ISIT 2009, Seoul, Korea, June 28 - July 3, 2009

Physical Modeling of Communication Systems in Information Theory

Michel T. Ivrlač and Josef A. Nossek

Institute for Circuit Theory and Signal Processing
Technische Universität München, D-80333 Munich, Germany
email: {ivrlac, nossek}@tum.de

***Abstract*—It is common in information theoretic channel models to rely on the average squared Euclidean norm of the channel input as being proportional to transmit power. Likewise, it is common to assume noise that is additive, Gaussian, and white. It is a legitimate question to ask, whether such a modeling approach has enough degrees of freedom to capture the physical constraints that are imposed on implementations of a communication system. In this paper, we show that in many, though not all, situations it is indeed possible to obtain a complete physical model, while nevertheless sticking with average squared Euclidean norm as power, and white Gaussian noise. Our systematic approach works in two steps. First, all channel inputs and outputs are replaced by ports, which are defined by two conjugated variables (like voltage and current). By this multi-port modeling approach, we can obtain a complete physical model. Secondly, we introduce linear transformations between the inputs and outputs of the information theoretic channel model on the one hand, and the physical inputs and physical outputs of the communication system, on the other. This approach gives us enough degrees of freedom to obtain a complete information theoretic model, which correctly reflects the physical constraints that are imposed upon the communication system by its environment. We apply the proposed approach to a multi-antenna communication system, and show that it is indeed possible that the channel capacity of multi-antenna systems can grow super-linearly with the number of antennas for large signal to noise ratios.**

I. Introduction

Information theoretic channel models comprise three parts: a transfer model, a noise model, and a power (or energy) model. The most prominent one is the transfer model, which relates the channel input with the noiseless channel output. Most of the channel modeling literature is concerned with this input-output relationship [1]. In the most widely used noise model, the noise is additive, Gaussian, and white. Finally, power (or energy) is usually used to describe the costs for the exchange of information, such that information theoretic measures, like channel capacity or cut-off rate are computed with a power constraint. It is all common to model power as the average squared Euclidean norm of the channel input [2].

Because such information theoretic channel models should be applicable for a variety of physical implementations, there is the legitimate question on how the constraints that are imposed by the particular physical environment are represented in the model. Clearly, it should make a difference whether the communication is based on a multi-antenna wireless system, a number of microphones and loudspeakers, or takes place over a bunch of wires in a cable or a multi-conductor bus on a chip. Is, for instance, the transmit power correctly captured by the information theoretic model for all the different physical environments? To see the problem more clearly, consider that, from a physics point of view, power can only by calculated from the (inner) product of *two* conjugated variables, like voltage and current[1] [3]. In general, power cannot be calculated by taking the squared magnitude of just one of the conjugated variables. Hence, the average squared Euclidean norm of the channel input is not at all guaranteed to have any relationship with physical power.

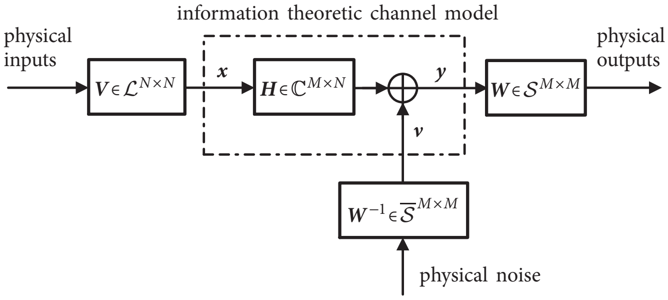

Figure 1: Bijective mappings (via V, and W) of the physical inputs and physical outputs, to the inputs and outputs of the information theoretic channel.

One way to handle this problem, is to replace the average squared Euclidean norm of the channel input, with a more elaborate power model that better reflects the physics [4]–[6]. However, this approach has the disadvantage that established and well-known results from information theory and signal processing cannot be directly applied anymore. For instance, the Waterfilling algorithm [7], has to be replaced by a »modified Waterfilling« algorithm, as in [4], which reflects the modified power model in use.

In this paper, we present a different approach. Our goal is to stick with the average squared Euclidean norm of the channel input as the transmit power, and also stick with the model of additive and *white* Gaussian noise. Because both the noise model and the power model are the same for different physical scenarios, the particular physical constraints have to be captured solely by the transfer model. At this point we encounter a problem that seems to be deep. In order to see why, let us consider a linear channel with N inputs, and M outputs. The noiseless input-output relationship is specified by the $M \times N$ channel matrix, with $M \cdot N$ components. On the other hand, in a physical modeling approach, each input and output is replaced by a *port*, which is defined by *two* conjugated variables (say voltage and current). Therefore, the linear channel requires for its complete (physical) description, a much larger $(M + N) \times (M + N)$ matrix, that relates one half of the port variables (for example, the voltages) with the other half of the port variables (for instance, the currents). Since

[1]Other examples of conjugated variables are: force and velocity (or torque and angular velocity) in mechanics, pressure and volume-flow in acoustics, temperature and entropy-flow in thermodynamics.

978-1-4244-4313-0/09/$25.00 ©2009 IEEE

$M \cdot N < (M+N)^2$, it is clear that the channel matrix simply does *not* have enough degrees of freedom to capture the complete physical model. In a *unilateral* channel, i.e., in a channel where the receiver does not influence the transmitter, we see in Section II-D, that the number of degrees of freedom for a complete physical description reduces to $M^2 + N^2 + MN$, which is still too large to be captured by the channel matrix. Therefore, the question is legitimate, whether it is possible to achieve our goal to obtain a complete model by proper definition of the transfer model alone. Fortunately, the answer is a conditional »yes«. In many, and practically relevant cases, though not in all cases, we can achieve our goal.

The key idea is to recognize that the $M \cdot N$ degrees of freedom offered by the channel matrix are not the only degrees of freedom that we have inside the transfer model. The reason is that we can have some mapping between the input and the output of the information theoretic channel on the one hand, and the physical input and the physical output variables, on the other. The only restriction to this mapping is, that it has to be bijective such that it preserves information. Hence, there is an $M \times M$ matrix, and an $N \times N$ matrix at our disposal, which define the mapping between the physical and the information theoretic channel inputs and outputs. Therefore, the true number of degrees of freedom that we can make use of in the transfer model equals $M^2 + N^2 + MN$, which is the same as the number of degrees of freedom of the complete physical description of the unilateral linear channel. In case of the bilateral channel, we also have enough degrees of freedom, if we supply a second channel matrix, which is to be used for the reverse direction of information flow.

Figure 1 illustrates this mapping process, where the $N \times N$ matrix $\boldsymbol{V}$ defines the mapping from the physical inputs to the inputs of the information theoretic channel, while the $M \times M$ matrix $\boldsymbol{W}$ defines the mapping from the outputs of the information theoretic channel to the physical outputs. Finally, the channel matrix $\boldsymbol{H} \in \mathbb{C}^{M \times N}$, defines the noiseless input-output relationship of the information theoretic channel.

In this paper, we present a systematic way to determine the three matrices $\boldsymbol{V}$, $\boldsymbol{W}$, and $\boldsymbol{H}$, such that both the input-output relationship, and the receiver noise of the *physical* communication system are captured by the information theoretic channel model, and that the physical transmit power is equal to the average squared Euclidean norm of the channel input $\boldsymbol{x}$. If obtaining the channel capacity is all we are looking for, it is already sufficient to know the channel matrix $\boldsymbol{H}$. However, when we compute signal processing solutions (e.g., beamforming vectors) within the information theoretic channel model, we need the matrix $\boldsymbol{V}^{-1}$, in order to obtain the physical channel input (for Tx-processing), and the matrix $\boldsymbol{W}^{-1}$, in order to obtain the channel output $\boldsymbol{y}$, of the information theoretic model from the physical channel output (Rx-processing).

We illustrate the proposed approach by showing an application to a multi-antenna communication system, where the physical antenna coupling of closely spaced antennas is taken care of by our modeling approach. We will see that bringing the physical constraints into information theory, can lead to new insights. Especially, it turns out that it is indeed possible for a multi-antenna system, to achieve a channel capacity that can grow *super-linearly* with the number of antennas, for large signal to noise ratio.

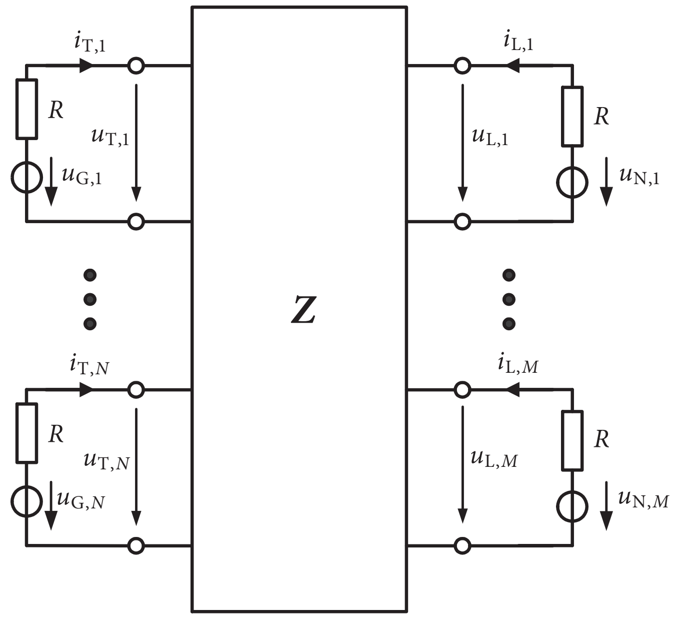

Figure 2: Circuit theoretic multi-port model of a linear multi-input multi-output communication system, with signal voltage generators, load terminations, and noise voltage sources.

II. Multi-Port Model

Let us now develop a physical model of a communication system with N inputs, and M outputs. Every input and output is replaced by a *port*, which is defined by two conjugated variables: voltage envelopes, and current envelopes. The resulting $(M+N)$-port is shown in Figure 2. The input signals are supplied by N voltage generators, which are modeled as ideal voltage sources with a series resistance R. The voltage envelopes $u_{G,n}$, with $n \in \{1, 2, \ldots, N\}$, contain the information that has to be transfered over the channel. In practice, we can view the voltage generators as a model for the high-power amplifiers located at the transmitting end of the channel. The voltage generators are connected to the first N ports of the multi-port. To the remaining M ports, we connect termination resistances R, that are put in series with M ideal voltage sources, which deliver the voltage envelopes $u_{N,m}$, with $m \in \{1, 2, \ldots, M\}$. These voltage envelopes model the thermal noise that is generated inside the termination resistances. This combination of a noiseless resistance, and a series noise voltage source, is used here to model the input of the receive amplifier. The voltage envelopes $u_{L,m}$, with $m \in \{1, 2, \ldots, M\}$, which appear at the M ports at the receiving end of the channel, are used as the *physical outputs*, while the N generator voltage envelopes $u_{G,n}$, with $n \in \{1, 2, \ldots, N\}$, are used as the *physical inputs* of the communication system.

Let us proceed by collecting all the voltage envelopes $u_{T,n}$, at the ports belonging to the transmitting end of the channel, into the vector $\boldsymbol{u}_T = \left[\, u_{T,1} \; \cdots \; u_{T,N} \,\right]^T \in \mathbb{C}^{N \times 1} \cdot \mathrm{V}$, and similarly let $\boldsymbol{u}_L = \left[\, u_{L,1} \; \cdots \; u_{L,M} \,\right]^T \in \mathbb{C}^{M \times 1} \cdot \mathrm{V}$, be defined as the vector of the voltage envelopes of the ports at the receiving end of the channel. We use the super-script $^{\mathrm{T}}$, to denote the transpose of a matrix or vector, while the symbol V, represents the physical unit »Volt«. Similarly, the corresponding conjugated variables, that is, the current envelopes $i_{\mathrm{T},n}$, and $i_{\mathrm{L},m}$, are conveniently collected into the vectors $\boldsymbol{i}_{\mathrm{T}} = \begin{bmatrix} i_{\mathrm{T},1} & \cdots & i_{\mathrm{T},N} \end{bmatrix}^{\mathrm{T}} \in \mathbb{C}^{N\times 1}\cdot \mathrm{A}$, and $\boldsymbol{i}_{\mathrm{L}} = \begin{bmatrix} i_{\mathrm{L},1} & \cdots & i_{\mathrm{L},M} \end{bmatrix}^{\mathrm{T}} \in \mathbb{C}^{M\times 1}\cdot \mathrm{A}$, respectively, where we use the symbol A, to denote the physical unit »Ampere«.

A. Noisy Input-Output Relationship

A complete description of the multi-port is obtained, when one half of the port variables are related with the other half. In this paper, we describe the multi-port by expressing the port voltage envelopes as a function of the port current envelopes:

$$\begin{bmatrix} \boldsymbol{u}_{\mathrm{T}} \\ \boldsymbol{u}_{\mathrm{L}} \end{bmatrix} = \boldsymbol{Z} \begin{bmatrix} \boldsymbol{i}_{\mathrm{T}} \\ \boldsymbol{i}_{\mathrm{L}} \end{bmatrix}, \tag{1}$$

where $\boldsymbol{Z} \in \mathbb{C}^{(M+N)\times(M+N)}\cdot \Omega$, denotes the *impedance matrix* of the multi-port [8], and the symbol Ω, denotes the physical unit »Ohm«, respectively. Kirchhoff's voltage law [8], yields $\boldsymbol{u}_{\mathrm{L}} = \boldsymbol{u}_{\mathrm{N}} - R\boldsymbol{i}_{\mathrm{L}}$, with $\boldsymbol{u}_{\mathrm{N}} = \begin{bmatrix} u_{\mathrm{N},1} & \cdots & u_{\mathrm{N},M} \end{bmatrix}^{\mathrm{T}} \in \mathbb{C}^{M\times 1}\cdot \mathrm{V}$, as the vector of the noise voltage envelopes. Similarly: $\boldsymbol{u}_{\mathrm{T}} = \boldsymbol{u}_{\mathrm{G}} - R\boldsymbol{i}_{\mathrm{T}}$, where $\boldsymbol{u}_{\mathrm{G}} = \begin{bmatrix} u_{\mathrm{G},1} & \cdots & u_{\mathrm{G},N} \end{bmatrix}^{\mathrm{T}} \in \mathbb{C}^{N\times 1}\cdot \mathrm{V}$, is the vector of the generators' voltage envelopes. When we substitute these two expression for $\boldsymbol{u}_{\mathrm{L}}$, and $\boldsymbol{u}_{\mathrm{T}}$, into (1), and solve for $\boldsymbol{u}_{\mathrm{L}}$, we obtain the following noisy input-output relationship:

$$\boldsymbol{u}_{\mathrm{L}} = \boldsymbol{E}\boldsymbol{u}_{\mathrm{G}} + \underbrace{\boldsymbol{F}\boldsymbol{u}_{\mathrm{N}}}_{\sqrt{R}\cdot\boldsymbol{\eta}}, \tag{2}$$

which expresses the physical channel output $\boldsymbol{u}_{\mathrm{L}} \in \mathbb{C}^{M\times 1}\cdot \mathrm{V}$, as a function of the physical channel input $\boldsymbol{u}_{\mathrm{G}} \in \mathbb{C}^{N\times 1}\cdot \mathrm{V}$, and the physical noise $\boldsymbol{\eta} \in \mathbb{C}^{M\times 1}\cdot\sqrt{\mathrm{W}}$, where the symbol W, denotes the physical unit »Watt«. Herein,

$$\boldsymbol{E} = -\widetilde{\boldsymbol{\Gamma}}\left(\mathbf{I}_{M+N} + R^{-1}\boldsymbol{Z}\right)^{-1}\boldsymbol{\Gamma}^{\mathrm{T}} \in \mathbb{C}^{M\times N}, \tag{3}$$

$$\boldsymbol{F} = \mathbf{I}_M - \widetilde{\boldsymbol{\Gamma}}\left(\mathbf{I}_{M+N} + R^{-1}\boldsymbol{Z}\right)^{-1}\widetilde{\boldsymbol{\Gamma}}^{\mathrm{T}} \in \mathbb{C}^{M\times M}, \tag{4}$$

and

$$\boldsymbol{\Gamma} = \begin{bmatrix} \mathbf{I}_N & \mathbf{O}_{N\times M} \end{bmatrix} \in \{0,1\}^{N\times(M+N)}, \tag{5}$$

$$\widetilde{\boldsymbol{\Gamma}} = \begin{bmatrix} \mathbf{O}_{M\times N} & \mathbf{I}_M \end{bmatrix} \in \{0,1\}^{M\times(M+N)}, \tag{6}$$

where $\mathbf{I}_k$, denotes the $k\times k$ identity matrix, and $\mathbf{O}_{p\times q}$ denotes the $p\times q$ zero matrix. Note that (3), and (4) can be written in many different ways, e.g., $\boldsymbol{F} = \widetilde{\boldsymbol{\Gamma}}\left(\mathbf{I}_{M+N} + R\boldsymbol{Z}^{-1}\right)^{-1}\widetilde{\boldsymbol{\Gamma}}^{\mathrm{T}}$, is another way to express $\boldsymbol{F}$, provided that $\boldsymbol{Z}^{-1}$ exists.

B. Transmit Power

It makes sense to define as the »transmit power«, the noiseless physical power that flows through the N ports of the transmitting end of the channel into the multi-port:

$$P_{\mathrm{Tx}} = \mathrm{E}\left[\mathrm{Re}\left\{\boldsymbol{u}_{\mathrm{T}}^{\mathrm{H}}\boldsymbol{i}_{\mathrm{T}}\right\} \,\middle|\, \boldsymbol{u}_{\mathrm{N}} = \mathbf{0}_M\right] \tag{7}$$

$$= \frac{1}{4R}\mathrm{E}\left[\boldsymbol{u}_{\mathrm{G}}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{u}_{\mathrm{G}}\right], \tag{7a}$$

where

$$\boldsymbol{B} = 2\left(\mathbf{I}_N + \boldsymbol{C}\right)^{-\mathrm{H}}\left(\boldsymbol{C} + \boldsymbol{C}^{\mathrm{H}}\right)\left(\mathbf{I}_N + \boldsymbol{C}\right)^{-1} \in \mathbb{C}^{N\times N}, \tag{8}$$

with the auxiliary matrix:

$$\boldsymbol{C} = \left(\boldsymbol{\Gamma}\left(\mathbf{I}_{M+N} + R^{-1}\boldsymbol{Z}\right)^{-1}\boldsymbol{\Gamma}^{\mathrm{T}}\right)^{-1} - \mathbf{I}_N \in \mathbb{C}^{N\times N}. \tag{9}$$

Herein, the superscript $^{\mathrm{H}}$, denotes complex conjugate transpose. Note from (7a) that, in general, the transmit power is not proportional to the average squared Euclidean norm of the generator voltage vector. There are only two special cases where the proportionality holds. The first case is when the components of $\boldsymbol{u}_{\mathrm{G}}$ are uncorrelated:

$$\exists\alpha:\ \mathrm{E}\left[\boldsymbol{u}_{\mathrm{G}}\boldsymbol{u}_{\mathrm{G}}^{\mathrm{H}}\right] = \alpha\mathbf{I}_N \quad\Longrightarrow\quad P_{\mathrm{Tx}} = \frac{\mathrm{tr}\,\boldsymbol{B}}{4R\cdot N}\cdot\mathrm{E}\left[\|\boldsymbol{u}_{\mathrm{G}}\|_2^2\right]. \tag{10}$$

This case usually occurs when the transmitter has no channel knowledge, and therefore treats all channel inputs the same. The second case is obvious:

$$\exists\alpha':\ \boldsymbol{B} = \alpha'\mathbf{I}_N \quad\Longrightarrow\quad P_{\mathrm{Tx}} = \frac{\mathrm{tr}\,\boldsymbol{B}}{4R\cdot N}\cdot\mathrm{E}\left[\|\boldsymbol{u}_{\mathrm{G}}\|_2^2\right]. \tag{11}$$

This can only happen in certain channels, which multi-port representation leads to uncoupled ports. We will elaborate on this later. In general however, there is a non-isotropic behavior of $\boldsymbol{u}_{\mathrm{G}}$ with respect to transmit power. A consequence is, for instance, that different vectors $\boldsymbol{u}_{\mathrm{G}}$, which have *the same* Euclidean norm, in general, produce *different* transmit power.

C. Receiver Noise Correlation

From (2), we see that the receiver noise voltage envelopes are, in general, linear superpositions of the voltage envelopes of *all* physical noise sources $u_{\mathrm{N},m}$, with $m \in \{1,2,\ldots,M\}$. This leads to the effect that the receiver noise is usually *correlated* even though all physical noise sources are uncorrelated, or even independent. In our case, the noise originates from thermal agitation of electrons inside the termination resistances at the receiver. Hence, the components of $\boldsymbol{u}_{\mathrm{N}}$ are uncorrelated, and we have [9]:

$$\mathrm{E}\left[\boldsymbol{u}_{\mathrm{N}}\boldsymbol{u}_{\mathrm{N}}^{\mathrm{H}}\right] = 4\underbrace{k_{\mathrm{B}}TW}_{\sigma^2}R\cdot\mathbf{I}_M, \tag{12}$$

where k_{B} is the Boltzmann constant, T is the absolute noise temperature, and W is the noise bandwidth. From (2) and (12), the correlation matrix of the receiver noise $\boldsymbol{\eta}$ is given by:

$$\boldsymbol{R}_{\boldsymbol{\eta}} = \mathrm{E}\left[\boldsymbol{\eta}\boldsymbol{\eta}^{\mathrm{H}}\right] \in \mathbb{C}^{M\times M}\cdot\mathrm{W} \tag{13}$$

$$= 4\sigma^2\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}. \tag{13a}$$

Consequently, only in case that $\boldsymbol{F}$ is a scaled unitary matrix, the receiver noise is uncorrelated. Again this only happens for special channels, which multi-port representation leads to uncoupled ports. We will elaborate on this later.

D. The Unilateral Channel

When energy and information can flow only in one direction, namely from the transmitting to the receiving end of the link, we have a *unilateral* channel. A radio communication channel operated in one direction is a very good approximation of a unilateral channel. The reason lies in the extremely large attenuation that the signals undergo on their way between the

ISIT 2009, Seoul, Korea, June 28 - July 3, 2009

transmitter and the receiver. Therefore, the transmitter essentially does not realize what is going on at the receiver, for instance, which impedance it is terminated with, or if the receiver exists at all. The unilateral channel is therefore a reasonable model for unidirectional radio communications. The impedance matrix of a unilateral channel has the form:

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{Z}_{1,1} & \mathbf{O}_{N\times M} \\ \boldsymbol{Z}_{2,1} & \boldsymbol{Z}_{2,2} \end{bmatrix}, \tag{14}$$

where $\boldsymbol{Z}_{1,1} \in \mathbb{C}^{N\times N}\cdot\Omega$, and $\boldsymbol{Z}_{2,2} \in \mathbb{C}^{M\times M}\cdot\Omega$, are the so-called transmit and receive impedance matrices, respectively, while $\boldsymbol{Z}_{2,1} \in \mathbb{C}^{M\times N}\cdot\Omega$, is called the trans-impedance matrix. Note that the complete description of the unilateral channel therefore requires $M^2 + N^2 + MN$ degrees of freedom. In the unilateral channel, (3), (8), and (13a) simplify to:

$$\boldsymbol{E} = \frac{1}{R}\left(\mathbf{I}_M + R^{-1}\boldsymbol{Z}_{2,2}\right)^{-1}\boldsymbol{Z}_{2,1}\left(\mathbf{I}_N + R^{-1}\boldsymbol{Z}_{1,1}\right)^{-1}, \tag{15}$$

$$\boldsymbol{B} = \frac{2}{R}\left(\mathbf{I}_N + R^{-1}\boldsymbol{Z}_{1,1}\right)^{-\mathrm{H}}\left(\boldsymbol{Z}_{1,1} + \boldsymbol{Z}_{1,1}^{\mathrm{H}}\right)\left(\mathbf{I}_N + R^{-1}\boldsymbol{Z}_{1,1}\right)^{-1}, \tag{15a}$$

$$\boldsymbol{R}_{\boldsymbol{\eta}} = \frac{4\sigma^2}{R^2}\left(\mathbf{I}_M + R^{-1}\boldsymbol{Z}_{2,2}\right)^{-1}\boldsymbol{Z}_{2,2}\boldsymbol{Z}_{2,2}^{\mathrm{H}}\left(\mathbf{I}_M + R^{-1}\boldsymbol{Z}_{2,2}\right)^{-\mathrm{H}}. \tag{15b}$$

The simplifications are not just of mathematical, but also of conceptual nature. As can be seen from (15a), the $\boldsymbol{B}$-matrix is only a function of transmit side properties of the channel (its transmit impedance matrix). Similarly, we see from (15b), that the noise correlation matrix is only a function of receiver side properties of the channel (most importantly, its receive impedance matrix). On the other hand, in the bilateral case (see (8), and (13a)), both the $\boldsymbol{B}$-matrix and the noise correlation matrix depend on the whole $\boldsymbol{Z}$-matrix, hence, on properties of both the transmitting and the receiving end of the channel. Take note from (15b), that in the case of $\boldsymbol{Z}_{2,2}$ being a scaled identity matrix, the receiver noise $\boldsymbol{\eta}$ is indeed uncorrelated. Also note that a $\boldsymbol{Z}_{2,2}$, which is a scaled identity, corresponds to a multi-port which receiver side ports are uncoupled. In many cases in practice, however, $\boldsymbol{Z}_{2,2}$ is not a scaled identity, like in multi-antenna radio communication systems, where the antennas are closely spaced.

III. Complete Information Theoretic Channel Model

Now that we have established a physical channel model, composed of the transfer model (2), the transmit power model (7a), and the noise model (13a), we return to our original problem of how to bring in the physical constraints into the standard information theoretic channel model:

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{v}. \tag{16}$$

Herein, the N-dimensional vector $\boldsymbol{x}$, is channel input, the M-dimensional vector $\boldsymbol{y}$, is channel output, while $\boldsymbol{v}$ denotes the M-dimensional channel noise vector, and $\boldsymbol{H} \in \mathbb{C}^{M\times N}$, is the channel matrix. As visualized in Figure 1, we define the following two bijective transformations:

$$\boldsymbol{x} = \boldsymbol{V}\boldsymbol{u}_{\mathrm{G}}, \tag{17}$$

$$\boldsymbol{y} = \boldsymbol{W}^{-1}\boldsymbol{u}_{\mathrm{L}}, \tag{17a}$$

between the physical input/output variables $(\boldsymbol{u}_{\mathrm{G}}, \boldsymbol{u}_{\mathrm{L}})$, and the information theoretic input and output variables $(\boldsymbol{x}, \boldsymbol{y})$. Our goal is to determine the tuple

$$\left(\boldsymbol{V}, \boldsymbol{W}, \boldsymbol{H}\right),$$

such that the following three conditions are met:

1) The input-output relationship in (16) is *equivalent* to the input-output relationship in (2).
2) The physical transmit power (7a) can be written as:
$$P_{\mathrm{Tx}} = \mathrm{E}\left[\|\boldsymbol{x}\|_2^2\right]. \tag{18}$$
3) The channel noise is Gaussian and white:
$$\mathrm{E}\left[\boldsymbol{v}\boldsymbol{v}^{\mathrm{H}}\right] = \sigma^2\cdot\mathbf{I}_M. \tag{19}$$

When we substitute (17) and (17a) into (16), we obtain

$$\boldsymbol{u}_{\mathrm{L}} = \boldsymbol{W}\boldsymbol{H}\boldsymbol{V}\boldsymbol{u}_{\mathrm{G}} + \boldsymbol{W}\boldsymbol{v}. \tag{20}$$

Comparing (20) with (2) shows that

$$\boldsymbol{H} = \boldsymbol{W}^{-1}\boldsymbol{E}\boldsymbol{V}^{-1}. \tag{21}$$

Since furthermore $\boldsymbol{W}\boldsymbol{v} = \sqrt{R}\boldsymbol{\eta}$, must hold, it follows that

$$\boldsymbol{W} = \frac{\sqrt{R}}{\sqrt{\sigma^2}}\boldsymbol{R}_{\boldsymbol{\eta}}^{1/2} \in \mathbb{C}^{M\times M}\cdot\sqrt{\Omega}, \tag{22}$$

yields the desired correlation matrix (19). In order that (17a) is bijective, we must require that $\boldsymbol{R}_{\boldsymbol{\eta}}$ is regular. When we solve (17) for $\boldsymbol{u}_{\mathrm{G}}$, substitute into (7a), and compare with (18), we see that the condition

$$\boldsymbol{V}\boldsymbol{B}^{-1}\boldsymbol{V}^{\mathrm{H}} = \frac{1}{4R}\mathbf{I}_N, \tag{23}$$

must hold. Note that as long as $\boldsymbol{B}$ is *positive definite*, its square root is Hermitian. Hence,

$$\boldsymbol{B}^{\mathrm{H}} = \boldsymbol{B} > \mathbf{0} \quad\Longrightarrow\quad \boldsymbol{V} = \frac{1}{2\sqrt{R}}\boldsymbol{B}^{1/2} \in \mathbb{C}^{N\times N}\cdot\sqrt{\Omega^{-1}}, \tag{24}$$

yields $\boldsymbol{V} = \boldsymbol{V}^{\mathrm{H}}$, and fulfills (23). Note that $\boldsymbol{H} \in \mathbb{C}^{M\times N}$, is dimensionless. With (24), (22), and (21) the problem is solved, provided that the following three conditions are met:

1) $\boldsymbol{B}$ is positive definite,
2) $\boldsymbol{R}_{\boldsymbol{\eta}}$ is regular,
3) $\boldsymbol{u}_{\mathrm{N}}$ contains Gaussian noise.

The third condition is necessary, because only Gaussian distributed random variables remain Gaussian distributed under arbitrary linear transformations. Under these conditions, the standard information theoretic channel model is in complete accordance with the physics governing the communication.

IV. Application: Multi-Antenna Communication

When antennas are arranged into antenna arrays, the antennas experience mutual near-field coupling because of spatial proximity [10]. These coupling effects are usually ignored in information theory [2]. On the other hand, by using the formalism presented in this paper, we can ensure that all relevant physical constraints, like, for instance, those arising from mutual antenna coupling, are taken care of inside the standard information theoretic vector channel model. Consequently, all information and signal theoretic results, like channel capacity or optimum beamforming, instantly become applicable to, and may take advantage of the physical constraints of the communication system. All we need is provide a multi-port description of the communication channel.

Let us briefly look at an illustrative example. As we have discussed in Section II-D, the unidirectional radio communication channel can be considered as unilateral. In [6], and [11], it is shown that for uniform linear arrays of isotropic radiators, which are equipped with an impedance matching network that compensates the imaginary part of the array's input impedance, the transmit impedance matrix can be written as:

$$(\boldsymbol{Z}_{1,1})_{m,n} = R_{\mathrm{r}} \cdot \mathrm{sinc}\left(2\pi \frac{\Delta}{\lambda}(m-n)\right), \tag{25}$$

where $\mathrm{sinc}(x)$ is the $\sin(x)/x$ function, λ is the wavelength, Δ is the spacing between adjacent antennas, and R_{r} is the radiation resistance [10], of the individual antennas. In [12], it is shown that a similar relationship as (25) also holds true for uniform linear arrays of Hertzian dipoles.

Notice from (25), that only for an antenna spacing of half of the wavelength, or integer multiples thereof, the transmit impedance matrix is a scaled identity. Hence, only for these special antenna spacings, the antennas are uncoupled, which clearly shows in $\boldsymbol{B} = \mathbf{I}_N$, and $\boldsymbol{R}_{\boldsymbol{\eta}} = \sigma^2 \mathbf{I}_M$, as can be seen from (15a), and (15b), for $R = R_{\mathrm{r}}$. For all other antenna separation, there is more or less strong mutual coupling, especially when antenna spacing is reduced below half the wavelength. As we will see in the following, mutual antenna coupling strongly influences both the transmit power (via $\boldsymbol{B}$), and the receiver noise covariance $\boldsymbol{R}_{\boldsymbol{\eta}}$, and hence, strongly impacts the channel capacity. The, perhaps, surprising result is, that the channel capacity can be substantially *increased* by the physical coupling effects! In order to demonstrate this effect, let us assume that the receiver is equipped with an identical antenna array, such that $\boldsymbol{Z}_{2,2} = \boldsymbol{Z}_{1,1}$, and that there is a correlated Rayleigh propagation channel [5], [6], [11]:

$$\boldsymbol{Z}_{2,1} = \frac{1}{\sqrt{\mathrm{tr}\,\boldsymbol{R}_{\mathrm{Tx}}}} \boldsymbol{R}_{\mathrm{Rx}}^{1/2} \boldsymbol{G} \boldsymbol{R}_{\mathrm{Tx}}^{1/2}, \tag{26}$$

where the matrices $\boldsymbol{R}_{\mathrm{Rx}} = \mathrm{E}\left[\boldsymbol{Z}_{2,1}\boldsymbol{Z}_{2,1}^{\mathrm{H}}\right]$, and $\boldsymbol{R}_{\mathrm{Tx}} = \mathrm{E}\left[\boldsymbol{Z}_{2,1}^{\mathrm{H}}\boldsymbol{Z}_{2,1}\right]$, denote the receive and the transmit fading correlation matrix, respectively, while $\boldsymbol{G} \in \mathbb{C}^{M\times N}$, contains i.i.d, zero-mean, unity-variance, complex, Gaussian entries. When we substitute (25) and (26), using $\boldsymbol{Z}_{2,2} = \boldsymbol{Z}_{1,1}$, into (15), (15a) and (15b), and the latter into (22), and (24), we obtain with (21) the channel matrix $\boldsymbol{H} \in \mathbb{C}^{M\times N}$, which captures all the governing physics. Because of (18), and (19), computing channel capacity is »business as usual« [13]. The fading correlation matrices are set up to reflect an angle spread of 120°, centered around the array axis (so-called »end-fire« direction) for both the transmitter and the receiver. For a constant ratio P_{Tx}/σ^2, which is large enough that the receiver operates in the high signal to noise ratio regime, Figure 3 shows the ergodic channel capacity as a function of the number of antennas for two different antenna separations. For $\Delta = \lambda/2$, the antennas are uncoupled, and we observe the well-known linear growth of channel capacity with the antenna number. However, when we reduce the antenna spacing to a quarter wavelength, we observe a *super-linear* growth. A signal processing explanation of this interesting effect can be found in [11]. From this briefly presented example we can learn that considering the governing physics in information theory is important and may even lead to better performance than expected otherwise.

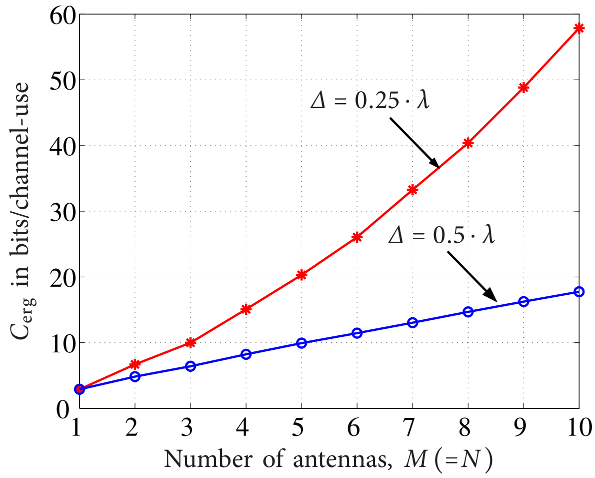

Figure 3: Ergodic channel capacity at high signal to noise ratio.

V. Conclusion

This paper makes the following contribution: a simple systematic approach is presented, which brings the physics that governs communications into information theory. Our approach is based on the proper definition of bijective transformations between the physical variables and the inputs and outputs of the information theoretic channel model.

References

[1] "Final report on link level and system level channel models," *IST-2003-507581 WINNER*, vol. D5.4, 2003.

[2] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[3] H. M. Paynter, *Analysis and Design of Engineering Systems*. MIT Press, Cambridge, MA, 1961.

[4] J. W. Wallace and M. A. Jensen, "Mutual Coupling in MIMO Wireless Systems: A Rigorous Network Theory Analysis," *IEEE Transactions on Wireless Communications*, vol. 3(4), pp. 1317–1325, July 2004.

[5] S. Lu, H. T. Hui, and M. Bialkowski, "Optimizing MIMO Channel Capacities Under the Influence of Antenna Mutual Coupling," *IEEE Antennas and Wireless Propagation Letters*, vol. 7, pp. 297–290, 2008.

[6] M. T. Ivrlač and J. A. Nossek, "The Maximum Achievable Array Gain under Physical Transmit Power Constraint," in *Proceedings of the IEEE International Symposium on Information Theory and its Applications*, Dec. 2008, pp. 1338–1343.

[7] R. G. Gallager, *Information Theory and Reliable Communication*. John Wiley & Sons, 1968.

[8] V. Belevitch, *Classical Network Theory*. Holden Day, San Francisco, 1968.

[9] M. J. Buckingham, *Noise in Electronic Devices and Systems*. New York: John Wiley & Sons Ltd., 1983.

[10] P. Russer, *Electromagnetics, Microwave Circuits and Antenna Design for Communications Engineering*, 2nd ed. Nordwood, MA: Artec House, 2006.

[11] M. T. Ivrlač and J. A. Nossek, "Receive Antenna Gain of Uniform Linear Arrays of Isotrops," in *Accepted for presentation at the IEEE International Conference on Communications*, June 2009.

[12] H. Yordanov, M. T. Ivrlač, P. Russer, and J. A. Nossek, "Arrays of Isotropic Radiators — A Field-theoretic Justification," in *Accepted for presentation at the ITG International Workshop on Smart Antennas, 2009*, Feb. 2009.

[13] I. E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," *European Transactions on Telecommunications*, vol. 10(6), pp. 585–596, Nov. 1999.