

Receive Antenna Selection for MIMO Systems over Correlated Fading Channels

Yangyang Zhang, Chunlin Ji, Wasim Q. Malik, *Senior Member, IEEE*,
Dominic C. O'Brien, and David J. Edwards

Abstract—In this letter, we propose a novel receive antenna selection algorithm based on cross entropy optimization to maximize the capacity over spatially correlated channels in multiple-input multiple-output (MIMO) wireless systems. The performance of the proposed algorithm is investigated and compared with the existing schemes. Simulation results show that our low complexity algorithm can achieve near-optimal results that converge to within 99% of the optimal results obtained by exhaustive search. In addition, the proposed algorithm achieves near-optimal results irrespective of the mutual relationship between the number of transmit and receive antennas, the statistical properties of the channel and the operating signal-to-noise ratio.

Index Terms—Channel capacity, correlated channel, cross entropy optimization (CEO), MIMO wireless systems, receive antenna selection.

I. INTRODUCTION

MULTIPLE - INPUT MULTIPLE - OUTPUT (MIMO) wireless systems can dramatically increase the channel capacity through the extra degrees of freedom provided by multiple antenna arrays. In [1], it was demonstrated that the capacity of MIMO systems increases linearly with $\min(N_T, N_R)$, where N_T and N_R denote the number of transmit and receive antennas. However, the higher performance of MIMO systems comes at the expense of increased hardware requirements and computational complexity due to multiple radio frequency (RF) chains required. In order to reduce the hardware cost and preserve the advantages of MIMO systems, a promising technique referred to as antenna selection is presented in [2]. With this method, the RF chains can be optimally connected to the best subset of the transmitter (or receiver) antennas. It has been demonstrated that the system performance using antenna selection techniques is better than the full-complexity systems with the same number of antennas but without selection [2]. However, the superior performance obtained by antenna selection is at the cost of additional

computational complexity which grows linearly with $\binom{M}{L}$, where M and L denote the total and selected number of antennas, respectively [3], [4].

Recently, a number of algorithms have been developed for selecting the optimal antenna subset in MIMO wireless systems. For example, in [5], Heath et al. derived a signal-to-noise ratio (SNR) based antenna selection criterion to improve the performance of MIMO systems with linear receivers. In [6], Gore et al. presented antenna selection algorithms to minimize the average probability of error (APE) and to maximize the average throughput. However, an exhaustive search method for antenna selection was used, which is computationally prohibitive for a large array size¹, and is not suitable for implementation in practical systems. To address this problem, some simplified antenna selection algorithms have also been developed, such as norm-based selection (NBS), which can be useful due to its low complexity [2], [7]. Sub-optimal algorithms were presented at a low complexity for receive antenna selection in [4]. Antenna selection approaches based on the theory of optimization were derived in [8]. However, the aforementioned studies have assumed that the MIMO channels are independently fading, which is not strictly true for real propagation environments. For example, in the case of insufficient spacing between antennas or scattering with a small angular spread, the channel capacity will be significantly degraded due to spatial correlation [10]. Thus far, only a small set of published literature investigates antenna selection for correlated channels [11], [12].

In this letter, we formulate the antenna selection problem as a combinatorial optimization problem. Cross entropy optimization (CEO) is used for antenna subset selection at the receiver to maximize the channel capacity². The CEO method is so named due to its relation with the Kullback-Leibler distance [13] which is also termed the cross entropy. It is a principled adaptive importance sampling technique devised by Rubinstein [14] to estimate the probabilities of rare events in complex stochastic networks. It was then extended to solve complicated combinatorial optimization problems by considering an optimal event as a rare event, such as non-deterministic polynomial time (NP) hard problems [15]. While most stochastic algorithms for combinatorial optimization are based on local search, the CEO method is a global random search procedure whose global convergence has been proven

Manuscript received December 14, 2007; revised June 17, 2008, January 13, 2009, and June 7, 2009; accepted June 22, 2009. The associate editor coordinating the review of this letter and approving it for publication was J. Andrews.

This work was supported by EPSRC grant GR/T21769/01 and a K. C. Wong Scholarship from the University of Oxford.

Y. Y. Zhang, D. C. O'Brien, and D. J. Edwards are with the Department of Engineering Science, University of Oxford, Parks Road Oxford OX1 3PJ, UK (e-mail: {yangyang.zhang, dominic.obrien, david.edwards}@eng.ox.ac.uk).

C. Ji is with the Institute of Statistics and Decision Sciences, Duke University, Durham, North Carolina 27708 (e-mail: chunlin.ji@duke.edu).

W. Q. Malik is with the Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139. He is also with the Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114 (e-mail: wqm@mit.edu).

Digital Object Identifier 10.1109/TWC.2009.071404

¹Choosing L out of M available antennas leads to a total of $\frac{M!}{L!(M-L)!}$ possible combinations for antenna selection at the transmitter or receiver. For example, if $L = 4$ and $M = 16$, 1820 combinations have to be examined to obtain the optimal antenna selection subset.

²The proposed CEO method can be also used for the transmit antenna selection with small revisions.

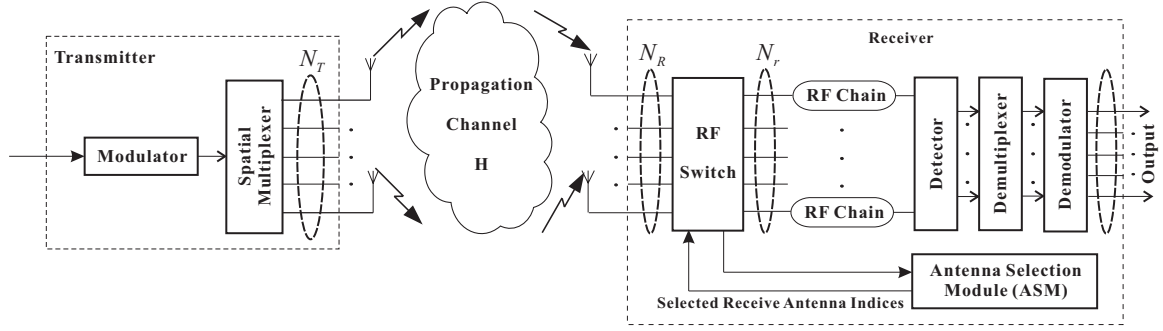


Fig. 1. Block diagram of the MIMO system with receive antenna selection.

in [15]. The main contribution of this letter is to present a novel receive antenna algorithm based on the CEO method to maximize the capacity over spatially correlated channels. Simulation results indicate that the near-optimal performance of the proposed antenna selection algorithm is not sensitive to the relationship between the number of transmit antennas and the number of selected receive antennas, the statistical properties of channels and the signal-to-noise ratio (SNR), as has been the case with previous approaches.

Notation: The following notation is used in this letter. Boldface uppercase and lowercase letters denote matrices and vectors. Plain lowercase letters denote scalars. The superscripts $(\cdot)^T$ and $(\cdot)^H$ represent the transpose and Hermitian operation. $\mathbb{E}[\cdot]$ denotes the statistical expectation. $\text{Tr}(\cdot)$ and $\|\cdot\|_F$ denote the trace and Frobenius norm. \mathbf{I}_m is an $m \times m$ identity matrix. $\mathcal{C}^{M \times N}$ refers to an $M \times N$ matrix with complex entries and $\det(\cdot)$ denotes the determinant operation.

II. SIGNAL MODEL

Consider a narrowband MIMO wireless system, shown in Figure 1, with N_T transmit and N_R receive antennas. The channel is assumed to be flat Rayleigh fading and slow varying with additive white Gaussian noise (AWGN) at the receiver. Then the corresponding received signal is given by [4]

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{v}, \quad (1)$$

which relates the received signal vector $\mathbf{y} = [y_1, \dots, y_{N_R}]^T \in \mathcal{C}^{N_R \times 1}$ to the transmitted signal vector $\mathbf{s} = [s_1, \dots, s_{N_T}]^T \in \mathcal{C}^{N_T \times 1}$ with covariance $\mathbf{Q} = \mathbb{E}[\mathbf{s}\mathbf{s}^H]$. The vector $\mathbf{v} \in \mathcal{C}^{N_R \times 1}$ represents additive complex Gaussian noise with zero mean, variance N_0 and independently and identically distributed (i.i.d.) entries. \mathbf{H} denotes the $N_R \times N_T$ fading channel matrix whose entries, h_{ij} ($i = 1 \dots N_R; j = 1 \dots N_T$), are the complex fading coefficients between the i^{th} receive and j^{th} transmit antenna.

In order to evaluate the performance of the proposed algorithm for correlated channels, the “one ring” model for Rayleigh channels [10] is adopted in this letter. Specifically, we assume that the correlation is present only at the receiver. In other words, the rows of \mathbf{H} are correlated while the columns of \mathbf{H} are independent. According to the Kronecker model, the corresponding channel matrix can be written as

$$\mathbf{H} = \mathbf{R}_r^{\frac{1}{2}} \mathbf{G}, \quad (2)$$

where $\mathbf{G} \in \mathcal{C}^{N_R \times N_T}$ is the spatially white MIMO channel matrix with zero-mean unit-variance i.i.d. complex Gaussian entries. $\mathbf{R}_r^{\frac{1}{2}}$ is the Hermitian square root of $\mathbf{R}_r \in \mathcal{C}^{N_R \times N_R}$ which is defined by

$$\mathbf{R}_r = \mathbb{E}[\mathbf{H}\mathbf{H}^H]. \quad (3)$$

According to the “one ring” model, the entries of the correlation matrix, $\mathbf{R}_r(i, j)$, represent the spatial correlation between the i^{th} and j^{th} receive antennas and can be approximated by $J_0(2\pi \Delta |i - j| d/\lambda_c)$, where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind, λ_c is the carrier wavelength, Δ is the angular spread and d is the antenna spacing.

We assume that perfect channel state information (CSI) is available at the receiver but not at the transmitter, and thus equal power allocation is used at the transmit array. Then, the capacity of the MIMO channel is given by [1]

$$C = \log_2 \det(\mathbf{I}_{N_R} + \frac{\eta}{N_T} \mathbf{H}\mathbf{H}^H), \quad (4)$$

where η is the average SNR.

III. RECEIVE ANTENNA SELECTION

A. Problem Statement

Let us denote the number of total and selected receive antennas by N_R and N_r respectively ($N_r \leq N_R$), the set of all $|\mathcal{A}| = \binom{N_R}{N_r}$ antenna subsets as $\Omega = \{\omega_1, \dots, \omega_{|\mathcal{A}|}\}$ and the indicators of the selected subset of receive antennas by

$$\omega_q = \{I_i\}_{i=1}^{N_R}, \quad \{I_i\} \in \{0, 1\}, \quad \text{for } q = 1, 2, \dots, |\mathcal{A}|, \quad (5)$$

where i is the index of the rows of \mathbf{H} and the indicator function I_i indicates that the i^{th} row of \mathbf{H} is selected, i.e., the i^{th} receive antenna is selected. The receive vector associated with the selection can be written as

$$\mathbf{y}_{\omega_q} = \mathbf{H}_{\omega_q} \mathbf{s}_{\omega_q} + \mathbf{v}_{\omega_q} = [\mathbf{R}_r^{\frac{1}{2}}]_{\omega_q} \mathbf{G}_{\omega_q} \mathbf{s}_{\omega_q} + \mathbf{v}_{\omega_q}, \quad (6)$$

where $\mathbf{y}_{\omega_q} \in \mathcal{C}^{N_r \times 1}$, $\mathbf{s}_{\omega_q} \in \mathcal{C}^{N_T \times 1}$ and $\mathbf{v}_{\omega_q} \in \mathcal{C}^{N_r \times 1}$ denote the received signal, transmitted signal and noise vectors associated with the selection, respectively. $\mathbf{H}_{\omega_q} \in \mathcal{C}^{N_r \times N_T}$, $\mathbf{G}_{\omega_q} \in \mathcal{C}^{N_r \times N_T}$ and $[\mathbf{R}_r^{\frac{1}{2}}]_{\omega_q} \in \mathcal{C}^{N_r \times N_r}$ denote the correlated channel, the spatially white channel and the receive correlation matrices after the selection, respectively.

B. Selection Criteria

In order to estimate instantaneous channels correctly, the coherence time of channels is assumed to be long enough that the fading coefficients are constant over the entire block and change independently from one block to the next according to the ‘‘one ring’’ spatial correlation model. Therefore, the optimal selected receive antenna index, ω^* , is selected out of Ω through the training sequence and changes from one block to another [2].

1) *Instantaneous CSI (ICSI) Selection Criterion:* Assuming that instantaneous CSI is only available at the receiver, the capacity associated with antenna selection is

$$C(\omega_q) = \log_2 \det(\mathbf{I}_{N_r} + \frac{\eta}{N_T} \mathbf{H}_{\omega_q} \mathbf{H}_{\omega_q}^H). \quad (7)$$

Given the ICSI, we can define the performance function as $S_{ICSI}(\omega_q) = \log_2 \det(\mathbf{I}_{N_r} + \frac{\eta}{N_T} \mathbf{H}_{\omega_q} \mathbf{H}_{\omega_q}^H)$. Therefore, maximizing the capacity associated with the receive antenna selection is equivalent to maximizing

$$\mathcal{P}_1 : \arg \max_{\omega_q \in \Omega} S_{ICSI}(\omega_q). \quad (8)$$

Since computing the ICSI selection criterion involves singular value decomposition, its complexity is $\mathcal{O}(\min\{N_r, N_T\} N_r N_T)$ [16].

2) *Norm-based Selection (NBS) Criterion:* At low SNR, (7) can be approximated by

$$\begin{aligned} C(\omega_q) &\approx \log_2 \left(1 + \frac{\eta}{N_T} \text{Tr}(\mathbf{H}_{\omega_q} \mathbf{H}_{\omega_q}^H) \right) \\ &= \log_2 \left(1 + \frac{\eta}{N_T} \|\mathbf{H}_{\omega_q}\|_F^2 \right). \end{aligned} \quad (9)$$

We define the performance function as $S_{NBS}(\omega_q) = \|\mathbf{H}_{\omega_q}\|_F$. Therefore, maximizing the capacity associated with the receive antenna selection is equivalent to maximizing

$$\mathcal{P}_2 : \arg \max_{\omega_q \in \Omega} S_{NBS}(\omega_q), \quad (10)$$

where $(\|\mathbf{H}_{\omega_q}\|_F)^{1/2}$ indicates the power of the channel matrix \mathbf{H}_{ω_q} . Although the NBS criterion cannot guarantee an optimal capacity performance, because of its low complexity ($\mathcal{O}(N_r N_T)$) [16], it is still a good candidate for antenna selection [2], [5], [7].

3) *Spatial Correlation Selection (SCS) Criterion:* When the channel is fast fading, channel estimation becomes a difficult task [17]. Moreover, in such a situation, a large number of training sequences have to be used to obtain the optimal receive antenna index, ω^* . These training sequences not only degrade the spectral efficiency but also increase the hardware complexity [18]. Compared with the ICSI, it is easier to estimate and track the spatial correlation because of its slow variation. This makes the SCS criterion desirable for practical MIMO systems with antenna selection. Specifically, at high SNR, (7) can be approximated as

$$C(\omega_q) \approx \log_2 \det \left(\frac{\eta}{N_T} \mathbf{H}_{\omega_q} \mathbf{H}_{\omega_q}^H \right). \quad (11)$$

Substituting (2) into (11) and using the eigenvalue decomposition (EVD) of $[\mathbf{R}_r]_{\omega_q}$, we have [11]

$$\begin{aligned} C(\omega_q) &\approx N_T \log_2 \left(\frac{\eta}{N_T} \right) + \log_2 \det(\mathbf{G}_{\omega_q} (\mathbf{G}_{\omega_q})^H) \\ &\quad + \log_2 \det([\mathbf{R}_r]_{\omega_q}). \end{aligned} \quad (12)$$

We define the performance function as $S_{SCS}(\omega_q) = \det([\mathbf{R}_r]_{\omega_q})$. Therefore, when instantaneous CSI is not available, maximizing the capacity is equivalent to maximizing

$$\mathcal{P}_3 : \arg \max_{\omega_q \in \Omega} S_{SCS}(\omega_q). \quad (13)$$

The computational complexity of the SCS criterion is $\mathcal{O}(N_r^2)$.

C. The Cross Entropy Optimization (CEO) Method

The most straightforward approach to obtain the optimal receive antenna subset, ω^* , is by exhaustive search. However, because of its high computational complexity, it becomes prohibitive for MIMO systems with large arrays. In order to reduce the complexity, we formulate the antenna selection problem as a combinatorial optimization problem as follows:

$$\omega^* = \arg \max_{\omega_q \in \Omega} S(\omega_q), \quad (14)$$

where ω^* denotes the global optimum of the objective function, $S(\omega_q)$. Here, $S(\omega_q)$ represents the performance functions of $S_{ICSI}(\omega_q)$, $S_{NBS}(\omega_q)$ or $S_{SCS}(\omega_q)$.

After transforming (14) into a combinatorial optimization problem, an iterative algorithm can be used to solve it. The idea of the CEO method is to associate a stochastic estimation problem with the optimization problem (14). Let us define a collection of indicator functions $\{I_{\{S(\omega_q) \geq r\}}\}$ in the solution space Ω for various thresholds (or levels) $r \in \{S(\omega_q) : \omega_q \in \Omega\}$, and a number of Bernoulli probability density functions given by

$$f(\omega_q, \mathbf{p}) = \prod_{i=1}^{N_R} p_i^{I_i(\omega_q)} (1 - p_i)^{1 - I_i(\omega_q)}, \quad (15)$$

where p_i indicates the probability of i^{th} receive antenna to be chosen. $I_i(\omega_q)$ is the indicator for the i^{th} element of ω_q . For a given probability distribution \mathbf{v} , we associate (14) with the following stochastic estimation

$$\begin{aligned} \ell(r) &= \mathbb{P}_{\mathbf{v}}(S(\omega_q) \geq r) = \sum_{\omega_q \in \Omega} I_{\{S(\omega_q) \geq r\}} f(\omega_q, \mathbf{v}) \\ &= \mathbb{E}_{\mathbf{v}}[I_{\{S(\omega_q) \geq r\}}], \end{aligned} \quad (16)$$

where $\ell(r)$ is the probability $S(\omega_q) \geq r$ and $I_{\{S(\omega_q) \geq r\}}$ is given by

$$I_{\{S(\omega_q) \geq r\}} = \begin{cases} 1, & \text{if } S(\omega_q) \geq r \\ 0, & \text{otherwise.} \end{cases} \quad (17)$$

A natural way to estimate ℓ in (16) is to use a crude Monte Carlo (CMC) simulation by drawing a set of random samples $\{\omega_q^{(n)}\}_{n=1}^N$ from $f(\cdot, \mathbf{v})$, and then the unbiased estimator of ℓ is

$$\hat{\ell} = \frac{1}{N} \sum_{n=1}^N I_{\{S(\omega_q^{(n)}) \geq r\}}. \quad (18)$$

For a large value of r (i.e. $r \rightarrow r^*$), the above problem is a rare event simulation, where $r^* = \max_{\omega_q \in \Omega} S(\omega_q)$. In order to obtain the optimum, a large number of samples ($N \rightarrow \infty$) have to be drawn to obtain an accurate estimation, because most of the samples are not effective in calculating $\hat{\ell}$. Therefore, the CMC method is not suitable for practical applications due to its high complexity.

An alternative way to estimate ℓ is through the importance sampling (IS) technique, drawing a set of random samples $\{\omega_q^{(n)}\}_{n=1}^N$ from an importance distribution $g(\omega_q)$. Then the unbiased estimator of ℓ is

$$\hat{\ell} = \frac{1}{N} \sum_{n=1}^N I_{\{S(\omega_q^{(n)}) \geq r\}} \frac{f(\omega_q^{(n)}, \mathbf{v})}{g(\omega_q^{(n)})}. \quad (19)$$

It is well known that the optimal $g^*(\omega_q)$ is given by [15]

$$g^*(\omega_q) = \frac{I_{\{S(\omega_q) \geq r\}} f(\omega_q, \mathbf{v})}{\ell}. \quad (20)$$

It is convenient to choose $g(\omega_q)$ from the parameterized family of densities $\{f(\cdot, \mathbf{p})\}$. The idea of CEO is to choose the parameter \mathbf{p}^* such that the Kullback-Leibler divergence³, which is also referred as the cross entropy, between g^* and f is minimal [15]. Minimizing the Kullback-Leibler divergence is equivalent to solving the following maximization problem [15]⁴

$$\max_{\mathbf{p}} \int_{\Omega} g^*(\omega_q) \ln f(\omega_q; \mathbf{p}) d\omega_q. \quad (21)$$

Substituting (20) into (21), we have

$$\max_{\mathbf{p}} \int_{\Omega} \frac{I_{\{S(\omega_q) \geq r\}} f(\omega_q, \mathbf{v})}{\ell} \ln f(\omega_q; \mathbf{p}) d\omega_q, \quad (22)$$

which is equivalent to

$$\mathbf{p}^* = \arg \max_{\mathbf{p}} \mathbb{E}_{\mathbf{v}} [I_{\{S(\omega_q^{(n)}) \geq r\}} \ln f(\omega_q; \mathbf{p})]. \quad (23)$$

Generally it is intractable to obtain a closed-form solution for the optimal parameter \mathbf{p}^* , as (23) involves an integration with respect to the density function $f(\omega_q, \mathbf{v})$. But \mathbf{p}^* can be estimated by the following stochastic program [15]

$$\hat{\mathbf{p}}^* = \arg \max_{\mathbf{p}} \frac{1}{N} \sum_{n=1}^N I_{\{S(\omega_q^{(n)}) \geq r\}} \ln f(\omega_q^{(n)}; \mathbf{p}), \quad (24)$$

where $\omega_q^{(n)}$ are the samples drawn from $f(\omega_q; \mathbf{v})$. Let $\hat{\mathcal{D}}(\mathbf{p}) = \frac{1}{N} \sum_{n=1}^N I_{\{S(\omega_q^{(n)}) \geq r\}} \ln f(\omega_q^{(n)}; \mathbf{p})$ and we have

$$\max_{\mathbf{p}} \hat{\mathcal{D}}(\mathbf{p}) = \frac{1}{N} \sum_{n=1}^N I_{\{S(\omega_q^{(n)}) \geq r\}} \ln(f(\omega_q^{(n)}, \mathbf{p})). \quad (25)$$

³The Kullback-Leibler divergence between two probability distributions $g(x)$ and $f(x)$ is defined as [13]

$$\mathcal{D}(g, f) = \mathbb{E}_g \left[\ln \frac{g(x)}{f(x)} \right] = \int g(x) \ln g(x) dx - \int g(x) \ln f(x) dx$$

⁴The integration with respect to $\omega_q \in \Omega$ is a summation when ω_q is discrete as in our case. But for generality, it is expressed in the form of integration.

To find the maximum of $\hat{\mathcal{D}}(\mathbf{p})$, we set $\frac{\partial \hat{\mathcal{D}}(\mathbf{p})}{\partial \mathbf{p}} = 0$. Consequently, we have the update rule as follow:

$$p_i = \frac{\sum_{n=1}^N I_{\{S(\omega_q^{(n)}) \geq r\}} I_i(\omega_q^{(n)})}{\sum_{n=1}^N I_{\{S(\omega_q^{(n)}) \geq r\}}} \quad \text{for } i = 1, 2, \dots, N_R. \quad (26)$$

The update equation (26) is iteratively used with the aim to generate a sequence of increasing thresholds $r^{(0)}, r^{(1)}$, until convergence to the global optimum r^* (or to a value close to it) is achieved. At the t^{th} iteration, a new vector $\mathbf{p}^{(t)}$ is used to draw a set of new samples, which provide better estimates of r . The vector $\mathbf{p}^{(t)}$ is then updated by these samples. This process stops when the stopping criterion is reached. A flowchart of the proposed receive selection algorithm based on the CEO method is described as follows:

Receive Antenna Selection Algorithm based on the CEO Method

Step 1: Start with an initial value $\mathbf{p}^{(0)} = \{p_i^{(0)}\}_{i=1}^{N_R}$, $p_i^{(0)} = \frac{1}{2}$. Set the iteration counter $t := 1$;

Step 2: Randomly generate samples $\{\omega_q^{(n)}\}_{n=1}^N$ from the density function $f(\cdot, \mathbf{p}^{(t-1)})$;

Step 3: Calculate the performance functions $\{S(\omega_q^{(n,t)})\}_{n=1}^N$ and order them from largest to smallest, $S^{(1)} \geq \dots \geq S^{(N)}$. Let $r^{(t)}$ be the $(1 - \rho)^{\text{th}}$ sample quantile of the performances: $r^{(t)} = S^{(\lceil (1-\rho)N \rceil)}$, where $\lceil \cdot \rceil$ is the ceiling operation.

Step 4: Update the parameter $\mathbf{p}^{(t)}$ via

$$p_i^{(t)} = \frac{\sum_{n=1}^N I_{\{S(\omega_q^{(n,t)}) \geq r^{(t)}\}} I_i(\omega_q^{(n,t)})}{\sum_{n=1}^N I_{\{S(\omega_q^{(n,t)}) \geq r^{(t)}\}}}. \quad (27)$$

Step 5: If the stopping criterion is satisfied⁶, then stop; otherwise set $t := t + 1$ and go back to step 2.

Note: In order to prevent occurrences of 0s and 1s in the parameter matrix \mathbf{p} , we introduce a *smoothing factor* λ and change the updating procedure to

$$\mathbf{p}^{(t)} := \lambda * \mathbf{p}^{(t)} + (1 - \lambda) * \mathbf{p}^{(t-1)}. \quad (28)$$

Clearly, when $\lambda = 1$, we have the original updating formulation. The convergence proof of the algorithm is shown in the Appendix.

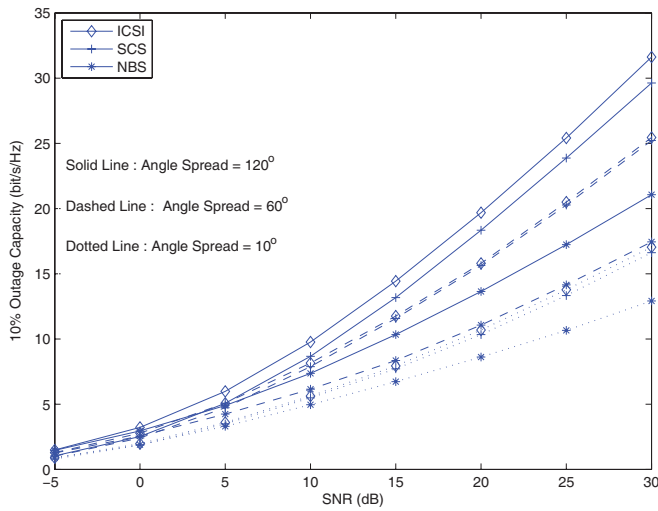
IV. SIMULATION RESULTS

In order to compare and validate the performance of the proposed CEO algorithm, simulations were performed over 10,000 channel realizations using algorithms based on the ICSI, NBS and SCS criteria. For the ‘‘one ring’’ correlated channel model, we assume that a broadside linear array is used at the receiver [10], the antenna spacing (d) is $\lambda/2$ and the directions of arrival (DOA) are uniformly distributed.

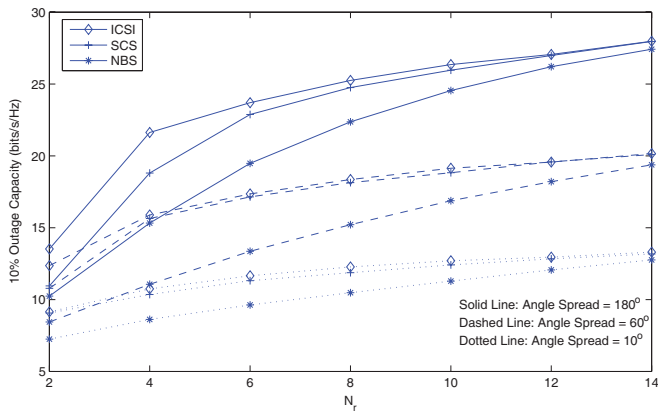
These three selection criteria offer a tradeoff between the performance and complexity. The ICSI selection criterion has

⁵The algorithm converges without the constraint of the starting point, but for simplicity we set $p_i^{(0)} = \frac{1}{2}$.

⁶The stopping criterion is $|r^{(t)} - r^{(t-1)}| \leq \beta$ where β is the stopping threshold and set as 10^{-2} in this letter.



(a) 10% outage capacity versus SNR with $N_R = 16$ and $N_r = N_T = 4$ for various angle spreads (Δ).



(b) 10% outage capacity versus N_r with $N_R = 16$, $N_T = 4$ and SNR = 20 dB for various angle spreads (Δ).

Fig. 2. Performance comparison between three selection criteria by exhaustive search.

the best performance but has the highest hardware and computational complexity, while the NBS criterion has the lowest complexity but this is achieved at the cost of performance. The SCS criterion is a possible compromise, but its performance should be close to the ICSI criterion if it is to be useful.

In order to investigate this, an exhaustive search is used to find the optimal antenna subset (ω^*) using each of the three criteria. Fig. 2 shows the 10% outage capacity as a function of the SNR and the number of selected receive antennas (N_r). From Fig. 2(a), it can be seen that the performance of the ICSI selection criterion nearly coincides with that of the SCS criterion over a wide range of SNR at small angle spread ($\Delta \leq 60^\circ$) and diverges at large angle spread, for example, $\Delta = 120^\circ$. Moreover, from Fig. 2(a), we find that the gap in outage capacity between the ICSI and SCS criteria is roughly fixed at various angle spread values for a wide range of SNR, which indicates that the performance difference between ICSI and SCS will be not significantly influenced by SNR. From

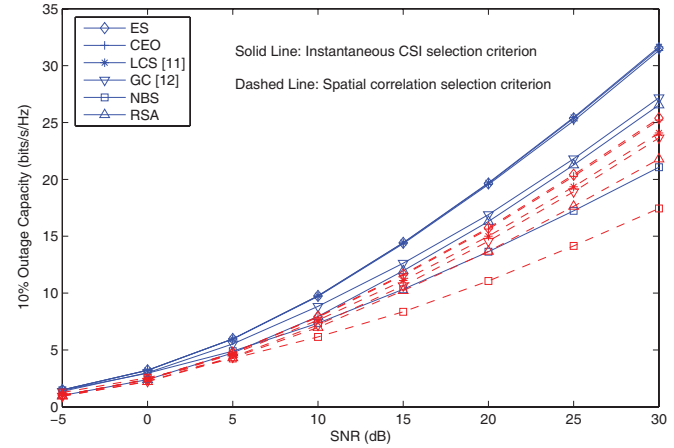


Fig. 3. 10% outage capacity versus SNR with $N_R = 16$ and $N_r = N_T = 4$ at $\Delta = 120^\circ$ based on the instantaneous CSI selection criterion (Solid Line) and $\Delta = 60^\circ$ based on the spatial correlation selection criterion (Dashed Line).

Fig. 2(b), it can also be seen that the gap in outage capacity between ICSI and SCS decreases as N_r increases regardless of the values of angle spread.

As a result, the simulation results from Fig. 2 show that the outage capacity performance of SCS is close to that of ICSI at large N_r or small angle spread. In this letter, we assume that SCS can replace ICSI for receive antenna selection at $N_r \geq 6$ or $\Delta \leq 60^\circ$ when $N_R = 16$ and $N_T = 4$.

Fig. 3 shows the 10% outage capacity versus SNR with $N_R = 16$ and $N_r = N_T = 4$ at $\Delta = 120^\circ$ and $\Delta = 60^\circ$. Based on the analysis in Fig. 2, SCS can replace ICSI to obtain near-optimal results at a small angle spread. Thus, SCS is used when $\Delta = 60^\circ$ while ICSI is used when $\Delta = 120^\circ$. The results indicate that the outage capacity achieved by the CEO algorithm is nearly the same as that by exhaustive search (ES) for a wide range of SNR. The NBS algorithm has near-optimal performance in the low SNR region ($\text{SNR} \leq 5\text{dB}$). However, when the value of SNR increases, the performance of the NBS algorithm is no longer optimal and even worse than the random selection algorithm (RSA) when $\text{SNR} \geq 10\text{dB}$. Hence, in the high SNR regime with spatial correlation, the NBS algorithm is not suitable for antenna selection. Fig. 3 also shows receive antenna selection by a low complexity selection (LCS) method [11] and Gerschgorin circles (GC) method [12] for comparison. From the figure, it can be seen that the LCS method obtains near-optimal capacity performance for the ICSI selection criterion but suffers a performance loss for the SCS criterion. Compared with the CEO algorithm and LCS method, the capacity performance obtained by the GC method is inferior for both the ICSI selection and SCS criteria.

The 10% outage capacity versus N_r with $N_R = 16$ and SNR = 20 dB at $\Delta = 120^\circ$ and $\Delta = 30^\circ$ is shown in Fig. 4. It can be seen that the CEO algorithm can obtain near-optimal performance for both the ICSI and SCS and this performance is independent of the selected receive antenna array size (N_r). The LCS method can also obtain near-optimal performance for the ICSI but not for the SCS, especially when $N_r \geq 6$. Compared with the LCS method, the GC method exhibits superior performance for the SCS when $N_r \leq 6$ and becomes inferior when $N_r \geq 8$. The results in Fig. 5 illustrate the

TABLE I
COMPLEXITY COMPARISONS FOR VARIOUS ANTENNA SELECTION ALGORITHMS WITH $N_R = 16$, $N_T = 4$, $\eta = 20$ dB

(N_R, N_r)	Number of samples (N)	Number of iterations (t)	CEO	LCS [11]	GC [12]	ES
(16, 2)	15	5	75	133	15	120
(16, 4)	18	5	90	126	78	1820
(16, 6)	20	5	100	115	165	8008
(16, 8)	20	5	100	100	252	12870
(16, 10)	20	5	100	81	315	8008
(16, 12)	18	5	90	58	330	1820
(16, 14)	15	5	75	31	273	120

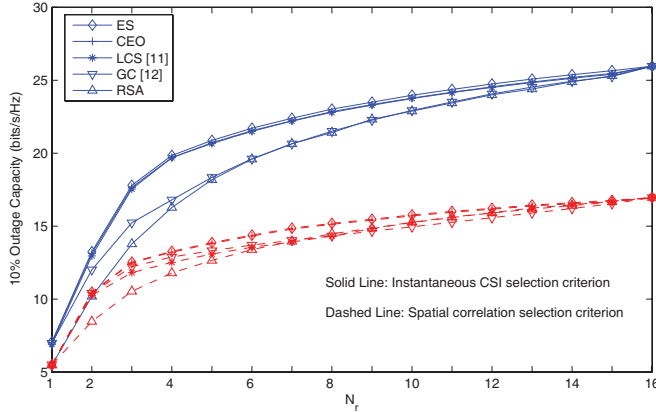


Fig. 4. 10% outage capacity versus N_r with $N_R = 16$ and SNR = 20 dB at $\Delta = 120^\circ$ based on the instantaneous CSI selection criterion (Solid Line) and $\Delta = 30^\circ$ based on the spatial correlation selection criterion (Dashed Line).

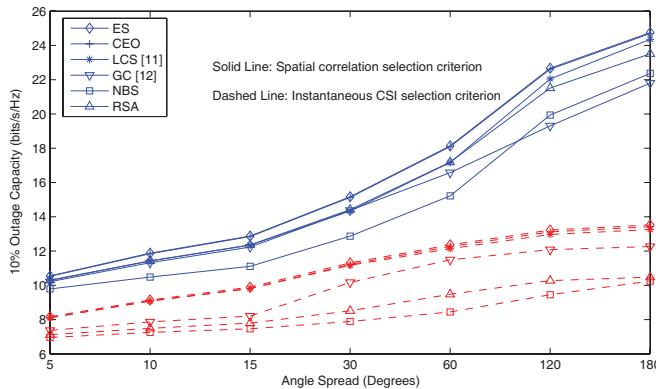


Fig. 5. 10% outage capacity versus the angle spread (Δ) with $N_R = 16$, $N_r = 2$, $N_T = 4$ based on the instantaneous CSI selection criterion (Dashed Line) and $N_R = 16$, $N_r = 8$, $N_T = 4$ based on the spatial correlation selection criterion (Solid Line) at SNR = 20 dB.

outage capacity versus the angle spread (Δ) with $N_r = 2$ and $N_r = 8$ at SNR = 20 dB. It can be seen that the CEO algorithm achieves nearly the same outage capacity as ES for both ICSI and SCS and this near-optimal performance is independent of the angle spread. The LCS method obtains near-optimal capacity for the ICSI but lower capacity for the SCS. Compared with the CEO algorithm, the capacity obtained by the GC method is considerably lower for both ICSI and SCS. Detailed complexity comparisons among the CEO, LCS, GC and ES methods are shown in Table I in terms of the total number of functional evaluations, $S(\omega_q)$. It can be seen that the CEO algorithm has much lower complexity than ES in all situations. In addition, according to results in Fig.

3- 5 and Table I, we can conclude that the proposed CEO algorithm can obtain better performance than the LCS [11] and GC [12] methods with comparable complexity.

V. CONCLUSION

In this letter, we have presented a novel receive antenna selection algorithm based on cross entropy optimization (CEO) to maximize the channel capacity over spatially correlated channels. Simulations demonstrate that the proposed algorithm can obtain near-optimal results with rapid convergence. In addition, we find that the proposed algorithm performs well irrespective of the SNR, the angle spread, the selected receive antenna array size and the mutual relationship between the transmit and selected receive antenna array size.

APPENDIX

CONVERGENCE PROOF OF THE PROPOSED RECEIVE ANTENNA SELECTION ALGORITHM

To begin, we define the Bernoulli p.d.f. for the t^{th} iteration of the i^{th} antenna (in ω) as

$$f_{i,t}(\omega, \mathbf{p}) \triangleq p_{i,t}^{\omega_{i,t}} (1 - p_{i,t})^{1 - \omega_{i,t}}, \quad (29)$$

where $\omega_{i,t}$ denotes the i^{th} element of ω at the t^{th} iteration.

Assume the following condition is satisfied

$$\lambda_t \geq \frac{t}{t+1} \quad (30)$$

for some $T \geq 0$. Without loss of generality, let $T \geq 1$. Then, according to (28) and $t \geq T$, we have

$$\begin{aligned} p_{i,t} &\geq \prod_{m=0}^{t-1} \lambda_m \cdot p_{i,0} \\ &\geq \prod_{m=0}^{T-1} \lambda_m \cdot \prod_{m=T}^{t-1} \frac{m}{m+1} \cdot p_{i,0} \\ &= \prod_{m=0}^{T-1} \lambda_m \cdot p_{i,0} \cdot \frac{T}{t} = \kappa \cdot \frac{p_{i,0}}{t}, \end{aligned} \quad (31)$$

where κ is a constant and equal to $\prod_{m=0}^{T-1} \lambda_m \cdot T$. Since $\kappa \geq 0$, we have $p_{i,t} \geq \frac{p_{i,0}}{t}$, which further implies, with probability one, that

$$f_{i,t}(\omega, \mathbf{p}) \geq \frac{f_{i,0}(\omega, \mathbf{p})}{t}, \quad \text{for } t = 1, 2, 3, \dots \quad (32)$$

The probability of lack of convergence to the optimal point ω^* is therefore bounded by

$$\begin{aligned} \text{Prob}(\omega^{(t)} \neq \omega^*) &= \prod_{t=1}^{\infty} \prod_{i=1}^{N_R} (1 - f_{i,t}(\omega^*, \mathbf{p})) \\ &\leq \prod_{t=1}^{\infty} \prod_{i=1}^{N_R} \left(1 - \frac{f_{i,0}(\omega^*, \mathbf{p})}{t}\right) \\ &\leq \prod_{t=1}^{\infty} \prod_{i=1}^{N_R} e^{-\frac{f_{i,0}(\omega^*, \mathbf{p})}{t}} \\ &= e^{-\sum_{i=1}^{N_R} f_{i,0}(\omega^*, \mathbf{p}) \sum_{t=1}^{\infty} \frac{1}{t}}. \end{aligned} \quad (33)$$

When $t \rightarrow \infty$, we can obtain

$$\sum_{t=1}^{\infty} \frac{1}{t} \rightarrow \infty. \quad (34)$$

Therefore, we finally have

$$\begin{aligned} 0 &\leq \lim_{t \rightarrow \infty} \text{Prob}(\omega^{(t)} \neq \omega^*) \\ &\leq \lim_{t \rightarrow \infty} e^{-\sum_{i=1}^{N_R} f_{i,0}(\omega^*, \mathbf{p}) \sum_{t=1}^{\infty} \frac{1}{t}} = 0, \end{aligned} \quad (35)$$

which implies that $\lim_{t \rightarrow \infty} \text{Prob}(\omega^{(t)} = \omega^*) = 1$. This completes the proof.

REFERENCES

- [1] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecom.*, vol. 10, pp. 585-595, Nov. 1999.
- [2] A. F. Molisch, M. Z. Win, Y. S. Choi, and J. H. Winters, "Capacity of MIMO systems with antenna selection," *IEEE Trans. Wireless Commun.*, vol. 4, pp.1759-1772, July 2005.
- [3] A. F. Molisch, "MIMO systems with antenna selection—an overview," *IEEE Commun. Mag.*, vol. 42, pp. 68-73, Oct. 2004.
- [4] A. G. Dhananjay, A. Gore, and A. J. Paulraj, "Receive antenna selection for MIMO spatial multiplexing: theory and algorithms," *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2796-2807, Nov. 2003.
- [5] R. W. Heath Jr., S. Sandhu, and A. Paulraj, "Antenna selection for spatial multiplexing systems with linear receivers," *IEEE Commun. Lett.*, vol. 5, no. 4, pp. 142-144, Apr. 2001.
- [6] D. A. Gore, R. W. Heath, and A. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Commun. Lett.*, vol.6, no. 11, pp. 491-493, Nov. 2002.
- [7] M. Z. Win and J. H. Winters, "Analysis of hybrid selection/maximal ratio combining in Rayleigh fading," *IEEE Trans. Commun.*, vol. 47, pp. 1773-1776, Dec. 1999.
- [8] A. Dua, K. Medepalli, and A. Paulraj, "Receive antenna selection in MIMO systems using convex optimization," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 2353- 2357, Sept. 2006.
- [9] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of spatially correlated MIMO Rayleigh fading channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2363-2371, Oct. 2003.
- [10] D. S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, pp. 502512, Mar. 2000.
- [11] L. Dai, S. Sfar, and K. B. Letaief, "Optimal antenna selection based on capacity maximization for MIMO systems in correlated channels," *IEEE Trans. Commun.*, vol. 54, pp. 563-573, Mar. 2006.
- [12] H. Zhang and H. Dai, "Fast MIMO transmit antenna selection algorithm: a geometric approach," *IEEE Commun. Lett.*, vol. 10, pp. 754-756, Nov. 2006.
- [13] S. Kullback and R. A. Leibler, "On information and sufficiency," *Annals of Mathematical Statistics*, vol. 22, pp. 79-86, 1951.
- [14] R. Y. Rubinstein, "Optimization of computer simulation models with rare events," *Eur. J. Operations Research*, pp. 89-112, 1997.
- [15] R. Y. Rubinstein and D. P. Kroese, *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation, and Machine Learning*. Springer Verlag, 2004.
- [16] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. The John Hopkins Univ. Press, 1996.
- [17] S. A. Jafar and A. Goldsmith, "Multiple-antenna capacity in correlated Rayleigh fading with channel covariance information," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 990- 997, May 2005.
- [18] H. Zhang, A. F. Molisch, D. Gu, D. Wang, and J. Zhang, "Antenna selection in high-throughput wireless LAN," *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (ISBMSB)*, June 2006.