Abstract—Exploiting signal space diversity (SSD) to improve the error performance of communications systems over fading channels has been shown to be an effective technique. The application of SSD in bit-interleaved coded modulation with iterative decoding (BICM-ID) is considered for cascaded Rayleigh fading channels, which are suitable for mobile-to-mobile communications. A tight bound on the asymptotic error performance is derived. The bound is then used to find the optimal rotation matrix. It is shown that employing SSD in a sufficiently large constellation can close the performance gap between a conventional Rayleigh fading channel and a cascaded Rayleigh fading channel. In fact, similar to the case of conventional Rayleigh fading, it is demonstrated that the error performance of BICM-ID with SSD over a cascaded Rayleigh fading channel can also closely approach the performance of BICM-ID over an AWGN channel. Various analytical and simulation results are provided to confirm the analysis.

I. INTRODUCTION

Signal space diversity (SSD) [1], [2] has recently been recognized as a very effective method to increase the diversity order, and therefore, the reliability of the wireless transmissions over fading channels. By applying the technique of SSD with high modulation diversity and good minimum product distance, it was shown that the error performance of rotated lattice constellations with or without channel coding over a classical Rayleigh fading channel becomes insensitive to fading, i.e., it closely approaches the error performance over an AWGN channel [2], [3].

Most of research works in the literature have only considered two simplified fading models, namely the Rayleigh fading and Rician fading models for the wireless channels. However, such channel models are only suitable to characterize the statistical properties of the conventional cellular radio systems where the mobile station is moving and the base-station is stationary. On the other hand, in several new applications such as mobile ad-hoc networks and dedicated short-range communication systems for intelligent highway, both transmitter and receiver are in motion, which results in different statistical properties of the wireless channel. Recently, both theoretical analysis and experimental results [4]–[6] indicate that cascaded Rayleigh fading is more suitable to model the mobile-to-mobile communications channels. This fact thus brings new challenges in performance analysis, and as a consequence, the design criterion for mobile-to-mobile communications. Surprisingly, to date, only a few efforts have been devoted to performance analysis of communications systems operating over cascaded Rayleigh fading channels [5]–[7]. For example, by studying the pairwise error probability of uncoded systems over cascaded Rayleigh fading channels, it was demonstrated in [7] that only a partial diversity order can be achieved, which implies a performance degradation compared to that over a conventional Rayleigh fading channel.

This paper is concerned with the application of SSD in BICM-ID to improve the error performance of communications systems over cascaded Rayleigh fading channels. At first, the tight bound on the asymptotic error performance is derived. Based on the error bound, the design criterion for the optimal rotation matrix is found. It turns out that the optimal rotation matrix (with respect to the asymptotic performance) for a cascaded Rayleigh fading channel coincides with that for a conventional Rayleigh fading channel. Furthermore, it is shown that employing the technique of SSD in a sufficiently large dimension can close the performance gap between the two fading channels. Moreover, it is observed that by employing the optimal rotation matrix, the error performance of BICM-ID with SSD over a cascaded Rayleigh fading channel can closely approach the performance of BICM-ID over an AWGN channel. This is similar to the observation made in [3] for the case of BICM-ID with SSD over a conventional Rayleigh fading channel.

II. SYSTEM MODEL

The block diagram of a BICM-ID system with SSD over a cascaded Rayleigh fading channel is shown in Fig. 1.
The block diagram of a BICM-ID system with SSD over a cascaded Rayleigh fading channel is shown in Fig. 1. The information sequence $\mathbf{u}$ is first encoded into a coded sequence $\mathbf{c}$. The coded sequence $\mathbf{c}$ is then interleaved by a bit-wise interleaver to become the interleaved sequence $\mathbf{c}$. Each group of $N_m$ coded bits in $\mathbf{c}$ is then mapped to one complex $N$-dimensional ($N$-dim) constellation symbol by some mapping rule $\xi$ to produce the symbol sequence $\mathbf{s} = [s_1, s_2, \ldots, s_N]$, where $s_i \in \Omega$, $\forall i$, with $\Omega$ is a two-dimensional constellation. For simplicity, assume that the mapping $\mathbf{c}$ is implemented independently and identically for each complex component $s_i$, i.e., the mapping rule $\xi$ is employed over the constellation $\Omega$. The “super” symbol $\mathbf{s}$ in the $N$-dim constellation $\Psi$ is then rotated by an $N \times N$ complex rotation matrix $\mathbf{G}$. The rotated symbol $\mathbf{x} = [x_1, x_2, \ldots, x_N]$ corresponding to a new rotated constellation $\Psi_t$ is given by $\mathbf{x}^\top = \mathbf{G} \mathbf{s}^\top$. The entries $g_{i,u}$, $1 \leq i, u \leq N$, of $\mathbf{G}$ satisfy the following power constraint

$$\sum_{i=1}^{N} \sum_{u=1}^{N} ||g_{i,u}||^2 = N.$$  

Consider the cascaded Rayleigh fading channel with the assumption that the channel changes independently in each component duration of the $N$-dim signal symbol. Furthermore, assume that the receiver has perfect channel state information. The $N$-dim received signal $r$ can then be represented as $r^\top = \mathbf{H} \mathbf{G} \mathbf{s}^\top + \mathbf{w}^\top$. Here, each entry of $\mathbf{w} = [w_1, \ldots, w_N]$ is a circularly symmetric complex Gaussian random variable of variance $N_0$. The matrix $\mathbf{H} = \text{diag}(h_1, \ldots, h_N)$ contains the fading coefficients in its diagonal. In the case of cascaded Rayleigh fading, each fading coefficient $h_i$ is represented by the product of two independent circularly symmetric complex Gaussian random variables $h_i = a_i b_i$, where both $a_i$ and $b_i$ have unit variance.

As shown in Fig. 1, the receiver in a BICM-ID system includes the soft-input soft-output (SISO) demodulator and the SISO channel decoder. The SISO channel decoder uses the MAP algorithm in [8], which can be efficiently implemented. On the other hand, the complexity of the optimal MAP demodulator grows exponentially with the number of coded bits per symbol $N_m$. To overcome this difficulty, the suboptimal, low-complexity but yet effective method proposed in [3] using the minimum mean-square error (MMSE) estimator and the sigma mapping, or the Gaussian approximations proposed in [9] can be attractive alternatives.

### III. PERFORMANCE EVALUATION

In [3], a tight union bound on the asymptotic bit error probability (BEP) of BICM-ID with SSD over a conventional Rayleigh fading channel is obtained based on the assumption of error-free feedback from the decoder to the demodulator. This section derives a similar bound but for a cascaded Rayleigh fading channel.

The union bound on the asymptotic BEP for a BICM-ID with SSD that employs a rate-$k_c/n_c$ convolutional code, a complex $N$-dim constellation $\Psi$ and a mapping rule $\xi$ is given as [3]:

$$P_b \leq \frac{1}{k_c} \sum_{d=d_H}^\infty c_d \int_{0}^{\pi/2} f(d, \Psi, \xi, \mathbf{G}) \, d\theta,$$

where $c_d$ is the total information weight of all error events at Hamming distance $d$ and $d_H$ is the free Hamming distance of the code.

The function $f(d, \Psi, \xi, \mathbf{G})$ is the average PEP, which can be computed from the PEP of two codewords as follows. Let $\mathbf{c}$ and $\mathbf{c}$ denote the input and estimated sequences, respectively, with Hamming distance $d$. These binary sequences correspond to the sequences $\mathbf{s}$ and $\mathbf{s}$. Without loss of generality, it is assumed that $\mathbf{c}$ and $\mathbf{c}$ differ in the first $d$ consecutive bits. Hence, $\mathbf{s}$ and $\mathbf{s}$ can be redefined as sequences of $d$ complex $N$-dim symbols as $\mathbf{s} = [s_1, \ldots, s_d]$ and $\mathbf{s} = [s_1, \ldots, s_d]$. Here, $s_e$ and $s_e$, $1 \leq e \leq d$, belong to $\Psi$. Also let $\mathbf{H} = [H_1, \ldots, H_d]$, where $H_e = \text{diag}(h_{e,1}, \ldots, h_{e,N})$, $1 \leq e \leq d$, represents the path gains that affect the transmitted symbol $s_e$. As mentioned before, over a cascaded Rayleigh fading channel, the path gain $h_{e,i}$, $1 \leq i \leq N$, is expressed as $a_{e,i} b_{e,i}$. The two symbols $s_e$ and $s_e$ correspond to the two rotated symbols $x_e^\top = \mathbf{G} s_e^\top$ and $x_e^\top = \mathbf{G} s_e^\top$. Similar to [3], the PEP conditioned on $\mathbf{H}$ can be computed as follows:

$$P(\mathbf{s} \rightarrow \tilde{\mathbf{s}}, \mathbf{H}) = \frac{1}{2N} \sum_{e=1}^{d} d^2(x_e, \tilde{x}_e|H_e),$$

where

$$d^2(x_e, \tilde{x}_e|H_e) = \sum_{i=1}^{N} ||a_{e,i}||^2 ||b_{e,i}||^2 ||G_i(s_e - \tilde{s}_e)^\top||^2$$

In (3), $G_i$ is the $i$th row of $\mathbf{G}$. Using the identity $Q(\gamma) = \frac{1}{\pi} \int_{0}^{\pi/2} \frac{\sin \theta}{\theta} \, d\theta$ and averaging over the Rayleigh random variables $||a_{e,i}||$ and $||b_{e,i}||$ as similar to [7], one obtains

$$P(\mathbf{s} \rightarrow \tilde{\mathbf{s}}) = \frac{1}{\pi} \int_{0}^{\pi/2} \left( \prod_{e=1}^{d} \Delta_e \right) \, d\theta,$$

where

$$\Delta_e = \prod_{i=1}^{N} E_1 \left( \frac{4N \sin^2 \theta}{\sqrt{||G_i(s_e - \tilde{s}_e)^\top||^2}} \right),$$

and $E_1(x) = \int_{x}^{\infty} \frac{\exp(-t)}{t} \, dt$ and $E_0(x) = \exp(-x)$ [10].

Thanks to the success of iterative decoding steps as normally seen in BICM-ID systems, one is most interested in the asymptotic performance to which the iterations converge [3], [11], [12]. Such asymptotic performance can be obtained by assuming that the iterations between the SISO decoder and the SISO demodulator work perfectly, or equivalently, one has perfect a priori information of the coded bits fed-back to the demodulator. Therefore, it can be assumed that the the labels of $s_e$ and $s_e$ differ in only 1 bit. Furthermore, observe that $\{\Delta_e\}$ are i.i.d. random variables. Then the union bound on $f(d, \Psi, \xi, \mathbf{G})$ in (1) can be computed by averaging over the constellation $\Psi$, instead of averaging over all pairs of codewords $\mathbf{s}$ and $\mathbf{s}$, as

$$f(d, \Psi, \xi, \mathbf{G}) \leq \frac{1}{\pi} \int_{0}^{\pi/2} \left[ \frac{E_{s,p}(\Delta(s, p, G))}{\gamma(\Psi, \xi, \mathbf{G})} \right] \, d\theta.$$
where
\[
\Delta(s, p, G) = \prod_{i=1}^{N} E_1 \left( \frac{4N_0 \sin^2 \theta}{\|G_i(s-p)\|^2} \right).
\] (7)

The expectation in (6) is over all the pairs of \(s\) and \(p\) in \(\Psi\) whose labels differ in only 1 bit, which is computed as:
\[
\gamma(\Psi, \xi, G) = \frac{1}{N m^2 \sum_{s \in \Psi} \sum_{k=1}^{N} \Delta(s, p, G)}. \tag{8}
\]

Similar to the analysis in [3], it can be verified that if \(\xi\) is implemented independently for each component of \(s\), then there is only one distinct component between \(s\) and \(p\) as long as their labels differ in only 1 bit. After some manipulations, the average in (8) can be computed by simply averaging over all pairs of signal points \((s, p) \in \Omega\) whose labels differ in only 1 bit at position \(j\), \(1 \leq j \leq m\). That is,
\[
\gamma(\Psi, \xi, G) = \frac{1}{m^2 \sum_{s \in \Omega} \sum_{j=1}^{m} \delta(s, p, G)}, \tag{9}
\]
where
\[
\delta(s, p, G) = \frac{1}{N} \prod_{u=1}^{N} \frac{E_1 \left( \frac{4N_0 \sin^2 \theta}{\|g_{u}(s-p)\|^2} \right)}{E_0 \left( \frac{4N_0 \sin^2 \theta}{\|g_{u}(s-p)\|^2} \right)}. \tag{10}
\]

Applying (9) to compute \(f(d, \Psi, \xi, G)\) in (6), the tight union bound on the asymptotic BEP in (1) can be efficiently computed via a single integral.

For a fixed constellation \(\Omega\) and mapping rule \(\xi\), the problem of choosing good rotation matrix \(G\) in order to minimize the asymptotic performance can be answered by applying the Chernoff bound \(\sqrt{\frac{2\pi}{d}} < \frac{1}{2} \exp(-d)\). Using the same analysis as before, one obtains the approximated version of the Chernoff bound as
\[
f(d, \Psi, \xi, G) \sim \frac{1}{2} \alpha^d(\Psi, \xi, G), \tag{11}
\]
where
\[
\alpha(\Psi, \xi, G) = \frac{1}{N} \sum_{s \in \Omega} \sum_{j=1}^{m} \kappa(s, p, G), \tag{12}
\]
with
\[
\kappa(s, p, G) = \frac{1}{N} \prod_{u=1}^{N} \frac{E_1 \left( \frac{4N_0 \sin^2 \theta}{\|g_{u}(s-p)\|^2} \right)}{E_0 \left( \frac{4N_0 \sin^2 \theta}{\|g_{u}(s-p)\|^2} \right)}. \tag{13}
\]

It follows from the above analysis that, as far as minimizing the asymptotic BEP is concerned, the optimal rotation matrix \(G\) is the one that minimizes \(\kappa(\Psi, \xi, G)\). In the next section, the optimal choice of \(G\) is addressed.

### IV. Optimal Rotation Matrix \(G\)

For given signal points \(s\) and \(p\) whose labels differ in only 1 bit at position \(j\), using the Cauchy inequality gives the following inequality on the parameter \(\kappa(s, p, G)\) in (13):
\[
(\kappa(s, p, G))^N \geq \prod_{u=1}^{N} \prod_{i=1}^{N} \frac{E_1(x_{i,u})}{E_0(x_{i,u})}, \tag{14}
\]
where \(x_{i,u} = \frac{4N_0}{\|g_{u}(s-p)\|^2}\). It is easy to see that the inequality holds when \(\{x_{i,u}\}\), or equivalently, \(\{|y_{i,u}|\}\) are equal for all \(1 \leq i, u \leq N\). This is because \(\frac{d}{dx} \left[ \frac{E_1(x)}{E_0(x)} \right] > 0\) [10, 5.1.21], which implies that \(\frac{E_1(x)}{E_0(x)}\) is an increasing function. Furthermore, from the power constraint \(\sum_{u=1}^{N} \sum_{i=1}^{N} \frac{1}{\|g_{i}(s-p)\|^2} = N\)

one has\(\sum_{u=1}^{N} \sum_{i=1}^{N} \frac{1}{\|g_{i}(s-p)\|^2} = A = \text{const.}\) Denote the right-hand side of (14) as \(q\{x_{i,u}\}\), and consider the function \(q\{\{x_{i,u}\}\}\) with variables \(y_{i,u} = 1/x_{i,u}\). From the fact that the function \(\frac{E_1(1/y)}{E_0(1/y)}\) is a log-convex function of \(y\ [13]\), one has [13]:
\[
q\{\{x_{i,u}\}\} = \prod_{u=1}^{N} \prod_{i=1}^{N} \frac{E_1(1/y_{i,u})}{E_0(1/y_{i,u})} \geq \left[ \frac{E_1(\frac{N^2}{A})}{E_0(\frac{N^2}{A})} \right]^{N^2}. \tag{15}
\]

From (15), it can be observed that the function \(q\{\{x_{i,u}\}\}\) and the parameter \(\kappa(s, p, G)\) achieve the minimum values if and only if \(x_{i,u}\) are equal for all \(i\) and \(u\). Equivalently, it can be stated that the optimal rotation matrix \(G\) with respect to the asymptotic performance is the matrix with all entries equal in magnitude. This turns out to be the same class of optimal rotation matrix obtained for a conventional Rayleigh fading channel in [3].

The above analysis is only concerned with the asymptotic performance. Typically, one is also interested in the convergence behavior of the performance with iterations. Unfortunately, dealing with this problem is not simple, since it involves many variables. Hence, similar to [3], we also restrict our attention to the class of unitary matrix with all entries equal in magnitude, which was well-studied in the literature. For example, the unitary matrix with size \(N = 2^t\) introduced in [2], [14] falls into the class of optimal \(G\), which is obtained as:
\[
G_N = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & \alpha_1 & \ldots & \alpha_1^{N-1} \\ \vdots & \vdots & & \vdots \\ 1 & \alpha_N & \ldots & \alpha_{N}^{N-1} \end{bmatrix}, \tag{16}
\]
where \(\alpha_1 = \alpha = \exp \left( \frac{2\pi j}{2^t} \right)\) and \(\alpha_0 = \alpha = \exp \left( \frac{2\pi j}{2^t} \right)\). More generally, the optimal unitary rotation matrix \(G\) can be constructed based on the inverse fast Fourier transform (IFFT) matrix as follows [14]:
\[
G = F_N^T \text{diag} \left(1, \beta, \ldots, \beta^{N-1}\right), \quad \beta = \exp \left( \frac{2\pi j}{B} \right), \tag{17}
\]
where \(B\) is an arbitrary integer and \(F_N\) is the \(N\)-point IFFT matrix whose \((i, u)\)th entry is given by \(\frac{1}{\sqrt{N}} \exp \left( j \frac{2\pi (i-1)(u-1)}{N} \right)\). Using this class of rotation matrix ensures that the performance after the first iteration of BICM-ID system with SSD is similar to the performance of a BICM-ID system without SSD at low SNR. Furthermore, following the same analysis as in the previous section, it can be shown that it will outperform the performance after 1 iteration for BICM-SD without SSD at high SNR.

Before closing this section, it is worth mentioning that the class of optimal rotation matrices discussed here does not guarantee the high modulation diversity constellation, which is crucial for the performance of uncoded systems over fading channels [1], [2]. However, as confirmed later by simulation...
results, the error performances for systems using different rotation matrices satisfying the unitary condition with all entries equal in magnitude can approach the same error bound, regardless of the modulation diversity of the constellations.

V. PERFORMANCE COMPARISON

By using the optimal rotation matrix $G$ with all elements equal in magnitude, i.e., $|g_{i,u}|^2 = \frac{1}{2}$ for all $i$ and $u$, the parameter $\delta(s, p, G)$ in (10) for given signal pair $s$ and $p$ whose label differs in only 1 bit at the $j$th position can be computed as:

$$\delta(s, p, G) = \left( \frac{E_1(N \cdot z)}{E_0(N \cdot z)} \right)^N, \quad (18)$$

where

$$z = \frac{4N_0 \sin^2 \theta}{||(s - p)||^2}. \quad (19)$$

In the case of BICM-ID with SSD over a conventional Rayleigh fading channel, it was shown in [3] that one can compute the function $f_{\text{Ray}}(d, \Psi, \xi, G)$, which is similar to $f(d, \Psi, \xi, G)$ in (6), as:

$$f_{\text{Ray}}(d, \Psi, \xi, G) \leq \frac{1}{\pi} \int_0^{\pi/2} [\gamma_{\text{Ray}}(\Psi, \xi, G)]^d d\theta, \quad (20)$$

where

$$\gamma_{\text{Ray}}(\Psi, \xi, G) = \frac{1}{2m^2} \sum_{s \in \Omega} \sum_{\xi = 1}^m \delta_{\text{Ray}}(s, p, G) \quad (21)$$

and $z$ is also given as in (19).

For a given value of $N$, the comparison between the asymptotic performances of the systems operating over two fading channel models can be made by using the parameters $\delta(s, p, G)$ and $\delta_{\text{Ray}}(s, p, G)$. From (18) and the fact that $E_0(x) = \exp(-x)/x$, one has:

$$\delta(s, p, G) = \left( N \cdot \exp(Nz) \cdot E_1(N \cdot z) \right)^N > \left( N \cdot \exp(Nz/(Nz + 1)) \right)^N = \delta_{\text{Ray}}(s, p, G) \quad (22)$$

where the inequality follows from [10, 5.1.19] with $1/(x + 1) < \exp(x)/xE_1(x)$. The inequality in (23) shows that for a given value of $N$, the error performance of the system over a cascaded Rayleigh fading channel is always poorer than the performance over a conventional Rayleigh fading channel.

To see the effect of $N$ on the difference between the two parameters $\delta(s, p, G)$ and $\delta_{\text{Ray}}(s, p, G)$, Fig. 2 plots these two parameters for various values of $N$ when $z$ is normalized to be 1. Observe from Fig. 2 that when $N$ increases, both $\delta_{\text{Ray}}(s, p, G)$ and $\delta(s, p, G)$ becomes smaller. This implies that the performance gains in both cases can be increased by implementing SSD in a larger dimension. Furthermore, it can be seen from Fig. 2 that when $N$ is small, there is a significant gap between $\delta(s, p, G)$ and $\delta_{\text{Ray}}(s, p, G)$. Therefore, it is expected that there is a severe performance degradation of the system over a cascaded Rayleigh fading channel when compared to a Rayleigh fading channel. This analysis agrees with the results obtained in [7] for uncoded systems. A more interesting observation from Fig. 2 is that the performance gap becomes negligible when $N$ is sufficiently large. It means that by employing the technique of SSD with a high enough dimension, the performance of the system over a cascaded Rayleigh fading channel can closely approach that over a Rayleigh fading channel.

The ultimate performance limit of the system under consideration using SSD is now examined. In [3], by letting $N$ goes to infinity, it was shown that the parameter $\delta_{\text{Ray}}(s, p, G)$ approaches $\exp(-1/z) = \exp\left(-\frac{||(s - p)||^2}{4N_0 \sin^2 \theta}\right)$. Equivalently, one has:

$$f_{\text{Ray}}(d, \Psi, \xi, G) = \frac{1}{\pi} \int_0^{\pi/2} \left\{ \frac{1}{2m^2 \sum_{s \in \Omega} \sum_{\xi = 1}^m \exp\left(-\frac{||(s - p)||^2}{4N_0 \sin^2 \theta}\right)} \right\}^d d\theta \quad (24)$$

where the right-hand side of (24) serves as the tight bound of the function in (1) for BICM-ID over an AWGN channel [12]. A similar result can be obtained for the case of BICM-ID over a cascaded Rayleigh fading channel, which is stated in the following Lemma.

**Lemma 1:** The asymptotic error performance of BICM-ID with SSD over a cascaded Rayleigh fading channel approaches that over an AWGN channel.

**Proof:** Proving the above lemma is equivalent to show that $\delta(s, p, G) = \exp(-1/z)$ when $N$ goes to infinity. This is also equivalent to prove that

$$\log \delta(s, p, G) = -1/z \quad (25)$$

when $N$ goes to infinity. By multiplying both sizes of (25) with $z$ and consider the new variable $X = Nz$, it can be seen that one needs to show that:

$$\lim_{X \to \infty} X \cdot \log (X \cdot \exp(X) \cdot E_1(X)) = -1. \quad (26)$$
the identity in (26) can be proved by making use of an asymptotic expression of $X \cdot \exp(X) \cdot E_1(X)$ when $X$ is sufficiently large. The detailed proof can be found in [13].

VI. ILLUSTRATIVE RESULTS

In the interest of space, results are provided only for the case of QPSK modulation with two-dimensional Gray mapping $\xi$. A rate-1/2, 4-state convolutional code with generator matrix $g = (5, 7)$ is applied as the outer channel code, along with a random interleaver of length 1920 coded bits. To compute the union bound from (1) and (6), the first 20 Hamming distances of the convolutional code are retained. Unless otherwise stated, the optimal matrix $G_N$ is chosen from (16) with different values of $N = 2^r$.

**Fig. 3.** Performance of BICM-ID systems without SSD and with SSD when $N = 2$ over a cascaded Rayleigh fading channel.

Fig. 3 shows the error performance of the systems without SSD (and 1 iteration) and with SSD (and 1 and 4 iterations) over a cascaded Rayleigh channel, when $N = 2$ and the optimal MAP demodulator is used at the receiver. We observe no performance improvement with iterative processing when Gray mapping is used for the system without SSD. The analytical error bounds are also provided for comparison. It can be seen that in case of the system with SSD, the derived error bound is very tight. For the system without SSD, the bound under-estimates the simulation result. This is because the asymptotic union bound is tight only at low levels of BER. Though not shown here, it was observed that the bound for the system without SSD is tight at higher SNR where the BER level of around $10^{-3}$ is reached. As can be observed from Fig. 3, a significant performance gain can be achieved by using the SSD technique with $N = 2$ compared to the system without SSD. By comparing the tight error bounds over a wider range of SNR, it is observed that the coding gain is about 4dB at the BER of $10^{-4}$.

**Fig. 4.** Performance of BICM-ID systems with SSD when $N = 2$ over a cascaded Rayleigh fading channel employing various rotation matrices.

Besides $G_2$ in (16), we consider the following class of real rotations:

$$G_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$$

(27)

More specifically, two cases of $G_\theta$ with $\theta = \pi/4$ and $\theta = \arctan(2)/2$ are examined. The two rotation matrices are denoted by $G_{\pi/4}$ and $G_{\theta,2}$, respectively. It should be mentioned that $G_{\theta,2}$ is the optimal rotation matrix for uncoded systems in the sense that it maximizes the modulation diversity and the minimum product distance [2], [15]. However, its entries are not equal in magnitude. On the other hand, it can be verified that the rotation $G_{\pi/4}$ does not achieve the full modulation diversity, but this rotation belongs to the optimal class of rotation matrices discussed earlier. In Fig. 4, the error bounds for all systems are also provided. The performance with 1 iteration of the system without SSD is plotted as a reference.

As can be seen from Fig. 4, the error performances after 4 iterations with $G_2$ and $G_{\pi/4}$ are almost the same and converge to the same error bound, which confirms the analytical results. Furthermore, these performances are better than the error performance after 4 iterations and the error bound of the system employing $G_{\theta,2}$. By plotting the error bounds over a wider range of SNR, we observed that a gain of 0.3dB is achieved at the BER level of $10^{-4}$. For demonstration purpose, thereafter, the rotation matrix $G_N$ in (16) is always chosen.

As mentioned earlier, the complexity of the MAP demodulator increases exponentially with $N$, which makes the receiver implementation impractical for large $N$. It is shown in [3] that the suboptimal low-complexity MMSE demodulator can be a good alternative. Fig. 5 shows the BER performance with 1 and 4 iterations for $N = 2$ over conventional and cascaded Rayleigh fading channels. The MMSE demodulator is employed for both systems. It can be observed that the BER performances converge to the error bounds in both cases. Fig. 5 also confirms that, for small values of $N$, the cascaded
Rayleigh fading induces a severe performance degradation as compared to Rayleigh fading at any number of iterations.

Finally, to see the benefit of SSD at a sufficiently large value of $N$, Fig. 6 presents the performances after 1 and 8 iterations of the systems over the two fading channel models for $N = 32$ and with the MMSE demodulator. The analytical error bounds are also presented to show the tightness of the bounds. Furthermore, both the analytical bound and the simulation result of system performance with MAP demodulator over an AWGN channel [12] are provided to serve as performance benchmarks for the fading channels. Since Gray mapping is used, the error performance over an AWGN channel is obtained with only 1 iteration. It can be observed that with $N = 32$, the error performances after 8 iterations for both fading models tightly converges to the error bounds. When convergence starts to happen, as assumed by the analysis, there is only a small gap between the error performances of the two systems after 8 iterations and they both closely approach the performance of BICM-ID over an AWGN channel. For example, at the BER level of $10^{-4}$, the gaps between the error performance over an AWGN channel and those over conventional and cascaded Rayleigh fading channels are only 0.12dB and 0.3dB, respectively.

VII. Conclusions

This paper investigates the application of SSD in BICM-ID systems over cascaded Rayleigh fading channels. A tight bound on the asymptotic bit error probability was derived and used to obtain the optimal rotation matrix $G$. It was shown that when SSD is employed with a small dimensional constellation, although the technique of SSD is still useful, there is always a performance degradation compared to the conventional Rayleigh fading channel. When SSD is implemented with a sufficiently large constellation, the error performance over a cascaded Rayleigh fading channel can closely approach the performance of BICM-ID over an AWGN channel, which also serves as the performance limit of BICM-ID with SSD over a conventional Rayleigh fading channel.

References