OPTIMAL MODULATION FOR KNOWN INTERFERENCE

Mikael Skoglund and Erik G. Larsson

Royal Institute of Technology (KTH), EE Communication Theory,
Osquldas väg 10, 100 44 Stockholm, Sweden. Email: `mikael.skoglund@ee.kth.se`

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. The demodulator picks the information symbol (as a function of the received symbol) which minimizes the average error probability. We find the optimal modulator–demodulator pair, in the minimum-probability-of-symbol-error sense, via an iterative optimization procedure, for fixed average transmit power. We illustrate that the new scheme can perform close to the no-interference bound, and in particular that it outperforms Tomlinson–Harashima precoding, which is a classical but suboptimal solution to the problem under study.

***Index Terms*—** interference suppression, modulation

1. INTRODUCTION

Costa [1] has shown 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 amount of transmit power used. Thus interference, no matter how strong, is never a limitation for a communication link, as long as the interfering signal is known to the transmitter.

Coding/modulation design for known interference at the transmitter is important in a number of contexts, for example when doing precoding for the downlink multiuser MIMO channel [2, 3, 8]. Consequently the problem has stimulated much research. The achievability proof in [1] is based on random binning over a set of codevectors that approximate 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. This strategy has inspired practical coding schemes, e.g. [4, 5], which can achieve rates close to capacity, however at a rather high computational cost.

In this paper we consider a new approach to interference cancellation at the transmitter. As opposed to [4,5] which aimed for achieving capacity, our goal is to examine what one can do under a very with tight constraint on complexity. Namely, we devise a method that performs coding over *one* single (complex) dimension—in fact, our focus is on *modulation* rather than on coding. More precisely, we 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 instance. We present a design methodology which attempts to find the jointly optimal modulator–demodulator pair for interference with given statistics, and show that the optimal transmitter and receiver in fact reduce to a table-lookup. Naturally, our strategy is inferior to that of [4, 5] (since suppressing the interference completely, as in [1], requires coding over infinitely many dimensions). We will see, however, that even with a computationally very simple method such as the one we suggest, we can achieve impressive performance.

This work was supported in part by the Swedish Research Council (VR), VINNOVA, and Wireless@KTH.

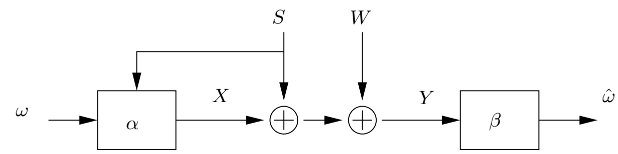

Fig. 1. System model.

Most previous work on transmitter interference cancellation is based on coding over high-dimensional lattices [4,5] or trellises [11]. Not much previous work exists on one-dimensional transmitter interference cancellation, i.e., using modulation rather than coding. A special case of the precoding structure that we propose here, however, is the Tomlinson–Harashima precoder (THP) [6, 7] (see Section 5). In a related previous paper [10] we considered the same problem but only for the special case of binary signaling with binary interference.

2. 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 a discrete random variable, uniformly distributed over $\mathcal{I}_M \triangleq \{1, \dots, M\}$. 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 transmitter and receiver. The random variables ω, S and W are assumed mutually independent.

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

$$E|X|^2 = \frac{1}{M} \sum_{i=1}^{M} E|x(i; S)|^2 \leq P \qquad (1)$$

where the second expectation is taken with respect to f_S.

At the receiver, the *demodulator* (detector) $\beta : \mathbb{C} \to \mathcal{I}_M$ looks at the received symbol

$$Y = X + S + W = x(\omega, S) + S + W \tag{2}$$

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

The problem we consider 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).

3. 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 [9], we formulate necessary conditions for optimality by presenting the optimal α given β, and vice versa.

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

$$\hat{\omega} = \arg\max_{i \in \mathcal{I}_M} f_{Y|\omega}(y|i) \tag{3}$$

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

$$\begin{aligned} f_{Y|\omega}(y|i) &= \int_{\mathbb{C}} f_S(s) f_{Y|\omega,S}(y|i,s)ds \\ &= \frac{1}{\pi\sigma^2}\int_{\mathbb{C}} f_S(s)\exp\left(-\frac{1}{\sigma^2}|y - s - x(i,s)|^2\right)ds \end{aligned} \tag{4}$$

and the optimal decision

$$\hat{\omega}(y) = \arg\max_{i \in \mathcal{I}_M}\int_{\mathbb{C}} f_S(s)\exp\left(-\frac{1}{\sigma^2}|y - s - x(i,s)|^2\right)ds \tag{5}$$

The Optimal Modulator for a Given Demodulator. To any given demodulator β, one can associate *decision regions* $\mathcal{B}_i$. The ith decision region is the set of y for which the demodulator decides on the information symbol $\omega = i$:

$$\mathcal{B}_i = \{y : \beta(y) = i\}, \;\; i = 1, \ldots, M \tag{6}$$

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

$$P_e = 1 - \frac{1}{M}\sum_{i=1}^{M}\int_{\mathcal{B}_i} f_{Y|\omega}(y|i)dy \tag{7}$$

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

$$\frac{1}{M}\sum_{i=1}^{M}\left(\int_{\mathcal{B}_i} f_{Y|\omega}(y|i)dy - \lambda_1\int_{\mathbb{C}}|x(i,s)|^2 f_S(s)ds\right) \tag{8}$$

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

$$\begin{aligned} \int_{\mathcal{B}_i} f_{Y|\omega}(y|i)dy &= \frac{1}{\pi\sigma^2}\int_{\mathbb{C}} f_S(s) \\ &\cdot\left\{\int_{\mathcal{B}_i}\exp\left(-\frac{1}{\sigma^2}|y - s - x(i,s)|^2\right)dy\right\}ds \end{aligned} \tag{9}$$

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

$$\int_{\mathcal{B}_i}\exp\left(-\frac{1}{\sigma^2}|y - s - x|^2\right)dy - \lambda_2|x|^2 \tag{10}$$

This leads to the following expression

$$\begin{aligned} &x(i,s) \\ &= \arg\max_{x\in\mathbb{C}}\left\{\Pr(Y \in \mathcal{B}_i|\omega = i, S = s, x(i,s) = x) - \lambda_3|x|^2\right\} \\ &= \arg\max_{x\in\mathbb{C}}\left\{\Pr(x + s + W \in \mathcal{B}_i) - \lambda_3|x|^2\right\} \end{aligned} \tag{11}$$

Equation (11) describes the optimal modulator α for a given demodulator (i.e., a given set of decision regions $\{\mathcal{B}_i\}$), and with $\lambda_3 > 0$. The interpretation of (11) is: Given $\omega = i$ 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 $\omega = i$ and $S = s$ can be obtained by taking derivatives of (10) with respect to x and equating the result to zero. Doing so gives

$$\frac{1}{\sigma^2}\int_{\mathcal{B}_i}(y - s - x)\exp\left(-\frac{1}{\sigma^2}|y - s - x|^2\right)dy = \lambda_2 x \tag{12}$$

or equivalently

$$\begin{aligned} \lambda_4 x &= \frac{\int_{\mathcal{B}_i} y\exp\left(-\frac{1}{\sigma^2}|y - s - x|^2\right)dy}{\int_{\mathcal{B}_i}\exp\left(-\frac{1}{\sigma^2}|y - s - x|^2\right)dy} - s \\ &= E[Y|S = s, \omega = i, x(i,s) = x, Y \in \mathcal{B}_i] - s \\ &= x + E[W|W \in \mathcal{B}_i - x - s] \end{aligned} \tag{13}$$

with $\lambda_4 \geq 1$ and where $\mathcal{B}_i - x - s$ is the set $\{w : w + x + s \in \mathcal{B}_i\}$. The expression (13) is useful when numerically determining the optimal modulator, since it can be iterated to solve for x; see Section 4.

4. IMPLEMENTATION ASPECTS

Modulation and Demodulation. In principle, the modulator and the demodulator can be implemented directly as formulated in Section 3: Modulation amounts to solving the optimization problem in (10)–(11) for an observed pair $(\omega, S) = (i, s)$. Likewise, the demodulator can be implemented by computing the integral over s in (5) numerically. However, the computational complexity associated with this approach is rather high.

An alternative procedure for computing the optimal modulator is based on (13), as follows. For $\lambda_4 \geq 1$, and for an observed pair $(\omega, S) = (i, s)$, first guess an initial value x_0. For $x = x_0$, compute the right hand side (rhs) of (13). 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 (13);

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.

Modulator Based on Quantization and Table-Lookup. While computing $x = \alpha(i, s)$ for $(\omega, S) = (i, s)$ 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 $J \in \mathcal{I}_K$ (assuming K quantization levels) and approximate the modulator α with a table-lookup, which maps ω and J onto a modulated symbol $x = \gamma(i, j)$ for $(\omega, J) = (i, j)$.

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 [9] 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 5, assume uniform quantization.

Demodulator Based on Quantization and Table-Lookup. The demodulator (5) 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.

Iterative System Design. The problem of jointly designing the optimal modulator α and the demodulator β is related to that of optimal quantizer design [9], and it is a non-convex problem. We propose an iterative approach based on (13). Using (13) we do not implement modulation directly, but instead we 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 $x(i, 0)$ and letting $x(i, s) = x(i, 0)$ for all s. Let $\mathcal{B}^{(0)}$ be the corresponding decision regions, defined using (5). Let $m = 1$.

1) Let $z(i, s)$ denote the rhs of (13), when evaluated for $\omega = i$ and $S = s$ using $\alpha^{(m-1)}$ and $\mathcal{B}^{(m-1)}$.
2) Specify a new modulator $\alpha^{(m)}$ as $x(i, s) = z(i, s)/\lambda_4$, and let $\mathcal{B}^{(m)}$ denote the optimal demodulator for $\alpha^{(m)}$.
3) Set $m \to m + 1$ and repeat from 1) until convergence.

5. 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 amplitude-levels $X \in \{\pm p, \pm 3p, \ldots, \pm(M-1)p\}$, 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 very small, however.)

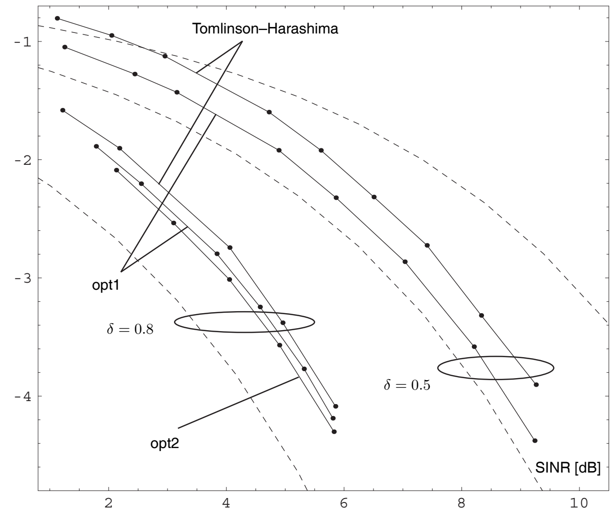

Fig. 2. Results with binary ($M = 2$) signaling, $\log_{10} P_e$ vs SINR in dB. Dashed curves from above: 2-PAM, 2-PAM with interference removed, $\delta = 0.5$ and $\delta = 0.8$, respectively. 'opt1' = optimized modulator–demodulator, s quantized with $K = 16$. 'opt2' = optimized modulator–demodulator, s quantized with $K = 256$.

In the simulations, we define the signal-to-noise-plus-interference ratio (SNIR) as $P/(E[S^2] + E[W^2]) = P/\eta^2$ 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 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. That is, δ is a measure of how much of the total noise-and-interference is known to the transmitter.

For optimizing the modulator–demodulator pair, we use the algorithm in Section 4. To initialize the design, that is, to provide an initial mapping $\alpha^{(0)}$, we used all of the following three different schemes: uniform M-PAM, uniform M-PAM + Gaussian noise (several different variances tested), and Tomlinson–Harashima precoding (see below). 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, though.

Benchmark: M-PAM with THP. As a reference, we have implemented traditional THP as defined in the following. Each information symbol ω is mapped onto an M-PAM symbol, $z \in \{\pm q, \pm 3q, \ldots, \pm(M-1)q\}$. Based on $S = s$, THP is then performed via $x(\omega, s) = (z - s) \mod \Lambda$ 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

$$\hat{\omega} = \arg\min_{\omega} |(y \mod \Lambda) - z(\omega)|^2 \tag{14}$$

Strictly speaking, this receiver is suboptimal (the optimal receiver is obviously the ML receiver (5)). However, the performance of (14) is close to the optimal receiver for all cases of practical interest.

Results with Binary ($M = 2$) Signaling. Figure 2 shows the performance in terms of P_e versus the SNIR for $M = 2$. The transmission is real-valued and S and W are independent, zero-mean Gaussian. The figure shows the performance for $\delta = 0.5$ and $\delta = 0.8$. The figure also demonstrates the performance of 2-PAM with THP. Additionally the figure shows the performance of 2-PAM

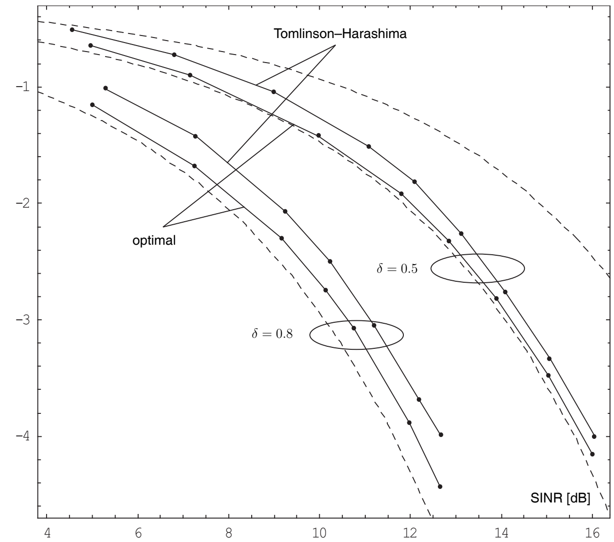

Fig. 3. Results with quaternary ($M = 4$) signaling, $\log_{10} P_e$ vs SINR in dB. Dashed curves from above: 4-PAM, 4-PAM with interference removed, $\delta = 0.5$ and $\delta = 0.8$, respectively. 'optimal' = optimized modulator–demodulator, s quantized with $K = 16$.

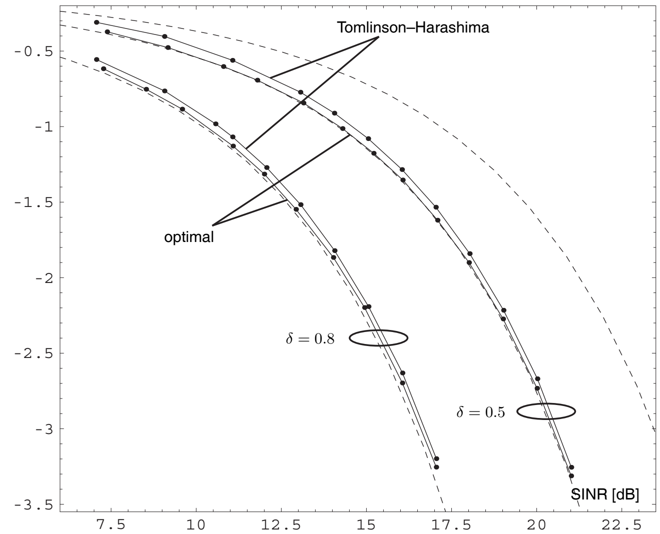

Fig. 4. Results with octonary ($M = 8$) signaling, $\log_{10} P_e$ vs SINR in dB. Dashed curves from above: 8-PAM, 8-PAM with interference removed, $\delta = 0.5$ and $\delta = 0.8$, respectively. 'optimal' = optimized modulator–demodulator, s quantized with $K = 64$.

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 4. For $\delta = 0.5$ we used $K = 2^4 = 16$ levels, and for $\delta = 0.8$ we used $K = 16$ as well as $K = 256$. 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.

At all values of the SINR, our optimized modulator outperforms THP. The price paid is a considerable increase in the design complexity, and some increase in demodulation complexity (evaluating (5) requires either a numerical integration or a table-lookup). We also observe that increasing from $K = 16$ to $K = 256$ quantization values for s results in a quite modest improvement. Hence, we expect that relatively low resolution for the quantization will suffice in practice.

Results with Quaternary ($M = 4$) Signaling. We next explore the performance with $M = 4$, see Figure 3. 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 (as opposed to infinitely-dimensional coding, as in [1, 4, 5]). For $\delta = 0.8$ the gap to optimal performance is somewhat larger. In general, similar conclusions as those for Figure 2 hold.

Results with Octonary ($M = 8$) Signaling. Finally, in Figure 4 we show performance with $M = 8$. The interference is quantized using resolution $K = 64$. Again, quite similar results apply. The major difference compared to the cases $M = 2$ and $M = 4$ is that the optimized system is now performing very close to the optimal, for both $\delta = 0.5$ and $\delta = 0.8$. However, this conclusion is also beginning to hold for the THP. In general, by comparing Figures 2, 3 and 4, 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].

6. CONCLUDING REMARKS

The main advantage of our proposed scheme is the low cost in complexity paid for a relatively good performance. The transmitter and the receiver can be implemented as a simple table-lookup. In terms of performance, our scheme outperforms Tomlinson–Harashima precoding and except for binary signaling, it also performs very close to the no-interference bound. An additional advantage of our method, which stands in contrast to most previous work, is that one can easily perform outer coding over the equivalent discrete channel defined by the modulator–demodulator pair in Figure 1.

References

[1] M. Costa, "Writing on dirty paper," *IEEE Trans. on Info. 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 multi-antenna Gaussian broadcast channel," *IEEE Trans. on Info. Theory*, vol. 49, pp. 1691–1706, July 2003.

[4] A. Bennatan, D. Burshtein, G. Caire and S. Shamai, "Superposition coding for side-information channels." Submitted to *IEEE Trans. on Info. Theory*, April 2004.

[5] U. Erez and S. ten Brink, "Approaching the dirty paper limit for cancelling known interference," in Proc. of *Allerton Conf. on Communications, Control and Computing*, Oct. 2003.

[6] M. Tomlinson, "New automatic equalizer employing modulo arithmetic," *Electronics Letters*, pp. 138-139, March 1971.

[7] H. Harashima, H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Trans. on Comm.*, vol. COM-20 No. 4, pp. 774-780, Aug. 1972.

[8] C. Windpassinger, R. F. H. Fischer, T. Vencel, and J. B. Huber, "Precoding in multiantenna and multiuser communications," *IEEE Transactions on Wireless Communications*, vol. 3, pp. 1305–1316, July 2004.

[9] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer 1992.

[10] J. Du, E. G. Larsson, and M. Skoglund, "Costa precoding in one dimension," in *Proc. of IEEE ICASSP*, 2006.

[11] W. Yu, D. P. Varodayan, and J. M. Cioffi, "Trellis and convolutional precoding for transmitter-based interference presubtraction," *IEEE Trans. on Comm.*, vol. 53, pp. 1220–1230, July 2005.