \{k+1, \ldots, N\}} \xi_x.$$  \hspace{1cm} (152)

Now we lower bound the determinant (152) by lower bounding the contribution of each column. Let us consider column $\tilde{h}_{a,M+b}$, $a = 0, \ldots, [s]M-1$, $b = 1, \ldots, M$. From Lemma 7 we know that $\tilde{h}_{a,M+b}$ occurs $N - M + 1$ times with $\{\tilde{h}_1, \ldots, \tilde{h}_{a,M+b}\}$ to its left, i.e. $b_{a,M+b}(a \cdot M + b - 1) = N - M + 1$. In addition, $\tilde{h}_{a,M+b}$ occurs in $\tilde{H}_{N-M+2v+1}$, $v = 1, \ldots, \min(M - l - 1, b - 1)$, with

$$\{\tilde{h}_1, \ldots, \tilde{h}_{a,M+b-1}\} \setminus \left\{ \bigcup_{z=0}^{a} \tilde{h}_{2z,M+1} \cdots \tilde{h}_{2z,M+v} \right\}.$$  \hspace{1cm} (153)

to its left, i.e. when $v$ is increased by one the number of columns to its left reduces by $a + 1$. Finally, $\tilde{h}_{a,M+b}$ occurs in $\tilde{H}_{N-M+2v}$, $v = 1, \ldots, \min(M - l - 1, M - b)$, with

$$\{\tilde{h}_1, \ldots, \tilde{h}_{a,M+b-1}\} \setminus \left\{ \bigcup_{z=1}^{a} \tilde{h}_{2z-M-v+1} \cdots \tilde{h}_{2z-M} \right\}.$$  \hspace{1cm} (154)

to its left (for $a = 0$ it occurs with $\{\tilde{h}_1, \ldots, \tilde{h}_{a-1}\}$ to its left), i.e. when $v$ is increased by one the number of columns to its left reduces by $a$. We wish to quantify the change in the determinant when reducing columns, and relate it to the PDF in (22). In order to analyze the performance we would like the set of columns in (153) to be a subset of the set of columns in (154), which is not the case. Hence, we assume a columns reduction that gives a lower bound on the determinant induced by the reduction in (153) and (154). We assume for $\tilde{H}_{N-M+2v}$, $v = 1, \ldots, \min(M - l - 1, M - b)$ that $\tilde{h}_{a,M+b}$ occurs with $\{\tilde{h}_1, \ldots, \tilde{h}_{a,M+b-1}\}$ to its left instead of (154). In this case, by adding columns to (154) we get a lower bound on the contribution of $\tilde{h}_{a,M+b}$ to the determinant in each of its occurrences, that equals to

$$\rho - \min_{x \in \{a,M+b,v, \ldots, N\}} \xi_x.$$  \hspace{1cm} (155)

for any $v = 1, \ldots, \min(M - l - 1, M - b)$. On the other hand for (153) we assume that only the left most column is reduced when increasing $v$, instead of the $a + 1$ columns. This leads to lower bound to the contribution of (153) to the determinant that equals to

$$\rho - \min_{x \in \{a,M+b-v, \ldots, N\}} \xi_x.$$  \hspace{1cm} (156)

Hence, we get that the set of columns corresponding to (156) is a subset of the set of columns corresponding to (155). Thus,
from (155), (156) we get the following lower bound on the determinant

\[ |H_{\text{eff}}^{(l)}|_{s^t \cdot H_{\text{eff}}^{(l)},|s|} \geq \prod_{a=0}^{s-1} \prod_{b=1}^{M} \rho^{-(N-M+1+\min(M-l-1,M-b)) \min_{z \in \{aM+b, \ldots, M\}} c_{z,aM+b}} \]

\[ \cdot \prod_{b'=2}^{M} \rho^{\min_{z \in \{aM+b', \ldots, N\}} c_{z,aM+b'}}. \]  \hspace{1cm} (157)

**APPENDIX K**

**PROOF OF THEOREM**

First let us denote \( l = |r_{\text{max}}| \). Recall from Theorem 6 that the upper bound on the error probability applies when \( \eta_i(s) \geq 0 \), for every \( i = 0, \ldots, |s| \cdot D_l \cdot T_l \) and for any \( s \subseteq \{1, \ldots, K\} \). In our analysis we assume that \( \xi_i,j \geq 0 \) for \( i = 1, \ldots, N \), \( j = 1, \ldots, K \cdot M \). We wish to show that it leads to \( \eta_i(s) \geq 0 \), i.e., we can use the upper bound on the error probability. We know that \( H_{\text{eff}}^{(l)} \) is a block diagonal matrix, where the set of columns of each block is a subset of \( \{h_1, \ldots, h_{K \cdot M}\} \). Let us denote the set of indices of the columns of \( H \) that take place in \( H_{\text{eff}}^{(l)} \) by \( a(s) \). In this case we get from trace considerations

\[ \sum_{i=1}^{N} \sum_{j \in a(s)} \rho^{\xi_{i,j}} = \sum_{i=1}^{N} \rho^{-\eta_i(s)} \forall s \subseteq \{1, \ldots, K\} \]

which leads to

\[ \sum_{i=1}^{N} \sum_{j \in a(s)} \rho^{\xi_{i,j}} = \sum_{i=1}^{N} \rho^{-\eta_i(s)} \forall s \subseteq \{1, \ldots, K\}. \]  \hspace{1cm} (158)

From (158) we get that \( \xi_{i,j} \geq 0 \) for \( i = 1, \ldots, N \), \( j = 1, \ldots, K \cdot M \) if and only if \( \eta_i(s) \geq 0 \) for any \( s \subseteq \{1, \ldots, K\} \) and \( i = 1, \ldots, \|s\| \cdot D_l \cdot T_l \). Therefore, we can use the upper bound from Theorem 6.

The upper bound on the error probability consists of the sum of \( P_e(\eta | s), \rho \) for all \( s \subseteq \{1, \ldots, K\} \). We wish to show that the DMT of each of the terms is lower bounded by \( d_{\text{max}}^{(s)} \left( r_{\text{max}} \right) \). First note that \( \forall s \subseteq \{1, \ldots, K\} \) we can write

\[ \overline{P}_e(\eta | s), \rho \leq \min \left( 1, D \left( |s| \cdot D_l \cdot T_l \right) \rho^{-|s| \cdot (N-M-1+2l) \cdot (|s| \cdot D_l \cdot T_l - r_{\text{max}})} \cdot |H_{\text{eff}}^{(l)}|_{s}, \rho \right)^{-1} \]

\[ \leq \min \left( 1, D \left( |s| \cdot D_l \cdot T_l \right) \rho^{-|s| \cdot (N-M-1+2l) \cdot (|s| \cdot D_l \cdot T_l - r_{\text{max}})} \cdot |H_{\text{eff}}^{(l)}|_{s}, \rho \right)^{-1} \]  \hspace{1cm} (159)

where the inequality comes from the fact that assuming all users transmit at the maximal multiplexing gain increases the error probability. By assigning \( D_l = \frac{MN-l(l+1)}{N-M-1-2l} \) and \( T_l = N + M - 1 - 2 \cdot l \) we get

\[ \overline{P}_e(\eta | s), \rho \leq \min \left( 1, D \left( |s| \cdot D_l \cdot T_l \right) \rho^{-|s| \cdot (MN-l(l+1)-(N+M-1-2l) \cdot r_{\text{max}})} \cdot |H_{\text{eff}}^{(l)}|_{s}, \rho \right)^{-1} \]  \hspace{1cm} (160)

From (18) we know that \( E_H \left( \overline{P}_e(\eta | s), \rho \right) = E_H \left( \overline{P}_e(\eta \left( 1, \ldots, |s| \right), \rho \right), \) i.e., the term corresponding to the first \( |s| \) users. Hence, for all terms with the same \( |s| \) we can consider

\[ \overline{P}_e(\eta \left( 1, \ldots, |s| \right), \rho) \leq \min \left( 1, D \left( |s| \cdot D_l \cdot T_l \right) \rho^{-|s| \cdot (MN-l(l+1)-(N+M-1-2l) \cdot r_{\text{max}})} \cdot |H_{\text{eff}}^{(l)}|_{s}, \rho \right)^{-1} \]  \hspace{1cm} (161)

Based on (157) let us define

\[ A(a \cdot M + b, l) = (N - b + 1) \min_{z \in \{aM+b, \ldots, N\}} c_{z,aM+b} \]

for \( b = 1, a = 0, \ldots, |s| - 1 \), and

\[ A(a \cdot M + b, l) = (N - b + 1) \min_{z \in \{aM+b, \ldots, N\}} c_{z,aM+b} + \sum_{i=1}^{\min(M-l-1,b-1)} \min_{z \in \{aM+b+i, \ldots, N\}} c_{z,aM+b} \]

for \( b = 2, \ldots, M \) and \( a = 0, \ldots, |s| - 1 \). From the bounds in (155), (156), (157) and also since \( N-M+1+\min(M-l-1,M-b) \leq N-b+1 \), we get that \( \rho^{-A(a \cdot M + b, l)} \) gives a lower bound on the contribution of \( b, a \cdot M + b, \ldots, N \) to the determinant. As a result we get the following upper bound

\[ |H_{\text{eff}}^{(l)}|_{s}, \rho \leq \prod_{a=0}^{s-1} \prod_{b=1}^{M} \rho^{A(a \cdot M + b, l)}, \]  \hspace{1cm} (164)
By assigning in the bound from (161) we get
\[ Pe(p^{(1,\ldots,s)}), \rho \leq \rho^{-\left(\|y\|_{\infty} - (N + M - 1 - 2l)r_{DMT}\right) - \sum_{i=1}^{s}|A(i,l)|} \] (165)
where \( (x)^+ \) equals \( x \) for \( x \geq 0 \) and 0 else.

Based on (165) the average over the channel realizations can be upper bounded by
\[ E_H \left( Pe(p^{(s)}), \rho \right) = E_H \left( Pe(p^{(1,\ldots,s)}), \rho \right) \]
\[ \leq \int_{\xi_{i,j} \geq 0} \rho^{-\left(\|y\|_{\infty} - (N + M - 1 - 2l)r_{DMT}\right) - \sum_{i=1}^{s}|A(i,l)|} \cdot \sum_{i=1}^{N} \sum_{j=1}^{K_M} \xi_{i,j} d\xi_{i,j} \] (166)
where \( \xi_{i,j} \geq 0 \) means \( \xi_{i,j} \geq 0 \) for \( i = 1, \ldots, N \) and \( j = 1, \ldots, K \cdot M \). We divide the integration range to two sets
\[ \int_{\xi_{i,j} \in A} \rho^{-\left(\|y\|_{\infty} - (N + M - 1 - 2l)r_{DMT}\right) - \sum_{i=1}^{s}|A(i,l)|} \cdot \sum_{i=1}^{N} \sum_{j=1}^{K_M} \xi_{i,j} d\xi_{i,j} + \int_{\xi_{i,j} \in \overline{A}} 1 \cdot \rho^{-\sum_{i=1}^{N} \sum_{j=1}^{K_M} \xi_{i,j}} d\xi_{i,j} \] (167)
where \( A = \bigcap_{i=1}^{N} \bigcap_{j=1}^{K_M} 0 \leq \xi_{i,j} \leq K \cdot M \cdot N \}, \overline{A} = \bigcup_{i=1}^{N} \bigcup_{j=1}^{K_M} \xi_{i,j} > K \cdot M \cdot N \} \), and for the second term in (167) we upper bounded the error probability per channel realization by 1.

We begin by lower bounding the DMT of the first term in (167). In a similar manner to (3), (8), for very large \( \rho \) and finite integration range, we can approximate the integral by finding the most dominant exponential term. Hence, for large \( \rho \) the first term in (167) equals
\[ \rho^{-\min_{\xi_{i,j} \in A}} \left(\|y\|_{\infty} - (N + M - 1 - 2l)r_{DMT}\right) \cdot \sum_{i=1}^{N} \sum_{j=1}^{K_M} \xi_{i,j} \] (168)
Hence, by showing that
\[ \min_{\xi_{i,j} \in A} \left(\|y\|_{\infty} - (N + M - 1 - 2l)r_{DMT}\right) \cdot \sum_{i=1}^{N} \sum_{j=1}^{K_M} \xi_{i,j} \]
greater than or equals to
\[ MN - l(l + 1) - (N + M - 1 - 2l)r_{DMT} \] (169)
we get that the first term attains DMT which is lower bounded by \( d_{M,N}^{(FC)}(r_{DMT}) \). In order to show (169) we use the following lemma.

**Lemma 9.** The solution for the minimization problem
\[ \min_{\xi_{i,j} \in A} \left(\|y\|_{\infty} - (N + M - 1 - 2l)r_{DMT}\right) \cdot \sum_{i=1}^{N} \sum_{j=1}^{K_M} \xi_{i,j} \]
equals to the solution of the following minimization problem
\[ \min_{\alpha \in A'} \sum_{i=1}^{\|y\|_{\infty} - M} (N - i + 1) \alpha_i \]
where \( \alpha = (\alpha_1, \ldots, \alpha_{\|y\|_{\infty}})^T \), and the set \( A' \) fulfills the following two conditions: \( 0 \leq \alpha_i \leq K \cdot M \cdot N \) for \( i = 1, \ldots, \|y\|_{\infty} \cdot M \) and also
\[ \sum_{a=0}^{\|y\|_{\infty} - M} \sum_{b=1}^{M} (N - b + 1) \alpha_{a,M+b} = \|y\|_{\infty} (MN - l(l + 1) - (N + M - 1 - 2l)r_{DMT}) \] (170)

**Proof:** The proof is in appendix [1].

Based on Lemma 9 we can see that by proving
\[ \min_{\alpha \in A'} \sum_{i=1}^{\|y\|_{\infty} - M} (N - i + 1) \alpha_i \geq MN - l(l + 1) - (N + M - 1 - 2l)r_{DMT} \] (170)
we also prove (169). Therefore, we wish to show that any vector \( \alpha \in A' \) fulfills this inequality. Consider a certain vector \( \alpha \in A' \). We define \( \beta_{a,M+b} = \frac{(N+1+b)\cdot a_{a,M+b}}{\|y\|_{\infty}} \) for \( a = 0, \ldots, \|y\|_{\infty} - M, b = 1, \ldots, M \). From this definition we get
\[ \sum_{a=0}^{\|y\|_{\infty} - M} \beta_{a,M+b} = \sum_{a=0}^{\|y\|_{\infty} - M} \sum_{b=1}^{M} (N - b + 1) \alpha_{a,M+b} = MN - l(l + 1) - (N + M - 1 - 2l)r_{DMT}. \] (171)
By assigning in (170) we get
\[
\sum_{a=0}^{\frac{s-1}{2}} \sum_{b=1}^{M} \frac{(N - a \cdot M - b + 1) \alpha_{a,M+b}}{(N - b + 1)} = \sum_{a=0}^{\frac{s-1}{2}} \sum_{b=1}^{M} \frac{(N - a \cdot M - b + 1) \beta_{a,M+b}}{(N - b + 1)}.
\] (172)

We use the following lemma to prove (170).

**Lemma 10.** Consider \( N \geq \frac{s}{s+1} M - 1 \), we get for any \( a = 0, \ldots, s - 1 \) and \( b = 1, \ldots, M \)
\[
\frac{|s| (N - (a \cdot M + b + 1))}{N - b + 1} \geq 1.
\]

**Proof:** The proof is in appendix M.

Since \( K \geq |s| \) and \( N \geq (K + 1) M - 1 \) we can assign the inequality of Lemma 10 in (172) to get
\[
\sum_{a=0}^{\frac{s-1}{2}} \sum_{b=1}^{M} (N - a \cdot M - b + 1) \alpha_{a,M+b} \geq \sum_{a=0}^{\frac{s-1}{2}} \sum_{b=1}^{M} \beta_{a,M+b} = MN - l (l + 1) - (N + M - 1 - 2l) r_{\text{max}}
\] (173)
where the equality results from (171). This proves (170) and so proves that the DMT of the first term in (167) is lower bounded by \( d_{M,N}^{*}(FC) (r_{\text{max}}) \).

Now we show that the second term in (167) is also lower bounded by \( d_{M,N}^{*}(FC) (r_{\text{max}}) \).

\[
\int_{\xi_{i,j} \in A} 1 \cdot \rho^{-\sum_{i=1}^{N} \sum_{j=1}^{K \cdot M} \xi_{i,j} d_{\xi}} \leq \int_{\xi_{i,1} > K \cdot M \cdot N} \rho^{-\sum_{i=1}^{N} \sum_{j=1}^{K \cdot M} \xi_{i,j}}
\]
Since \( d_{M,N}^{*}(FC) (r_{\text{max}}) \leq K \cdot M \cdot N \) the DMT of the second term in (167) is also lower bounded by \( d_{M,N}^{*}(FC) (r_{\text{max}}) \).

We have shown that for \( l = \lfloor r_{\text{max}} \rfloor \) the DMT of \( E_{\text{H}}(\overline{P_{e}(\eta(s), \rho)}) \) is lower bounded by \( d_{M,N}^{*}(FC) (r_{\text{max}}) = MN - l (l + 1) - (M + N - 1 - 2l) r_{\text{max}} \) for any \( s \in \{1, \ldots, K\} \). Since
\[
\overline{P_{e}(\overline{H}_{\text{eff}}^{(l)}, \rho)} \leq \sum_{s \in \{1, \ldots, K\}} \overline{P_{e}(\eta(s), \rho)}
\]
we get that the DMT of the \( K \) sequences of IC’s is also lower bounded by \( d_{M,N}^{*}(FC) (r_{\text{max}}) \). This concludes the proof.

**Appendix L**

**Proof of Lemma 9**

Recall that the optimization problem
\[
\min_{\xi_{i,j} \in A} \left( |s| \cdot (MN - l (l + 1) - (N + M - 1 - 2l) r_{\text{max}}) - \sum_{i=1}^{\frac{|s|M}{2}} A(i,l) \right)^{+} + \sum_{i=1}^{N} \sum_{j=1}^{K \cdot M} \xi_{i,j}
\] (174)
where
\[
A(a \cdot M + b, l) = (N - b + 1) \min_{z \in \{aM+b, \ldots, N\}} \xi_{z,aM+b}
\] (175)
for \( b = 1 \) and \( a = 0, \ldots, s - 1 \), and
\[
A(a \cdot M + b, l) = (N - b - 1) \min_{z \in \{aM+b+1, \ldots, N\}} \xi_{z,aM+b} + \sum_{i=1}^{\min(M-l-1, b-1)} \min_{z \in \{aM+b-1, \ldots, N\}} \xi_{z,aM+b}
\] (176)
for \( b = 2, \ldots, M \) and \( a = 0, \ldots, |s| - 1 \). When \( |s| \cdot M + 1 \leq j \leq K \cdot M \) and \( 1 \leq i \leq N \), we get that \( \xi_{i,j} \) occurs only in the term \( \sum_{i=1}^{N} \sum_{j=1}^{K \cdot M} \xi_{i,j} \) in (174), where \( \xi_{i,j} \geq 0 \). Thus, the minimization is obtained when
\[
\xi_{i,j} = 0 \quad |s| \cdot M + 1 \leq j \leq K \cdot M, \quad 1 \leq i \leq N.
\] (177)

Therefore, we can rewrite the optimization problem
\[
\min_{\xi_{i,j} \in A} \left( |s| \cdot (MN - l (l + 1) - (N + M - 1 - 2l) r_{\text{max}}) - \sum_{i=1}^{\frac{|s|M}{2}} A(i,l) \right)^{+} + \sum_{i=1}^{N} \sum_{j=1}^{\frac{|s|M}{2}} \xi_{i,j}
\] (178)

Now we wish to show that \( \xi_{i,j} = 0 \) for \( j = 1, \ldots, |s| \cdot M \) and \( i = 1, \ldots, j - 1 \). Essentially, we show for \( i < j \) that reducing \( \xi_{i,j} \) has larger effect on (178) than the effect of \(-\min_{z \in \{i, \ldots, N\}} \xi_{z,j}\). First let us observe \( \xi_{i,aM+b} \) for \( i = 1, \ldots, a \cdot M + b - \min(M-l-1, b-1) - 1 \), where \( a = 0, \ldots, |s| - 1 \), \( b = 1, \ldots, M \). Note that this values do not have
any representation in $A(a \cdot M + b, l)$. Therefore, they do not effect $(\cdot)^+$ and only effect $\sum_{s=1}^{N} \sum_{j=1}^{M} \xi_{i,j}$. Thus, in order to obtain the minimum we must choose

$$\xi_{i,a,M+b} = 0 \quad i = 1, \ldots, a \cdot M + b - \min(M - l - 1, b - 1) - 1$$

for any $a = 0, \ldots, |s| - 1$ and $b = 1, \ldots, M$. Note that the function in (178) is continuous. In case $(\cdot)^+ = 0$ the function in (178) can be written as

$$\sum_{a=0}^{M} \sum_{b=1}^{N} \xi_{i,a,M+b}$$

(179)

In this case as long as $(\cdot)^+ = 0$ reducing $\xi_{i,a,M+b}$ for $a \cdot M + b - \min(M - l - 1, b - 1) \leq i \leq a \cdot M + b - 1$ and $a = 0, \ldots, |s| - 1$, $b = 2, \ldots, M$ also reduces (179). When $(\cdot)^+ > 0$ (178) can be written as

$$|s| \cdot (MN - l(l+1) - (N + M - 1 - 2l) r_{max}) + \sum_{a=0}^{M} \sum_{b=2}^{N} \left(\xi_{a,M+b-i,a,M+b} - \min_{z \in \{a,M+b-i,\ldots,N\}} \xi_{z,a,M+b}\right)$$

$$+ \sum_{a=0}^{M} \sum_{b=1}^{N} \left(\xi_{z,a,M+b} - (N - b + 1) \min_{z \in \{a,M+b,\ldots,N\}} \xi_{z,a,M+b}\right).$$

(180)

Since $\xi_{a,M+b-i,a,M+b} \geq \min_{z \in \{a,M+b-i,\ldots,N\}} \xi_{z,a,M+b}$ reducing $\xi_{a,M+b-i,a,M+b}$ also reduces (180). Since the function is continuous, considering these two cases is sufficient in order to state that the minimum is obtained when

$$\xi_{i,j} = 0 \quad j = 1, \ldots, |s| \cdot M, \quad i = 1, \ldots, j - 1.$$  

(181)

This is due to the fact that for any value of $\xi_{z,a,M+b} \geq 0$, $a = 0, \ldots, |s| - 1$, $b = 1, \ldots, M$ and $z = a \cdot M + b, \ldots, N$ the terms in (179), (180) are reduced when decreasing $\left\{\xi_{z,a,M+b-i,a,M+b}\right\}_{i=1}^{\min(M-l-1,b-1)}$, and also since the function is continues. Note that from (180) we can see that decreasing $\sum_{z=a,M+b}^{N} \xi_{z,a,M+b}$ does not necessarily decrease the function. This is due to the fact that $N - b + 1 \geq N - (a \cdot M + b) + 1$, and so the contribution of $(N - b + 1) \min_{z \in \{a,M+b,\ldots,N\}} \xi_{z,a,M+b}$ may be more significant than $\sum_{z=a,M+b}^{N} \xi_{z,a,M+b}$.  

Based on (181) we can rewrite the function in the following manner

$$\left(|s| \cdot (MN - l(l+1) - (N + M - 1 - 2l) r_{max}) - \sum_{a=0}^{M} \sum_{b=1}^{N} (N - b + 1) \min_{z \in \{a,M+b,\ldots,N\}} \xi_{z,a,M+b}\right)^+ + \sum_{a=0}^{M} \sum_{b=1}^{N} \xi_{z,a,M+b}.$$  

(182)

From (182) we can see that the minimum is obtained when

$$\xi_{z,a,M+b} = \alpha_{a,M+b} \quad a \cdot M + b \leq z \leq N$$

(183)

for $a = 0, \ldots, |s| - 1$, $b = 1, \ldots, M$. This is due to the fact that when the values are not equal, reducing the values to the minimal value will reduce $\sum_{z=a,M+b}^{N} \xi_{z,a,M+b}$ while not changing $\min_{z \in \{a,M+b,\ldots,N\}} \xi_{z,a,M+b}$. Therefore, we can write (182) as follows

$$\left(|s| \cdot (MN - l(l+1) - (N + M - 1 - 2l) r_{max}) - \sum_{a=0}^{M} \sum_{b=1}^{N} (N - b + 1) \alpha_{a,M+b}\right)^+ + \sum_{a=0}^{M} \sum_{b=1}^{N} (N - (a \cdot M + b) + 1) \alpha_{a,M+b}$$

(184)

where $0 \leq \alpha_{i} \leq K \cdot M \cdot N$, $i = 1, \ldots, |s| \cdot M$.

We wish to show that the minimum is obtained when

$$\sum_{a=0}^{M} \sum_{b=1}^{N} (N - b + 1) \alpha_{a,M+b} = \left|s\right| \cdot (MN - l(l+1) - (N + M - 1 - 2l) r_{max}).$$

(185)

Again, note that the function is continues. For the case where $(\cdot)^+ = 0$ we get

$$\sum_{a=0}^{M} \sum_{b=1}^{N} (N - (a \cdot M + b) + 1) \alpha_{a,M+b}.$$
On the other hand for the case \((\cdot)^+ > 0\) we get
\[
|s| \cdot (MN - l(l+1) - (N + M - 1 - 2l) r_{\text{max}}) - \sum_{a=0}^{s-1} \sum_{b=1}^{M} (a \cdot M) \alpha_{a,M+b}. \tag{186}
\]

Hence increasing \(\sum_{a=0}^{s-1} \sum_{b=1}^{M} (a \cdot M) \alpha_{a,M+b}\) decreases the function as long as \((\cdot)^+ > 0\) which means
\[
\sum_{a=0}^{s-1} \sum_{b=1}^{M} (N - b+1) \alpha_{a,M+b} < |s| \cdot (MN - l(l+1) - (N + M - 1 - 2l) r_{\text{max}}). \tag{187}
\]

Hence, based on the fact that the function is continuous we get again that for this case the minimal values occur at
\[
\sum_{a=0}^{s-1} \sum_{b=1}^{M} (N - b+1) \alpha_{a,M+b} = |s| \cdot (MN - l(l+1) - (N + M - 1 - 2l) r_{\text{max}}). \tag{188}
\]

The event \(\sum_{a=0}^{s-1} \sum_{b=1}^{M} (N - b+1) \alpha_{a,M+b} = |s| \cdot (MN - l(l+1) - (N + M - 1 - 2l) r_{\text{max}})\), where \(\alpha_i \geq 0, i = 1, \ldots, |s| \cdot M\), is within the range \(0 \leq \alpha_i \leq K \cdot M \cdot N, i = 1, \ldots, |s| \cdot M\). This is because in order to fulfill the equality we get
\[
\max (\alpha_1, \ldots, \alpha_{|s|M}) \leq \frac{|s| \cdot (MN - l(l+1) - (N + M - 1 - 2l) r_{\text{max}})}{N - b+1} \leq K \cdot M \cdot N.
\]

Therefore, the minimization problem solution is obtained for
\[
\min_{\alpha \in A} \sum_{a=0}^{s-1} \sum_{b=1}^{M} (N - (a \cdot M + b) + 1) \alpha_{a,M+b}
\]
where the set \(A^*\) is defined by the following two conditions: \(0 \leq \alpha_i \leq K \cdot M \cdot N, i = 1, \ldots, |s| \cdot M\), and
\[
\sum_{a=0}^{s-1} \sum_{b=1}^{M} (N - b+1) \alpha_{a,M+b} = |s| \cdot (MN - l(l+1) - (N + M - 1 - 2l) r_{\text{max}}). \tag{189}
\]

**APPENDIX M**

**PROOF OF LEMMA 10**

We begin by analyzing the case \(a = |s| - 1\) and \(b = M\). For this case let us consider \(N = (|s|+1) M - 1\). In this case we get
\[
\frac{|s| (N - |s| \cdot M + 1)}{N - M + 1} = \frac{|s| (M)}{|s| M} = 1. \tag{187}
\]

Note that for \(c \geq d \geq 0\) and \(x_2 > x_1 \geq c\) we get
\[
\frac{x_2 - c}{x_2 - d} \geq \frac{x_1 - c}{x_1 - d}. \tag{188}
\]

Hence, based on (188), (187), we get for \(N > (|s|+1) M - 1\)
\[
\frac{|s| (N - (|s| \cdot M - 1))}{N - (M - 1)} \geq \frac{|s| (M)}{|s| M} = 1. \tag{189}
\]

So far we have proved the lemma for \(a = |s| - 1, b = M\) and \(N \geq (|s|+1) M - 1\). For the general case we consider \(\frac{|s|(N-(a \cdot M + b-1))}{N-(b-1)}\). In this case we get
\[
\frac{|s|(N-(a \cdot M + b-1))}{N-(b-1)} = \frac{|s| ((N+|s| M - a \cdot M - b) - (|s| M - 1))}{(N+M-b) - (M-1)} \geq \frac{|s| ((N+|s| M - a \cdot M - b) - (|s| M - 1))}{(N+|s| M - a \cdot M - b) - (M-1)} \geq 1. \tag{190}
\]

where the inequality results from the fact that \(M - b \leq |s| M - a \cdot M - b\). From (188) and (189) we get that
\[
\frac{|s| ((N+|s| M - a \cdot M - b) - (|s| M - 1))}{(N+|s| M - a \cdot M - b) - (M-1)} \geq \frac{|s| (N - (|s| M - 1))}{N - (M - 1)} \geq 1. \tag{191}
\]

From (190), (191) we get the proof of the lemma also for any \(a = 0, \ldots, |s| - 1\) and \(b = 1, \ldots, M\). This concludes the proof.

**APPENDIX N**

**PROOF OF THEOREM 5**

We prove that there exists \(K\) sequences of \(2 \cdot D_1 \cdot T_1\)-real dimensional lattices (as a function of \(\rho\)) that attains the optimal DMT for \(N \geq (K+1) M - 1\). We rely on the extension of the Minkowski-Hlawaka Theorem to the multiple-access channels
presented in (10) Theorem 2]. We upper bound the error probability of the ensemble of lattices for each channel realization, and average the upper bound over all channel realizations to obtain the optimal DMT.

We consider $K$ ensembles of $2 \cdot D_1 \cdot T_1$-real dimensional lattices, one for each user, transmitted using $G_{i}^{(1, \ldots, K)}$ defined in [LV-B]. For user $i$, the first $D_1 \cdot T_1$ dimensions of the lattice are spread on the real part of the non-zero entries of $G_{i}^{(1)}$, and the other $D_1 T_1$ dimensions of the lattice on the imaginary part of the non-zero entries of $G_{i}^{(1)}$. The volume of the Voronoi region of the lattice of user $i$ equals $V_{f}^{(i)} = (\alpha_{f}^{(i)})^{-1} = \rho^{-r_{f} T_{1}}$, i.e. multiplexing gain $r_{i}$. Since the users are distributed, the effective lattice in the transmitter can be written as $\Lambda_{tr} = \Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_K$, where $\Lambda_i$ is the lattice transmitted by user $i$.

In the receiver the channel induces a new lattice $H_{eff}^{(i)} \cdot \hat{x}$, where $\hat{x} \in \Lambda_{tr}$. For lattices with regular lattice decoding, the error probability is equal among all codewords. Hence, it is sufficient to analyze the lattice’s zero codeword error probability. Without loss of generality let us assume that the receiver rotates $y_{ex}$ such that the channel can be rewritten as

$$y_{ex} = B \cdot \hat{x} + \tilde{n}_{ex}$$

where $B^\top B = H_{eff}^{(i), K} H_{eff}^{(i), K}$, and $\tilde{n}_{ex} \sim CN(0, \rho^{-1} \cdot \frac{D_{eff}}{2 \cdot r_{f}} \cdot I_{K \cdot D_1 \cdot T_1})$.

We define the indication function of a $2 \cdot K \cdot D_1 \cdot T_1$ dimensional ball with radius $2R$ centered around zero by

$$I_{Ball(2R)}(\hat{x}) = \left\{ \begin{array}{ll} 1, & \| \hat{x} \| \leq 2R \\ 0, & \text{else} \end{array} \right.$$  

In addition let us define the continues function of bounded support $f_{rec}(\hat{x}) = I_{Ball(2R_{eff})}(\hat{x}) \cdot Pr(\|\tilde{n}_{ex}\| > \|\hat{x} - \tilde{n}_{ex}\|)$. Based on (141) we can state that for each lattice induced in the receiver, $\Lambda_{rc}$, the lattice zero codeword error probability is upper bounded by

$$\sum_{\hat{x} \in \Lambda_{rc}, \hat{x} \neq 0} f_{rec}(\hat{x}) + Pr(\|\tilde{n}_{ex}\| > R_{eff}).$$

where $R_{eff}^{2} = \rho^{-1} \cdot \frac{D_{eff}}{2 \cdot r_{f}}$.

For regular lattice decoding we can equivalently consider

$$\tilde{y}_{ex} = B^{-1} \cdot y_{ex} = \hat{x} + \tilde{n}_{ex}.$$

where $\tilde{n}_{ex} \sim CN(0, (H_{eff}^{(i), K} H_{eff}^{(i), K})^{-1})$, i.e. the lattice in the receiver remains $\Lambda_{tr}$ and the effect of the channel realization is passed on to the additive noise. In addition let us denote an indication function over an ellipse centered around zero by

$$I_{ellipse(B,2R)}(\hat{x}) = \left\{ \begin{array}{ll} 1, & \| B \cdot \hat{x} \| \leq 2R \\ 0, & \text{else} \end{array} \right.$$  

By defining the continues function $g_{rec}(\hat{x}) = I_{ellipse(B,2R_{eff})}(\hat{x}) \cdot Pr(\|B\tilde{n}_{ex}\| > \|B(\hat{x} - \tilde{n}_{ex})\|)$ we get the following upper bound on the error probability

$$\sum_{\hat{x} \in \Lambda_{rc}, \hat{x} \neq 0} g_{rec}(\hat{x}) + Pr(\|B \cdot \tilde{n}_{ex}\| \geq R_{eff})$$

that equals to the upper bound in (193). In addition, since $f_{rec}(B \cdot \hat{x}) = g_{rec}(\hat{x})$, and based on the fact that $H_{eff}^{(i), K}$ is a block diagonal matrix we get

$$|H_{eff}^{(i), (S)} H_{eff}^{(i), (S)}|^{-1} \cdot \int_{\mathbb{R}^{2 \cdot |S|} \cap D_{1} \cdot T_{1}} f_{rec}(\hat{x}(S)) d\hat{x}(S) = \sum_{\hat{x} \in \mathbb{R}^{2 \cdot |S|} \cap D_{1} \cdot T_{1}} g_{rec}(\hat{x}(S)) d\hat{x}(S) \quad \forall S \subseteq \{1, \ldots, K\}$$

where $\hat{x}(S)$ equals zero in the entries corresponding to $\{1, \ldots, K\} \setminus S$ and the other entries are in $\mathbb{R}^{2 \cdot |S| \cdot D_{1} \cdot T_{1}}$.

In (10) Theorem 2] Nam and El Gamal extended the Minkowski-Hlawka theorem to multiple-access channels by using Loeliger ensembles of lattices [13] for each user. From this theorem we get that for a certain Riemann integrable function of bounded support $f(x)$

$$E_{\Lambda_{tr}} \left( \sum_{\hat{x} \in \Lambda_{tr}, \hat{x} \neq 0} f(\hat{x}) \right) = \sum_{S \subseteq \{1, \ldots, K\}} \prod_{s \in S} \frac{1}{V_{f}^{(s)}} \int_{\mathbb{R}^{2 \cdot |S|} \cap D_{1} \cdot T_{1}} f(x(S)) d\hat{x}(S).$$

For each channel realization $B$, the function $g_{rec}(\hat{x})$ is bounded, and so by averaging over the Loeliger ensembles for the multiple-access channels, we get based on (195), (197) that the upper bound on the error probability using regular lattice decoding is

$$\sum_{S \subseteq \{1, \ldots, K\}} \prod_{s \in S} \frac{1}{V_{f}^{(s)}} \int_{\mathbb{R}^{2 \cdot |S|} \cap D_{1} \cdot T_{1}} g_{rec}(\hat{x}(S)) d\hat{x}(S) + Pr(\|B \cdot \tilde{n}_{ex}\| \geq R_{eff}).$$
By assigning the relation of (196) in (198), we get
\[ \sum_{S \subseteq \{1, \ldots, K\}} \rho T_1 \sum_{S \subseteq \{1, \ldots, K\}} r_s |H_{eff}^{(1)}(S)| \frac{1}{H_{eff}^{(1)}(S)} - 1 \cdot \int_{S \subseteq \{1, \ldots, K\}} f_{\rho T_1}(x) \left( \frac{d_{\rho T_1}(S)}{H_{eff}^{(1)}(S)} \right) \rho T_1 \sum_{S \subseteq \{1, \ldots, K\}} r_s |H_{eff}^{(1)}(S)| \frac{1}{H_{eff}^{(1)}(S)} - 1. \]

Based on the bounds derived in [8, Theorem 3], we can upper bound the integral of the first term in (199) by
\[ \sum_{S \subseteq \{1, \ldots, K\}} \frac{|S| - D_1 - \sum_{S \subseteq \{1, \ldots, K\}} r_s |H_{eff}^{(1)}(S)| \frac{1}{H_{eff}^{(1)}(S)} - 1. \]

Since we consider radius of \( R_{eff} \), for large values of \( \rho \) the second term in (199) is negligible compared to the first term [8, Theorem 3]. Hence, the remaining step is calculating the average over all channel realizations. We divide the average into two ranges \( A \) and \( \overline{A} \) as depicted in Theorem 7. Hence, the remaining step is calculating the average over all channel realizations. We divide the average into two ranges \( A \) and \( \overline{A} \) as depicted in Theorem 7. This bound coincides with the upper bound in Theorem 7 which was shown to obtain the optimal DMT. This concludes the proof.

APPENDIX O
PROOF OF COROLLARY 2

We first consider the symmetric case \( r_1 = \cdots = r_K = r_{\text{max}} \). Similarly to [8, Corollary 3] we can state that if a sequence of \( K \) lattices attains diversity order \( d \) for symmetric multiplexing gain \( r_{\text{max}} = 0 \), it also attains diversity order
\[ d \left( 1 - \frac{r_{\text{max}}}{D_{r_{\text{max}}} T_{r_{\text{max}}}} \right) \]
for any symmetric multiplexing gain \( 0 < r_{\text{max}} \leq D_{r_{\text{max}}} T_{r_{\text{max}}} \). This is due to the fact that changing \( r_{\text{max}} \) merely has the effect of scaling the effective lattice in the receiver. From Theorem 3 we get that there exists a sequence of \( K \) lattices (one for each user) that attains for symmetric multiplexing gain \( r_{\text{max}} = l \) the optimal DMT \( d^*_{M,N}(FC) \mid l \), where \( l = 0, \ldots, M - 1 \). In this case we also get from (200) and Theorem 8 that this sequence also attains the optimal DMT \( d^*_{M,N}(FC) \mid r_{\text{max}} \), when the symmetric multiplexing gain is in the range \( l \leq r_{\text{max}} \leq l + 1 \).

Now consider for the same sequence of lattices a multiplexing gains tuple \( (r_1, \ldots, r_K) \) with \( r_{\text{max}} \) as its maximal multiplexing gain. The performance can only improve compared to the symmetric case since some of the multiplexing gains of the users are smaller than \( r_{\text{max}} \). Since the DMT can not be any larger than \( d^*_{M,N}(FC) \mid r_{\text{max}} \), which is already obtained in the symmetric case, we get that \( d^*_{M,N}(FC) \mid r_{\text{max}} \) is obtained by any multiplexing gains tuple with \( r_{\text{max}} \) as its maximal value.

REFERENCES

[1] I. Telatar, “Capacity of multi-antenna gaussian channels,” Europ. Trans. Telecommun., vol. 10, pp. 585 –595, Nov. 1999.
[2] G. Foschini, “Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas,” Bell Labs Tech. J., vol. 1, no. 2, pp. 41 –59, 1996.
[3] L. Zheng and D. Tse, “Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels,” IEEE Trans. on Inf. Theory, vol. 49, no. 5, pp. 1073 –1096, 2003.
[4] H. El Gamal, G. Caire, and M. Damen, “Lattice coding and decoding achieve the optimal diversity-multiplexing tradeoff of mimo channels,” IEEE Trans. on Inf. Theory, vol. 50, no. 6, pp. 986 –985, 2004.
[5] Y. Yona and M. Feder, “Fundamental limits of infinite constellations in mimo fading channels,” Submitted to IEEE Trans. on Information Theory. Available on arxiv.org.
[6] D. Tse, P. Viswanath, and L. Zheng, “Diversity-multiplexing tradeoff in multiple-access channels,” Information Theory, IEEE Transactions on, vol. 50, no. 9, pp. 3869 –3884, sep. 2006.
[7] J. Jalden and P. Elia, “Dmt optimality of lr-aided linear decoders for a general class of channels, lattice designs, and system models,” IEEE Trans. on Inf. Theory, vol. 56, no. 10, pp. 4765 –4780, oct. 2010.
[8] Y. Yona and M. Feder, “Fundamental limits of infinite constellations in mimo fading channels,” Submitted to IEEE Trans. on Information Theory. Available on arxiv.org.
[9] Y.-H. Nam and H. El Gamal, “On the optimality of lattice coding and decoding in multiple access channels,” in Information Theory, 2007. ISIT 2007. IEEE International Symposium on, june 2007, pp. 211 –215.
[10] H.-F. Lu, C. Hollanti, R. Vehkalahti, and J. Lahtonen, “Dmt optimal codes constructions for multiple-access mimo channel,” Information Theory, IEEE Transactions on, vol. 57, no. 6, pp. 3594–3617, June.
[11] G. Poltyrev, “On Coding Without Restrictions for the AWGN Channel,” IEEE Trans. on Inf. Theory, vol. 40, no. 2, pp. 409–417, 1994.
[12] H.-A. Loeliger, “Averaging bounds for lattices and linear codes,” IEEE Trans. on Inf. Theory, vol. 43, no. 6, pp. 1767 –1773, nov 1997.
[13] V. Tarokh, H. Jafarkhani, and A. Calderbank, “Space-time block codes from orthogonal designs,” IEEE Trans. on Inf. Theory, vol. 45, no. 5, pp. 1456 –1467, jul. 1999.
[14] S. Alamouti, “A simple transmit diversity technique for wireless communications,” Selected Areas in Communications, IEEE Journal on, vol. 16, no. 8, pp. 1451 –1458, oct 1998.