Physical limits to the capacity of wide-band Gaussian MIMO channels

Anna Martini, Andrea Massa Member IEEE, Massimo Franceschetti Member IEEE

Abstract—In this letter a physical limit to the information capacity of a multiple-input multiple-output (MIMO) Gaussian channel is presented exploiting the theory of non-redundant sampling of scattered fields. For a MIMO narrow-band system of arbitrarily large spatial extension, the information capacity limit is the same as the one of a single-input single-output (SISO) ultra-wide band (UWB) system. For MIMO systems of finite size, transmitting over a range of frequencies, space and frequency diversities can be optimally combined by allocating the signal power uniformly across space, and increasing linearly across frequency.

Index Terms: Multiple-input multiple-output (MIMO), ultra-wide band (UWB), physical channel modeling, frequency diversity, space diversity.

I. INTRODUCTION

In the last years, there has been considerable interest in MIMO wireless communication, mostly due to the promise of capacity boost due to the spatial diversity provided by multiple antennas transmitting in a scattering environment, see [1], [2], [3], [4] for an overview.

In order to precisely estimate the performance of a space-time communication system, it is necessary to accurately characterize the wireless multipath channel. A common approach consists in defining a random matrix \( \mathbf{H} \) of i.i.d. channel gains, of dimension \( n_T \times n_R \), \( n_T \) and \( n_R \) being the number of transmitting and receiving antennas. The MIMO channel can then be decomposed into \( n = \min(n_T, n_R) \) parallel SISO subchannels, subject to a total power constraint [4]. However, practical experience shows that the channel capacity does not increase proportionally to \( n \), but it is limited by the physical properties of the propagation channel [5]. In other words, the effective number of degrees of freedom (ENF), i.e., the number of SISO subchannels actively participating in conveying information, can be much smaller than \( n \). Clearly, it is worth allocating power only to these active channels, and knowledge of their number is crucial to predict the ultimate system performance.

One way to more realistically model the effect of the propagation environment in the computation of the capacity is to introduce correlations in the matrix \( \mathbf{H} \), see for example [1], [6], [7], [8], [9], [10], however this approach fails to capture the inherent geometric features of the problem and leads to capacity computations that depend on the specific modeling assumptions. On the other hand, one can try to find a direct connection between the natural process of wave propagation and the number of independent information-theoretic channels between transmitters and receivers. This approach that blends the physics of propagation with the theory of information, has gained recent attention in both the electromagnetic and the communication theory community [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24]. In this framework, the recent works in [11], [12], [18] present physical MIMO communication limits in the narrow-band regime, obtained by applying the theory of non-redundant sampling of a monochromatic electromagnetic field radiated by a bounded domain. They rely on electromagnetic theory results that appeared in [25], [26]. These works have shown that scattered fields radiated by a domain of finite size are essentially space-bandlimited functions, in the sense that they can be approximated by space-bandlimited functions with an error that undergoes a sharp decay after a prescribed value of the spatial bandwidth of the field. Such a value, which is called the effective spatial bandwidth of the field, is proportional to the size of the scattering system

This work was in part supported in Italy by the Progettazione di un Livello Fisico ‘Intelligente’ per Reti Mobili ad Elevata Riconfigurabilità, Progetto di Ricerca di Interesse Nazionale - MIUR Project COFIN 2005099984; and in U.S. by NSF grant number number is CCF-0635048 and ONR grant YIP-N00014-07-1-0870.

A shorter version of this paper appeared in the proceedings of the 41st Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, California, 2007.

The authors are with the the Department of Information Engineering and Computer Science, University of Trento, via Sommarive 14, 38050 Trento, Italy. Massimo Franceschetti is also with the California Institute for Telecommunications and Information Technology, Department of Electrical and Computer Engineering, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92037-0407.
normalized to the wavelength. The number of spatial degrees of freedom of the electromagnetic field, i.e., the infimum of the number of dimensions of all the subspaces that can represent the field with the required degree of accuracy over a given observation domain, can then be approximated by the Nyquist number, thereby being the product of the effective bandwidth of the field times the angular extension of the observation domain. Exploiting these previous results, the work in [18] has shown that the number of spatial degrees of freedom of the electromagnetic field represents an upper bound for the ENDF of a MIMO channel. Thus, the number of equivalent SISO subchannels cannot be made arbitrarily large by increasing the number of transmitting and receiving antennas, but it is limited by the physical properties of the electromagnetic field. Similar results, based on different models but reaching the same conclusions also appeared in [13], [19], [21], [22], [23].

We observe that all of the above results hold under the assumption of narrow-band signals. In the present work, we address the general case of wide-band signals. The main difficulty in doing so is the mutual coupling that arises between the time and space spectra of the field. Accordingly, we decompose the transmitted signal into a combination of narrow-band frequency bins for which it is possible to apply the electromagnetic theory of [25], [26]. Then, by applying superposition and letting the bin width tend to zero, we obtain a physical limit to the capacity of an AWGN MIMO channel with arbitrary frequency bandwidth, along with a strategy for the optimal power allocation in the frequency domain. While the obtained capacity formula relies on the assumption that the channel is flat in space and frequency, the obtained power allocation strategy holds more generally in the absence of channel state information. It shows an optimal uniform allocation of power across space and linearly increasing across frequencies, due to the larger number of spatial modes occurring at high frequencies. This result is not based on stochastic modeling, but only on physical considerations on wave diversity limits.

The remainder of this letter is organized as follows. In Section II the geometry of the problem is described and the main results of previous works that are used throughout the letter are briefly recalled. Section III presents the mathematical formulation of the problem. In Section IV some considerations on the optimal signal power allocation in the frequency domain are drawn, while Section V provides an analysis of the channel capacity for different system configurations. Final comments and conclusions are drawn in Section VI.

II. Problem Statement

Let us consider the scattering system depicted in Figure 1. All the electromagnetic sources, along with the environmental obstacles, are enclosed in a two-dimensional ball \( B \) of radius \( a \). The receiving antennas are located on an observation domain \( M \) external to \( B \). The observation domain \( M \) is an arc of circumference concentric with the radiating ball \( B \) and it spans the range \([-S/2, S/2]\). This ideal geometric configuration is chosen for ease of presentation, but a generalization to three-dimensional radiating systems and two-dimensional receiving domains of arbitrary shape can also be obtained, see [25], [26], [27].

We now briefly recall some main results of electromagnetic theory and their application in the framework of MIMO systems. All of these results hold under the assumption of narrow-band signals.

In [25], it has been formally established that the field radiated by the ball \( B \) has effective spatial bandwidth

\[ W_s = \beta a, \]  

(1)

where \( a \) is the radius of \( B \), and \( \beta \) is the wavenumber defined as

\[ \beta = \frac{2\pi}{\lambda}, \]  

(2)

\( \lambda \) being the wavelength.

Starting from this result, the number of spatial degrees of freedom of the scattered field can be approximated
from below, as \( a/\lambda \to \infty \), by the Nyquist number [26],
\[
\eta_s \approx 2 \left( \frac{S}{r_m} \right) \left( \frac{W_s}{2\pi} \right) = 2\Omega \left( \frac{a}{\lambda} \right),
\] (3)
where \( r_m \) is the distance of \( \mathcal{M} \) from the center of \( \mathcal{B} \) and \( \Omega \) is the angle subtended by the observation domain \( \mathcal{M} \) to the center of \( \mathcal{B} \), see Figure 1.

The concepts above have been recently successfully applied to MIMO communication theory. Migliore [18] has shown that \( \eta_s \) represents an upper bound to the ENDF of the MIMO channel. Such a result is obtained by expressing the channel matrix \( \mathbf{H} \) in terms of the physical radiation operator (Green’s function) that relates all the electromagnetic sources inside \( \mathcal{B} \) to the scattered field measured on \( \mathcal{M} \). Then, due to the compactness of the radiation operator, it is possible to perform a singular value decomposition (SVD) of the channel matrix \( \mathbf{H} \) and view the continuous space MIMO channel as a finite number of independent SISO subchannels which cannot exceed \( \eta_s \). The evaluation of the channel capacity then follows from standard information-theoretic arguments [2], and assuming that the number of receiving and transmitting antennas is large enough to convey the whole information content that the electromagnetic field can carry, we have
\[
C = \sum_{i=1}^{\eta_s} \log \left( 1 + \frac{P_i^r t_i^r}{N_N} \right) \text{[bits/Hz]},
\] (4)
where \( N_N \) is the power spectral density of the additive white Gaussian noise in the time evolution, \( P_i^r \), \( i = 1, \ldots, \eta_s \), are the water-filling power allocations [28], and \( t_i^r \), \( i = 1, \ldots, \eta_s \), are the singular values of the channel matrix \( \mathbf{H} \), i.e., the power gains of each of the \( \eta_s \) SISO subchannels. We let the overall transmit power constraint be
\[
\sum_{i=1}^{\eta_s} P_i^r \leq P.
\] (5)

Now, if the communication modes are the same, i.e., \( t_i^r = t_j^r \), \( i, j = 1, \ldots, \eta_s \), then the water-filling power allocation requires that equal power is allocated to all SISO subchannels and (4) becomes,
\[
C = \eta_s \log \left( 1 + \frac{P}{\eta_s N_N} \right) \text{[bits/Hz]},
\] (6)

Above equation shows that the increase in capacity in MIMO communication is due to the exploitation of the electromagnetic spatial wave diversity; it is clear that even if one adds more and more receiving and transmitting antennas, at a given frequency and for the prescribed system geometry, i.e., for a given \( \eta_s = 2\Omega a/\lambda \), equation (6) represents a physical limit.

The above results hold under the assumption of narrow-band signals. The aim of this work is to deal with the more general case where the electromagnetic field radiated by \( \mathcal{B} \) is defined on a non-zero frequency range, \( W_t = f_{\text{max}} - f_{\text{min}} \).

Before proceeding further, let us write the power constraint in the frequency domain. The primary field sources, along with the polarization currents, are uniformly bounded. It follows that the power associated to the electromagnetic wave radiated by the ball \( \mathcal{B} \) is bounded as well, and we can write:
\[
\int_{f_{\text{min}}}^{f_{\text{max}}} N_S(f) df \leq P,
\] (7)

\( N_S(f) \) being the power spectral density of the radiated field, i.e., the signal conveying the information in the MIMO channel.

### III. Computation of the Channel Capacity

In order to estimate the channel capacity, let us decompose the signal outgoing the ball \( \mathcal{B} \) as a sum of narrowband frequency bins. The \( i \)-th component is defined on a range \( \Delta f = f_{i+1} - f_i \) and carries power \( P_i = N_S(f_i) \Delta f \) according to the following constraint:
\[
\lim_{\Delta f \to 0} \sum_{i=1}^{n(\Delta f)} N_S(f_i) \Delta f \leq P,
\] (8)

where \( n(\Delta f) = W_t/\Delta f \), \( f_1 = f_{\text{min}} \), and \( f_n(\Delta f) = f_{\text{max}} - 1 \).

We now process each one of these narrowband components separately according to the outline provided in the previous section. The capacity of \( i \)-th frequency bin according to (6) is
\[
C_i = \eta_i \Delta f \log \left( 1 + \frac{P_i}{\eta_i \Delta f N_N} \right) \text{[bits/s]},
\] (9)
\( \eta_i \) being the number of spatial degrees of freedom of the \( i \)-th frequency bin. We explicitly observe that \( \eta_i \) depends on the given frequency bin we are considering,
\[
\eta_i = 2\Omega \left( \frac{a}{\lambda_i} \right) = 2\Omega \left( \frac{a}{c} \right) f_i,
\] (10)
c being the velocity of propagation of the wave. Accordingly, it is convenient to rewrite (9) as follows,
\[
C_i = \alpha f_i \Delta f \log \left( 1 + \frac{P_i}{\alpha f_i \Delta f N_N} \right) \text{[bits/s]},
\] (11)
where
\[ \alpha = 2\Omega \left( \frac{a}{c} \right) \]
(12)
is a system parameter that takes into account the size of the radiating ball \(B\) and of the observation domain \(\mathcal{M}\).

We now want to find the optimal power allocation that maximizes the sum of the capacities of each frequency bin, when the bin-width tends to zero. Accordingly, assuming the channel to be flat in frequency, we have

\[ C = \max_{N_S(f)} \left\{ \lim_{\Delta f \to 0} \sum_{i=1}^{n(\Delta f)} C_i \right\} \]

\[ = \max_{N_S(f)} \left\{ \lim_{\Delta f \to 0} \sum_{i=1}^{n(\Delta f)} \alpha f_i \Delta f \log \left( 1 + \frac{P_i}{\alpha f_i \Delta f N_N} \right) \right\} \text{[bits/s]}, \]

(13)

\(N_S(f)\) being constrained by (7). Now, in order to write (13) as an integral, we can write \(P_i = N_S(f_i) \Delta f\), leading to

\[ C = \max_{N_S(f)} \left\{ \lim_{\Delta f \to 0} \sum_{i=1}^{n(\Delta f)} \alpha f_i \Delta f \log \left( 1 + \frac{N_S(f_i)}{\alpha f_i N_N} \right) \right\} \]

\[ = \max_{N_S(f)} \int_{f_{\min}}^{f_{\max}} \alpha f \log \left( 1 + \frac{N_S(f)}{\alpha f N_N} \right) df \text{[bits/s]}. \]

Thus, taking into account the power constraint (7), the original problem of computing the channel capacity can be recast as the following constrained optimization problem:

\[ \begin{aligned}
C &= \max_{N_S(f)} \int_{f_{\min}}^{f_{\max}} \alpha f \log \left( 1 + \frac{N_S(f)}{\alpha f N_N} \right) df, \\
\text{s.t.} \quad & \int_{f_{\min}}^{f_{\max}} N_S(f) df = P, \quad N_S(f) \geq 0.
\end{aligned} \]

(14)

The above problem can be solved by applying standard variational calculus methods [29]. We define \(\overline{\eta}_s\) and \(\overline{W}_s\) as the number of degrees of freedom and the spatial bandwidth at the average frequency \(f_{av} = (f_{\max} + f_{\min})/2\) respectively, i.e.,

\[ \overline{\eta}_s = \alpha f_{av} = 2\Omega \left( \frac{a}{c} \right) f_{av}, \]

(15)

and

\[ \overline{W}_s = \frac{2\pi a}{\lambda_{av}} = \left( \frac{2\pi a}{c} \right) f_{av}. \]

(16)

The optimal power allocation is given by

\[ N_S^{opt}(f) = \frac{2P}{f_{\max} - f_{\min}} f = \frac{P}{f_{\max} W_t} f, \]

(17)

where \(f_{\min} \leq f \leq f_{\max}\), which results in the capacity

\[ C = \overline{\eta}_s W_t \log \left( 1 + \frac{P}{\overline{\eta}_s W_t N_N} \right) \]

\[ = \frac{\Omega W_s W_t}{\pi} \log \left( 1 + \frac{\pi P}{\Omega W_s W_t N_N} \right). \]

(18)

Above capacity is achieved by superposition of different space codes at each frequency within the band. The power allocation increases with the frequency, due to the larger number of spatial modes occurring at high frequencies, and for each frequency it is allocated uniformly across the different spatial modes.

IV. PHYSICAL CONSIDERATIONS ON THE POWER ALLOCATION STRATEGY

Let us first fix the bandwidth \(W_t\), i.e., the time diversity of the system, and study the behavior of \(N_S^{opt}(f)\) for different values of the average frequency \(f_{av}\). We want to analyze how a change in the spatial diversity affects the optimal power allocation. The situation is depicted in subfigure (a) of Figure 2, where we observe that the slope of \(N_S^{opt}(f)\) decreases as \(f_{av}\) increases, becoming flat in the limit of \(f_{av} \to \infty\). In order to provide some intuition for this behavior, we can quantify the gap between the number of degrees of freedom associated to the frequencies \(f_{\min}\) and \(f_{\max}\) as,

\[ \Delta \eta = \frac{\eta_s(f_{\max}) - \eta_s(f_{\min})}{\overline{\eta}_s} = \frac{W_t}{f_{av}}. \]

(19)

By (19) it is clear that for low values of \(f_{av}\), the gap is larger and it is convenient to favor the high frequencies within the band, so that the power is allocated linearly in frequency and with a very high slope. On the other hand, when \(f_{av}\) increases, \(\Delta \eta\) decreases, so the gain obtained by allocating more power to the frequencies approaching \(f_{\max}\) decreases and the signal power allocation becomes flat in frequency.

Next, we fix the frequency \(f_{\min}\) and analyze the behavior of \(N_S^{opt}(f)\) when \(f_{\max}\) increases. This situation is depicted in subfigure (b) of Figure 2. The slope of \(N_S^{opt}(f)\) decreases as \(W_t\) becomes larger. Such behavior is due to the need of spreading the limited power over the whole channel’s frequency spectrum. While for low
$W_t$ values we are still able to exploit the space diversity of the system by allocating more power at the higher frequencies, as $W_t$ increases it becomes more and more difficult to simultaneously exploit time and space diversities. In the limit case of $W_t \to \infty$ we obtain a uniform power allocation, spread over the whole bandwidth, as in the case of UWB SISO systems [30].

V. PHYSICAL CONSIDERATIONS ON THE CHANNEL CAPACITY

It is easy to see that when either the average space ($\overline{W}_s$) or the frequency ($W_t$) band tends to infinity, (18) is bounded by

$$C_\infty = \frac{P}{N N \ln(2)}, \quad (20)$$

where $C_\infty$ also corresponds to the capacity of a SISO UWB channel [28].

We now focus on the role of the average spatial bandwidth in (18). With reference to Figure 3, let us fix a given frequency bandwidth $W_t$. We observe that the capacity increases compared to the corresponding SISO system by increasing $\overline{W}_s$, i.e., by increasing the size of the radiating ball with respect to $\lambda_{av} = c/f_{av}$. In the limit $\overline{W}_s \to \infty$, the spatial information content of the transmitted signal is maximized and the capacity tends to $C_\infty$. This is the corresponding of the UWB channel in the spatial domain.

It is worth pointing out that as $\overline{W}_s \to \infty$ the capacity tends to $C_\infty$ for any fixed frequency bandwidth $W_t$, meaning that in the limit of large spatial band there is no additional gain in exploiting a wide frequency band $W_t$.

We finally want to spend a few words on the role of the angular size of the observation domain $M$. Since $\Omega$ can not exceed $2\pi$, we have

$$C \leq 2\overline{W}_s W_t \log \left( 1 + \frac{P}{2\overline{W}_s W_t N N} \right) \text{[bits/sec]} \quad (21)$$

This inequality also holds for observation domains of arbitrary shape. Indeed, with reference to narrowband signals, the work of [27] introduced the notion of local
bandwidth measured in terms of the angular coordinate of the observation domain to the center of \( B \), and has shown that a variable sampling rate, which changes according to the value of the local bandwidth on the observation domain, leads to a minimum total number of significative samples which is independent of the shape of the observation domain and cannot exceed \( 2\pi(2a/c) f \). This upper bound on the number of degrees of freedom is achieved by any closed line embracing the whole scattering system, confirming the intuition that a closed line allows to intercept all the information flow outgoing the radiating ball \( B \).

VI. CONCLUSIONS

In this letter, the physical limit of the information capacity of a MIMO system transmitting over an arbitrary frequency band has been studied. Our results rely on the electromagnetic wave diversity limits of [25], [26], and build upon the recent works [11], [18], [21], [22] dealing with the narrowband case.

It has been shown that the maximum physically achievable information rate is bounded by \( C_{\infty} \) and can be achieved in two limiting cases. In the first case, time diversity is maximized as the frequency bandwidth used for transmission tends to infinity. This is the UWB SISO case, for which there is no additional gain in using more than one transmitting and one receiving antenna. On the other hand, \( C_{\infty} \) can also be achieved by increasing the spatial bandwidth, i.e., the size of the scattering system with respect to the wavelength at the average signal frequency. Such channel configuration can be seen as a spatial UWB channel, for which no further increase in the capacity can be obtained by increasing the frequency bandwidth.

In contrast, when neither the spatial bandwidth nor the frequency bandwidth tend to infinity, information transmission can rely on both spatial and temporal diversity, and there is an advantage with respect to channels exploiting only one kind of diversity.

REFERENCES


