THE CAPACITY OPTIMALITY OF BEAM STEERING IN LARGE MILLIMETER WAVE MIMO SYSTEMS

Omar El Ayach, Robert W. Heath, Jr.[*]

The University of Texas at Austin
Austin, TX, 78712

Shadi Abu-Surra, Sridhar Rajagopal, Zhouyue Pi

Dallas Technology Lab, Samsung Electronics
Richardson, TX, 75082

ABSTRACT

Millimeter wave (mmWave) systems must overcome the heavy attenuation at high frequency to support high-throughput wireless communication. The small wavelength in mmWave systems enables beamforming using *large antenna arrays* to combat path loss with large array gain. Beamforming in traditional microwave systems is often done at baseband for maximum flexibility. Such baseband processing requires a dedicated transceiver chain per antenna element. The high cost of radio frequency (RF) chains in mmWave systems, however, makes supporting each antenna with a dedicated RF chain expensive. This mismatch between the number of antennas and transceiver chains makes baseband processing infeasible; thus mmWave systems typically rely on a traditional approach known as beam steering which can be done at RF using inexpensive phase shifters. Unlike baseband precoding, however, traditional beam steering is not explicitly designed to achieve the capacity of the mmWave channel. In this paper, we consider both beamforming and multi-stream precoding in single user systems with large mmWave antenna arrays at both transmitter and receiver. Using a realistic channel model, we show that the unconstrained capacity-achieving precoding solutions converge to simple beam steering solutions. Therefore, in large mmWave systems, no rate loss is incurred by adopting the traditional lower-complexity solution.

1. INTRODUCTION

Millimeter wave (mmWave) communication has enabled gigabit per second data rates in indoor wireless systems [1, 2] and has been recently proposed for outdoor cellular deployments [3]. Millimeter wave's high carrier frequency facilitates packing many antenna elements in small form factors, thus enabling multiple-input multiple-output (MIMO) processing with *very large arrays*.

Beamforming in mmWave systems is essential for achieving gigabit per second data rates by counteracting high path loss with high array gain [3]. Precoding multiple data streams could also allow users to approach capacity and has been shown to be both feasible and beneficial in mmWave systems [4]. Precoding for mmWave systems, however, differs from traditional MIMO precoding in a variety of ways. First, MIMO processing often assumes hardware complexity that is impractical in mmWave devices, such as a dedicated transceiver chain per antenna element. Since transceiver chains are limited due to their high cost [3], most precoding must be done after upconversion using analog phase shifters. This system architecture places constant modulus constraints on the used precoders. To address this problem, antenna selection [5] and equal gain transmission [6–8] have been proposed. Second, MIMO precoding analysis often assumes rich fading environments, e.g. Rayleigh fading. However, the high pathloss at mmWave frequencies significantly limits scattering. The combination of *limited scattering* and *large tightly packed antenna arrays* makes idealized fading distributions unrealistically rich for mmWave systems. Realistic models such as [9, 10] have been proposed and used for capacity analysis [11], though their implications on precoder design is not yet understood.

In this paper, we focus on the *insight into constrained single user precoding* that can be derived from explicitly accounting for the characteristics of *large antenna arrays* in *limited scattering environments*. To do so, we adopt a realistic finite ray channel model based on the extended Saleh Valenzuela model [9,10] which captures both the limited scattering at high frequency and the antenna correlation in tightly packed arrays. In this environment, we show that the unconstrained capacity-achieving MIMO precoders, based on the channel's singular value decomposition (SVD), converge in the limit of large arrays to traditional beam steering, a well known RF beamforming strategy that is not explicitly designed to approach capacity. Thus we demonstrate that optimal precoders naturally satisfy RF precoder constraints, and low-complexity hardware incurs no loss in data rate performance. In addition to proving that beam steering is conveniently optimal, our results shed light onto the proper design of training and feedback in mmWave systems. For example, the derived results imply that the optimal precoders have very few free variables, i.e., the beam steering angles. Thus, in limited feedback systems, beam steering optimality obviates the need for intricate vector codebooks, typically used in MIMO limited feedback [12], and replaces it with simple scalar feedback. Finally, we present simulation results to verify the optimality of beam steering in large systems and demonstrate its performance for realistic array sizes.

We use the following notation: $\mathbf{A}$ is a matrix; $\mathbf{a}$ is a vector; a is a scalar; $\mathbf{A}^{(i)}$ is the i^{th} column of $\mathbf{A}$ and a_i is the i^{th} element of $\mathbf{a}$; $\mathbf{A}^{(-i)}$ is the matrix $\mathbf{A}$ with the i^{th} column removed; $(.)^T$ and $(.)^*$ denote transpose and conjugate transpose respectively; $\|\mathbf{A}\|_F$ is the Frobenius norm of $\mathbf{A}$ and $|\mathbf{A}|$ is its determinant; $\|\mathbf{a}\|$ is the 2-norm of $\mathbf{a}$; $\mathbf{A} \otimes \mathbf{B}$ is the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$; $[\mathbf{A} \mid \mathbf{B}]$ denotes horizontal concatenation; $\mathrm{diag}(\mathbf{a})$ is a matrix with the elements of $\mathbf{a}$ on its diagonal; $\mathbf{I}_N$ and $\mathbf{0}_{N\times N}$ are the $N \times N$ identity and zero matrices respectively; $\mathcal{CN}(\mathbf{a}; \mathbf{A})$ is a complex Gaussian vector with mean $\mathbf{a}$ and covariance matrix $\mathbf{A}$.

2. SYSTEM MODEL

In this section we present the signal and channel model for a single user system with large antenna arrays and limited RF chains.

[*]This work was done while the first author was with the Dallas Technology Lab at Samsung Electronics. The authors at the University of Texas at Austin were supported in part by the Office of Naval Research (ONR) under grant N000141010337.

978-1-4673-0971-4/12/$31.00 © 2012 IEEE

2.1. System Model

Consider the single user mmWave system shown in Fig. 1 in which a base station with N_t antennas communicates N_s data streams to a user with N_r antennas [13]. The transmitter is equipped with N_{RF}^t RF chains such that $N_s \leq N_{RF}^t \leq N_t$. Using the N_{RF}^t transmit chains, the transmitter can apply a "small" $N_{RF}^t \times N_s$ baseband precoder, $\mathbf{F}_{BB}$, followed by a "large" $N_t \times N_{RF}^t$ RF precoder, $\mathbf{F}_{RF}$. Therefore, the sampled transmitted signal is

$$\mathbf{x} = \mathbf{F}_{RF}\mathbf{F}_{BB}\mathbf{s},$$

where $\mathbf{s}$ is the $N_s \times 1$ symbol vector such that $\mathbb{E}\left[\mathbf{s}\mathbf{s}^*\right] = \frac{1}{N_s}\mathbf{I}_{N_s}$. Since the RF precoder is implemented using analog phase shifters, its elements are constrained to satisfy $(\mathbf{F}_{RF}^{(i)}\mathbf{F}_{RF}^{(i)*})_{\ell,\ell} = N_t^{-1}$, where $(.)_{\ell,\ell}$ denotes the ℓ^{th} diagonal element, i.e., all elements of $\mathbf{F}_{RF}$ have equal norm. A total power constraint is enforced by setting $\|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2 = N_s$; no other constraints are placed on $\mathbf{F}_{BB}$.

We consider a narrowband block fading propagation channel which yields a received signal

$$\mathbf{y} = \sqrt{\rho}\mathbf{H}\mathbf{F}_{RF}\mathbf{F}_{BB}\mathbf{s} + \mathbf{n}, \tag{1}$$

where $\mathbf{y}$ is the $N_r \times 1$ received vector, $\mathbf{H}$ is the $N_r \times N_t$ channel matrix such that $\mathbb{E}\left[\|\mathbf{H}\|_F^2\right] = N_t N_r$, ρ represents the average received power, and $\mathbf{n}$ is the vector of i.i.d $\mathcal{CN}(0, \sigma_n^2)$ noise. In writing (1), we implicitly assume perfect timing and frequency recovery. Throughout this paper, we assume that $\mathbf{H}$ is known to both transmitter and receiver. Efficiently designing training and feedback to obtain accurate channel knowledge in mmWave systems is left for future work; training and feedback design can in fact benefit from the results shown under the perfect channel knowledge assumption.

At the receiver, $N_{RF}^r \geq N_s$ RF chains are used to receive the N_s data streams. The processed signal is

$$\widetilde{\mathbf{y}} = \sqrt{\rho}\mathbf{W}_{BB}^*\mathbf{W}_{RF}^*\mathbf{H}\mathbf{F}_{RF}\mathbf{F}_{BB}\mathbf{s} + \mathbf{W}_{BB}^*\mathbf{W}_{RF}^*\mathbf{n}, \tag{2}$$

where $\mathbf{W}_{RF}$ is the $N_r \times N_{RF}^r$ RF combining matrix with unit norm entries, and $\mathbf{W}_{BB}$ is the $N_{RF}^r \times N_s$ baseband combining matrix. The achieved rate assuming Gaussian signaling is

$$R = \log_2\left(\left|\ \mathbf{I}_{N_s} + \frac{\rho}{N_s}\mathbf{R}_n^{-1}\mathbf{W}_{BB}^*\mathbf{W}_{RF}^*\mathbf{H}\mathbf{F}_{RF}\mathbf{F}_{BB}\right.\right.$$
$$\left.\left.\mathbf{F}_{BB}^*\mathbf{F}_{RF}^*\mathbf{H}^*\mathbf{W}_{RF}\mathbf{W}_{BB}\ \right|\right),$$

where $\mathbf{R}_n = \sigma_n^2\mathbf{W}_{BB}^*\mathbf{W}_{RF}^*\mathbf{W}_{RF}\mathbf{W}_{BB}$ is the covariance matrix of the colored Gaussian noise at the output of the baseband combiner.

2.2. Channel Model

Analytical models such as Rayleigh fading represent scattering levels that are too rich for mmWave channels. In this paper, we consider a clustered channel model based on the extended Saleh Valenzuela model that is often used in mmWave channel modeling [9] and standardization [10]. For simplicity of exposition, each scattering cluster is assumed to contribute a single propagation path [11]. This assumption can be relaxed to account for clusters with a finite angular spread and the derived results can be readily extended. Using this model, the channel is given by

$$\mathbf{H} = \sqrt{\frac{N_t N_r}{L}}\sum_{\ell=1}^{L}\alpha_\ell \mathbf{a}_r(\phi_\ell^r,\ \theta_\ell^r)\mathbf{a}_t(\phi_\ell^t,\ \theta_\ell^t)^*. \tag{3}$$

where L is the number of propagation paths, α_ℓ is the "complex gain" of the ℓ^{th} ray, and ϕ_ℓ^r (θ_ℓ^r) and ϕ_ℓ^t (θ_ℓ^t) are its random azimuth

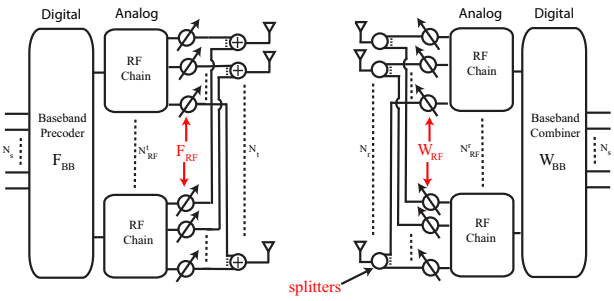

Fig. 1. Simplified hardware block diagram of mmWave single user system with digital baseband precoding followed by constrained radio frequency precoding implemented using RF phase shifters.

(elevation) angles of arrival and departure respectively. The complex gains α_ℓ are assumed to be $\mathcal{CN}(0,1)$ in simulation [9]. Note that the independence of α_ℓ's from ϕ_ℓ^r, θ_ℓ^r, ϕ_ℓ^t and θ_ℓ^t implies the use of omnidirectional antenna elements. Directional antenna elements can be considered by letting $\alpha_\ell \sim \mathcal{CN}(0, \Lambda^r(\phi_\ell^r, \theta_\ell^r)\Lambda^t(\phi_\ell^t, \theta_\ell^t))$ where $\Lambda(.)$ represents the directional element gain. The vectors $\mathbf{a}_r(\phi_\ell^r, \theta_\ell^r)$ and $\mathbf{a}_t(\phi_\ell^t, \theta_\ell^t)$ are the normalized receive/transmit array response vectors at the corresponding angles of arrival/departure.

For an N-element uniform linear array (ULA) on the y-axis, the array response vector is [14]

$$\mathbf{a}^{ULAy}(\phi) = \frac{1}{\sqrt{N}}\left[1,\ e^{jkd\sin(\phi)},\ \ldots,\ e^{j(N-1)kd\sin(\phi)}\right]^T \tag{4}$$

where $k = \frac{2\pi}{\lambda}$ and d is the inter-element spacing. We do not include θ in the arguments of $\mathbf{a}^{ULAy}$ as the response for such a ULA is independent of the elevation angle. All results naturally carry over to arrays on the z or x axes; considering the y-axis is without loss of generality. Since small form factors may make large linear arrays impractical, we also consider uniform planar arrays (UPA) which yield realistic antenna dimensions and enable beamforming in elevation. For a uniform planar array in the yz-plane with W and H elements on the y and z axes respectively, the array response vector is given by [14]

$$\mathbf{a}^{UPA}(\phi,\theta) = \frac{1}{\sqrt{N}}\left(\left[1,\ \ldots,\ e^{jkd(m\sin(\phi)sin(\theta)+n\cos(\theta))},\ \ldots,\right.\right.$$
$$\left.\left.\ldots,\ e^{jkd((W-1)\sin(\phi)sin(\theta)+(H-1)\cos(\theta))}\right]\right)^T \tag{5}$$

where $0 < m < W-1$ and $0 < n < H-1$ are the y and z indexes of an antenna element respectively and $N = WH$.

3. SINGLE USER BEAMFORMING IN LARGE MMWAVE SYSTEMS

For the case of single stream transmission, designing transmit and receive filters to maximize data rate simplifies to maximizing effective SNR. This can be written as

$$\left(\mathbf{w}^{opt}, \mathbf{f}^{opt}\right) = \arg\max |\mathbf{w}^*\mathbf{H}\mathbf{f}|^2 \tag{6}$$
$$s.t.\ \mathbf{w} \in \mathcal{W},\ \mathbf{f} \in \mathcal{F}$$

where $\mathcal{W}$ and $\mathcal{F}$ are the feasible sets of transmit and receive vectors which reflect the beamforming constraints imposed by the system's hardware.

In beamforming systems that can control both the amplitude and phase of a signal, the only real beamforming constraint is a total power constraint. This translates into defining the feasible sets as $\mathcal{W} = \{\mathbf{w} \mid \|\mathbf{w}\|^2 = 1\}$ and $\mathcal{F} = \{\mathbf{f} \mid \|\mathbf{f}\|^2 = 1\}$. It is well known that this least-constrained version admits a simple solution based on singular value decomposition. Letting the channel's singular value decomposition be $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$, the optimal vectors are $\mathbf{w}^{opt} = \mathbf{U}^{(1)}$ and $\mathbf{f}^{opt} = \mathbf{V}^{(1)}$.

3.1. Steering Vector Beamforming and Combining

The solutions generated by SVD decomposition will in general require beamforming vectors to control both the signal's phase and amplitude [7]. Since beamforming is done at RF using phase shifters, amplitude control is infeasible (especially when $N_{RF}^t = N_{RF}^r = 1$). So, mmWave systems are constrained to using vectors with constant gain elements.

To maintain well behaved beam patterns, additional constraints are often imposed on the RF phase shifters. In such cases, systems are often limited to steering their signal in a single physical direction. In this case the optimization problem is given by

$$\begin{aligned}\left(\mathbf{w}^{opt}, \mathbf{f}^{opt}\right) &= \arg\max |\mathbf{w}^* \mathbf{H} \mathbf{f}|^2 \\ s.t.\ w_i &= \sqrt{N_r^{-1}} e^{j f_r(\phi_r, \theta_r, i)}\ \forall i, \\ f_\ell &= \sqrt{N_t^{-1}} e^{j f_t(\phi_t, \theta_t, \ell)}\ \forall \ell\end{aligned} \tag{7}$$

where $f_r(\phi_r, \theta_r, i)$ and $f_t(\phi_t, \theta_t, \ell)$ are linear functions designed to steer the array in a predefined direction. For example, for a receive uniform linear array steered in azimuth to an angle ϕ_r the phase function is $f_r(\phi_r, \theta_r, i) = ikd\sin(\phi_r)$.

To the best of our knowledge, no analytical results are available for the above problem in general and beamforming algorithms are limited to searches over a fixed number of predefined directions. Moreover, the performance loss due to this highly constrained yet highly practical beamforming solution has not been quantified. In what follows, we show that for the realistic channel model of Section 2.2, such beam steering becomes optimal for large N_t, i.e., the additional hardware constraints cause no performance degradation. This result is both theoretically and practically promising since mmWave systems often employ large antenna arrays. The proof of this asymptotic optimality is based mainly on the following two properties of the ULA and UPA response vectors in poor scattering channels.

Lemma 1 *For a ULA system with azimuth angles of arrival or departure drawn independently from a continuous distribution, the transmit and receive array response vectors are orthogonal, i.e., we have* $\mathbf{a}^{ULAy}(\phi_k) \perp \text{span}(\{\mathbf{a}^{ULAy}(\phi_\ell) \mid \forall \ell \neq k\})$ *as the number of antenna elements, N, tends to infinity and the number of paths in the channel is $L = o(N)$.*

Proof: This is given in the appendix. □

The same is true for a uniform planar array for paths arriving or departing at random azimuth and elevation angles. This is given in the following corollary.

Corollary 2 *For a UPA with azimuth and elevation angles of arrival or departure drawn independently from a continuous distribution, the array response vectors are orthogonal. That is* $\mathbf{a}^{UPA}(\phi_k, \theta_k) \perp \text{span}(\{\mathbf{a}^{UPA}(\phi_\ell, \theta_\ell) \mid \forall \ell \neq k\})$*, as the number of antenna elements, N, tends to infinity and the number of paths in the channel is $L = o(N)$.*

Proof: This is given in the appendix. □

As a result of Lemma 1 and Corollary 2, we have the following theorem relating the *array response vectors* to the *singular vectors* of the channel matrix $\mathbf{H}$, regardless of array structure.

Theorem 3 *Each right singular vector of the matrix channel $\mathbf{H}$ given by the model in (3) with $L = o(N_t)$ and $L = o(N_r)$ converges in chordal distance to an array response vector $\mathbf{a}_t(\phi_\ell^t, \theta_\ell^t)$. Each left singular vector similarly converges to $\mathbf{a}_r(\phi_\ell^r, \theta_\ell^r)$. The singular values in turn converge to $\frac{N_t N_r}{L} |\alpha_\ell|^2$.*

Proof: Consider the channel model in (3) in which $\mathbf{H}$ is naturally decomposed into a summation of L rank-one contributions each corresponding to a propagation ray. The channel can be written in matrix form as

$$\mathbf{H} = \sqrt{\frac{N_t N_r}{L}} \sum_{\ell=1}^{L} \alpha_\ell \mathbf{a}_r(\phi_\ell^r,\ \theta_\ell^r) \mathbf{a}_t(\phi_\ell^t,\ \theta_\ell^t)^* = \mathbf{A}_r \mathbf{D} \mathbf{A}_t^* \tag{8}$$

where we define the matrix $\mathbf{A}_r = [\mathbf{a}_r(\phi_1^r,\ \theta_1^r),\ \ldots,\ \mathbf{a}_r(\phi_L^r, \theta_L^r)]$, and $\mathbf{A}_t = \left[\mathbf{a}_t(\phi_1^t, \theta_1^t),\ \ldots,\ \mathbf{a}_t(\phi_L^t, \theta_L^t)\right]$, and the diagonal matrix $\mathbf{D} = \text{diag}([\alpha_1 \sqrt{N_t N_r / L},\ \ldots,\ \alpha_L \sqrt{N_t N_r / L}])$. To see the direct link between (8) and the channel's SVD, we make two simple modifications to the channel representation in (8); (i) we "extract" the phases of the matrix $\mathbf{D}$ and apply them to $\mathbf{A}_t$, (ii) we extend (8) by including a basis for the channel's nullspace. Effectively, we note that (8) can be rewritten as

$$\mathbf{H} = \left[\mathbf{A}_r \mid \mathbf{A}_r^\perp\right] \widetilde{\mathbf{D}} \left[\text{diag}([e^{j\angle\alpha_1},\ \ldots,\ e^{j\angle\alpha_L}]) \mathbf{A}_t \mid \mathbf{A}_t^\perp\right]^* \tag{9}$$

where $\mathbf{A}_t^\perp$ and $\mathbf{A}_r^\perp$ are any unitary matrices spanning the nullspace of $\mathbf{H}$ and $\mathbf{H}^T$ respectively[1], $\angle\alpha_\ell$ is the phase of the ℓ^{th} path and $\widetilde{\mathbf{D}} = \left[\text{diag}\left(\left[\sqrt{\frac{N_t N_r}{L}} |\alpha_1|,\ \ldots,\ \sqrt{\frac{N_t N_r}{L}} |\alpha_L|\right]\right) \mid \mathbf{0}_{N_r - L \times N_t - L}\right]$.

By lemma 1 however, we know that the vector $\mathbf{a}_r(\phi_k^r, \theta_k^r) \perp \text{span}\{\mathbf{a}_r(\phi_\ell^r, \theta_\ell^r) | \forall \ell \neq k\}$ and similarly the vectors $\mathbf{a}_t(\phi_k^t, \theta_k^t) \perp \text{span}\{\mathbf{a}_t(\phi_\ell^t, \theta_\ell^r) | \forall \ell \neq k\}$ for large N_t and N_r. Therefore, in the limit of large systems we have $\mathbf{A}_t^* \mathbf{A}_t = \mathbf{I}_L$ and $\mathbf{A}_r^* \mathbf{A}_r = \mathbf{I}_L$. Thus $\mathbf{A}_r$ and $\mathbf{A}_t$ are asymptotically unitary.

Since the matrices $\left[\mathbf{A}_r,\ \mathbf{A}_r^\perp\right]$ and $\left[\mathbf{A}_t,\ \mathbf{A}_t^\perp\right]$ are asymptotically unitary, and $\widetilde{\mathbf{D}}$ is a diagonal matrix of non-negative elements, the channel representation in (9) *is itself a singular value decomposition of the channel* $\mathbf{H}$. Moreover, since the singular vectors of non-degenerate singular values are *unique* to within a rotation by $e^{j\theta}$ (such as the rotation by $e^{j\angle\alpha_\ell}$ in (9)), the vectors $\mathbf{a}_r(\phi_k^r, \theta_k^r)$ and $\mathbf{a}_t(\phi_k^t, \theta_k^t)$ are themselves singular vectors of $\mathbf{H}$ when the number of antennas grows. In other words, the left and right singular vectors of the channel converge in chordal distance to the array response vectors $\mathbf{a}_r(\phi_k^r, \theta_k^r)$ and $\mathbf{a}_t(\phi_k^t, \theta_k^t)$. □

Theorem 3 in summary states that the channel representation "converges" to its singular value decomposition and as a result, the *optimal SVD beamforming vector* are in fact given in closed form and *equal to the array response vector* in the strongest direction. This is given in the following corollary.

Corollary 4 *In the limit of large N_t and N_r, the optimal beamforming and combining vectors for the single user single stream mmWave channel given in (3) are $\mathbf{a}_t(\phi_{k^*}^t, \theta_{k^*}^t)$ and $\mathbf{a}_r(\phi_{k^*}^r, \theta_{k^*}^r)$ respectively where $k^* = \arg\max_\ell |\alpha_\ell|$.*

[1]For random path gains, angles of arrival, and angles of departure, it is easy to verify that $\mathbf{H}$ has rank L with probability 1 and thus has left and right nullspaces of dimension $N_r - L$ and $N_t - L$ respectively.

Proof: This is immediate from Theorem 3 □

It is promising to see that as array sizes grow, the optimal SVD solutions satisfy the constraints given in (7) meaning that only one RF chain is needed since these vectors can be implemented in RF. In Section 5 we show that the loss incurred by using the asymptotically optimal beamforming vectors in mmWave systems with a typical number of antennas is limited. Moreover, it is promising to see that $(\phi^t_{k^*}, \theta^t_{k^*})$ uniquely identify the optimal beamforming vector. Therefore, only 2 scalars need to be given to the transmitter to enable beamforming. As a result, limited feedback systems should use codebooks that simply target scalar quantization as opposed to more complex vector quantizers typically used in MIMO [12]. In fact, the receiver itself need not learn the entire matrix $\mathbf{H}$, which is rather large in mmWave system. Channel training need not go beyond estimating the dominant ($\phi^t_{k^*}$ and $\theta^t_{k^*}$).

4. SINGLE USER PRECODING IN LARGE MMWAVE SYSTEMS

Though Section 3 considers the case of $N_s = 1$, Theorem 3 can be used in the case of $N_s > 1$ where the optimal precoder again relies on the channel's singular vectors. In this section, we present the results on the optimality of beam steering, and layered precoding, for $N_s > 1$.

When $N_s > 1$, the well known optimal precoder is $\mathbf{F}^{opt} = \left[\mathbf{V}^{(1)}, \ldots, \mathbf{V}^{(N_s)}\right] \Gamma^{1/2}$, where Γ is a diagonal power allocation matrix, found by waterfilling. From Section 3, however, we know that each $\mathbf{V}^{(i)}$ converges to a vector $\mathbf{a}_t(\phi^t_\ell, \theta^t_\ell)$.

Corollary 5 *For large systems with $L = o(N_t)$, the optimal SVD precoder can be perfectly decomposed into $\mathbf{F}^{opt} = \mathbf{F}_{RF}\mathbf{F}_{BB}$ where $\mathbf{F}_{RF} = \left[\mathbf{a}_t(\phi^t_1, \theta^t_1), \ldots, \mathbf{a}_t(\phi^t_{N_s}, \theta^t_{N_s})\right]$ and $\mathbf{F}_{BB} = \Gamma$. This assumes $|\alpha_i| \geq |\alpha_j|$, $\forall i < j$ without loss of generality to simplify notation. Similarly the optimal receiver is $\mathbf{W}^{opt} = \mathbf{W}_{RF} = \left[\mathbf{a}_r(\phi^r_1, \theta^r_1), \ldots, \mathbf{a}_r(\phi^r_{N_s}, \theta^r_{N_s})\right]$ and can be implemented fully in RF using phase shifters.*

Proof: This is immediate from the Theorem 3 which establishes the equivalence between the natural channel representation and its singular value decomposition. □

Corollary 5 implies that optimal precoding converges to a form of *multi-directional beam steering* in which the baseband precoder is only needed for power allocation. Note if unitary precoding is assumed, i.e., no power allocation is allows, baseband processing is entirely needless for optimal precoding.

5. SIMULATION RESULTS

In this section we validate the asymptotic optimality results of Sections 3 and 4 for practically sized arrays with a half-wavelength inter-element spacing. To that end, Fig. 2 shows the data rates achieved by the optimal SVD beamforming vector versus the asymptotically optimal beam steering vector from Corollary 4. The propagation environment is modeled as a 6 ray channel with random angles of arrival and departure. This seemingly poor level of scattering is in fact considered typical in richer scattering microwave systems [15]. Inter-element spacing is assumed to be half-wavelength. Fig. 2 shows that as array size increases, the loss due to beam steering decreases and is negligible for a 256 element transmitter. This is promising since most mmWave systems consider tens or hundreds of transmit and receive antennas [3]. While Fig. 2 presents single stream performance,

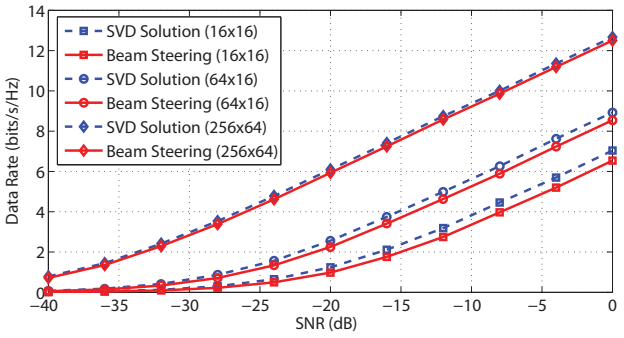

Fig. 2. Rates achieved by the optimal SVD solution and the asymptotically optimal steering solution in a 6 ray channel in an $N_t \times N_r$ ($N_t \times N_r$ is given in the legend). All UPA's are square, i.e. $W = H = \sqrt{N}$.

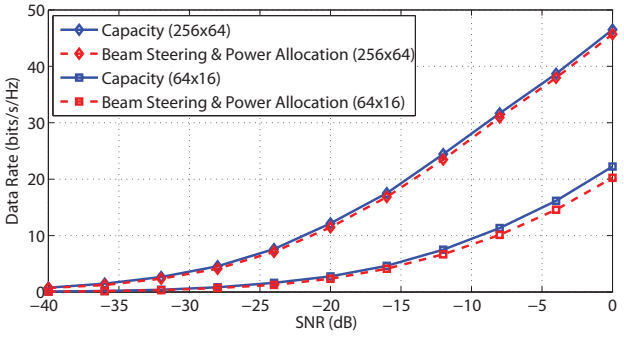

Fig. 3. A comparison between capacity and the rates achieved by beam steering in a 6 ray environment using square UPA's and 4 baseband chains.

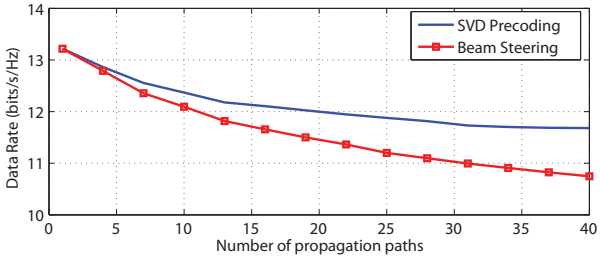

Fig. 4. Rates achieved at 0dB by the optimal SVD solution and the beam steering solution for $N_s = 1$ in a 256×64 square UPA system.

the same results hold for $N_s > 1$. To demonstrate this, Fig. 3 shows the small gap between *channel capacity* and the rate achieved by the beam steering and power allocation solution in Section 4. Note that the system is constrained to send $N_s \leq \min\{N^t_{RF}, N^r_{RF}\}$, therefore when the capacity achieving number of streams exceeds $\min\{N^t_{RF}, N^r_{RF}\}$, achieving capacity is impossible. Finally, Fig. 4 plots the achieved rate vs. the number of propagation rays in the channel to illustrate the effects of richer scattering. As can be seen, even in the significantly scattering environments with 40 paths, beam steering still achieves more than 90% of the maximum rate.

6. CONCLUSION

In this paper we considered single user precoding in large mmWave systems. Using a realistic spatial channel model, we established the

optimality of beam steering for both single and multi-stream transmission. We presented results on the performance of beam steering that validate the analytical results derived.

7. APPENDIX

7.1. Proof of Lemma 1

Let $\mathbf{A} = \left[\mathbf{a}^{ULAy}(\phi_1),\ \ldots,\ \mathbf{a}^{ULAy}(\phi_L)\right]$, then in the limit we have

$$\begin{aligned}
\lim_{N\longrightarrow\infty}\left|\mathbf{a}^{ULAy}(\phi_k)^*\mathbf{A}^{(-k)}\right| &\leq \lim_{N\longrightarrow\infty}\sum_{\ell\neq k}\left|\mathbf{a}^{ULAy}(\phi_k)^*\mathbf{a}^{ULAy}(\phi_\ell)\right| \\
&= \lim_{N\longrightarrow\infty}\sum_{\ell\neq k}\frac{1}{N}\left|\sum_{n=0}^{N-1}e^{jkdn(\sin(\phi_\ell)-\sin(\phi_k))}\right| \\
&\overset{(a)}{=} \lim_{N\longrightarrow\infty}\sum_{\ell\neq k}\frac{1}{N}\frac{\left|1-e^{jkd(\sin(\phi_\ell)-\sin(\phi_k))N}\right|}{\left|1-e^{jkd(\sin(\phi_\ell)-\sin(\phi_k))}\right|} \\
&\overset{(b)}{\leq} \lim_{N\longrightarrow\infty}\sum_{\ell\neq k}\frac{1}{N}\frac{2}{\left|1-e^{jkd(\sin(\phi_\ell)-\sin(\phi_k))}\right|} = 0.
\end{aligned}$$

Since ϕ_k and ϕ_ℓ are chosen independently from a continuous distribution, we have $(\sin(\phi_\ell)-\sin(\phi_k)) \neq 0$ with probability one, and thus each term $\mathbf{a}^{ULAy}(\phi_k)^*\mathbf{a}^{ULAy}(\phi_\ell)$ is a geometric series with ratio $e^{jkd(\sin(\phi_\ell)-\sin(\phi_k))} \neq 1$ which can be written as in (a). Finally (b) is since each of the $L-1=o(N)$ terms in the summation decays with N.

Therefore, $\mathbf{a}^{ULAy}(\phi_k) \perp \text{span}(\{\mathbf{a}^{ULAy}(\phi_\ell)|\ \forall \ell \neq k\})$ since the projection on the spanning set $\left\{\mathbf{a}^{ULAy}(\phi_\ell)|\quad \forall \ell \neq k\right\}$ is zero.

7.2. Proof of Corollary 2

The UPA array response vector in (5), can be decomposed into a Kronecker product of two linear array responses. The array response vectors can therefore be written as

$$\begin{aligned}
\mathbf{a}^{UPA}(\phi,\theta) = \frac{1}{\sqrt{N}}\left[1,\ e^{jkd\cos(\theta)},\ \ldots,\ e^{j(H-1)kd\cos(\theta)}\right]^T \otimes \\
\left[1,\ e^{jkd\sin(\phi)\sin(\theta)},\ldots,\ e^{j(W-1)kd\sin(\phi)\sin(\theta)}\right]^T.
\end{aligned}$$

Thus the magnitude of the projection of the vector $\mathbf{a}^{UPA}(\phi_k,\theta_k)$ on to $\text{span}(\{\mathbf{a}^{UPA}(\phi_\ell,\theta_\ell)|\quad \forall \ell \neq k\})$ is upper bounded by

$$\begin{aligned}
&\lim_{N\longrightarrow\infty}\sum_{\ell\neq k}\left|\mathbf{a}^{UPA}(\phi_k,\ \theta_k)^*\mathbf{a}^{UPA}(\phi_\ell,\ \theta_\ell)\right| \\
&= \lim_{N\longrightarrow\infty}\sum_{\ell\neq k}\Big|\ \left(\mathbf{a}^{ULAz}(\phi_k,\ \theta_k)\otimes\mathbf{a}^{ULAy}(\phi_k,\ \theta_k)\right)^* \\
&\qquad\qquad \left(\mathbf{a}^{ULAz}(\phi_\ell,\ \theta_\ell)\otimes\mathbf{a}^{ULAy}(\phi_\ell,\ \theta_\ell)\right)\ \Big| \\
&= \lim_{N\longrightarrow\infty}\sum_{\ell\neq k}\Big|\ \left(\mathbf{a}^{ULAz}(\phi_k,\ \theta_k)^*\mathbf{a}^{ULAz}(\phi_\ell,\ \theta_\ell)\right)\otimes \\
&\qquad\qquad \left(\mathbf{a}^{ULAy}(\phi_k,\ \theta_k)^*\mathbf{a}^{ULAy}(\phi_\ell,\ \theta_\ell)\right)\ \Big|
\end{aligned}$$

where $\mathbf{a}^{ULAz}(\phi_k,\ \theta_k)^* = \left[1,\ \ldots,\ e^{j(H-1)kd\cos(\theta)}\right]^T$ is the array response of a ULA on the z-axis. So, after writing $\mathbf{a}^{UPA}(\phi_k,\ \theta_k) = \mathbf{a}^{ULAz}(\phi_k,\ \theta_k)\otimes\mathbf{a}^{ULAy}(\phi_k,\ \theta_k)$ the remainder of the proof proceeds as in Appendix 7.1.

8. REFERENCES

[1] S. Yong and C. Chong, "An overview of multigigabit wireless through millimeter wave technology: potentials and technical challenges," *EURASIP Journal on Wireless Comm. and Net.*, vol. 2007, no. 1, 2007.

[2] R. Daniels and R. W. Heath, Jr., "60 GHz wireless communications: emerging requirements and design recommendations," *IEEE Vehicular Technology Magazine*, vol. 2, no. 3, pp. 41–50, 2007.

[3] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Comm. Mag.*, vol. 49, no. 6, pp. 101–107, 2011.

[4] C. Sheldon, M. Seo, E. Torkildson, M. Rodwell, and U. Madhow, "Four-channel spatial multiplexing over a millimeter-wave line-of-sight link," *IEEE Int. Microwave Symposium Digest*, pp. 389–392, Jun. 2009.

[5] P. Sudarshan, N. Mehta, A. Molisch, and J. Zhang, "Channel statistics-based RF pre-processing with antenna selection," *IEEE Transactions on Wireless Communications*, vol. 5, no. 12, pp. 3501 –3511, Dec. 2006.

[6] X. Zhang, A. Molisch, and S.-Y. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," *IEEE Transactions on Signal Processing*, vol. 53, no. 11, pp. 4091–4103, Nov. 2005.

[7] X. Zheng, Y. Xie, J. Li, and P. Stoica, "MIMO transmit beamforming under uniform elemental power constraint," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5395–5406, 2007.

[8] D. Love and R. W. Heath, Jr., "Equal gain transmission in multiple-input multiple-output wireless systems," *IEEE Transactions on Communications*, vol. 51, no. 7, pp. 1102–1110, Jul. 2003.

[9] H. Xu, V. Kukshya, and T. Rappaport, "Spatial and temporal characteristics of 60-GHz indoor channels," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 3, pp. 620–630, 2002.

[10] IEEE 802.15 WPAN Millimeter Wave Alternative PHY Task Group 3c. [Online]. Available: www.ieee802.org/15/pub/TG3c.html, Sept., 2011.

[11] V. Raghavan and A. Sayeed, "Multi-antenna capacity of sparse multipath channels," *submitted to IEEE Trans. Inform. Theory*, Aug. 2006. [Online]. Available: dune.ece.wisc.edu/pdfs/sp_mimo_cap.pdf

[12] D. Love and R. W. Heath, Jr., "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Transactions on Information Theory*, vol. 51, no. 8, pp. 2967 –2976, Aug. 2005.

[13] P. Xia, S. Yong, J. Oh, and C. Ngo, "Multi-stage iterative antenna training for millimeter wave communications," *IEEE GlOBECOM*, 2008.

[14] C. Balanis, *Antenna theory*. Wiley New York, 1997.

[15] 3GPP, "Spatial channel model for MIMO simulations," 3GPP, TR 25.996 V6.1.0, Sep. 2003. [Online]. Available: http://www.3gpp.org/