Beamforming with no instantaneous feedback for mmWave transmission

Yuan-Pei Lin and Shang-Ho (Lawrence) Tsai
Dept. Electrical Engr., National Chiao Tung Univ., Hsinchu, Taiwan

***Abstract*—In this paper we consider statistical beamforming for millimeter wave (mmWave) channels. For a commonly used geometry channel model of uniform planar array (UPA), we show that the average channel Gram matrix can be approximated in a closed form, which allows the statistical beamformer to be computed directly. The sparse nature of the mmWave channel lends itself to statistical beamforming as the optimal beamformer that uses full channel information is typically steered in some spatial frequency determined by the statistics. Simulations are given to show that the performance of the proposed statistical beamformer that needs no instantaneous feedback is comparable to a limited feedback system with 1/8 feedback bits per antenna.**

I. Introduction

The performance of a MIMO system is known to improve with the number of antennas on the two transmission ends. Recent advances show that it is feasible to pack a large number of antennas in a small area, particularly in millimeter wave communication systems that use small wavelengths [1]. For frequency division duplex systems, the transmitter relies on the receiver to feedback the channel information. As the size of the channel increases with the number of antennas, the feedback of the channel information becomes a more daunting task.

Statistical beamforming and precoding [2]-[4], designed based on channel statistics, requires only infrequent update of channel statistics but not instantaneous feedback. Optimal beamforming for maximizing the average capacity of Rayleigh fading channels is designed in [2]. Optimization of the precoder for minimum pairwise error probability is consider in [3] and the optimal precoder that minimizes the sum of mean squared error in [4]. These earlier works, based on Kronecker Rayleigh channels, can not be used for mmWave channels as Kronecker model is not valid in this case [5]. Long-term beamforming is considered for multicasting application using gradient projection algorithm in [6]. The statistics of the vector received at the basestation is exploited in [7] to design the basestation beamformer for users with a single antenna.

Although a large number of antenna may be employed in a mmWave system, cost and power constraints often prohibit having one dedicated RF (radio frequency) chain for each antenna [1]. The beamforming coefficients are typically implemented using analog phase shifters, thus the coefficients are usually constrained to be of unit modulus. More recently, it is shown in [8]-[10] that if two phase shifters are used to implement each coefficient we can obtain an arbitrary beamformer. Thus the design of RF beamformer or precoder can be free from unit modulus constraint.

In this paper we consider statistical beamforming for mmWave channels. The transmitter knows only the statistics of the channel but not instantaneous channel information. For the commonly used geometry based clustered channel model, we derive the statistical beamformer from the average channel Gram matrix for UPA. It turns out that the average channel Gram matrix depends only on the statistics on the transmitter side. When there are multiple clusters in the channel, the 2D magnitude response of the beamformer shows that the power is allocated in an uneven manner. The clusters with angle of departures away from the normal of the antenna plane are allocated more power; power allocation is done over the spatial frequency in this case. The proposed statistical beamformer, though designed without unit modulus constraint, has a small degradation of around 0.4dB when each is implemented as the sum of two phase shifters of three-bit resolution. Simulations show that the performance of the proposed statistical beamformer is comparable to that of a limited feedback system with 1/8 instantaneous feedback bits per antenna.

Notation. The variance of a random variable x is denoted as σ_x^2 and the expectation of x by $E[x]$. The 2-norm of a vector $\mathbf{f}$ is denoted as $||\mathbf{f}||$. The notation $\mathbf{A}^\dagger$ denotes the transpose and conjugate of a matrix $\mathbf{A}$.

II. System Model

Consider a MIMO channel with N_t transmit antennas and N_r receive antennas, represented by an $N_r \times N_t$ matrix $\mathbf{H}$. We adopt the geometric channel representation that is useful for modelling mmwave propagation [11][12],

$$\mathbf{H} = \frac{1}{\sqrt{L}} \sum_{\ell=1}^{L} \alpha_\ell \mathbf{a}_r(\phi^r_{a,\ell}, \phi^r_{e,\ell}) \mathbf{a}_t^\dagger(\phi^t_{a,\ell}, \phi^t_{e,\ell}), \quad (1)$$

where L is the number of rays in the channel, and α_ℓ, denoting the complex gain of the ℓth rays, are assumed to be independent Gaussian random variables of zero mean. The

This work was supported by Ministry of Science and Technology, Taiwan, R. O. C., under MOST 101-2221-E-009 -091 -MY3.

978-1-5090-3009-5/17/$31.00 ©2017 IEEE

azimuth (elevation) angle[1] of departure (AoD) $\phi^t_{a,\ell}$ ($\phi^t_{e,\ell}$) and the azimuth (elevation) angle of arrival (AoA) $\phi^r_{a,\ell}$ ($\phi^r_{e,\ell}$) are of the truncated Laplacian distribution, which is demonstrated in [13][11] to be a good model in this case. The angle of departure $\phi^t_{a,\ell}$ ($\phi^t_{e,\ell}$) are independent and of mean $\overline{\phi}^t_{a,\ell}$ ($\overline{\phi}^t_{e,\ell}$). The standard deviations of angles of departure in azimuth and elevation, also called angular spreads, are denoted, respectively, by $\sigma_{\phi^t_{a,\ell}}$ and $\sigma_{\phi^t_{e,\ell}}$. When the channel is clustered [12], the rays in the same cluster are assumed to be of the same variance and mean. The vectors $\mathbf{a}^\dagger_t(\phi^t_{a,\ell}, \phi^t_{e,\ell})$ and $\mathbf{a}_r(\phi^r_{a,\ell}, \phi^r_{e,\ell})$ are, respectively, the transmit and receive antenna array response vectors. The array response vector for a uniform planar array arranged on the yz-plane with size $N_z \times N_y$ (N_z in the z-direction and N_y in the y-direction) is given by

$$[\mathbf{a}(\phi_a, \phi_e)]_{m+nN_z} = \frac{1}{\sqrt{N_y N_z}} e^{j\xi(m\cos(\phi_e)+n\sin(\phi_e)\sin(\phi_a))}, \tag{2}$$

for $0 \leq m < N_z$ and $0 \leq n < N_y$, where $\xi = 2\pi d$ and d is the antenna spacing normalized by the wavelength. The channel in (1) can be written in a matrix form as

$$\mathbf{H} = \mathbf{A}_r \mathbf{D}_\alpha \mathbf{A}^\dagger_t, \tag{3}$$

where $\mathbf{A}_r$ is the $N_r \times L$ matrix whose column vectors are the L receive antenna array response vectors

$$\mathbf{a}_r(\phi^r_{a,1}, \phi^r_{e,1}), \mathbf{a}_r(\phi^r_{a,2}, \phi^r_{e,2}), \cdots \mathbf{a}_r(\phi^r_{a,L}, \phi^r_{e,L})$$

whereas $\mathbf{A}_t$ is the $N_t \times L$ matrix whose column vectors are the L transmit antenna array response vectors. The matrix $\mathbf{D}_\alpha$, of size $L \times L$, is diagonal with diagonal elements $\alpha_1, \alpha_2, \cdots, \alpha_L$.

Let the transmitted symbol be s and the transmit beamformer be $\mathbf{f}$, then the transmission power is $P_t = ||\mathbf{f}||^2\sigma^2_s$. The output of the receiver is $r = \mathbf{gHs} + \mathbf{gn}$, where $\mathbf{g}$ is the $1 \times N_r$ receive combiner, $\mathbf{n}$ is the $N_r \times 1$ channel noise vector. Assuming the elements of the channel noise vector $\mathbf{n}$ are independent, of variance N_0 and zero mean, the resulting SNR is

$$SNR = \frac{|\mathbf{gHf}|^2}{||\mathbf{g}||^2||\mathbf{f}||^2}\frac{P_t}{N_0}. \tag{4}$$

III. Statistical beamforming

Assume the transmitter knows only the statistics and the beamformer $\mathbf{f}$ is designed based only on the statistics. The receiver has the channel information and knows the statistical beamformer used at the transmitter. In this case, the optimal combiner is $\mathbf{g} = (\mathbf{Hf})^\dagger$ and the resulting SNR is $SNR = \frac{P_t}{N_0}||\mathbf{Hf}||^2/||\mathbf{f}||^2$. Averaging the SNR over the random channel, we get $E[SNR] = \frac{P_t}{N_0}E[||\mathbf{Hf}||^2/||\mathbf{f}||^2]$. To maximize the average SNR, we design the statistical beamformer to solve the following problem:

$$\max_{\mathbf{f} \text{ s.t. } ||\mathbf{f}||=1} \mathbf{f}^\dagger \mathbf{B} \mathbf{f}, \quad \text{where } \mathbf{B} = E[\mathbf{H}^\dagger\mathbf{H}].$$

The optimal beamformer is the unit eigen vector of the average channel Gram matrix $\mathbf{B}$ that corresponds to the largest eigen value. In what follows we show that for the channel model given in (1), the matrix $\mathbf{B}$ can be approximated in a closed form when the angle spread is small.

Lemma 1: When the angle of departures $\{\phi^t_{a,\ell}\}$ and $\{\phi^t_{e,\ell}\}$ and the complex gains $\{\alpha_\ell\}$ are independent, the average channel Gram matrix $\mathbf{B} = E[\mathbf{H}^\dagger\mathbf{H}]$ is given by

$$\mathbf{B} = E[\mathbf{A}_t \mathbf{\Lambda} \mathbf{A}^\dagger_t], \tag{5}$$

where $\mathbf{\Lambda} = \text{diag}\{\sigma^2_{\alpha_1}, \sigma^2_{\alpha_1}, \cdots \sigma^2_{\alpha_{N_{ray}}}\}$ and in particular

$$[\mathbf{B}]_{m+nN_z, \ell+kN_z} = \sum_{i=1}^{L} \sigma^2_{\alpha_i}\mu_i(m-\ell, n-k), \tag{6}$$

for $0 \leq m, \ell < N_z$ and $0 \leq n, k < N_y$, where $\mu_i(m,n) = E[e^{j\xi(m\cos\phi^t_{e,i}+n\sin\phi^t_{e,i}\sin\phi^t_{a,i})}]$.

See Appendix A for a proof. The function $\mu_i(m,n)$ depends on the statistics of the AoD. When the angular spreads of these angles are small, $\mu_i(m,n)$ can be approximated as follows.

Lemma 2: When angular spreads are small at the transmitter side, $\mu_i(m,n)$ can be approximated as

$$\begin{aligned} \mu_i(m,n) \;\approx\; & e^{j\xi(m\cos\overline{\phi}^t_{e,i}+n\sin\overline{\phi}^t_{e,i}\sin\overline{\phi}^t_{a,i})} \\ \times\; & p(-m\sin\overline{\phi}^t_{e,i} + n\cos\overline{\phi}^t_{e,i}\sin\overline{\phi}^t_{a,i}, \Delta\phi^t_{e,i}) \\ \times\; & p(n\sin\overline{\phi}^t_{e,i}\cos\overline{\phi}^t_{a,i}, \Delta\phi^t_{a,i}) \end{aligned} \tag{7}$$

where $\Delta\phi^t_{e,i} = \phi^t_{e,i} - \overline{\phi}^t_{e,i}$, $\Delta\phi^t_{a,i} = \phi^t_{a,i} - \overline{\phi}^t_{a,i}$, and $p(\beta, x) = E[e^{j\xi\beta x}]$.

A proof is given in Appendix B. Note that both $\Delta\phi^t_{e,i}$ and $\Delta\phi^t_{a,i}$ are truncated Laplacian random variables with zero mean. For a random variable x obtained by truncating a Laplacian random variable with zero mean and variance σ for the range $[-\pi, \pi)$, the pdf is $1/(\sqrt{2}c_0\sigma)e^{-\sqrt{2}|x|/\sigma}$. In this case we can verify that

$$p(\beta, x) = \frac{1}{c_0(1+\frac{\beta^2\sigma^2}{2})}\left[1 + e^{-\frac{\pi\sqrt{2}}{\sigma}}\left(\frac{\beta\sigma}{\sqrt{2}}\sin(\beta\pi) - \cos(\beta\pi)\right)\right], \tag{8}$$

where $c_0 = 1 - e^{-j\pi\sqrt{2}/\sigma}$ is a constant. Combining the above two lemmas, we arrive at the following theorem.

Theorem 1: When the angular spreads are small at the transmitter side, the average channel Gram matrix can be approximated by

$$\begin{aligned} &[\mathbf{B}]_{m+nN_z, \ell+kN_z} \\ &\approx \sum_{i=1}^{L} \sigma^2_{\alpha_i} e^{j\xi((m-\ell)\cos\overline{\phi}^t_{e,i}+(n-k)\sin\overline{\phi}^t_{e,i}\sin\overline{\phi}^t_{a,i})} \\ &\times p(-(m-\ell)\sin\overline{\phi}^t_{e,i} + (n-k)\cos\overline{\phi}^t_{e,i}\sin\overline{\phi}^t_{a,i}, \Delta\phi^t_{e,i}) \end{aligned}$$

[1]The elevation angle of a ray is the angle between the ray and the z-axis whereas the azimuth angle is the angle between the x-axis and the orthogonal projection of the ray on the xy-plane.

$$\times p((n-k)\sin\overline{\phi}^t_{e,i}\cos\overline{\phi}^t_{a,i}, \Delta\phi^t_{a,i}), \quad (9)$$

for $0 \le m, \ell < N_z$ and $0 \le n, k < N_y$, where $p(\beta, x)$ is as given in (8)

With the above closed-form approximation of $\mathbf{B}$, the beamforming vector can be computed accordingly. Although the above derivation of $\mathbf{B}$ requires the assumption that the angular spreads are small, simulation examples will be given to demonstrate that the approximation is accurate and the resulting beamformer useful even for larger angular spread. From Lemma 1, we see that the optimal statistical beamformer depends only on the statistics on the transmitter side, but not those on the receiver side. Therefore the statistical beamformer can be determined by the statistics of the AoD. To understand the above results better, we look into the beamformer design from the frequency domain point of view.

Magnitude response of the beamformer. In the beamforming system, each antenna sends out the symbol that is weighted by a coefficient in the beamforming vector $\mathbf{f}$. Let us re-index the beamforming coefficients as an array $f_{m,n}$, for $0 \le m < N_z$ and $0 \le m < N_y$, such that the coefficients are arranged according to the planar antenna array on the yz-plane, with m corresponding to the index in the z-axis and n corresponding to the y-axis. Let

$$F(\omega_1, \omega_2) = \sum_{m=0}^{N_z-1}\sum_{n=0}^{N_y-1} f_{m,n} e^{-j(m\omega_1 + n\omega_2)}$$

be the Fourier transform for the two-dimensional sequence $f_{m,n}$. The expression of $\mathbf{B}$ given in (5) means that the average SNR is of the form $E[SNR] = E[\mathbf{f}^\dagger \mathbf{A}_t \mathbf{\Lambda} \mathbf{A}_t^\dagger \mathbf{f}]$, which can be written as $E[SNR] = E[||\mathbf{\Lambda}^{1/2}\mathbf{A}_t^\dagger \mathbf{f}||^2]$. Observe that the i-th element of the vector $\mathbf{A}_t^\dagger \mathbf{f}$ is equal to

$$\frac{1}{\sqrt{N_y N_z}} \sum_{m=0}^{N_z-1}\sum_{n=0}^{N_y-1} f_{m,n} e^{-j\xi(m\cos(\phi^t_{e,i}) + n\sin(\phi^t_{e,i})\sin(\phi^t_{a,i}))},$$

$$= F(\xi\cos(\phi^t_{e,i}), \xi\sin(\phi^t_{e,i})\sin(\phi^t_{a,i})).$$

Therefore we have

$$E[SNR] = E\left[\sum_{i=1}^{L} \sigma^2_{\alpha_i} |F(\xi\cos(\phi^t_{e,i}), \xi\sin(\phi^t_{e,i})\sin(\phi^t_{a,i}))|^2\right].$$

This is the weighted sum of the mean square of the magnitude response $|F(\omega_1, \omega_2)|$ evaluated at frequencies corresponding to the AoD, i.e., $(\xi\cos(\phi^t_{e,i}), \xi\sin(\phi^t_{e,i})\sin(\phi^t_{a,i}))$. This means that the design of the optimal statistical beamformer is to maximize the magnitude response $F(\omega_1, \omega_2)$ at frequencies determined from the AoD, irrespective of the AoA. We will see in simulations that $|F(\omega_1, \omega_2)|$ has peaks at frequencies corresponding to the means of the AoD.

Uniform linear array (ULA). When $N_y = 1$, the antennas form a uniform linear array along the z-axis. The above results are valid for ULA as well. In this case, only the elevation AoD is relevant and (6) reduces to $[\mathbf{B}]_{m,\ell} = \sum_{i=1}^{L} \sigma^2_{\alpha_i} E[e^{j\xi((m-\ell)\cos\phi^t_{e,i}}]$. We have

$$[\mathbf{B}]_{m,\ell} \approx \sum_{i=1}^{L} \sigma^2_{\alpha_i} e^{j\xi((m-\ell)\cos\overline{\phi}^t_{e,i}} p(-(m-\ell)\sin\overline{\phi}^t_{e,i}, \Delta\phi^t_{e,i}),$$

where $p(\beta, x)$ is as given in (8). Let us define the Fourier transform of the beamformer $F(\omega) = \sum_{m=0}^{N_t-1} f_m e^{-jm\omega}$, where f_m denotes the m-th coefficient of $\mathbf{f}$. Then $E[SNR] = E\left[\sum_{i=1}^{L}\sigma^2_{\alpha_i}|F(\xi\cos(\phi^t_{e,i}))|^2\right]$. That is, maximizing the average SNR is in effect maximizing the weighted sum of mean squares of $|F(\omega)|$ at $\xi\cos(\phi^t_{e,i})$, a result similar to the UPA case.

IV. Simulation Examples

Consider the geometry channel model in (1) with 64 transmit antennas (arranged in 8×8) and 16 receive antennas (arranged in 4×4). We assume there are three clusters and each has ten rays. The antenna spacing is half wavelength, thus $d = 1/2$. The complex gains $\{\alpha_\ell\}$are assumed to Gaussian random variables of zero mean and unity variance. We have used 10^5 channels in the simulations.

Example 1. Consider the case when the cluster means are fixed. The means of the AoD in elevation and azimuth $(\bar{\phi}^t_e, \bar{\phi}^t_a)$ for the three clusters are respectively $(0.3\pi, 0.2\pi)$, $(0.1\pi, 0.48\pi)$ and $(0.55\pi, 0.1\pi)$. Paths in the same cluster are of the same mean. The angular spreads of AoD and AoA in azimuth and elevation are 8° for each cluster. Fig. 1(a) shows the magnitude response $|F(\omega_1, \omega_2)|$ as a function of ω_1 and ω_2, where we have marked the cluster number at frequencies corresponding to the mean of AoD, i.e., $(\xi\cos(\bar{\phi}^t_{e,\ell}), \xi\sin(\bar{\phi}^t_{e,\ell})\sin(\bar{\phi}^t_{a,\ell}))$. In Fig. 1(b) the magnitude response $|F(\xi\cos(\phi_e), \xi\sin(\phi_e)\sin(\phi_a))|$ is shown as a function of ϕ_e and ϕ_a. We see that the peaks of the magnitude response $|F(\omega_1, \omega_2)|$ agree with the marked frequencies. Somewhat surprisingly, $|F(\omega_1, \omega_2)|$ is not of similar height in these frequencies. From Fig. 1(b), we observe the largest peak at the second mean $(0.1\pi, 0.48\pi)$, a smaller peak at the first mean $(0.3\pi, 0.2\pi)$ and a even smaller one at the third $(0.55\pi, 0.1\pi)$. The largest peak occurs at the frequency that corresponds to the second cluster, whose AoD is more away from the normal of the plane of the antenna. That is, more power is allocated to the spatial frequency away from the normal of the antenna plane.

Example 2. In this example, we evaluate the performance of the statistical beamformer. The means of the AoD in elevation and azimuth are uniformly distributed over $[0, 2\pi]$. Fig. 2 shows the average of the SNR in (4) as a function of the angular spread of AoD for $P_t/N_0 = 0$ dB. Consider the performance of the statistical beamformer (labeled as 'stat') obtained through the approximation in Theorem 1 and the optimal one (labeled as 'optimal stat') that is obtained by numerically

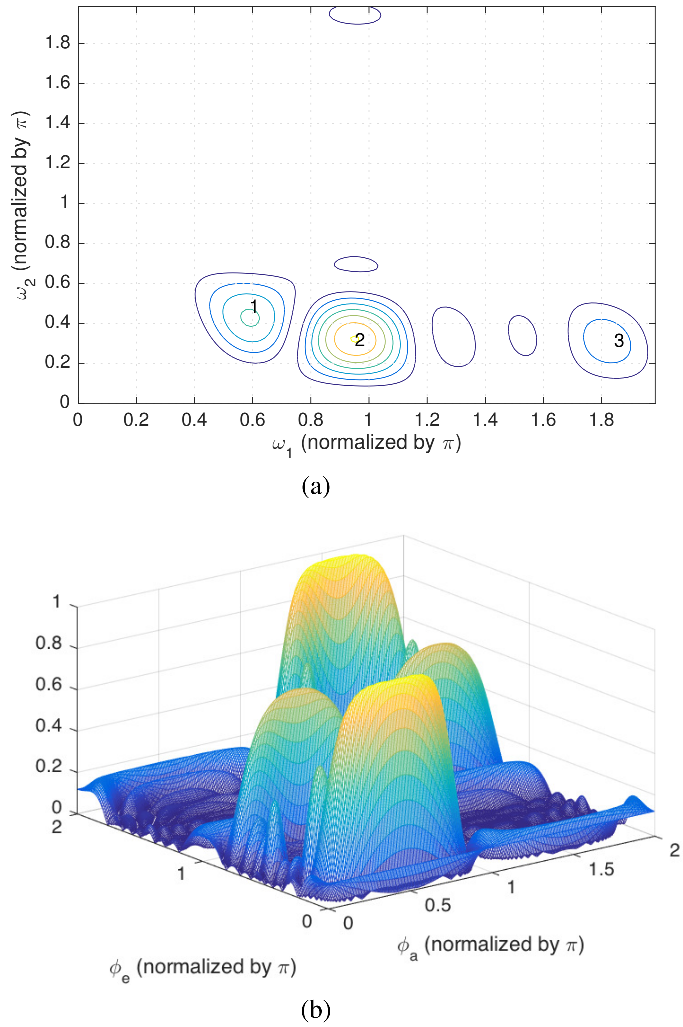

Fig. 1. Magnitude response of the beamformer; (a) $|F(\omega_1,\omega_2)|$; (b) $|F(\xi\cos(\phi_e^t),\xi\sin(\phi_e^t)\sin(\phi_a^t))|$ as a function of ϕ_e and ϕ_a.

evaluating the average of the 10^5 channel Gram matrices used in the simulation. The two curves are close, especially for angular spread smaller than $5°$. The formula in Theorem 1, though obtained under small angular spread assumption, is useful even for moderate spread. As a bench mark, we also show in the figure the case when the transmitter has full CSI and employs the optimal unconstrained beamformer. The difference is around 2.3 dB for a small angular spread and increases to around 3.5 dB when the angular spread is $10°$. Also shown in Fig. 2 is the average SNR of the statistical beamformer when the coefficients are quantized due to the use of finite resolution phase shifters. Each beamforming weight is implemented as the sum of two phase shifters that are of three-bit resolution [8], [10]. We see that the degradation due to quantization is around 0.4 dB. Fig. 3 shows the transmission rate as a function of P_t/N_0 when the angular spread is $5°$. For comparison, we have shown the rate of sparse precoding and combining (SPC) [14] for two cases of feedback bits, 8 and 10 bits. The beamforming vector therein is in the form of an array response vector and only the elevation and azimuth angles need to be fed back to the transmitter. The phase shifters are also quantized using three bits. We see that the performance of the quantized beamformer is comparable to SPC with 8 feedback bits, i.e, 1/8 feedback bits per antenna.

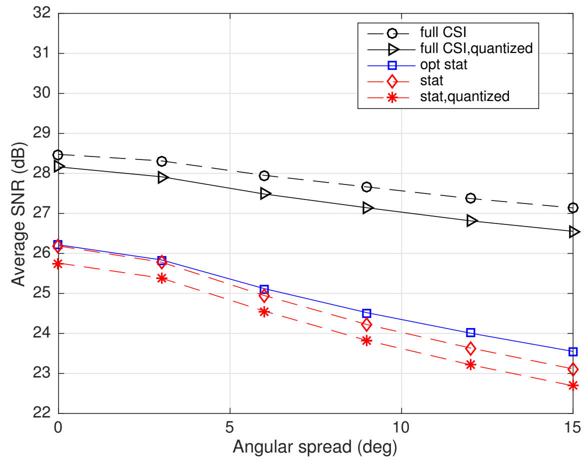

Fig. 2. Average SNR vs. angular spread

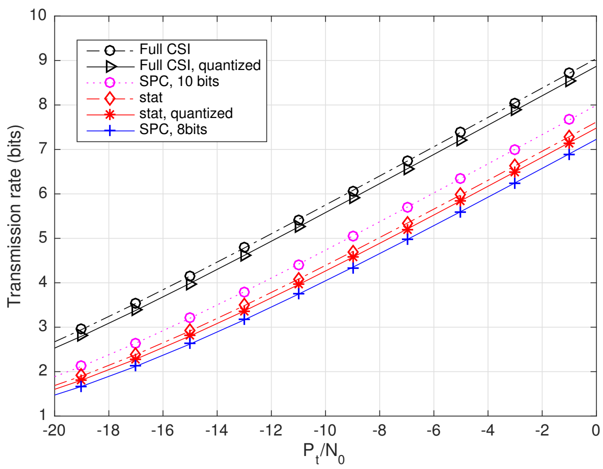

Fig. 3. Transmission rate performance.

Appendix A. Proof of Lemma 1.

Using (3), we can express $\mathbf{B}$ as $\mathbf{B}=E_{\mathbf{H}}[\mathbf{A}_t\mathbf{D}^\dagger\mathbf{A}_r^\dagger\mathbf{A}_r\mathbf{D}\mathbf{A}_t^\dagger]$, where the subscript of the expectation means the average is performed over $\mathbf{H}$. As the AoD, AoA and the complex gains $\{\alpha_\ell\}$ are independent we get $\mathbf{B}=E_{\{\phi_{a,\ell}^t,\phi_{e,\ell}^t\}}[\mathbf{A}_t\mathbf{\Lambda}\mathbf{A}_t^\dagger]$, where $\mathbf{\Lambda}$ is given by $E_{\{\alpha_\ell\}}[\mathbf{D}_\alpha^\dagger E_{\{\phi_{a,\ell}^r,\phi_{e,\ell}^r\}}[\mathbf{A}_r^\dagger\mathbf{A}_r]\mathbf{D}_\alpha]$. Notice that $\{\alpha_\ell\}$ are independent and of zero mean, so $\mathbf{\Lambda}$ is a diagonal matrix. Only the diagonal elements of the matrix $E_{\{\phi_{a,\ell}^r,\phi_{e,\ell}^r\}}[\mathbf{A}_r^\dagger\mathbf{A}_r]$ matter and the diagonal elements are all equal to 1 because the product $\mathbf{a}_r^\dagger(\phi_{a,\ell}^r,\phi_{e,\ell}^r)\mathbf{a}_r(\phi_{a,\ell}^r,\phi_{e,\ell}^r)=$

1. Thus we have $\mathbf{\Lambda} = \text{diag}\{\sigma_{\alpha_1}^2, \sigma_{\alpha_1}^2, \cdots \sigma_{\alpha_{N_{ray}}}^2\}$. Using the expression of the antenna response vector in (2), we can obtain the elements of $\mathbf{B}$ as given in (6).

Appendix B. Proof of Lemma 2.

With the definition of $\Delta\phi_{e,i}^t = \phi_{e,i}^t - \overline{\phi}_{e,i}^t$, we can write the term $\cos\phi_{e,i}^t$ as $\cos(\overline{\phi}_{e,i}^t)\cos(\Delta\phi_{e,i}^t) - \sin(\overline{\phi}_{e,i}^t)\sin(\Delta\phi_{e,i}^t)$. When the angular spread is small, $\Delta\phi_{e,i}^t$ is small and we have $\cos(\Delta\phi_{e,i}^t) \approx 1$ and $\sin(\Delta\phi_{e,i}^t) \approx \Delta\phi_{e,i}^t$. So $\cos\phi_{e,i}^t \approx \cos(\overline{\phi}_{e,i}^t) - \sin(\overline{\phi}_{e,i}^t)\Delta\phi_{e,i}^t$. Similarly, we have $\sin\phi_{e,i}^t \approx \sin(\overline{\phi}_{e,i}^t) + \cos(\overline{\phi}_{e,i}^t)\Delta\phi_{e,i}^t$. Thus

$$\sin\phi_{e,i}^t \sin\phi_{a,i}^t \approx \sin(\overline{\phi}_{e,i}^t)\sin(\overline{\phi}_{a,i}^t)$$
$$+\cos(\overline{\phi}_{e,i}^t)\sin(\overline{\phi}_{a,i}^t)\Delta\phi_{e,i}^t + \sin(\overline{\phi}_{e,i}^t)\cos(\overline{\phi}_{a,i}^t)\Delta\phi_{a,i}^t,$$

where we have ignore the second order term of $\Delta\phi_{e,i}^t\Delta\phi_{a,i}^t$. Pulling out the constant, we get

$$\begin{aligned}\mu_i(m,n) &\approx e^{j\xi(m\cos\overline{\phi}_{e,i}^t + n\sin\overline{\phi}_{e,i}^t\sin\overline{\phi}_{a,i}^t)}\\ &\times E[e^{j\xi(-m\sin\overline{\phi}_{e,i}^t + n\cos\overline{\phi}_{e,i}^t\sin\overline{\phi}_{a,i}^t)\Delta\phi_{e,i}^t}]\\ &\times E[e^{j\xi n\sin\overline{\phi}_{e,i}^t\cos\overline{\phi}_{a,i}^t\Delta\phi_{a,i}^t}].\end{aligned}$$

It can be further expressed as in (7).

References

[1] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 101-107, June 2011.

[2] S. A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Trans. Wireless Communi.*, vol. 3, no. 4, pp. 1165-1175, July 2004.

[3] Y. Li, H. Liu, and L. Jiang, "Statistical precoder design for spatial correlated MIMO channels in V-BLAST systems," *International Conference on Neural Networks and Signal Processing*, 2008.

[4] T. Liu, J.-K. Zhang, and K. M. Wong, "Optimal precoder design for correlated MIMO communication systems using zero-forcing decision feedback equalization," *IEEE Trans. Signal Processing*, vol. 57, no. 9, pp. 3600-3612, Sept. 2009.

[5] D. Ying, F. W. Vook, T. A. Thomas, D. J. Love and A. Ghosh, "Kronecker product correlation model and limited feedback codebook design in a 3D channel model," IEEE International Conference on Communications, 2014.

[6] A. Lozano, "Long-Term Transmit Beamforming for Wireless Multicasting," IEEE International Conference on Acoustics, Speech and Signal Processing, 2007.

[7] G. G. Raleigh, S. N. Diggavi, V. K. Jones and A. Paulraj, "A blind adaptive transmit antenna algorithm for wireless communication," IEEE International Conference on Communications, 1995.

[8] E. Zhang and C. Huang, "On achieving optimal rate of digital precoder by RF-baseband codesign for MIMO systems," Proceedings of IEEE Vehicular Technology Conference, Sept. 2014

[9] T. E. Bogale, L. B. Le, A. Haghighat, and L. Vandendorpe, "On the number of RF chains and phase shifters, and scheduling design with hybrid analog-digital beamforming," IEEE Trans. Wireless Communications, vol. 15, no. 5, pp. 3311-3326, May 2016.

[10] Yuan-Pei Lin and Shang-Ho Tsai, "THIC structures for RF precoding in mmwave communications," submitted to 2017 IEEE International Workshop on Signal Processing Advances in Wireless Communications.

[11] V. Erceg, et al., "TGn channel models," IEEE 802.11-03/940r4, May 2004.

[12] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter Wave Channel Modeling and Cellular Capacity Evaluation," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1164 - 1179, June 2014.

[13] K. I. Pedersen, P. E. Mogensen, and B. H. Fleury, "The Power azimuth spectrum in outdoor environments," *IEE Electronics Letters*, vol. 33, no. 18, pp. 1583-1584, Aug. 1997.

[14] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath Jr, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 99, pp. 1-15, Jan. 2014.