Unitary Nonuniform Space-Time Constellations for the Broadcast Channel

Erik G. Larsson, *Member, IEEE*

***Abstract*—We propose a new class of unitary space-time matrix constellations that have nonuniform error performance. Such codes are important for transmission over a broadcast, or point-to-multipoint, communication channel and when scalable source coding is used. Our new codes can be encoded differentially and are based entirely on phase-shift keying. Analytical design criteria are discussed and a simulation result is presented to exemplify the designs.**

***Index Terms*—Broadcast channels, space-time coding.**

I. Introduction

THE USE of multiple antennas at the receiver and transmitter, or so-called *multiple-input multiple-output* (MIMO) technology, along with appropriate coding over space and time, has shown to be very promising both to increase the throughput and to counteract fading in communication systems.

In this letter, we propose a new class of space-time codes suitable for data encoded using *layered source coding*. With layered source coding, different information bits have different importance for the reconstruction of a message, and therefore the target error rate for different classes of bits is different. For instance, the image and video coding standards JPEG-2000 and MPEG-4 use so-called "fine granularity scalability" as a way of trading error-free throughput against the quality of the reconstructed image or video sequence [1]. A particularly important application of such layered coding techniques is transmission over *point-to-multipoint*, or *broadcast*, communication links where distinct receivers typically experience very different radio link qualities and where the receivers also often have *inherently* different capabilities to decode the message.

Our codes are constructed starting from a criterion that attempts to minimize the error rate. They also satisfy two other important design goals. First, they can be encoded (and detected) differentially. Second, they are entirely based on phase-shift keying (PSK) modulation and consequently the transmitted signal has constant envelope at all times. They can, therefore, be seen as a multidimensional extension of a technique in [2].

II. Differential Modulation for MIMO Systems

We consider a MIMO system with n_t transmit antennas, n_r receive antennas, and assume for simplicity that the associated propagation channel is linear, time-invariant and frequency flat. Let $\boldsymbol{H}$ be an $n_r \times n_t$ matrix whose (m,n)th element contains the channel gain between transmit antenna n and receive antenna m, and suppose that a code matrix $\boldsymbol{X}$ of dimension $n_t \times N$ is taken from a matrix constellation $\mathcal{X}$ and transmitted by sending its N columns via the n_t antennas during N time epochs. If the signal received during the same N time intervals is arranged in an $n_r \times N$ matrix $\boldsymbol{Y}$, then $\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{X} + \boldsymbol{E}$ where $\boldsymbol{E}$ is an $n_r \times N$ matrix of noise. Throughout this letter, we shall also make the somewhat standard assumption that the elements of $\boldsymbol{H}$ and $\boldsymbol{E}$ are independent and zero-mean complex Gaussian with variances ρ^2 and σ^2, respectively; hence the channel is Rayleigh fading.

Given $\boldsymbol{Y}$, detection of $\boldsymbol{X}$ in a maximum-likelihood (ML) sense amounts to minimizing the Euclidian distance $\|\boldsymbol{Y} - \boldsymbol{H}\boldsymbol{X}\|^2$ with respect to $\boldsymbol{X} \in \mathcal{X}$, as well as $\boldsymbol{H}$ (unless it is known). The average (over $\boldsymbol{H}$) probability P_e that the ML detector mistakes a transmitted code matrix $\boldsymbol{X}_0$ for an incorrect (and different) code matrix $\boldsymbol{X} \neq \boldsymbol{X}_0$ can be bounded by [3]–[5]

$$P_e\Big|_{\mathrm{coh}} \leq \left|(\boldsymbol{X} - \boldsymbol{X}_0)(\boldsymbol{X} - \boldsymbol{X}_0)^H\right|^{-n_r} \cdot \left(\frac{\rho^2}{4\sigma^2}\right)^{-n_r n_t} \quad (1)$$

$$P_e\Big|_{\mathrm{noncoh}} \leq \left|\boldsymbol{X}_0 \Pi^{\perp}_{\boldsymbol{X}^H} \boldsymbol{X}_0^H\right|^{-n_r} \cdot \left(\frac{\rho^2}{4\sigma^2}\right)^{-n_r n_t} \quad (2)$$

for the two cases that the channel is known respectively unknown, in the latter case assuming $\boldsymbol{H}$ to be a deterministic unknown. Here $|\cdot|$ denotes the determinant of a matrix and $(\cdot)^H$ stands for the complex conjugate transpose. Also, $\Pi_{\boldsymbol{X}} = \boldsymbol{X}(\boldsymbol{X}^H\boldsymbol{X})^{-1}\boldsymbol{X}^H$ and $\Pi^{\perp}_{\boldsymbol{X}} = \boldsymbol{I} - \Pi_{\boldsymbol{X}}$ stand for the projections onto the range space of $\boldsymbol{X}$ and its orthogonal complement, and $\boldsymbol{I}$ is the identity matrix.

Space-time modulation matrices that can be encoded *differentially* are often of special importance since such codes can be demodulated noncoherently. Differential codes with uniform error probabilities have been studied by many authors (see, e.g., [6]–[8] for some prominent examples) and some of them can be seen as an extension of differential PSK to MIMO systems. In general, if t is the time index and if $\{\boldsymbol{U}(t)\}$ is a sequence of (square) information-bearing matrices, then differential encoding obtains the transmitted code matrix $\boldsymbol{X}(t)$ at time t via

$$\boldsymbol{X}(t) = \boldsymbol{X}(t-1)\boldsymbol{U}(t) \quad (3)$$

where $\boldsymbol{X}(t-1)$ is the matrix transmitted at time $t-1$. Such an encoding usually is only meaningful under certain circumstances, for instance if $\boldsymbol{U}(t)$ is unitary [in which case $\boldsymbol{X}(t)$ becomes unitary for all t as well, given a unitary "initial matrix" $\boldsymbol{X}(0)$]. By considering two (or more) received matrices $\boldsymbol{Y}(t)$ and $\boldsymbol{Y}(t-1)$ simultaneously, noncoherent detection is possible. For example, if we concatenate the matrices received at time $t-1$ and t, and assume that the channel $\boldsymbol{H}$ remains constant

Manuscript received July 16, 2002. The associate editor coordinating the review of this letter and approving it for publication was Prof. H. Jafarkhani.

The author is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA (e-mail: larsson@dsp.ufl.edu).

Digital Object Identifier 10.1109/LCOMM.2002.807432

over these two time intervals (to within practical accuracy), we can write [using (3)]

$$[\boldsymbol{Y}(t-1) \quad \boldsymbol{Y}(t)] = \boldsymbol{HX}(t-1)\cdot[\boldsymbol{I} \quad \boldsymbol{U}(t)] + [\boldsymbol{E}(t-1) \quad \boldsymbol{E}(t)] \tag{4}$$

where $\boldsymbol{HX}(t-1)$ can be seen as an unknown "effective" channel matrix and $[\boldsymbol{I} \quad \boldsymbol{U}(t)]$ is an effective code matrix. Provided that $[\boldsymbol{I} \quad \boldsymbol{U}(t)]$ is such that the corresponding determinant in (2) is nonzero, noncoherent demodulation is possible and the equation referred to can be used to assess the error rate.

III. A New Class of Nonuniform Space-Time Codes

The principal idea behind nonuniform codes is that two messages, a "basic" and an "additional" such, are simultaneously encoded into a transmitted matrix $\boldsymbol{X}$ at a given time, and that by construction the error rates associated with the basic and the additional message are different. The additional message requires a more capable receiver, or a higher signal-to-noise (SNR) ratio to be decoded than the basic message does. Since uniform constellations constitute a subclass of nonuniform ones, the design of nonuniform space-time constellations appears to be even more difficult than the design of uniform such.

Our new nonuniform space-time codes are based on differential encoding of the product of a square and unitary code matrix $\boldsymbol{U}_b \in \mathcal{X}_b$ associated with a basic message, and another square and unitary code matrix $\boldsymbol{U}_a \in \mathcal{X}_a$ corresponding to an additional message. Thus the transmitted code matrix at time t is given by [cf. (3)]

$$\boldsymbol{X}(t) = \boldsymbol{X}(t-1)\boldsymbol{U}_b(t)\boldsymbol{U}_a(t). \tag{5}$$

By using (2) along with the fact that all involved code matrices are unitary we find that the average probability that a transmitted message pair $(\boldsymbol{U}_b^0, \boldsymbol{U}_a^0)$ is mistaken for another pair $(\boldsymbol{U}_b, \boldsymbol{U}_a)$ can be bounded by

$$E\left[P(\boldsymbol{U}_b^0, \boldsymbol{U}_a^0 \to \boldsymbol{U}_b, \boldsymbol{U}_a)\right] \leq \left|\boldsymbol{I} - \frac{1}{2}\left(\boldsymbol{U}_b\boldsymbol{U}_a\boldsymbol{U}_a^{0H}\boldsymbol{U}_b^{0H} + \boldsymbol{U}_b^0\boldsymbol{U}_a^0\boldsymbol{U}_a^H\boldsymbol{U}_b^H\right)\right|^{-n_r}\left(\frac{\rho^2}{4\sigma^2}\right)^{-n_r n_t}. \tag{6}$$

A bound such as (6) was used in, for instance, [6] for the design of (uniform) differential space-time codes. However, although it may be thought of as a feasible approach, an attempt to minimize this bound in the context of *nonuniform* space-time modulation may *not* produce the desired result since the target error rates for $\boldsymbol{U}_b$ and $\boldsymbol{U}_a$ are different.

IV. Design Criteria for Nonuniform Codes

Since the additional message can be decoded only at high SNR, the matrices $\{\boldsymbol{U}_a\}$ associated with the additional message should, loosely speaking, be close to the identity matrix. Inspired by this fact, we suggest to first consider the design of $\{\boldsymbol{U}_b\}$, treating the presence of $\boldsymbol{U}_a$ as an unmodeled noise-like disturbance term. Doing so, we can approximately bound the error probability for $\boldsymbol{U}_b$ alone (assuming differential detection) by

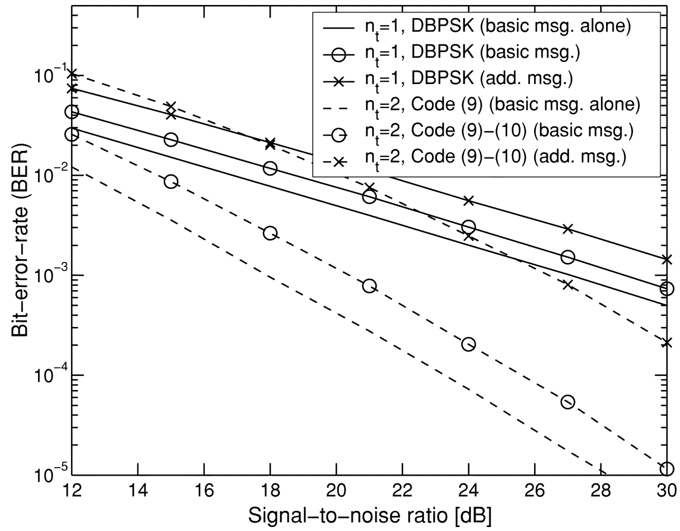

Fig. 1. Example of a nonuniform space-time constellation.

$$E\left[P(\boldsymbol{U}_b^0 \to \boldsymbol{U}_b)\right] \lesssim \left|\boldsymbol{I} - \frac{1}{2}\left(\boldsymbol{U}_b^0\boldsymbol{U}_b^H + \boldsymbol{U}_b\boldsymbol{U}_b^{0H}\right)\right|^{-n_r} \cdot\left(\frac{\rho^2}{4\sigma^2+\alpha^2}\right)^{-n_r n_t} \tag{7}$$

where α^2 is a factor that incorporates the noise-like effect of the presence of $\boldsymbol{U}_a$. Although the bound (7) is somewhat heuristic (and probably neither tight nor very accurate), we believe that it may serve a purpose as a meaningful design criterion for $\{\boldsymbol{U}_b\}$.

Next, for the design of $\{\boldsymbol{U}_a\}$ we proceed as follows. Suppose that the SNR is in a region such that $\boldsymbol{U}_b$ can be reliably decoded. For the design of $\{\boldsymbol{U}_a\}$ this should be a reasonable assumption, since if it is not true then decoding of $\boldsymbol{U}_a$ is probably of less interest anyway. Assuming that $\boldsymbol{U}_b$ is known, the demodulation of $\boldsymbol{U}_a$ is essentially another noncoherent detection problem. To obtain a criteria for the design of the constellation $\{\boldsymbol{U}_a\}$, we want to form an error bound on $\boldsymbol{U}_a$ and average it over all possible basic messages $\{\boldsymbol{U}_b\}$. After some simplification, we find that $E\left[P(\boldsymbol{U}_a^0 \to \boldsymbol{U}_a)\right]$ is bounded by

$$\frac{1}{|\mathcal{X}_b|}\sum_{\boldsymbol{U}_b\in\mathcal{X}_b}\left|\boldsymbol{I} - \frac{1}{2}\left(\boldsymbol{U}_b\boldsymbol{U}_a\boldsymbol{U}_a^{0H}\boldsymbol{U}_b^H + \boldsymbol{U}_b\boldsymbol{U}_a^0\boldsymbol{U}_a^H\boldsymbol{U}_b^H\right)\right|^{-n_r} \cdot\left(\frac{\rho^2}{4\sigma^2}\right)^{-n_r n_t} = \left|\boldsymbol{I} - \frac{1}{2}\left(\boldsymbol{U}_a\boldsymbol{U}_a^{0H} + \boldsymbol{U}_a^0\boldsymbol{U}_a^H\right)\right|^{-n_r} \cdot\left(\frac{\rho^2}{4\sigma^2}\right)^{-n_r n_t} \tag{8}$$

Note that in the last step $\boldsymbol{U}_b$ disappears since it is unitary. In (8), $|\cdot|$ denotes the cardinality of a set.

V. Design Example

Let us consider a basic message $\boldsymbol{U}_b$ taken from the following set:

$$\mathcal{X}_b = \left\{\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}, \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}, \begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}\right\}. \tag{9}$$

The constellation in (9), which is uniform and possesses certain optimality properties, is due to [6] and was called "BPSK"

$$\frac{1}{|\mathcal{X}_b|} \sum_{\substack{(\boldsymbol{U}_b^{(k)}, \boldsymbol{U}_b^{(n)}) \in \mathcal{X}_b \\ k \neq n}} \frac{1}{|\mathcal{X}_a|} \sum_{(\boldsymbol{U}_a^{(r)}, \boldsymbol{U}_a^{(s)}) \in \mathcal{X}_a} \left| \boldsymbol{I} - \frac{1}{2}\left(\boldsymbol{U}_b^{(k)} \boldsymbol{U}_a^{(r)} \boldsymbol{U}_a^{(s)H} \boldsymbol{U}_b^{(n)H} + \boldsymbol{U}_b^{(n)} \boldsymbol{U}_a^{(s)H} \boldsymbol{U}_a^{(r)H} \boldsymbol{U}_b^{(k)H} \right) \right|^{-n_r} \cdot \left(\frac{\rho^2}{4\sigma^2} \right)^{-n_r n_t} \quad (12)$$

therein. There are other constellations with higher rate, but for illustration purposes we choose a simple example. Based on the design rules in Section IV, we have handcrafted the following constellation of matrices for the additional message $\boldsymbol{U}_a$:

$$\mathcal{X}_a = \left\{ \begin{bmatrix} e^{i\pi\lambda} & 0 \\ 0 & e^{i\pi\gamma} \end{bmatrix}, \begin{bmatrix} e^{i\pi\gamma} & 0 \\ 0 & e^{i\pi\lambda} \end{bmatrix}, \begin{bmatrix} e^{-i\pi\lambda} & 0 \\ 0 & e^{-i\pi\gamma} \end{bmatrix}, \begin{bmatrix} e^{-i\pi\gamma} & 0 \\ 0 & e^{-i\pi\lambda} \end{bmatrix} \right\} \quad (10)$$

where (λ, γ) are design parameters (to be discussed below). If we take, somewhat arbitrarily

$$\boldsymbol{X}(0) = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \quad (11)$$

it follows that $\boldsymbol{X}^H(t)\boldsymbol{X}(t) = 2\boldsymbol{I}$ for all t and hence differential encoding of the code is meaningful. Furthermore, we can verify that all elements of $\boldsymbol{X}$ have always unit magnitude: $\left|X_{k,l}(t)\right| = 1$ and hence all transmitted symbols are obtained by constant envelope modulation, which was one of our design goals.

The error performance (of both the basic and the additional message) will depend on λ and γ, and typically the error performance of the additional message can be traded against that of the basic message. The values (λ, γ) must be chosen carefully. For instance, at first glance intuition would perhaps suggest us to take $\lambda > 0$ and $\gamma = 0$; we can show that such a "simple" choice leads to a code that "works" but that does *not* provide maximal diversity.

"Optimal" values of (λ, γ) can be obtained via simulation, but we have also used the following analytical tool. By the union bound, the probability for the basic message to be in error can be bounded by (12), shown at the top of the page. Likewise, the error rate for the additional message can be bounded by a similar expression (not presented here).

Fig. 1 (on previous page) shows the empirical bit-error-rate (BER) for the code described above using $\lambda = 0.23$, $\gamma = 0.03$ (obtained via optimization of the above error bounds) and ML decoding. The solid lines ("——") show the performance of differential BPSK for a conventional system with $n_t = 1$ transmit antenna and a single receive antenna; the coding for $n_t = 1$ is essentially a special case of [2]. The dashed lines ("– –") show the performance for a system with $n_t = 2$ (and a single receive antenna) using the code presented above. For the curves without marks, only a basic message is transmitted. The curves with marks show the performance when both a basic and an additional message are transmitted: the curves marked with "$\circ$" show the BER for the basic message, and the curves with "$\times$" show the BER for the additional message.

Clearly, the transmit diversity system outperforms the conventional one – observe, in particular, the different slopes of the BER curves. Also, it is clear that the transmission of an additional message incurs a performance degradation, which is natural since the minimum distance in the constellation decreases (cf. the above remark about performance tradeoff).

VI. Concluding Remarks

We have presented new nonuniform space-time codes that can be encoded and detected differentially, and that are based entirely on phase-shift keying. We also discussed analytical criteria for code construction and optimization. More results and codes with higher rates, along with a more complete discussion of the code design criteria, will be presented elsewhere.

It may be argued that using nonuniform constellations for a single transmit antenna (see, e.g., [2]) together with known linear space-time codes (see, e.g., [5]) should be a natural way of designing nonuniform MIMO constellations. However, this is in general suboptimal since we optimize over the class of space-time codes and the class of nonuniform single-antenna constellations *separately*, instead of *jointly*. Also, constraints like the constant-modulus property may be harder to incorporate. Therefore, we believe that nonuniform space-time constellations should be designed by approaching the problem from first principles.

References

[1] W. Li, "Overview of fine granularity scalability in the MPEG-4 video standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, pp. 301–317, Mar. 2001.

[2] M. B. Pursley and J. M. Shea, "Nonuniform phase-shift key modulation for multimedia multicast transmission in mobile wireless networks," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 774–783, May 1999.

[3] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.

[4] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 46, pp. 543–564, Mar. 2000.

[5] E. G. Larsson and P. Stoica, Space-Time Block Coding for Wireless Communications. Cambridge, U.K.: Cambridge Univ. Press, to be published.

[6] B. L. Hughes, "Differential space-time modulation," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2567–2578, Nov. 2000.

[7] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1169–1174, July 2000.

[8] B. M. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Trans. Commun.*, vol. 48, pp. 2041–2052, Dec. 2000.