



Title	Non-orthogonal Opportunistic Beamforming: Performance Analysis and Implementation
Author(s)	Xia, M; Wu, YC; Aissa, S
Citation	IEEE Transactions on Wireless Communications, 2012, v. 11, p. 1424-1433
Issued Date	2012
URL	http://hdl.handle.net/10722/164032
Rights	Creative Commons: Attribution 3.0 Hong Kong License

Non-Orthogonal Opportunistic Beamforming: Performance Analysis and Implementation

Minghua Xia, Yik-Chung Wu, and Sonia Aïssa, *Senior Member, IEEE*

Abstract—Aiming to achieve the sum-rate capacity in multi-user multi-antenna systems where N_t antennas are implemented at the transmitter, opportunistic beamforming (OBF) generates N_t orthonormal beams and serves N_t users during each channel use, which results in high scheduling delay over the users, especially in densely populated networks. Non-orthogonal OBF with more than N_t transmit beams can be exploited to serve more users simultaneously and further decrease scheduling delay. However, the inter-beam interference will inevitably deteriorate the sum-rate. Therefore, there is a tradeoff between sum-rate and scheduling delay for non-orthogonal OBF. In this context, system performance and implementation of non-orthogonal OBF with $N > N_t$ beams are investigated in this paper. Specifically, it is analytically shown that non-orthogonal OBF is an interference-limited system as the number of users $K \rightarrow \infty$. When the inter-beam interference reaches its minimum for fixed N_t and N , the sum-rate scales as $N \ln \left(\frac{N}{N-N_t} \right)$ and it degrades monotonically with the number of beams N for fixed N_t . On the contrary, the average scheduling delay is shown to scale as $\frac{1}{N} K \ln K$ channel uses and it improves monotonically with N . Furthermore, two practical non-orthogonal beamforming schemes are explicitly constructed and they are demonstrated to yield the minimum inter-beam interference for fixed N_t and N . This study reveals that, if user traffic is light and one user can be successfully served within a single transmission, non-orthogonal OBF can be applied to obtain lower worst-case delay among the users. On the other hand, if user traffic is heavy, non-orthogonal OBF is inferior to orthogonal OBF in terms of sum-rate and packet delay.

Index Terms—Multiple-input single-output (MISO), multi-user scheduling, non-orthogonal transmission, opportunistic beamforming.

I. INTRODUCTION

FOR different scheduling strategies in the downlink of multi-user wireless systems, there are two conflicting goals. One is to satisfy the quality of service (QoS) requirements of different users, such as scheduling delay, and the other aims to maximize system throughput [1], [2]. Round-robin scheduling follows a strict order to serve each user once

in each round such that it guarantees minimum scheduling delay over the users. On the contrary, opportunistic scheduling exploits multi-user diversity gain and achieves the sum-rate capacity as the number of users K approaches infinity [3]. When the base station (BS) is equipped with $N_t > 1$ transmit antennas, opportunistic scheduling can be implemented by opportunistic beamforming (OBF), which generates $N = N_t$ beams and serves up to N_t users for each channel use [4], [5]. While it achieves the sum-rate capacity, OBF can result in significant scheduling delay over the users, especially in densely populated networks. Specifically, when the number of users K is large and all users have the same average signal-to-noise ratio (SNR), the average scheduling delay for which each user is served at least once is $(K \ln K)/N_t$ channel uses [1]. In practice, shorter scheduling delay is much more desirable for wireless applications in general and for delay-sensitive traffic such as audio/video streaming, and is even a must for critical data such as alarm applications [6]–[10].

Conventionally, opportunistic and round-robin scheduling can be combined to keep a balance between scheduling delay and sum-rate [11]. On the other hand, for the OBF, if $N > N_t$ beams are generated for each channel use, more users can be simultaneously served and scheduling delay will be further reduced. Unfortunately, in this case, inter-beam interference will inevitably deteriorate the sum-rate. Therefore, there is a tradeoff between sum-rate and scheduling delay when the number of transmit beams N increases. This motivates us to analytically investigate the performance and implementation of OBF with $N > N_t$ beams, by using the same amount of system resource. Hereafter, the conventional OBF with $N = N_t$ beams is denoted by “orthogonal OBF” while the scheme with $N > N_t$ beams is referred to as “non-orthogonal OBF”.

In order to simultaneously transmit on more than N_t beams at the N_t -antenna BS, non-orthogonal beamforming vectors have to be designed and analyzed. In [12], [13], a Grassmannian beamforming scheme with $N = N_t + 1$ beams is proposed, where beamforming vectors can be constructed from the vertices of a regular simplex in the space \mathbb{C}^{N_t} [14]. However, this construction cannot be extended to the general case with $N > N_t + 1$. In [15], a greedy algorithm is proposed to construct non-orthogonal beamforming matrices for arbitrary N . Unfortunately, the prior work [12], [13], [15] show only some simulation results and no rigorous analysis on the system performance has been made. Filling this important gap, in this work we analytically investigate system performance in terms of achievable sum-rate and scheduling delay, and design optimal beamforming matrices

Manuscript received February 18, 2011; revised June 22, August 27, and November 29, 2011; accepted January 13, 2012. The associate editor coordinating the review of this paper and approving it for publication was K. B. Lee.

The material in this paper was presented in part at the IEEE International Conference on Communications, Kyoto, Japan, 2011. This work was supported in part by the HKU Seed Funding Programme, Project No. 200911159084, and by King Abdullah University of Science and Technology (KAUST).

M. Xia is with the Division of Physical Sciences and Engineering, KAUST, Thuwal, Saudi Arabia (e-mail: minghua.xia@kaust.edu.sa).

Y.-C. Wu is with the Department of Electrical and Electronic Engineering, The University of Hong Kong (e-mail: ycwu@eee.hku.hk).

S. Aïssa is with INRS, University of Quebec, Montreal, QC, Canada, and with KAUST, Thuwal, Saudi Arabia (e-mail: sonia.aissa@ieee.org).

Digital Object Identifier 10.1109/TWC.2012.020812.110311

achieving minimum inter-beam interference. In particular, our results include previously reported results under orthogonal transmission as special cases.

Specifically, we focus on the general non-orthogonal OBF with $N > N_t$ transmit beams. Increasing N is essentially equivalent to increasing the spatial multiplexing gain. On the contrary, the inter-beam interference due to non-orthogonal beams will inevitably deteriorate the data-rate on each beam. Therefore, the key question is: how does the number of beams N and inter-beam interference affect the final sum-rate? In this paper, the achievable sum-rate and the sum-rate scaling law are established. In particular, the sum-rate scaling law reveals that non-orthogonal OBF is an interference-limited technique. Moreover, when inter-beam interference is minimized for fixed N_t and N , the sum-rate monotonically decreases with the number of beams N .

On the other hand, increasing N allows transmitting to more users during each channel use, thus decreasing scheduling delay. Our results show that the average scheduling delay, which guarantees that each user is served at least once, improves monotonically with increasing number of beams N . Moreover, when practical constraints on the received SINR and outage probability are imposed, the maximum number of supported transmit beams is derived.

As aforementioned, in order to maximize the sum-rate for fixed N_t and N , the inter-beam interference shall be kept as small as possible. But, in practice, how to explicitly construct the non-orthogonal beamforming vectors to achieve the smallest inter-beam interference remains an open problem. Thus, in the last part of this paper, two practical non-orthogonal beamforming matrices are proposed and they are shown to achieve the minimum inter-beam interference.

This study reveals that, if user traffic is light and one user can be successfully served within a single transmission, e.g., in alarm applications, non-orthogonal OBF can be applied to obtain lower worst-case delay among the users [1]. On the other hand, if multiple transmissions are needed to successfully serve a user, orthogonal transmission is preferable due to its higher achievable sum-rate and lower packet delay.

The rest of this paper is organized as follows. Section II describes the system model and the scheduling strategy. In Section III, the distribution function of the received SINR is developed. Section IV focuses on the performance analysis based on the obtained SINR distributions. Section V illustrates two practical non-orthogonal beamforming matrices. Simulation results and discussions are presented in Section VI and, finally, Section VII concludes the paper.

II. SYSTEM MODEL AND SCHEDULING STRATEGY

We consider the downlink transmission from a BS equipped with N_t antennas to K single-antenna users. The number of users, K , is assumed to be larger than the number of transmit beams, N . The block-fading channels of different users are mutually independent and identically distributed. The transmit power is equally allocated among different transmit antennas, and the average SNRs for all users are assumed to be identical.

At each time slot, N different symbols in the vector $\mathbf{x} = [x_1, \dots, x_N]^H$ are simultaneously transmitted onto N differ-

ent beams, where $(\cdot)^H$ denotes the Hermitian transpose operator. Prior to transmission, \mathbf{x} is multiplied by a $N_t \times N$ beamforming matrix $\mathbf{B} = [\mathbf{b}_1, \dots, \mathbf{b}_N]$, where \mathbf{b}_n is the beamforming vector for beam n . Therefore, the received symbol of user k is given by $y_k = \sqrt{\frac{\rho}{N}} \sum_{n=1}^N \mathbf{h}_k^H \mathbf{b}_n x_n + z_k$, where ρ is the average SNR at the BS, $\mathbf{h}_k = [h_1^k, h_2^k, \dots, h_{N_t}^k]^H$ is the complex Gaussian channel vector between the BS and user k , with zero mean and unit covariance matrix \mathbf{I} , and z_k denotes the additive white Gaussian noise (AWGN) with zero mean and unit variance.

As to the scheduling strategy at the BS, each user calculates the received SINRs on N different beams, and feeds back the maximum SINR and its corresponding beam index to the BS. Specifically, the received SINR of user k on beam n is given by

$$\gamma_{k,n} = \frac{|\mathbf{h}_k^H \mathbf{b}_n|^2}{\frac{N}{\rho} + \sum_{l=1, l \neq n}^N |\mathbf{h}_k^H \mathbf{b}_l|^2}, \quad (1)$$

where $|x|$ denotes the amplitude of x . For user k , the maximum SINR among N beams is determined as $\hat{\gamma}_k = \max_{n=1, \dots, N} \gamma_{k,n}$, and the corresponding beam-index is $\hat{n}_k = \arg \max_{n=1, \dots, N} \gamma_{k,n}$. Therefore, the feedback pertaining to user k is the pair $(\hat{\gamma}_k, \hat{n}_k)$.

After receiving all the feedbacks at the BS, the user with the maximum SINR among all users whose $\hat{n}_k = n$ is chosen to be served on beam n . More precisely, the maximum SINR achieved through beam n is determined as

$$\hat{\gamma}_{\max, n} = \max \{ \hat{\gamma}_k \} \quad \text{s.t.} \quad \hat{n}_k = n. \quad (2)$$

III. DISTRIBUTION FUNCTION OF RECEIVED SINR

In general, increasing the number of transmit beams enhances spatial multiplexing gain but also increases the total inter-beam interference, thus degrading the achievable sum-rate. Aiming at assessing this tradeoff, we analytically investigate the distribution function of the received SINR in this section. The system performance analyses are provided in the next section.

With the principle of orthogonal projection, the beamforming vector \mathbf{b}_l can be expressed in terms of \mathbf{b}_n via their cross-correlation coefficient $\delta_{l,n} \triangleq |\mathbf{b}_l^H \mathbf{b}_n|$, that is, $\mathbf{b}_l = \delta_{l,n} \mathbf{b}_n + \sqrt{1 - \delta_{l,n}^2} \mathbf{b}_{l,n}^\perp$, $1 \leq l, n \leq N$, where $\mathbf{b}_{l,n}^\perp$ is the orthonormal vector of \mathbf{b}_l to \mathbf{b}_n . Substituting the previous equation into (1) and performing some algebraic manipulations, we obtain

$$\gamma_{k,n} = X \left\{ \frac{N}{\rho} + \sum_{l=1, l \neq n}^N \left[\delta_{l,n}^2 X + (1 - \delta_{l,n}^2) Y_l + \delta_{l,n} \sqrt{1 - \delta_{l,n}^2} Z_l \right] \right\}^{-1}, \quad (3)$$

where $X \triangleq |\mathbf{h}_k^H \mathbf{b}_n|^2$, $Y_l \triangleq |\mathbf{h}_k^H \mathbf{b}_{l,n}^\perp|^2$, and

$$Z_l \triangleq \mathbf{h}_k^H \mathbf{b}_n (\mathbf{h}_k^H \mathbf{b}_{l,n}^\perp)^H + \mathbf{h}_k^H \mathbf{b}_{l,n}^\perp (\mathbf{h}_k^H \mathbf{b}_n)^H. \quad (4)$$

Based on the theory of optimal Grassmannian line packing [17], in order to guarantee that the correlation between any

two beamforming vectors is as small as possible, beamforming vectors must be symmetric, i.e., $\delta_{l,n} = \delta_0$, where $1 \leq l \leq N$ and $l \neq n$. Define $\alpha \triangleq (N-1)\delta_0^2$ as the total inter-beam interference from the other $N-1$ beams on beam n . Also, let $\beta \triangleq 1 - \delta_0^2$, $\eta \triangleq \delta_0\sqrt{1 - \delta_0^2}$, $Y \triangleq \sum_{l=1, l \neq n}^N Y_l$, and $Z \triangleq \sum_{l=1, l \neq n}^N Z_l$. Then, when the inter-beam interference are symmetric, (3) can be rewritten as

$$\gamma_{k,n} = \frac{X}{\frac{N}{\rho} + \alpha X + \beta Y + \eta Z}. \quad (5)$$

Since \mathbf{h}_k is a complex Gaussian random vector and that \mathbf{b}_n and $\mathbf{b}_{l,n}^\perp$ are normalized constant vectors, it is clear that X is of exponential distribution with unit mean and unit variance, and Y is of chi-square distribution with $2(N-1)$ degrees of freedom. Furthermore, the PDF of Z in (5) is summarized in the following lemma.

Lemma 1: The PDF of Z in (5) is given by

$$f(z) = \frac{1}{2^N \Gamma(N)} z^{N-1} W_{0, N-\frac{1}{2}}(2z), \quad -\infty < z < +\infty, \quad (6)$$

with zero mean and variance $2(N-1)$, where $W_{\kappa, \mu}(z)$ denotes the Whittaker function [19, Eq.(13.14.3)].

Proof: See Appendix A. ■

Due to the high complexity of the PDF in (6), the exact distribution function of $\gamma_{k,n}$ in (5) is hard to obtain. However, after carefully examining the distributions of αX , βY and ηZ in the denominator of (5), we find that ηZ is negligible. Specifically, βY has mean $N-1$ and variance $(N-1)(1-\delta_0^2)^2$, and ηZ has zero mean and small variance $2(N-1)\delta_0^2(1-\delta_0^2)$, since δ_0 is usually very small ($\delta_0^2 \ll 1$) in order to avoid inter-beam interference. For example, when $N=5$ and $\delta_0=0.25$, the PDF and cumulative density function (CDF) of αX , βY and ηZ are plotted in Fig. 1. It is observed that the PDF of ηZ is symmetric with respect to zero and its values are generally much smaller than βY . Therefore, the effect of ηZ on the received SINR $\gamma_{k,n}$ is negligible, and (5) can be approximated using

$$\gamma_{k,n} \approx \frac{X}{\frac{N}{\rho} + \alpha X + \beta Y}. \quad (7)$$

Consequently, the CDF and PDF of $\gamma_{k,n}$ in (7) can be, respectively, expressed as

$$F_{\Gamma_{k,n}}(\gamma) = 1 - \exp\left(-\frac{N\gamma}{\rho(1-\alpha\gamma)}\right) \left(1 + \frac{\beta\gamma}{1-\alpha\gamma}\right)^{-(N-1)} \quad (8)$$

and

$$f_{\Gamma_{k,n}}(\gamma) = \frac{\exp\left(-\frac{N\gamma}{\rho(1-\alpha\gamma)}\right)}{(1-\alpha\gamma) \left(1 + \frac{\beta\gamma}{1-\alpha\gamma}\right)} \times \left[\frac{N}{\rho} \left(1 + \frac{\beta\gamma}{1-\alpha\gamma}\right) + (N-1)\beta \right]. \quad (9)$$

Note that, as a special case, the orthogonal OBF with $N = N_t$ implies $\delta_0 = 0$ and, thus, $\alpha = 0$ and $\beta = 1$. Putting $\alpha = 0$ and $\beta = 1$ into (7), (8), and (9), the expressions reduce to the exact counterparts under orthogonal case [18]. Consequently, our analysis is general and applicable to either orthogonal or non-orthogonal OBF. Also, the accuracy of the approximation

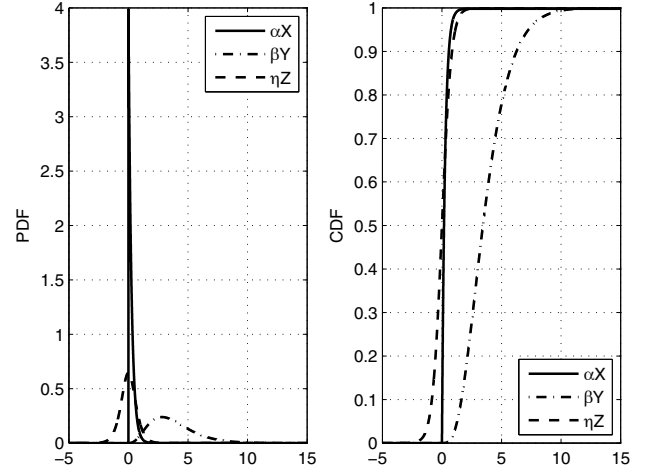


Fig. 1. Distribution functions of αX , βY and ηZ in the denominator of the received SINR in (5), with $N=5$ and $\delta_0=0.25$.

in (7) will be further illustrated in Fig. 2, and is detailed in Section VI.

Since the maximum SINR in (2) is given by $\hat{\gamma}_{\max,n} = \max_{k=1, \dots, K} \gamma_{k,n}$, by using the results from the theory of order statistics, the CDF and PDF of $\hat{\gamma}_{\max,n}$ in (2) is, respectively, given by

$$F_{\Gamma_{\max,n}}(\gamma) = F_{\Gamma_{k,n}}^K(\gamma) \quad (10)$$

and

$$f_{\Gamma_{\max,n}}(\gamma) = K f_{\Gamma_{k,n}}(\gamma) F_{\Gamma_{k,n}}^{K-1}(\gamma). \quad (11)$$

Note that when a user simultaneously obtains the maximum SINR on different beams, the transmitter has to choose another user who has the second largest SINR. This probability is very small when the number of users is large [4], and it is neglected herein.

IV. PERFORMANCE ANALYSES

Based on the distribution function of the received SINR, the sum-rate scaling law is first investigated in this section. The analysis of the average scheduling delay over the users then follows. Moreover, when practical constraints on the received SINR and outage probability are imposed, the maximum number of supported transmit beams and access probability are derived.

A. Sum-Rate Scaling Law

For beam n with received SINR $\hat{\gamma}_{\max,n}$, the instantaneous data rate can be calculated by the Shannon formula $\ln(1 + \hat{\gamma}_{\max,n})$ in the unit of nat/s/Hz. Moreover, since there are N beams in total, the achievable sum-rate is given by

$$R = \sum_{n=1}^N \mathcal{E} \{ \ln(1 + \hat{\gamma}_{\max,n}) \} \quad (12)$$

$$= \sum_{n=1}^N \int_0^{+\infty} \ln(1 + \gamma) K f_{\Gamma_{k,n}}(\gamma) F_{\Gamma_{k,n}}^{K-1}(\gamma) d\gamma \quad (13)$$

$$= KN \int_0^{+\infty} \ln(1+\gamma) f_{\Gamma_{k,n}}(\gamma) F_{\Gamma_{k,n}}^{K-1}(\gamma) d\gamma, \quad (14)$$

where (11) was exploited to reach (13). Due to the complicated expressions of $f_{\Gamma_{k,n}}$ and $F_{\Gamma_{k,n}}$ in (8) and (9) respectively, the integration in (14) has no closed-form expression and has to be evaluated numerically. In order to gain insights into the sum-rate, we instead derive the limiting distribution of $\hat{\gamma}_{\max,n}$ as $K \rightarrow \infty$. However, notice that the limiting distribution cannot be obtained by directly applying $K \rightarrow \infty$ in (10) since, for any $F_{\Gamma_{k,n}}(\gamma) < 1$, (10) reduces to 0 as $K \rightarrow \infty$ and hence (10) is a degenerate distribution. In the following lemma, the asymptotic theory of extreme order statistics is exploited to attain a non-degenerate limiting distribution for $\hat{\gamma}_{\max,n}$ such that the sum-rate scaling law is obtained.

Lemma 2: For the received SINR $\gamma_{k,n}$ of user k with respect to beam n , with CDF $F_{\Gamma_{k,n}}(\gamma)$ and PDF $f_{\Gamma_{k,n}}(\gamma)$ given by (8) and (9) respectively, as the number of users $K \rightarrow \infty$, the limiting distribution of the maximum SINR in (2) is of the Gumbel distribution. That is,

$$\lim_{K \rightarrow \infty} F_{\Gamma_{\max,n}}(\gamma) = H_{3,0} \left(\frac{\gamma - a}{b} \right), \quad (15)$$

where $H_{3,0}(\gamma) = \exp(-e^{-\gamma})$, and the normalizing parameters a and b are given by

$$a = \frac{\rho \ln K}{c + \rho \alpha \ln K} + \mathcal{O}(\ln \ln K), \quad (16)$$

where the notation $f(x) = \mathcal{O}(g(x))$ is defined as $\lim_{x \rightarrow \infty} |f(x)/g(x)| < \infty$, and

$$b = \frac{c^2 (c + \rho \beta \ln K)}{(c + \rho \alpha \ln K)^2 \left[\frac{N}{\rho} (c + \rho \beta \ln K) + c(N-1)\beta \right]}, \quad (17)$$

respectively. In (16) and (17), the constant $c \triangleq N + \rho(N-1)\beta$.

Proof: See Appendix B. ■

Applying Lemma 2 to (12) yields:

$$R \leq \sum_{n=1}^N \ln(1 + \mathcal{E}\{\hat{\gamma}_{\max,n}\}) \quad (18)$$

$$= N \ln(1 + a + b\Upsilon), \quad (19)$$

where the Jensen's inequality was exploited to derive (18) from (12), the limiting distribution (15) was exploited to reach (19) and $\Upsilon = 0.5772 \dots$ is the Euler's constant. Moreover, substituting (16) and (17) into (19) and examining it as $K \rightarrow \infty$, the expression (19) reduces to a simple sum-rate scaling law as given in the following theorem.

Theorem 1: When the total inter-beam interference α for each beam is identical, and as $K \rightarrow \infty$, the sum-rate of non-orthogonal OBF scales as

$$R \sim N \ln \left(1 + \frac{1}{\alpha} \right), \quad (20)$$

where the notation $f(K) \sim g(K)$ is defined as $\lim_{K \rightarrow \infty} f(K)/g(K) = 1$.

Special Case: The preceding analysis can be applied to orthogonal OBF without inter-beam interference, i.e., where $\delta_0 = 0$. In particular, putting $\alpha = 0$, $\beta = 1$, and $N = N_t$ into (16), (17), and (19) results in the following formula for the

sum-rate as $K \rightarrow \infty$

$$R' \sim N_t \ln \ln K, \quad (21)$$

which is exactly the sum-rate scaling law under orthogonal transmission [4]. Furthermore, comparing (20) with (21), it is observed that, as $K \rightarrow \infty$, for non-orthogonal OBF, multi-user diversity gain vanishes and the sum-rate is dominated by N and α while, for orthogonal OBF, multi-user diversity gain always benefits the sum-rate since there is no inter-beam interference.

The theorem above reveals that the sum-rate of non-orthogonal OBF increases proportionally to the number of transmit beams N , but it is offset by the total inter-beam interference α . That is, for fixed N_t and N , non-orthogonal OBF is an interference-limited system. Clearly, in order to achieve maximum R with fixed N and N_t , α must be kept as small as possible. Practical non-orthogonal beamforming matrices achieving the smallest α will be illustrated in Section V.

B. Average Scheduling Delay

For delay-sensitive traffic, the delay is usually analyzed in terms of the waiting time of packets in the input/output queues at the network layer, on average and over all users. However, in order to provide delay guarantee for all users, we have to study the delay for *the most unfortunate user* in the network. Therefore, in [1], the scheduling delay is defined as the minimum number of channel uses that guarantees each of the users is served at least one packet. This definition reflects the worst-case delay among the users, which roughly determines the time scale required for the system to behave fairly in a homogeneous network [1]. The guarantee of the worst-case delay for all users is increasingly important in commercial wireless systems.

Based on [1], the scheduling delay $D_{1,K}$ in this paper is defined as the minimum number of channel uses that guarantees each of the K users is served at least once. For round-robin scheduling, it is clear that $D_{1,K} = K$. For opportunistic single-user scheduling, the average scheduling delay scales as $K \ln K$, while for the orthogonal OBF, it scales as $K \ln K / N_t$ [1]. Based on the probabilistic arguments similar to [22, Theorem 3], the following theorem shows the average scheduling delay of non-orthogonal OBF with $N > N_t$ transmit beams.

Theorem 2: For the non-orthogonal OBF with $N > N_t$ transmit beams, as $K \rightarrow \infty$, the average scheduling delay is given by

$$\mathcal{E}\{D_{1,K}\} = \frac{K}{N} \ln K + \omega \left(\frac{K}{N} \right), \quad (22)$$

where the notation $f(K) = \omega(g(K))$ is defined as $\lim_{K \rightarrow \infty} |f(K)/g(K)| = \infty$.

Proof: For the non-orthogonal OBF with N beams, let the event E denote a successful service that each of the K users is served at least once after D channel uses. Moreover, let \mathcal{N}_k , where $k = 1, \dots, K$, refer to the actual number of service for user k , that is, $\sum_{k=1}^K \mathcal{N}_k = DN$. Then, the probability that the event E occurs is given by

$$Pr\{E\} = Pr\{\mathcal{N}_1 > 0, \mathcal{N}_2 > 0, \dots, \mathcal{N}_K > 0\}$$

$$\begin{aligned}
&= 1 - Pr \left\{ \bigcup_{k=1}^K \mathcal{N}_k = 0 \right\} \\
&\geq 1 - K Pr \{ \mathcal{N}_k = 0 \} \\
&= 1 - K \left(1 - \frac{N}{K} \right)^D. \tag{23}
\end{aligned}$$

On the other hand, we have

$$\begin{aligned}
Pr\{E\} &= Pr \{ \mathcal{N}_1 > 0, \mathcal{N}_2 > 0, \dots, \mathcal{N}_K > 0 \} \\
&\leq \prod_{k=1}^K Pr \{ \mathcal{N}_k > 0 \} \\
&= \left[1 - \left(1 - \frac{N}{K} \right)^D \right]^K \\
&\leq \exp \left[-K \left(1 - \frac{N}{K} \right)^D \right], \tag{24}
\end{aligned}$$

where in (24) we used the inequality $1 - x \leq \exp(-x)$ for $x < 1$ [19, Eq.(4.5.7)]. Combining (23) with (24) yields $Pr\{E\} \rightarrow 1$ if $K \left(1 - \frac{N}{K} \right)^D \rightarrow 0$. To this end, when $K \rightarrow \infty$, we write

$$\begin{aligned}
K \left(1 - \frac{N}{K} \right)^D &= K \exp \left[D \ln \left(1 - \frac{N}{K} \right) \right] \\
&\sim \exp \left[-\frac{N}{K} \left(D - \frac{K}{N} \ln K \right) \right] \\
&\quad \times \left(1 + \mathcal{O} \left(\frac{N}{K} \right) \right). \tag{25}
\end{aligned}$$

Therefore, $K \left(1 - \frac{N}{K} \right)^D \rightarrow 0$ is equivalent to $\frac{N}{K} \left(D - \frac{K}{N} \ln K \right) \rightarrow \infty$, which yields the intended (22). ■

This theorem implies that, for the non-orthogonal OBF with N transmit beams, the minimum average scheduling delay over the users scales as $\frac{K}{N} \ln K$. In other words, the average scheduling delay improves with increasing N and, hence, non-orthogonal OBF with $N > N_t$ beams yields shorter scheduling delay than that of the orthogonal counterpart with $N = N_t$ beams. Furthermore, when $N = N_t$ or $N = 1$, (22) reduces to the existing results mentioned above.

C. Maximum Number of Supported Transmit Beams and Access Probability

In principle, there could be infinite number of transmit beams if no constraint is imposed on the received SINR of served users. In practice, however, if one user is to be successfully served, a minimum SINR or equivalently, error probability requirement has to be satisfied at the receiver side [27]. On the other hand, an outage probability threshold regarding the minimum SINR is demanded by the BS [28].

Assuming the thresholds of received SINR and outage probability are γ_{th} and P_{th} , respectively, in view of (8) and (10), the maximum number of transmit beams is the largest N satisfying $F_{\Gamma_{k,n}}^K(\gamma_{th}) < P_{th}$. After performing some algebraic manipulations, we obtain

$$N \leq \underbrace{\frac{\beta \gamma_{th} - (1 - \alpha \gamma_{th}) \ln \left(-\frac{1}{K} \ln P_{th} \right)}{\left(\beta + \frac{1}{\rho} \right) \gamma_{th}}}_{N_{max}}. \tag{26}$$

Moreover, once N is determined, the access probability of a given user among the total number of users K is expressed as [29]

$$P_{access} = \frac{N}{K} \left(1 - F_{\Gamma_{k,n}}^K(\gamma_{th}) \right) \geq \frac{N}{K} (1 - P_{th}). \tag{27}$$

In practice, P_{th} is usually very small, for instance 0.01. Thus, (27) implies that the access probability of each user increases linearly with the number of transmit beams N .

Notice that, the maximum number of supported transmit beams N_{max} in (26) depends not only on the users' requirements (γ_{th} and P_{th}) but also on the total number of users K . Clearly, larger K leads to larger N_{max} , which means more flexible choices of N . As mentioned above, for fixed N_t , orthogonal OBF aims to maximize sum-rate and it always generates $N = N_t$ beams regardless of the QoS requirements of users. However, for non-orthogonal OBF, the number of transmit beams N can be configured to satisfy QoS requirements of users and decrease the scheduling delay, which shows the flexibility of non-orthogonal OBF over the conventional orthogonal scheme. This will be further illustrated in Fig. 7 of Section VI.

V. PRACTICAL NON-ORTHOGONAL BEAMFORMING SCHEMES

In this section, two practical non-orthogonal beamforming schemes are illustrated. First, an optimality criterion on non-orthogonal beamforming is established. Then, two practical beamforming matrices are elaborated and they are demonstrated to achieve the optimality criterion. The first scheme is based on the optimal Grassmannian frame and applies only to certain values of N . The second scheme relies upon the Fourier matrix and applies to any N .

A. Optimality Criterion on Non-orthogonal Beamforming

In general, the $N_t \times N$ non-orthogonal beamforming matrix with $\delta_{l,n} = \delta_0$, where $1 \leq l \leq N$ and $l \neq n$, is equivalent to the optimal Grassmannian frame in the complex space \mathbb{C}^{N_t} [17]. The optimal Grassmannian frame is equiangular tight frame and its frame correlation achieves the lower bound provided by the Rankin inequality and, thus, $\delta_0 = \sqrt{\frac{N - N_t}{N_t(N - 1)}}$ [17]. Accordingly, for fixed N , the minimum total inter-beam interference is given by

$$\alpha_{min} = (N - 1) \times \left(\sqrt{\frac{N - N_t}{N_t(N - 1)}} \right)^2 = \frac{N}{N_t} - 1. \tag{28}$$

Applying (28) to *Theorem 1* yields the following corollary.

Corollary 1: For fixed N_t and N , when the inter-beam interference reaches the minimum $\frac{N}{N_t} - 1$, the sum-rate of non-orthogonal OBF scales as

$$R \sim N \ln \left(\frac{N}{N - N_t} \right). \tag{29}$$

Based on *Corollary 1*, it can be shown that the sum-rate monotonically decreases with increasing N for fixed N_t , but it increases with N_t for fixed N . More precisely, when N_t is fixed, the first-order derivative of R in (29) with respect to N is given by $\frac{dR}{dN} = \ln \left(\frac{N}{N - N_t} \right) - \frac{N_t}{N - N_t} < 0$, where we

exploited the property that $\ln x < x - 1$ whenever $x > 1$ [19, Eq.(4.5.4)]. On the other hand, when N is fixed and $N > N_t$, we have $\frac{dR}{dN_t} = \frac{N}{N-N_t} > 0$. An intuitive explanation of the monotonicity of R with respect to N and N_t is as follows. In the complex space \mathbb{C}^{N_t} , for fixed N_t , increasing N yields larger cross-correlation among N vectors. Nevertheless, if N is fixed, increasing N_t reduces the cross-correlation among N vectors.

In the following, we illustrate two non-orthogonal beamforming matrices which achieve the minimum total inter-beam interference in (28).

B. Grassmannian-Based Construction

Although the optimal Grassmannian frame achieves the minimum total inter-beam interference in (28), a general explicit construction of the optimal Grassmannian frame with $N > N_t + 1$ is almost impossible and it is usually obtained by exhaustive searching [23]. Fortunately, for wireless systems where the number of transmit antennas N_t is usually small for practical implementation purposes, an explicit construction of the optimal Grassmannian frame with, for example, $N_t = 2, 3, 4, 5, 6$, can be obtained by the following lemma.

Lemma 3: (König [24]) Let p be a prime number and l a positive integer such that $N_t = p^l + 1$ and $N = N_t^2 - N_t + 1$. Then, there exist integers $0 \leq d_1 < \dots < d_{N_t} < N$ with all numbers $1, \dots, N-1$ occurring as residues of $\text{mod}(d_i - d_q, N)$ where $i \neq q$ and $1 \leq i, q \leq N_t$, such that $\{\mathbf{b}_n\}_{n=1}^{N_t}$ with \mathbf{b}_n given by (30) forms an optimal Grassmannian frame:

$$\mathbf{b}_n = \frac{1}{\sqrt{N_t}} [e^{j2\pi nd_1/N} \quad e^{j2\pi nd_2/N} \quad \dots \quad e^{j2\pi nd_{N_t}/N}]^T. \quad (30)$$

As an application of *Lemma 3*, consider $N_t = 4$. In this case, we have $N = 13$. Then through simple searching we obtain $d_1 = 0, d_2 = 1, d_3 = 3, d_4 = 9$, and substituting them into (30) yields the beamforming matrix \mathbf{B}_G . Moreover, the cross-correlation between any two columns of \mathbf{B}_G is $\sqrt{3}/4$ and, thus, the total inter-beam interference for any beam is $\alpha_{\min} = 9/4$. Similarly, the non-orthogonal beamforming matrix for other values of N_t can be designed off-line, with the total inter-beam interference easily computed.

Note that, *Lemma 3* is not applicable for any number of transmit antennas, for instance, when $N_t = 7$. Furthermore, the optimal Grassmannian frame does not exist for arbitrary N_t and N [17]. Therefore, in the next sub-section, we propose another non-orthogonal beamforming which is based on the Fourier matrix. The proposed design is applicable for arbitrary N_t and N , and it is proven to obtain the minimum total inter-beam interference.

C. Fourier-Based Construction

For any N , it is well known that the Fourier matrix \mathbf{F} , whose $(p, q)^{\text{th}}$ entry is ω^{pq} with $\omega = \exp(-j2\pi/N)$ and $p, q = 0, 1, \dots, N-1$, is an orthonormal basis in a N -dimensional complex space. Its projection onto a N_t -dimensional complex space forms a tight frame whose elements have the broadest scattering [25]. This projection simply retains any N_t rows of \mathbf{F} . More specifically, assuming r_1, r_2, \dots, r_{N_t} rows of \mathbf{F} are selected, where $0 \leq r_1, r_2, \dots, r_{N_t} \leq$

$N-1$, we denote the resultant $N_t \times N$ beamforming matrix as $\mathbf{B}_F = \frac{1}{\sqrt{N_t}} \mathbf{F}_{N_t}$, where \mathbf{F}_{N_t} is the truncated version of \mathbf{F} with its r_1, r_2, \dots, r_{N_t} rows. The magnitude square of cross-correlation between l^{th} and n^{th} beamforming vectors is given by

$$\begin{aligned} \delta_{l,n}^2 &= \frac{1}{N_t^2} \left| \sum_{p \in \{r_1, \dots, r_{N_t}\}} \omega^{-pl} \omega^{pn} \right|^2 \\ &= \frac{1}{N_t} + \frac{1}{N_t^2} \sum_{\substack{p, q \in \{r_1, \dots, r_{N_t}\} \\ p \neq q}} \omega^{(p-q)(l+n)}. \end{aligned} \quad (31)$$

Clearly, (31) implies that the correlation coefficient $\delta_{l,n}$ depends upon the indices of selected rows from \mathbf{F} . In total, there are $\binom{N}{N_t} = \frac{N!}{N_t!(N-N_t)!}$ alternatives to select N_t rows out of \mathbf{F} . Therefore, the next question is which rows should be selected so as to obtain the minimum total inter-beam interference? The following theorem answers this question.

Theorem 3: For the $N \times N$ Fourier matrix, its any N_t rows can be selected as the beamforming matrix and each beam suffers the minimum total inter-beam interference $\frac{N}{N_t} - 1$.

Proof: According to (31), for beam n , the total inter-beam interference from the other $N-1$ beams is given by

$$\begin{aligned} \delta_n &= \sum_{l=0}^{N-1} \left[\frac{1}{N_t} + \frac{1}{N_t^2} \sum_{\substack{p, q \in \{r_1, \dots, r_{N_t}\} \\ p \neq q}} \omega^{(p-q)(l+n)} \right] - 1 \quad (32) \\ &= \frac{N}{N_t} + \frac{1}{N_t^2} \sum_{\substack{p, q \in \{r_1, \dots, r_{N_t}\} \\ p \neq q}} \omega^{n(p-q)} \sum_{l=0}^{N-1} \omega^{l(p-q)} - 1 \\ &= \frac{N}{N_t} - 1, \end{aligned} \quad (33)$$

where the constant 1 in the second term of the right-hand side of (32) denotes the square of the auto-correlation pertaining to the n^{th} column of \mathbf{B}_F , and we used the fact that $\sum_{l=0}^{N-1} \omega^{l(p-q)} = 0$ whenever $p \neq q$ to reach (33), since ω is the N^{th} root of unity [26, Eq.(5.8.2)]. Comparing (33) with (28), it is evident that the truncated Fourier matrix with any N_t rows achieves the minimum total inter-beam interference, which completes the proof. ■

Since the Fourier scheme achieves the minimum total inter-beam interference as the Grassmannian scheme when they have the same N_t and N , based on *Theorem 1*, both schemes achieve the same sum-rate as $K \rightarrow \infty$. This will be further illustrated in Fig. 8 in the next section.

VI. SIMULATION RESULTS AND DISCUSSIONS

In this section, numerical results based on the above analyses and Monte-Carlo simulation results are presented. All transmissions are over Rayleigh fading channels with zero mean and unit variance.

As justified from (3) to (7), the received SINR in (3) is well-approximated by (7) and then its closed-form distribution functions follow in (8) and (9). The accuracy of this approximation is also demonstrated in Fig. 2, where the simulated PDF of the received SINR in (3) is compared with

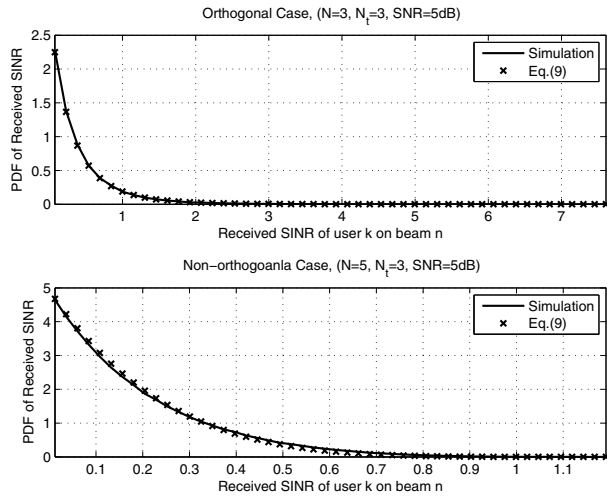


Fig. 2. PDF of the received SINR of user k on beam n with Fourier beamforming.

the analytical PDF in (9), based on the Fourier beamforming with $N_t = 3$ and $\text{SNR} = 5$ dB. From the upper panel, which corresponds to the orthogonal OBF with $N = N_t = 3$, it is seen that the numerical results of (9) coincide perfectly with the simulation results, since the PDF in (9) with $\alpha = 0$ and $\beta = 1$ is the exact PDF of the received SINR under orthogonal transmission. Furthermore, the lower panel of Fig. 2 shows that, for the non-orthogonal OBF with $N_t = 3$ and $N = 5$, the numerical results of (9) are almost indistinguishable from the simulation results.

The sum-rate of orthogonal and non-orthogonal transmissions is illustrated in Fig. 3, where the orthogonal OBF is implemented with $N = N_t = 3$ and the non-orthogonal scheme operates with $N = 5$ and $N_t = 3$. For the orthogonal case, it is observed that the numerical results of the integration in (14) coincides perfectly with the simulation results. Moreover, it is seen that the analytical expression (19) is very tight with (14), in which the tiny difference is due to the Jensen's loss introduced in (18). On the other hand, for the non-orthogonal transmission, the numerical results based on (14) are only slightly smaller (about 6%) than the simulation results, due to the approximation introduced in (7). The analytical expression (19) is still very tight with (14). Furthermore, as expected, Fig. 3 shows that non-orthogonal OBF underperforms the orthogonal case in terms of sum-rate, as inter-beam interference exists. However, non-orthogonal transmission benefits the scheduling delay, which is shown in Fig. 4. Figure 4 illustrates the simulation results of the average scheduling delay versus the number of users for different number of transmit beams $N = 3, 4, 5$, and 6 with $N_t = 3$. It is observed that, for fixed K , the orthogonal case with $N = N_t = 3$ has the longest scheduling delay, while the scheduling delay decreases with N . Moreover, as N increases, the scheduling-delay improvement from each of additional beam becomes smaller and smaller. These observations coincide with *Theorem 2*.

Figure 5 illustrates the tradeoff between sum-rate and scheduling delay, when considering $N_t = 3$ and $K = 50$,

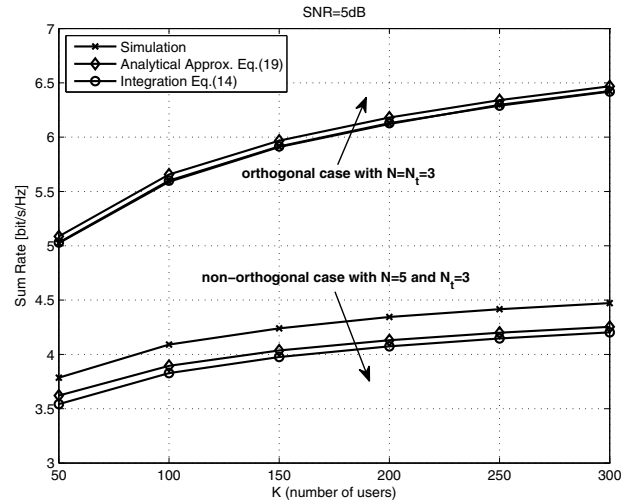


Fig. 3. Sum-rates of orthogonal and non-orthogonal OBFs with $N_t = 3$.

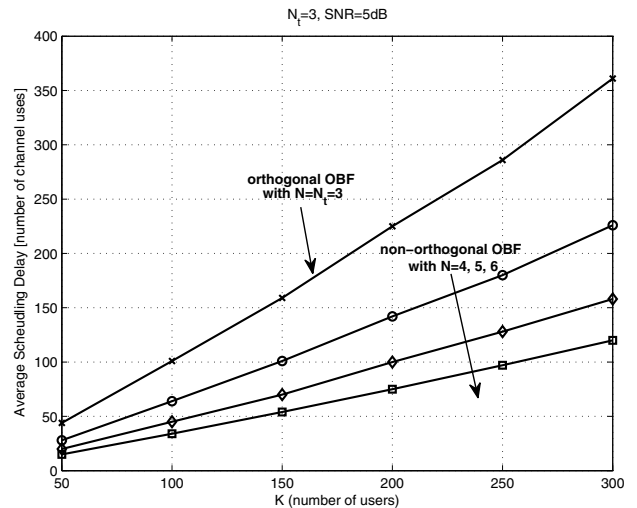


Fig. 4. Average scheduling delay $\mathcal{E}\{D_{1, K}\}$ versus the number of users K .

100, and 200. When K is fixed, the number of transmit beams varies from $N = N_t = 3$ (corresponding to the orthogonal scenario) to $N = 4, 5, 6$, and 7 (corresponding to non-orthogonal scenarios). It is observed that, for fixed K , the orthogonal case with $N = N_t$ yields the largest sum-rate but with the longest scheduling delay. Moreover, we can see that the scheduling delay decreases with increasing N at the cost of sum-rate. On the other hand, Fig. 6 shows the ratio of sum-rate to scheduling delay versus the number of beams N . It is seen that, for fixed number of users K , the ratio monotonically increases with N , which means that increasing N reduces the scheduling delay significantly more than the sum-rate.

Figure 7 shows the maximum number of supported transmitted beams, i.e., N_{\max} in (26), versus the threshold of the received SINR (γ_{th}), with $N_t = 4$ and outage probability $P_{\text{th}} = 0.01$. It is observed that, for non-orthogonal OBF with fixed number of users, N_{\max} increases as the threshold of the received SINR decreases, but the orthogonal case always transmits $N = N_t$ beams regardless of the users'

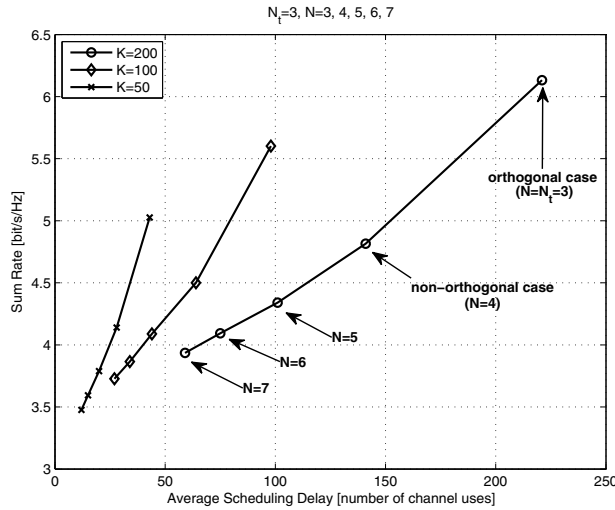


Fig. 5. Tradeoff between sum-rate and average scheduling delay.

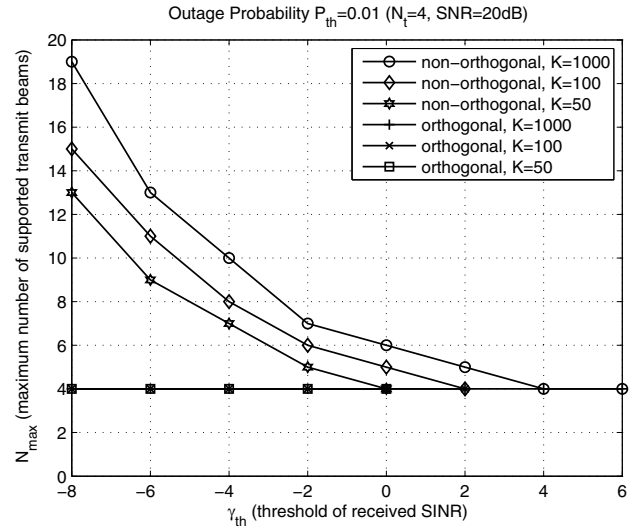


Fig. 7. Maximum number of supported transmit beams, i.e., N_{max} in (26), versus the threshold of received SINR.

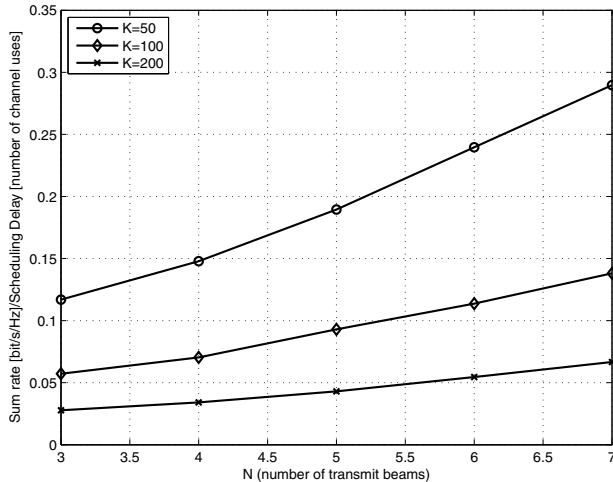


Fig. 6. The ratio of sum-rate to scheduling delay versus the number of transmit beams N .

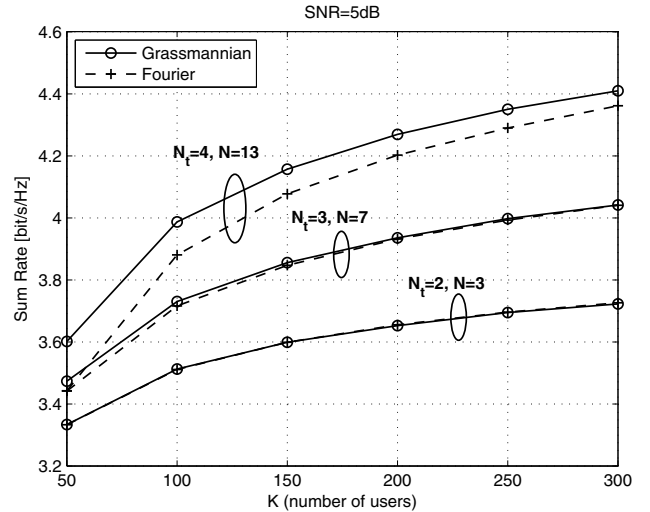


Fig. 8. Sum-rates of Grassmannian and Fourier schemes with same total inter-beam interference.

SINR requirements. Furthermore, for a fixed threshold of the received SINR, more users allows more transmit beams. Clearly, non-orthogonal OBF can be configured with different number of transmit beams to satisfy different received SINR requirements of users, which provides a flexibility over conventional orthogonal scheme.

Finally, according to the sum-rate scaling law in *Theorem 1*, different non-orthogonal beamforming schemes result in the same sum-rate as $K \rightarrow \infty$, if their total inter-beam interference for each beam are identical. When K is finite, this assertion is examined in Fig. 8, which depicts the simulation results of sum-rate of Grassmannian and Fourier schemes. It is observed that, when $N_t = 2$ and $N = 3$, the sum-rates of both schemes coincide perfectly in the entire range of K under consideration. When $N_t = 3$ and $N = 7$, both schemes obtain almost the same sum-rates when $K > 100$. When $N_t = 4$ and $N = 13$, the sum-rate of Grassmannian beamforming is higher than that of Fourier beamforming, however, their sum-

rate difference becomes smaller and smaller as K increases. Therefore, Grassmannian and Fourier beamforming results in the same sum-rate as $K \rightarrow \infty$, in accordance with *Theorem 1*.

VII. CONCLUSIONS

The system performance of non-orthogonal OBF with $N > N_t$ transmit beams were investigated in this paper. Our results show that non-orthogonal OBF is an interference-limited system. When the inter-beam interference attains the minimum for fixed N_t and N , the sum-rate decreases monotonically with the number of beams N , while the average scheduling delay improves monotonically with increasing N . Furthermore, two practical non-orthogonal beamforming matrices were proposed, and they were shown to achieve the minimum inter-beam interference and thus obtain the same asymptotic sum-rate. Due to the flexibility of the number of beams N related to the users' QoS requirements and the number of users

K , non-orthogonal OBF achieves the tradeoff between sum-rate and scheduling delay, and it is applicable to a wide range of wireless networks, especially densely populated networks with delay-sensitive traffic.

APPENDIX A PROOF OF LEMMA 1

According to (4), Z_l can be rewritten as $Z_l = \mathbf{h}_k^H (\mathbf{b}_n \mathbf{b}_{l,n}^{\perp H} + \mathbf{b}_{l,n}^{\perp} \mathbf{b}_n^H) \mathbf{h}_k$. Let $\mathbf{W} \triangleq \mathbf{b}_n \mathbf{b}_{l,n}^{\perp H} + \mathbf{b}_{l,n}^{\perp} \mathbf{b}_n^H = \mathbf{b}_n \mathbf{b}_{l,n}^{\perp H} + (\mathbf{b}_n \mathbf{b}_{l,n}^{\perp H})^H$, it is straightforward that \mathbf{W} is a Hermitian matrix with rank two. Since $\text{trace}(\mathbf{W}) = \text{trace}(\mathbf{b}_n \mathbf{b}_{l,n}^{\perp H}) + \text{trace}(\mathbf{b}_{l,n}^{\perp} \mathbf{b}_n^H) = 0$, we have $\text{trace}(\mathbf{W}) = \lambda_1 + \lambda_2 = 0$, where λ_1 and λ_2 denote the two real eigenvalues of \mathbf{W} . Thus, we obtain $\lambda_1 = -\lambda_2$. Furthermore, in the following, it is shown that $|\lambda_1| = |\lambda_2| = 1$.

Assuming that \mathbf{x}_i is an eigen-vector of \mathbf{W} with respect to λ_i , where $i = 1, 2$, we have

$$(\mathbf{b}_n \mathbf{b}_{l,n}^{\perp H} + \mathbf{b}_{l,n}^{\perp} \mathbf{b}_n^H) \mathbf{x}_i = \lambda_i \mathbf{x}_i. \quad (34)$$

Multiplying both sides of (34) with \mathbf{b}_n^H and in view of the facts that $\mathbf{b}_n^H \mathbf{b}_n = 1$ and $\mathbf{b}_n^H \mathbf{b}_{l,n}^{\perp} = 0$, we get $\mathbf{b}_{l,n}^{\perp H} \mathbf{x}_i = \lambda_i \mathbf{b}_n^H \mathbf{x}_i$. Similarly, multiplying both sides of (34) with $\mathbf{b}_{l,n}^{\perp H}$ yields $\mathbf{b}_n^H \mathbf{x}_i = \lambda_i \mathbf{b}_{l,n}^{\perp H} \mathbf{x}_i$. Substituting the former into the latter results in $\mathbf{b}_n^H \mathbf{x}_i = \lambda_i^2 \mathbf{b}_n^H \mathbf{x}_i$ and, finally, taking account of the fact that λ_i is a real number, we obtain that $\lambda_i \in \{\pm 1\}$, where $i = 1, 2$.

Applying the spectral decomposition theorem over \mathbf{W} , Z_l can be rewritten as

$$Z_l = \mathbf{h}_k \mathbf{U} \mathbf{D} \mathbf{U}^H \mathbf{h}_k^H = \sum_{i=1}^2 \lambda_i \underbrace{|\mathbf{h}_k^H \mathbf{u}_i|^2}_{E_i} \quad (35)$$

where $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_N]$ is a unitary matrix and \mathbf{D} is a diagonal matrix with the first two diagonal elements as λ_1 and λ_2 and other elements being zero. Since \mathbf{h}_k is a complex Gaussian random vector, E_i , $i = 1, 2$ in (35), are two independent exponentially-distributed variables. Therefore, no matter $(\lambda_1 = 1, \lambda_2 = -1)$ or $(\lambda_1 = -1, \lambda_2 = 1)$ holds, Z_l is of Laplace distribution with zero mean and variance two.

Furthermore, it is clear that the Laplace transform of $Z \triangleq \sum_{l=1, l \neq n}^N Z_l$ is given by $1/(1 - s^2)^{(N-1)}$. Accordingly, exploiting the inverse Laplace transform [20, vol.5, Eq.(2.1.2.72)] and performing some algebraic manipulations, we get the PDF of Z as shown in (6).

APPENDIX B PROOF OF LEMMA 2

We exploit the von Mises criteria [21, Theorem 2.7.2] to determine the limiting distribution of $\hat{\gamma}_{\max, n}$. Accordingly, two conditions need to be satisfied. The first one requires that the PDF $f_{\Gamma_{k,n}}(\gamma)$ in (8) is differentiable for all $\gamma \in (0, 1/\alpha)$, where $1/\alpha$ is the supremum of $\gamma_{k,n}$ in (7). This is obviously true. Furthermore, substituting (8) and (9) into the function $g(\gamma) = (1 - F_{\Gamma_{k,n}}(\gamma))/f_{\Gamma_{k,n}}(\gamma)$ and performing some algebraic manipulations, it can be shown that $\lim_{\gamma \rightarrow 1/\alpha} \frac{dg(\gamma)}{d\gamma} = 0$, which is the second condition. Consequently, the limiting

CDF of $\hat{\gamma}_{\max, n}$ is of the Gumbel distribution [21, p.54], as shown in (15). Moreover, the normalizing parameter a is the solution to the equation $1 - F_{\Gamma_{k,n}}(a) = 1/K$ [21, Theorem 2.1.3]. Substituting (9) into the preceding equation yields a transcendental equation:

$$\exp\left(-\frac{Na}{\rho(1-a\alpha)}\right) \left(1 + \frac{a(1-\delta_0^2)}{1-a\alpha}\right)^{-(N-1)} = \frac{1}{K}. \quad (36)$$

Now, taking the logarithm on both sides of (36), applying the first-order Taylor approximation of $\ln(1+x)$ with small x , and performing further derivations, a can be given by (16). Once a is determined, the value of b is given by $b = g(a)$ [21, Remark 2.7.1], which results in the desired (17).

REFERENCES

- [1] M. Sharif and B. Hassibi, "Delay considerations for opportunistic scheduling in broadcast fading channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 9, pp. 3353–3363, Sep. 2007.
- [2] G. Aniba and S. Aïssa, "Adaptive scheduling for MIMO wireless networks: cross-layer approach and application to HSDPA," *IEEE Trans. Wireless Commun.*, vol. 6, no. 1, pp. 259–268, Jan. 2007.
- [3] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1277–1294, June 2002.
- [4] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channels with partial side information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 506–522, Feb. 2005.
- [5] M. Xia, W. Wen, and S.-C. Kim, "Opportunistic cophasing transmission in MISO systems," *IEEE Trans. Commun.*, vol. 57, no. 12, pp. 3764–3770, Dec. 2009.
- [6] M. Agrawal and A. Puri, "Base station scheduling of requests with fixed deadlines," in *Proc. 2002 IEEE INFOCOM*, pp. 487–496.
- [7] B. T. Cadmus, B. Hohensee, A. C. Accardo, and D. Perez, "Platform alarming through multiple agencies," *IEEE Trans. Industry Applications*, vol. 43, no. 1, pp. 172–179, Jan.-Feb. 2007.
- [8] T. Heikkinen and A. Hottinen, "Delay-differentiated scheduling in a fading channel," *IEEE Trans. Wireless Commun.*, vol. 7, no. 3, pp. 848–856, Mar. 2008.
- [9] D. Shah, D. N. C. Tse, and J. N. Tsitsiklis, "Hardness of low delay network scheduling," in *Proc. 2010 IEEE Information Theory Workshop*, pp. 1–13.
- [10] G. R. Gupta and N. B. Shroff, "Delay analysis and optimality of scheduling policies for multihop wireless networks," *IEEE Trans. Networking*, vol. 19, no. 1, pp. 129–141, Jan. 2011.
- [11] E. A. Jorswieck, A. Sezgin, and X. Zhang, "Throughput versus fairness: channel-aware scheduling in multiple antenna downlink," *EURASIP J. Wireless Commun. and Networking*, vol. 2009, article ID: 271540, pp. 1–13, 2009.
- [12] N. Zorba and A. I. Pérez-Neira, "A multiple user opportunistic scheme: the Grassmannian approach," in *Proc. 2006 Int. Zürich Seminar Commun.*, pp. 134–137, Feb. 2006.
- [13] N. Zorba and A. I. Pérez-Neira, "Opportunistic Grassmannian beamforming for multiuser and multiantenna downlink communications," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1174–1178, Apr. 2008.
- [14] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc.: packing in Grassmannian spaces," *Experimental Math.*, vol. 5, no. 2, pp. 139–159, 1996.
- [15] P. Svedman, E. A. Jorswieck, and B. Ottersten, "Reduced feedback SDMA based on subspace packings," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1329–1339, Mar. 2009.
- [16] E. A. Jorswieck, P. Svedman, and B. Ottersten, "Performance of TDMA and SDMA based opportunistic beamforming," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4058–4063, Nov. 2008.
- [17] T. Strohmer and R. W. Heath, "Grassmannian frames with applications to coding and communication," *Applied and Computational Harmonic Analysis*, vol. 14, no. 3, pp. 257–275, May 2003.
- [18] M. Xia, Y. Zhou, J. Ha, and H. K. Chung, "Opportunistic beamforming communication with throughput analysis using asymptotic approach," *IEEE Trans. Veh. Technol.*, vol. 58, no. 5, pp. 2608–2614, June 2009.
- [19] F. W. J. Olver, *NIST Handbook of Mathematical Functions*. Cambridge University Press, 2010.
- [20] A. P. Prudnikov, Y. A. Brychkov, and O. I. Marichev, *Integrals and Series*. Gordon and Breach Science Publishers, 1986.

- [21] J. Galambos, *The Asymptotic Theory of Extreme Order Statistics*, 2nd edition. Robert E. Krieger Publishing Co., 1987.
- [22] A. Bayesteh, M. A. Sadrabadi, and A. K. Khandani, "Throughput and fairness maximization in wireless downlink systems," Technical Report, University of Waterloo, 2007. Available: http://shannon2.uwaterloo.ca/t/2007-23_Bayesteh.pdf.
- [23] N. Sloane, *How to Pack Lines, Planes, 3-Spaces, Etc.* Available: <http://www2.research.att.com/~njas/grass/>.
- [24] H. König, "Cubature formulas on spheres," *Advances in Multivariate Approximation (Witten-Bommerholz, 1998)*, vol. 107 of *Math. Res.*, pp. 201–211. Wiley-VCH, 1999.
- [25] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellation," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 1962–1973, Sep. 2000.
- [26] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. SIAM, 2000.
- [27] C. Anton-Haro, P. Svedman, M. Bengtsson, A. Alexiou, and A. Gameiro, "Cross-layer scheduling for multi-user MIMO systems," *IEEE Commun. Mag.*, vol. 44, no. 9, pp. 39–45, Sep. 2006.
- [28] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.
- [29] M. J. Hossain, V. K. Bhargava, and M.-S. Alouini, "Rate adaptive hierarchical modulation-assisted two-user opportunistic scheduling," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2076–2085, Jan. 2007.



Minghua Xia obtained his Ph.D. degree in Telecommunications and Information Systems from Sun Yat-sen University, Guangzhou, China, in 2007. From Mar. 2007 to July 2009, he was with the Electronics and Telecommunications Research Institute (ETRI) of South Korea, Beijing R&D Center, Beijing, China, where he worked as a member of engineering staff and participated in the projects on the physical layer design of 3GPP LTE mobile communications. From Aug. 2009 to Feb. 2011, he was with The University of Hong Kong (HKU), Hong Kong, as

a Postdoctoral Fellow. Currently, he is with King Abdullah University of Science and Technology (KAUST), Saudi Arabia. His research interests are in the area of network information theory and space-time signal processing, and in particular the design and performance analysis of multi-user multi-antenna systems, cooperative relaying systems, and cognitive radio networks.



Yik-Chung Wu received the B.Eng. (EEE) degree in 1998 and the M.Phil. degree in 2001 from the University of Hong Kong (HKU). He received the Ph.D. degree from Texas A&M University, College Station, in 2005.

From August 2005 to August 2006, he was with the Thomson Corporate Research, Princeton, NJ, as a Member of Technical Staff. Since September 2006, he has been with the HKU as an Assistant Professor. He was a visiting scholar at Princeton University, Princeton, in summer 2011. His research interests

are in general area of signal processing and communication systems. He is currently serving as an associate editor for the IEEE COMMUNICATIONS LETTERS.



Sonia Aïssa (S'93-M'00-SM'03) received her Ph.D. degree in Electrical and Computer Engineering from McGill University, Montreal, QC, Canada, in 1998. Since then, she has been with the National Institute of Scientific Research-Energy, Materials, and Telecommunications (INRS-EMT), University of Quebec, Montreal, QC, Canada, where she is a Professor.

From 1996 to 1997, she was a Researcher with the Department of Electronics and Communications of Kyoto University, Kyoto, Japan, and with the Wireless Systems Laboratories of NTT, Kanagawa, Japan. From 1998 to 2000, she was a Research Associate at INRS-EMT, Montreal. From 2000 to 2002, while she was an Assistant Professor, she was a Principal Investigator in the major program of personal and mobile communications of the Canadian Institute for Telecommunications Research (CITR), leading research in radio resource management for code division multiple access systems. From 2004 to 2007, she was an Adjunct Professor with Concordia University, Montreal. In 2006, she was Visiting Invited Professor with the Graduate School of Informatics, Kyoto University, Kyoto, Japan. Her research interests lie in the area of wireless and mobile communications, and include radio resource management, cross-layer design and optimization, design and analysis of multiple antenna (MIMO) systems, cognitive and cooperative transmission techniques, and performance evaluation, with a focus on Cellular, Ad Hoc, and Cognitive Radio networks.

Dr. Aïssa was the Founding Chair of the Montreal Chapter IEEE Women in Engineering Society in 2004-2007, acted or is currently acting as Technical Program Leading Chair or Cochair for the Wireless Communications Symposium of the IEEE International Conference on Communications (ICC) in 2006, 2009, 2011 and 2012, as PHY/MAC Program Chair for the 2007 IEEE Wireless Communications and Networking Conference (WCNC), and as Technical Program Committee Cochair of the 2013 IEEE Vehicular Technology Conference - spring (VTC). She has served as a Guest Editor of the *EURASIP Journal on Wireless Communications and Networking* in 2006, and as Associate Editor of the *IEEE Wireless Communications Magazine* in 2006-2010. She is currently an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON COMMUNICATIONS and the *IEEE Communications Magazine*, and Associate Editor of the *Wiley Security and Communication Networks Journal*. Awards and distinctions to her credit include the Quebec Government FQRNT Strategic Fellowship for Professors-Researchers in 2001-2006; the INRS-EMT Performance Award in 2004 and 2011 for outstanding achievements in research, teaching and service; the IEEE Communications Society Certificate of Appreciation in 2006-2011; and the Technical Community Service Award from the FQRNT Center for Advanced Systems and Technologies in Communications (SYTACom) in 2007. She is also co-recipient of Best Paper Awards from IEEE ISCC 2009, WPMC 2010, IEEE WCNC 2010 and IEEE ICCIT 2011; and recipient of NSERC (Natural Sciences and Engineering Research Council of Canada) Discovery Accelerator Supplement Award.