Abstract—Using the random coding argument, we derive in this paper the effective cutoff rate, $R_e$, of multiple-input-multiple-output (MIMO) space-time (ST) codes in Rayleigh fast fading channels with imperfect channel state information at the receiver (CSIR). In contrast to conventional cutoff rate analysis where the codeword length $N$ is infinity and the frame error probability $P_f$ is implicitly zero, we loosen the definition to include the more realistic case of finite $N$ and finite $P_f$. Using the tight upper and lower bounds on the pair-wise error probability (PEP) in [18], we are able to derive in turn tight lower and upper bounds on the cutoff rate of MIMO ST coding systems. The results are very general and can be applied to any linear modulation scheme, like MPSK and MQAM. Numerically we found that for the 2 transmit and 2 receive antenna configuration, the non-asymptotic segments of our cutoff rate upperbounds for 4, 16, and 64 MQAM actually coincide with the ergodic capacity curve. Furthermore, we found that the performance of the Smart-Greedy codes proposed in the seminal work in [13] falls within the range predicted by our cutoff rate bounds. Finally, we show in the paper that for MIMO systems employing pilot-symbol assisted channel estimation, the asymptotic cutoff rate no longer increases linearly with the number of transmit antennas. Instead, for a normalized fade rate of $f_o$ and a signal constellation of size $M$, the maximum effective cutoff rate is $(8f_o)^{-1}\log M$ and this is achieved when the number of transmit antennas is the integer closest to $1/(4f_o)$. The result suggests that with a large antenna array and high user mobility, a more bandwidth efficient channel estimation strategy is desired.

Index Terms—MIMO systems, cutoff rate, space-time codes, pilot-symbol assisted modulation, Rayleigh fast fading channel

Copyright © 2009 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permission@ieee.org.

Manuscript received April 21, 2008; revised August 24, 2008. The paper was presented in parts at the IEEE MILCOM 08, 17-19 Nov 2008, San Diego. This work was supported in part by the Singapore MoE AcRF Tier 2 Grant T206B2101.

Paul Ho is with the School of Engineering Science Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada (Tel: 778-782-3822; fax: 778-782-4951; e-mail: paul@cs.sfu.ca).

Kar Peo Yar is with the Institute for Infocomm Research, (A*STAR) Agency for Science, Technology and Research, Fusionopolis, Singapore 138632 (e-mail: kpyar@i2r.a-star.edu.sg)

Pooi Yuen Kam is with the Department of Electrical and Computer Engineering National University of Singapore, Singapore 117576 (e-mail: py.kam@nus.edu.sg)

I. INTRODUCTION

MULTIPLE-INPUT-MULTIPLE-OUTPUT (MIMO) is a technology that uses multiple antennas at the transmitter and the receiver to achieve space-time (ST) diversity, spatial multiplexing, or beamforming (precoding) in wireless communication channels. As demonstrated in [1]-[3], MIMO can provide dramatic capacity improvement over conventional single-input-single-output (SISO) transmission in a fading environment without increasing transmission bandwidth. Since the report of these seminal works, there is an explosion of research activities in the field, many of which address MIMO from an information theoretic perspective, namely, determining the capacity of MIMO under different modes and operating conditions; see for example the research in [4]-[10]. While much insights into the design of practical MIMO systems can be gained from these performance limits, there are also some limitations associated with capacity analysis. Firstly, with imperfect CSI, the capacity of MIMO in fading channels is still not known precisely. Only bounds are available; see for example [11] and [12]. Because of fading, noise, and limited bandwidth available for channel sounding and feedback, one would expect that the CSI estimates available at the transmitter and/or receiver to be imperfect all the times. Secondly, even when the CSI is perfect, the capacity-achieving modulation (usually) follows a Gaussian distribution, which is challenging to implement in practice. In actual deployment of MIMO, fixed constellations will always be used. Thirdly, in any capacity analysis, the block length of the code is implicitly taken to be infinity. In actual systems though, the codeword length is always finite. For example, in the seminal work by Tarokh et.al. [13], the authors focus on the frame error rates of ST trellis codes with a frame size of 130 symbols.

Because of the aforementioned issues, we consider the adoption of a cutoff rate-like parameter, $R_e$, as an alternative measure of the performance limit of MIMO with finite codeword length and fixed frame error probability (FEP). To the best of our knowledge, there are only a limited number of cutoff rate analyses for MIMO systems in the literature; see for example [14]-[15]. Furthermore, given that the modulation is fixed in any cutoff rate analysis, the incorporation of imperfect CSI in the analysis is straightforward; see for example the analyses in [16]-[17] for SISO systems. In
contrast, the capacity of MIMO with imperfect CSI is still not known. The cutoff rate that we derive is based on the new bounds for the pairwise-error probability (PEP) of ST codes in [18]. We prefer these bounds over the exact PEP [19] in the cutoff rate analysis because they are tight, simple, and intuitive. More importantly, the results are implicitly functions of the codeword length $N$.

The paper is organized as follows. The MIMO system model and the corresponding upper and lower bounds on the PEP are summarized in Section II. Based on these PEP bounds, upper and lower bounds on the effective cutoff rate, as functions of the codeword length and fixed PEP, are derived in Section III. These bounds on the effective cutoff rate are expressed in analytical as well as integral forms, with the latter more suitable for long codeword lengths. An application of the cutoff rate analysis to pilot-symbol assisted modulation (PSAM) is provided in Section IV, where we highlight the potential rate loss caused by a lowering of the effective channel signal-to-noise ratio because of imperfect channel estimation and by the insertion of pilot symbols. Numerical results for different MIMO configurations and CSI assumptions are provided in Section V. Comparisons with the cutoff rates obtained from the Chernoff bound on the PEP and with capacity are made when applicable. Finally conclusions of this investigation are provided in Section VI.

II. SYSTEM MODEL AND PAIRWISE ERROR PROBABILITY

We consider in this paper a space-time (ST) coded MIMO system using $T$ antennas to transmit modulation symbols from multiple users over a Rayleigh fading channel using time division multiplexing (TDM) [18]. The symbol transmitted by Antenna $i$ of any user during the $k$-th TDM frame is denoted by $v_i(k)$, and is taken either from an unit-energy $M$-ary PSK (MPSK) constellation $S_{\text{MPSK}} = \{ \exp(j2\pi m/M) \}_{m=0}^{M-1}$, or from an unit-energy $M$-ary Quadrature Amplitude Modulation (MQAM) constellation

$$ S_{\text{MQAM}} = \left\{ \left( 2m - 1 - \sqrt{M} \right) \mu - j \left( 2n - 1 - \sqrt{M} \right) \mu \right\}_{m,n=1}^{\sqrt{M}} , $$

$\mu = \sqrt{1.5/(M-1)}$. Each receiver is equipped with $R$ receive antennas and the fading gains in the $T \times R$ links between the transmitter and receiver are independent and identically distributed (iid) zero mean complex Gaussian processes with variance $\sigma_i^2$ in both the real and imaginary components. Mathematically, the $k$-th transmitted ST codeword segment $v(k) = (v_1(k), v_2(k), ..., v_N(k))^T$, $k = 1, 2, ..., N$, of each user is related to its received version $r(k) = (r_1(k), r_2(k), ..., r_R(k))^T$ according to

$$ r(k) = \sqrt{E_s} H(k)v(k) + n(k) , \quad (1) $$

where $N$ is the codeword length, $\sqrt{E_s}$ is the average symbol energy, $H(k) = [h_{ij}(k)]$ is the channel fading matrix, where for any given $k$, $h_{ij}(k)$’s are all iid and $\mathcal{CN}(0,2\sigma_i^2)$, and $n(k)$ is an additive white Gaussian noise vector and is $\mathcal{CN}(0,N_0I_R)$ distributed. Since we assume that the TDM frame is longer than the coherent time of the channel, different codeword segments of the same user experience independent fading.

In [18], the authors derived upper and lower bounds on the PEP for the ST coded system in (1), with maximum likelihood (ML) detection and pilot symbol assisted channel estimation. Specifically, for an erroneous codeword $\{\hat{v}(k)\}_{k=1}^N$ having distances $\{d_{jk}^2; j = 1, 2, ..., T\}_{j=1}^T$ with the transmitted codeword $\{v(k)\}_{k=1}^N$, the PEP, $P_2$, is upper and lower bounded as follows [18, (9) and (11)]:

$$ \frac{(2q-1)!}{2(2q)!!} \prod_{j=1}^{2(2q-1)} \left( 1 + \frac{E_s \sum_{j=1}^{2q} \hat{\sigma}_j^2 d_{jk}^2}{2(N_0 + E_s \sum_{j=1}^{2q} \hat{\sigma}_j^2)} \right)^{-R} \leq P_2 , \quad (2) $$

where $N$ is the codeword length, $\kappa$ is the set of symbol positions where the two codewords are different, $q = |\kappa|$, $R$ is the total diversity order, $2q!! = 2q \cdot (2q - 2) \cdot ... \cdot 2$, $(2q-1)! = (2q-1) \cdot (2q-3) \cdot ... \cdot 1$, $2\sigma_j^2$, $j = 1, 2, ..., T$, the variances of the channel estimates, and $2\hat{\sigma}_j^2$, $j = 1, 2, ..., T$, the variances of the estimation errors. Since we assume identical links, $2\sigma_j^2 = 2\sigma_k^2$ and $2\hat{\sigma}_j^2 = 2\hat{\sigma}_k^2$; $j = 1, 2, ..., T$. Furthermore, by defining the transmit SNR to be $\Gamma = N_0 / E_s$ and the effective signal-to-noise ratio (SNR) to be

$$ \gamma = \frac{2E_s \Gamma \sigma_i^2}{N_0 + 2E_s \Gamma \sigma_i^2} = \left( \frac{2\sigma_i^2}{1 + 2\sigma_i^2} \right) \Gamma , \quad (3) $$

the upper and lower bounds in (2) can be simplified to

$$ P_{\text{UB}} \leq \frac{(2q-1)!}{2(2q)!!} \prod_{j=1}^{2q} \left( 1 + \gamma / 4T \sum_{j=1}^{2q} d_{jk}^2 \right)^{-R} \quad (4) $$

and

$$ P_{\text{LB}} \geq \frac{(2q-1)!}{2(2q)!!} \prod_{j=1}^{2q} \left( 1 + \gamma / 4T \sum_{j=1}^{2q} d_{jk}^2 \right)^{-R} \quad (5) $$

respectively. It is evident from these equations that the upper and lower bounds are tight at large SNR, as $\gamma /(4T) \sum_{j=1}^{2q} d_{jk}^2 \approx 1 + \gamma /(4T) \sum_{j=1}^{2q} d_{jk}^2$ when $\gamma \to \infty$. 

VT-2008-00347.R1 – accepted Jan 27, 2009
III. CUTOFF RATE

As stated earlier, we are interested in determining the cut-off rate of the MIMO system described in the previous section, under the conditions of a finite codeword length \( N \) and a finite targeted FEP \( P_f \). Since traditionally, the cutoff rate, \( R_0 \), is defined only for systems with infinite \( N \) and implicitly zero \( P_f \), our result is thus not exactly the cutoff rate in the conventional sense. In the absence of a better terminology, we simply refer to it as the effective cutoff rate \( R_e \).

Consider first the PEP \( P_2 \) in (2). The average of this parameter, over all random choices of codeword pairs, can be written as

\[
E[P_2] = 2^{-R \cdot N}, \quad (6)
\]

where \( E[\cdot] \) is the expectation operator and

\[
R_e = -\frac{1}{N} \log_2 \left( E[P_2] \right) \quad (7)
\]

is the error exponent of the average PEP. It follows that the average word/frame error probability takes on the form [21]

\[
E[P_s^f] = 2^{N \cdot E[P_2]} = 2^{-(R_e - R) \cdot N}, \quad (8)
\]

where \( R \) is the actual coding rate (in bits/symbol) and \( E(R, N) - R_e - R \) is the corresponding error exponent [22]. If \( N \) is infinite, then any coding rate up to \( R_e \) will yield a zero \( E[P_s^f] \). Consequently

\[
R_0 = \lim_{N \to \infty} R_e
\]

is the conventional definition of the cutoff rate [20]. On the other hand when \( N \) is finite, there is no guarantee that \( E[P_s^f] = 0 \) even when the actual coding rate \( R \) is less than \( R_e \). This means the error exponent \( R_e \) alone cannot be regarded as a rate limit. However, it can be converted into one through the following 2-step procedure.

As pointed out in [13], coded systems are typically designed based on a targeted frame FEP, \( P_f \). By setting the right hand side (RHS) of (8) to be less than \( P_f \), we can see that this FEP target is met when the transmission rate \( R \) is less than

\[
R_e' = R_e - \varepsilon; \quad \varepsilon = -\frac{1}{N} \log_2 \left( \frac{P_f}{P_s^f} \right) \quad (9)
\]

where \( \varepsilon \) represents the rate reduction due to the finite word length effect. We list in Table 1 the value of this reduction for various combinations of \( N \) and \( P_f \). As an example, with a relatively stringent requirement of \( P_f = 10^{-3} \), the rate reduction is 0.1557 bits/symbol for a short codeword length of 64. To put this rate loss in perspective, we compare it against the asymptotic value of \( R_c \). As shown later in the section, at large SNR, \( R_c \) approaches \( T \cdot \log_2(M) + 1/N \). This means for \( T = 2 \) and \( M = 2 \), the rate loss mentioned above is 7.8% of \( R_c \). The loss as a percentage becomes even less significant when more antennas and/or a larger signal constellation are used. Note also that when \( N \to \infty \), the absolute value of the rate loss itself approaches zero for any practical FEP of interest. Consequently, \( R_e' \) approaches \( R_c \), which itself approaches the true cutoff rate \( R_0 \).

<table>
<thead>
<tr>
<th>Word Length ( N )</th>
<th>Targeted Frame Error Probability ( P_f )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( 10^{-1} )</td>
</tr>
<tr>
<td>16</td>
<td>0.2076</td>
</tr>
<tr>
<td>64</td>
<td>0.0519</td>
</tr>
<tr>
<td>246</td>
<td>0.0130</td>
</tr>
<tr>
<td>1024</td>
<td>0.0032</td>
</tr>
</tbody>
</table>

The second and final step required to convert \( R_c \) into a meaningful rate limit is the quantization of \( R_e' \). Assume an \( M \)-ary modulation over each of the \( T \) transmit antennas. Then there are altogether \( M^{TN} \) code patterns and we can choose \( I \) out of these \( M^{TN} \) patterns to form a code of rate \( R = \log_2(I)/N \) bits/symbol, \( 1 < I \leq M^{TN} \). For example, when \( M = 2 \), \( T = 1 \), \( N = 3 \), there are altogether 8 binary patterns of length 3, and the only possible code rates are 0.33, 0.53, 0.67, 0.77, 0.86, 0.94, and 1. With this observation in mind, it becomes clear that for finite \( N \), the maximum transmission rate in (9) should be modified to

\[
R_e = \frac{1}{N} \log_2 \left( 2^{N \cdot R_e'} \right), \quad \text{finite } N, \text{ non-zero FEP} \quad (10)
\]

where \( [\cdot] \) is the floor operator. We will refer to (10) as the effective cut-off rate, as it take into account all the finite word length effect. We next proceed to determine lower and upper bounds on \( R_e \) in (7), which in turn will be used to obtain the corresponding bounds on \( R_e \).

First consider the upper bound on the PEP in (4). Using a random coding argument, the expected value of this bound is

\[
E[P_{ub}] = \sum_{k=0}^{N} \Pr \left[ q = kR \right] E[P_{ub} | q = kR] = \frac{p^N}{2} \left[ 1 + \sum_{j=1}^{N} \left( \frac{N}{k} \right) \frac{1}{(1-p)^j} (2kR-1)!! (2kR)! \alpha^j \right] \quad (11)
\]

where

\[
p = M^{-T} \quad (12)
\]

is the probability that \( \sum_{j=1}^{T} d_j(k) = 0 \), and

\[
\alpha = E \left[ \left( \frac{\gamma}{4T} \sum_{j=1}^{T} d_j(k) \right)^R \left( \sum_{j=1}^{T} d_j(k) \neq 0 \right) \right] \quad (13)
\]
In the Appendix, we provide analytical expressions for the term $\alpha$ for MPSK and MQAM modulations; refer to (A3), (A6), and (A11). Now, since the "$N$ chooses $k$" function in (11) increases rapidly with $N$, calculating the sum in (11) directly is problematic when $N$ is large (e.g. $N > 170$) in Matlab 6 running on a Intel Core 2 CPU based PC. A more "robust" approach is to view the sum in (11) as the $z$-transform of the product of two signals in the time domain. Specifically, let us define

$$a[n] = \begin{cases} \frac{p^N}{2^n} \left( \frac{N}{n/R} \right), & n = 0, R, 2R, ..., NR, \\ 0, & \text{otherwise}, \end{cases}$$

$$A(v) = \sum_{n=0}^{NR} a[n] v^{-n} = \frac{1}{2} p^N (1 + v^{-R})^N,$$

$$b[n] = \frac{(2n-1)!!}{(2n)!!}, \quad n = 0, 1, ..., ;$$

$$B(v) = \sum_{n=0}^{N} b[n] v^{-n} = \frac{1}{\sqrt{1-v^{-1}}}, \quad |v| > 1,$$

and

$$z^{-1} = \left( \frac{1}{p} - 1 \right)^{v/R} = \left( (M^T - 1) \alpha \right)^{1/R},$$

where the transform in (15) is taken from [23]. This means (11) can be written as $E[P_{b,h}] = \sum_{n=0}^{N} a[n] b[n] z^{-n}$, which is simply the $z$-transform, $C(z)$, of the product signal $c[n] = a[n] b[n]$. This transform can be written as a contour integral

$$E[P_{b,h}] = C(z) = \frac{1}{2\pi j} \oint_{C} A(v) B(z/v) v^{-1} dv = \frac{1}{2\pi j} \oint_{C} p^N (1 + v^{-R})^N \frac{1}{\sqrt{1-v^{-1}}} v^{-1} dv,$$

where the contour must lie within the region of convergence of $B(z/v)$, i.e. $|v| < z$. Once $E[P_{b,h}]$ is obtained, it can be readily be used in place of $E[P_{c,h}]$ in (7) to obtain a lower bound on $R_e$. The latter bound is further used in (10) to provide a lower bound on the effective cutoff rate $R_e$.

Similarly, we can express the average of the lower bound on PEP (5) over random codes in integral form as

$$E[P_{b,h}] = \frac{1}{2\pi j} \oint_{C} p^N (1 + v^{-R})^N \frac{1}{\sqrt{1-v^{-1}/Z}} v^{-1} dv, \quad |v| < Z,$$

where

$$Z^{-1} = \left( \frac{1}{p} - 1 \right)^{v/R} \beta,$$

$$\beta = E\left[ \left( 1 + \frac{\gamma}{4T \sum_{j=1}^{T} d_j^2(k)} \right)^{-\frac{R}{2}} \sum_{j=1}^{T} d_j^2(k) \neq 0 \right].$$

The reader is again referred to the Appendix for the analytical expressions of $\beta$ for various modulations. Finally, an upper bound on $R_e$ can be obtained by first replacing the average PEP in (7) by $E[P_{b,h}]$, and then substitute the resultant upper bound on $R_e$ into (9) and (10).

IV. PILOT-SYMBOL ASSISTED MODULATION AND RATE REDUCTION

The cutoff rate analysis presented in the previous sections is based on pilot-symbol assisted channel estimation, where imperfection in the channel estimates is reflected in the effective SNR $\gamma$ obtained from the variances of the CSI $2\sigma^2_k$ and the estimation error $2\sigma^2_e$; refer to (3). The pilot-assisted estimator itself can operate based on the least square [19] or the minimum mean square error (MMSE) principle [24]. In the following, we derive the effective SNR of a MMSE-based MIMO channel estimator similar to the one used in the SISO pilot-symbol assisted modulation (PSAM) system in [24]. Channel estimation in the system is performed prior to de-interleaving and this requires the transmitted data stream be organized into frames of $K$ symbols, $T$ of which are pilot, and the rest being data. The pilot insertion frequency, $1/K$, must be at least $2f_p$, where $f_p$ is the Doppler frequency of the fading channel normalized by the symbol rate. For simplicity, we assume the $T$ pilot symbols in each frame take on the values of $\sqrt{T}u_1, \sqrt{T}u_2, ..., \sqrt{T}u_T$, where $u_i$ is a unit-basis vector with its $i$-th component equal to unity and zero otherwise. In other word, the transmitter only engages one transmit antenna at a time during the pilot interval, but each antenna is to transmit at $T$ times the power level it normally does during data transmission. This pilot-encoding schemes ensures proper energy normalization across the pilot and data transmission phases, and more importantly it guarantees that the pilot symbols arriving at the receiver from different transmit antennas are separable. The $T$ pilot symbols within each frame are transmitted in succession and the fading variation within the each $T$-length pilot interval is assumed negligible. Once channel estimates at these pilot insertion instants are obtained, they are fed to a bank of interpolation filters for generating the channel estimates at the data positions. Wiener filters designed based on a Jakes power spectrum and the normalized Doppler frequency $f_p$ are used for this purpose. For convenience, we use the mean-square value of the channel estimate and mean-square estimation error at the mid-frame position as representative values of $2\sigma^2_k$ and $2\sigma^2_e$ respectively. In practice, there will be a small degree of variation in the CSI quality across a frame (in the order of $1/4$
As demonstrated in [24] for PSAM, \( 2\sigma_0^2 = 2\sigma_h^2 \cdot \rho^2 \), \( 2\sigma_h^2 = 2\sigma_h^2 \cdot (1 - \rho^2) \), where \( \rho \) is the correlation coefficient between \( h_y \) and \( \hat{h}_y \) (when the estimator input SNR is \( \Gamma = \frac{TE_y}{N_0} \)). Consequently, the effective SNR in (3) can be rewritten as
\[
\gamma = \frac{\rho^2 \cdot \Gamma}{1 + (1 - \rho^2) \cdot \Gamma}. \quad \text{(PSAM-MIMO)} \quad (21)
\]
The reader is refer to [24] for an expression of \( \rho \). All the numerical results presented in the next section on PSAM-MIMO were generated using the above definition of SNR. Furthermore, the rate \( R_c \) in (10) is multiplied by the factor
\[
\eta = 1 - \frac{T}{K} \quad (22)
\]
to reflect the drop in throughput due to the insertion of pilot symbols.

The multiplication of the rate in (10) by the above factor has an interesting implication on the maximum rate that MIMO-PSAM can achieve. Consider the following asymptotic analysis. When \( \Gamma \to \infty \), then \( \rho \to 1 \), \( \gamma = \Gamma \), and the parameters \( \alpha \) and \( \beta \) in (13) and (20) both equal to zero. As a consequence, both the upper and lower bounds on the PEP equal \( \rho^4 \cdot 2 \), and (10) approaches \(-\log \rho = T \log M \) when \( N \) is large. Multiplying \( T \log_2 M \) by the factor in (22) allows us to obtain the maximum effective rate of MIMO-PSAM as
\[
R_c^\text{EPSAM} = \frac{T \log_2 M}{\log_2 T} \quad (23)
\]
By treating the number of transmit antennas as a continuous variable, (23) becomes a quadratic function in \( T \) and is largest at \( T = K/2 \). At this value of \( T \), \( R_c^\text{EPSAM} = K \cdot \log_2 M / 4 \).

Since \( 1/K \geq 2f_D \) because of the Nyquist sampling requirement in PSAM, this means the effective rate of MIMO-PSAM is always upperbound by
\[
R_c^\text{PSAM} = \log_2 M / (8f_D) \quad (24)
\]
independent of the number of transmit antenna. In comparison, the capacity and the asymptotic cutoff rate of MIMO with perfect CSI increase indefinitely with the number of transmit antennas. This implies that PSAM is not an efficient channel estimation strategy for high mobility applications involving a large antenna array. For low mobility and a small number of antennas though, we may still get a substantial increase in throughput by adding an extra antenna. The reader is further reminded that in PSAM, we are trading transmission efficiency for a low implementation complexity. If we are willing to scarify implementation complexity instead, it is possible to perform blind channel estimation without resorting to pilot symbols, allowing the rate to (theoretically) increase with the number of transmit antennas. However, these blind channel estimators usually iterate between data detection and channel estimation and are thus computationally intensive.

Furthermore, any erroneous decision on the data may severely degrade the quality of the subsequence channel estimates. The result is a drop in the effective SNR \( \gamma \), and hence a lowering of the maximum achievable rate from the theoretical limit.

V. NUMERICAL RESULTS

We present in this section numerical results for the effective cutoff rate, \( R_c \), of ST coded MIMO systems operating in fast fading channels. Both MPSK and MQAM modulations are considered. The cutoff rates of these MIMO systems are plotted against the transmit SNR, \( \Gamma \). The received SNR, \( \gamma \), used in the numerical calculations is obtained under the condition that \( 2\sigma_h^2 = 1 \). As can be seen from (3), when there is no channel estimation error, i.e. when \( 2\sigma_h^2 = 2\sigma_0^2 \) and \( 2\sigma_h^2 = 0 \), then \( \gamma = \Gamma \). We would also like to point out that the numerical integrations in (17) and (18) were carried out using circular contours of radius 0.99\( \pi \) and 0.99Z respectively. These choices provide numerical stability over all \( N \) and SNR considered. In general, the smaller \( N \) is, the less stringent it is in choosing the contour’s radius.

Included in all the figures of this section are the ergodic capacity curves of MIMO with perfect CSIR. These curves were obtained from the simulation of [1]
\[
C = E_h \left[ \log_2 \left( I_r + \frac{\gamma}{T} H H^T \right) \right], \quad (25)
\]

where \( H \) is the channel matrix in (1), \( I_r \) an identity matrix of size \( R \), \( || \) and \( (\cdot)^T \) denote respectively the determinant and conjugate transpose of a matrix. We caution the reader that when comparing the cutoff rate results against capacity, one has to bear in mind that the former is based on a fixed FEP while the latter implicitly assumes zero error probability.

A. SISO-BPSK

We first consider the cutoff rate of the simplest system: single input single output, BPSK signaling and perfect channel estimation. Our upper and lower bounds on \( R_c \) for \( N = 16 \) and 1024 (“short” and “long” codewords”) are shown in Fig. 1 for a targeted FEP of \( P_f = 10^{-3} \). Also shown in the figure is the Chernoff lower bound on \( R_c \), obtained by substituting the result from [15] below
\[
R_c = 1 - \log_2 \left( 1 + \frac{1}{1 + \gamma} \right); \quad \text{Chernoff-based SISO-BPSK} \quad (26)
\]
into (9) and (10). It is clear from the figures that our bounds are much tighter than the Chernoff bound when \( N \) is small. For larger \( N \) though, like \( N=1024 \), the Chernoff lower bound almost coincides with or equals our upper bound at practical channel SNR of interest. So both are tight! At a rate of 0.5 bits/symbol, the gap between the cutoff rate of the \( N=1024 \) dB).
system and capacity is approximately 1.5 dB. The gap is widened to 5 dB when the rate is 0.8 bits/symbol.

Fig. 1: Our upper/lower bounds and the Chernoff lower bound on the effective cutoff rate for SISO-BPSK at a FEP of $P_f = 10^{-3}$ and perfect CSIR.

Fig. 2 plots our upper bound on $R_c$ with $N$ as a parameter. It is evident from the figure that the cutoff rate increases quickly with $N$. At large enough $N$, the rate loss due to finite codeword length becomes insignificant.

Fig. 2: Upper bounds on the effective cutoff rate for SISO-BPSK with different codeword lengths at a FEP of $P_f = 10^{-3}$ and perfect CSIR.

B. Effect of MIMO Configuration

We next consider the cases of 2 transmit antennas with first and second order receive diversity. These configurations are denoted as $2 \times 1$ and $2 \times 2$ MIMO respectively. The upperbounds on the cutoff rate for the $2 \times 2$ configuration with QPSK modulation, perfect CSIR, and a FEP of $P_f = 10^{-3}$ are shown in Fig. 3 with the codeword length $N$ as a parameter. For rate less than 2.5 bits/symbol, these cutoff rate curves run parallel to the capacity curve. In particular, when $N = 1024$, the cutoff rate upperbound and the capacity curve are almost indistinguishable. On the other hand when $N = 16$, the difference between capacity and cutoff rate is approximately 2 dB in SNR.

Fig. 3: Effective cutoff rate of MIMO-QPSK with 2 transmit and 2 receive antennas, perfect CSIR, and a FEP of $P_f = 10^{-3}$.

Focusing on a codeword length of $N = 128$, we show in Fig. 4 the capacity and cutoff rates of $2 \times 1$ and $2 \times 2$ MIMO-QPSK with perfect CSIR and $P_f = 10^{-3}$. Also shown are performance of the smart-greedy codes from [13, pp. 763] with the same FEP and approximately the same $N$. It is observed that in the $2 \times 2$ case, the performance of the smart-greedy code lies within our upper and lower bounds on the effective cutoff rate. In the $2 \times 1$ case though, the smart-greedy code is 2.5 dB worse than that predicted by our lower bound. We do not consider this 2.5 dB deviation a discrepancy though, it just alludes to the fact that this simple 2-state $2 \times 1$ Smart-Greedy is not “optimal” and more efficient codes do exist and awaited to be found.

Fig. 4: Effective cutoff rate of MIMO-QPSK with 2 transmit antennas, a codeword length of $N = 128$, perfect CSIR, and a FEP of $P_f = 10^{-3}$. 
C. Effect of Constellation Size

We show in Fig. 5 the cutoff rates of $2 \times 2$ MIMO with 4QAM, 16QAM, and 64QAM codes of length $N = 64$ and perfect CSIR. It is observed that the non-asymptotic segments of the cutoff rate upperbounds superimpose on one another and coincide with the capacity curve over ranges of SNRs. While the superposition suggests that 64QAM is always the best choice among the three modulations, the intersection of the lowerbounds hints that a smaller constellation is favored over a larger constellation at lower SNRs. Specifically when the SNR is below 12 dB, then 4QAM is preferred over the two other modulations. On the other hand when the SNR is above 19 dB, then 64QAM should be used. Finally, when the SNR is between the two SNR thresholds, 16QAM is the best choice.

D. Effect of Imperfect CSIR

The numerical results provided in the last three sections are for the ideal case of perfect CSI at the receiver. We consider next the effect of imperfect channel estimation using PSAM with interpolators of size 12 and pilot frames of size $1/(2.5 f_D)$. Fig. 6 shows the effects a normalized Doppler frequency of $f_D = 0.03$ has on the $2 \times 2$ MIMO-MQAM systems in Fig. 5. In comparing the two figures, we notice that asymptotically, there is a 10-12% drop in rate because of imperfect CSI and the insertion of pilot symbols. In the non-asymptotic operating regions, the gap between perfect (Fig. 5) and imperfect CSI (Fig. 6) is in the range of 3 to 5 dB.

Finally in Fig. 7, we vary the normalized Doppler frequency as a parameter.

VI. CONCLUSION

As opposed to most of the information-theoretic work on MIMO which focus on the system capacity, we consider in this investigation an alternative measure, namely, the cutoff rate of the system. The main rationales of adopting the cut off rate as the performance limit are (1) in any practical system, the modulation is fixed, and (2) incorporating imperfect CSI in the analysis is straightforward. In terms of cutoff rate analysis, our work deviates from those in the literature through loosening of the cutoff rate definition to include finite codeword length and finite frame error probability. Through using the tight upper and lower bounds on the pair-wise error probability (PEP) in [18], we are able to derive in turn tight lower and upper bounds on the cutoff rate of MIMO ST coding systems. Specifically, we found that segments of the
cutoff rate upperbounds of MIMO-MQAM coincide with the ergodic capacity curve. Furthermore, we found that the performance of the Smart-Greedy codes proposed by Tarokh et al. [13] lies within the range predicted by our cutoff rate bounds. Finally, we show in the paper that for MIMO systems employing pilot-symbol assisted channel estimation, the asymptotic cutoff rate no longer increases indefinitely with the number of transmit antennas. Instead, for a normalized fade rate of $f_d$ and a signal constellation of size $M$, the maximum effective cutoff rate is $\log_2 M/(8 f_d)$ and this is achieved when the number of transmit antennas is the integer to closest to $1/(4 f_d)$. The result suggests that with a large antenna array and high user mobility, a more bandwidth efficient channel estimation strategy should be used instead.

**APPENDIX**

We derive in this appendix the terms $\alpha$ and $\beta$ in (13) and (20) respectively. Consider first

$$\Delta = \sum_{j=1}^{T} d^2_j(k), \quad (A1)$$

the square Euclidean distance between the a pair of coded symbols. With BPSK signaling and the random coding argument, each $d^2_j(k)$ is a binary random variable with a probability density function (pdf) of

$$p_{d^2_j}(x) = \frac{1}{2} \left( \delta(x) + \delta(x-4) \right), \quad \text{where } \delta(x) \text{ is the discrete-time unit impulse function.}$$

The characteristic function (CF) of $d^2_j(k)$ can be described in term of the z-transform of its pdf, yielding $P_{d^2_j}(z) = \frac{1}{2} (1 + z^{-4})$. Since $\Delta$ is the sum of $T$ number of iid $d^2_j(k)$ s, its CF is simply

$$P_{\Delta}(z) = 2^{-T} \left( 1 + z^{-4} \right)^T. \quad \text{In other word, the parameter } \Delta \text{ is binomial distributed with a pdf}$$

$$p_{\Delta}(x) = \frac{1}{2^T} \sum_{t=0}^{T} \binom{T}{t} \delta(x-4t). \quad (BPSK) \quad (A2)$$

Since $\alpha$ and $\beta$ are the averages of $(\gamma \Delta / 4)^R$ and $(1 + \gamma \Delta / 4)^{-R}$ given that $\Delta \neq 0$, respectively, we can deduce from (A2) that

$$\alpha = \frac{1}{2^T-1} \sum_{t=0}^{T} \binom{T}{t} \left( \frac{\gamma t}{T} \right)^{-R}, \quad (BPSK) \quad (A3)$$

and

$$\beta = \frac{1}{2^T-1} \sum_{t=0}^{T} \binom{T}{t} \left( 1 + \frac{\gamma t}{T} \right)^{-R}. \quad (BPSK) \quad (A4)$$

Next for QPSK, the random coding argument leads to a pdf of

$$p_{d^2_j}(x) = \frac{1}{4} \left( \delta(x) + 2 \delta(x-2) + \delta(x-4) \right) \text{ and a CF of}$$

$$P_{d^2_j}(z) = \frac{1}{4} \left( 1 + z^{-2} \right)^2 \text{ for } d^2_j(k). \text{ The CF of } \Delta \text{ is thus}$$

$$P_{\Delta}(z) = 2^{-2T} \left( 1 + z^{-2} \right)^{2T}. \text{ Taking the inverse } z \text{ transform of } P_{\Delta}(z) \text{ yields the pdf of }$$

$$p_{\Delta}(x) = \frac{1}{2^{2T}} \sum_{t=0}^{2T} \binom{2T}{t} \delta(x-2t) \quad \text{(QPSK)} \quad (A5)$$

As a result, the average values of $(\gamma \Delta / 4)^R$ and $(1 + \gamma \Delta / 4)^{-R}$, given that $\Delta \neq 0$, are respectively

$$\alpha = \frac{1}{2^{2T-1}} \sum_{t=0}^{2T} \binom{2T}{t} \left( \frac{\gamma t}{2T} \right)^{-R}, \quad \text{(QPSK)} \quad (A6)$$

and

$$\beta = \frac{1}{2^{2T-1}} \sum_{t=0}^{2T} \binom{2T}{t} \left( 1 + \frac{\gamma t}{2T} \right)^{-R}. \quad \text{(QPSK)} \quad (A7)$$

In comparing (A6)-(A7) with (A3)-(A4), we conclude that the $\alpha$ and $\beta$ parameters for QPSK are those of BPSK with twice as many transmit antennas.

How about MQAM? This signal constellation is the Cartesian product of two identical $Q$-ary PAM constellations $\{ A_i \}_{i=1}^{Q}$ and $\{ B_j \}_{j=1}^{Q}$, where $Q = \sqrt{M}$. The $i$-th signal levels in the two $Q$-ary PAM constellations are $A_i = B_i = (2i-1-Q)\mu, \ i = 1,2,\ldots,Q$. For the MQAM constellation to have unit average energy, it is required that $2E[A_i^2] = 1$, or $\mu = \sqrt{1.5/(M-1)}$. Since the square distance between the $i$-th and $j$-th signal levels in $\{ A_i \}_{i=1}^{Q}$ is $4\mu^2(i-j)^2$, we can deduce that the square distance between symbol pairs drawn randomly from the constellation, after normalized by $(2\mu)^2$, has a pdf of

$$p_{\Delta}(x) = \frac{1}{Q^2} \left[ (Q-2) \delta(x) + 2 \sum_{d=1}^{Q-1} (Q-d) \cdot \delta(x-d^2) \right] \quad (A8)$$

and a CF of

$$P_{\Delta}(z) = \frac{1}{Q^2} \left[ Q + 2 \sum_{d=1}^{Q-1} (Q-d)z^{-d^2} \right] = \sum_{k=0}^{Q^2-1} c_k z^{-k}, \quad (A9)$$

where the $a_k$ s can be obtained by equating coefficients with the same powers of $z$ on both sides of the second equality. The result means that the CF of $\Delta/(2\mu)^2$ is simply

$$P_{\Delta}^{2T}(z) = \sum_{k=0}^{2T(Q^2-1)} c_k z^{-k}, \quad (A10)$$

where the sequence $(c_0,c_1,\ldots,c_{2T(Q-1)^2})$ is the sequence $(a_0,a_1,\ldots,a_{(Q-1)^2})$ convolving with itself $2T$ times. Once the $c_k$ s are found, the terms $\alpha$ and $\beta$ in (13) and (20) become
\[ \alpha = \frac{1}{1-c_0} \sum_{k=1}^{2T} c_k \left( \frac{\gamma}{4T} (4k \mu^2) \right)^R \]

\[ = \frac{1}{1-c_0} \sum_{k=1}^{2T} c_k \left( \frac{3\gamma}{2(M-1)T} \right)^R, \quad \text{(MQAM)} \quad (A11) \]

and

\[ \beta = \frac{1}{1-c_0} \sum_{k=1}^{2T} c_k \left( 1 + \frac{3\gamma}{4T} (4k \mu^2) \right)^R \]

\[ = \frac{1}{1-c_0} \sum_{k=1}^{2T} c_k \left( 1 + \frac{3\gamma}{2(M-1)T} \right)^R, \quad \text{(MQAM)} \quad (A12) \]

Finally, for any signal constellation with square distances between symbol pairs that can not be expressed as integer multiples of a basic quantity, the term \( \Delta \) has a multinomial distribution. Suppose there are \( J \) distinct square distances, \( D_1, D_2, \ldots, D_J \), between symbol pairs in the signal constellation. If under the random coding argument, these square distances occur with probabilities \( p_1, p_2, \ldots, p_J \), then the probability that

\[ \Delta = \sum_{i=1}^J n_i D_i \]

is

\[ \text{Pr} \left[ \Delta = \sum_{i=1}^J n_i D_i \right] = \frac{T!}{n_1! n_2! \cdots n_J!} p_1^{n_1} p_2^{n_2} \cdots p_J^{n_J}, \quad \sum_{i=1}^J n_i = T, \]

0, \quad \text{otherwise,} \]

(A13)

for non-negative integers \( n_1, n_2, \ldots, n_J \). The result can then be readily used in (13) and (20) to compute \( \alpha \) and \( \beta \).

REFERENCES


has been a consultant to a number of companies in Canada and abroad. Paul is a registered Professional Engineer in the province of British Columbia.

Kar-Peo Yar received the B.Eng. (Hons) and M.Sc. degrees in electrical engineering from the National University of Singapore in 1998 and 2001, respectively. In 2001 he was awarded the National Science Scholarship by the Agency for Science, Technology and Research to study at the University of Michigan, Ann Arbor, where he obtained his Ph.D. degree in electrical and computer engineering in 2007. Since 2007, he has been with the Institute for Infocomm Research, Singapore, where he is currently a research engineer. His research interests consist of coding and modulation, energy analysis and cross-layer design.


From 1976 to 1978, he was a member of the technical staff at the Bell Telephone Laboratories, Holmdel, N. J., where he was engaged in packet network studies. Since 1978, he has been with the Department of Electrical and Computer Engineering, National University of Singapore, where he is now a professor. He served as the Deputy Dean of Engineering and the Vice Dean for Academic Affairs, Faculty of Engineering of the National University of Singapore, from 2000 to 2003. His research interests are in detection and estimation theory, digital communications and coding. He spent the sabbatical year 1987 to 1988 at the Tokyo Institute of Technology, Tokyo, Japan, under the sponsorship of the Hitachi Scholarship Foundation. In year 2006, he was invited to the School of Engineering Science, Simon Fraser University, Burnaby, B.C., Canada as the David Bested Fellow.

Dr. Kam is a member of Eta Kappa Nu, Tau Beta Pi, and Sigma Xi. Since 1996, he has been the Editor for Modulation and Detection for Wireless Systems of the IEEE Transactions on Communications. He won the Best Paper Award at the IEEE VTC 2004 – Fall.