the interrelations between the infinite linear array and the isolated cluster setups, it can be shown that

$$C^{\rm IC}_{M_{\rm opt}} \le \mathcal{C}_M \le \left(1+\frac{2}{M}\right) C^{\rm IC}_{(M+2)_{\rm opt}}.$$

But from (A-3), both upper and lower bounds on $\mathcal{C}_M$ converge to the same limit $C^C_{\rm opt}$ as $M \to \infty$. This completes the proof of the theorem, since the term $\frac{2}{M}\mathcal{C}_I$ in (3–6) vanishes with M. □

REFERENCES

[1] A. D. Wyner, "Shannon-theoretic approach to a Gaussian cellular multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 1713–1727, Nov. 1994.

[2] S. Hanly and P. Whiting, "Information-theoretic capacity of multi-receiver networks," *Telecommun. Syst.*, vol. 1, pp. 1–42, 1993.

[3] O. Somekh and S. Shamai (Shitz), "Shannon-theoretic approach to a Gaussian cellular multi-access channel with fading," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1401–1425, Jul. 2000.

[4] S. Shamai (Shitz) and A. D. Wyner, "Informationtheoretic considerations for symmetric, cellular, multiple-access fading channels—Parts I & II," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1877–1911, Nov. 1997.

[5] S. Verdú and S. Shamai (Shitz), "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 622–640, Mar. 1999.

[6] S. Shamai (Shitz) and S. Verdú, "The impact of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1302–1327, May 2001.

[7] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," in *Foundations and Trends in Communications and Information Theory*. Hanover, MA: Now Publishers, 2004, vol. 1, pp. 1–182.

[8] B. M. Zaidel, S. Shamai (Shitz), and S. Verdú, "Multi-cell uplink spectral efficiency of coded DS-CDMA with random signatures," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 8, pp. 1556–1569, Aug. 2001, see also "Spectral efficiency of randomly spread DS-CDMA in a multi-cell model," *Proc. 37th Annu. Allerton Conf. Communications, Control, and Computing*, Monticello, IL, Sep. 1999, pp. 841–850.

[9] B. M. Zaidel, S. Shamai (Shitz), and S. Verdú, "Multi-cell uplink spectral efficiency of randomly spread DS-CDMA in Rayleigh fading channels," in *Proc. 6th Int. Symp. Communication Techniques and Applications (ISCTA'01)*, Ambleside, U.K., Jul. 2001, pp. 499–504, see also "Random CDMA in the multiple cell uplink environment: The effect of fading on various receivers," *Proc. 2001 IEEE Information Theory Workshop*, Cairns, Australia, Sep. 2001, pp. 42–45.

[10] B. M. Zaidel, S. Shamai (Shitz), and S. Verdú, "Impact of out-of-cell interference on strongest-users-only CDMA detectors," in *Proc. Int. Symp. Spread Spectrum Techniques and Applications (ISSSTA)*, Prague, Czech Republic, Sep. 2002, vol. 1, pp. 258–262.

[11] S. Shamai (Shitz), B. M. Zaidel, and S. Verdú, "Strongest-users-only detectors for randomly spread CDMA," in *Proc. IEEE Int. Symp. Information Theory*, Lausanne, Switzerland, Jun./Jul. 2002, p. 20.

[12] H. Li and H. V. Poor, "Power allocation and spectral efficiency of DS-CDMA systems in fading channels with fixed QoS—Part I: Single-rate case," *IEEE Trans. Wireless Commun.*, vol. 5, no. 9, pp. 2516–2528, Sep. 2006.

[13] H. Li and H. V. Poor, "Power allocation and spectral efficiency of DS-CDMA systems in fading channels with fixed QoS—Part II: Multiple-rate case," *IEEE Trans. Wireless Commun.*, vol. 5, no. 9, pp. 2529–2538, Sep. 2006.

[14] A. M. Tulino, S. Verdú, and A. Lozano, "Capacity of antenna arrays with space polarization and pattern diversity," in *Proc. IEEE Information Theory Workshop*, Paris, France, Mar./Apr. 2003, pp. 324–327.

[15] A. M. Tulino, A. Lozano, and S. Verdú, "Impact of antenna correlation on the capacity of multiantenna channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2491–2509, Jul. 2005.

[16] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1329–1343, Jun. 2002.

[17] S. V. Hanly and D. N. C. Tse, "Resource pooling and effective bandwidths in CDMA networks with multiuser receivers and spatial diversity," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1328–1351, May 2001.

[18] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.

[19] O. Somekh, B. M. Zaidel, and S. Shamai (Shitz), Spectral Efficiency of Joint Multiple Cell-Site Processors for Randomly Spread DS-CDMA Systems, Technion–Israel Institute of Technology, Haifa, Israel, 2004, CCIT Rep. 480.

Degrees of Freedom for the MIMO Interference Channel

Syed Ali Jafar, *Member, IEEE*, and Maralle Jamal Fakhereddin

***Abstract*—In this correspondence, we show that the exact number of spatial degrees of freedom (DOF) for a two user nondegenerate (full rank channel matrices) multiple-input–multiple-output (MIMO) Gaussian interference channel with M_1, M_2 antennas at transmitters 1, 2 and N_1, N_2 antennas at the corresponding receivers, and perfect channel knowledge at all transmitters and receivers, is $\min\{M_1+M_2, N_1+N_2, \max(M_1,N_2), \max(M_2,N_1)\}$. A constructive achievability proof shows that zero forcing is sufficient to achieve all the available DOF on the two user MIMO interference channel. We also show through an example of a share-and-transmit scheme how the gains of transmitter cooperation may be entirely offset by the cost of enabling that cooperation so that the available DOF are not increased.**

***Index Terms*—Broadcast, degrees of freedom (DOF), interference, multiple–input–multiple–output (MIMO), multiple access, zero forcing.**

I. INTRODUCTION

Multiple-input–multiple-output (MIMO) systems have assumed great importance in recent times because of their remarkably higher capacity compared to single-input–single-output (SISO) systems. It is well known [1]–[3] that capacity of a point–to–point (PTP) MIMO system with M inputs and N outputs increases linearly as $\min(M,N)$ at high signal–to–noise power ratio (SNR). For power and bandwidth limited wireless systems, this opens up another dimension – "space" that can be exploited in a similar way as time and frequency. Similar to time division and frequency division multiplexing, MIMO systems present the possibility of multiplexing signals in space. Spatial dimensions are especially interesting for how they may be limited by distributed processing as well the amount of channel knowledge. Previous work has shown that in the absence of channel knowledge, spatial degrees of freedom (DOF) are lost [4], [5]. Multiuser systems, with constrained cooperation between inputs/outputs distributed among multiple users, are especially challenging since, unlike PTP case, joint processing is not possible at inputs and outputs. The available spatial DOF are affected by the inability to jointly process the signals at the distributed inputs and outputs. The two user interference channel with

Manuscript received May 22, 2006; revised December 12, 2006. This work was supported in part by the National Science Foundation by CAREER Grant 0546860. The material in this correspondence was presented in part at the IEEE International Symposium on Information Theory, Seattle, WA, July 2006 .

S. A. Jafar is with the Department of Electrical Engineering and Computer Science, University of California, Irvine, CA 92697 USA (e-mail: syed@uci.edu).

M. J. Fakhereddin was a research intern with the California Institute of Technology, Pasadena, CA 91125 USA. She is now with VMware, Palo Alto, CA 91125 USA (e-mail: maralle@systems.caltech.edu).

Communicated by Y. Steinberg, Associate Editor for Shannon Theory.
Digital Object Identifier 10.1109/TIT.2007.899557

single antennas at all nodes is considered by Host-Madsen [6], [7]. It is shown that the maximum multiplexing gain is only equal to one even if cooperation between the two transmitters or the two receivers is allowed via a noisy communication link. Nosratinia and Høst-Madsen [8] show that even if communication links are introduced between the two transmitters as well as between the two receivers the highest multiplexing gain achievable is equal to one. These results are somewhat surprising as it can be shown that with ideal cooperation between transmitters (broadcast channel) or with ideal cooperation between receivers (multiple access channel) the maximum multiplexing gain is equal to 2.

In this correspondence, we focus on the two user (M_1, N_1, M_2, N_2) MIMO interference channel where transmitter 1 with M_1 antennas has a message for receiver 1 with N_1 antennas, and transmitter 2 with M_2 antennas has a message for receiver 2 with N_2 antennas. We develop a MIMO multiple-access channel (MAC) outer bound on the sum capacity of this MIMO interference channel. The outer bound is used to prove a converse result for the maximum number of DOF. We also provide a constructive proof of achievability of the DOF based on zero forcing. We show that the inner bound and the outer bound are tight, thereby establishing the precise number of DOF on the MIMO interference channel as $\min\{M_1 + M_2, N_1 + N_2, \max(M_1, N_2), \max(M_2, N_1)\}$. We also consider a simple cooperative scheme to understand why transmitter cooperation may not increase DOF. Through this simple scheme, we are able to show how the benefits of cooperation can be completely offset by the cost of enabling it.

II. Degrees of Freedom Measure

We assume that channel state is fixed and perfectly known at all transmitters and receivers. Also, we assume that the channel matrices are sampled from a rich scattering environment. Therefore we can ignore the measure zero event that some channel matrices are rank deficient. It is well known that the capacity of a *scalar* additive white Gaussian noise (AWGN) channel scales as $\log$ (SNR) at high SNR. On the other hand, for a single user MIMO channel with M inputs and N outputs, the capacity growth rate can be shown to be $\min(M, N) \log$ (SNR) at high SNR. This motivates the natural definition of spatial DOF as

$$\eta \triangleq \lim_{\rho \to \infty} \frac{C_\Sigma(\rho)}{\log(\rho)} \tag{1}$$

where $C_\Sigma(\rho)$ is the sum capacity (just capacity in case of PTP channels) at SNR ρ. In other words, DOF η represent the maximum *multiplexing gain* [3] of the generalized MIMO system. For PTP case, $\min(M, N)$ DOF are easily seen to correspond to the parallel channels that can be separated using the singular value decomposition (SVD) of the channel matrix, involving joint processing at the M inputs and N outputs, i.e.,

$$\eta(\text{PTP}) = \min(M, N). \tag{2}$$

A. The Multiple Access Channel (MAC)

The MAC channel is an example of a MIMO system where cooperation is allowed only between the channel outputs. Let the MAC consist of N outputs controlled by the same receiver and two users, each controlling M_1 and M_2 inputs for a total of $M = M_1 + M_2$ inputs. For the MAC, the available DOF are the same as with perfect cooperation between all users

$$\eta(\text{MAC}) = \eta(\text{PTP}) = \min(M_1 + M_2, N). \tag{3}$$

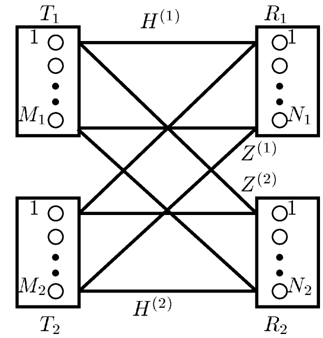

Fig. 1. (M_1, N_1) , (M_2, N_2) Interference channel.

The converse is straightforward because, for the same number of inputs and outputs, $\eta(\text{MAC}) \le \eta(\text{PTP}) = \min(M_1 + M_2, N)$. In other words, the lack of cooperation at the inputs cannot increase DOF. For achievability, it is interesting to note that zero forcing (ZF), which is normally a suboptimal strategy, is easily seen to be sufficient to utilize all DOF.

B. The Broadcast Channel

The BC channel is an example of a MIMO system where cooperation is allowed only between the channel inputs. Let the BC consist of M inputs controlled by the same transmitter and two users, each controlling N_1 and N_2 outputs for a total of $N = N_1 + N_2$ outputs. In a similar fashion as the MAC, it is easy to show that by ZF at the BC transmitter, $\min(M, N)$ parallel channels can be created, so that the total DOF are the same as with perfect cooperation between all the users

$$\eta(\text{BC}) = \eta(\text{MAC}) = \eta(\text{PTP}) = \min(M, N). \tag{4}$$

III. Interference Channel

Consider an $(M_1, N_1), (M_2, N_2)$ interference channel with two transmitters T_1 and T_2, and two receivers R_1 and R_2, where T_1 has a message for R_1 only and T_2 has a message for R_2 only. T_1 and T_2 have M_1 and M_2 antennas, respectively. R_1 and R_2 have N_1 and N_2 antennas, respectively. The interference channel is characterized by the following input–output relationships:

$$Y^{(1)} = H^{(1)} X^{(1)} + Z^{(1)} X^{(2)} + W^{(1)} \tag{5}$$
$$Y^{(2)} = H^{(2)} X^{(2)} + Z^{(2)} X^{(1)} + W^{(2)} \tag{6}$$

where we denote the $N_1 \times M_1$ channel matrix between T_1 and R_1 by $H^{(1)}$, the $N_2 \times M_2$ channel matrix between T_2 and R_2 by $H^{(2)}$, the $N_2 \times M_1$ channel matrix between T_1 and R_2 by $Z^{(2)}$, and the $N_1 \times M_2$ channel matrix between T_2 and R_1 by $Z^{(1)}$. $X^{(1)}, X^{(2)}$ are the M_1 and M_2 dimensional inputs vectors, $Y^{(1)}, Y^{(2)}$ are the N_1 and N_2 dimensional output vectors, and $W^{(1)}, W^{(2)}$ are the N_1 and N_2 dimensional additive white Gaussian noise (AWGN) vectors, respectively. As mentioned before, we assume that the channels are nondegenerate, i.e., all channel matrices are full rank. Fig. 1 shows an illustration of this interference channel. Without loss of generality we arrange the links so that link 1 always has the most number of antennas either at its transmitter or receiver, i.e., $\max(M_1, N_1) \ge \max(M_2, N_2)$.

A. Achievability: Inner Bound on the Degrees of Freedom

For the $(M_1, N_1), (M_2, N_2)$ interference channel we prove the following inner bound on the available DOF.

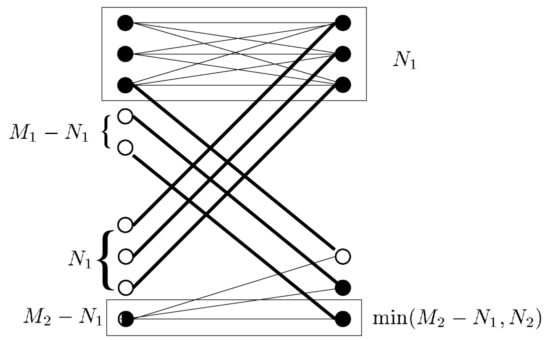

Fig. 2. Achievability proof for (M_1, N_1), (M_2, N_2) Interference channel when $M_1 \ge M_2, N_1, N_2$.

Lemma 1:

$$\eta(\text{INT}) \ge \min(M_1, N_1) + \min(M_2 - N_1, N_2)^+ \, 1(M_1 > N_1) + \min(M_2, N_2 - M_1)^+ \, 1(M_1 < N_1) \quad (7)$$

where $1(\cdot)$ is the indicator function and $(x)^+ = \max(0, x)$.

Sketch of Achievability Proof: According to our model, either $M_1 \ge N_1, M_2, N_2$ or $N_1 \ge M_1, M_2, N_2$. We explain the zero forcing based constructive achievability argument for the case when $M_1 \ge N_1, M_2, N_2$. The case with $N_1 \ge M_1, M_2, N_2$ follows similarly and is omitted to avoid repetition.

Based on (7), when $M_1 \ge N_1, M_2, N_2$, we need to show the achievability of $N_1 + \min(M_2 - N_1, N_2)^+ \, 1(M_1 > N_1)$ DOF. If either $M_1 = N_1$ or $M_2 \le N_1$ then we need to show the achievability of only N_1 DOF which can be trivially achieved by only allowing communication between T_1 and R_1. Therefore, we consider the remaining case of $M_1 > N_1$ and $M_2 > N_1$. In this case, we need to show the achievability of $N_1 + \min(M_2 - N_1, N_2)$ DOF. Fig. 2 illlustrates the scheme described in the remainder of this section with the example of an interference channel with $M_1 = 5, M_2 = 4, N_1 = 3, N_2 = 3$ where a total of 4 DOF are achieved.

Step 1. Let the singular value decomposition (SVD), $Z^{(1)} = U^{(1)}\Lambda^{(1)}V^{(1)\dagger}$ and $Z^{(2)} = U^{(2)}\Lambda^{(2)}V^{(2)\dagger}$, where $U^{(1)}$, $V^{(1)}$, $U^{(2)}$, $V^{(2)}$ are $N_1 \times N_1$, $M_2 \times M_2$, $N_2 \times N_2$, and $M_1 \times M_1$ unitary matrices, respectively. $\Lambda^{(1)}, \Lambda^{(2)}$ are $N_1 \times M_2$ and $N_2 \times M_1$ matrices with singular values of $Z^{(1)}$, $Z^{(2)}$ respectively on the main diagonal and zeros elsewhere. Using the standard MIMO SVD approach, we absorb the unitary matrices into the corresponding input and output vectors to obtain

$$Y^{(1)'} = H^{(1)'}X^{(1)'} + \Lambda^{(1)}X^{(2)'} + W^{(1)'} \quad (8)$$
$$Y^{(2)'} = H^{(2)'}X^{(2)'} + \Lambda^{(2)}X^{(1)'} + W^{(2)'} \quad (9)$$

where $Y^{(1)'} = U^{(1)\dagger}Y^{(1)}, Y^{(2)'} = U^{(2)\dagger}Y^{(2)}, X^{(1)'} = V^{(2)\dagger}X^{(1)}$, $X^{(2)'} = V^{(1)\dagger}X^{(2)}, W^{(1)'} = U^{(1)\dagger}W^{(1)}, W^{(2)'} = U^{(2)\dagger}W^{(2)}$, $H^{(1)'} = U^{(1)\dagger}H^{(1)}V^{(2)}$ and $H^{(2)'} = U^{(2)\dagger}H^{(2)}V^{(1)}$. In particular note that only the first N_1 columns of $\Lambda^{(1)}$ are nonzero.

$$\Lambda^{(1)} = \left[\text{Diag}(\lambda_1^{(1)}, \ldots, \lambda_{N_1}^{(1)}) \;\; \mathbf{0}_{N_1 \times (M_2 - N_1)}\right]. \quad (10)$$

Therefore, only the inputs $X_1^{(2)'}, X_2^{(2)'}, \ldots, X_{N_1}^{(2)'}$ present interference at R_1 from T_2. Similarly, only the inputs $X_1^{(1)'}, X_2^{(1)'}, \ldots, X_{N_2}^{(1)'}$ present interference at R_2 from T_1. In Fig. 2 the bold channels represent the interference paths after the diagonalization achieved through the SVD as there are $\min(5, 3) = 3$ parallel paths from T_1 to R_2 and $\min(4, 3) = 3$ parallel paths from T_2 to R_1.

Step 2. At transmitter T_1 we set inputs $X_1^{(1)'}, X_2^{(1)'}, \ldots, X_{M_1-N_1}^{(1)'}$ to zero, i.e., we do not transmit on these inputs. This leaves N_1 available inputs $X_{M_1-N_1+1}^{(1)'}, \ldots, X_{M_1}^{(1)'}$ at T_1. For the example of Fig. 2 the $M_1 - N_1 = 2$ transmit antennas indicated by white circles have their inputs set to zero and the dark circles indicate the three available inputs at T_1.

Step 3. At transmitter T_2 we set inputs $X_1^{(2)'}, X_2^{(2)'}, \ldots, X_{N_1}^{(2)'}$ to zero, i.e., we do not transmit on these inputs. This leaves $M_2 - N_1$ available inputs $X_{N_1+1}^{(2)'}, \ldots, X_{M_2}^{(2)'}$ at T_2. Fig. 2 illustrates this step as the three unused inputs are indicated by white circles and the remaining $M_2 - N_1 = 1$ input by a dark circle.

Step 4. The previous step eliminates any interference from T_2 to R_1 since all the interfering inputs have been set to 0. Therefore, communication between T_1 and R_1 takes place over an $N_1 \times N_1$ MIMO channel with no interference from T_2. N_1 DOF are achieved through this communication.

Step 5. At receiver R_2 we consider only outputs $Y_1^{[2]'}, Y_2^{[2]'}, \ldots, Y_{\min(M_1-N_1,N_2)}^{[2]'}$ and discard the rest. Note that because of Step 2, these outputs do not contain any interference from T_1.

Step 6. From Step 3, we have $M_2 - N_1$ available inputs at T_2. From Step 5, we have $\min(M_1 - N_1, N_2)$ outputs at R_2 with no interference from T_1. Therefore, the communication between T_2 and R_2 takes place over a MIMO channel with $\min(M_2 - N_1, \min(M_1 - N_1, N_2)) = \min(M_2 - N_1, N_2)$ DOF.

Combining Steps 4 and 6, we have established the achievability of the required total of $N_1 + \min(M_2 - N_1, N_2)$ DOF. Fig. 2 illustrates the proof with white circles indicating discarded inputs and outputs and black circles indicating the inputs and outputs used for the achievability scheme aforementioned.

B. Converse: Outer Bound on the Degrees of Freedom

For the $(M_1, N_1), (M_2, N_2)$ interference channel, we prove the following outer bound on the available DOF.

Lemma 2:

$$\eta(\text{INT}) \le \min\{M_1+M_2, N_1+N_2, \max(M_1,N_2), \max(M_2,N_1)\}.$$

To start with, notice that a trivial outer bound is obtained from the PTP case, i.e., $\eta(\text{INT}) \le \min(M_1+M_2, N_1+N_2)$. Indeed this outer bound coincides with the inner bound when either $\min(M_1, M_2) \ge N_1 + N_2$ or $\min(N_1, N_2) \ge M_1 + M_2$. In general, while the capacity region of the interference channel is not known even with single antennas at all nodes, various outer bounds have been obtained [9]–[11] that have been useful in finding the capacity region in some special cases [12], [13]. Most of the existing outer bounds are for single antenna systems.

For our purpose, we develop a genie-based outer bound for MIMO interference channel where the number of antennas at either receiver is $\ge$ the number of transmit antennas at the interfering transmitter, i.e., either $N_1 \ge M_2$ or $N_2 \ge M_1$. This outer bound is the key to the tight converse needed to establish the number of DOF. Note that for this section, since we do not need the assumption that $\max(M_1, N_1) \ge \max(M_2, N_2)$, the proof for the cases $N_1 \ge M_2$ or $N_2 \ge M_1$ is identical.

Theorem 1: For the $(M_1, N_1), (M_2, N_2)$ interference channel with $N_1 \ge M_2$, the sum capacity is bounded above by that of the corresponding (M_1, M_2, N_1) MAC channel with additive noise $W^{(1)} \sim \mathcal{N}(0, I_{N_1})$ modified to $W^{(1)'} \sim \mathcal{N}(0, K')$ where

$$K' = I_{N_1} - Z^{(1)}\left(Z^{(1)\dagger}Z^{(1)}\right)^{-1}Z^{(1)\dagger} + \alpha Z^{(1)}Z^{(1)\dagger}$$
$$\alpha = \min\left(\frac{1}{\sigma_{\max}^2(Z^{(1)})}, \frac{1}{\sigma_{\max}^2(H^{(2)})}\right).$$

TABLE I
THE SAME NUMBER OF DEGREES OF FREEDOM ARE OBTAINED FROM THE UPPER BOUND AND THE LOWER BOUND IN ALL CASES

$M_1>(M_2,N_1,N_2)$			$N_1\geq(M_1,M_2,N_2)$		
$N_1\geq M_2$	$N_1<M_2$		$N_2\geq M_1+M_2$	$N_2<M_1+M_2$	
$D=N_1$	$M_2\leq N_1+N_2$	$M_2>N_1+N_2$	$D=M_1+M_2$	$N_2\geq M_1$	$N_2<M_1$
	$D=M_2$	$D=N_1+N_2$		$D=N_2$	$D=M_1$

where $\sigma_{\max}(A)$ represents the principal singular value of matrix A.

Proof: Let us define

$$W_a^{(1)} \sim \mathcal{N}\left(0, I_{N_1} - Z^{(1)}\left(Z^{(1)\dagger}Z^{(1)}\right)^{-1}Z^{(1)\dagger}\right)$$
$$W_b^{(1)} \sim \mathcal{N}\left(0, Z^{(1)}\left(Z^{(1)\dagger}Z^{(1)}\right)^{-1}Z^{(1)\dagger} - \alpha Z^{(1)}Z^{(1)^\dagger}\right)$$
$$W_c^{(1)} \sim \mathcal{N}\left(0, \alpha Z^{(1)}Z^{(1)^\dagger}\right)$$

as three $N_1 \times 1$ jointly Gaussian and mutually independent random vectors. The positive semidefinite property of the respective covariance matrices is easily established from the definition of α.

Without loss of generality we assume

$$W^{(1)} = W_a^{(1)} + W_b^{(1)} + W_c^{(1)}$$
$$W^{(1)'} = W_a^{(1)} + W_c^{(1)}.$$

Since a part of the proof is similar to the corresponding proof for the single antenna case, we will summarize the common steps, and emphasize only the part that is unique to MIMO interference channel. Consider any achievable scheme for any rate point within the capacity region of the interference channel, so that R_1 and R_2 can correctly decode their intended messages from their received signals with sufficiently high probability.

Step 1. We replace the original additive noise $W^{(1)}$ at R_1 with $W^{(1)'}$ as defined in Theorem 1. We argue that this does not make the capacity region smaller because the original noise statistics can easily be obtained by locally generating and adding noise $W_b^{(1)}$ at R_1. Therefore, since R_1 was originally capable of decoding its intended message with noise $W^{(1)}$, it is still capable of decoding its intended message with $W^{(1)'}$.

Step 2. Suppose that a genie provides R_2 with side information containing the entire codeword $X^{(1)}$. Since $X^{(2)}$ is independent of $X^{(1)}$, R_2 simply subtracts out the interference from its received signal. Thus, the channel $Z^{(2)}$ can be eliminated without making the capacity region smaller.

Step 3. By our assumption, R_1 can decode its own message and therefore it can subtract $X^{(1)}$ from its own received signal as well. In this manner, after the interfering signals have been subtracted out we have

$$Y^{(1)} = Z^{(1)}X^{(2)} + W^{(1)'}$$
$$Y^{(2)} = H^{(2)}X^{(2)} + W^{(2)}.$$

To complete the proof we need to show that if R_2 can decode $X^{(2)}$ then so can R_1. This would imply that R_1 can decode both messages, hence giving us the MAC outer bound.

Step 4. Without loss of generality, let us perform SVD $H^{(2)} = F^{(2)}\Sigma^{(2)}G^{(2)\dagger}$ on the channel between T_2 and R_2. This is a lossless operation that leads to

$$Y^{(2)\text{new}} = X^{(2)\text{new}} + W^{(2)''} \tag{11}$$

where $X^{(2)\text{new}} = G^{(2)\dagger}X^{(2)}$ and $W^{(2)''}$ is additive noise that consists of independent zero mean complex Gaussian random variables with variances $\frac{1}{\sigma_i^2(H^{(2)})}$ and $\sigma_i(H^{(2)})$ are the singular values of $H^{(2)}$. Note that we have dropped dimensions that correspond to zero channel gains as these channels are useless for R_2.

Step 5. Next, we show that R_1 can obtain a stronger channel to $X^{(2)\text{new}}$ so that if R_2 can decode it, so can R_1. To this end, let R_1 use ZF to obtain

$$Y^{(1)\text{new}} = X^{(2)\text{new}} + V^{(2)}\left(Z^{(1)\dagger}Z^{(1)}\right)^{-1}Z^{(1)\dagger}W^{(1)'}$$
$$= X^{(2)\text{new}} + W^{(1)''}$$

where $W^{(1)''}$ is a vector of AWGN with i.i.d. elements and variance α.

Now both R_1 and R_2 have a diagonal channel with input $X^{(2)\text{new}}$ and uncorrelated additive white noise components on each diagonal channel. Moreover, the strongest channel for R_2 has noise $\frac{1}{\sigma_{\max}^2(H^{(2)})}$. However the noise on any channel for R_1 is only α which is smaller. Thus, we argue once again that R_1 can locally generate noise and add it to its received signal to create a statistically equivalent noise signal as seen by R_2. In other words, R_1 has a less noisy channel to T_2 and therefore can decode any signal that R_2 can. Since R_1 can decode T_1s message by assumption, we have the MAC outer bound. □

The previous theorem leads directly to the following corollary.

Corollary 1: For the $(M_1, N_1), (M_2, N_2)$ interference channel the number of spatial DOF $\eta(\text{INT}) \leq \max(M_2, N_1)$.

Proof: If $M_2 \leq N_1$ the sum capacity of the interference channel is upper-bounded by the multiple access channel with N_1 receive antennas. Therefore, for $M_2 \leq N_1$ we must have $\eta(\text{INT}) \leq N_1$. Now, if $M_2 > N_1$, then let us add more antennas to receiver 1 so that it has a total of M_2 receive antennas. Additional receive antennas cannot hurt, so the converse argument is not violated. However, with M_2 receive antennas at receiver 1, once again the multiple-access upper bound applies to the new interference channel. The number of DOF is, therefore, upper-bounded as $\eta(\text{INT}) \leq M_2$ when $M_2 > N_1$. Combining the two cases, we have the result of the corollary $\eta(\text{INT}) \leq \max(M_2, N_1)$. □

Simply by switching the arguments to user 2 instead of user 1, Corollary 1 leads to another upper bound: $\eta(\text{INT}) \leq \max(M_1, N_2)$ that holds for all M_1, M_2, N_1, N_2. Combining the two upper bounds of the corollary and the trivial PTP upper bounds we have the converse result.

Finally, we show that the achievable inner bound and the converse outer bound are always tight. The following theorem presents the main result of this correspondence.

Theorem 2: For the $(M_1, N_1), (M_2, N_2)$ interference channel the number of spatial DOF

$$\begin{aligned}\eta(\text{INT}) = {} & \min(M_1, N_1) \\ & + \min(M_2 - N_1, N_2)^+ \, 1(M_1 > N_1) \\ & + \min(M_2, N_2 - M_1)^+ \, 1(M_1 < N_1) \\ = {} & \min\{M_1+M_2, N_1+N_2, \max(M_1, N_2), \max(M_2, N_1)\}.\end{aligned}$$

Proof: The proof is found by verifying directly that the number of DOF obtained from the inner and outer bounds always match. The resulting number D from the $\eta(\text{INT})$ inner and outer bounds is listed for all cases in Table I. □

TABLE II
DOF OF MIMO INTERFERENCE CHANNELS FOR VARIOUS M_1, M_2, N_1, N_2

(M_1, N_1)	(M_2, N_2)	$\eta(INT)$
$(1,1)$	$(1,1)$	1
$(1,2)$	$(1,2)$	2
$(2,1)$	$(2,1)$	2
$(1,2)$	$(2,1)$	1
$(3,2)$	$(2,3)$	2
$(2,3)$	$(2,3)$	3
$(2,3)$	$(1,3)$	3
$(2,2)$	$(3,2)$	2
(n,m)	(m,n)	$\min(m,n)$
(m,n)	(m,n)	$\min(2m,n)(n \geq m)$

Thus, we have the exact number of DOF for all possible M_1, M_2, N_1, N_2. Some examples are provided in Table II. A couple of observations can be made about the spatial DOF. First, there is a reciprocity in that $\eta(INT)$ is unaffected if M_1 and M_2 are switched with N_1 and N_2, respectively. In other words, the DOF are unaffected if the directions of the messages are reversed. However, notice that $\eta(INT)$ may change if only M_1 and N_1 are switched while M_2 and N_2 are not switched. Finally, from the constructive achievability proof one can see that the available DOF can be divided among the two users in all possible ways so that the sum is $\eta(INT)$ and the individual DOF allocations are within the individual maxima of $\max(M_1, N_1)$ for user 1 and $\max(M_2, N_2)$ for user 2.

IV. EFFECT OF TRANSMIT COOPERATION ON THE NUMBER OF DOF

Comparing the interference channel and the BC channel obtained by full cooperation between the transmitters, it is clear that the available DOF are severely limited by the lack of transmitter cooperation in the interference channel. As an example, consider the interference channel with $(M_1, N_1) = (n, 1)$ and $(M_2, N_2) = (1, n)$. From the preceding section we know there is only one available DOF in this channel. However, if full cooperation between the transmitters is possible the resulting BC channel has $(M, N_1, N_2) = (n+1, 1, n)$. The number of DOF is now $n+1$. Therefore, transmitter cooperation would seem highly desirable. Rather surprisingly, it has been shown recently [6] that for the $(1,1), (1,1)$ interference channel, allowing the transmitters to cooperate through a wireless link between them (even with full duplex operation), does not increase DOF. For MIMO interference channels, as suggested by the example above, the potential benefits of cooperation are even stronger and it is not known if transmitter cooperation can increase DOF. The capacity results of [6] do not seem to allow direct extensions to MIMO interference channels.

To gain insights into the cost and benefits of cooperation in a MIMO interference channel, we consider a specific scheme where transmitters first share their information in a full duplex mode as a MIMO channel (Step 1) and subsequently transmit together as BC channel. We will refer to this scheme as the share-and-transmit scheme.

A. *Degrees of Freedom With Share-and-Transmit*

Consider an $(M, N), (M, N)$ interference channel ($M \leq N$). Also assume that each transmitter is sending information with rate R. Note that while we make the preceding simplifying assumptions for simplicity of exposition, the following analysis and the main result extend directly to the general case of unequal number of antennas and unequal rates.

From (7), we know that the number of DOF for this interference channel with no transmitter cooperation is $\min(M, N) + \min(M, N - M) = \min(2M, N)$. For the share-and-transmit scheme, we compute DOF as follows. We first find the capacity of the sharing link C_s and the capacity of transmission C_t. Then, we find the total capacity of the system C by evaluating the total amount of data transmitted divided by the total time it requires to transmit this data, i.e.

$$C = \frac{2R}{\frac{R}{C_s} + \frac{2R}{C_t}}. \tag{12}$$

Dividing by log(SNR) where SNR is large, we obtain the total number of DOF as

$$\lim_{\mathrm{SNR}\to\infty} \frac{C}{\log \mathrm{SNR}} = \frac{2}{\frac{1}{\mathrm{DOF(sharing)}} + \frac{2}{\mathrm{DOF(transmit)}}}. \tag{13}$$

The number of DOF for the sharing link is that of MIMO PTP channel with M transmit and receive antennas $= min(M, M) = M$. After transmitters share their information, they can fully cooperate as a $(2M, N, N)$ BC channel. The number of DOF for this channel is $\min(2M, 2N) = 2\min(M, N)$. Therefore, (13), which gives the total number of degrees of freedom for the share-and-transmit scheme, becomes $\frac{2M \min(M,N)}{M + \min(M,N)} = M$. Note that

$$M + \min(M, N - M)^+ \geq M. \tag{14}$$

Therefore, we conclude that (for this specific scheme) transmitter cooperation in the high SNR regime does not provide any advantage to the number of degrees of freedom in the MIMO interference channel.

V. CONCLUSION

We investigate the degrees of freedom for the MIMO interference channel. The distributed nature of the antennas significantly limits degrees of freedom. For an interference channel with a total of N transmit antennas and a total of N receive antennas, the available number of DOF can vary from N to 1 based on how the antennas are distributed among the two transmitters and receivers. Through an example of a share-and-transmit scheme, we show how the gains of transmitter cooperation can be entirely offset by the cost of enabling that cooperation so that the available degrees of freedom are not increased. Our result is in a sense a negative result, because similar to [7] it shows that on the MIMO interference channel there is nothing beyond zero forcing as far as spatial multiplexing is concerned.

REFERENCES

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, no. 6, pp. 311–335, Mar. 1998, published by Kluwer Academic.

[2] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecomm. (ETT)*, vol. 10, pp. 585–596, Nov. 1999.

[3] L. Zheng and D. N. Tse, "Packing spheres in the Grassmann manifold: A geometric approach to the noncoherent multi-antenna channel," *IEEE Trans. Inf. Theory*, vol. 48, pp. 359–383, Feb. 2002.

[4] S. Jafar, "Isotropic fading vector broadcast channels: The scalar upper dbound and loss in degrees of freedom," *IEEE Trans. Inf. Theory* , vol. 51, no. 3, pp. 848–857, Mar. 2005.

[5] A. Lapidoth, "On the high-SNR capacity of noncoherent networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3025–3036, Sep. 2005.

[6] A. Host-Madsen and Z. Yang, "Interference and cooperation in multisource wireless networks," in *IEEE Commun. Theory Workshop*, Jun. 2005.

[7] A. Host-Madsen, "Capacity bounds for cooperative diversity," *IEEE Trans. Inf. Theory*, vol. 52, pp. 1522–1544, Apr. 2006.

[8] A. Host-Madsen and A. Nosratinia, "The multiplexing gain of wireless networks," in *Proc. IEEE Int. Symp. Inf. Theory*, 2005.

[9] A. B. Carliel, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 29, pp. 602–606, Jul. 1983.

[10] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 50, pp. 581–586, Mar. 2004.

[11] S. Vishwanath and S. Jafar, "On the capacity of vector Gaussian interference channels," in *Proc. IEEE Inf. Theory Workshop*, Oct. 2004.

[12] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *Ann. Prob.*, pp. 805–814, Oct. 1974.
[13] A. B. Carliel, "A case where interference does not reduce capacity," *IEEE Trans. Inf. Theory*, vol. 21, pp. 569–570, Sep. 1975.

Derivatives of Entropy Rate in Special Families of Hidden Markov Chains

Guangyue Han and Brian Marcus, *Fellow, IEEE*

***Abstract*—Consider a hidden Markov chain obtained as the observation process of an ordinary Markov chain corrupted by noise. Recently Zuk *et al.* showed how, in principle, one can explicitly compute the derivatives of the entropy rate of at extreme values of the noise. Namely, they showed that the derivatives of standard upper approximations to the entropy rate actually stabilize at an explicit finite time. We generalize this result to a natural class of hidden Markov chains called "Black Holes." We also discuss in depth special cases of binary Markov chains observed in binary-symmetric noise, and give an abstract formula for the first derivative in terms of a measure on the simplex due to Blackwell.**

***Index Terms*—Analyticity, entropy, entropy rate, hidden Markov chain, hidden Markov model, hidden Markov process.**

I. Introduction

Let $Y = \{Y_{-\infty}^{\infty}\}$ be a stationary Markov chain with a finite state alphabet $\{1, 2, \ldots, B\}$. A function $Z = \{Z_{-\infty}^{\infty}\}$ of the Markov chain Y with the form $Z = \Phi(Y)$ is called a hidden Markov chain; here Φ is a finite-valued function defined on $\{1, 2, \ldots, B\}$, taking values in $\{1, 2, \ldots, A\}$. Let Δ denote the probability transition matrix for Y; it is well known that the entropy rate $H(Y)$ of Y can be analytically expressed using the stationary vector of Y and Δ. Let W be the simplex, comprising the vectors

$$\{w = (w_1, w_2, \ldots, w_B) \in \mathbb{R}^B : w_i \geq 0, \sum_i w_i = 1\}$$

and let W_a be all $w \in W$ with $w_i = 0$ for $\Phi(i) \neq a$. For $a \in A$, let Δ_a denote the $B \times B$ matrix such that $\Delta_a(i,j) = \Delta(i,j)$ for j with $\Phi(j) = a$, and $\Delta_a(i,j) = 0$ otherwise. For $a \in A$, define the scalar-valued and vector-valued functions r_a and f_a on W by

$$r_a(w) = w\Delta_a \mathbf{1}$$

and

$$f_a(w) = w\Delta_a / r_a(w).$$

Note that f_a defines the action of the matrix Δ_a on the simplex W.

If Y is irreducible, it turns out that the entropy rate

$$H(Z) = -\int \sum_a r_a(w) \log r_a(w) dQ(w) \tag{1.1}$$

where Q is *Blackwell's measure* [1] on W. This measure is defined as the limiting distribution $p(y_0 = \cdot | z_{-\infty}^0)$.

Recently, there has been a great deal of work on the entropy rate of a hidden Markov chain. Jacquet *et al.* [6] considered entropy rate of the hidden Markov chain Z, obtained by passing a binary Markov chain through a binary-symmetric channel with crossover probability ε, and computed the derivative of $H(Z)$ with respect to ε at $\varepsilon = 0$. For the same channel, Ordentlich and Weissman used Blackwell's measure to bound the entropy rate [11] and obtained an asymptotical formula for entropy rate [12]. For certain more general channels, Zuk *et al.* [16], [17] proved a "stabilizing" property of the derivatives of entropy rate of a hidden Markov chain and computed the Taylor series expansion for a special case. Several authors have observed that the entropy rate of a hidden Markov chain can be viewed as the top Lyapunov exponent of a random matrix product [5], [6], [3]. Under mild positivity assumptions, Han and Marcus [4] showed the entropy rate of a hidden Markov chain varies analytically as a function of the underlying Markov chain parameters.

In Section II, we establish a "stabilizing" property for the derivatives of the entropy rate in a family we call "Black Holes." Using this property, one can, in principle, explicitly calculate the derivatives of the entropy rate for this case, generalizing the results of [16], [17].

In Section III, we consider binary Markov chains corrupted by binary-symmetric noise. For this class, we obtain results on the support of Blackwell's measure, and for a special case, that we call the "nonoverlapping" case, we express the first derivative of the entropy rate as the sum of terms, involving Blackwell's measure, which have meaningful interpretations.

II. Stabilizing Property of Derivatives in Black Hole Case

Suppose that for every $a \in A$, Δ_a is a rank one matrix, and every column of Δ_a is either strictly positive or all zeros. In this case, the image of f_a is a single point and each f_a is defined on the whole simplex W. Thus, we call this the *Black Hole* case. Analyticity of the entropy rate at a Black Hole follows from Theorem 1.1 of [4].

As an example, consider a binary-symmetric channel with crossover probability ε. Let $\{X_n\}$ be the input Markov chain with the transition matrix

$$\Pi = \begin{bmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{bmatrix}. \tag{2.2}$$

At time n the channel can be characterized by the following equation:

$$Z_n = X_n \oplus E_n$$

where $\oplus$ denotes binary addition, E_n denotes the independent and identically distributed (i.i.d.) binary noise with $p_E(0) = 1 - \varepsilon$ and $p_E(1) = \varepsilon$, and Z_n denotes the corrupted output. Then $Y_n = (X_n, E_n)$ is jointly Markov, so $\{Z_n = \Phi(Y_n)\}$ is a hidden Markov chain with the corresponding

$$\Delta = \begin{bmatrix} \pi_{00}(1-\varepsilon) & \pi_{00}\varepsilon & \pi_{01}(1-\varepsilon) & \pi_{01}\varepsilon \\ \pi_{00}(1-\varepsilon) & \pi_{00}\varepsilon & \pi_{01}(1-\varepsilon) & \pi_{01}\varepsilon \\ \pi_{10}(1-\varepsilon) & \pi_{10}\varepsilon & \pi_{11}(1-\varepsilon) & \pi_{11}\varepsilon \\ \pi_{10}(1-\varepsilon) & \pi_{10}\varepsilon & \pi_{11}(1-\varepsilon) & \pi_{11}\varepsilon \end{bmatrix};$$

here, Φ maps states 1 and 4 to 0 and maps states 2 and 3 to 1 (the reader should not confuse Π with the 4×4 matrix Δ, which defines the hidden Markov chain via a deterministic function). When $\varepsilon = 0$

$$\Delta = \begin{bmatrix} \pi_{00} & 0 & \pi_{01} & 0 \\ \pi_{00} & 0 & \pi_{01} & 0 \\ \pi_{10} & 0 & \pi_{11} & 0 \\ \pi_{10} & 0 & \pi_{11} & 0 \end{bmatrix}.$$

Manuscript received April 14, 2006; revised January 20, 2007. The material in this correspondence was presented at the IEEE International Symposium on Information Theory, Seattle, WA, July 2006.

The authors are with the Department of Mathematics, University of British Columbia, Vancouver, BC V6T 1Z2, Canada (e-mail: ghan@math.ubc.ca; marcus@math.ubc.ca).

Communicated by Y. Steinberg, Associate Editor for Shannon Theory.
Digital Object Identifier 10.1109/TIT.2007.899467