Optimal Modulation for Known Interference

Mikael Skoglund, *Senior Member, IEEE*, and Erik G. Larsson, *Member, IEEE*

***Abstract*—We present a symbol-by-symbol approach to the problem of canceling known interference at the transmitter in a communication system. In the envisioned system, the modulator maps an information symbol (taken from a finite alphabet) and an interference symbol (from the complex field) onto a transmitted constellation point. Our scheme is based on joint optimization of a modulator and demodulator, subject to a constraint on the average transmit power. The demodulator picks the information symbol (as a function of the received symbol) that minimizes the average error probability. We emphasize that our focus is on transmission in a single (complex) dimension, and hence the proposed technique is a "modulation" rather than a "coding" scheme. We illustrate that the new scheme outperforms Tomlinson–Harashima precoding, which is a classical but suboptimal solution to the one-dimensional known-interference precoding problem. In our simulations, the new approach is able to perform close to the no-interference bound.**

***Index Terms*—Modulation, dirty paper coding, Costa precoding, interference cancellation, joint design.**

I. INTRODUCTION

COSTA showed in his 1983 paper [1] that the achievable rates of a Gaussian channel remain unchanged if the decoder observes the transmitted codeword in the presence of additive Gaussian interference, provided that the *encoder knows the interference signal non-causally*. That is, the transmitter can effectively cancel the interference, without increasing the transmit power. This result, somewhat astonishing at first glance, shows that interference (no matter how strong) is never a limitation for a communication link, as long as the interfering signal is known to the transmitter. Note that the result is far from trivial: a naive solution like subtracting the interference at the transmitter in general would violate the power budget.

The problem of transmitter design in the presence of known interference is often called the "Costa (precoding)" or "dirty paper" problem (after the title of [1]). This problem is important because the known-interference scenario arises in a number of contexts, notably, when doing precoding for ISI channels and for the downlink multiuser MIMO channel [2], [3]. Consequently the Costa problem has stimulated much research. The achievability proof in [1] is based on random binning over a set of codevectors that are correlated with the possible interference sequences, with each bin representing a value for the information variable to be sent. The codeword actually transmitted is then obtained as a linear combination of a codevector inside the bin specified by the data, and the observed sequence of interference samples. Recently, the fundamental results presented in [1] has inspired several sophisticated coding schemes, for example those in [4]–[6], based on lattice strategies [4], lattices and trellis shaping [5] and superposition coding [6].

While the coding-based approaches of [4]–[6] achieve rates close to capacity, they are rather complex, both conceptually and computationally. In this paper we consider instead a conceptually very simple, *one-dimensional*[1] scheme for transmitter interference cancellation. Our scheme attempts to cancel interference on a symbol-by-symbol basis, by choosing the transmitted symbol at a given time to be a function of an information symbol and of the interference symbol that affects the channel at that given time instant. Thereby, our focus is on *modulation* rather than on coding. We design the modulator and the demodulator *jointly*, using an iterative design strategy and taking the symbol error probability as optimality criterion. Naturally, with our strategy we cannot achieve the same performance as without interference (since suppressing the interference completely, as in [1], requires coding over infinitely many dimensions). However, we will see that even with a computationally simple method as the one we suggest, we can cancel a large fraction of the interference in the channel.

We are not aware of much previous work that systematically treats the transmitter interference cancellation problem in a small, or a single, dimension. A special case of the precoding structure that we propose here (and which we also take as a benchmark) is the Tomlinson–Harashima precoder (THP), originally proposed for ISI channels [7], [8]. The THP effectively subtracts the interference at the transmitter, and then performs a nonlinear (more precisely, modulo) operation on the resulting signal to ensure that the transmitter operates within the power budget. THP is a popular method in many papers that deal with precoding for the multiuser MIMO downlink channel [9], [10]. In [12] we considered the same problem using a similar approach, but only for the case of binary signaling with binary interference and using mutual information as optimization criterion instead of error probability. The present paper is more general in the sense that we allow a symbol alphabet of arbitrary size, and interference with an arbitrary distribution.

While in this paper we only consider uncoded transmission,

Paper approved by S. A. Jafar, the Editor for Wireless Communication Theory and CDMA of the IEEE Communications Society. Manuscript received June 26, 2006; revised February 15, 2007. Parts of this paper were presented at IEEE ICASSP 2007.

M. Skoglund is with the School of Electrical Engineering and the ACCESS Linnaeus Center, Royal Institute of Technology, Stockholm, Sweden (e-mail: mikael.skoglund@ee.kth.se).

E. G. Larsson was with the Royal Institute of Technology. He is now with the Department of Electrical Engineering, Linköping University, Linköping, Sweden. He is a Royal Swedish Academy of Sciences Research Fellow, supported by a grant from the Knut and Alice Wallenberg Foundation. This work was also supported in part by the Swedish Research Council (VR) and VINNOVA.

Digital Object Identifier 10.1109/TCOMM.2008.060382.

[1]Throughout the paper, one dimension refers to one complex dimension.

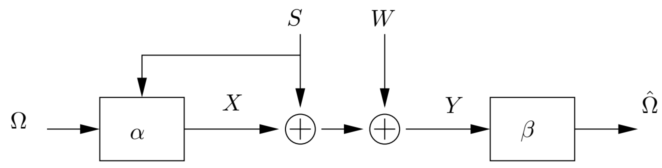


Fig. 1. System model.

using the raw error probability as performance measure, one can use outer (non-adaptive) coding over the effective channel created by our modulator–demodulator pair. For this, one could use a standard convolutional or Turbo code, or perhaps a code optimized for the particular characteristics of the effective channel that arises with our optimized scheme. This stands in contrast to previous works, for example [13] (see also [2]), that have considered extensions of THP to multiple dimensions by various combinations of the basic THP principle with trellis coding techniques.

II. Model and Problem Formulation

Consider the system depicted in Figure 1. The goal is to communicate an information symbol Ω over a discrete-time channel (one realization of Ω is transmitted per channel use), as reliably as possible. We model Ω as an M-ary discrete random variable, uniformly distributed over the set $\{\omega_1, \dots, \omega_M\}$. The transmission is subject to additive interference S and additive noise W. The interference S is a continuous complex-valued random variable with probability density function (pdf) f_S. The noise W is zero-mean complex Gaussian with variance σ^2. The transmitter (but not the receiver) observes the interference symbol S. The noise symbol W, however is unknown to both the transmitter and the receiver. The random variables Ω, S and W are assumed mutually independent.[2]

At the transmitter, the *modulator* $\alpha : \{\omega_1, \dots, \omega_M\} \times \mathbb{C} \to \mathbb{C}$ maps an information symbol $\Omega = \omega \in \{\omega_1, \dots, \omega_M\}$ and an interference symbol $S = s$ onto a transmitted symbol $x = \alpha(\omega, s) \in \mathbb{C}$. The modulator mapping α satisfies the following average power constraint

$$E|X|^2 = \frac{1}{M}\sum_{i=1}^{M} E|\alpha(\omega_i, S)|^2 \le P \qquad (1)$$

where the second expectation is taken with respect to f_S.

At the receiver, the *demodulator* (detector) $\beta : \mathbb{C} \to \{\omega_1, \dots, \omega_M\}$ looks at the received symbol

$$Y = X + S + W = \alpha(\Omega, S) + S + W \qquad (2)$$

and produces an estimate $\hat{\Omega} = \beta(Y)$ of the transmitted symbol. The resulting average symbol error probability is

$$P_e = \Pr(\hat{\Omega} \neq \Omega) \qquad (3)$$

The problem we consider in this paper is the optimal design of the modulator-demodulator pair (α, β) in the following sense: find the pair (α, β) which *minimizes P_e subject to the power constraint* (1).

Note that the presence of the interference can never improve the performance compared to the case $S = 0$. This is clear for the following reason: If the receiver in Figure 1 were provided perfect knowledge of S as well, performance would improve, or at the very least stay unchanged. But if S were known at both the transmitter and the receiver, the transmitter could simply pretend that $S = 0$ and let the receiver subtract S from Y. Therefore, a system where S is known at both ends is equivalent to a system without interference.

III. Optimal Modulator and Demodulator Design

The modulator–demodulator design problem is nonlinear and non-convex. Similar to classical designs for related problems, e.g. optimal design of vector quantizers [11], we formulate necessary conditions for optimality by presenting the optimal α given β, and vice versa.

A. The Optimal Demodulator for a Fixed Modulator

Assume that α is given and satisfies the power constraint (1). Then it is clear that the optimal demodulator is defined by the maximum-likelihood decision rule

$$\hat{\omega} = \beta(y) = \underset{\omega \in \{\omega_1, \dots, \omega_M\}}{\operatorname{argmax}} f_{Y|\Omega}(y|\omega) \qquad (4)$$

where $f_{Y|\Omega}$ is the conditional pdf of Y given Ω. (If soft decisions are desired, for example when the system in Figure 1 is concatenated with an "outer" code, then the actual values of $\{f_{Y|\Omega}\}$, not only the index of the maximizing ω, will be of interest.) Since the noise is Gaussian, we get

$$\begin{aligned} f_{Y|\Omega}(y|\omega) &= \int_{\mathbb{C}} f_S(s) f_{Y|\Omega,S}(y|\omega, s) ds \\ &= \frac{1}{\pi\sigma^2}\int_{\mathbb{C}} f_S(s) \exp\left(-\frac{1}{\sigma^2}|y - s - \alpha(\omega, s)|^2\right) ds \end{aligned} \qquad (5)$$

and the optimal decision

$$\hat{\omega} = \underset{\omega \in \{\omega_1, \dots, \omega_M\}}{\operatorname{argmax}} \int_{\mathbb{C}} f_S(s) \exp\left(-\frac{1}{\sigma^2}|y - s - \alpha(\omega, s)|^2\right) ds \qquad (6)$$

The integral in this expression does not admit a closed-form solution in general.

B. The Optimal Modulator for a Given Demodulator

To any given demodulator β, one can associate *decision regions* $\mathcal{B}_\omega$, $\omega \in \{\omega_1, \dots, \omega_M\}$. The decision region $\mathcal{B}_\omega$ is the set of y for which the demodulator decides on the information symbol ω:

$$\mathcal{B}_\omega = \{y : \beta(y) = \omega\}, \;\; \omega \in \{\omega_1, \dots, \omega_M\} \qquad (7)$$

For a given demodulator with decision regions $\{\mathcal{B}_\omega\}$, the average symbol error probability is

$$P_e = 1 - \frac{1}{M}\sum_{i=1}^{M} \int_{\mathcal{B}_{\omega_i}} f_{Y|\Omega}(y|\omega_i) dy \qquad (8)$$

[2]We shall use the following convention: Uppercase Ω, S, W, X, Y refer to the random variables and lowercase ω, s, w, x, y denote one particular possible value of each random variable. Occasionally, we will emphasize the finite-alphabet-nature of ω by enumerating its possible values according to ω_i where $i \in \{1, \dots, M\}$.

Taking the power constraint (1) into account via a positive Lagrange multiplier λ_1, a proper objective function to determine the optimal modulator hence is

$$\frac{1}{M}\sum_{i=1}^{M}\left(\int_{\mathcal{B}_{\omega_i}} f_{Y|\Omega}(y|\omega_i)dy - \lambda_1\int_{\mathbb{C}}|\alpha(\omega_i,s)|^2 f_S(s)ds\right) \tag{9}$$

The optimal modulator maximizes this objective function, for $\lambda_1 > 0$. (The Lagrange multiplier λ_1 must be chosen such that the power constraint is satisfied.) Hence, observing that

$$\int_{\mathcal{B}_\omega} f_{Y|\Omega}(y|\omega)dy = \frac{1}{\pi\sigma^2}\int_{\mathbb{C}} f_S(s) \cdot\left\{\int_{\mathcal{B}_\omega}\exp\left(-\frac{1}{\sigma^2}|y-s-\alpha(\omega,s)|^2\right)dy\right\}ds \tag{10}$$

with $f_S(s) \geq 0$, and absorbing positive constants into a new multiplier $\lambda_2 > 0$, it is clear that given $\Omega = \omega$ and $S = s$ the modulator should choose $x \in \mathbb{C}$ to maximize

$$\int_{\mathcal{B}_\omega}\exp\left(-\frac{1}{\sigma^2}|y-s-x|^2\right)dy - \lambda_2|x|^2 \tag{11}$$

This leads to the following, somewhat simpler, expression for the optimal modulator:

$$\begin{aligned}\alpha(\omega,s) &= \underset{x\in\mathbb{C}}{\operatorname{argmax}}\left\{\Pr(Y\in\mathcal{B}_\omega|\Omega=\omega,S=s,\alpha(\omega,s)=x)-\lambda_3|x|^2\right\}\\ &= \underset{x\in\mathbb{C}}{\operatorname{argmax}}\left\{\Pr(x+s+W\in\mathcal{B}_\omega)-\lambda_3|x|^2\right\}\end{aligned} \tag{12}$$

Equation (12) describes the optimal modulator α for a given demodulator (i.e., a given set of decision regions $\{\mathcal{B}_\omega\}$), and with $\lambda_3 > 0$. The interpretation of (12) is: Given $\Omega = \omega$ and $S = s$, x should be transmitted such that *the probability of observing Y in the correct decision region is maximized*, subject to limiting $|x|^2$ to satisfy the power constraint.

An alternative expression for x as a function of $\Omega = \omega$ and $S = s$ can be obtained by taking derivatives of (11) with respect to x and equating the result to zero. Doing so gives

$$\frac{1}{\sigma^2}\int_{\mathcal{B}_\omega}(y-s-x)\exp\left(-\frac{1}{\sigma^2}|y-s-x|^2\right)dy = \lambda_2 x \tag{13}$$

or equivalently

$$\begin{aligned}\lambda_4 x &= \frac{\int_{\mathcal{B}_\omega} y\exp\left(-\frac{1}{\sigma^2}|y-s-x|^2\right)dy}{\int_{\mathcal{B}_\omega}\exp\left(-\frac{1}{\sigma^2}|y-s-x|^2\right)dy} - s\\ &= E[Y|S=s,\Omega=\omega,\alpha(\omega,s)=x,Y\in\mathcal{B}_\omega]-s\\ &= x+E[W|W\in\mathcal{B}'_\omega]\end{aligned} \tag{14}$$

with $\lambda_4 \geq 1$ and where $\mathcal{B}'_\omega$ is the set of w for which $w+x+s \in \mathcal{B}_\omega$. As a side-result, we get the condition $x = -E[W|W\in\mathcal{B}'_\omega]$ in the case $\lambda_4 = 0$ (no power constraint). The expression (14) is also useful when numerically determining the optimal modulator, since it can be iterated to solve for x; see Section IV.

IV. Implementation Aspects

Here we discuss various aspects of the implementation of the system in Figure 1.

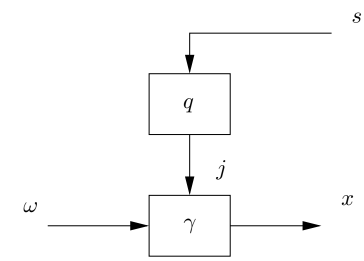


Fig. 2. Quantization in the modulator. The quantizer maps s onto an integer $j = q(s)$. The pair (ω, j) is then used to select a value $x = \gamma(\omega, j)$.

A. Modulation and Demodulation

In principle, the modulator and the demodulator can be implemented directly as formulated in Section III: Modulation amounts to solving the optimization problem in (11)–(12) for an observed pair $(\Omega, S) = (\omega, s)$. Likewise, the demodulator can be implemented by computing the integral over s in (6) numerically. However, the computational complexity associated with this approach is rather high.

An alternative procedure for computing the optimal modulator is based on (14), as follows. For $\lambda_4 \geq 1$, and for an observed pair $(\Omega, S) = (\omega, s)$, first guess an initial value x_0. For $x = x_0$, compute the right hand side (rhs) of (14). Denote the result z_1, and set $x_1 = z_1/\lambda_4$. Repeat this computation by using $x = x_1$ in the rhs of (14); denote the result z_2, and set $x_2 = z_2/\lambda_4$. Iterate until $x_{m+1} \approx x_m$. Our numerical experiments have shown that this approach converges relatively fast.

B. Modulator Based on Quantization and Table-Lookup

While computing $x = \alpha(\omega, s)$ for $(\Omega, S) = (\omega, s)$ is conceptually straightforward, it may be infeasible if computational complexity is a concern since a new x must be computed for each observed s. An alternative is to first quantize S into an integer $J \in \{1, \ldots, K\}$ (assuming K quantization levels) and approximate the modulator α with a table-lookup, which maps Ω and J onto a modulated symbol $x = \gamma(\omega, j)$ for $(\Omega, J) = (\omega, j)$. This approach is illustrated in Figure 2.

The main advantage of the table-lookup approach is that the table γ can be pre-computed, and the complexity of modulation hence effectively reduces to that of the quantization of s. For a sufficiently large K, the quantizer can without loss be implemented as uniform in the real and imaginary components of s. For small K, using a two-dimensional vector quantizer [11] may give substantial performance gains. In fact, for a given K, one can design an optimal quantizer for s, this however is outside the scope of the paper. Our simulation results, in Section V, assume uniform quantization.

C. Demodulator Based on Quantization and Table-Lookup

The demodulator (6) can also be implemented via a table-lookup, thereby avoiding to perform a numerical integration for each received sample. This is somewhat simpler than to implement the modulator, since all that has to be done is to quantize y; $f_{Y|\Omega}$ is then a function of the quantization index. However, the lookup-table must be recalculated when the encoder mapping α, the interference distribution f_S or the noise variance σ^2 changes. We have not pursued this approach further.

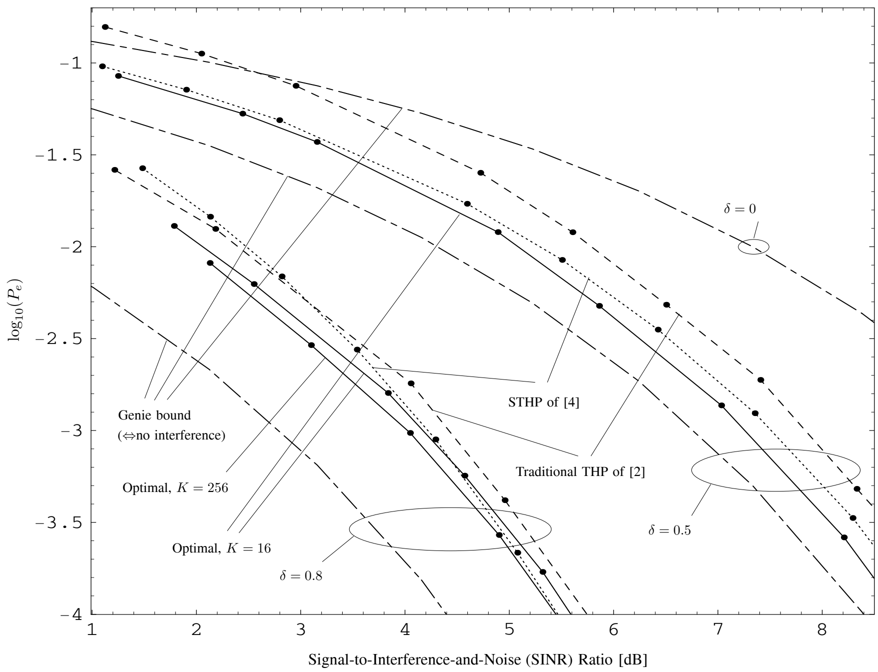


Fig. 3. Results with binary ($M = 2$) signaling.

D. Iterative System Design

The problem of jointly designing the optimal modulator α and the demodulator β is related to that of optimal quantizer design [11], and it is a non-convex problem. We propose the following two different algorithms for iterative design:

I) The first approach is based on updating the decision regions $\mathcal{B} = \{\mathcal{B}_\omega\}$. Let $\mathcal{B}^{(0)}$ be an initial choice for these regions, based on the nearest-neighbor regions of some well-known two-dimensional constellation (QAM), for example. Let $m = 1$.

1) Let $\mathcal{B}^{(m-1)}$ specify an optimal modulator $\alpha^{(m)}$, subject to the power constraint, as discussed in Sec. III-B.
2) Use $\alpha^{(m)}$ to specify $\alpha(\omega, s)$ in (6), and denote the resulting decision regions $\mathcal{B}^{(m)}$.
3) Set $m \to m+1$ and repeat from 1), until convergence.

Convergence in 3) can be monitored based on the resulting error probability P_e in each iteration. Since each update of the decision regions cannot lead to a system that performs worse, the algorithm will converge to a (local) optimum.

II) The second approach is based on using the expression (14) not to implement modulation directly (as in Section IV-A), but instead to indirectly specify an overall system iteration. Let $\alpha^{(0)}$ be an initial choice for the modulator, based for example on a standard constellation to define $\alpha(\omega, 0)$ and letting $\alpha(\omega, s) = \alpha(\omega, 0)$ for all s. Let $\mathcal{B}^{(0)}$ be the corresponding decision regions, defined using (6). Let $m = 1$.

1) Let $z(\omega, s)$ denote the rhs of (14), when evaluated for $\Omega = \omega$ and $S = s$ using $\alpha^{(m-1)}$ and $\mathcal{B}^{(m-1)}$.
2) Specify a new modulator $\alpha^{(m)}$ as $\alpha(\omega, s) = z(\omega, s)/\lambda_4$, and let $\mathcal{B}^{(m)}$ denote the optimal demodulator for $\alpha^{(m)}$.
3) Set $m \to m+1$ and repeat from 1) until convergence.

The above two algorithms are based on the optimal modulator, with $\alpha(\omega, s)$ continuous in s. When instead implementing (an approximation of) the modulator via quantized measurements of s, the only essential difference in the design is to use $\alpha(\omega, s) = \gamma(\omega, q(s))$ and update the table γ in the design iterations, for the K possible values for $q(s)$.

V. Numerical Results

Here we present numerical results comparing our optimized modulator to 1) regular PAM without precoding, 2) PAM with THP, and 3) PAM over a channel without interference, or equivalently a system where the interference is known at the demodulator. In all results, the information symbol Ω takes on M equally probable values. We consider only one-dimensional modulation (transmission on the I-carrier). However, since the I and Q channels are independent if the receiver maintains perfect phase synchronization, all results extend directly to I/Q modulation. The M-PAM reference scheme, used in cases 1) and 3), is uniform and has the amplitude-levels

$$X \in \{\pm p, \pm 3p, \ldots, \pm (M-1)p\} \tag{15}$$

with p chosen such that $E[X^2] = P$. (Strictly speaking, a uniform constellation is suboptimal for $M > 2$, and a gain could be obtained by using non-uniform M-PAM instead. This potential gain is relatively small, however.)

In the simulations, we define the signal-to-noise-plus-interference ratio (SINR) as

$$\frac{P}{E[S^2] + E[W^2]} = \frac{P}{\eta^2} \tag{16}$$

where $\eta^2 = \sigma^2 + E[S^2]$ is the total noise-plus-interference power. For a given η^2, we can write $E[S^2] = \delta\eta^2$ and $E[W^2] = \sigma^2 = (1-\delta)\eta^2$ for some $\delta \in [0, 1]$. By varying

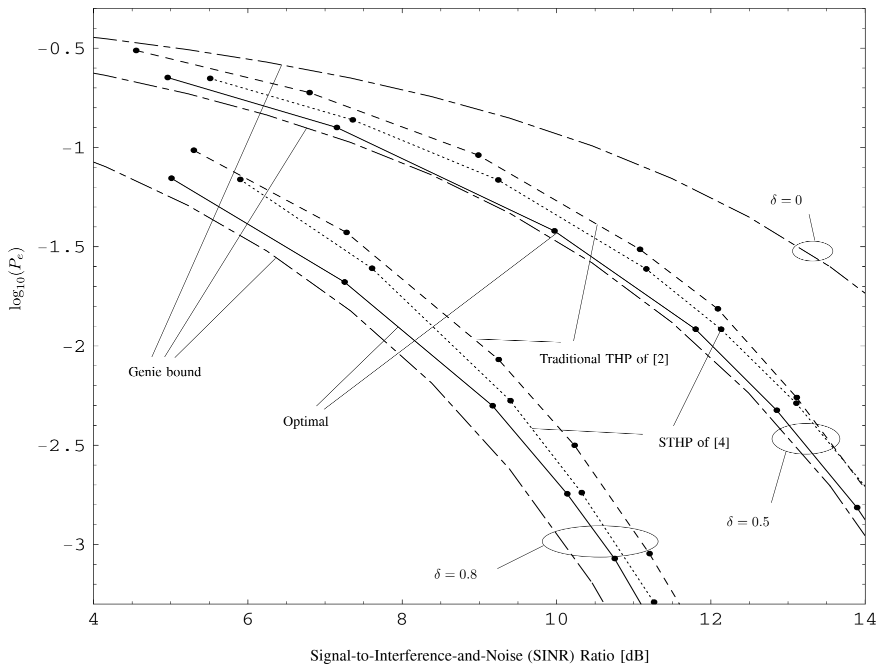


Fig. 4. Results with quaternary ($M = 4$) signaling.

δ from 0 to 1, we can choose how much of the total noise-and-interference power (η^2) that is allocated to S and to W, respectively. In other words, for a fixed SINR, δ quantifies how much of the total noise-and-interference that is known to the transmitter. In particular, taking $\delta = 0$ means that there is only noise; no interference is known to the transmitter. Conversely, if $\delta = 1$ then the transmitter fully knows all interference.

We will show results in terms of symbol error probability as a function of SINR, for different but fixed values of δ. This means that we fix the interference-to-noise ratio, and vary the signal power.[3] Note that the optimal modulator and demodulator designs depend on the power constraint (P), the noise power (σ^2; needed by the demodulator) and the statistical distribution of the interference ($f_S(s)$). In the design we assume these parameters are perfectly known, that is, each modulator/demodulator pair is tailored to a certain set of parameters. In practice, the iterative design can be carried out offline for a set of different parameters.

For optimizing the modulator–demodulator pair, we use Algorithm II in Section IV-D. To initialize the design, i.e., to provide an initial mapping $\alpha^{(0)}$ in Algorithm II, we used all of the following three different schemes: uniform M-PAM, uniform M-PAM + Gaussian noise (several different variances tested), and THP (cf., Section V-A). Then the modulator that performed the best after the iterative design was implemented for the simulations. We noted that these different initializations resulted in quite similar performance, however.

A. Benchmark 1: M-PAM with Traditional Tomlinson–Harashima Precoding (THP)

As a reference, we have implemented traditional THP [2] as defined in the following. Each information symbol Ω is mapped onto an M-PAM symbol

$$Z \in \{\pm q, \pm 3q, \ldots, \pm(M-1)q\} \tag{17}$$

Based on $\Omega = \omega$, $S = s$, and $Z = z$, THP is then performed via

$$\alpha(\omega, s) = (z - s) \mod \Lambda \tag{18}$$

where $\Lambda = [-Mq, +Mq]$. The constant q is chosen such that $E[x^2] = P$ (the expectation is over Ω and S). As in the traditional scheme, demodulation is implemented via

$$\hat{\Omega} = \underset{\omega}{\operatorname{argmin}} \left|(y \mod \Lambda) - z(\omega)\right|^2 \tag{19}$$

Strictly speaking, this receiver is suboptimal (the optimal receiver is obviously the ML receiver (6)). However, the performance of (19) is close to that of the optimal receiver for all cases of practical interest.

B. Benchmark 2: M-PAM with Scaled THP

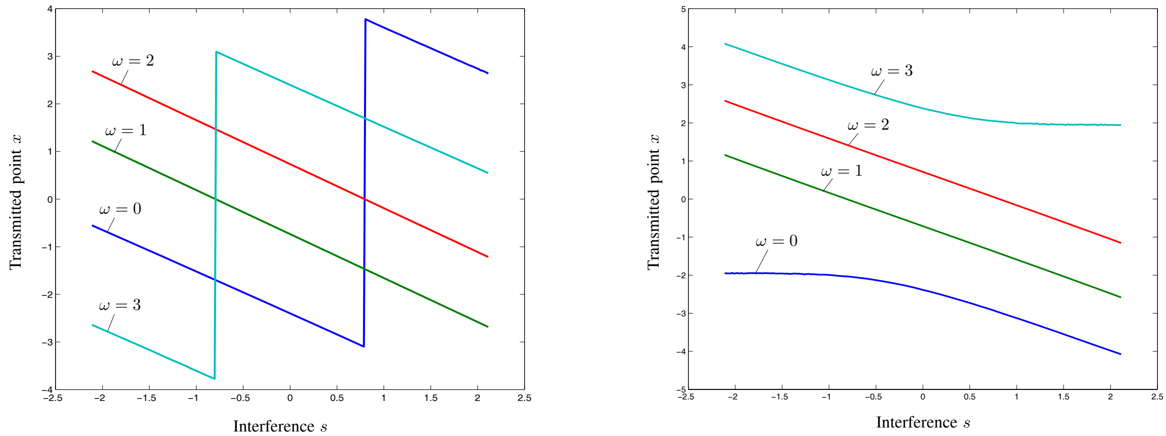
We also compare our new scheme with a modern version of THP which we refer to as "scaled THP" (STHP). This scheme was proposed in [4] (cf., the "inflated lattice strategy" in [4, Sec. III.F], with no dithering) and is described as follows. As in traditional THP, each information symbol Ω is mapped onto an M-PAM symbol $Z \in \{\pm q, \pm 3q, \ldots, \pm(M-1)q\}$. Then for $\Omega = \omega$, $S = s$ and $Z = z$, $\alpha(\omega, s)$ is formed as

$$\alpha(\omega, s) = (z - ks) \mod \tilde{\Lambda} \tag{20}$$

In (20), k is a constant, and $\tilde{\Lambda} = [-\Delta, +\Delta]$ with $\Delta = \sqrt{3P}$, where P is the transmit power.[4] The parameter k satisfies $0 < k \leq 1$ and can be chosen in different ways. Inspired by [1], [4] we set $k = P/(P + \sigma^2)$ (c.f., [4, Eq. (74)]). The

[3]Note that in some works on "dirty-paper precoding" the performance is instead measured for fixed P and $E[S^2]$ and a varying σ^2.

[4]Note that $\Delta > (M-1)q$.



(a) Optimization initialized with THP.

(b) Optimization initialized with uniform PAM.

Fig. 5. Illustration of the mapping $\gamma(\omega, j)$, c.f. Section IV-B, that approximates $\alpha(\omega, s)$ based on quantizing s into K levels. The figure shows the resulting x for $M = 4$ as a function of s when quantized into one out of $K = 256$ levels.

constant q is chosen such that $E[x^2] = P$. Detection at the receiver is then implemented as

$$\hat{\Omega} = \underset{\omega}{\operatorname{argmin}} |(k\, y \mod \tilde{\Lambda}) - z(\omega)|^2 \tag{21}$$

C. Results with Binary ($M = 2$) Signaling

Figure 3 shows the performance in terms of P_e versus the SINR for $M = 2$. The transmission is real-valued and S and W are independent, zero-mean Gaussian random variables. The figure shows the performance for $\delta = 0.5$ and $\delta = 0.8$. The figure also demonstrates the performance of 2-PAM with THP and STHP, implemented as described in Secs. V-A and V-B. Additionally the figure shows the performance of 2-PAM with no precoding as well as with perfect interference cancellation in the receiver. To implement the modulator mapping, the interference was quantized as discussed in Section IV-B. For $\delta = 0.5$ we used $K = 2^4 = 16$ levels. For $\delta = 0.8$ we used, and compared, the two choices $K = 16$ and $K = 256$. Notice that this way of implementing the modulator results in KM possible values for the transmitted x. Optimization of the modulator and demodulator was performed for several combinations of (δ, η^2); each combination corresponds to one solid dot in the figure.

Our optimized modulator outperforms traditional THP. Scaled THP generally performs better than traditional THP, still our new scheme is able to outperform also STHP (for $\delta = 0.8$ and high SINR's a higher resolution in the quantization ($K = 256$) is however needed). The price paid is a considerable increase in the design complexity, and some increase in demodulation complexity (evaluating (6) requires either a numerical integration or a table-lookup).

D. Results with Quaternary ($M = 4$) Signaling

We next explore the performance with $M = 4$, see Figure 4. For $\delta = 0.5$ our optimal modulator–demodulator performs close to the system for which the receiver knows the interference. This is a very encouraging result as it shows that at least for quaternary modulation (per dimension), almost perfect transmitter interference cancellation can actually be achieved via one-dimensional processing. For $\delta = 0.8$ the gap to optimal performance is somewhat larger. In general, similar conclusions as those for Figure 3 hold.

It is interesting to explore the behavior of the mapping $\alpha(\omega, s)$ when optimized using K-level quantization for s. Figure 5 illustrates this, for $M = 4$, $K = 256$ and $\delta = 0.5$. The left plot shows the resulting modulator when initializing the iterative optimization procedure using traditional THP. Similarly, the right plot shows the result when using uniform 4-PAM for the initial values. The range for s (from approx. -2 to $+2$) is set to $[-4\sigma_S, +4\sigma_S]$, with $\sigma_S^2 = E[S^2]$ (note that $\Pr(|S| > 4\sigma_S) < 10^{-4}$).[5] The modulator in a) resulted in $P_e \approx 0.202$ at SINR ≈ 5.54 dB. The corresponding values for the plot in b) are $P_e \approx 0.202$ at SINR ≈ 5.51 dB, that is, the system in b) performs slightly better. Notice, however, that even though the two systems "look different" they have very similar performance. This conclusion extends to many of the cases illustrated in the performance plots. In Figure 5 we can clearly see that when initializing with THP (plot a), the final result is reminiscent of THP. In fact the optimized modulator in a) appears to correspond to THP with an adaptive modulo region, i.e. THP with Λ as a function of s. For larger s the modulo region is shifted to allow for a bias toward positive x-values. From Figure 5 it is also clear that the mapping $\alpha(\omega, s)$ exhibits symmetries that can be exploited to reduce the design and implementation complexity.

E. Results with Octonary ($M = 8$) Signaling

Finally, in Figure 6 we show the performance with $M = 8$, comparing our optimized scheme with traditional THP (STHP was not implemented in this simulation). The interference is quantized using $K = 64$. The results and conclusions are quite similar to those in Figure 3 and Figure 4. The major difference is that the optimized system now performs very

[5]The modulator is implemented such that in the unlikely event $|s| > 4\sigma_S$ the output values for x remain constant as a function of s (not illustrated in Figure 5).

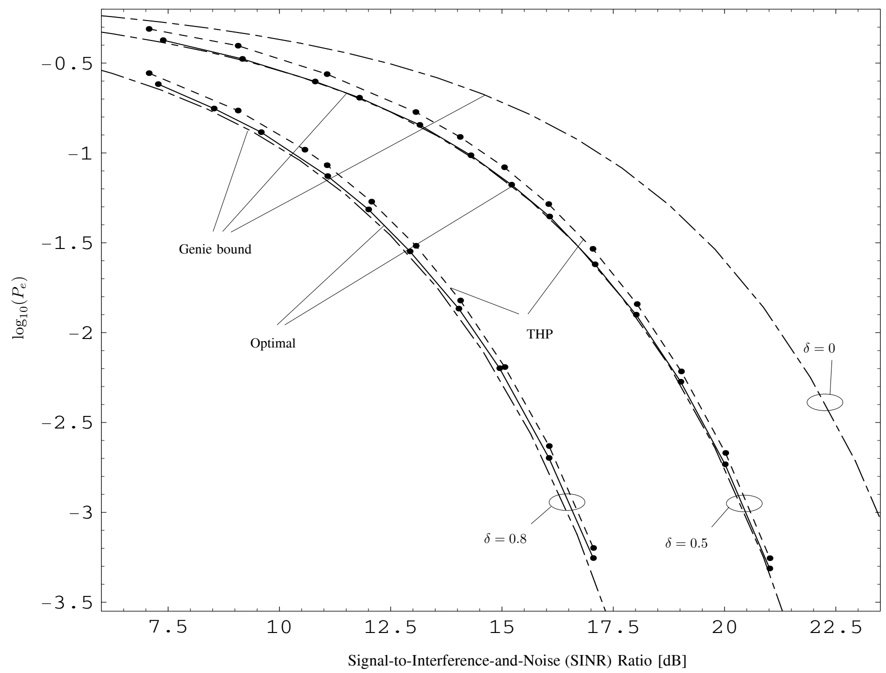


Fig. 6. Results with octonary ($M = 8$) signaling.

close to the genie-bound (perfect interference cancellation), both for $\delta = 0.5$ and for $\delta = 0.8$. However, this conclusion is also beginning to hold for THP. In general, by comparing Figures 3, 4 and 6, we see that the relative gain of using our optimized modulator instead of THP is larger for small constellations. This is natural, as it is known that THP is closer to optimal for large M [2].

VI. Conclusions and Further Work

We have proposed a scheme to cancel known interference at the transmitter, using a symbol-by-symbol processing approach. We addressed the problem of designing a modulator–demodulator pair with minimum error probability by using an iterative design strategy: optimizing the modulator for a given demodulator, and vice versa.

The performance of our new approach was compared with Tomlinson–Harashima precoding, and with a system without interference. We were able to demonstrate relatively good results. In particular, our proposed scheme in general outperforms Tomlinson–Harashima precoding, and it also is able to perform close to the no-interference bound.

One can think of the system in Figure 1 as an "inner" channel, that can be concatenated with outer (long, non-adaptive) coding. Soft demodulation for the outer code can then be directly implemented based on $f_{Y|\Omega}$ as given by (5), or a corresponding lookup-table. Future work may include quantification of how such a concatenated system performs. Another avenue for future work is to extend the proposed framework to vector-valued X, S and Y (for example, two or three dimensions). It is conceivable that this will bring a significant performance gain for small M.

VII. Acknowledgments

The authors would like to thank two anonymous reviewers for valuable comments that helped improve the paper. We also wish to thank one of the reviewers for suggesting the comparison with STHP (as proposed in [4]).

References

[1] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, pp. 439-441, May 1983.
[2] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge University Press, 2005.
[3] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1691–1706, July 2003.
[4] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice strategies for canceling known interference," *IEEE Trans. Inform. Theory*, vol. 51, pp. 3820–3833, Nov. 2005.
[5] U. Erez and S. ten Brink, "A close-to-capacity dirty paper coding scheme," *IEEE Trans. Inform. Theory*, vol. 51, pp. 3417–3432, Oct. 2005.
[6] A. Bennatan, D. Burshtein, G. Caire, and S. Shamai, "Superposition coding for side-information channels." *IEEE Trans. Inform. Theory*, vol. 52, pp. 1872–1889, May 2006.
[7] M. Tomlinson, "New automatic equalizer employing modulo arithmetic," *Electron. Lett.*, pp. 138–139, Mar. 1971.
[8] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Trans. Commun.*, vol. COM-20 No. 4, pp. 774–780, Aug. 1972.
[9] M. Airy, A. Forenza, R. W. Heath Jr., and S. Shakkottai, "Practical Costa precoding for the multiple antenna broadcast channel," in *Proc. GLOBECOM*, Dec. 2004.
[10] C. Windpassinger, R. F. H. Fischer, T. Vencel, and J. B. Huber, "Precoding in multiantenna and multiuser communications," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1305–1316, July 2004.
[11] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Kluwer 1992.
[12] J. Du, E. G. Larsson, and M. Skoglund, "Costa precoding in one dimension," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2006.
[13] W. Yu, D. P. Varodayan, and J. M. Cioffi, "Trellis and convolutional precoding for transmitter-based interference presubtraction," *IEEE Trans. Commun.*, vol. 53, pp. 1220–1230, July 2005.

Mikael Skoglund (S'93-M'97-SM'04) received the Ph.D. degree in 1997 from Chalmers University of Technology, Göteborg, Sweden. In 1997, he joined the Royal Institute of Technology, Stockholm, Sweden. Here he held various positions until he was appointed Professor of Communication Theory in October 2003.

Dr. Skoglund's research interests are in information theory, communications, and detection and estimation. He has worked on problems in source–channel coding, coding and transmission for wireless communications, Shannon theory and statistical signal processing. He has authored more than 130 scientific papers, including papers that have received awards, invited conference presentations, and papers ranking as highly cited according to the ISI Essential Science Indicators. He has also consulted for industry, and he holds six patents. Dr. Skoglund is an Associate Editor with the IEEE TRANSACTIONS ON COMMUNICATIONS.

Erik G. Larsson is Professor and Head of the Division for Communication Systems in the Department of Electrical Engineering (ISY) at Linköping University (LiU) in Linköping, Sweden. He joined LiU in September 2007. He has previously been Associate Professor (Docent) at the Royal Institute of Technology (KTH) in Stockholm, Sweden, and Assistant Professor at the University of Florida and the George Washington University, USA. His main professional interests are within the areas of wireless communications and signal processing. He has published some 50 papers on these topics, he is co-author of the textbook *Space-Time Block Coding for Wireless Communications* (Cambridge Univ. Press, 2003) and he holds 10 patents on wireless technology. He is Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE SIGNAL PROCESSING LETTERS and a member of the IEEE Signal Processing Society SAM and SPCOM technical committees.