Entropy **2008**, *10*, 722-735; DOI: 10.3390/e10040722

OPEN ACCESS

entropy

ISSN 1099-4300

www.mdpi.com/journal/entropy

Article

Intercept Capacity: Unknown Unitary Transformation

John Kitchen [1,*], **Bill Moran** [1] **and Stephen D. Howard** [2]

[1] Department of Electrical and Electronic Engineering, University of Melbourne, Australia. E-mails: john.kitchen@dsto.defence.gov.au; b.moran@ee.unimelb.edu.au

[2] EWRD, DSTO, PO Box 1500, Edinburgh, Australia. E-mail: stephen.howard@dsto.defence.gov.au.

* Author to whom correspondence should be addressed.

Received: 22 May 2008 / Accepted: 10 November 2008/ Published: 20 November 2008

Abstract: We consider the problem of intercepting communications signals between Multiple-Input Multiple-Output (MIMO) communication systems. To correctly detect a transmitted message it is necessary to know the gain matrix that represents the channel between the transmitter and the receiver. However, even if the receiver has knowledge of the message symbol set, it may not be possible to estimate the channel matrix. Blind Source Separation (BSS) techniques, such as Independent Component Analysis (ICA) can go some way to extracting independent signals from individual transmission antennae but these may have been preprocessed in a manner unknown to the receiver. In this paper we consider the situation where a communications interception system has prior knowledge of the message symbol set, the channel matrix between the transmission system and the interception system and is able to resolve the transmissionss from independent antennae. The question then becomes: what is the mutual information available to the interceptor when an unknown unitary transformation matrix is employed by the transmitter.

Keywords: MIMO, Communications, Unitary, BSS.

---

1. Introduction

In this paper we are interested in differential entropy and mutual information as it applies to wireless communication systems employing antenna arrays at both the transmission and receiving sites. Systems of this type are more commonly known as Multiple-Input Multiple-Output (MIMO) communication systems. MIMO communication techniques are known to provide increased information capacity over that

Figure 1. MIMO Wireless Intercept Model.

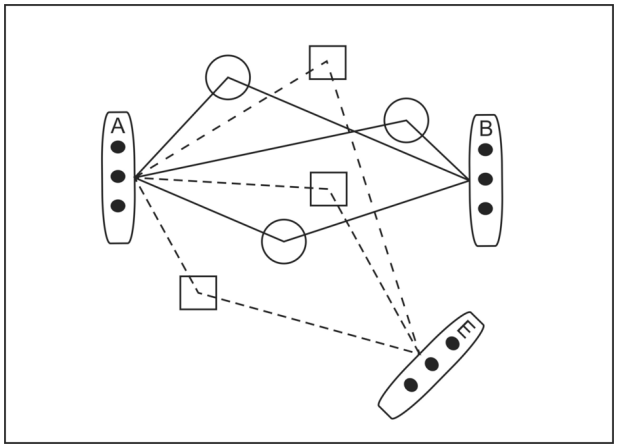
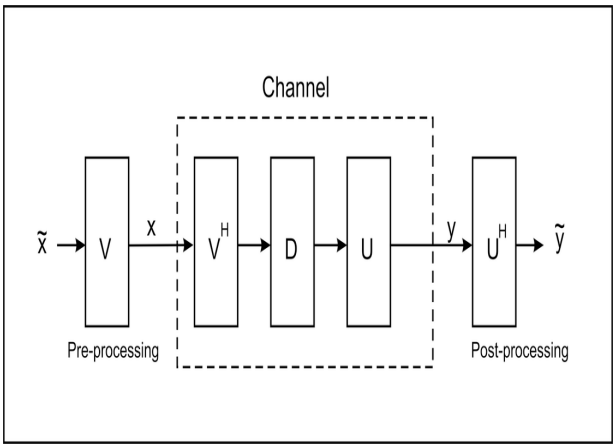


Figure 2. Converting MIMO channel to parallel channel via SVD.



obtainable via a single transmit antenna to single receive antenna system [1, 2]; however this extra capacity comes at the expense of increased system complexity and additional processing. To correctly receive and detect the transmitted message, the receive system must know the channel, or mixing, matrix as well as the message symbol set being used. The channel matrix may be estimated when a predetermined, known message sequence is incorporated into the transmitted message and the receiver knows where in the message this sequence occurs. However this training sequence may not always be available and this presents a blind source estimation problem where neither the message nor the channel matrix are known to the receiver. One possible solution to this problem is to employ a Blind Source Separation (BSS) technique such as Independent Component Analysis (ICA) [3] which can go some way to extracting the signals from individual transmission antennae with the caveat that all but one of the transmitted signals must have a non-gaussian probability distribution. In some cases the transmitted signals may have been preprocessed in a manner unknown to the receiver. In this paper we consider the situation where a communications receiving system has prior knowledge of the message symbol set, the channel matrix between the transmission system and the receiving system, is able to resolve the transmissions from the, assumed independent, transmitter antennae but does not know the unitary transformation that has been applied at the transmitter. The question then becomes: what is the mutual information available to the receiver when an unknown unitary transformation matrix is employed by the transmitter?

In the following sections we derive expressions for differential entropy, mutual information and hence capacity for a two-element transmit array to two-element receive array system which we shall refer to as a 2-Dimensional (2D) system. The 3D case is studied in the appendix giving a basis for a high snr approximation for the general N-Dimensional (ND) case. The general snr, ND case is derived and the resulting intended-receiver and intercept receiver mutual informations are compared.

2. Problem and Assumptions

The model that we shall employ for a MIMO system is the simple linear transformation

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{1}$$

where $\mathbf{y}$ is the received signal vector, $\mathbf{x}$ is the transmitted vector, $\mathbf{n}$ is additive receiver noise and $\mathbf{H}$ is the channel gain or mixing matrix between the transmitter and receiver. The standard MIMO channel model [11] assumes independent identically distributed (i.i.d.), frequency-flat Rayleigh fading between the transmit and receive antennae. Consequently the components $\mathbf{H}_{i,j}$ of $\mathbf{H}$ are typically modelled with a complex Normal density i.e. $\mathbf{H}_{i,j} \sim \mathcal{CN}(0,1)$. Here we shall assume $\mathbf{H}$ to be constant for both the intended and eavesdropper channels. In [11] the authors show that, for the case where the channel is unknown and with block coding over a coherence time T, the signal structure that achieves capacity is formed by the product of an isotropically distributed unitary matrix and a independent real, nonnegative diagonal matrix. For the purpose of this study we shall treat all of $\mathbf{y}$, $\mathbf{x}$, and $\mathbf{n}$ as real random variables. The benefit of this approach will be to simplify the derivations whilst recognising that, if the real and imaginary parts of the variables are independent, the results may be readily extended to the complex case by increasing the dimensionality of the vectors.

Figure 1 illustrates the scenario that we are studying. Employing a well-known cryptographic convention [4], the transmission source array is labelled Alice (A), the intended cooperative receiver array is labelled Bob (B) and the unintended, passive intercept receiver is labelled Eve (E). The lines represent the paths that signals take from transmitter antennae to receiver antennae. Shapes in the signal paths represent objects that cause signal scattering. An important point to realise here is that the paths (channel $\mathbf{H}_B$) between A and B are different to those between A and E (channel $\mathbf{H}_E$).

The channel matrix can be factorized using Singular Value Decomposition (SVD) as : $\mathbf{H} = \mathbf{UDV}^\dagger$ and we can then use:

$$\begin{aligned} \mathbf{U}^\dagger\mathbf{y} &= \mathbf{DV}^\dagger\mathbf{x} + \mathbf{U}^\dagger\mathbf{n} \\ \text{or} \quad \tilde{\mathbf{y}} &= \mathbf{D}\tilde{\mathbf{x}} + \tilde{\mathbf{n}}. \end{aligned}$$

This allows us to view the MIMO system as if it were composed of a set of parallel channels and the input data vector can be designed with this in mind. Figure 2 shows how this channel, with pre and post-processing, may be configured. For such an approach to work the transmitter requires precise knowledge of the channel matrix and it is a simple matter for the intended receiver to obtain the (scaled) message, since $\mathbf{D}$ is a real diagonal matrix. However for an unintended receiver, with a different (known) channel matrix, an unknown unitary transformation has been applied. In this case we desire to know how the mutual information, is affected. We make the following assumptions:

- $\mathbf{y}$ is a real $N \times 1$ observation vector.
- $\mathbf{U}$ is a real $N \times N$ unitary (orthogonal) matrix.
- $\mathbf{x}$ is a real $N \times 1$ random Gaussian signal vector, $x_i \sim \mathcal{N}(0, \sigma_x^2)$.
- $\mathbf{n}$ is a real $N \times 1$ random Gaussian noise vector, $n_i \sim \mathcal{N}(0, \sigma_n^2)$.
- $||\mathbf{x}|| = \sqrt{\sum_{i=1}^{N} x_i^2} = A =$ constant.
- the intended channel ($\mathbf{H}_B$) is known to both Alice and Bob.
- Eve knows the intercept channel ($\mathbf{H}_E$) but not the intended channel.

- the channels $\mathbf{H}_B$ and $\mathbf{H}_E$ vary slowly with time (or over many symbol periods) and may be assumed constant for the present study.

Based on the last assumption, Eve attempts to estimate the signal vector by applying the channel inverse as

$$\hat{\mathbf{x}} = \mathbf{H}_E^{-1}\mathbf{y} = \mathbf{V}\tilde{\mathbf{x}} + \mathbf{H}_E^{-1}\mathbf{n}_E. \tag{2}$$

Eve is therefore unable to directly obtain $\tilde{\mathbf{x}}$ due to the unknown unitary matrix $\mathbf{V}$. In applying the channel inverse, the noise vector has also been scaled and the modified noise covariance term $\mathbf{H}_e^{-1}\Sigma_n\mathbf{H}_e^{-T}$ shows that the intercept receiver may be operating with a different signal to noise ratio to that of the intended receiver. This also indicates that Eve could obtain better mutual information with a better channel.

Optimal power allocation to the parallel channels between Alice and Bob would typically be implemented via a technique called waterfilling, see [5] chapter 5, and hence lead to optimal system capacity. We have not taken waterfilling into account in this study and simply assume that equal power is assigned to each of the parallel channels.

We could proceed to derive the eavesdropper mutual information in a cartesian or a polar coordinate system. Of course it doesn't matter which coordinate system we choose - we should get the same answer. It is well known that differential entropy involves a Jacobian (J) in the transformation of coordinates [6] , leading to a $\ln\det(J)$ term but this will cancel in the mutual information calculations because mutual information is a relative entropy i.e. the difference between two entropies. For the purpose of this study our derivations will be based on a cartesian coordinate system. We shall derive differential entropies according to the definitions given by Cover and Thomas in [7], i.e. the differential entropy $h(Y)$ of a continuous random variable Y with a probability density $p(y)$ is defined as

$$h(Y) \triangleq -\int_{\mathbb{Y}} p(y)\ln(p(y))dy, \tag{3}$$

where $\mathbb{Y}$ is the support set of the random variable. When we have two random variables Y, X with joint probability density $p(y,x)$, the conditional differential entropy is defined as

$$h(Y|X) \triangleq -\int_{\mathbb{Y},\mathbb{X}} p(y,x)\ln(p(y|x))dy\,dx, \tag{4}$$

where $\mathbb{X}$ is the support set of the random variable X. The Mutual Information (MI) between the two random variables Y and X is defined as

$$\begin{aligned} I(Y;X) &= \int_{\mathbb{Y},\mathbb{X}} p(y,x)\ln\frac{p(y,x)}{p(y)p(x)}dy\,dx \\ &= h(Y) - h(Y|X) = h(X) - h(X|Y). \end{aligned} \tag{5}$$

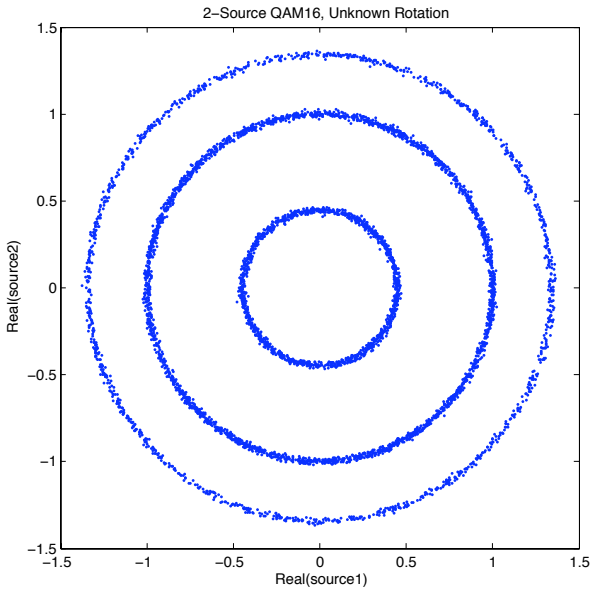
The capacity C is then obtained by maximizing the mutual information over all probability distributions for the source i.e. over $p(x)$:

$$C = \sup_{p(\mathbf{x})} I(Y;X). \tag{6}$$

It is well known [7] that a Gaussian source distribution is an entropy maximizer (for a given variance) so that, by treating $\mathbf{x}$ as a vector with i.i.d Gaussian components, the resulting differential entropy expressions will determine the capacity. Since the channels are assumed known we may consider $\mathbf{y} = \mathbf{x} + \mathbf{n}$

Figure 3. 2D Transmitter message symbol set.

Figure 4. Received ring distribution caused by unknown rotation on message symbol set.



to represent the fully informed (unitary transformation known) case and $\mathbf{y} = \mathbf{V}\mathbf{x} + \mathbf{n}$ to represent the partially informed (unitary transformation unknown) case. We can write $\mathbf{x} = \frac{\mathbf{x}}{||\mathbf{x}||}||\mathbf{x}||$ to obtain

$$\mathbf{y} = \mathbf{V}\frac{\mathbf{x}}{||\mathbf{x}||}||\mathbf{x}|| + \mathbf{n} = \mathbf{v}A + \mathbf{n} \tag{7}$$

where $A = ||\mathbf{x}||$ and $\mathbf{v} = V\frac{x}{||x||}$ is a unit vector for which we may or may not know the rotations. For the random vectors $\mathbf{y}$ and $\mathbf{x}$ the mutual information for the fully informed model is given by:

$$I_F = h(\mathbf{y}) - h(\mathbf{y}|\mathbf{x}, \mathbf{V}) \tag{8}$$

and for the partially informed model the mutual information is obtained from:

$$I_P = h(\mathbf{y}) - h(\mathbf{y}|A) \tag{9}$$

where the message amplitude A is known but not the rotation angles.

3. 2D Capacity

To illustrate the consequence of not knowing the rotation imposed by the orthogonal transformation in the 2D case, figure 3 shows a message symbol set where each of the two transmitters can set one of four possible values. Thus a constellation containing 16 points may be realised at the receiver and the density of these points is determined by the additive noise. If the rotation is unknown but the amplitude levels are known then the receiver might obtain a message that looks something like figure 4 where the density of the rings is determined by the additive noise.

3.1. 2D Density Function

We can construct the joint density function beginning with

$$p(y_1, y_2|x_1, x_2) = p(y_1|x_1)p(y_2|x_2) = \frac{1}{2\pi\sigma_n^2} \exp\left\{\frac{-[\mathbf{y}-\mathbf{x}]^T[\mathbf{y}-\mathbf{x}]}{2\sigma_n^2}\right\} \tag{10}$$

and then letting $|\mathbf{x}|^2 = x_1^2 + x_2^2$, $x_1 = |\mathbf{x}|\cos\alpha$ and $x_2 = |\mathbf{x}|\sin\alpha$ i.e. $|\mathbf{x}|$ is the magnitude of the vector $[x_1 x_2]^T$ and α is the angle of this vector relative to the origin. Similarly $|\mathbf{y}|^2 = y_1^2 + y_2^2$, $y_1 = |\mathbf{y}|\cos\phi$ and $y_2 = |\mathbf{y}|\sin\phi$, where $|\mathbf{y}|$ is the magnitude of the vector $[y_1 y_2]^T$ and ϕ is the angle of this vector relative to the origin.

so that

$$p(y_1, y_2|x_1, x_2) = \frac{1}{2\pi\sigma_n^2} \exp\left\{\frac{-|\mathbf{y}|^2 - |\mathbf{x}|^2 - 2|\mathbf{x}||\mathbf{y}|\cos(\phi-\alpha)}{2\sigma_n^2}\right\}. \tag{11}$$

3.2. $\mathbf{x}$ *and* $\mathbf{V}$ *known*

In this case $\mathbf{V}$ rotates the original vector $\mathbf{x}_o$ through a known angle to a new, known $\mathbf{x}$ and we can treat this case with the probability density function (pdf)

$$p(y_1, y_2|x_1, x_2) = \frac{1}{\sqrt{(2\pi\sigma_n^2)}} \exp\left\{\frac{-(y_1-x_1)^2}{2\sigma_n^2}\right\} \frac{1}{\sqrt{(2\pi\sigma_n^2)}} \exp\left\{\frac{-(y_2-x_2)^2}{2\sigma_n^2}\right\} \tag{12}$$

and the differential entropy is

$$h(\mathbf{y}|\mathbf{x}) = -\int_0^\infty p(\mathbf{y}|\mathbf{x}) \ln p(\mathbf{y}|\mathbf{x}) d\mathbf{y} = \frac{1}{2}\ln(2\pi e\sigma_n^2) + \frac{1}{2}\ln(2\pi e\sigma_n^2) = \ln(2\pi e\sigma_n^2). \tag{13}$$

3.3. A *known,* $\mathbf{V}$ *unknown*

In this case $\mathbf{V}$ rotates the original vector $\mathbf{x}_o$ through an unknown angle γ so that $x_1 = A\cos\gamma$ and $x_2 = A\sin\gamma$, giving the pdf

$$p(y_1, y_2|A, \gamma) = \frac{1}{(2\pi\sigma_n^2)} \exp\left\{\frac{-|\mathbf{y}|^2 - A^2}{2\sigma_n^2}\right\} \exp\left\{\frac{A|\mathbf{y}|\cos(\phi-\gamma)}{\sigma_n^2}\right\}. \tag{14}$$

Now, with $\beta \triangleq \phi - \gamma$ and $p(\beta) = \frac{1}{2\pi}$,

$$p(y_1, y_2|A) = \int_0^{2\pi} p(y_1, y_2|A, \beta)p(\beta)d\beta = \frac{1}{2\pi\sigma_n^2} \exp\left\{\frac{-|y|^2 - A^2}{2\sigma_n^2}\right\} I_0\left(\frac{A|y|}{\sigma_n^2}\right). \tag{15}$$

At high enough SNR we may approximate the Bessel function as

$$I_0(Kx) \approx \frac{1}{\sqrt{Kx}} \exp\{Kx\}. \tag{16}$$

Therefore

$$\begin{aligned} p(y_1, y_2|A) &\approx \frac{\sigma_n}{2\pi\sigma_n^2\sqrt{(2\pi A|\mathbf{y}|)}} \exp\left\{\frac{-(|\mathbf{y}|-A)^2}{2\sigma_n^2}\right\} \\ &\approx \frac{1}{2\pi A} \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left\{\frac{-(|\mathbf{y}|-A)^2}{2\sigma_n^2}\right\}. \end{aligned} \tag{17}$$

and the differential entropy is

$$h(\mathbf{y}|A) = -\int_0^\infty p(\mathbf{y}|A) \ln p(\mathbf{y}|A) dy \approx \ln(2\pi A) + \frac{1}{2}\ln(2\pi e\sigma_n^2) \tag{18}$$

3.4. $\mathbf{x}$ *and* $\mathbf{V}$ *unknown*

In this case we assume that we only have knowledge of the variance of $\mathbf{x}$ and $\mathbf{n}$ and hence the variance of $\mathbf{y}$. With the components of both $\mathbf{x}$ and $\mathbf{n}$ treated as zero-mean Gaussian, then the components of $\mathbf{y}$ will also be zero-mean Gaussian with variance equal to the sum of the variances of $\mathbf{x}$ and $\mathbf{n}$ i.e. $y_i \sim \mathcal{N}(0, \sigma_y^2)$ where $\sigma_y^2 = \sigma_x^2 + \sigma_n^2$. The joint pdf for $\mathbf{y}$ is

$$p(y_1, y_2) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left\{\frac{-y_1^2}{2\sigma_y^2}\right\} \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left\{\frac{-y_2^2}{2\sigma_y^2}\right\} \tag{19}$$

which leads us to the differential entropy

$$h(\mathbf{y}) = -\int_0^\infty p(\mathbf{y}) \ln p(\mathbf{y}) dy = \frac{1}{2}\ln(2\pi e\sigma_y^2) + \frac{1}{2}\ln(2\pi e\sigma_y^2) = \ln(2\pi e\sigma_y^2). \tag{20}$$

3.5. Capacity

The fully informed mutual information was defined in equation (8) and so when both $\mathbf{x}$ and $\mathbf{V}$ are given, with Gaussian distributions for the source and noise, we have the fully informed capacity

$$C_{F_2} = \ln(2\pi e\sigma_y^2) - \ln(2\pi e\sigma_n^2) = \ln\left(\frac{\sigma_y^2}{\sigma_n^2}\right). \tag{21}$$

Similarly partially informed mutual information was defined in equation (9) so that, when the rotation matrix is unknown, we obtain the partially informed capacity

$$C_{P_2} \approx \ln(2\pi e\sigma_y^2) - \ln(2\pi A) - \frac{1}{2}\ln(2\pi e\sigma_n^2) = \ln\left(\frac{\sigma_y^2}{A\sigma_n}\right) + \frac{1}{2}\ln\left(\frac{e}{2\pi}\right) \tag{22}$$

In a similar fashion we may derive the entropies and mutual information for the 3D case. The derivation is given in Appendix A, where we find that

$$C_{F_3} = \ln\left(\frac{\sigma_y^3}{\sigma_n^3}\right). \tag{23}$$

and

$$C_{P_3} \approx \ln\left(\frac{\sigma_y^3}{A^2\sigma_n}\right) + \ln\left(\frac{e}{2}\right) \tag{24}$$

4. ND Capacity

4.1. High SNR Case

At high snr we found that the partially informed probability density functions factored into two parts:

$$\begin{aligned} \text{2D case: } p(\mathbf{y}|A) &\approx \left(\frac{1}{2\pi A}\right) \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left\{\frac{-(|\mathbf{y}| - A)^2}{2\sigma_n^2}\right\} \\ \text{3D case: } p(\mathbf{y}|A) &\approx \left(\frac{1}{4\pi A^2}\right) \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left\{\frac{-(|\mathbf{y}| - A)^2}{2\sigma_n^2}\right\}. \end{aligned} \tag{25}$$

The first part appears to have the form of a uniform density on the surface of an N-dimensional sphere. The second part appears to represent a Gaussian distribution across an N-dimensional shell. Therefore $p(\mathbf{y}|A)$ may be viewed as an N-dimensional, variable density, shell with mean radius A. From Wikipedia ("Sphere") [8] the general equations for the surface area and volume of an N-dimensional sphere, with radius $A = \sqrt{\sum_i^N x_i^2}$, are given by:

$$\text{Surface Area} = \frac{2\pi^{\frac{N}{2}} A^{N-1}}{\Gamma\left(\frac{N}{2}\right)} \tag{26}$$

and

$$\text{Volume} = \frac{2\pi^{\frac{N}{2}} A^{N}}{N\Gamma\left(\frac{N}{2}\right)} \tag{27}$$

Thus the required N-D, high SNR entropies, may be written as:

$$h(\mathbf{y}|\mathbf{x}) = \frac{N}{2}\ln(2\pi e\sigma_n^2), \tag{28}$$

$$h(\mathbf{y}|A) \approx \ln\left(\frac{2\pi^{\frac{N}{2}} A^{N-1}}{\Gamma\left(\frac{N}{2}\right)}\right) + \frac{1}{2}\ln\left(2\pi e\sigma_n^2\right), \tag{29}$$

$$h(\mathbf{y}) = \frac{N}{2}\ln(2\pi e\sigma_y^2). \tag{30}$$

The densities $p(\mathbf{y})$ and $p(\mathbf{y}|\mathbf{x})$ could be pictured as N-dimensional, probability spheres. Hence the fully informed capacity becomes the difference in entropy between an N-dimensional probability sphere, representing the signal plus noise vector distribution, and an N-dimensional sphere, representing the noise vector distribution. In the partially informed case the capacity becomes the difference in entropy between an N-dimensional probability sphere, representing the signal plus noise vector distribution, and an N-dimensional probability shell, representing the amplitude known plus noise distribution. The ND fully informed capacity may be written as

$$C_{F_N} = h(\mathbf{y}) - h(\mathbf{y}|\mathbf{x}) = \frac{N}{2}\ln\left(\frac{\sigma_y^2}{\sigma_n^2}\right) \tag{31}$$

and the partially informed capacity may be approximated by

$$C_{P_N} \approx h(\mathbf{y}) - h(\mathbf{y}|A) = \frac{1}{2}\ln\left(\frac{\sigma_y^2}{\sigma_n^2}\right) + \frac{1}{2}\ln\left(\pi^{-1}2^{N-3}e^{N-1}\right) + \ln\left(\Gamma\left(\frac{N}{2}\right)\right).$$

Defining the signal-to-noise ratio as $\rho = \frac{A^2}{\sigma_n^2}$ and with $\sigma_y^2 = \sigma_x^2 + \sigma_n^2 = A^2 + \sigma_x^2$, then the capacities may be expressed as

$$C_{F_N} \approx \frac{N}{2}\ln(1+\rho) \tag{32}$$

$$C_{P_N} \approx \frac{1}{2}\ln(1+\rho) + \frac{1}{2}\ln\left(\pi^{-1}2^{N-3}e^{N-1}\right) + \ln\left(\Gamma\left(\frac{N}{2}\right)\right). \tag{33}$$

4.2. General Case

In this section we derive the general form for $p(\mathbf{y}|A)$ thus allowing us to obtain the capacity for any dimension and snr. The derivation utilises a result by Vesely [9] that shows how the "mass" of an N dimensional spherical shell is distributed along one sphere axis. This result greatly simplifies the multidimensional integrals that we require to solve. The surface area, $S_N(r_0)$, of an N dimensional sphere, as a function of radius, may be represented as

$$S_N(r_0) = \int_{-r_0}^{r_0} \frac{r_0 S_{N-1}(r_2)}{r_2} dr_1 \tag{34}$$

where $r_2 = \sqrt{r_0^2 - r_1^2}$. We can rewrite the above as

$$1 = \int_{-r_0}^{r_0} \frac{r_0 S_{N-1}(r_2)}{r_2 S_N(r_0)} dr_1 = \int_{-r_0}^{r_0} p_N(r_1) dr_1 \tag{35}$$

which shows how the "mass"of the shell is distributed along the single sphere axis r_1.

$$p_N(r_1) = \frac{r_0 S_{N-1}(r_2)}{r_2 S_N(r_0)} = \frac{(N-1)C_{N-1} r_2^{N-3}}{N C_N r_0^{N-2}} = \frac{(N-1)C_{N-1}}{N C_N} \frac{1}{r_0} \left[1 - \frac{r_1^2}{r_0^2}\right]^{\frac{N-3}{2}} \tag{36}$$

where

$$C_N = \frac{2\pi^{N/2}}{N\Gamma(N/2)}. \tag{37}$$

The integrals that we are dealing with take the form

$$\begin{aligned} p(\mathbf{y}|\mathbf{x}) &= (2\pi\sigma_n^2)^{-N/2} \exp\left\{\frac{-|\mathbf{y}|^2 - |\mathbf{x}|^2}{2\sigma_n^2}\right\} \exp\left\{\frac{\sum_{i=1}^N x_i y_i}{\sigma_n^2}\right\} \\ &= (2\pi\sigma_n^2)^{-N/2} \exp\left\{\frac{-|\mathbf{y}|^2 - |\mathbf{x}|^2}{2\sigma_n^2}\right\} \exp\left\{\frac{\mathbf{x} \cdot \mathbf{y}}{\sigma_n^2}\right\} \end{aligned} \tag{38}$$

from which we wish to obtain $p(\mathbf{y}|\,|\mathbf{x}|)$. Assuming now that $|\mathbf{x}|$ is given we have

$$p(\mathbf{y}|\mathbf{x}, |\mathbf{x}|) = (2\pi\sigma_n^2)^{-N/2} \exp\left\{\frac{-|\mathbf{y}|^2 - |\mathbf{x}|^2}{2\sigma_n^2}\right\} \exp\left\{\frac{\mathbf{x} \cdot \mathbf{y}}{\sigma_n^2}\right\} \tag{39}$$

and so to obtain $p(\mathbf{y}|\,|\mathbf{x}|)$ we must integrate over the x_i as follows

$$\begin{aligned} p(\mathbf{y}|\,|\mathbf{x}|) &= \int_{|\mathbf{x}|=A} p(\mathbf{y}|\mathbf{x}, |\mathbf{x}|) p(\mathbf{x}) d\mathbf{x} \\ &= (2\pi\sigma_n^2)^{-N/2} \exp\left\{\frac{-|\mathbf{y}|^2 - |\mathbf{x}|^2}{2\sigma_n^2}\right\} \int_{|\mathbf{x}|=A} \exp\left\{\frac{\mathbf{x} \cdot \mathbf{y}}{\sigma_n^2}\right\} p(\mathbf{x}) d\mathbf{x} \end{aligned} \tag{40}$$

We proceed to calculate this integral by first noting that the x_i are uniformly distributed over the surface of an N-dimensional sphere and we only need to perform the integral along a single dimension, e.g. x_1 and replace $p(\mathbf{x})$ with $p(x_1)$ using the results derived earlier. To better understand this, consider the dot product $\mathbf{x} \cdot \mathbf{y}$. The dot product will be unchanged if both vectors are operated on by the same rotation matrix. Let the rotation matrix be $\mathcal{M} \in R^{N \times N}$, then

$$(\mathcal{M}\mathbf{x}) \cdot (\mathcal{M}\mathbf{y}) = (\mathcal{M}\mathbf{x})^T (\mathcal{M}\mathbf{y}) = \mathbf{x}^T \mathcal{M}^T \mathcal{M} \mathbf{y} = \mathbf{x}^T \mathbf{y} = \mathbf{x} \cdot \mathbf{y} \tag{41}$$

since $\mathcal{M}\mathcal{M}^T = \mathcal{M}\mathcal{M}^{-1} = I$. So we are free to choose any rotation matrix and the integral will be unaffected. Let us choose $\mathcal{M}$ such that $\mathcal{M}\mathbf{y} = |\mathbf{y}|[1, 0, \ldots, 0] = |\mathbf{y}|\mathbf{e}$, where $\mathbf{e}$ is a unit vector, i.e. the vector $\mathbf{y}$ is rotated to lie along the y_1 axis. Let $\mathbf{x}' = (\mathcal{M}\mathbf{x})$ then we have

$$\mathbf{x}' \centerdot (\mathcal{M}\mathbf{y}) = \mathbf{x}' \centerdot |\mathbf{y}|\mathbf{e} = |\mathbf{y}|(\mathbf{x}')^T\mathbf{e} = |\mathbf{y}|x_1'. \tag{42}$$

Hence, with $|\mathbf{x}| = A$,

$$\begin{aligned}\int_{|\mathbf{x}|=A} \exp\left\{\frac{\mathbf{x}\centerdot\mathbf{y}}{\sigma_n^2}\right\} p(\mathbf{x})d\mathbf{x} &= \int_{-A}^{A} p_N(x_1') \exp\left\{\frac{|\mathbf{y}|x_1'}{\sigma_n^2}\right\} \\ &= \frac{(N-1)C_{N-1}}{NC_N}\frac{1}{A}\int_{-A}^{A}\left[1-\frac{x_1'^2}{A^2}\right]^{\frac{N-3}{2}} \exp\left\{\frac{|\mathbf{y}|x_1'}{\sigma_n^2}\right\} dx_1'\end{aligned} \tag{43}$$

We may make a change of variable by letting $z = \frac{x_1'}{A}$ to get

$$\int_{|\mathbf{x}|=A} \exp\left\{\frac{\mathbf{x}\centerdot\mathbf{y}}{\sigma_n^2}\right\} p(\mathbf{x})d\mathbf{x} = \frac{(N-1)C_{N-1}}{NC_N}\int_{-1}^{1}\left[1-z^2\right]^{\frac{N-3}{2}} \exp\left\{\frac{|\mathbf{y}|Az}{\sigma_n^2}\right\} dz \tag{44}$$

The general form for the density, given A, is therefore

$$p(\mathbf{y}|A) = (2\pi\sigma_n^2)^{-\frac{N}{2}} \exp\left\{\frac{-A^2-|\mathbf{y}|^2}{2\sigma_n^2}\right\}\frac{(N-1)C_{N-1}}{NC_N}\int_{-1}^{1}\left[1-z^2\right]^{\frac{N-3}{2}} \exp\left\{\frac{|\mathbf{y}|Az}{\sigma_n^2}\right\} dz. \tag{45}$$

The entropy calculation involves a multidimensional integration over the components in $\mathbf{y}$:

$$h(\mathbf{y}|A) = -\int_{\mathbf{y}} p(\mathbf{y}|A)\ln p(\mathbf{y}|A)d\mathbf{y}. \tag{46}$$

With the general form for the differential entropies we are now able to derive the capacity for both the fully informed cases and the partially informed (amplitude only) cases. The capacity for dimensions two to five have been calculated for both cases and the results are presented in figures 5 and 6. Comparing the two figures we note that the partially informed curves have a smaller slope than their fully informed counterparts. If both receivers were operating with the same snr then we could also make the observation that the partially informed values are always less than their fully informed counterparts. However, as indicated in section 2 earlier, due to the channel matrix inversion required by Eve and a possibly different local (local to the receivers) noise environment, this may not be the case.

5. Summary

The problem of determining the information intercept capacity, available to a receiving system which knows its channel matrix but has no prior knowledge of a unitary transformation that has been applied at the transmitter, has been analysed. Entropy derivations were carried out for two dimensions and three dimensions giving some insight to the general dimensional, high snr case. The exact capacity for the N-Dimensional case has been obtained but requires numerical integration to derive the differential entropy for the partially informed case.

Figure 5. Capacity: fully informed Vs SNR.

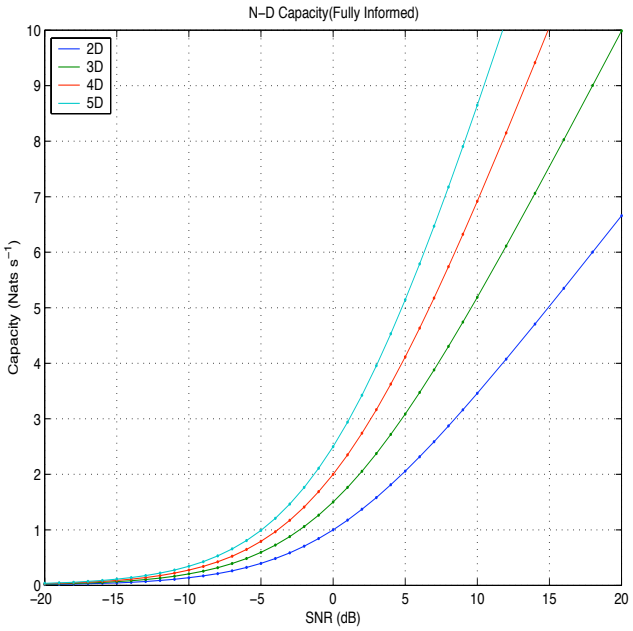
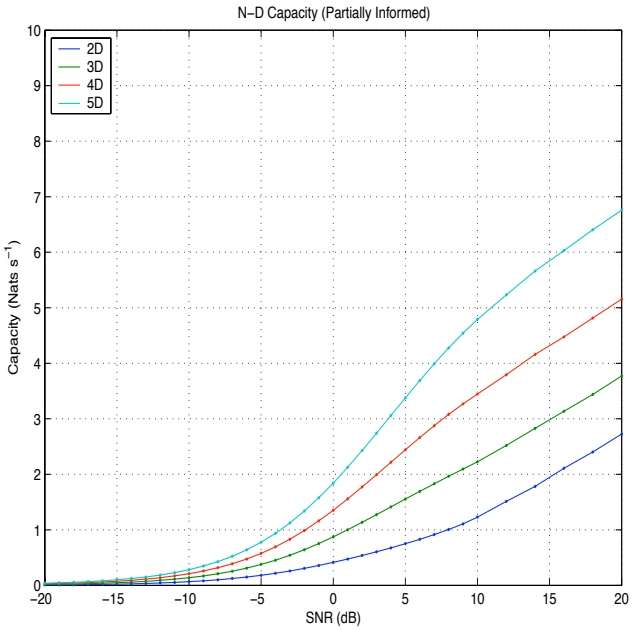


Figure 6. Capacity: partially informed Vs SNR.



The fully informed capacity has been likened to the difference in entropy between two N-dimensional probability spheres: the larger sphere, representing the distribution of the signal plus noise vector, and the smaller sphere, representing the distribution of the noise vector. At high snr, the partially informed capacity was found to be equal to the difference in entropy between an N-dimensional probability sphere, representing the distribution of the signal plus noise vector, and an N-dimensional probability shell, representing the distribution of the amplitude plus noise vector.

Acknowledgements

We would like to thank the anonymous reviewers for their observations and helpful suggestions, which improved the original manuscript.

References and Notes

1. Foschini, G.J.; Gans, M.J. On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas, *Wireless Pers. Commun.* **1998**, *6*, 311-335.
2. Telatar, E. Capacity of Multi-antenna Gaussian Channels, AT&T-Bell Lab. *Internal Tech. Memo.* **1995**.
3. Comon, P. Independent Component Analysis, A new concept?, *Signal Process.* **1994**, *36*, 287-314.
4. Maurer, U.M. Secret Key Agreement by Public Discussion From Common Information, *IEEE Trans. Inform. Theory* **1993**, *39*, 733-742.
5. Tse, D.;Viswanath, P. *Fundamentals of Wireless Communication*; Cambridge University Press: U.K., 2005.
6. Papoulis, A. *Probability, Random Varables, and Stochastic Processes*; McGraw-Hill: 1989.

7. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; John Wiley & Sons, Inc.: 1991.
8. Wikipedia, *N-sphere — Wikipedia, The Free Encyclopedia*, 2008, http://en.wikipedia.org/wiki/N-sphere.
9. Vesely, F. *From Hyperspheres to Entropy*, http://homepage.univie.ac.at/franz.vesely/sp.english/sp/sp.html.
10. Prudnikov, A.P.; Brychkov, Yu.A.; Marichev, O.I. *Integrals and Series*; Gordon and Breach Sci. Publ.: New York, 1986; 464, .
11. Marzetta, T.L.; Hochwald, B.M. Capacity of a Mobile Multiple-Antenna Communication Link in Rayleigh Flat Fading, *IEEE Trans. Inform. Theory* **1999**; *45*, 139-157.

Appendix: Derivation of 3D Mutual Information

We can construct the joint probability density function in the following manner. Beginning with

$$p(y_1, y_2, y_3|x_1, x_2, x_3, V) = \frac{1}{(2\pi\sigma_n^2)^{\frac{3}{2}}} \exp\left\{\frac{-[\mathbf{y}-\mathbf{V}\mathbf{x}]^T[\mathbf{y}-\mathbf{V}\mathbf{x}]}{2\sigma_n^2}\right\} \tag{47}$$

x *known*

In the case where **x** is known after the transformation the pdf is given by

$$p(y_1, y_2, y_3|x_1, x_2, x_3) = \frac{\exp\left\{\frac{-(y_1-x_1)^2}{2\sigma_n^2}\right\}}{\sqrt{(2\pi\sigma_n^2)}} \frac{\exp\left\{\frac{-(y_2-x_2)^2}{2\sigma_n^2}\right\}}{\sqrt{(2\pi\sigma_n^2)}} \frac{\exp\left\{\frac{-(y_3-x_3)^2}{2\sigma_n^2}\right\}}{\sqrt{(2\pi\sigma_n^2)}} \tag{48}$$

and the entropy is

$$h(\mathbf{y}|\mathbf{x}) = -\int_0^\infty p(\mathbf{y}|\mathbf{x}) \ln p(\mathbf{y}|\mathbf{x}) dy = \frac{1}{2}\ln(2\pi e\sigma_n^2) + \frac{1}{2}\ln(2\pi e\sigma_n^2) + \frac{1}{2}\ln(2\pi e\sigma_n^2) = \frac{3}{2}\ln(2\pi e\sigma_n^2). \tag{49}$$

A *known,* α, β *unknown*

For the vector $[x_1 x_2 x_3]^T$ there are two rotation angles to consider: α, β and so, with

$$\begin{aligned} |x|^2 &= A^2 \\ x_1 &= A\sin\alpha\cos\beta \\ x_2 &= A\sin\alpha\sin\beta \\ x_3 &= A\cos\alpha, \end{aligned} \tag{50}$$

we have the joint probability of the two angles $p(\alpha, \beta) = \frac{\sin\alpha}{4\pi}$. Therefore $p(\mathbf{y}|\mathbf{x}) \rightarrow p(\mathbf{y}|A, \alpha, \beta)$ becomes

$$\begin{aligned} p(\mathbf{y}|\mathbf{x}) &= \frac{1}{(2\pi\sigma_n^2)^{\frac{3}{2}}} \exp\left\{\frac{-|\mathbf{y}|^2 - |\mathbf{x}|^2}{2\sigma_n^2}\right\} \exp\left\{\frac{x_1y_1 + x_2y_2 + x_3y_3}{\sigma_n^2}\right\} \\ p(\mathbf{y}|A, \alpha, \beta) &= \frac{1}{(2\pi\sigma_n^2)^{\frac{3}{2}}} \exp\left\{\frac{-|\mathbf{y}|^2 - A^2}{2\sigma_n^2}\right\} \exp\left\{\frac{A}{\sigma_n^2}[\sin\alpha\cos\beta y_1 + \sin\alpha\sin\beta y_2 + \cos\alpha y_3]\right\} \end{aligned} \tag{51}$$

The integral

$$p(y|A) = \int_0^{2\pi} \int_0^{\pi} p(y|A,\alpha,\beta)p(\alpha,\beta)d\alpha d\beta \tag{52}$$

is obtained by using the form given in Prudnikov et al [10] :

$$\int_0^{2\pi} \int_0^{\pi} \sin\alpha \exp\left\{\frac{A}{\sigma_n^2}[\sin\alpha\cos\beta y_1 + \sin\alpha\sin\beta y_2 + \cos\alpha y_3]\right\} d\alpha d\beta = \frac{2\pi\sigma_n^2}{A|\mathbf{y}|}\exp\left\{\frac{A|\mathbf{y}|}{\sigma^2}\right\} \tag{53}$$

and so

$$p(\mathbf{y}|A) = \frac{1}{4\pi A|\mathbf{y}|}\frac{1}{\sqrt{2\pi\sigma_n^2}}\exp\left\{\frac{-(|\mathbf{y}|-A)^2}{2\sigma_n^2}\right\}, \tag{54}$$

which may be approximated, at high SNR, as:

$$p(\mathbf{y}|A) \approx \frac{1}{4\pi A^2}\frac{1}{\sqrt{2\pi\sigma_n^2}}\exp\left\{\frac{-(|\mathbf{y}|-A)^2}{2\sigma_n^2}\right\}, \tag{55}$$

The differential entropy is

$$h(\mathbf{y}|A) = -\int_0^{\infty} p(\mathbf{y}|A)\ln p(\mathbf{y}|A)dy \approx \ln(4\pi A^2) + \frac{1}{2}\ln(2\pi e\sigma_n^2) \tag{56}$$

x *unknown*

In this case we assume that we only have knowledge of the variance of $\mathbf{x}$ and $\mathbf{n}$ and hence the variance of $\mathbf{y}$. With both $\mathbf{x}$ and $\mathbf{n}$ treated as zero-mean Gaussian, then $\mathbf{y}$ will also be zero-mean Gaussian with variance equal to the sum of the variances of $\mathbf{x}$ and $\mathbf{n}$ i.e. $y_i \sim \mathcal{N}(0,\sigma_y^2)$ where $\sigma_y^2 = \sigma_x^2 + \sigma_n^2$.

$$p(y_1,y_2,y_3) = \frac{\exp\left\{\frac{-y_1^2}{2\sigma_y^2}\right\}}{\sqrt{2\pi\sigma_y^2}}\frac{\exp\left\{\frac{-y_2^2}{2\sigma_y^2}\right\}}{\sqrt{2\pi\sigma_y^2}}\frac{\exp\left\{\frac{-y_3^2}{2\sigma_y^2}\right\}}{\sqrt{2\pi\sigma_y^2}} \tag{57}$$

giving the differential entropy as

$$h(\mathbf{y}) = -\int_0^{\infty} p(\mathbf{y})\ln p(\mathbf{y})dy = \frac{1}{2}\ln(2\pi e\sigma_y^2) + \frac{1}{2}\ln(2\pi e\sigma_y^2) + \frac{1}{2}\ln(2\pi e\sigma_y^2) = \frac{3}{2}\ln(2\pi e\sigma_y^2). \tag{58}$$

Capacity

$$\text{Define} \quad I_F \triangleq h(\mathbf{y}) - h(\mathbf{y}|\mathbf{x}) \tag{59}$$

$$\text{and} \quad I_P \triangleq h(\mathbf{y}) - h(\mathbf{y}|\mathbf{A}) \tag{60}$$

where I_F is the mutual information in the best case where both $\mathbf{x}$ and $\mathbf{V}$ are given. I_P is the mutual information when the rotation matrix is unknown. Since the source and noise distributions are Gaussian, and assuming constant source variance, we then obtain the fully informed and partially informed capacities as

$$C_{F_3} = \frac{3}{2}\ln(2\pi e\sigma_y^2) - \frac{3}{2}\ln(2\pi e\sigma_n^2) = \ln\left(\frac{\sigma_y^3}{\sigma_n^3}\right) \tag{61}$$

and

$$C_{P_3} = \frac{3}{2}\ln(2\pi e\sigma_y^2) - \ln(4\pi A^2) - \frac{1}{2}\ln(2\pi e\sigma_n^2) = \ln\left(\frac{\sigma_y^3}{A^2\sigma_n}\right) + \ln\left(\frac{e}{2}\right) \tag{62}$$