Low Complexity Energy Efficiency Analysis in Millimeter Wave Communication Systems

Pan Cao and John Thompson
Institute for Digital Communications, The University of Edinburgh, EH9 3JL, UK
Email: {p.cao, john.thompson}@ed.ac.uk

***Abstract*—Millimeter wave (mm-wave) system performance may be degraded if the operation mechanism is not properly designed, because mm-wave systems suffer severe path loss and very short coherence time. Thanks to the sparse channel model and directional transmission property, it is usually sufficient to use analog beam codebooks in beam training to estimate dominant channel components instead of complete instantaneous channel matrices. With this viewpoint, we first characterize the achievable beam gain by the number of antennas and beamwidth, and then propose a low complexity mechanism that employs the offline designed analog beam codebooks for both beam training and data transmission. This mechanism not only avoids high overhead and delay caused by the online beamforming design based on instantaneous channels but also enables a much longer *quasi coherence time*, which is the new concept proposed in this work. In addition, it can realize a theoretical analysis of the role of system parameters in energy efficiency. We consider a phase-controlled point-to-point (P2P) mm-wave system example to illustrate the proposed concepts and mechanism. Numerical simulations also verify the effectiveness of the theoretical analysis result and also provide a suggestion for system design.**

I. Introduction

Based on Friis' free space path loss (FSPL) equation [1]:

$$\mathrm{FSPL}(f_G, d) = 32.4 + 20\log_{10}(f_G) + 20\log_{10}(d)\ [dB] \quad (1)$$

where f_G and d denote the carrier frequency in [GHz] and transmission distance [meter], respectively, the transmission at millimeter wave (mm-wave) frequencies suffers much more severe FSPL compared with microwave frequency (below 5 GHz) transmission. To combat this severe path loss at mm-wave frequencies, large scale antenna arrays are required to achieve high beamforming gain. However, it is usually challenging to deploy a large number of radio frequency (RF) chains and data converters to serve the massive antennas equipped in a very compact physical space due to both the high cost/power consumption and high complexity requirement in hardware design [2]. This motivates a wide application of the hybrid Tx/Rx structure in mm-wave communication systems that only a few radio frequency (RF) chains and data converters are used to serve a large scale antenna array [3]. This architecture can greatly reduce the complexity and cost, and the beamforming design is split into a high dimensional analog beamforming design and a low dimensional digital beamforming design [3].

Regarding the beamforming design in mm-wave communications, most existing algorithms can be roughly categorized into two types. The first type of approach is online design for both analog beamforming and digital beamforming strategies. For example, it is popular to first online generate both digital and analog beamforming strategies [4]–[7]. However, the resulting high time delay, overhead and computation load may degrade the spectral and energy efficiency. The second type of approaches is to use offline designed analog beam codebooks and online designed digital beamforming strategies, e.g., [8]. In particular, analog beam codebooks can be stored at both the transmitter (Tx) and receiver (Rx), and then efficient beam training is used to determine the best Tx-Rx analog beam pair and estimate the dominant components instead of the complete channels. Compared with the first approach, this method can greatly reduce the online complexity and time delay. Therefore, offline designed analog beam codebooks play a significant role in latency reduction.

In beam codebooks, the main beam lobe gain becomes greater when the beamwidth becomes narrower and thus the received SNR if the beam alignment is perfect. However, a narrower beamwidth leads to more overhead for beam training such that less time is left for the data transmission, which may degrade the spectral efficiency [9]. Therefore, a trade-off exists between the beamwidth, the number of antennas and system spectral and energy efficiency performance.

To the best knowledge of the authors, few previous works provides a theoretical analysis of the role of the analog beamwidth and the number of antennas in the system energy efficiency. We propose to use the best beam pairs determined in the beam training for data transmission. In this method, we propose a new concept of the *quasi coherence time* based on beamforming adjustment, which is much longer than the instantaneous channel coherence time. Futhermore, based on the offline designed analog beam codebooks, the achievable beam gain is characterized by the main lobe beamwidth and the number of antennas. This parameterization result allows us to analyze the role of both the main beamwidth and the number of antennas in energy efficiency. A case study for the P2P system is done to illustrate and evaluate the proposed concepts and mechanism.

II. System Model

In this work, we consider a mm-wave P2P system, denoted by Tx $\mapsto$ Rx., where each node is equipped with an N-antenna uniform linear array (ULA) associated with a single RF chain and data converter. For the narrow-band and far-field

The work was financially supported by the UK EPSRC grant EP/L026147/1.

978-3-901882-90-6/17 ©2017 IFIP

transmission, the array response vector (steering vector) for an N-antenna ULA at Tx or Rx can be defined as

$$\boldsymbol{a}(\theta) := \left[1, e^{j\pi\cos\theta}, \cdots, e^{j(N-1)\pi\cos\theta}\right]^T. \quad (2)$$

This is based on half wavelength spacing between two adjacent elements, and $\theta \in [0, \pi]$ denotes the angle of departure (AoD) of the signal from Tx or the angle of arrival (AoA) to Rx, where $[0, \pi]$ is defined as the physical angular range that covers the entire (one-sided) spatial horizon for a linear antenna array.

The directional propagation property at mm-wave frequencies leads to limited spatial selectivity or scattering. For this reason, the narrowband clustered channel representation can be used to model a mm-wave channel for the link Tx $\mapsto$ Rx as follows [10], [11]:

$$\boldsymbol{H} := \underbrace{\alpha_1 \boldsymbol{a}(\theta_1^r)\boldsymbol{a}^H(\theta_1^t)}_{LOS\ ray} + \underbrace{\sum_{m=2}^{M} \alpha_m \boldsymbol{a}(\theta_m^r)\boldsymbol{a}^H(\theta_m^t)}_{NLOS\ rays}, \quad (3)$$

which consists of M transmission rays – one LOS ray and $M-1$ NLOS rays. Based on the statistical model for mm-wave channels suggested in [10], each complex channel gain α_m can be expressed as $\alpha_m := \sqrt{\rho_m} h_m$, where ρ_m and h_m denote the large scale fading and small scale fading of the m-th ray. And θ_m^t and θ_m^r denote the AoD from Tx and AoA to Rx of the m-th ray. These angles depend on the number and locations of the clusters and in statistical it is usually assumed that $\theta_m^t, \theta_m^r, \forall m$ are independently and randomly distributed in $[0, \pi]$ for simplicity.

Based on the channel measurement in [10, Table I], the path loss is roughly expressed as:

$$\gamma := \sum_{m=1}^{M} |\alpha_m|^2 = 10^{-(a+\zeta)/10} d^{-b}, \quad (4)$$

which is based on the non-coherent combination of all the channel gains. More precisely, ζ accounts for variance in shadowing and satisfies a uniform distribution with zero mean and variance σ_{sh}^2. Thus, γ satisfies a log-normal distribution due to ζ. As we assume that the LOS transmission exists in the link Tx $\mapsto$ Rx, the values of the parameters a, b and σ_{sh}^2 in LOS scenario are suggested in [10, Table I].

In particular, the LOS ray experience the FSPL as (1):

$$\gamma_{LOS} := |\alpha_1|^2 = 10^{-3.24} f_G^{-2} d^{-2}, \quad (5)$$

and the combining channel gain of the rest $M-1$ NLOS rays of the LOS channel is approximately expressed as

$$\gamma_{NLOS} := \sum_{m=2}^{M} |\alpha_m|^2 = (10^{-(a+\zeta)/10} - 10^{-3.24} f_G^{-2}) d^{-2}, \quad (6)$$

which is based on $b = 2$ for LOS transmission, and γ_{NLOS} still satisfies log-normal distribution due to ζ.

Let $\boldsymbol{g}_t, \boldsymbol{g}_r$ be the two analog beamforming vectors employed for Tx and Rx, respectively, in the set $\mathcal{A}$:

$$\mathcal{A} := \{\boldsymbol{x} : \boldsymbol{x} \in \mathbb{C}^{N\times 1}, |\boldsymbol{x}(n)|^2 = 1/N, \forall n\}. \quad (7)$$

Then, the received signal at Rx is

$$y = \sqrt{p}\boldsymbol{g}_r^H \boldsymbol{H} \boldsymbol{g}_t s + \boldsymbol{g}_r^H \boldsymbol{z} \quad (8)$$

where s with $\mathbb{E}\{|s|^2\} = 1$ denotes the desired transmit symbol from Tx to Rx with the transmit power p. The vector $\boldsymbol{z} \in \mathbb{C}^{N\times 1}$ denotes the terminal noise at Rx and satisfies the distribution of $\mathcal{CN}(\boldsymbol{0}, \sigma_0^2 \boldsymbol{I})$, where σ_0^2 denotes the received noise power.

III. Analog Beam Codebook Design

A. Analog Beam Codebook Based Beam Training

Consider an analog beam codebook $\mathcal{W} := \{\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_L\}$ where $\boldsymbol{w}_i \in \mathcal{A}, \forall i$. Many papers assume that each beam codeword $\boldsymbol{w}_i$ enables the following shaped beam pattern [3], [10], [12]:

$$G_i(\theta) := |\boldsymbol{w}_i^H \boldsymbol{a}(\theta)|^2 = \begin{cases} G_m(N, \overline{\theta}), & \text{when } \theta \in \Theta_m^i \\ G_s(N, \overline{\theta}), & \text{when } \theta \in \Theta_s^i, \end{cases} \quad (9)$$

where $G_m(N, \overline{\theta}), G_s(N, \overline{\theta})$ and Θ_m^i, Θ_s^i denote the constant beam gain, and angular range for the main lobe and side lobe, respectively, when the main lobe width is $\overline{\theta}$. The main lobe of each beam pattern covers a unique angular range Θ_m^i. All the beam codewords support the entire angular range $[0, \pi]$. For example, Fig. 1 shows a six-beam codebook generated in [13], where each top-flat main lobe covers an individual $\pi/6$ angular range. Furthermore, we also conclude in [13] the variance of beam gain with the system parameters N and $\overline{\theta}$.

Lemma 1 *The main lobe gain and the side lobe gain can be parameterized as a function of N and $\overline{\theta}$ as follows:*

$$G_q(N, \overline{\theta}) \approx c_{00} + c_{10}N + c_{01}\overline{\theta} + c_{20}N^2 + c_{11}N\overline{\theta} + c_{02}\overline{\theta}^2 + c_{21}N^2\overline{\theta} + c_{12}N\overline{\theta}^2 + c_{03}\overline{\theta}^3 \quad (10)$$

where $q \in \{m, s\}$ refers to either main lobe ($q = m$) or side lobe ($q = s$) and the coefficients are given in Table I.

TABLE I: Coefficient values for the parameterization (10)

coefficients	main lobe $q = m$	side lobe $q = s$
c_{00}	12.11	-1.723
c_{10}	0.02365	-0.01646
c_{01}	-0.4.01	-0.4791
c_{20}	-0.0001302	-0.0004497
c_{11}	-0.0004741	0.001808
c_{02}	0.006375	0.01271
c_{21}	1.481e-06	1.505e-05
c_{12}	3.444e-06	-5.144e-05
c_{03}	-3.998e-05	-9.127e-05

□

In practice, analog beam codebooks can be designed offline and used as a *look-up table* in the beam training. The beam

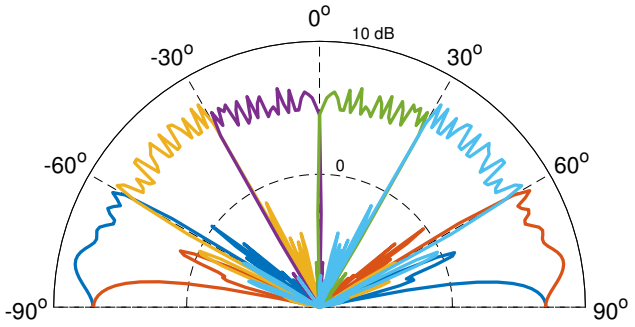

Fig. 1: A six-beam codebook example for $N = 64$

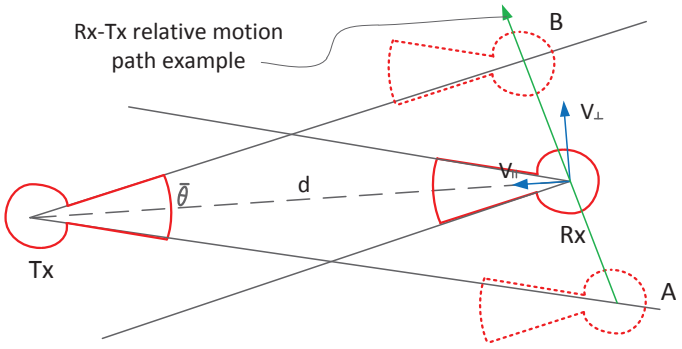

Fig. 2: Concept illustration for quasi coherence time

training is a Tx-Rx beam swapping procedure to find the best Tx-Rx beam pair by measuring the received signal strength with different Tx-Rx beam pair [14]. Therefore, this *exhaustive beam swapping* takes L^2+1 symbol intervals consisting – L^2 beam swapping symbol intervals and 1 symbol interval for feedback of the best Tx beam index from Rx to Tx.

Remark 1 *With a single resolution beam codebook, where each beam pattern has the same main beamwidth $\overline{\theta} = \pi/L$, the beam codebook based beam training takes*

$$T_{channel} := ((\pi/\overline{\theta})^2 + 1)\Delta \quad [second], \tag{11}$$

where Δ denotes a symbol transmission interval. □

B. Analog Beam Codebook based Operation Mechanism

In wireless communication, transmission data rate can be expressed as follows:

$$R = B\left(1 - \frac{T_{channel} + T_{online}}{T_{ct}}\right)\log_2(1 + \text{SNR}) \tag{12}$$

where $T_{channel}, T_{online}$ and T_{ct} denote beam training time, online strategy design time and channel coherence time for the system in [second], and SNR denotes the received SNR. The portion of coherence time for the desired data transmission is $T_{data} = T_{ct} - T_{channel} - T_{online}$, and thus the parameter $1 - \frac{T_{channel}+T_{online}}{T_{ct}} < 1$ denotes the spectral efficiency loss factor caused by the time delay.

Based on (8), the SNR for the k-th user in (12) can be further expressed as:

$$\text{SNR} = \frac{p}{\sigma_0^2}|\boldsymbol{g}_r^H \boldsymbol{H} \boldsymbol{g}_t|^2. \tag{13}$$

If the beamforming vectors $\boldsymbol{g}_t$ and $\boldsymbol{g}_r$ are designed based on the instantaneous channel matrices $\boldsymbol{H}$ to maximize the SNR in (13), the system coherence time T_{ct} is equivalent to the instantaneous channel coherence time. However, the efficient analog beam codebook based beam training is not able to estimate complete channel matrices but dominant channel components. In this work, *we propose to employ the aligned Tx-Rx beam pairs determined in the beam training procedure to serve as the beamforming strategies in data transmission, which can greatly reduce T_{online} by avoiding the online beamforming design procedure. Furthermore, the mechanism can also lead to a much longer system coherence time compared with the instantaneous channel coherence time.*

Under this mechanism, we introduce a new concept – *quasi coherence time* for a mm-wave channel, defined as follows:

Definition 1 *The quasi coherence time of a mm-wave channel refers to the time period during which the dominant transmission rays (including LOS ray) go through both the main lobes of a fixed Tx-Rx beam pair.* □

Remark 2 *The instantaneous channel may still change during a quasi coherence time, but we can keep using the same Tx-Rx beam pair because it is still the best beam pair to support dominant transmission rays.* □

As an example shown in Fig. 2, the quasi coherence time refers to the time period for Rx moving from Position A to Position B relative to Tx. Thus, the quasi coherence time is much longer than the traditional coherence time for the instantaneous channel depending on the small scale fading. We assume that Rx has a velocity $(V_\perp, V_\parallel)$ relative to Tx. Then, we can compute the quasi coherence time as follows:

Lemma 2 *When the initial distance from Tx to Rx is d_0, the average quasi coherence time for the mm-wave P2P system in Fig. 2 can be approximately computed as*

$$T_{ct} = \frac{\sin(\overline{\theta}/2)}{\sin(\pi - \overline{\theta}/2 - \text{asin}(V_\perp/V_\parallel))}\frac{d_0}{\sqrt{V_\perp^2 + V_\parallel^2}}. \tag{14}$$

□

Proof: In practice, Rx may be at different any location in the coverage of Tx when the P2P system starts beam training. From statistical point of view, we can simply assume that Rx is located in the middle line of the angular coverage of Tx. ■

Under the proposed mechanism, we can *avoid the online beamforming design* ($T_{online} = 0$) as the best Tx-Rx beam pair has been determined in the beam training procedure. As a result, the proposed mechanism takes

$$T_{channel} + T_{online} = ((\pi/\overline{\theta})^2 + 1)\Delta \quad [second], \tag{15}$$

before data transmission.

IV. Low Complexity Energy Efficiency Optimization

In this work, we consider the role of analog beamwidth and the number of antennas in the energy efficiency, which will provide a suggestion for the green system design. More precisely, the *energy efficiency performance metric is defined as average rate to power consumption ratio [bits/Joule]*. In the following, the average rate and power consumption will be computed, respectively.

A. Average Rate Computation

After Tx-Rx beam pairs are determined in the beam training procedure, the SINR_k in (13) can be rewritten as

$$\text{SNR} = \frac{p}{\sigma_0^2}\sum_{k=1}^{K}|\alpha_k|^2 G(\theta_k^t)G(\theta_k^r), \tag{16}$$

where $G(\theta_k^t)$ and $G(\theta_k^r)$ denote the Tx and Rx beam gain experienced by the k-th ray with AoD θ_k^t and AoA θ_k^r, respectively. As we assume that the LOS ray always exists between Tx k and Rx k, the LOS ray gain $|\alpha_1|^2$ will experience the large beam gain after beam alignment, i.e., $G(\theta_1^t)G(\theta_1^r) = G_m^2(N,\overline{\theta})$ with the probability (w.p.) of one. For each NLOS ray when $k > 1$, it experiences the Tx-Rx beam gain as:

$$G(\theta_k^t)G(\theta_k^r) = \begin{cases} G_m^2(N,\overline{\theta}) & \text{w.p. } \left(\frac{\overline{\theta}}{\pi}\right)^2 \\ G_m(N,\overline{\theta})G_s(N,\overline{\theta}) & \text{w.p. } 2\frac{\overline{\theta}(\pi-\overline{\theta})}{\pi^2} \\ G_s^2(N,\overline{\theta}) & \text{w.p. } \left(\frac{\pi-\overline{\theta}}{\pi}\right)^2. \end{cases} \tag{17}$$

We desire to investigate the impact of the beamwidth on the average spectral efficiency for the link Tx $\mapsto$ Rx. When the initial distance between Tx and Rx is d_0 [meters], we define $\widetilde{R} := \mathbb{E}_{T_{ct},\{\theta_m^t,\theta_m^r\},\zeta}[R]$ as the performance metric averaged over AoAs and AoDs of NLOS rays, log-normal shadowing and the whole quasi coherence time period, which is approximately derived in [13] as follows.

Proposition 1 *When the initial Tx-Rx distance d_0 is known, the average rate $\widetilde{R}$ can be approximately expressed as*

$$\begin{aligned}\widetilde{R} \approx & B\left(1-\frac{((\pi/\overline{\theta})^2+1)\Delta}{T_{ct}}\right)\frac{B}{T_{ct}\log(2)}\times \\ & \Big(\Gamma\left(V_\perp^2+V_\parallel^2,-2d_0V_\perp,d_0^2+\beta,0,T_{ct}\right) \\ & \quad -\Gamma\left(V_\perp^2+V_\parallel^2,-2d_0V_\perp,d_0^2,0,T_{ct}\right)\Big)\end{aligned} \tag{18}$$

where T_{ct} is given in (14) *and the definitions are given as:*

$$\begin{aligned}\beta := & \frac{p}{\sigma_0^2}(G_m^2(N,\overline{\theta})10^{-3.24}f_G^{-2} \\ & + \widetilde{G}(N,\overline{\theta})(e^{(\widehat{\mu}+\widehat{\sigma}_{sh}^2/2)} - 10^{-3.24}f_G^{-2})) \\ \Gamma(c_2,c_1,c_0,\underline{b},\overline{b}) := & \ \Omega(c_2,c_1,c_0,\overline{b}) - \Omega(c_2,c_1,c_0,\underline{b})\end{aligned}$$

$$\begin{aligned}\Omega(c_2,c_1,c_0,x) := & \frac{1}{c_2}\sqrt{4c_2c_0-c_1^2}\tan^{-1}\left(\frac{2c_2x+c_1}{\sqrt{4c_2c_0-c_1^2}}\right) \\ & -2x+\left(\frac{c_1}{2c_2}+x\right)\log\left(c_2x^2+c_1x+c_0\right)\end{aligned}$$

□

B. Power Consumption Model

Within a quasi coherence time period, the power consumption in the P2P system can be expressed as:

$$P_{total} := \frac{1}{\eta}p + 2\left(1-\frac{((\pi/\overline{\theta})^2+1)\Delta}{T_{ct}}\right)\rho BN + 2P_c, \tag{19}$$

where the three terms in (19) denote the consumption by transmit power, signal processing power and constant circuit power, respectively. More precisely, $\eta \in (0,1)$ and ρ denotes the power amplifier efficiency and the unit signal processing power for per antenna and per Hz transmission bandwidth. The total power P_{total} includes the circuit power and signal processing power in both Tx and Rx. Furthermore, as the signal processing power like AD and DA conversion is only needed in data transmission, the factor $\left(1-\frac{((\pi/\overline{\theta})^2+1)\Delta}{T_{ct}}\right)$ is used in the signal processing power term.

C. Energy Efficiency Maximization

As $\widetilde{R}$ in (18) is a closed-form function of both $\overline{\theta}$ and N, it is possible to determine the best system parameters achieving the best energy efficiency performance by doing a simple two-dimensional grid search of $\overline{\theta} \in [\pi/30,\pi]$ and $N \in [2, N_{\max}]$ instead of high complexity online beamforming design.

$$\max_{\boldsymbol{g}_t,\boldsymbol{g}_r,N}\frac{R}{P_{total}} \quad \Rightarrow \quad \max_{\overline{\theta},N}\frac{\widetilde{R}}{P_{total}}. \tag{20}$$

Although the low complexity search on the average performance may degrade instantaneous energy efficiency, yet it could greatly reduce the channel acquisition overhead and online strategy design time delay.

V. Numerical Results

We evaluate energy efficiency performance of a P2P system, where Tx and Rx have different initial distance d_0. Tx has a transmit power $p = 23$ dBm, and the terminal noise floor is set to be $\sigma_0^2 = -174 \times B + 40$ dBm (including 10 dB noise figure). The carrier frequency is 73 GHz and $B = 2$ GHz. The channel fading coefficients in (6) are set as $a = 69.8$ and $\zeta = 5.8$. Set $\rho = 10$ Watt/GHz, $P_c = 10$ Watt, and symbol interval in beam training $\Delta = 1$ microsecond.

The analysis result in (20) is based on the assumption of Rx is located at middle line of the angular coverage. In the Monte Carlo simulations, we consider 10000 random angular positions for Rx. This aims to see the accuracy of the coherence time computaton in (14).

When the initial distance between Tx and Rx is $d_0 = 50$ meter, Fig. 3 shows analysis result in (20) based on (14) has the same variance of energy efficiency with main beamwidth $\overline{\theta}$ and the number of antennas N, and in addition the gap between them is very small. From 3, we observe that energy efficiency is decreasing with both N and $\overline{\theta}$ and thus its maximum value

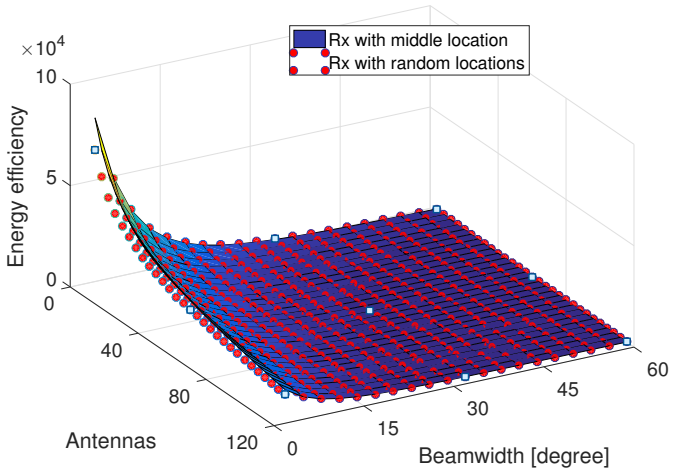

Fig. 3: Energy efficiency vs. $\overline{\theta}$ and N when $d_0 = 50$ meter

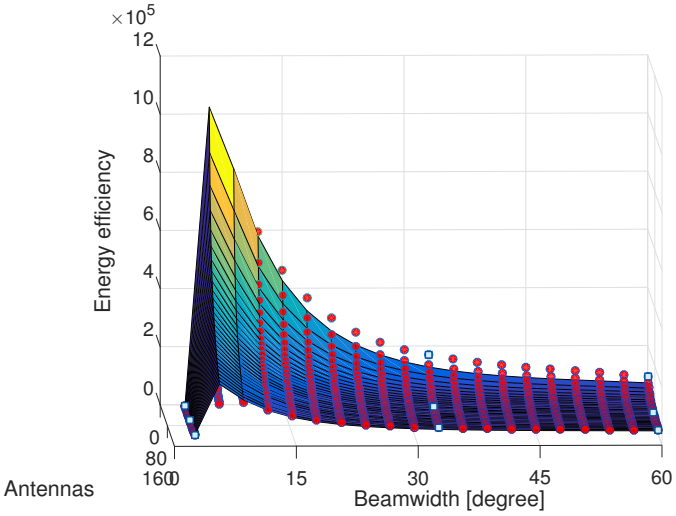

Fig. 4: Energy efficiency vs. $\overline{\theta}$ and N when $d_0 = 10$ meter

is achieved when $N = 2$ and $\overline{\theta} = 3^o$. The reason for result may be the coherence time in this case is long enough even when $\overline{\theta} = 3$ and fewer antennas are used to save signal processing power. However, when the initial Tx-Rx distance is $d_0 = 10$ meter, Fig. 4 shows the energy efficiency achieves the maximum when $N = 2$ but $\overline{\theta} = 6^o$. As when the Tx is closed to Rx, a wider main lobe beamwidth is needed to enable a longer quasi coherence time.

Simulation results imply that we can *semi-dynamically* adjust the best system parameters to achieve the most energy efficient performance once the initial Tx-Rx distance is given.

VI. Conclusions

In this work, we propose an analog beam codebook based mechanism mm-wave systems, which has a very low complexity, avoids the high overhead and online computation, and prolongs the system coherence time (i.e., the proposed *quasi coherence time*) compared with the previous approaches. In addition, this mechanism based on the beam gain parameterization enables a theoretical analysis for energy efficiency performance. As an application example, we analyze the energy efficiency in a mm-wave P2P system and efficiently determine the best solution by a two-dimensional grid search, which provides a suggestion on green system parameters design.

References

[1] T. S. Rappaport, *Wireless Communications: Principles and Practice*, Upper Saddle River, NJ: Prentice Hall,, 2nd ed. edition, 2002.

[2] C. H. Doan, S. Emami, D. A. Sobel, A. M. Niknejad, and R. W. Brodersen, "Design considerations for 60 GHz CMOS radios," *IEEE Commun. Mag.*, vol. 42, no. 12, pp. 132–140, Dec. 2004.

[3] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.

[4] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[5] Y. Y. Lee, C. H. Wang, and Y. H. Huang, "A hybrid RF/baseband precoding processor based on parallel-index-selection matrix-inversion-bypass simultaneous orthogonal matching pursuit for millimeter wave MIMO systems," *IEEE Transactions on Signal Processing*, vol. 63, no. 2, pp. 305–317, Jan. 2015.

[6] X. Yu, J. C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 485–500, Apr. 2016.

[7] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, Apr. 2016.

[8] A. Alkhateeb and R. W. Heath, "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1801–1818, May 2016.

[9] P. Cao and J. Thompson, "The role of analogue beamwidth in spectral efficiency of mmWave Ad Hoc networks," in *EAI SmartGIFT*, May 2016, pp. 98–104.

[10] 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 J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.

[11] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!," *IEEE Access*, vol. 1, pp. 335–349, 2013.

[12] Tianyang Bai, A. Alkhateeb, and R. Heath, "Coverage and capacity of millimeter-wave cellular networks," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 70–77, Sep. 2014.

[13] P. Cao, Y. Jiang, J. Thompson, and H. Haas, "Analog beam codebook design and its applications in Millimeter wave communication systems," *submitted to IEEE Trans. Wirel. Commun.*, 2017.

[14] Joongheon Kim and A.F. Molisch, "Fast millimeter-wave beam training with receive beamforming," *Journal of Commun. and Net.*, vol. 16, no. 5, pp. 512–522, Oct. 2014.