Adaptive selection of antennas for optimum transmission in spatial modulation

Xiping Wu, Marco Di Renzo, Harald Haas

Adaptive Selection of Antennas for Optimum Transmission in Spatial Modulation

Xiping Wu, *Member, IEEE*, Marco Di Renzo, *Senior Member, IEEE*, and Harald Haas, *Member, IEEE*

Abstract—In this paper, we propose an optimum transmit structure for spatial modulation (SM), a unique single-stream multiple-input multiple-output (MIMO) transmission technique. As a three-dimensional modulation scheme, SM enables a trade-off between the size of the spatial constellation diagram and the size of the signal constellation diagram. Based on this fact, the novel method, named transmission optimized spatial modulation (TOSM), selects the best transmit structure that minimizes the average bit error probability (ABEP). Unlike the traditional antenna selection methods, the proposed method relies on statistical channel state information (CSI) instead of instant CSI, and feedback is only needed for the optimal number of transmit antennas. The overhead for this, however, is negligible. In addition, TOSM has low computational complexity as the optimization problem is solved through a simple closed-form objective function with a single variable. Simulation results show that TOSM significantly improves the performance of SM at various channel correlations. Assuming Rayleigh fading channels, TOSM outperforms the original SM by up to 9 dB. Moreover, we propose a single radio-frequency (RF) chain base station (BS) based on TOSM, which achieves low hardware complexity and high energy efficiency. In comparison with multi-stream MIMO schemes, TOSM offers an energy saving of at least 56% in the continuous transmission mode, and 62% in the discontinuous transmission mode.

Index Terms—Spatial modulation (SM), channel correlation, transmit antenna selection, MIMO.

I. INTRODUCTION

THE need to curtail the carbon footprint and the operation cost of wireless networks requires an overall energy reduction of base stations (BSs) in the region of two to three orders of magnitude [3]. At the same time, a significant increase in spectrum efficiency from currently about 1.5 bit/s/Hz to at least 10 bit/s/Hz is required to cope with the exponentially increasing traffic loads [4]. This challenges the design

of multiple-input multiple-output (MIMO) systems associated with the BS. A typical long-term evolution (LTE) BS consists of radio-frequency (RF) chains, baseband interfaces, direct current to direct current (DC-DC) converters, cooling fans, etc. Each RF chain contains a power amplifier (PA), and PAs contribute around 65% of the entire energy consumption [5]. The efficiency of state-of-the-art (SOTA) PAs is about 30% only [6], i.e. more than two thirds of the energy is consumed in quiescent power. This drives research on minimizing the overall BS energy consumption instead of the energy required for the RF output stage only. As a result, power optimization of PAs has been studied. In [7], cell discontinuous transmission (DTX) was proposed to enable the BSs to fall into a sleep mode when there is no data to convey, so that the overall energy consumption can be reduced. Based on that concept, another optimization method using on/off PAs was reported in [8], and a similar work was conducted for MIMO orthogonal frequency division multiple access (OFDMA) systems [9]. However, those studies have the following limitations: i) they focus on the operation of RF chains, while modulation schemes are not considered; ii) the optimization is implemented within each individual RF chain; and iii) the benefit is inversely proportional to the traffic load. When the BS has to be operated in the active mode continuously, the above methods would fail to achieve any energy-saving gain. Therefore it is necessary to study energy reduction on a more comprehensive level, including not only hardware operations, but also modulation schemes.

While multi-stream MIMO schemes, such as vertical Bell Labs layered space-time (V-BLAST) and space-time block coding (STBC), offer high spectrum efficiency, unfortunately, they need multiple RF chains that heavily compromise the energy efficiency. Meanwhile, spatial modulation (SM) is a unique single-stream MIMO technique [10]–[12], where the bit stream is divided into blocks and each block is split into two parts: i) the first part activates one antenna from the antenna array while the remaining antennas do not emit a signal; ii) the bits in the second part are modulated by a signal constellation diagram, and sent out through the activated antenna. The use of a single active antenna makes SM a truly energy-efficient MIMO transmission technique, because only one RF chain is required, regardless of the number of transmit antennas used. At the same time, SM ensures spatial multiplexing gains as information is encoded in the antenna index. However, like all other MIMO schemes, SM suffers performance degradation caused by channel correlations [13]. Trying to improve the performance of SM against channel variations, an adaptive method was proposed in [14], where one candidate is selected from several optional SM structures. Although the performance

of SM can be improved to some extent, this method has the following weaknesses: i) it requires instant channel state information (CSI), and therefore it is not suitable for fast fading channels; ii) the relation between the adaptive selection and the channel correlation has not been exploited; and iii) despite using a simplified modulation order selection criterion, it still requires significant processing power.

In this context, we propose a novel adaptive antenna selection method for optimum transmission in SM. As a three-dimensional modulation scheme, SM enables a trade-off between the size of the spatial constellation diagram and the size of the signal constellation diagram, while achieving the same spectrum efficiency. Based on this unique characteristic, transmission optimized spatial modulation (TOSM) aims to select the best combination of these two constellation sizes, which minimizes the average bit error probability (ABEP). To avoid the prohibitive complexity caused by exhaustive search, a two-stage optimization strategy is proposed. The first step is to determine the optimal number of transmit antennas, and this is performed at the receiver. In the second step, the required number of antennas are selected at the transmitter. In addition to low computational complexity, TOSM needs very limited feedback because of two aspects: i) since it is based on statistical CSI, the frequency of updating is relatively low; and ii) feedback is required only to inform the transmitter of the number of selected antennas, instead of the index of each selected antenna. Therefore, the feedback overhead is negligibly low and not considered in this paper. In addition, assuming the SOTA 2010 power model [15], the overall BS energy consumption is studied for TOSM. The DTX technique is combined with TOSM to further improve the energy efficiency. Compared with our preceding studies in [1] and [2], the contributions in this paper are four-fold: i) a two-stage optimization method is proposed to balance the spatial modulation order with the signal modulation order in SM systems; ii) a complete derivation of a simplified ABEP bound for SM over generalized fading channels is presented; iii) a direct antenna selection method based on circle packing is proposed; and iv) the energy efficiency of TOSM in terms of the BS energy consumption is evaluated for both the continuous mode and the DTX mode.

The remainder of this paper is organized as follows. Section II describes the system model, including the SM transceiver, the channel model, and the BS power model. In Section III, a two-stage antenna selection method is proposed for optimum transmission in SM. Section IV studies the proposed method in the case of Rayleigh fading channels. Simulation results are presented in Section V to validate the optimization accuracy and the bit error rate (BER) performance of the proposed method. Finally, the paper is concluded in Section VI.

II. SYSTEM MODEL

A. MIMO System and Signal Model

A $N_t \times N_r$ SM-MIMO system is considered, where N_t and N_r are the number of transmit antennas and the number of receive antennas, respectively. Unlike the original SM, only a subset of the transmit antennas is used. The size of the spatial constellation diagram, i.e. the number of utilized transmit antennas, is denoted by N, while the size of the signal constellation diagram is denoted by M. The bit stream is divided into blocks with the length of η_s bits, where $\eta_s = \log_2(N) + \log_2(M)$ is the number of bits per symbol. Each block is then split into two units of $\log_2(N)$ and $\log_2(M)$ bits. The first part activates a single transmit antenna from the spatial constellation diagram, and the currently active antenna is denoted by t_{act}. The second part chooses the corresponding symbol χ_l $(1 \le l \le M)$ from a specific signal constellation diagram, such as phase shift keying (PSK) or quadrature amplitude modulation (QAM), and sends it out through the activated antenna. The transmitted signal of SM is represented by a vector $\mathbf{x} = [0, \ldots, x_{t_{\text{act}}}, \ldots, 0]^T$ of N elements, where the t_{act}-th element is χ_l and all other elements are zero.

The received signal is given by $\mathbf{y} = \mathbf{HX} + \mathbf{w}$, where $\mathbf{H}$ stands for the channel matrix and it is described in Section II-B; the vector $\mathbf{w} = [w_1, w_2, \ldots, w_{N_r}]^T$ and w_r, the noise at the r-th receive antenna, is a sample of complex additive white Gaussian noise with distribution $\mathcal{CN}(0, N_0)$. Across receive antennas, the noise components are statistically independent. The signal-to-noise ratio (SNR) is defined as $\gamma = E_m L / N_0$, where E_m is the average energy per symbol transmission and L denotes the path loss without shadowing. In addition, the required RF output energy per bit is denoted by $E_b = E_m / \eta_s$. The transmitted information bits are decoded by the joint maximum likelihood (ML) detection in [12] as follows:

$$[\hat{t}_{\text{act}}, \hat{l}] = \arg\min_{t,l} \|\mathbf{h}_t \chi_l\|^2 - 2\Re\{\mathbf{y}^H \mathbf{h}_t \chi_l\}, \qquad (1)$$

where $\mathbf{h}_t$ is the t-th column of $\mathbf{H}$; $\Re\{\cdot\}$ is the real part of a complex number; $\hat{t}_{\text{act}}$ and $\hat{l}$ are the detection results of t_{act} and l, respectively.

B. Channel Model

1) Channel Distribution: The fading coefficient of the link from the t-th transmit antenna to the r-th receive antenna is denoted by $h_{t,r} = \beta_{t,r} \exp(j\varphi_{t,r})$, where $\beta_{t,r}$ and $\varphi_{t,r}$ are the amplitude and the phase, respectively. The channel fading distribution as well as the CSI is assumed to be known at the receiver. Nakagami-m fading is considered in this paper, i.e. $\beta_{t,r} \sim \text{Nakagami}(m_{t,r}, \Omega_{t,r})$, where $m_{t,r}$ is the shape parameter (when $m_{t,r} = 1$, the channel is Rayleigh fading) and $\Omega_{t,r}$ is the spread controlling parameter. The phase $\varphi_{t,r}$ is uniformly distributed between $(-\pi, \pi]$.

2) Channel Correlation: Since we focus on selecting the transmit antennas, the receive antennas are assumed to be independent without loss of the generality. The correlation coefficient between the amplitudes of the two propagation paths from the transmit antennas t_i and t_j to the r-th receive antenna is denoted by $\rho_{t_i, t_j, r}$. The exponential correlation matrix model in [16] is considered, which is based on the fact that the channel correlation decreases with increasing the distance between antennas. As shown in Fig. 1, the transmit antennas are located in a normalized square area, i.e. the distance between t_1 and t_A is unity. The correlation coefficient between t_1 and t_A with respect to the r-th receive antenna is denoted by $\rho_{s(r)}$.
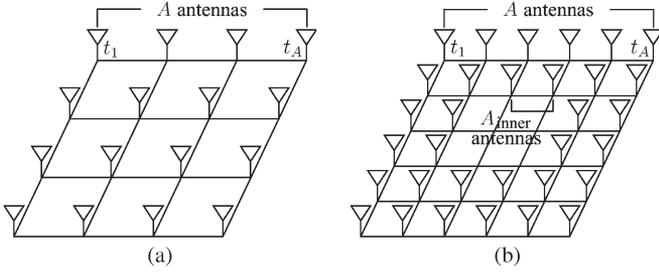
Fig. 1.   Examples of the transmit antenna array. (a) $N_t = 16$. (b) $N_t = 32$.

The number of antennas on each side of the antenna array is formulated as follows:

$$A = \begin{cases} \sqrt{N_t} & \text{if } \log_2 N_t \text{ is even} \\ 3 \times \sqrt{\frac{N_t}{8}} & \text{if } \log_2 N_t \text{ is odd} \end{cases}. \qquad (2)$$

When $\log_2 N_t$ is even, the antennas form a square array with the dimension of $A \times A$. If $\log_2 N_t$ is odd, the antennas are placed in the shape shown in Fig. 1(b), where $A_{\text{inner}} = \sqrt{\frac{N_t}{8}}$. The absolute distance between t_i and t_j is denoted by d_{t_i,t_j}, and the correlation between those two antennas is given by [16, Eq. (10)]:

$$\rho_{t_i,t_j,r} = \rho_{s(r)}^{d_{t_i,t_j}}, \quad 0 \le \rho_{s(r)} \le 1. \qquad (3)$$

The average degree of the channel correlations, denoted by ρ_{av}, is calculated by:

$$\rho_{\text{av}} = \frac{1}{N_r} \sum_{r=1}^{N_r} \left(\frac{1}{N_t(N_t-1)} \sum_{t_i=1}^{N_t} \sum_{t_j \ne t_i = 1}^{N_t} \rho_{t_i,t_j,r} \right). \qquad (4)$$

C. Base Station Power Model

In [17], a linear relationship between the RF output power and the overall consumed power of a multi-sector BS was established. The overall BS power consumption is divided into two parts: the load-dependent portion and the constant portion. The former is dependent on the RF output power, while the latter is invariant. In addition, when no data is transmitted, a sleep mode is enabled to reduce the consumption by switching off unneeded components. In this section, a practical BS power model is introduced for the purpose of evaluating the energy efficiency of the proposed method.

1) Power Model: In [15], based on the above literature, a power model named SOTA 2010 was proposed for a one-sector, single-RF-chain BS. Table I specifies the parameters: P_{max} is the maximum RF output power; P_0 and P_s denote the constant power consumption for the active mode and the sleep mode, respectively; ζ stands for the slope that quantifies the load dependence. The instantaneous BS power consumption, denoted by P_{in}, is formulated as a function of the RF output power P_{out} [15, Eq. (1)]:

$$P_{\text{in}} = \begin{cases} P_0 + \zeta P_{\text{out}} & \text{if } 0 < P_{\text{out}} \le P_{\text{max}} \\ P_s & \text{if } P_{\text{out}} = 0 \end{cases}. \qquad (5)$$

Also, the ratio of the time consumed in the active mode and the total period is referred to as the activation ratio μ. The

TABLE I
BS POWER MODEL PARAMETERS [15]

Power model	P_{max} (W)	P_0 (W)	P_s (W)	ζ
SOTA 2010	40	119	63	2.4

average power consumption of a single-RF-chain BS is then computed by:

$$P_{\text{BS}} = \mu(P_0 + \zeta P_{\text{out}}) + (1 - \mu)P_s. \qquad (6)$$

Unlike SM using a single RF chain, multi-stream MIMO schemes require multiple RF chains to activate the transmit antennas simultaneously. For N_{act} *activated* antennas, the RF output power of each antenna is $1/N_{\text{act}}$ of P_{out}. As a result, the overall power consumption of a BS with multiple RF chains is calculated by:

$$P_{\text{BS}} = N_{\text{act}} \left[\mu \left(P_0 + \zeta \frac{P_{\text{out}}}{N_{\text{act}}} \right) + (1 - \mu)P_s \right]. \qquad (7)$$

2) Continuous Mode and DTX Mode: Two modes are considered to operate the RF chains: the continuous mode and the DTX mode [7]. In the continuous mode, the RF chains are always delivering output power of the same level. As a result, P_{out} is equal to the average RF output power $\bar{P}_{\text{out}}$, and $\mu = 1$. Substituting those conditions into (7), the overall BS power consumption in the continuous mode is obtained by:

$$P_{\text{BS·cont.}} = N_{\text{act}} P_0 + \zeta \bar{P}_{\text{out}}. \qquad (8)$$

The data rate of the continuous mode is equal to the average data rate, which is denoted by $\bar{R}_b = \bar{P}_{\text{out}}/E_b$. Conversely, the DTX mode conveys data with full load, and the instantaneous data rate is $R_{b\max} = P_{\text{max}}/E_b$ that is higher than $\bar{R}_b$. Then the system is enabled into sleep mode during the saved time to maintain the same average data rate. The parameter μ of the DTX mode is computed by:

$$\mu = \frac{\bar{R}_b}{R_{b\max}} = \frac{\bar{P}_{\text{out}}}{P_{\text{max}}}. \qquad (9)$$

Substituting (9) and $P_{\text{out}} = P_{\text{max}}$ into (7), the overall BS power consumption in the DTX mode is expressed as:

$$P_{\text{BS·DTX}} = N_{\text{act}} P_s + \left(\zeta + \frac{N_{\text{act}}(P_0 - P_s)}{P_{\text{max}}} \right) \bar{P}_{\text{out}}. \qquad (10)$$

III. TOSM OVER GENERALIZED FADING CHANNELS

As mentioned, SM enables to balance the size of the signal constellation diagram with the size of the spatial constellation diagram. To achieve a certain spectrum efficiency η_s, there are $\eta_s + 1$ possible combinations of (M, N), based on the requirement of a power of two to provide a full usage in the constellation diagrams. To supply a complete selection range of N, N_t is set to be equal to 2^{η_s}. In this section, an optimization algorithm is proposed to select the best (M, N) as well as the specific antennas. The context is arranged in four portions: i) an ABEP upper bound is introduced for SM over generalized fading channels; ii) the ABEP upper bound is simplified

to suit the minimization on the subject of either M or N; iii) the optimum (M, N) is obtained by solving the minimization problem; and iv) the required number of transmit antennas are selected from the antenna array.

A. ABEP Upper Bound

The ABEP upper bound of SM with the joint ML detector is given by [18, Eq. (6)]:

$$\text{ABEP} \leq \text{ABEP}_{\text{spatial}} + \text{ABEP}_{\text{signal}} + \text{ABEP}_{\text{joint}}, \quad (11)$$

and:

$$
\begin{cases}
\text{ABEP}_{\text{spatial}} = \frac{\log_2(N)}{\eta_s M} \sum_{l=1}^{M} \text{ABEP}_{\text{SSK}}(l) & (12a) \\[2mm]
\text{ABEP}_{\text{signal}} = \frac{\log_2(M)}{\eta_s N} \sum_{t=1}^{N} \text{ABEP}_{\text{MOD}}(t) & (12b) \\[2mm]
\text{ABEP}_{\text{joint}} = \frac{1}{\eta_s M N} \sum_{t_i=1}^{N} \sum_{l_i=1}^{M} \sum_{t_j=1}^{N} \sum_{l_j=1}^{M} \text{ABEP}_{\text{MIX}} & (12c)
\end{cases}
$$

with the terms in summations expressed in (13), shown at bottom of the page, where: i) $P_s(l, t)$ is the average symbol error rate (SER) of the l-th signal symbol χ_l emitted from the t-th transmit antenna; ii) $N_H(\cdot)$ denotes the Hamming distance between two symbols; iii) the average pair-wise error probability (APEP) is defined as the probability of pair-wise error event which is calculated by (14), shown at bottom of the page. The denotation $E[\cdot]$ is the expectation operator, and $Q(\cdot)$ is the Q-function.

B. Simplification

Due to the relaxation of linearity requirements, unlike QAM, PSK can work in PA saturation [19]. This makes PSK a more energy efficient modulation scheme. Moreover, results in [18] have shown that, for SM, PSK is not worse than QAM in many cases and in some cases it is even better. Therefore,

PSK modulation is applied in the rest of this paper to keep a moderate level of analytical complexity. Every symbol has an equal weight and the average of different signal symbols in (12) can be neglected. In addition, $P_s(t, l)$ reduces to $P_s(t)$.

1) ABEP of Spatial Part: At first, we focus on $\text{ABEP}_{\text{spatial}}$ in the union bound. Assuming a high SNR, the APEP of space shift keying (SSK) modulation over correlated Nakagami-m fading channels (with the assumption of $m_{t,r} = m_r$ for $t = 1, 2, \ldots, N_t$) can be obtained based on the moment generate function (MGF) as follows [20, Eq. (15)]:

$$\text{APEP}(t_j \to t_i) = \gamma^{-N_r} \frac{2^{3N_r-1}\Gamma(N_r + 0.5)}{\sqrt{\pi}\Gamma(N_r + 1)} \prod_{r=1}^{N_r} f(r), \quad (15)$$

where $f(r)$ is computed in (16), shown at bottom of the page. We define:

$$C_{t_i, t_j} = \frac{1}{2} \times \frac{2^{3N_r-1}\Gamma(N_r + 0.5)}{\sqrt{\pi}\Gamma(N_r + 1)} \prod_{r=1}^{N_r} f(r). \quad (17)$$

Note that C_{t_i, t_j} is constant when m_r, $\Omega_{t,r}$ and $\rho_{t_i, t_j, r}$ are given. Consequently, (15) can be rewritten as:

$$\text{APEP}(t_j \to t_i) = 2C_{t_i, t_j}\gamma^{-N_r}. \quad (18)$$

The $\log_2(N)$ bits provided to the spatial constellation is assumed to be encoded by Gray codes. When a certain antenna is activated, there are $N/2$ other antennas that may cause an error for any particular bit. Therefore the total Hamming distance from one symbol to the other symbols is calculated by:

$$\sum_{t_j \neq t_i=1}^{N} N_H(t_j \to t_i) = \frac{N}{2}\log_2(N). \quad (19)$$

Combining (19) and (13a), we have:

$$\text{ABEP}_{\text{SSK}}(l) = \frac{1}{2(N-1)} \sum_{t_i=1}^{N} \sum_{t_j \neq t_i=1}^{N} \text{APEP}(t_j \to t_i). \quad (20)$$

$$
\begin{cases}
\text{ABEP}_{\text{SSK}}(l) = \frac{1}{N\log_2(N)} \sum_{t_i=1}^{N} \sum_{t_j \neq t_i=1}^{N} N_H(t_j \to t_i)\text{APEP}\left((t_j, \chi_l) \to (t_i, \chi_l)\right) & (13a) \\[3mm]
\text{ABEP}_{\text{MOD}}(t) = \frac{1}{M\log_2(M)} \sum_{l=1}^{M} P_s(l, t) & (13b) \\[3mm]
\text{ABEP}_{\text{MIX}} = N_H\left((t_j, \chi_{l_j}) \to (t_i, \chi_{l_i})\right)\text{APEP}\left((t_j, \chi_{l_j}) \to (t_i, \chi_{l_i})\right) & (13c)
\end{cases}
$$

$$\text{APEP}\left((t_j, \chi_{l_j}) \to (t_i, \chi_{l_i})\right) = E\left[Q\left(\sqrt{\gamma\left|\alpha_{t_j}\chi_{l_j} - \alpha_{t_i}\chi_{l_i}\right|^2/4}\right)\right] \quad (14)$$

$$f(r) = \frac{m_r\left(\Omega_{t_i,r} + \Omega_{t_j,r}\right)^{1-2m_r}\left(1 - \rho_{t_i,t_j,r}\right)^{m_r-1}}{\left(\Omega_{t_i,r}\Omega_{t_j,r}\right)^{1-m_r}\Gamma(m_r)} \sum_{k_r=0}^{+\infty} \frac{\rho_{t_i,t_j,r}^{k_r}\Gamma(2m_r + 2k_r - 1)}{(k_r!)\Gamma(m_r + k_r)}\left(\frac{\sqrt{\Omega_{t_i,r}\Omega_{t_j,r}}}{\Omega_{t_i,r} + \Omega_{t_j,r}}\right)^{2k_r} \quad (16)$$

Substituting (18) and (20) into (12a), this gives $\text{ABEP}_{\text{spatial}}$ as a function of N:

$$\text{ABEP}_{\text{spatial}} = \frac{\gamma^{-N_r} \log_2(N)}{\eta_s(N-1)} \sum_{t_i=1}^{N} \sum_{t_j \neq t_i = 1}^{N} C_{t_i, t_j}. \quad (21)$$

2) ABEP of Signal Part: The average SER of the PSK modulation over Nakagami-m fading channels is given by [21, Eq. (9.16)]:

$$P_s(t) = \frac{1}{\pi} \int_0^{\frac{M-1}{M}\pi} \prod_{r=1}^{N_r} \left(1 + \frac{\overline{\gamma}_{t,r} \sin^2\left(\frac{\pi}{M}\right)}{2m_r \sin^2(\theta)} \right)^{-m_r} d\theta, \quad (22)$$

where $\overline{\gamma}_{t,r} = \Omega_{t,r}\gamma$ is the average SNR of the symbol sent from the t-th transmit antenna at the input of the r-th receive antenna. The assumption of a high SNR results in $\frac{\gamma \sin^2(\pi/M)}{2m_r \sin^2(\theta)} \gg 1$. Hence (22) can be rewritten as:

$$P_s(t) = \frac{\int_0^{\frac{M-1}{M}\pi} (\sin\theta)^{\sum_{r=1}^{N_r} 2m_r} d\theta}{\pi} \prod_{r=1}^{N_r} \left(\frac{2m_r/\Omega_{t,r}}{\gamma \sin^2\left(\frac{\pi}{M}\right)} \right)^{m_r}. \quad (23)$$

When $M \geq 4$, we have: i) $\int_0^{\frac{M-1}{M}\pi} (\sin\theta)^{\sum_{r=1}^{N_r} 2m_r} d\theta \approx \int_0^{\pi} (\sin\theta)^{\sum_{r=1}^{N_r} 2m_r} d\theta$, which is independent of the signal constellation size M; ii) $\sin(\pi/M) \approx \pi/M$. The average shape parameter of the fading distributions across all receive antennas is denoted by $\overline{m}_r = 1/N_r \sum_{r=1}^{N_r} m_r$. Then the average SER of PSK can be formulated as follows:

$$P_s(t) = \frac{M^{2\overline{m}_r N_r} \int_0^{\pi} (\sin\theta)^{2\overline{m}_r N_r} d\theta}{\gamma^{\overline{m}_r N_r} \pi^{2\overline{m}_r N_r + 1}} \prod_{r=1}^{N_r} \left(\frac{2m_r}{\Omega_{t,r}} \right)^{m_r}. \quad (24)$$

Substituting (24) into (12b), a simplified $\text{ABEP}_{\text{signal}}$ is obtained by:

$$\text{ABEP}_{\text{signal}} = \frac{\overline{B}_N}{\eta_s} M^{2\overline{m}_r N_r} \gamma^{-\overline{m}_r N_r}, \quad (25)$$

with

$$\overline{B}_N = \frac{\int_0^{\pi} (\sin\theta)^{2\overline{m}_r N_r} d\theta}{\pi^{2\overline{m}_r N_r + 1} N} \sum_{t=1}^{N} \prod_{r=1}^{N_r} \left(\frac{2m_r}{\Omega_{t,r}} \right)^{m_r}. \quad (26)$$

Note that similar to $\text{ABEP}_{\text{spatial}}$ in (21), $\text{ABEP}_{\text{signal}}$ is also a function of N after replacing M by $2^{\eta_s}/N$.

3) ABEP of Joint Part: The symbols of the PSK modulation are expressed as $\chi_l = \exp(j\varphi_l)$, where $\varphi_l = 2\pi(l-1)/M$. Thus, (14) can be rewritten as (27), shown at bottom of the page, where φ_t is uniformly distributed between $(-\pi, \pi]$ and φ_l is a constant. The new random variable $\varphi_t + \varphi_l$ complies with a uniform distribution over the same region as φ_t. As a result, $\text{APEP}((t_j, \chi_{l_j}) \to (t_i, \chi_{l_i}))$ is reduced to $\text{APEP}(t_j \to t_i)$. Similar to (19), the Hamming distance in (13c) can be expressed as (28), shown at bottom of the page. Substituting (18) and (28) into (12c), $\text{ABEP}_{\text{joint}}$ is simplified in (29), shown at bottom of the page.

4) Simplified ABEP: Combining (21), (25) and (29), a simplified ABEP of SM is derived as follows:

$$\text{ABEP} = \frac{M^{2\overline{m}_r N_r}}{\eta_s \gamma^{\overline{m}_r N_r}} \overline{B}_N + \frac{\eta_s 2^{\eta_s} - M \log_2(M)}{\eta_s \gamma^{N_r}} \overline{C}_N, \quad (30)$$

with

$$\overline{C}_N = \frac{1}{N(N-1)} \sum_{t_i=1}^{N} \sum_{t_j \neq t_i = 1}^{N} C_{t_i, t_j}. \quad (31)$$

The term $\overline{C}_N$ reflects the average degree that the selected antenna array is affected by the fading distribution, the channel correlation, and the method of transmit antenna selection. Since the production of M and N is fixed to be 2^{η_s}, the optimization problem in (30) is actually subjected to one variable, M or N. The aim of TOSM is to determine the optimal transmit structure, including the number of antennas and their locations. However, jointly optimizing both of them requires an exhaustive search of prohibitive computational complexity. Instead, we propose a two-stage approach which separates the optimization problem in two steps: i) find the best combination

$$\text{APEP}\left((t_j, \chi_{l_j}) \to (t_i, \chi_{l_i})\right) = E\left[Q\left(\sqrt{\gamma \left| \beta_{t_j} \exp\left(j\left(\varphi_{t_j} + \varphi_{l_j}\right)\right) - \beta_{t_i} \exp\left(j\left(\varphi_{t_i} + \varphi_{l_i}\right)\right) \right|^2 / 4} \right) \right] \quad (27)$$

$$\sum_{t_j \neq t_i = 1}^{N} \sum_{l_j \neq l_i = 1}^{M} N_H\left((t_j, \chi_{l_j}) \to (t_i, \chi_{l_i})\right) = \frac{(M-1)N}{2} \log_2(N) + \frac{(N-1)M}{2} \log_2(M) \quad (28)$$

$$\text{ABEP}_{\text{joint}} = \frac{\eta_s 2^{\eta_s} - M \log_2 M - N \log_2(N)}{\eta_s} \left(\frac{1}{N(N-1)} \sum_{t_i=1}^{N} \sum_{t_j \neq t_i = 1}^{N} C_{t_i, t_j} \right) \gamma^{-N_r} \quad (29)$$

of (M, N) that minimizes the ABEP of SM; and ii) select the specific antennas from the antenna array.

C. Optimal Selection of the Number of Transmit Antennas

In this step, the minimization of the simplified ABEP with respect to N (or M) is implemented for a given scenario, which is comprised of the spectrum efficiency, the number of receive antennas, the SNR, the fading distribution, and the correlation coefficient. The term $1/\eta_s$ in (30) is a positive constant, hence it can be removed without affecting the optimization result. In addition, the difference between antennas is not considered in this step. As a result, $\overline{B}_N$ and $\overline{C}_N$ are replaced by $\overline{B}_{N_t}$ and $\overline{C}_{N_t}$ to avoid the dependence on the antenna dissimilarity. The optimization problem can thus be formulated as (32), shown at bottom of the page.

Note that both $\overline{C}_{N_t}\gamma^{-N_r}$ and $\overline{B}_{N_t}\gamma^{-\overline{m}_r N_r}$ are constant for a certain scenario. In other words, they are independent of the variable M. The optimization problem in (32) is a non-linear programming problem and can be solved numerically [22]. The optimization result of $\log_2(M)$ could be, and most likely is, a non-integer. However, without considering special encoding methods such as fractional bit encoding [23], both M and N must be a power of two to supply a full usage in the constellation diagram. This can readily be achieved by comparing the ABEP values of the two nearest integers around the optimal M. Afterwards, the best combination of (M, N) is obtained and denoted by $(M_{\text{opt}}, N_{\text{opt}})$.

D. Direct Antenna Selection

The second step is to select a subarray of N_{opt} antennas from the size-N_t antenna array. The chosen subset should achieve the minimum ABEP of all subarrays with the same size. Since $\overline{B}_N$ in (30) is irrelevant to the channel correlations, the problem is equivalent to finding the subarray with a minimum $\overline{C}_N$. Like the traditional transmit antenna selection (TAS) methods, this issue can be solved by an exhaustive search. However, this results in an unaffordable complexity for a large η_s. Taking $\eta_s = 6$ and $N_{\text{opt}} = 16$ as an example, the full search space is about 5×10^{14}, which is prohibitive for practical implementations. Here we propose a novel TAS method based on circle packing, which can directly determine the selection.

As the correlation coefficient ρ_{t_i, t_j} is inversely proportional to the distance d_{t_i, t_j}, a rational solution is to maximize the minimum geometric distance between any pair of the chosen antennas. This is equivalent to the circle packing problem in mathematics which can be worked out numerically [24]. Fig. 2 shows the circle packing solutions for various numbers of antennas, where the antennas are located at the circle centers. In the original problem, each circle must fit inside the square boundary. The problem at hand is slightly different where the circle centers are restricted to be inside the boundary, and in
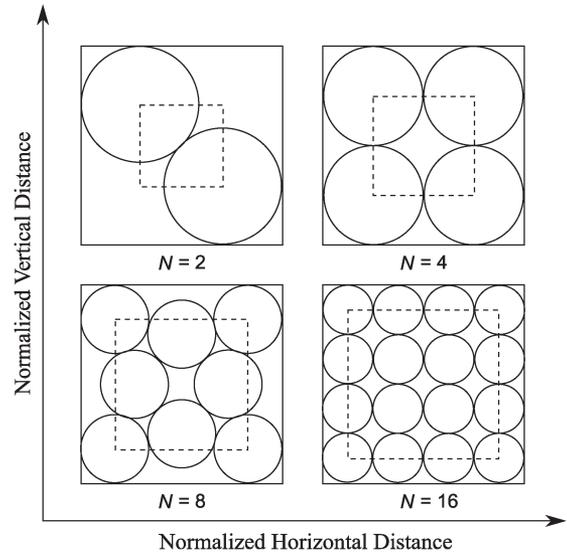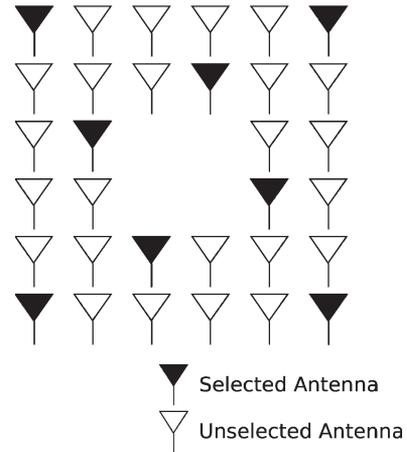

Fig. 2. Examples of circle packing problems.

Fig. 3. RCP for selecting 8 out of 32 antennas.

Fig. 2 this is shown by dashed lines. It is worth noting that this solution requires fully flexible positions. Thus, we refer to it as ideal circle packing (ICP). However, the antenna positions are fixed in practice, and the subarray cannot be perfectly allocated by ICP. Instead, a realistic circle packing (RCP) is developed by selecting those antennas closest to the ideal positions. In Fig. 3, an RCP solution is demonstrated for the case of $N_t = 32$ and $N = 8$. As can be observed, the selection presents a similarity to the solution for $N = 8$ in Fig. 2. With an increase of N_t, the RCP solution becomes closer to ICP as the antenna array supplies a larger flexibility in positions.

IV. TOSM Over c.i.d. Rayleigh Fading Channels

In this section, we study TOSM in the special case of channel fading, Rayleigh fading. We present that TOSM is independent

$$M_{\text{opt}} = \arg\min_{M} \quad \frac{\overline{B}_{N_t}}{\gamma^{\overline{m}_r N_r}} M^{2\overline{m}_r N_r} + \frac{\overline{C}_{N_t}}{\gamma^{N_r}} \left(2^{\eta_s}\eta_s - M\log_2(M)\right), \quad \text{subject to :} \quad 1 \leq M \leq 2^{\eta_s} \qquad (32)$$

of SNR in this particular scenario, and a look-up table can be built to quickly determine the best choice of (M, N). In addition, given a target bit error rate (BER) and transmission rate, a closed-form expression of the BS energy consumption is derived for TOSM. This allows us to evaluate the proposed scheme analytically.

A. ABEP of TOSM4

For correlated and identically distributed (c.i.d.) Rayleigh fading channels, we have $m_r = 1$ and $\Omega_{t,r} = \Omega$ for all t and r. As a result, $\overline{B}_N$ in (26) and $\overline{C}_N$ in (31) reduce to:

$$B = \left(\frac{2}{\Omega}\right)^{N_r} \frac{1}{\pi^{2N_r+1}} \int_0^\pi (\sin\theta)^{2N_r} d\theta, \quad (33)$$

and

$$C = \frac{4^{N_r-1}\Gamma(N_r+0.5)}{\Omega^{N_r}\sqrt{\pi}\Gamma(N_r+1)} \left(\sum_{k=0}^{+\infty} \frac{\Gamma(2k+1)\rho_{av}^k}{4^k(k!)\Gamma(k+1)}\right)^{N_r}. \quad (34)$$

Correspondingly, the ABEP expression in (30) reduces to:

$$\text{ABEP} = \frac{B(M)^{2N_r} + C\left(2^{\eta_s}\eta_s - M\log_2(M)\right)}{\eta_s\gamma^{N_r}}. \quad (35)$$

The term $\eta_s\gamma^{N_r}$ is a positive constant, and therefore can be removed without influencing the optimization result. After further removing the constant item $C2^{\eta_s}\eta_s$, the optimization problem is reduced to:

$$M_{\text{opt}} = \arg\min_M \quad B(M)^{2N_r} - CM\log_2(M)$$
$$\text{subject to}: \quad 1 \le M \le 2^{\eta_s}. \quad (36)$$

It is not difficult to find that the optimal solution of TOSM over c.i.d. Rayleigh fading is independent of SNR. Also, M_{opt} is independent of η_s when η_s is large enough. In other words, for a certain N_r and a large enough η_s, the optimal solution is only dependent on ρ_{av}. As shown in Table II, this allows us to build a look-up table to quickly decide the best SM deployment against different channel correlation degrees.

B. Base Station Energy Consumption based on TOSM

Substituting $\gamma = E_m L/N_0$ into (35), the required E_m using TOSM is computed by:

$$E_m = \frac{N_0}{L}\left(\frac{F(M_{\text{opt}})}{\eta_s R_e}\right)^{\frac{1}{N_r}}, \quad (37)$$

where $F(M) = B(M)^{2N_r} + C(2^{\eta_s}\eta_s - M\log_2(M))$ and R_e denotes the value of the target BER. Substituting (37) into $P_m = E_m R_b/\eta_s$, the required RF output power is obtained by:

$$P_m = \frac{R_b N_0}{\eta_s L}\left(\frac{F(M_{\text{opt}})}{\eta_s R_e}\right)^{\frac{1}{N_r}}. \quad (38)$$

The BS energy consumption per bit E_{BS} is given by:

$$E_{\text{BS}} = \frac{P_{\text{BS}}}{R_b}. \quad (39)$$

TABLE II
TOSM DEPLOYMENT OF (M, N) OVER C.I.D. RAYLEIGH FADING

$\rho_{av} \backslash \eta_s$	$\eta_s = 4$	$\eta_s = 5$	$\eta_s = 6$	$\eta_s = 7$	$\eta_s = 8$
		$N_r = 1$			
$\rho_{av} = 0.0$	(16,1)	(16,2)	(16,4)	(16,8)	(16,16)
$\rho_{av} = 0.1$	(16,1)	(16,2)	(16,4)	(16,8)	(16,16)
$\rho_{av} = 0.2$	(16,1)	(16,2)	(16,4)	(16,8)	(16,16)
$\rho_{av} = 0.3$	(16,1)	(16,2)	(16,4)	(16,8)	(16,16)
$\rho_{av} = 0.4$	(16,1)	(16,2)	(16,4)	(16,8)	(16,16)
$\rho_{av} = 0.5$	(16,1)	(16,2)	(16,4)	(16,8)	(16,16)
$\rho_{av} = 0.6$	(16,1)	(16,2)	(16,4)	(16,8)	(16,16)
$\rho_{av} = 0.7$	(16,1)	(32,1)	(32,2)	(32,4)	(32,8)
$\rho_{av} = 0.8$	(16,1)	(32,1)	(32,2)	(32,4)	(32,8)
$\rho_{av} = 0.9$	(16,1)	(32,1)	(64,1)	(64,2)	(64,4)
$\rho_{av} = 1.0$	(16,1)	(32,1)	(64,1)	(128,1)	(256,1)

$\rho_{av} \backslash \eta_s$	$\eta_s = 4$	$\eta_s = 5$	$\eta_s = 6$	$\eta_s = 7$	$\eta_s = 8$
		$N_r = 2$			
$\rho_{av} = 0.0$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.1$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.2$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.3$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.4$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.5$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.6$	(8,2)	(8,4)	(8,8)	(8,16)	(8,32)
$\rho_{av} = 0.7$	(8,2)	(8,4)	(8,8)	(8,16)	(8,32)
$\rho_{av} = 0.8$	(8,2)	(8,4)	(8,8)	(8,16)	(8,32)
$\rho_{av} = 0.9$	(8,2)	(8,4)	(8,8)	(8,16)	(8,32)
$\rho_{av} = 1.0$	(16,1)	(32,1)	(32,2)	(32,4)	(32,8)

$\rho_{av} \backslash \eta_s$	$\eta_s = 4$	$\eta_s = 5$	$\eta_s = 6$	$\eta_s = 7$	$\eta_s = 8$
		$N_r = 3$			
$\rho_{av} = 0.0$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.1$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.2$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.3$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.4$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.5$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.6$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.7$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.8$	(8,2)	(8,4)	(8,8)	(8,16)	(8,32)
$\rho_{av} = 0.9$	(8,2)	(8,4)	(8,8)	(8,16)	(8,32)
$\rho_{av} = 1.0$	(16,1)	(16,2)	(16,4)	(16,8)	(16,16)

$\rho_{av} \backslash \eta_s$	$\eta_s = 4$	$\eta_s = 5$	$\eta_s = 6$	$\eta_s = 7$	$\eta_s = 8$
		$N_r = 4$			
$\rho_{av} = 0.0$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.1$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.2$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.3$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.4$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.5$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.6$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.7$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.8$	(4,4)	(4,8)	(4,16)	(4,32)	(4,64)
$\rho_{av} = 0.9$	(8,2)	(8,4)	(8,8)	(8,16)	(8,32)
$\rho_{av} = 1.0$	(16,1)	(16,2)	(16,4)	(16,8)	(16,16)

1) Continuous Mode: In the continuous mode, the energy consumption per bit of a BS based on TOSM is obtained by substituting (8) and (38) into (39) with $N_{\text{act}} = 1$:

$$E_{\text{BS}} = \frac{P_0}{R_b} + \frac{\zeta N_0}{\eta_s L}\left(\frac{F(M_{\text{opt}})}{\eta_s R_e}\right)^{\frac{1}{N_r}}. \quad (40)$$

2) DTX Mode: Similarly, based on (10), we can compute E_{BS} in the DTX mode as follows:

$$E_{\text{BS}} = \frac{P_s}{R_b} + \frac{N_0}{\eta_s L}\left(\zeta + \frac{P_0 - P_s}{P_{\max}}\right)\left(\frac{F(M_{\text{opt}})}{\eta_s R_e}\right)^{\frac{1}{N_r}}. \quad (41)$$
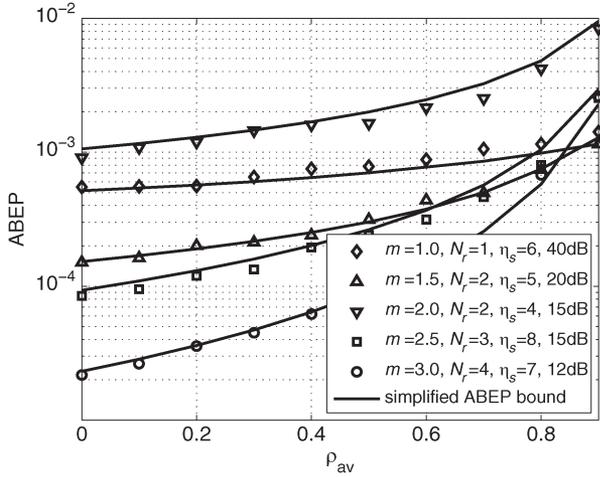
Fig. 4.   The simplified ABEP bound of SM versus simulations.

Fig. 5.   BER performance of fixed-SM schemes over c.i.d. Rayleigh fading.

V.   SIMULATION RESULTS

In this section, Monte Carlo results are presented to validate the performance of our scheme. Two cases are studied: i) c.i.d. Nakagami-m fading; and ii) c.i.d. Rayleigh fading. In the first case, we focus on verifying the accuracy of the simplified ABEP bound. The reason for choosing the second case is three-fold: i) multipath fading is typically modeled by a Rayleigh distribution; ii) by fixing the parameter m, the main trends of TOSM in relation to the channel correlation can be studied; and iii) a look-up table is formed within this scenario to exhibit the optimum transmit structure. In addition, the performance of TOSM is compared with some other schemes including fixed-SM, V-BLAST and STBC. Furthermore, the BS energy consumption and the maximum transmission rate are analyzed for the proposed scheme. To ensure the fairness among different values of η_s, the transmission energy per bit against noise, i.e. E_b/N_0, is used instead of SNR.

A.   Accuracy of the Simplified ABEP

In Fig. 4, the simplified ABEP expression of SM in (32) is verified against simulation results. To present an extensive comparison, several scenarios are considered by varying the shape factor, the number of receive antennas, the spectrum efficiency and E_b/N_0. A unit spread controlling factor is assumed. For a certain scenario, the BER curve of SM is shown as a function of the average correlation degree. As can be seen, in general, the theoretical curves match the simulation results well. Since the simplified ABEP is an approximation of the ABEP upper bound in [18], we expect some deviations especially at high channel correlations. Despite this, the simplified ABEP is still very close to the simulation results.

B.   Optimality of the Look-up Table

The deployment of (M, N) for TOSM over c.i.d. Rayleigh fading channels is shown in Table II. We use gray colour to highlight the situations in which the optimal number of transmit antennas is one, i.e. SM is not applied. The following outcomes are observed: i) given a certain η_s and N_r, the optimal N
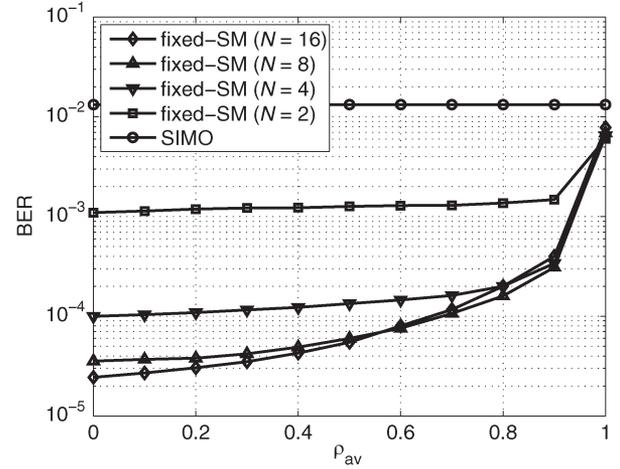
decreases as ρ_{av} increases. In other words, it is better to use SM with fewer transmit antennas at high channel correlations; ii) if $N_r = 1$, the best choice regresses to a single transmit-antenna scheme for an extremely high ρ_{av}; iii) when N_r is increased, SM is suitable for more cases of (η_s, ρ_{av}), and the optimal number of transmit antennas becomes larger; and iv) for a certain N_r and ρ_{av}, the best selection of M maintains a constant when η_s is large enough.

The look-up table can be extended to any needed situation and easily used to configure the deployment of the TOSM system. In practice, N_r and η_s are usually fixed for a BS. As a result, the only parameter that needs to be determined is the correlation coefficient, which can be obtained through the structured correlation estimator [25]. Using $\eta_s = 6$ and $N_r = 2$, for example, Fig. 5 shows the simulation results of various fixed-SM schemes at $E_b/N_0 = 25$ dB. As shown, SM using $N = 16$ outperforms $N = 8$ when ρ_{av} is below 0.5. However, the opposite result happens for $0.6 \leq \rho_{av} \leq 0.9$. At an extremely high correlation of $\rho_{av} = 1$, two transmit-antenna SM achieves the lowest BER. It can be observed that Table II fits the simulation results.

C.   BER Performance of Direct Antenna Selection

The BER performance of the proposed RCP approach is evaluated against two baseline schemes: i) the exhaustive search (ES); and ii) the worst case where the neighbouring antennas are selected. We refer to this scheme as worst selection (WS) in the sequel.

Fig. 6 and Fig. 7 present the BER performance of RCP for $\eta_s = 4$ and 5, respectively. Due to the intractable complexity of ES, the results when $\eta_s > 5$ are not presented. In addition, the antenna area is assumed to be the same to ensure a fair comparison for different η_s. Therefore, ρ_s is used instead of ρ_{av}. As shown, the RCP scheme achieves almost the same performance as ES with a gap of less than 0.3 dB. Furthermore, the negligible difference between RCP and ES is barely affected by the channel correlations, whereas the performance of WS becomes much worse as the correlation increases. To achieve the same BER value of 1×10^{-4} in the case of selecting 8 out
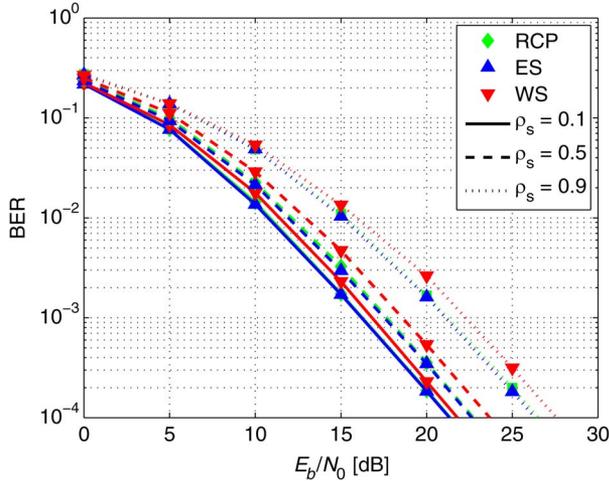
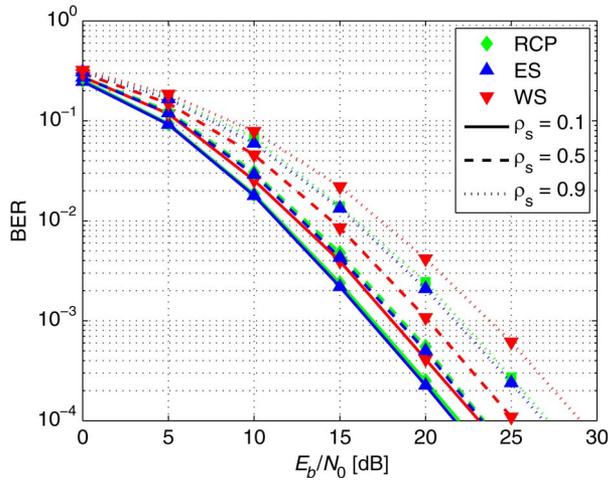Fig. 6.   BER performance of RCP for $N_t = 16$ and $N = 8$.

Fig. 7.   BER performance of RCP for $N_t = 32$ and $N = 8$.

of 32 antennas, in comparison with WS, RCP obtains an energy saving of 1.1 dB and 2.0 dB at $\rho_s = 0.1$ and 0.9, respectively.

D. BER Performance of TOSM

The complexity of the MIMO system depends on the number of RF chains rather than the total number of transmit antennas. Despite the requirement of large antennas at the transmitter, TOSM needs only one RF chain. For this reason, it is reasonable to compare our approach to fixed-SM schemes with the same η_s. Based on the obtained optimal N, we evaluate the BER performance of TOSM. Assuming $N_r = 2$ and $E_b/N_0 = 25$ dB, Figs. 8–10 show the BER results against the channel correlation for $\eta_s = 4$, 5 and 6, respectively. The case of $N = 1$ is referred to as single-input multiple-output (SIMO).

The following trends are observed: i) fixed-SM with more antennas is not always better than those using fewer antennas. This signifies that the benefit does not simply come from employing more transmit antennas; ii) TOSM always performs better than or equal to fixed-SM schemes, which validates the optimization results; and iii) when η_s increases, TOSM employs more transmit antennas and performs much better
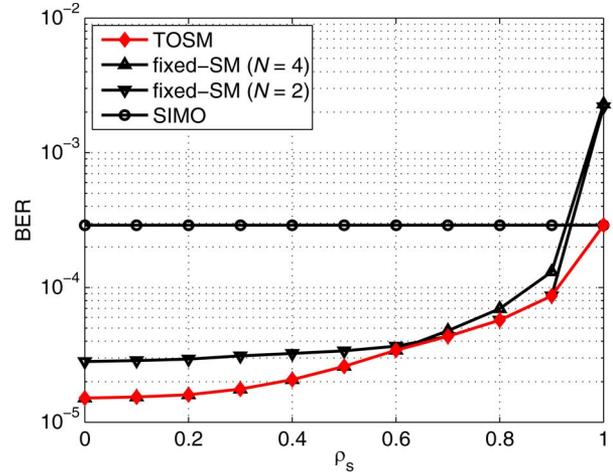

Fig. 8.   BER performance of TOSM against channel correlation for $\eta_s = 4$.
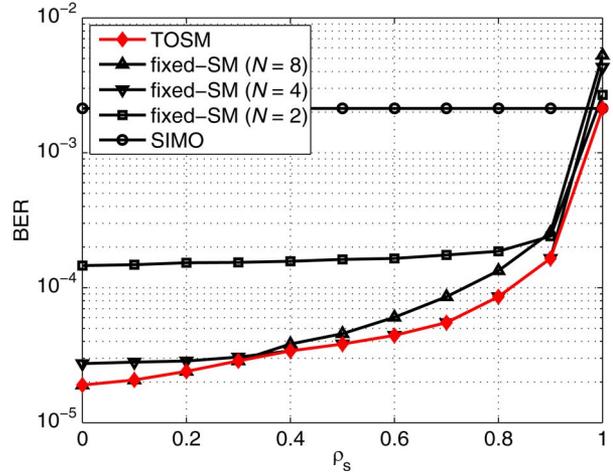

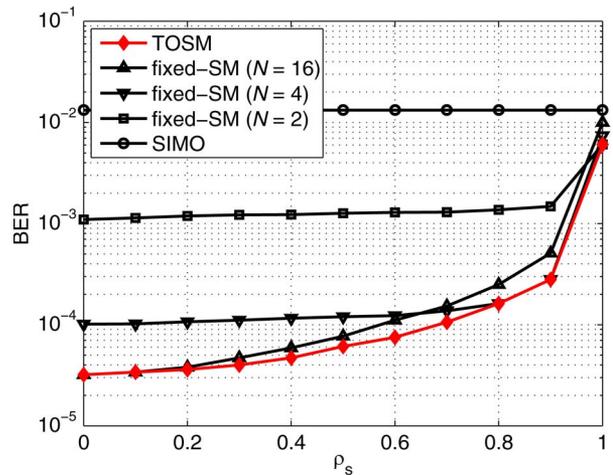Fig. 9.   BER performance of TOSM against channel correlation for $\eta_s = 5$.

Fig. 10.   BER performance of TOSM against channel correlation for $\eta_s = 6$.

than the fixed-SM with a small N. Specifically, TOSM slightly outperforms fixed-SM with $N = 2$ at both low and high correlations for $\eta_s = 4$. However, for $\eta_s = 5$ and 6, TOSM can always achieve a significant gain except when the channel correlation is extremely high. Similar, but less pronounced trends are
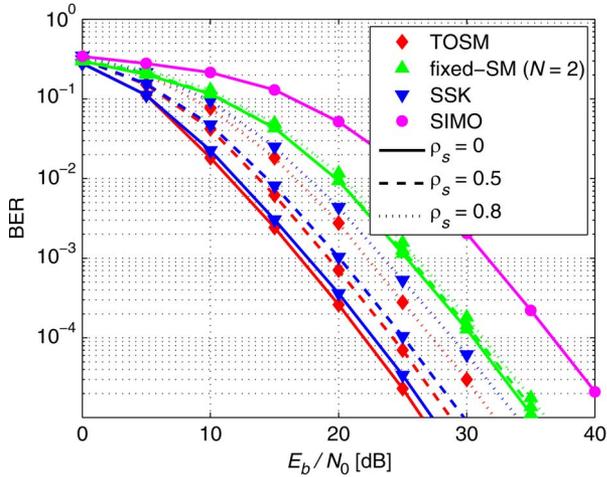
Fig. 11.   BER performance of TOSM against E_b/N_0 for $\eta_s = 6$.

Fig. 12.   Transmit energy consumption of TOSM.

noticed at lower SNRs. In Fig. 11, the BER performance of TOSM is shown as a function of E_b/N_0 for $\eta_s = 6$. As can be seen, TOSM significantly outperforms the other schemes for all presented SNRs and various channel correlation degrees. When the channels are independent, i.e. $\rho_s = 0$, TOSM saves energy in the regions of 0.8 dB, 8.7 dB, and 15.1 dB relative to SSK, fixed-SM with $N = 2$, and SIMO, respectively. As ρ_s increases, TOSM outperforms SSK more significantly. Conversely, fixed-SM with $N = 2$ is only slightly affected by the channel correlation, and the advantage of TOSM is diminishing with an increase of ρ_s. However, the gain of TOSM over fixed-SM with $N = 2$ still exceeds 4 dB at $\rho_s = 0.8$.

E. Energy Consumption

The overall BS energy consumption of TOSM is studied in contrast to V-BLAST and STBC as well as fixed-SM. The simulations are restricted in two aspects. On the one hand, the number of transmit antennas for V-BLAST is limited due to the constraint $N_r \geq N_t$. On the other hand, when using more transmit antennas, SM can save more energy in RF chains. The motivation here is to validate the energy efficiency of TOSM. To carry out a relatively fair comparison, the case of $\eta_s = 4$ and $N_r = 4$ is chosen, where TOSM employs four transmit antennas for the channel correlation varied from 0 to 0.8. The same 4×4 MIMO structure is arranged for both V-BLAST in [26] and rate 3/4 STBC in [27]. In addition, the path loss is assumed to be 100 dB without considering the shadowing.

Fig. 12 gives the transmit energy consumption results for a target BER value of 1×10^{-4}. It is noticed that, TOSM provides a remarkable and stable gain of around 5 dB in comparison with STBC and more with V-BLAST. The overall BS energy consumptions in both the continuous mode and the DTX mode are shown in Fig. 13. To maintain a certain E_b, P_{max} leads to a ceiling of the transmission rate. For this reason, a bit rate of 30 Mbit/s is chosen to ensure the BS works physically, and we compare TOSM with STBC to show the trends. The following outcomes are observed: i) using the DTX mode for TOSM provides a gain of around 1.4 dB over the continuous mode; ii) TOSM significantly outperforms
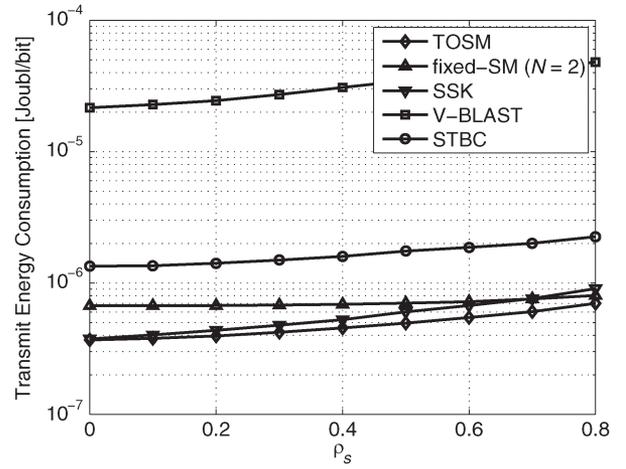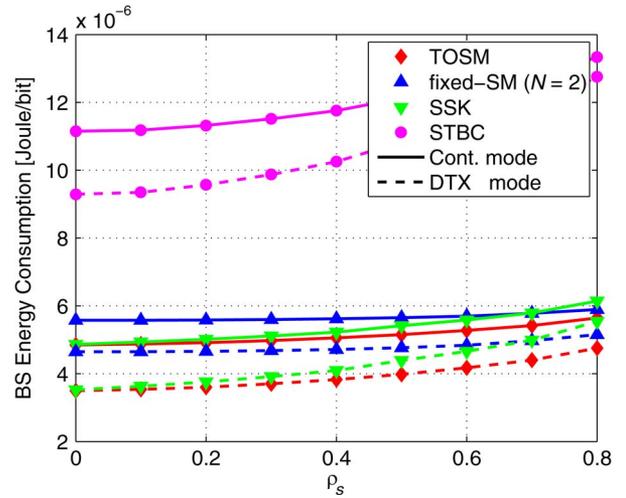

Fig. 13.   BS energy consumption of TOSM for $R_b = 30$ Mbit/s.

STBC for both the continuous mode and the DTX mode with energy-saving gains of at least 6.3 μJoule/bit (56%) and 5.8 μJoule/bit (62%), respectively; iii) compared with multi-stream MIMO schemes, TOSM requests much less energy because of the single RF chain requirement; iv) in both modes, the gain obtained by TOSM increases as ρ_s increases; and v) TOSM saves more energy in the DTX mode than the continuous mode, especially at high channel correlations. When $\rho_s = 0.8$, for example, TOSM outperforms STBC by 7.7 μJoule/bit (57%) and 8.0 μJoule/bit (63%) in the continuous mode and the DTX mode, respectively.

F. Maximum Transmission Rate

In contrast to multi-stream MIMO schemes, SM requires only one RF chain, and therefore less energy is requested to drive it. However, the maximum RF output power of SM is $1/N_{act}$ of the MIMO scheme with N_{act} active antennas, and this restricts the maximum transmission rate of SM. Fig. 14 shows $R_{b\,max}$ as a function of the channel correlation in the same scenario of the previous subsection. As can be seen, the maximum transmission rate of fixed-SM outperforms V-BLAST, but
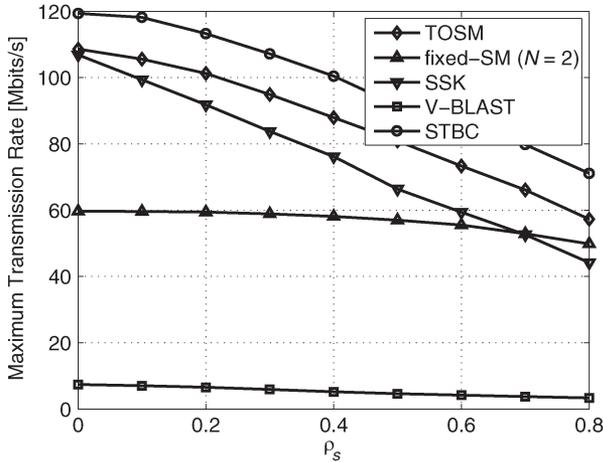
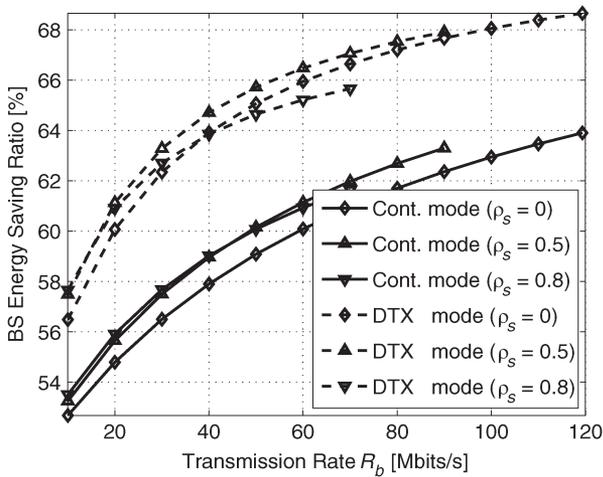Fig. 14.   Maximum transmission rate of TOSM.

Fig. 15.   BS energy saving ratio between TOSM and STBC.

is much smaller than STBC. Meanwhile, TOSM offers a great improvement of maximum transmission rate to the original SM and overcomes this disadvantage. In Fig. 15, the BS energy consumption gains between TOSM and STBC are presented in terms of the transmission rate. As shown, the gain of TOSM obtained in the DTX mode is larger than the continuous mode for various correlation degrees. Also, DTX is much better than the continuous mode when R_b increases to the full load. This signifies that TOSM is more robust and energy-efficient when combined with DTX.

VI. CONCLUSION

In this paper, we proposed an optimum transmit structure for SM, which balances the size of the spatial constellation diagram and the size of the signal constellation diagram. Instead of using exhaustive search, a novel two-stage TAS method has been proposed for reducing the computational complexity, where the optimal number of transmit antennas and the specific antenna positions are determined separately. The first step is to obtain the optimal number of transmit antennas by minimizing a simplified ABEP bound for SM. In the second step, a direct antenna selection method, named RCP, was developed to select

the required number of transmit antennas from an antenna array. In addition, a look-up table was built in the case of c.i.d. Rayleigh fading, which can readily be used to determine the optimum transmit structure. Results show that TOSM improves the BER performance of the original SM significantly, and outperforms V-BLAST and STBC greatly in terms of the overall BS energy consumption. A further study shows that TOSM is more energy efficient when combined with the DTX mode than the continuous mode. Furthermore, the issue with respect to the maximum transmission rate in the SM systems has been addressed, which is caused by the limited output power of a single RF chain. It was shown that TOSM uplifts the maximum transmission rate of the original SM greatly, and diminishes the gap between SM and STBC significantly. All these merits make TOSM a highly energy-efficient, low-complexity scheme to satisfy the requirement of high data rate transmission, and an ideal candidate for massive MIMO [28]. Further research will extend the optimum transmit structure to generalized SM [29], where several antennas are activated simultaneously.

REFERENCES

[1] X. Wu, S. Sinanovic, M. Di Renzo, and H. Haas, "Structure optimisation of spatial modulation over correlated fading channels," in *Proc. IEEE GLOBECOM*, Dec. 2012, pp. 4049–4053.

[2] X. Wu, S. Sinanovic, M. Di Renzo, and H. Haas, "Base station energy consumption for Transmission Optimised Spatial Modulation (TOSM) in correlated channels," in *Proc. 17th IEEE CAMAD Commun. Links Netw.*, Sep. 2012, pp. 261–265.

[3] A. Fehske, G. Fettweis, J. Malmodin, and G. Biczok, "The global footprint of mobile communications: The ecological and economic perspective," *IEEE Commun. Mag.*, vol. 49, no. 8, pp. 55–62, Aug. 2011.

[4] H. Haas and S. McLaughlin, Eds., *Next Generation Mobile Access Technologies: Implementing TDD*.   Cambridge, U.K.: Cambridge Univ. Press, Jan. 2008.

[5] M. Gruber *et al.*, "EARTH—Energy aware radio and network technologies," in *Proc. 20th IEEE Int. Symp. PIMRC*, Sep. 2009, pp. 1–5.

[6] G. Auer, V. Giannini, M. Olsson, M. Gonzalez, and C. Desset, "Framework for energy efficiency analysis of wireless networks," in *Proc. 2nd Int. Conf. Wireless VITAE Commun. Syst. Technol.*, Feb. 2011, pp. 1–5.

[7] P. Frenger, P. Moberg, J. Malmodin, Y. Jading, and I. Godor, "Reducing energy consumption in LTE with cell DTX," in *Proc. 73rd IEEE VTC Spring*, May 2011, pp. 1–5.

[8] A. Chatzipapas, S. Alouf, and V. Mancuso, "On the minimization of power consumption in base stations using on/off power amplifiers," in *Proc. IEEE Online Conf. GreenCom*, Sep. 2011, pp. 18–23.

[9] Z. Xu *et al.*, "Energy-efficient MIMO-OFDMA systems based on switching off RF chains," in *Proc. 74th IEEE VTC Fall*, Sep. 2011, pp. 1–5.

[10] M. Di Renzo, H. Haas, and P. M. Grant, "Spatial modulation for multiple-antenna wireless systems: A survey," *IEEE Commun. Mag.*, vol. 49, no. 12, pp. 182–191, Dec. 2011.

[11] R. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228–2241, Jul. 2008.

[12] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Spatial modulation: Optimal detection and performance analysis," *IEEE Commun. Lett.*, vol. 12, no. 8, pp. 545–547, Aug. 2008.

[13] M. Di Renzo and H. Haas, "A general framework for performance analysis of Space Shift Keying (SSK) modulation for MISO correlated Nakagami-m fading channels," *IEEE Trans. Commun.*, vol. 58, no. 9, pp. 2590–2603, Sep. 2010.

[14] P. Yang, Y. Xiao, Y. Yu, and S. Li, "Adaptive spatial modulation for wireless MIMO transmission systems," *IEEE Commun. Lett.*, vol. 15, no. 6, pp. 602–604, Jun. 2011.

[15] H. Holtkamp, G. Auer, and H. Haas, "On minimizing base station power consumption," in *Proc. 74th IEEE VTC Fall*, Sep. 2011, pp. 1–5.

[16] S. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 369–371, Sep. 2001.

[17] G. Auer *et al.*, "Cellular energy efficiency evaluation framework," in *Proc. 73rd IEEE VTC Spring*, May 2011, pp. 1–6.

[18] M. Di Renzo and H. Haas, "Bit error probability of SM-MIMO over generalized fading channels," *IEEE Trans. Veh. Technol.*, vol. 61, no. 3, pp. 1124–1144, Mar. 2012.

[19] Z. Hasan, H. Boostanimehr, and V. Bhargava, "Green cellular networks: A survey, some research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 4, pp. 524–540, 4th Quart. 2011.

[20] M. Di Renzo and H. Haas, "Bit error probability of space modulation over Nakagami-m fading: Asymptotic analysis," *IEEE Commun. Lett.*, vol. 15, no. 10, pp. 1026–1028, Oct. 2011.

[21] M. K. Simon and M. S. Alouini, *Digital Communication over Fading Channels*, 2nd ed. Piscataway, NJ, USA: Wiley/IEEE Press, 2005.

[22] M. S. Bazaraa, H. D. Sherali, and C. Shetty, *Nonlinear Programming: Theory and Algorithms*, 3rd ed. Hoboken, NJ, USA: Wiley-Interscience, 2006.

[23] N. Serafimovski, M. Di Renzo, S. Sinanović, R. Y. Mesleh, and H. Haas, "Fractional Bit Encoded Spatial Modulation (FBE-SM)," *IEEE Commun. Lett.*, vol. 14, no. 5, pp. 429–431, May 2010.

[24] K. Stphenson, *Introduction to Circle Packing: The Theory of Discrete Analytic Function*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[25] G. Matz, "Recursive MMSE estimation of wireless channels based on training data and structured correlation learning," in *Proc. 13th IEEE Workshop Stat. Signal Process.*, Jul. 2005, pp. 1342–1347.

[26] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. ISSSE*, Sep. 1998, pp. 295–300.

[27] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.

[28] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[29] A. Younis, N. Serafimovski, R. Mesleh, and H. Haas, "Generalised spatial modulation," in *Proc. 44th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2010, pp. 1498–1502.

Marco Di Renzo (S'05–AM'07–M'09–SM'14) received the Laurea (cum laude) and the Ph.D. degrees in electrical and information engineering from the Department of Electrical and Information Engineering, University of L'Aquila, Italy, in April 2003 and in January 2007. In October 2013, he received the Habilitation á Diriger des Recherches (HDR) from the University of Paris-Sud XI, Paris, France. Since January 2010, he has been a Tenured Academic Researcher ("Chargé de Recherche Titulaire") with the French National Center for Scientific Research (CNRS), as well as a faculty member of the Laboratory of Signals and Systems (L2S), a joint research laboratory of the CNRS, the école Supérieure d'électricité (SUPéLEC) and the University of Paris-Sud XI, Paris, France. His main research interests are in the area of wireless communications theory.

Dr. Di Renzo is a recipient of several awards, which include a special mention for the outstanding five-year (1997–2003) academic career, University of L'Aquila, Italy; the THALES Communications fellowship (2003–2006), University of L'Aquila, Italy; the 2004 Best Spin-Off Company Award, Abruzzo Region, Italy; the 2006 DEWS Travel Grant Award, University of L'Aquila, Italy; the 2008 Torres Quevedo Award, Ministry of Science and Innovation, Spain; the "Dérogation pour l'Encadrement de Théese" (2010), University of Paris-Sud XI, France; the 2012 IEEE CAMAD Best Paper Award; the 2012 IEEE WIRELESS COMMUNICATIONS LETTERS Exemplary Reviewer Award; the 2013 IEEE VTC-Fall Best Student Paper Award; the 2013 Network of Excellence NEWCOM# Best Paper Award; the 2013 IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY Top Reviewer Award; the 2013 IEEE-COMSOC Best Young Researcher Award for Europe, Middle East and Africa (EMEA Region); and the 2014 IEEE ICNC Single Best Paper Award Nomination (Wireless Communications Symposium). Currently, he serves as an Editor of the IEEE COMMUNICATIONS LETTERS and of the IEEE TRANSACTIONS ON COMMUNICATIONS (Heterogeneous Networks Modeling and Analysis).

Xiping Wu (S'12–M'15) received the B.Sc. degree from Southeast University, Nanjing, China, in 2008 and the M.Sc. degree (with distinction) from the University of Edinburgh, Scotland, U.K., in 2011. From September 2011 to August 2014, he was a Marie-Curie Early-Stage Researcher (ESR) as well as a Ph.D. candidate at the University of Edinburgh, funded by the European Union's Seventh Framework Programme (FP7) project GREENET. From December 2013 to April 2014, he was on secondment to the Department of Electrical and Information Engineering, University of L'Aquila, Italy. Since September 2014, he has been a Research Associate with the Institute for Digital Communications (IDCOM), University of Edinburgh, funded by the British Engineering and Physical Sciences Research Council (EPSRC). His main research interests are in the areas of wireless communication theory, optical wireless communications, and wireless networking. In 2010, he was granted the Scotland Saltire Scholarship by the Scottish Government.

Harald Haas (S'98-AM'00-M'03) received the Ph.D. degree from the University of Edinburgh in 2001. He currently holds the Chair of Mobile Communications at the University of Edinburgh. His main research interests are in optical wireless communications, hybrid optical wireless and RF communications, spatial modulation, and interference coordination in wireless networks. He first introduced and coined "spatial modulation" and "Li-Fi". Li-Fi was listed among the 50 best inventions in TIME Magazine 2011.

He was an invited speaker at TED Global 2011, and his talk has been watched online more than 1.5 million times. He is Co-founder and Chief Scientific Officer (CSO) of pureLiFi Ltd. He holds 31 patents and has more than 30 pending patent applications. He has published 300 conference and journal papers including a paper in *Science*. He was co-recipient of a Best Paper Award at the IEEE Vehicular Technology Conference in Las Vegas, NV, USA, in 2013. In 2012, he was the only recipient of the prestigious Established Career Fellowship from the Engineering and Physical Sciences Research Council (EPSRC) within Information and Communications Technology in the U.K. He is recipient of the Tam Dalyell Prize 2013 awarded by the University of Edinburgh for excellence in engaging the public with science. In 2014, he was selected by EPSRC as one of ten Recognising Inspirational Scientists and Engineers (RISE) Leaders.