International Symposium on Information Theory and its Applications, ISITA2004
Parma, Italy, October 10–13, 2004

On the Sublinear Behavior of MIMO Channel Capacity at Low SNR

Siddharth Ray, Muriel Médard, Lizhong Zheng and Jinane Abounadi

Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, MA 02139, USA
E-mail: {sray, medard, lizhong, jinane}@mit.edu

Abstract

We consider wideband wireless communication over a multiple-input, multiple-output (MIMO) Rayleigh fading channel where the transmitter has no channel state information (CSI). We study the channel with perfect receiver CSI (coherent channel) versus the channel with no CSI at the receiver (non-coherent channel). A channel with partial CSI at the receiver has a capacity between the coherent and non-coherent channels and in this paper, we study the two extremes.
The capacities of the coherent and non-coherent channels differ in their sublinear term and in [6], Verdú computed the sublinear term of the coherent channel and gave a lower bound for the sublinear term of the non-coherent channel. However, without an upper bound for this term, we cannot quantify the loss in capacity (or increase in bandwidth required) due to lack of receiver CSI. In this paper, we compute an upper bound for the sublinear term of the non-coherent MIMO channel and show that it is approximately $O(\mathsf{SNR})$. We therefore quantify the maximum penalty for not having CSI at the receiver.
For the i.i.d. Rayleigh fading channel, we study the effect of the number of receive antennas on the loss in capacity due to not having CSI at the receiver. We show that this loss increases monotonically with the number of receive antennas. Therefore, as the number of receive antennas increases, channel estimation becomes more and more important. This also shows that the capacity increase due to power gain from the additional receive antennas is more if we know the additional paths at the receiver.

1. Introduction and Preliminaries

It is well known from [1]-[5] that the use of large bandwidth helps in improving the power efficiency of time varying point to point channels. In this paper, we consider wideband wireless communication over a Rayleigh fading channel using multiple antennas at the transmitter and the receiver. We will denote the number of transmit antennas as t and the number of receive antennas as r. A wideband channel can be modelled as a set of N parallel narrowband channels with independent and identical statistics, such that each channel is flat faded. We will denote SNR as the signal-to-noise ratio per symbol time per narrowband channel. By symmetry, the capacity of the wideband channel with power constraint P is N times the capacity of each such narrowband channel with power constraint P/N. As $N \to \infty$, $\mathsf{SNR} \to 0$ for the narrowband channel and hence our interest is in the low signal to noise ratio (SNR) regime.
We consider two cases. The first is when the receiver has perfect estimate of the channel state and uses that for decoding what is being transmitted. This is called a coherent channel and we will denote its ergodic capacity as $C_c(\mathsf{SNR})$. The second is the channel where the receiver has no CSI. This channel is known as a non-coherent channel. We will denote its capacity as $C_{nc}(\mathsf{SNR})$. We will denote the type of channel by using the appropriate subscript. Note that for both cases, the transmitter does not have any knowledge of the channel state.
Thus, the coherent and non-coherent channels can be viewed as two extreme cases which we study in this paper. Any channel with partial channel knowledge will have a capacity $C(\mathsf{SNR})$ that will lie between the two, i.e.

$$C_{nc}(\mathsf{SNR}) \leq C(\mathsf{SNR}) \leq C_c(\mathsf{SNR}).$$

As the receiver's knowledge of the channel state improves, the capacity increases from $C_{nc}(\mathsf{SNR})$ to $C_c(\mathsf{SNR})$.
For a MIMO channel with t transmit antennas and r

[†]This research is supported by HP/MIT Award 008542-008, NSF Ultra Wideband Wireless Award ANI0335256 and NSF Career award CCR 0093349.

receive antennas, it has been shown in [7] that

$$\lim_{\mathsf{SNR}\to 0} \frac{C_c(\mathsf{SNR})}{\mathsf{SNR}} = \lim_{\mathsf{SNR}\to 0} \frac{C_{nc}(\mathsf{SNR})}{\mathsf{SNR}} = r. \quad (1)$$

Moreover, we can expand the capacity function as a Taylor's series at $\mathsf{SNR} = 0$ as

$$C(\mathsf{SNR}) = C^{'}(0)\mathsf{SNR} + \frac{C^{''}(0)}{2}\mathsf{SNR}^2 + o(\mathsf{SNR}^2).$$

From (1), we have for the coherent as well as the non-coherent channel, $C_c^{'}(0) = C_{nc}^{'}(0) = r$. Therefore, the difference in capacity for the coherent and non-coherent channels lies in the sublinear term of the capacity expression. In order to study this, let us denote the sublinear term for a MIMO channel with t transmit antennas, r receive antennas and capacity $C(\mathsf{SNR})$ as $\Delta^{(t,r)}(\mathsf{SNR})$, where

$$\Delta^{(t,r)}(\mathsf{SNR}) \triangleq r\mathsf{SNR} - C(\mathsf{SNR}).$$

This sublinear term can be bounded as

$$\Delta_c^{(t,r)}(\mathsf{SNR}) \leq \Delta^{(t,r)}(\mathsf{SNR}) \leq \Delta_{nc}^{(t,r)}(\mathsf{SNR}).$$

As the knowledge of channel state improves at the receiver, the sublinear capacity term decreases from $\Delta_{nc}^{(t,r)}(\mathsf{SNR})$ to $\Delta_c^{(t,r)}(\mathsf{SNR})$. For the coherent channel, it is shown in [6] that

$$\Delta_c^{(t,r)}(\mathsf{SNR}) = \frac{r(r+t)}{2t}\mathsf{SNR}^2 + o(\mathsf{SNR}^2).$$

However, the sublinear term $\Delta_{nc}(\mathsf{SNR})$ for the non-coherent case is not well studied. The only result that we are aware of is [6] that shows that

$$\lim_{\mathsf{SNR}\to 0} \frac{\Delta_{nc}^{(t,r)}(\mathsf{SNR})}{\mathsf{SNR}^2} = \infty,$$
$$\equiv \Delta_{nc}^{(t,r)}(\mathsf{SNR}) \gg O(\mathsf{SNR}^2).$$

Though this provides a clear distinction between the coherent and non-coherent channels by saying that the sublinear term in the capacity expression for the non-coherent channel is much larger than $O(\mathsf{SNR}^2)$, it does not tell us how large it can be. Thus, from [6] we know that the penalty we pay in terms of bandwidth for not knowing the channel is extremely large but we do not know by how much. In this work, we find out more precisely what the right order of the sub-linear term for the non-coherent channel is. We compute an upper bound for $\Delta_{nc}^{(t,r)}(\mathsf{SNR})$ as a function of SNR. This tells us how large the bandwidth penalty can be for not knowing the channel. Note that as the receiver estimate of the channel state improves, the bandwidth penalty goes to 0.

This problem has recently been studied in [10] for the single-input, single-output channel. In this paper, we generalize the results to the MIMO channel and aim at understanding the effect of multiple antennas on the sublinear term of the channel capacity.

We study the loss in capacity from not having receiver CSI. This loss can be expressed as

$$C_c(\mathsf{SNR}) - C_{nc}(\mathsf{SNR}) = \Delta_{nc}^{(t,r)}(\mathsf{SNR}) + O(\mathsf{SNR}^2).$$

In the low SNR regime, we may neglect terms of $O(\mathsf{SNR}^2)$ and the loss in capacity is given by $\Delta_{nc}^{(t,r)}(\mathsf{SNR})$.

We present the channel model in Section 2 and compute an upper bound to the sublinear term for the non-coherent channel in Section 3. Section 4 looks at the loss in capacity from not having CSI at the receiver. We conclude in Section 5.

2. Channel Model

We consider a multiple-input, multiple-output (MIMO) channel where SNR is low and neither the transmitter nor the receiver have knowledge of the channel. The signal to noise ratio at each receive antenna will be denoted as SNR. Let the number of transmit and receive antennas be t and r, respectively. We deal with a linear model in which the received vector $\mathbf{y} \in \mathcal{C}^r$ depends on the transmit vector $\mathbf{x} \in \mathcal{C}^t$ via

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w},$$

where $\mathbf{H}$ is a $r \times t$ complex matrix and $\mathbf{w}$ is zero-mean complex Gaussian noise vector with independent, equal variance real and imaginary parts. We assume $\mathcal{E}[\mathbf{w}\mathbf{w}^\dagger] = I_r$, that is, the noises corrupting the different receivers are independent. We assume that the channel is flat faded. The entries of $\mathbf{H}$ are independent and each entry is zero-mean complex Gaussian, with independent real and imaginary parts, each with variance $1/2$. Equivalently, each entry of $\mathbf{H}$ has uniformly distributed phase and Rayleigh distributed magnitude. The coherence time of the channel spans l symbols within which the channel matrix is assumed constant. We thus model a Rayleigh block fading channel with enough separation within the transmitting and receiving antennas to achieve independence in the entries of $\mathbf{H}$.

3. Computing the Upper Bound

We know from [8] that as the coherence length of a non-coherent channel increases, its capacity increases and tends towards the coherent capacity. Thus, the sublinear term is a non-increasing function of the coherence length l. Hence to upper bound this term, it suffices to consider the i.i.d.. Rayleigh fading channel ($l = 1$). Denoting the capacity and sublinear term for the i.i.d. Rayleigh fading non-coherent channel as $C_{iid}(\mathsf{SNR})$ and $\Delta_{iid}^{(t,r)}(\mathsf{SNR})$ respectively, we have

$$\Delta_{nc}^{(t,r)}(\mathsf{SNR}) \leq \Delta_{iid}^{(t,r)}(\mathsf{SNR}). \quad (2)$$

In [8], it is shown that capacity of a non-coherent MIMO channel does not increase by increasing the number of transmit antennas beyond the coherence length. Hence, while considering the non-coherent i.i.d. Rayleigh fading channel, it suffices to let $t = 1$ from a capacity point of view.

Let us choose on-off signaling to communicate over the channel. Therefore, the input distribution can be specified as

$$x = \begin{cases} A & \text{w.p. } \delta \\ 0 & \text{w.p. } 1-\delta \end{cases} \qquad \delta = \frac{\mathsf{SNR}}{A^2}$$

for $A \in \mathcal{R}$. The capacity achievable using on-off signaling, $C_{\mathsf{on-off}}(\mathsf{SNR})$, may be less than $C_{iid}(\mathsf{SNR})$ since on-off signaling may not be optimal. Though the structure of the capacity achieving input distribution for the MIMO non-coherent channel has been specified in [8], the optimal input distribution for the i.i.d. Rayleigh fading MIMO non-coherent channel is still an open problem[1]. Thus by assuming such on-off signaling, we get a lower bound on the capacity and an upper bound on the sublinear capacity term for the non-coherent i.i.d. Rayleigh fading channel. Let us denote the non-linear capacity term for on-off signaling as $\Delta_{\mathsf{on-off}}^{(t,r)}(\mathsf{SNR})$. We have

$$\Delta_{iid}^{(t,r)}(\mathsf{SNR}) \leq \Delta_{\mathsf{on-off}}^{(t,r)}(\mathsf{SNR}). \quad (3)$$

With the on-off signaling described, we have the following probability distributions

$$p_{\mathbf{y}|x=0}(\vec{y}) = \frac{1}{\pi^r} \exp(-\|\vec{y}\|^2),$$

$$p_{\mathbf{y}|x=A}(\vec{y}) = \frac{1}{\pi^r (1+A^2)^r} \exp\left(-\frac{\|\vec{y}\|^2}{1+A^2}\right).$$

The mutual information $I(x, \mathbf{y})$ can be written as $I(x, \mathbf{y}) = h(\mathbf{y}) - h(\mathbf{y}|x)$. Now,

$$\begin{aligned} h(\mathbf{y}) &= -\int p_{\mathbf{y}}(\vec{y}) \log(p_{\mathbf{y}}(\vec{y})) d\vec{y} \\ &= -\log(1-\delta) + h(\mathbf{y}|x) + \delta D(p_{\mathbf{y}|x=A} || p_{\mathbf{y}|x=0}) \\ &\quad -(1-\delta) \int p_{\mathbf{y}|x=0}(\vec{y}) \log\left[1 + \frac{\delta}{1-\delta} \frac{p_{\mathbf{y}|x=A}(\vec{y})}{p_{\mathbf{y}|x=0}(\vec{y})}\right] d\vec{y} \\ &\quad -\delta \int p_{\mathbf{y}|x=A}(\vec{y}) \log\left[1 + \frac{\delta}{1-\delta} \frac{p_{\mathbf{y}|x=A}(\vec{y})}{p_{\mathbf{y}|x=0}(\vec{y})}\right] d\vec{y}. \end{aligned}$$

(Note: For brevity, we provide the outline of the mutual information calculation. The detailed proof follows similar lines as the proof for the single-input, single-output case considered in [10].)

The divergence $D(p_{\mathbf{y}|x=A} || p_{\mathbf{y}|x=0})$ is the divergence between the distributions of two Gaussian random vectors and is therefore

$$D(p_{\mathbf{y}|x=A} || p_{\mathbf{y}|x=0}) = r(A^2 - \log(1+A^2))$$

Computing the integrals we obtain the expression for the mutual information in the low SNR regime as

$$\begin{aligned} I(x, \mathbf{y}) &= \\ & r\mathsf{SNR} - r\mathsf{SNR}\frac{\log(1+A^2)}{A^2} \\ &\quad - rA^{-\frac{2(r+1)}{A^2}} \mathsf{SNR}^{1+\frac{1}{A^2}} + o(\mathsf{SNR}^2). \end{aligned}$$

Therefore the capacity of the channel with on-off signaing is given by

$$\begin{aligned} C_{\mathsf{on-off}}(\mathsf{SNR}) &= \max_A I(x, \mathbf{y}) \\ &= r\mathsf{SNR} - \frac{r\mathsf{SNR}}{\log(\frac{r}{\mathsf{SNR}})} + o(\mathsf{SNR}^2). \end{aligned}$$

The sublinear term for the non-coherent channel is therefore

$$\Delta_{\mathsf{on-off}}^{(t,r)}(\mathsf{SNR}) = \frac{r\mathsf{SNR}}{\log(\frac{r}{\mathsf{SNR}})} + o(\mathsf{SNR}^2).$$

Using (2,3), we can therefore upper bound the sublinear term for any non-coherent channel as

$$\Delta_{nc}^{(t,r)}(\mathsf{SNR}) \leq \frac{r\mathsf{SNR}}{\log(\frac{r}{\mathsf{SNR}})} + o(\mathsf{SNR}^2). \quad (4)$$

Thus, the upper bound to the sublinear term for the non-coherent channel can be said to be almost $O(\mathsf{SNR})$. This tells us the maximum penalty we pay in terms of capacity or bandwidth by not having CSI at the receiver. Note that the lower bound for the non-coherent MIMO capacity becomes

$$C_{nc}(\mathsf{SNR}) \geq r\mathsf{SNR} - \frac{r\mathsf{SNR}}{\log(\frac{r}{\mathsf{SNR}})} + o(\mathsf{SNR}^2).$$

[1] Recent research by the authors show that on-off signaling is optimal for the i.i.d. Rayleigh fading MIMO channel.

In [9], by trying to estimate the channel at the receiver at low SNR, a loose capacity lower bound of $O(\mathsf{SNR}^2)$ is obtained for the Rayleigh faded MIMO channel. Our results provide a much tighter bound than [9] since, our lower bound is $O(\mathsf{SNR})$ even without channel state knowledge at the receiver.

4. Effect of receive antennas on capacity loss due to no receiver CSI

In this section, we study the effect of the number of receive antennas on the loss in capacity due to no CSI at the receiver for an i.i.d. Rayleigh fading channel. We have seen in Section 1 that this loss is captured by the sublinear capacity term, $\Delta_{nc}^{(t,r)}(\mathsf{SNR})$.
Now using (4), $\forall r \geq 1$ and $\mathsf{SNR} \to 0$,

$$\Delta_{iid}^{(t,r+1)}(\mathsf{SNR}) - \Delta_{iid}^{(t,r)}(\mathsf{SNR}) \approx \frac{\mathsf{SNR}}{\log(\frac{r}{\mathsf{SNR}})} > 0.$$

Thus, $\Delta_{iid}^{(t,r)}(\mathsf{SNR})$ is a monotonically increasing function of the number of receive antennas r. Therefore as we increase the number of receive antennas, the gap between the coherent capacity and the non-coherent capacity for the i.i.d. Rayleigh fading channel increases. This happens because as we increase r, we increase the number of paths between the transmitting and receiving antennas. Both the coherent and non-coherent channels get a capacity gain due to extra power received at the additional antennas. However, the capacity boost from knowing the additional paths and having a power gain outweighs not knowing them and having a power gain. Hence, though the capacities for both the coherent and non-coherent i.i.d. Rayleigh fading MIMO channels increase with r, the latter moves further away from the former as r increases[2].

5. Conclusion

In this paper we quantify the maximum capacity loss or increased bandwidth needed due to lack of receiver CSI for a wideband MIMO channel with Rayleigh fading. We show that the sublinear term for the capacity of the non-coherent channel is at most $O(\mathsf{SNR})$ which tells us the maximum penalty we pay in terms of capacity or bandwidth by not having CSI at the receiver. We also see that as the number of receive antennas increases for the i.i.d. Rayleigh fading channel, the gain from receiver CSI is more since the non-coherent capacity moves further and further away from the coherent capacity with increase in the number of receive antennas.

[2]Recent research by the authors shows that this holds for any coherence length.

References

[1] R. S. Kennedy, “Fading Dispersive Communication Channels”, New York: *Wiley Interscience*, 1969.

[2] E. Telatar and D. N. C. Tse, “Capacity and mutual information of wideband multipath fading channels”, *IEEE Transactions on Information Theory*, Vol. 46, July 2000.

[3] E. Telatar, “Coding and multiaccess for the energy limited Rayleigh fading channel”, *Master’s Thesis*, MIT, 1988.

[4] M. Médard and R.G. Gallager, “Bandwidth scaling for fading multipath channels”, *IEEE Transactions on Information Theory*, Vol. 48, April 2002.

[5] V. G. Subramanian and B. Hajek, “Broad-band fading channels: Signal burstiness and Capacity”, *IEEE Transactions on Information Theory*, Vol. 48, pp. 908-27, April 2002.

[6] S. Verdú, “Spectral efficiency in the wideband regime”, *IEEE Transactions on Information Theory*, Vol. 48, pp. 1319-1343, June 2002.

[7] L. Zheng and D. N. C. Tse, “Communication on the Grassman Manifold: A geometric approach to the noncoherent multiple-antenna channel”, *IEEE Transactions on Information Theory*, Vol. 48, pp. 359-83, February 2002.

[8] T. L. Marzetta and B. M. Hochwald, “Capacity of a mobile multiple-antenna communication link in flat fading”, *IEEE Transactions on Information Theory*, Vol. 45, pp. 139-157, Jan. 1999.

[9] B. Hassibi and B. M. Hochwald, “How much training is needed in multiple-antenna wireless links ?”, *IEEE Transaction on Information Theory*, Vol. 49, pp. 951-963, April 2003.

[10] L. Zheng, D. N. C. Tse and M. Médard, “Channel Coherence in the low SNR Regime”, private communication to be submitted to *IEEE Transactions on Information Theory.*