

RESEARCH

Open Access

# Adaptive antenna selection and Tx/Rx beamforming for large-scale MIMO systems in 60 GHz channels

Ke Dong<sup>1</sup>, Narayan Prasad<sup>2</sup>, Xiaodong Wang<sup>3\*</sup> and Shihua Zhu<sup>1</sup>

## Abstract

We consider a large-scale MIMO system operating in the 60 GHz band employing beamforming for high-speed data transmission. We assume that the number of RF chains is smaller than the number of antennas, which motivates the use of antenna selection to exploit the beamforming gain afforded by the large-scale antenna array. However, the system constraint that at the receiver, only a linear combination of the receive antenna outputs is available, which together with the large dimension of the MIMO system makes it challenging to devise an efficient antenna selection algorithm. By exploiting the strong line-of-sight property of the 60 GHz channels, we propose an iterative antenna selection algorithm based on discrete stochastic approximation that can quickly lock onto a near-optimal antenna subset. Moreover, given a selected antenna subset, we propose an adaptive transmit and receive beamforming algorithm based on the stochastic gradient method that makes use of a low-rate feedback channel to inform the transmitter about the selected beams. Simulation results show that both the proposed antenna selection and the adaptive beamforming techniques exhibit fast convergence and near-optimal performance.

**Keywords:** 60 GHz communication, MIMO, Antenna selection, Stochastic approximation, Gerschgorin circle, Beamforming, Stochastic gradient

## 1 Introduction

The 60 GHz millimeter wave communication has received significant recent attention, and it is considered as a promising technology for short-range broadband wireless transmission with data rate up to multi-giga bits/s [1-4]. Wireless communications around 60 GHz possess several advantages including huge clean unlicensed bandwidth (up to 7 GHz), compact size of transceiver due to the short wavelength, and less interference brought by high atmospheric absorption. Standardization activities have been ongoing for 60 GHz Wireless Personal Area Networks (WPAN) [5] (i.e., IEEE 802.15) and Wireless Local Area Networks (WLAN) [6] (i.e., IEEE 802.11). The key physical layer characteristics of this system include a large-scale MIMO system (e.g.,  $32 \times 32$ ) and the use of both transmit and receive beamforming techniques.

To reduce the hardware complexity, typically, the number of radio-frequency (RF) chains employed (consisting of amplifiers, AD/DA converters, mixers, etc.) is smaller than the number of antenna elements, and the antenna selection technique is used to fully exploit the beamforming gain afforded by the large-scale MIMO antennas. Although various schemes for antenna selection exist in the literature [7-10], they all assume that the MIMO channel matrix is known or can be estimated. In the 60 GHz WPAN system under consideration, however, the receiver has no access to such a channel matrix, because the received signals are combined in the analog domain prior to digital baseband due to the analog beamformer or phase shifter [11]. But rather, it can only access the scalar output of the receive beamformer. Hence, it becomes a challenging problem to devise an antenna selection method based on such a scalar only rather than the channel matrix. By exploiting the strong line-of-sight property of the 60 GHz channel, we propose a low-complexity iterative antenna selection technique based on the Gerschgorin circle and the

\* Correspondence: wangx@ee.columbia.edu

<sup>3</sup>Electrical Engineering Department, Columbia University, New York, NY, 10027, USA

Full list of author information is available at the end of the article

stochastic approximation algorithm. Given the selected antenna subset, we also propose a stochastic gradient-based adaptive transmit and receive beamforming algorithm that makes use of a low-rate feedback channel to inform the transmitter about the selected beam.

The remainder of this paper is organized as follows. The system under consideration and the problems of antenna selection and beamformer adaptation are described in Section 2. The proposed antenna selection algorithm is developed in Section 3. The proposed transmit and receive adaptive beamforming algorithm is presented in Section 4. Simulation results are provided in Section 5. Finally Section 6 concludes the paper.

## 2 System description and problem formulation

Consider a typical indoor communication scenario and a MIMO system with  $N_t$  transmit and  $N_r$  receive antennas both of omni-directional pattern operating in the 60 GHz band. The radio wave propagation at 60 GHz suggests the existence of a strong line-of-sight (LOS) component as well as the multi-cluster multi-path components because of the high path loss and inability of diffusion [3,4]. Such a near-optical propagation characteristic also suggests a 3-D ray-tracing technique in

channel modeling (see Figure 1), which is detailed in [12]. In our analysis, the transceiver can be any device, defined in IEEE 802.15.3c [5] or 802.11ad [6], located in arbitrary positions within the room. For each location, possible rays in LOS path and up to the second-order reflections from walls, ceiling, and floor are traced for the links between the transmit and receive antennas. In particular, the impulse response for one link is given by

$$h(t, \phi_{tx}, \theta_{tx}, \phi_{rx}, \theta_{rx}) = \sum_i A^{(i)} C^{(i)}(t - T^{(i)}, \phi_{tx} - \Phi_{tx}^{(i)}, \theta_{tx} - \Theta_{tx}^{(i)}, \phi_{rx} - \Phi_{rx}^{(i)}, \theta_{rx} - \Theta_{rx}^{(i)}) \quad (1)$$

where  $A^{(i)}$ ,  $T^{(i)}$ ,  $\Phi_{tx}^{(i)}$ ,  $\Theta_{tx}^{(i)}$ ,  $\Phi_{rx}^{(i)}$ ,  $\Theta_{rx}^{(i)}$  are called the inter-cluster parameters that are the amplitude, delay, departure, and arrival angles (in azimuth and elevation) of ray cluster  $i$ , respectively, and

$$C^{(i)}(t, \phi_{tx}, \theta_{tx}, \phi_{rx}, \theta_{rx}) = \sum_k \alpha^{(i,k)} \delta(t - \tau^{(i,k)}) \delta(\phi_{tx} - \phi_{tx}^{(i,k)}) \delta(\theta_{tx} - \theta_{tx}^{(i,k)}) \delta(\phi_{rx} - \phi_{rx}^{(i,k)}) \delta(\theta_{rx} - \theta_{rx}^{(i,k)}) \quad (2)$$

denotes the cluster constitution by rays therein, where  $\alpha^{(i,k)}$ ,  $\tau^{(i,k)}$ ,  $\phi_{tx}^{(i,k)}$ ,  $\theta_{tx}^{(i,k)}$ ,  $\phi_{rx}^{(i,k)}$ ,  $\theta_{rx}^{(i,k)}$  are the intra-cluster parameters for  $k$ th ray in cluster  $i$ . Some inter-cluster parameters are usually location related, e.g., the severe path loss in cluster amplitude; some are random

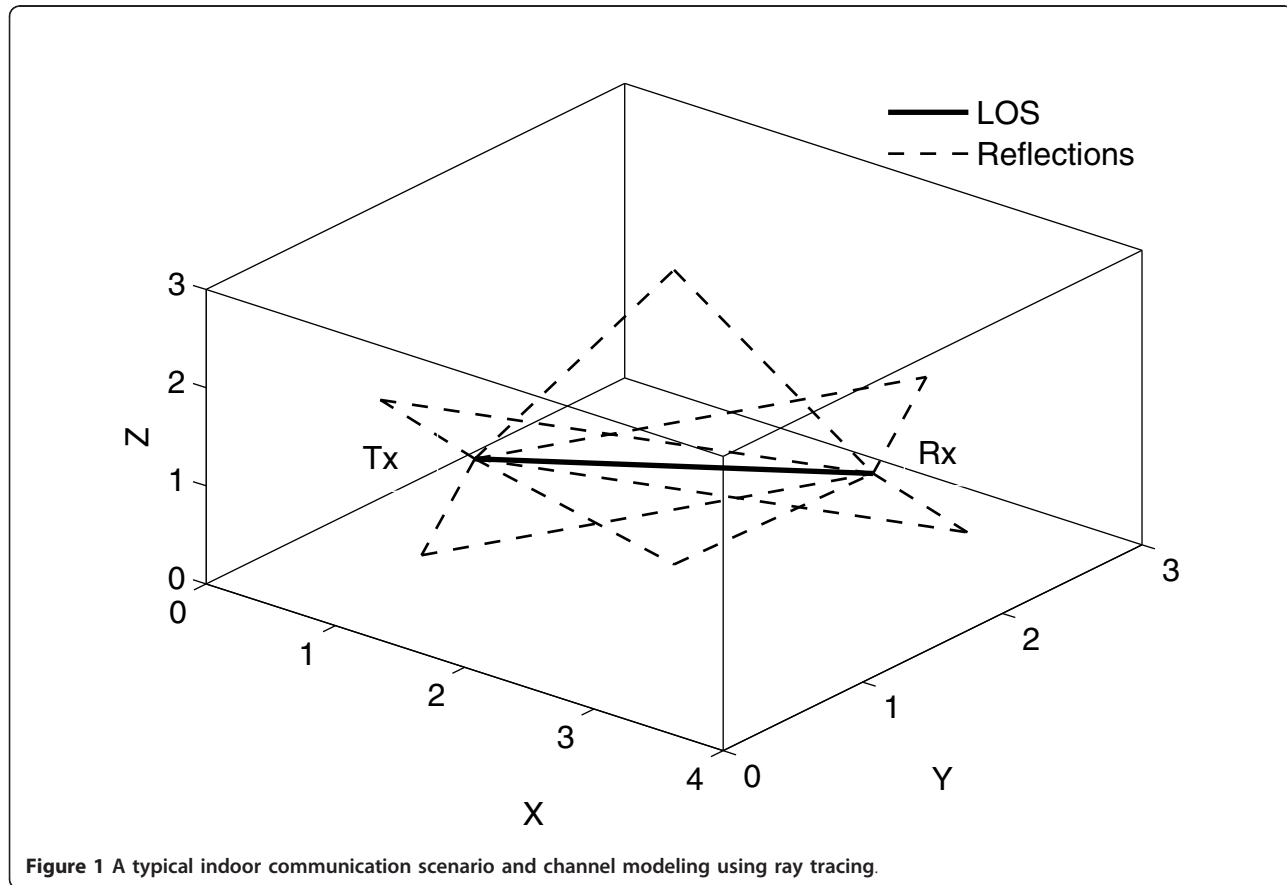


Figure 1 A typical indoor communication scenario and channel modeling using ray tracing.

variables, e.g., reflection loss, which is typically modeled as a truncated log-normal random variable with mean and variance associated with the reflection order [12], if linear polarization is assumed for each antenna. Besides, most intra-cluster parameters are randomly generated. On the other hand, for the short wavelength, it is reasonable to assume that the size of antenna array is much smaller than the size of the communication area, which leads to a similar geographic information for all links. It naturally accounts for the strong and near-deterministic LOS component and the independent realizations from reflection paths in modeling the overall channel response.

In OFDM-based systems, the narrowband subchannels are assumed to be flat fading. Thus, the equivalent channel matrix between the transmitter and receiver is given by

$$\mathbf{H} = [h_{ij}], \quad \text{with } h_{ij} = \sum_{\ell=1}^{N_{\text{rays}}} \alpha_{ij}^{(\ell)} \delta(t - \tau_0)|_{t=\tau_0} \quad (3)$$

for  $i = 1, 2, \dots, N_r$  and  $j = 1, 2, \dots, N_t$ , where the entry  $h_{ij}$  denotes the channel response between transmitter  $j$  and receiver  $i$  by aggregating all  $N_{\text{rays}}$  traced rays between them at the delay of the LOS component,  $\tau_0$ ; and  $\alpha_{ij}^{(\ell)}$  is the amplitude of  $\ell$ th ray in the corresponding link. Analytically, we can further separate the channel matrix in (3) into  $\mathbf{H}_{\text{LOS}}$  and  $\mathbf{H}_{\text{NLOS}}$  accounting for the LOS and non-LOS components, respectively

$$\mathbf{H} = \sqrt{\frac{1}{K+1}} \mathbf{H}_{\text{NLOS}} + \sqrt{\frac{K}{K+1}} \mathbf{H}_{\text{LOS}} \quad (4)$$

where the Rician  $K$ -factor indicates the relative strength of the LOS component.

We assume that the numbers of transmit and receive antennas, i.e.,  $N_t$  and  $N_r$ , are large. However, the numbers of available RF chains at the transmitter and receiver,  $n_t$  and  $n_r$ , are such that  $n_t \ll N_t$  and/or  $n_r \ll N_r$ . Hence, we need to choose a subset of  $n_t \times n_r$  transmit and receive antennas out of the original  $N_t \times N_r$  MIMO system and employ these selected antennas for data transmission (see Figure 2). Denote  $\omega$  as the set of indices corresponding to the chosen  $n_t$  transmit antennas and  $n_r$  receive antennas, and denote  $\mathbf{H}_\omega$  as the submatrix of the original MIMO channel matrix  $\mathbf{H}$  corresponding to the chosen antennas.

For data transmission over the chosen MIMO system  $\mathbf{H}_\omega$ , a transmit beamformer  $\mathbf{w} = [w_1, w_2, \dots, w_{n_t}]^T$ , with  $\|\mathbf{w}\| = 1$ , is employed. The received signal is then given by

$$\mathbf{r} = \sqrt{\rho} \mathbf{H}_\omega \mathbf{w} \mathbf{s} + \mathbf{n} \quad (5)$$

where  $s$  is the transmitted data symbol;  $\rho = \frac{E_s}{n_t N_0}$  is the system signal-to-noise ratio (SNR) at each receive antenna;  $E_s$  and  $N_0$  are the symbol energy and noise power density, respectively;  $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$  is additive white Gaussian noise vector. At the receiver, a receive beamformer  $\mathbf{u} = [u_1, u_2, \dots, u_{n_r}]^T$ , with  $\|\mathbf{u}\| = 1$ , is applied to the received signal  $\mathbf{r}$ , to obtain

$$y(\omega, \mathbf{w}, \mathbf{u}) = \mathbf{u}^H \mathbf{r} = \sqrt{\rho} \mathbf{u}^H \mathbf{H}_\omega \mathbf{w} \mathbf{s} + \mathbf{u}^H \mathbf{n}. \quad (6)$$

For a given antenna subset  $\omega$  and known channel matrix  $\mathbf{H}_\omega$ , the optimal transmit beamformer  $\mathbf{w}$  and receive beamformer  $\mathbf{u}$ , in the sense of maximum received SNR, are given by the right and left singular vectors of  $\mathbf{H}_\omega$  corresponding to the principal singular value  $\sigma_1(\mathbf{H}_\omega)$ , respectively. The optimal antenna subset  $\hat{\omega}$  is then given by the antennas whose corresponding channel submatrix has the largest principal singular value. Letting  $\mathcal{S}$  be a set each element of which corresponds to a particular choice of  $n_t$  transmit antennas and  $n_r$  receive antennas, we have

$$\hat{\omega} = \arg \max_{\omega \in \mathcal{S}} \sigma_1(\mathbf{H}_\omega). \quad (7)$$

One variation to the above antenna selection problem is that instead of the numbers of available RF chains ( $n_t, n_r$ ), we are given a minimum performance requirement, e.g.,  $\sigma_1 \geq v$ . The problem is then to find the antenna subset with the minimum size such that its performance meets the requirement.

### Problem statement

Our problem is to compute the optimal antenna set  $\hat{\omega}$  and the corresponding transmit and receiver beamformers  $\mathbf{w}$  and  $\mathbf{u}$  for a ray-traced MIMO channel realization  $\mathbf{H}$ . However, for the system under consideration,  $\mathbf{H}$  is not available to us, but rather, we only have access to the receive beamformer output  $y(\omega, \mathbf{w}, \mathbf{u})$ . This makes the straightforward approach of computing the singular value decomposition (SVD) of  $\mathbf{H}_\omega$  to obtain the beamformers impossible. Furthermore, the brute-force approach to antenna selection in (7) involves an exhaustive search over  $\binom{N_t}{n_t} \binom{N_r}{n_r}$  possible antenna subsets, which is computationally expensive.

In this paper, we propose a two-stage solution to the above problem of joint antenna selection and transmit-receive beamformer adaptation. In the first stage, we employ a discrete stochastic approximation algorithm to perform antenna selection. By setting the transmit and receive beamformers to some specific values, this method computes a bound on the principal singular value of  $\mathbf{H}_\omega$  corresponding to the current antenna subset  $\omega$ , and then iteratively updates  $\omega$  until it converges.

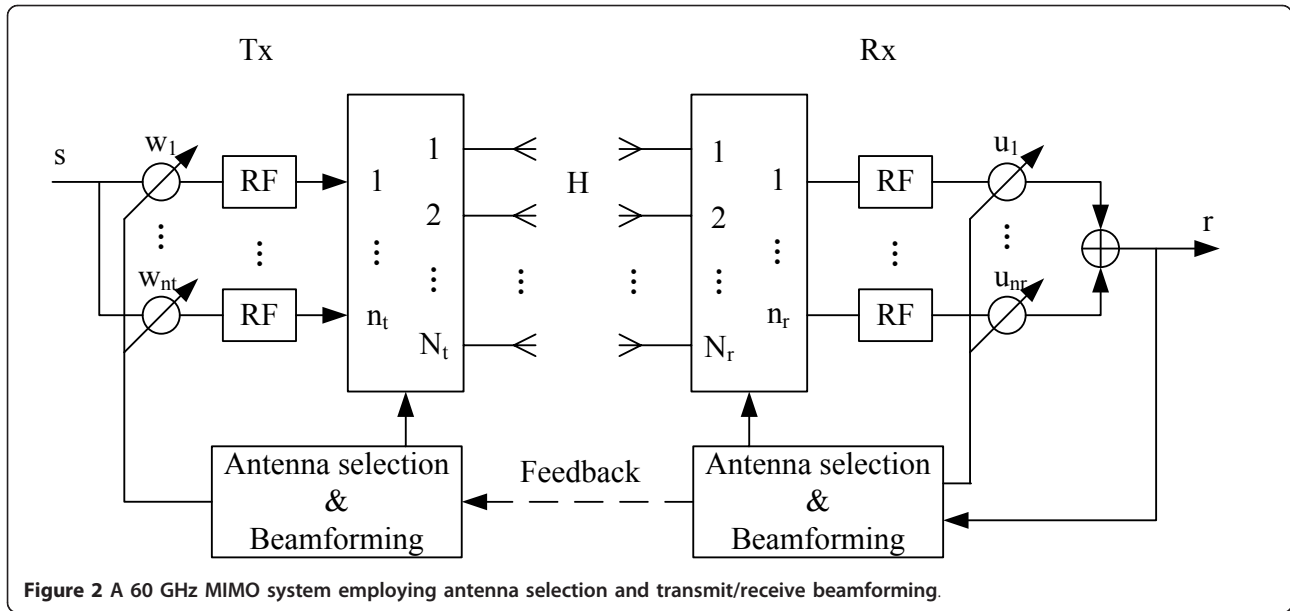


Figure 2 A 60 GHz MIMO system employing antenna selection and transmit/receive beamforming.

Once the antenna subset  $\omega$  is selected, in the second stage, we iteratively update the transmit and receive beamformers  $\mathbf{w}$  and  $\mathbf{u}$  using a stochastic gradient algorithm. At each iteration, some feedback bits are transmitted from the receiver to the transmitter via a low-rate feedback channel to inform the transmitter about the updated transmit beamformer.

In the next two sections, we discuss the detailed algorithms for antenna selection and beamformer adaptation, respectively.

### 3 Antenna selection using stochastic approximation and Gerschgorin circle

#### 3.1 The stochastic approximation algorithm

As mentioned earlier, we can only observe  $y(\omega, \mathbf{w}, \mathbf{u})$  in (6), which is a noisy function of the channel submatrix  $\mathbf{H}_\omega$ . On the other hand, the objective function to be maximized for antenna selection is the principal singular value of  $\mathbf{H}_\omega$  as in (7). If we could find a function  $\phi(\cdot)$  of  $y$  such that it is an unbiased estimate of  $\sigma_1(\mathbf{H}_\omega)$ , then we can rewrite the antenna selection problem (7) as

$$\hat{\omega} = \arg \max_{\omega \in \mathcal{S}} \mathbb{E}\{\phi(y(\omega, \mathbf{w}, \mathbf{u}))\}. \quad (8)$$

In [10], the stochastic approximation method is introduced to solve the problem of the form (8). The basic idea is to generate a sequence of the estimates of the optimal antenna subset where the new estimate is based on the previous one by moving a small step in a good direction towards the global optimizer. Through the iterations, the global optimizer can be found by means of maintaining an occupation probability vector  $\boldsymbol{\pi}$ , which indicates an estimate of the occupation

probability of one state (i.e., antenna subset). Under certain conditions, such an algorithm converges to the state that has the largest occupation probability in  $\boldsymbol{\pi}$ . Compared with the exhaustive search approach, in this way, more computations are performed on the “promising” candidates, that is, the better candidates will be evaluated more than the others.

Due to the potentially large search space in the present problem, which not only limits the convergence speed but also makes it difficult to maintain the occupation probability vector, the algorithms in [10] can become inefficient. Here, we propose a modified version of the stochastic approximation algorithm that is more efficient to implement, and more importantly, it fits naturally to a procedure for estimating the principal singular value of  $\mathbf{H}_\omega$  based on the receive beamformer output  $y(\omega, \mathbf{w}, \mathbf{u})$  only.

Specifically, we start with an initial antenna subset  $\omega^{(0)}$  and an occupation probability vector  $\boldsymbol{\pi}^{(0)} = [\omega^{(0)}, 1]^T$ , which has only one element, with the first entry serving as the index of the antenna subset and the other entry indicating the corresponding occupation probability. We divide each iteration into  $n_t + n_r$  subiterations, and in each sub-iteration, we replace one antenna in the current subset  $\omega$  with a randomly selected antenna outside  $\omega$ , resulting in a new subset  $\tilde{\omega}$  that differs from  $\omega$  by one element. By comparing their corresponding objective functions, the better subset is updated as well as the occupation probability vector. This procedure is repeated until all  $n_t + n_r$  antennas are updated.

Instead of keeping records for all candidates, we dynamically allocate and maintain record in  $\boldsymbol{\pi}$  for the new subset found in each iteration. If a subset already has a

record in  $\pi$ , the corresponding occupancy probability will be updated. Otherwise, a new element is appended in  $\pi$  with the subset index and its occupation probability. Such a dynamic scheme avoids the huge memory requirement, since typically in practice, only a small fraction of the all possible subsets is visited.

We replace the selected subset with the current subset if the current subset has a larger occupation probabilities in  $\pi$ . Otherwise, keep the selected subset unchanged, thus completes one iteration.

In general, the convergence is achieved when the number of iterations goes to infinity. In practice, when it happens that one subset is selected in a large number, say 100, consecutive iterations, the algorithm is regarded as convergent and terminated, and the last selected subset is the global (sub)optimizer. Since most of the evaluations and decisions are generally made at the receiver, a low-rate and error-free feedback channel is assumed to coordinate the transmitter via feedback information. In each subiteration, the transmitter should know in advance which transmit antennas have been left in the current subset (i.e.,  $\omega^{(n)}$ ) from last subiteration (because the current subset might have been changed in the previous subiteration), and then could generate a new subset by replacing the one with a random transmit antenna outside  $\omega^{(n)}$ . Without feedback an invalid situation might happen such that a transmit antenna, which is already assigned to one RF chain in the current subset, is selected again for another RF chain. In other words, feedback is necessary only in subiterations in which the current subset has changed for the transmit antennas during the last update in the previous subiteration. This implies that the amount of feedbacks is rather limited.

The modified stochastic approximation algorithm for antenna selection is summarized in Algorithm 1. In what follows we discuss the form of the objective function  $\varphi(\cdot)$  in (8) and its calculation.

### 3.2 Estimating the principal singular value using Gerschgorin circle

The Gerschgorin circle theorem [13] gives a range on a complex plane within which all the eigenvalues of a square matrix lie. In this section, we show that a good approximation to the largest eigenvalue can be calculated as long as the Rician  $K$ -factor is high enough. By calculating the  $G$ -circles, a simple estimator  $\varphi(\cdot)$  of the objective function in (8) is developed and employed in the stochastic approximation algorithm for antenna selection, i.e., Algorithm 1.

Denote the channel submatrix of the selected antenna subset by  $H_\omega = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_{n_t}]$ , where  $\mathbf{h}_k \in \mathbb{C}^{n_r \times 1}$  is the SIMO channel between the  $k$ th transmit antenna and the  $n_r$  receive antennas in the subset  $\omega$ . The correlation

matrix of  $H_\omega$  is then

$$\mathbf{R}_\omega = H_\omega^H H_\omega = \begin{bmatrix} \mathbf{h}_1^H \mathbf{h}_1 & \mathbf{h}_1^H \mathbf{h}_2 & \dots & \mathbf{h}_1^H \mathbf{h}_{n_t} \\ \mathbf{h}_2^H \mathbf{h}_1 & \mathbf{h}_2^H \mathbf{h}_2 & \dots & \mathbf{h}_2^H \mathbf{h}_{n_t} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{h}_{n_t}^H \mathbf{h}_1 & \mathbf{h}_{n_t}^H \mathbf{h}_2 & \dots & \mathbf{h}_{n_t}^H \mathbf{h}_{n_t} \end{bmatrix}. \quad (9)$$

Denote the eigenvalues of  $\mathbf{R}_\omega$  in descending order as  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{n_t}$ . Then, according to the Gerschgorin circles theorem [13], these  $n_t$  eigenvalues lie in at least one of the following circles

$$\{\lambda : |\lambda - \mathbf{h}_k^H \mathbf{h}_k| \leq \rho_k\}, \quad k = 1, \dots, n_t, \quad (10)$$

with the radius of the  $k$ th circle being

$$\rho_k = \sum_{\ell=1, \ell \neq k}^{n_t} |\mathbf{h}_k^H \mathbf{h}_\ell|, \quad k = 1, \dots, n_t. \quad (11)$$

The above  $n_t$  circles are centered along the positive real axis. Since the correlation matrix  $\mathbf{R}_\omega$  is positive semi-definite, all eigenvalues are located along the positive real axis within these circles, as illustrated in Figure 3. Note that from (10) to (11), a circle with a larger center coordinate implies a larger channel gain for the corresponding transmit antenna; and a circle with a smaller radius implies a smaller channel correlation between the corresponding antenna and the other selected antennas. As seen from Figure 3, the right-most point among the  $n_t$  circles is the upper bound for all eigenvalues and such a point can be used as the estimate of the largest eigenvalue of  $\mathbf{R}_\omega$ . That is,

$$\lambda_1 \leq \max_{k=1, \dots, n_t} \{ \|\mathbf{h}_k\|^2 + \sum_{\ell=1, \ell \neq k}^{n_t} |\mathbf{h}_k^H \mathbf{h}_\ell| \} \triangleq B_1. \quad (12)$$

Since the principal singular value  $\sigma_1$  of  $H_\omega$  is related to  $\lambda_1$  through  $\lambda_1 = \sigma_1^2$ , we can rewrite (7) as

$$\hat{\omega} = \arg \max_{\omega \in \mathcal{S}} \lambda_1(\mathbf{R}_\omega). \quad (13)$$

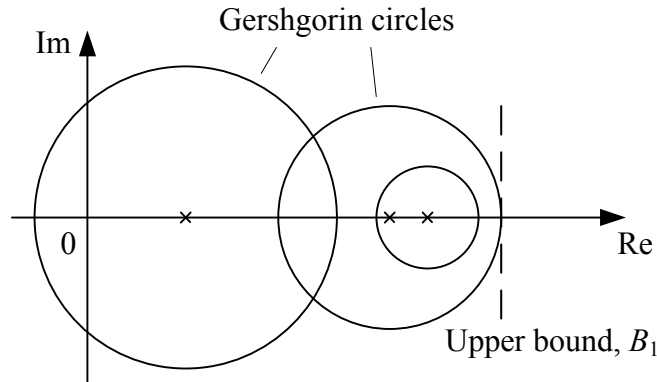
Note that,  $B_1$  is the maximum over the  $\lambda_1$  norms of the rows of  $\mathbf{R}_\omega$ . In particular, letting  $\mathbf{R}_\omega = [r_{ij}]$  we have

$$B_1 = G(\mathbf{R}_\omega) \triangleq \max_i \left\{ \sum_{j=1}^{n_t} |r_{ij}| \right\} \quad (14)$$

Next we prove a lemma that provides a useful bound on  $B_1$  and  $\lambda_1$ .

**Lemma 1** For any semi-unitary matrix  $\mathbf{U} \in \mathbb{C}^{n_r \times r}$  such that  $\mathbf{U}^H \mathbf{U} = \mathbf{I}$ , we have

$$B_1 \geq \lambda_1(\mathbf{R}_\omega) \geq \frac{1}{n_t \sqrt{\min\{n_t, r\}}} F(\mathbf{H}_\omega^H \mathbf{U} \mathbf{U}^H \mathbf{H}_\omega) \quad (15)$$



**Figure 3** An illustration of the Gerschgorin circle.

where  $F(A)$  is defined upon matrix  $A = [a_{ij}]$  such that

$$F(A) \triangleq \sum_i \sum_j |a_{ij}| \quad (16)$$

To prove the lemma, we define  $\tilde{\mathbf{R}}_\omega = \mathbf{H}_\omega^H \mathbf{U} \mathbf{U}^H \mathbf{H}_\omega$  and let  $\tilde{\mathbf{R}}_\omega = [\tilde{r}_{ij}]$ . We offer the following inequalities.

$$B_1 = G(\mathbf{R}_\omega) \geq \lambda_1(\mathbf{R}_\omega) \geq \lambda_1(\tilde{\mathbf{R}}_\omega), \quad (17)$$

where the last inequality follows upon noting the positive semi-definite ordering  $\mathbf{R}_\omega \succcurlyeq \tilde{\mathbf{R}}_\omega$ . Next, we let  $\|\tilde{\mathbf{R}}_\omega\|_F \triangleq \sqrt{\text{tr}(\tilde{\mathbf{R}}_\omega \tilde{\mathbf{R}}_\omega)}$  denote the Frobenius norm of  $\tilde{\mathbf{R}}_\omega$ . Then, since the rank of  $\tilde{\mathbf{R}}_\omega$  is no greater than  $\min\{n_b, r\}$ , it can be readily verified that

$$\lambda_1(\tilde{\mathbf{R}}_\omega) \geq \frac{1}{\sqrt{\min\{n_t, r\}}} \|\tilde{\mathbf{R}}_\omega\|_F. \quad (18)$$

Further, we have

$$\|\tilde{\mathbf{R}}_\omega\|_F = \sqrt{\sum_{i=1}^{n_t} \sum_{j=1}^{n_t} |\tilde{r}_{ij}|^2} \geq \frac{1}{n_t} \sum_{i=1}^{n_t} \sum_{j=1}^{n_t} |\tilde{r}_{ij}| \quad (19)$$

Combining (18) with (19) we have the desired result.

In our problem, only the receive beamformer output  $y(\omega, \mathbf{w}, \mathbf{u})$  in (6) is available. We will obtain an approximation to the lower bound on  $B_1, \lambda_1$  given in the right-hand side (RHS) of (15) in the following way. For each transmit antenna in the subset  $\omega, k = 1, \dots, n_b$ , we set the transmit and receive beamformers as

$$\mathbf{w} = \mathbf{e}_k, \quad \text{and} \quad \mathbf{u} = \frac{1}{\sqrt{n_r}} \mathbf{1},$$

respectively, where  $\mathbf{e}_k$  is a length- $n_t$  column vector of all zeros, except for the  $k$ -th entry which is one; and  $\mathbf{1}$  is a length- $n_r$  column vector of all ones. The transmitted symbol is set as  $s = 1$ . Then by (5)-(6), we have the

corresponding receive beamformer output given by<sup>1</sup>

$$y(k) = \sqrt{\frac{1}{n_r}} \mathbf{1}^T \mathbf{h}_k + v(k), \quad \text{with } v(k) \sim \mathcal{CN}(0, 1), \quad k = 1, \dots, n_t. \quad (20)$$

We now use the following expression to approximately lower bound  $B_1, \lambda_1$ .

$$B_2 \triangleq \frac{1}{n_t} \sum_{k=1}^{n_t} \beta(k), \quad \text{with } \beta(k) \triangleq |y(k)|^2 + \sum_{\ell=1, \ell \neq k}^{n_t} |y(k)^H y(\ell)|. \quad (21)$$

Substituting (20) into (21), we have

$$B_2 = \frac{1}{n_t} \sum_{k=1}^{n_t} \sum_{j=1}^{n_t} |y(k)^H y(j)| = \frac{1}{n_t} F(|y(1), \dots, y(n_t)|^H |y(1), \dots, y(n_t)|). \quad (22)$$

Note that in the noiseless case, we have that  $B_2$  in (22) is equal to  $\hat{B}_2$ , where

$$\hat{B}_2 = \frac{1}{n_t} \sum_{k=1}^{n_t} \sum_{j=1}^{n_t} |\mathbf{h}_k^H \mathbf{u} \mathbf{u}^H \mathbf{h}_j| = \frac{1}{n_t} F(\mathbf{H}_\omega^H \mathbf{u} \mathbf{u}^H \mathbf{H}_\omega). \quad (23)$$

Then, using Lemma 1 and its proof, we see that  $\hat{B}_2$  is indeed a lower bound on  $B_1$  as well as  $\lambda_1(\mathbf{R}_\omega)$ .

In order to mitigate the noise, for each transmit beamformer  $\mathbf{e}_k$ , we will make multiple, say  $M$  transmissions, and denote the corresponding receive beamformer outputs as  $y(k)^{(m)}, m = 1, \dots, M$ . A smoothed version of the estimator  $\beta(k)$  is then given by

$$\tilde{\beta}(k) \triangleq \frac{1}{M} \left\{ |y(k)^{(1)} y(k)^{(2)} + y(k)^{(2)} y(k)^{(3)} + \dots + y(k)^{(M)} y(k)^{(1)}| \right. \\ \left. + \sum_{\ell=1, \ell \neq k}^{n_t} \left| \sum_{m=1}^M y(k)^{(m)} y(\ell)^{(m)} \right| \right\}. \quad (24)$$

The final estimator of the lower bound on the principal eigenvalue of  $\mathbf{R}_\omega$  is then given by

$$\tilde{B}_2 \triangleq \frac{1}{n_t} \sum_{k=1}^{n_t} \tilde{\beta}(n_t) \quad (25)$$

It is easily seen that both the 1st-order and 2nd-order noise terms are averaged out in  $\tilde{B}_2$ , so that as  $M \rightarrow \infty$  we have

$$\tilde{B}_2 \rightarrow \hat{B}_2. \quad (26)$$

Recall that in the stochastic approximation algorithm for antenna selection, at each iteration, we sequentially update the transmit and receive antennas and compute the corresponding objective functions. The above approach for calculating the objective function fits naturally in this framework, since for each transmit antenna candidate, we only need to transmit a pilot signal from it and then compute the corresponding  $\tilde{\beta}(k)$ . The complete antenna selection algorithm is now summarized in Algorithm 1.

*Remark-1:* We note that a typical scenario in 60 GHz has a strongly LOS channel with  $K \gg 1$  and one dominant path, so that  $\mathbf{H}_{\text{LOS}} = \mathbf{a}\mathbf{b}^H$  is a rank one matrix. Moreover, in many applications, it is feasible to retain all receive antenna elements, so that the task reduces to selection of the optimal transmit antenna subset. In this case, neglecting  $\mathbf{H}_{\text{NLOS}}$  and the background noise (which holds for  $K, M \gg 1$ ), it can be verified that *the transmit antenna subset which maximizes  $\tilde{B}_2$  also results in the largest eigenvalue. In particular*

$$\hat{\omega} = \arg \max_{\omega \in S} \lambda_1(\mathbf{R}_\omega) \approx \arg \max_{\omega \in S} \tilde{B}_2(\omega). \quad (27)$$

where we use  $\tilde{B}_2(\omega)$  to denote the  $\tilde{B}_2$  evaluated for a particular subset and where the approximation becomes exact in the limit of large  $K, M$ .

*Remark-2:* So far, we have assumed that only one receive beamformer  $\mathbf{u} = \frac{1}{\sqrt{n_r}}\mathbf{1}$  is employed for a given choice of receive antenna subset. Suppose upto  $r$  receive beamformers  $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$  (which are columns of a  $n_r \times n_r$  unitary matrix) could be used for each transmit beamformer  $\mathbf{e}_k$ ,  $k = 1, \dots, n_t$ . Then, invoking Lemma 1 and defining  $\gamma(v, \mathbf{u}_j) = [\gamma(\omega, \mathbf{e}_1, \mathbf{u}_j), \dots, \gamma(\omega, \mathbf{e}_{n_t}, \mathbf{u}_j)]$ ,  $j = 1, \dots, r$ , we see that a better approximation can be obtained as

$$\frac{1}{n_t \sqrt{\min\{r, n_t\}}} F \left( \sum_{j=1}^r \gamma(\omega, \mathbf{u}_j)^H \gamma(\omega, \mathbf{u}_j) \right), \quad (28)$$

or its smoother version

$$\frac{1}{n_t \sqrt{\min\{r, n_t\}}} \sum_{k=1}^{n_t} \tilde{\gamma}(k) \quad (29)$$

where

$$\tilde{\gamma}(k) \triangleq \frac{1}{M} \left\{ \sum_{j=1}^r [\gamma(\omega, \mathbf{e}_k, \mathbf{u}_j)^{(1)}] \gamma(\omega, \mathbf{e}_k, \mathbf{u}_j)^{(2)} + \dots + \gamma(\omega, \mathbf{e}_k, \mathbf{u}_j)^{(M)} \gamma(\omega, \mathbf{e}_k, \mathbf{u}_j)^{(1)} \right. \\ \left. + \sum_{\ell=1, \ell \neq k}^{n_t} \left| \sum_{j=1}^r \sum_{m=1}^M \gamma(\omega, \mathbf{e}_\ell, \mathbf{u}_j)^{(m)} \gamma(\omega, \mathbf{e}_\ell, \mathbf{u}_j)^{(m)} \right| \right\}. \quad (30)$$

Finally, we note that for a given  $n_t, n_r, r$ , the channel-independent constant can be omitted when computing the metric in (25) or (30).

#### 4 Adaptive Tx/Rx beamforming with low-rate feedback

Once the antenna subset  $\mathbf{H}_\omega$  is chosen, the transmit and receive beamformers  $\mathbf{w}$  and  $\mathbf{u}$  will be computed. As mentioned in Section 2,  $\mathbf{w}$  and  $\mathbf{u}$  should be chosen to maximize the received SNR, or alternatively, to maximize the power of the receive beamformer output in (6),  $|\gamma(\omega, \mathbf{w}, \mathbf{u})|^2$ , i.e.,

$$(\hat{\mathbf{w}}, \hat{\mathbf{u}}) = \arg \max_{\substack{\mathbf{w} \in \mathbb{C}^{n_t}, \|\mathbf{w}\|=1; \\ \mathbf{u} \in \mathbb{C}^{n_r}, \|\mathbf{u}\|=1}} |\gamma(\omega, \mathbf{w}, \mathbf{u})|^2. \quad (31)$$

Since the channel matrix  $\mathbf{H}_\omega$  is not available, we resort to a simple stochastic gradient method for updating the beamformers.

##### 4.1 Stochastic gradient algorithm for beamformer update

The algorithm for the beamformer update is a generalization of [14] and is described as follows. At each iteration, given the current beamformers  $(\mathbf{w}, \mathbf{u})$ , we generate  $K_t$  perturbation vectors for the transmit beamformer,  $\mathbf{p}_j \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ ,  $j = 1, \dots, K_t$ , and  $K_r$  perturbation vectors for the receive beamformer,  $\mathbf{q}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ ,  $i = 1, \dots, K_r$ . Then for each of the normalized perturbed transmit-receive beamformer pairs

$$\left( \frac{\mathbf{w} + \beta \mathbf{p}_j}{\|\mathbf{w} + \beta \mathbf{p}_j\|}, \frac{\mathbf{u} + \beta \mathbf{q}_i}{\|\mathbf{u} + \beta \mathbf{q}_i\|} \right), \quad (32)$$

**Algorithm 1** Adaptive antenna selection using stochastic approximation and G-circle

INITIALIZATION:

$n \leftarrow 0$ ;

Select initial antenna subset  $\omega^{(0)}$  and set  $\boldsymbol{\pi}^{(0)} = [\omega^{(0)}, \mathbf{1}]^T$ ;

Transmit pilot signals from each selected transmit antenna and obtain the received signals using the selected receive antennas  $\{\gamma(k)^{(m)}, m = 1, \dots, M; k = 1, \dots, n_t + n_r\}$ ;

Compute the objective function  $\varphi(\omega^{(0)})$  using (24)-(25);

Set selected antenna subset  $\hat{\omega} = \omega^{(0)}$ .

For  $n = 1, 2, \dots$

For  $k = 1, 2, \dots, n_t + n_r$

**SAMPLING AND EVALUATION:**

Replace the  $k$ th element in  $\omega^{(n)}$  by a randomly selected antenna that is not in  $\omega^{(n)}$  to obtain a new subset  $\tilde{\omega}^{(n)}$  that differs with  $\omega^{(n)}$  by only one element;

For a newly selected transmit antenna, transmit pilot signals from it and obtain the received signals  $\{y(k)^{(m)}, m = 1, \dots, M\}$ ;

For a newly selected receive antenna, sequentially transmit pilot signals from all transmit antennas and obtain the received signals;

Recalculate the objective function  $\phi(\tilde{\omega}^{(n)})$  using (24)-(25).

**ACCEPTANCE:**

If  $\phi(\tilde{\omega}^{(n)}) > \phi(\omega^{(n)})$  Then

Update  $\omega^{(n)} = \tilde{\omega}^{(n)}$ ;

If  $\tilde{\omega}^{(n)}$  is NOT recorded in  $\pi$  Then

Append the column  $[\tilde{\omega}^{(n)}, 0]^T$  to  $\pi$  ;

EndIf

Feed back  $\omega^{(n)}$  if the update affects any transmit antenna therein

EndIf

**ADAPTIVE FILTERING:**

Set forgetting factor:  $\mu(n) = 1/n$ ;

$\pi^{(n)} = [1 - \mu(n + 1)] \pi^{(n)}$ ;

$\pi^{(n)}(\omega^{(n)}) = \pi^{(n)}(\omega^{(n)}) + \mu(n + 1)$ ;

EndFor ( $k$ )

**SELECTION:**

If  $\pi^{(n)}(\omega^{(n)}) > \pi^{(n)}(\hat{\omega})$  Then

Set  $\hat{\omega} = \omega^{(n)}$ ;

EndIf

$\omega^{(n+1)} = \omega^{(n)}$ ;  $\pi^{(n+1)} = \pi^{(n)}$ ;

EndFor ( $n$ )

where  $\beta$  is a step-size parameter, the corresponding received output power  $|y|^2$  are measured, and the effective channel gain  $|u^H H_\omega w|^2$  can be used as a performance metric independent of transmit power. Finally, the beamformers are updated using the perturbation vector pair that gives the largest output power at the receiver. The transmitter is informed about the selected perturbation vector by a  $\lceil \log K_t \rceil$ -bit message from the

receiver. The algorithm is regarded as convergent, and the iteration terminates when the performance metric fluctuates below a tolerance threshold. The algorithm is summarized as follows.

**Algorithm 2** Stochastic gradient algorithm for beamformer update

**INITIALIZATION:**

Initialize  $w^{(0)}$  and  $u^{(0)}$

For  $n = 0, 1, \dots$

**PROBING:**

Generate  $K_t$  and  $K_r$  new beamformer vectors using (32) based on  $w^{(n)}$  and  $u^{(n)}$ , respectively;

Evaluate the received power  $|y|^2$  for each one of the  $K_t K_r$  perturbed beamformer pairs;

**UPDATE AND FEEDBACK:**

Let  $p_{j^*}$  and  $q_{i^*}$  be the perturbation vectors that give the largest received power;

Feedback the index of the best transmit perturbation vector using  $\lceil \log K_t \rceil$  bits;

Update the beamformers:

$$w^{(n+1)} = (w^{(n)} + \beta p_{j^*}) / \|w^{(n)} + \beta p_{j^*}\|, \quad u^{(n+1)} = (u^{(n)} + \beta q_{i^*}) / \|u^{(n)} + \beta q_{i^*}\|.$$

EndFor

**4.2 Implementation issues**

We next discuss some implementation issues related to the above stochastic gradient algorithm for beamformer update.

**Initialization**

A good initialization can considerably speed up the convergence of the above stochastic gradient algorithm compared with random initialization. For the application considered in this paper, recall that the channel consists of a deterministic LOS component  $H_{LOS}$  and a random component. When the  $K$ -factor is high, the LOS component mostly determines the largest singular mode. Hence, we can initialize the transmit and receive beamformers as the right and left singular vectors of  $H_{LOS}$ , respectively, which we will call it a hot start.

**Parameterization**

Since both  $w$  and  $u$  have unit norms, we can represent them using spherical coordinates. Consider  $w = [w_1, w_2, \dots, w_{n_t}]^T$  as an example. Expanding  $v = [\text{Re}\{w^T\}, \text{Im}\{w^T\}]^T$ , it is equivalent to a point on the surface of the  $2n_t$ -dimensional unit sphere. Thus,  $v$  can be parameterized by  $(2n_t - 1)$ -dimensional vector  $\psi$  as follows [15]



$$v_1 = \cos \psi_1, \quad (33)$$

$$v_2 = \sin \psi_1 \cos \psi_2, \quad (34)$$

$$\vdots \quad (35)$$

$$v_{2n_t-1} = \sin \psi_1 \sin \psi_2 \cdots \sin \psi_{2n_t-2} \cos \psi_{2n_t-1}, \quad (36)$$

$$v_{2n_t} = \sin \psi_1 \sin \psi_2 \cdots \sin \psi_{2n_t-2} \sin \psi_{2n_t-1}, \quad (37)$$

$$\text{with } 0 < \psi_i < \pi, \quad 1 \leq i \leq 2n_t - 2; \quad \text{and } 0 < \psi_{2n_t-1} \leq 2\pi. \quad (38)$$

Given the vector  $\nu$  or equivalently  $\psi$ , to obtain a new perturbed weight vector near  $\nu$ , we can set an arbitrary small  $\varepsilon > 0$  and generate i.i.d. random variables  $\{\delta_i\}_{i=1}^{2n_t-2}$ , which are uniformly distributed within  $[-\frac{\varepsilon}{2}, \frac{\varepsilon}{2}]$  and another independent uniform random variable  $\delta_{2n_t-1} \in [-\varepsilon, \varepsilon]$ . Then, new parameters are obtained within some predefined boundaries, given by

$$\hat{\psi}_i = [\psi_i + \delta_i]_{a_i}^{b_i}, \quad 1 \leq i \leq 2n_t - 1, \quad (39)$$

where  $[x]_a^b$  denotes that  $x$  is confined in the interval of  $[a, b]$ , i.e.,  $[x]_a^b = x$  if  $a \leq x \leq b$ ,  $[x]_a^b = b$  if  $x > b$  and  $[x]_a^b = a$  if  $x < a$ . As a result, uniform search for the better weight vector is confined within a fixed space defined by  $[a_i, b_i]$ ,  $1 \leq i \leq 2n_t - 1$  and the range of the perturbation depends on the definition of  $\{\delta_i\}$ . For example, given a hot start, the current weight vector maybe very close to the optimizer, and it is necessary to set a smaller search region and a finer perturbation.

#### Parallel reception

Since at each iteration, the best beamformer pair is chosen out of  $K_t K_r$  combinations based on the corresponding output powers  $|y|^2$ , it would require  $K_t K_r$  transmissions. In practice, instead of switching to different the receive beamformers and making the corresponding transmissions for each transmit beamformer, we can set up  $K_r$  parallel receiver beamformers to obtain  $K_r$  receiver outputs simultaneously. Then, only  $K_t$  transmissions are needed for each iteration.

#### Conservative update

If all candidate  $K_t + K_r$  beamformers at each iteration are generated anew, then the algorithm is termed aggressive. On the other hand, a conservative strategy keeps the best transmit and receive beamformers from the previous iteration and generates  $K_t - 1$  new transmit and  $K_r - 1$  new receive beamformers for the current iteration. With a fixed step size and a single feedback bit, the advantage of the aggressive update is the quicker convergence. But with multiple feedback bits, such an advantage is less significant. Therefore, the conservative

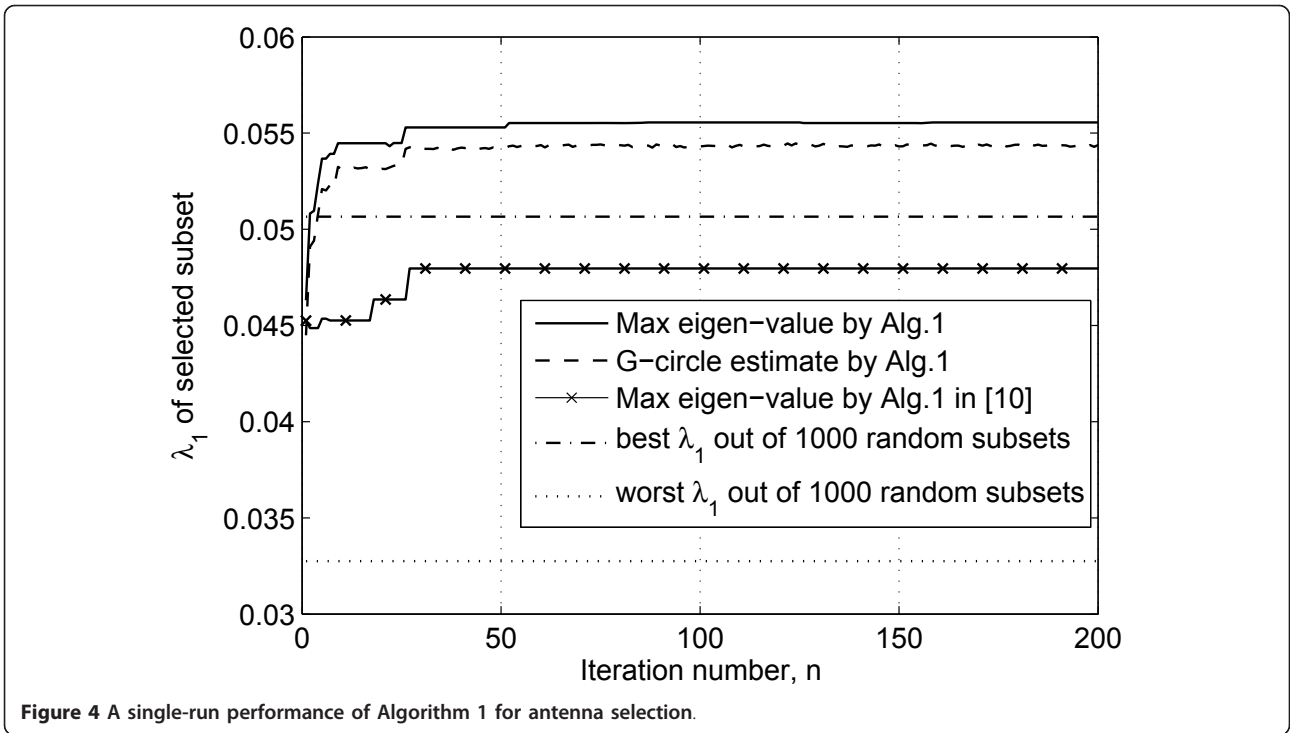
update is preferable for a finer performance upon convergence.

## 5 Simulation results

We consider an empty conference room with dimension 4m(L)  $\times$  3m(W)  $\times$  3m(H) for analysis, in which a large-scale MIMO system with  $N_t = 32$  and  $N_r = 10$  transmit and receive antennas operating at the 60 GHz band is randomly located. All the antennas are omni-directional with 20 dBi gain and vertical linear polarized. There are 10 available RF chains at both the transmitter and the receiver, i.e.,  $n_t = 10$  and  $n_r = 10$ . To generate the channel realizations, 3-D ray tracing is performed between the transceiver using the inter- and intra-cluster parameters specified for the conference room scenario in [12]. By the result of ray tracing, the  $32 \times 10$  channel matrix is gathered using (3). The channel remains static during antenna selection and beamformer update. Note that the channels simulated in the sequel are covered by Remark-1 in Section 3.2. Also, OFDM-based PHY is used as suggested in [5], where 512 subchannels divide total 2.16 GHz bandwidth. The default system SNR is assumed as  $\rho = 60$ dB. The insertion loss on signal power due to the switches between the RF chains and antennas is considered as an extra 5 dB increase in noise figure.

#### Performance of antenna selection with fixed size

The performance of Algorithm 1 for antenna selection in a single run is shown in Figure 4. Both the G-circle estimates  $\tilde{B}_2$  given by (24)-(25) as well as the actual largest eigenvalues of the selected antenna subsets are plotted for the first 200 iterations as a zoom-in view. The number of transmissions for obtaining the smoothed estimate in (24) is  $M = 20$ . Since the search space is quite large, i.e.,  $\binom{32}{10} = 64512240$ , in the same figure, we also plot the largest eigenvalues of the best and the worst subsets among 1,000 randomly selected antenna subsets. Moreover, the single-run performance of the antenna selection algorithm in [10] is also shown. In Figure 5, the average performance of 100 runs for the above schemes is plotted in a larger span of iterations. Several observations are in order. First, it is seen that the G-circle estimates are quite close to the actual largest eigenvalues, which validates the use of G-circle as a metric for antenna selection in strong line-of-sight channels. Secondly, Algorithm 1 has a much faster convergence rate than the algorithm in [10], which at each iteration picks the next candidate subset randomly and independent of the current subset, whereas Algorithm 1 searches for the next candidate subset in the neighborhood of the current subset. Thirdly, Algorithm 1 can lock onto a near-optimal antenna subset very quickly, e.g., in 10-20 iterations, and it significantly outperforms

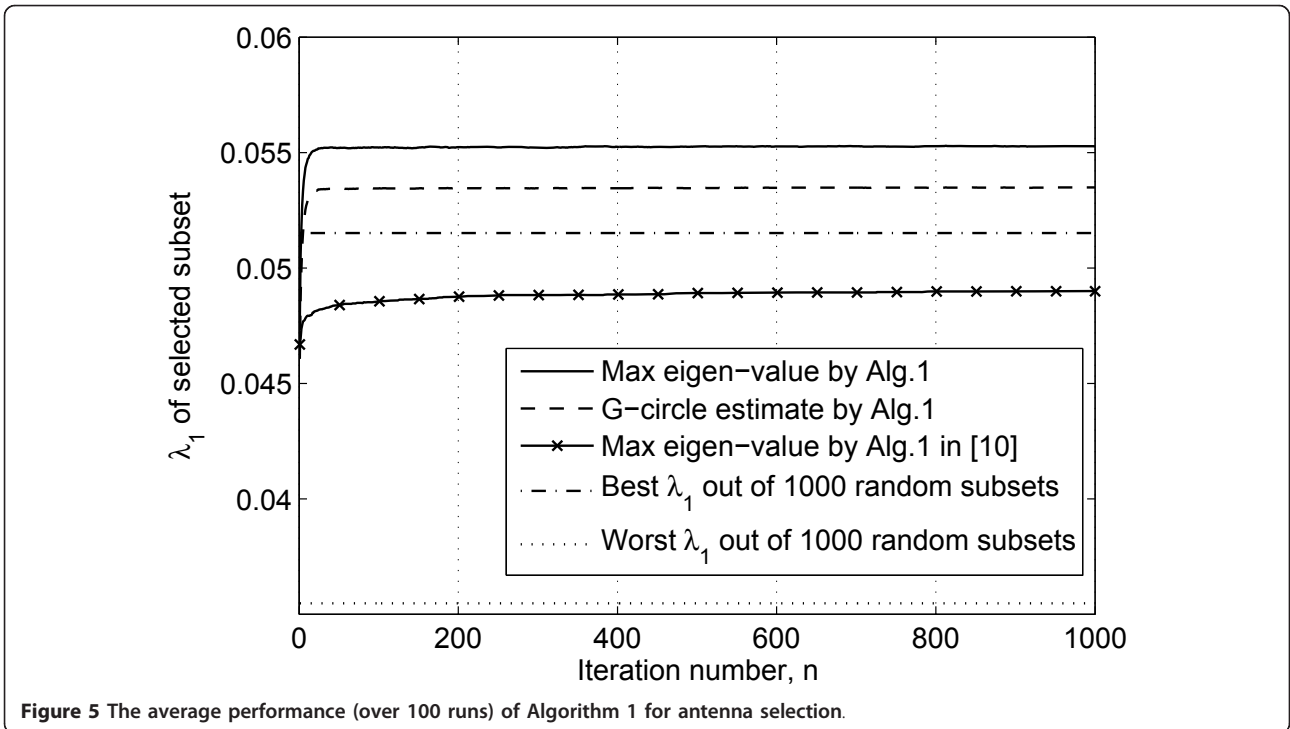


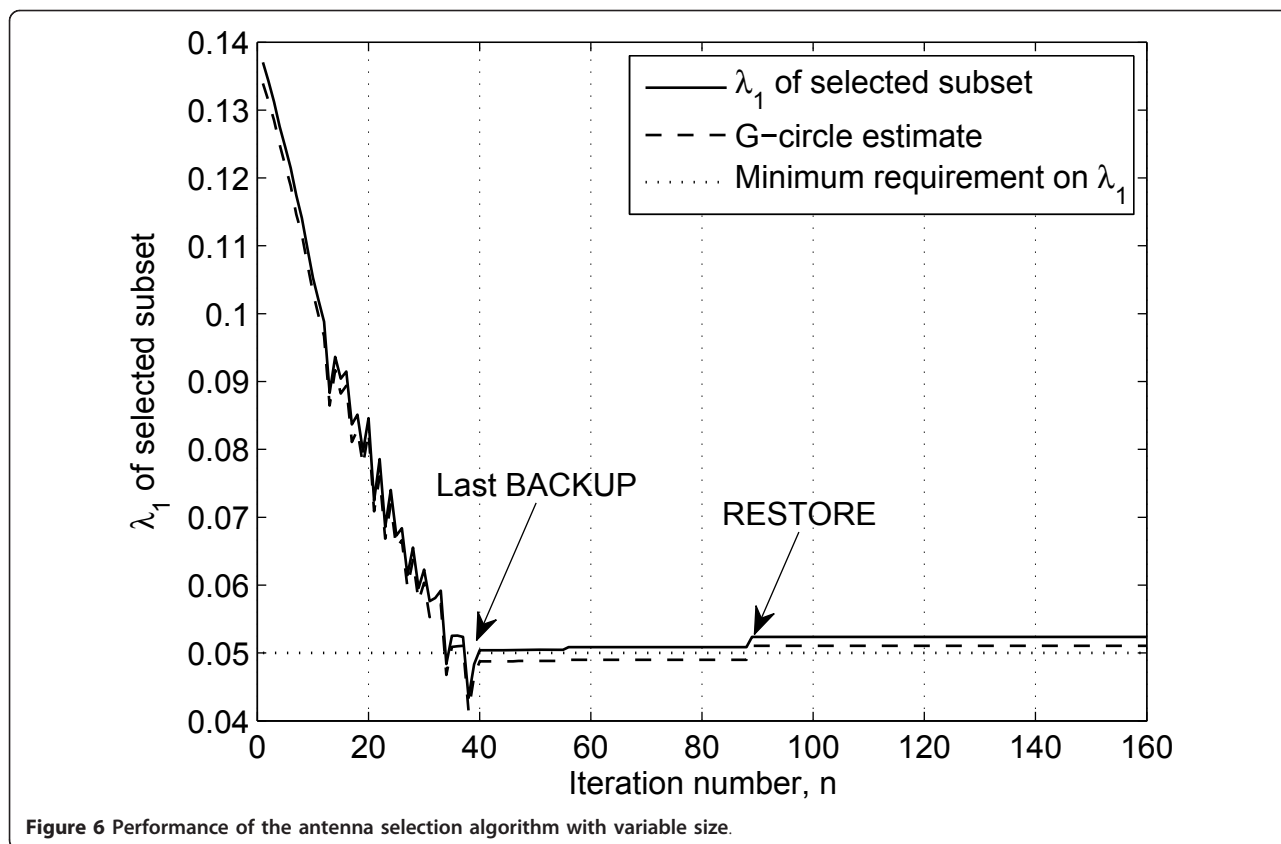
the exhaustive search over a large number (e.g., 1,000) of subsets.

**Performance of antenna selection with variable size**

Figure 6 shows the performance of the adaptive antenna selection given a minimum requirement, and

Figure 7 shows the selected subset sizes. The simulation starts with the largest subset containing all the 32 transmit antennas. The number of selected antennas is then decreased by one at each step. For a given size of the selected subset, say  $n_s$ , Algorithm 1 is performed to generate a sequence of, e.g., 20, antenna subsets. If





all of them meet the requirement, i.e.,  $\lambda_1 \geq 0.05$ , we backup the current parameters (i.e., current iteration number, selected subset, probability vector, etc.), and then terminate the current iteration and set  $n_t \leftarrow n_t - 1$ . If again the condition is met, a new backup is performed to simply replace the previous one. As shown in Figures 6 and 7,  $n_t$  keeps decreasing until the selected subsets do not meet the requirement for a number of iterations, e.g., 50, which means the last  $n_t$  is the desired minimum size  $n_t^*$ . Therefore, by restoring the last backup data, the terminated iteration in Algorithm 1 is resumed till the optimal antenna subset with size  $n_t^*$  is found. In Figure 6, we show both the G-circle estimates and the exact largest eigenvalues of the selected subsets. Since the estimation provides a lower bound to the largest eigenvalue and G-circle, a margin should be taken into consideration when setting the minimum performance requirement in order to guarantee that the actual performance of the selected subset meets the requirement with minimum number of selected antenna.

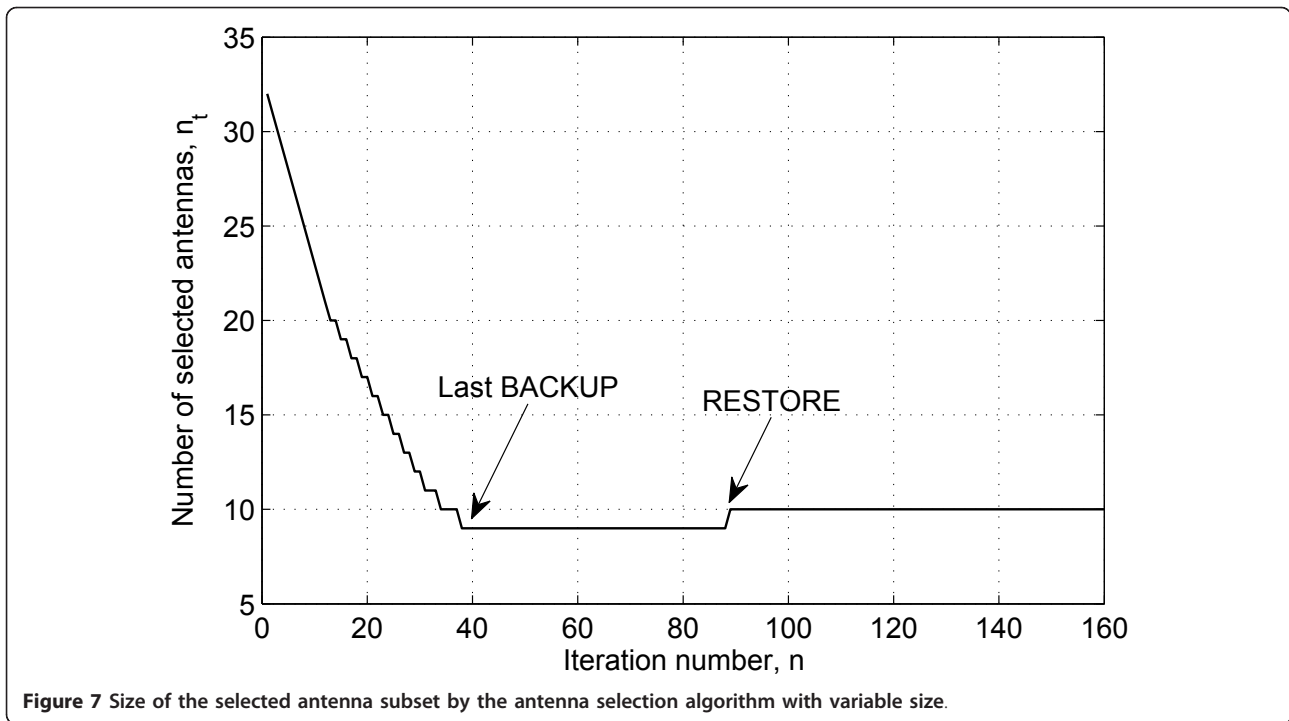
#### Performance of adaptive beamforming

Figure 8 shows one run of Algorithm 2 for adaptive transmit and receive beamforming upon a selected channel submatrix. The number of perturbations at

the transmitter and receiver are  $K_t = 16$  and  $K_r = 16$ , respectively; hence, the number of feedback bits is  $\log 16 = 4$ . The conservative update with step size 0.05 is used. The performance of the Algorithm 2 with a random initialization and hot start is plotted, as well as the exact largest eigenvalue of the channel obtained by SVD. It is seen that the hot start can significantly speed up the convergence. In Figure 9, we compare the performance of Algorithm 2 with different number of feedback bits, i.e.,  $K_t = 2, 4, 8, 16$  and fixed  $K_r = 16$ . It is seen that by employing more feedback bits, the convergence rate can be significantly increased. Similar behavior can be seen if we fix  $K_t$  and vary  $K_r$ .

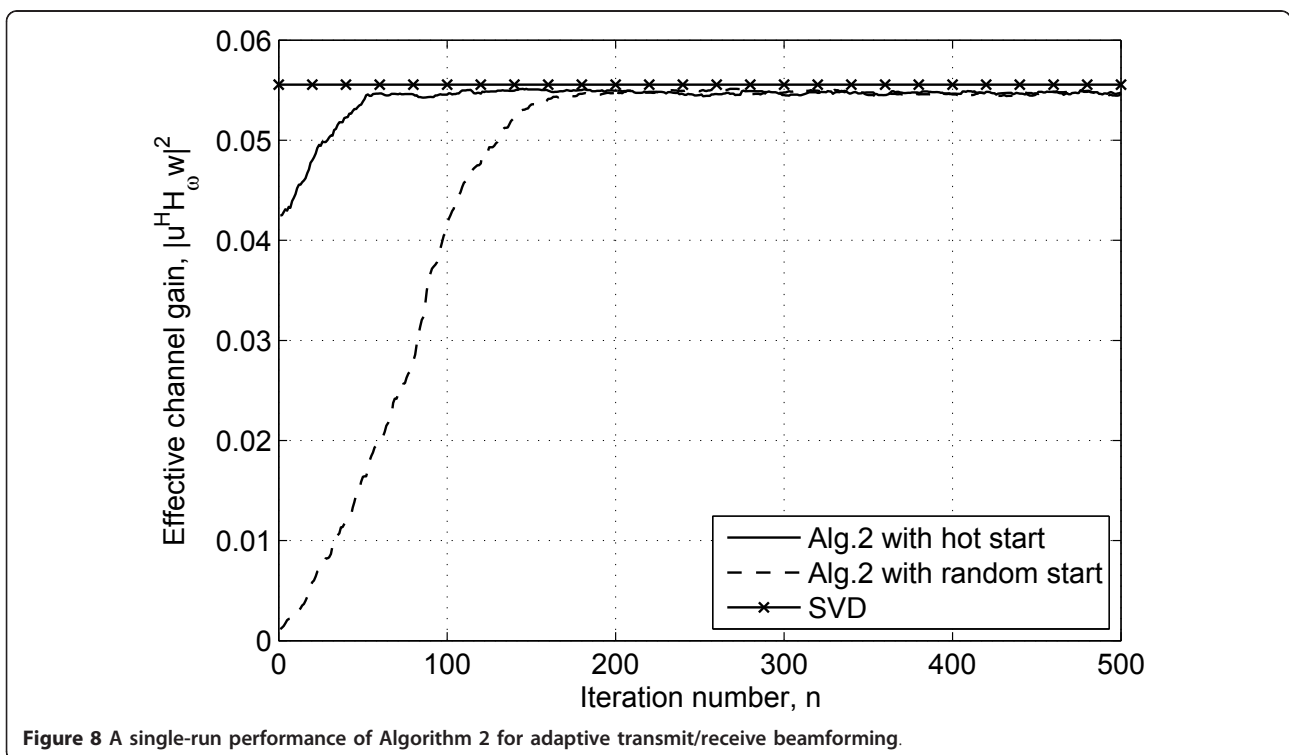
#### Overall performance of adaptive antenna selection and beamforming

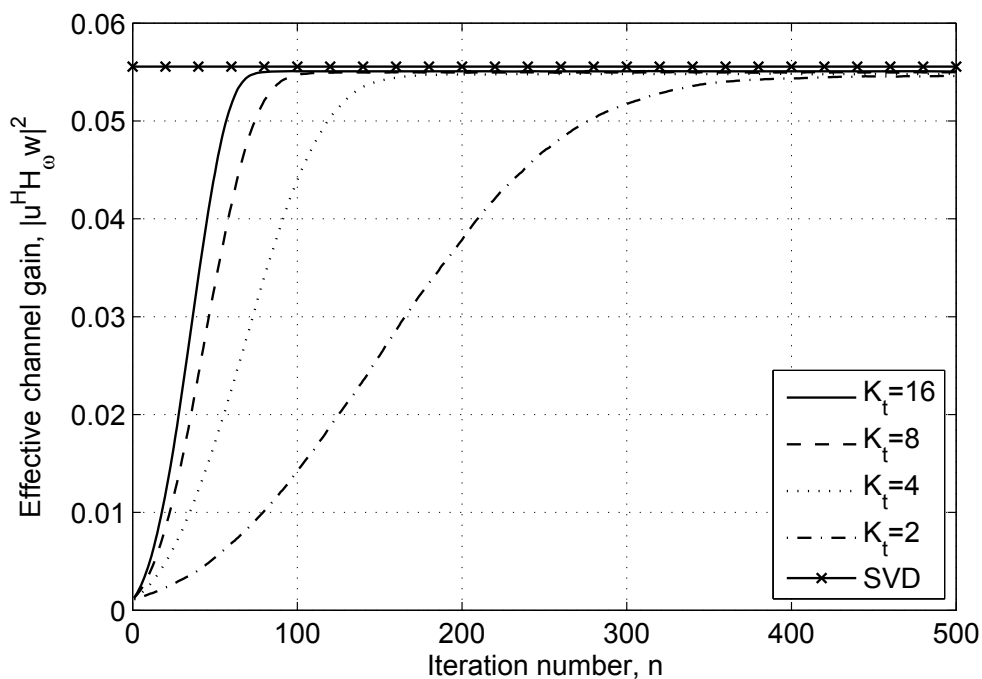
The effective channel gain,  $|u^H H_o w|^2$ , is a metric indicating the overall performance by associating the adaptive antenna selection with beamforming. In this simulation, the transceiver is dropped at 100 random locations with minimum distance 30 cm in the room independently, and we generate the channel realizations therein using 3-D ray-tracing technique. By running the proposed adaptive algorithms for these drops, Figure 10 shows the averaged effective channel gain against



different system SNR. For comparison, the non-adaptive solutions, i.e., the best out of 1,000 random subsets and SVD are also investigated. We have several observations. First, for both beamforming algorithm (Algorithm 2 and SVD), Algorithm 1 outperforms the best out of random

1,000 subsets at the high SNR region, but its performance is inferior at the lower SNR. This is because when the SNR is low, the accuracy and reliability can not be guaranteed in estimating the objective function value and ranking the subsets, which prevents the

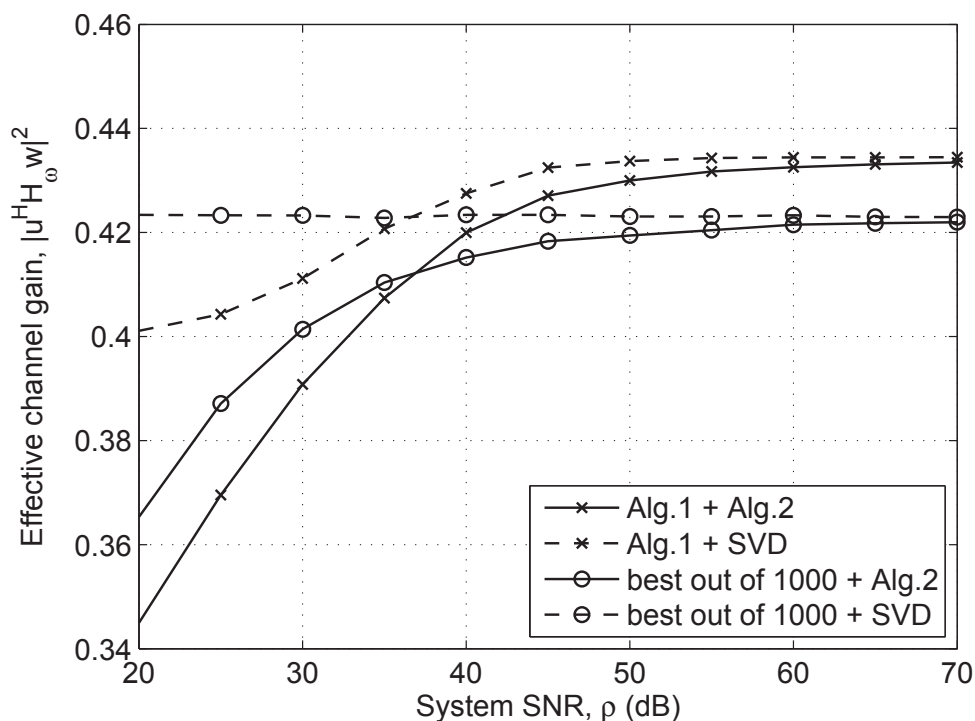




**Figure 9** Performance of Algorithm 2 with different feedback bits.

adaptive algorithms from converging to better solutions. Second, for the same reason, Algorithm 2 is inferior to SVD at lower SNR, but approaching SVD at high SNR by using both antenna selection strategies. It implies that the accuracy in objective function estimation is a

key factor that largely affects the convergence and overall performance. From (24), we see that it is feasible to increase  $M$  in order to guarantee the estimation accuracy and maintain the overall performance in the low SNR region.



**Figure 10** Overall performance for Algorithm 1 and 2. Average over 100 random Tx/Rx locations and channel realizations.

## 6 Conclusions

We have proposed a sequential antenna selection algorithm and an adaptive transmit/receive beamforming algorithm for large-scale MIMO systems in the 60 GHz band. One constraint of the system under consideration is that the receiver can only access a linear combination of the receive antenna outputs, which makes the traditional antenna selection schemes based on the channel matrix not applicable. The proposed antenna selection method uses a bound on the largest singular value of the channel matrix based on the Gerschgorin circle. The method is particularly useful over the 60 GHz channel, which has a strong line-of-sight component, and it employs a discrete stochastic approximation technique to quickly lock onto a near-optimal antenna subset. We have also proposed an adaptive joint transmit and receive beamforming technique based on the stochastic gradient method that makes use of a low-rate feedback channel to inform the transmitter about the selected beam. Simulation results show that both the proposed antenna selection and the adaptive beamforming techniques exhibit fast convergence and near-optimal performance.

## Note

<sup>1</sup>Note that in obtaining (20) without loss of generality we have absorbed  $\rho$  into  $H$ .

## Acknowledgements

The authors wish to acknowledge financial support from the National Key Specialized Project of China (2009ZX03003-008-02) and the National Science Foundation of China (60902043).

## Author details

<sup>1</sup>Xi'an Jiao Tong University, Xi'an, China <sup>2</sup>NEC Labs America, Princeton, NJ, 08540, USA <sup>3</sup>Electrical Engineering Department, Columbia University, New York, NY, 10027, USA

## Competing interests

The authors declare that they have no competing interests.

Received: 24 November 2010 Accepted: 11 August 2011

Published: 11 August 2011

## References

1. CH Doan, S Emami, DA Sobel, AM Niknejad, RW Brodersen, Design considerations for 60 GHz CMOS radios. *IEEE Commun Mag.* **42**(12), 132–140 (2004)
2. SK Yong, CC Chong, An overview of multigigabit wireless through millimeter wave technology: potentials and technical challenges. *EURASIP J Wirel Commun Netw.* **2007**(1), 50–50 (2007)
3. A Seyedi, On the capacity of wideband 60 GHz channels with antenna directionality, in *IEEE GLOBECOM 2007* (Washington D.C., 2007), pp. 4532–4536
4. J Nsenga, W Van Thillo, F Horlin, A Bourdoux, R Lauwereins, Comparison of OQPSK and CPM for communications at 60 GHz with a nonideal front end. *EURASIP J Wirel Commun Netw.* **2007**(1), 51–51 (2007)
5. IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements. Part 15.3: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for High Rate Wireless Personal Area Networks (WPANs) Amendment 2: Millimeter-wave-based Alternative Physical Layer Extension, *IEEE Std 802.15.3c-2009 (Amendment to IEEE Std 802.15.3-2003)* (2009)

6. Draft standard for IEEE 802.11ad, IEEE P802.11ad/D0.1, June 2010
7. M Gharavi-Alkhanjari, AB Gershman, Fast antenna subset selection in MIMO systems. *IEEE Trans Sig Proc.* **52**(2), 339–347 (2004). doi:10.1109/TSP.2003.821099
8. A Gorokhov, D Gore, A Paulraj, Receive antenna selection for MIMO flat-fading channels: Theory and algorithms. *IEEE Trans Inform Theory* **49**(10), 2687–2696 (2003). doi:10.1109/TIT.2003.817458
9. AF Molisch, MZ Win, YS Choi, JH Winters, Capacity of MIMO systems with antenna selection. *IEEE Trans Wirel Commun.* **4**(4), 1759–1772 (2005)
10. I Berenguer, X Wang, V Krishnamurthy, Adaptive MIMO antenna selection via discrete stochastic optimization. *IEEE Trans Sig Proc.* **53**(11), 4315–4329 (2005)
11. C Choi, E Grass, R Kraemer, T Derham, S Roblot, L Cariou, P Christin, Beamforming training for IEEE 802.11ad. doc.: IEEE 802.11-10/0493r1. (May 2010)
12. A Maltsev, Channel models for 60 GHz WLAN systems. doc.: IEEE 802.11-09/0334r8. (2010)
13. R Horn, C Johnson, *Matrix Analysis* (Cambridge University Press, New York, NY, 1985)
14. BC Banister, JR Zeidler, A simple gradient sign algorithm for transmit antenna weight adaptation with feedback. *IEEE Trans Sig Proc.* **51**(5), 1156–1171 (2003). doi:10.1109/TSP.2002.808104
15. X Wang, V Krishnamurthy, J Wang, Stochastic gradient algorithms for design of minimum error-rate linear dispersion codes in MIMO wireless systems. *IEEE Trans Sig Proc.* **54**(4), 1242–1255 (2006)

doi:10.1186/1687-1499-2011-59

**Cite this article as:** Dong et al.: Adaptive antenna selection and Tx/Rx beamforming for large-scale MIMO systems in 60 GHz channels.

*EURASIP Journal on Wireless Communications and Networking* 2011 **2011**:59.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)