MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

# Universal Multi-Stage Precoding with Monomial Phase Rotation for Full-Diversity M2M Transmission

Koike-Akino, T.; Kim, K.J.; Orlik, P.V.; Pajovic, M.

TR2015-130 December 2015

**Abstract**

Machine-to-machine (M2M) communications have been considered as an important application to connect a massively large number of different devices in networks. In particular for M2M wireless networks, low latency and high reliability are of great importance. To fulfill the requirements, a diversity technique exploiting limited resources is proposed in this paper for short-message transmissions. The proposed method uses multiple stages of fast unitary transforms and diagonal phase rotations to achieve full-diversity gain. A monomial phase rotation is also proposed to facilitate an optimization of the precoding matrix. It is verified that the proposed four-stage precoding provides universal diversity gain irrespective of channel selectivity in time and frequency, without changing monomial parameters. In addition, it is shown that the four-stage precoding based on the discrete Haar transform (DHT) achieves a full diversity while the computational complexity is significantly reduced from a loglinear order to a linear order compared to the other unitary transforms such as the discrete Fourier transform (DFT).

*2015 IEEE Global Communications Conference (GLOBECOM)*

# Universal Multi-Stage Precoding with Monomial Phase Rotation for Full-Diversity M2M Transmission

Toshiaki Koike-Akino, Keyong Jin Kim, Milutin Pajovic, and Philip V. Orlik
Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA 02139, USA

***Abstract*—Machine-to-machine (M2M) communications have been considered as an important application to connect a massively large number of different devices in networks. In particular for M2M wireless networks, low latency and high reliability are of great importance. To fulfill the requirements, a diversity technique exploiting limited resources is proposed in this paper for short-message transmissions. The proposed method uses multiple stages of fast unitary transforms and diagonal phase rotations to achieve full-diversity gain. A monomial phase rotation is also proposed to facilitate an optimization of the precoding matrix. It is verified that the proposed four-stage precoding provides *universal* diversity gain irrespective of channel selectivity in time and frequency, without changing monomial parameters. In addition, it is shown that the four-stage precoding based on the discrete Haar transform (DHT) achieves a full diversity while the computational complexity is significantly reduced from a log-linear order to a linear order compared to the other unitary transforms such as the discrete Fourier transform (DFT).**

## I. Introduction

Machine-to-machine (M2M) communications systems are intended to enable machines to exchange short command and control messages over wireless links. One of the major applications of M2M communications are factories wherein the automated production processes would benefit if machines communicate with each other wirelessly. Such applications need to achieve both low latency and high reliability for transmitting short messages over wireless channels [1]. It is often preferred to use only one or two antennas for small and simple devices.

We consider a single-antenna system with block transmission as a candidate for practical M2M communications. The low latency requirement may call for short block sizes, e.g., less than $100$ symbols. For such short-message communications, it is difficult to fulfill the high reliability requirement. It is because we cannot use capacity-achieving error correction codes including low-density parity-check (LDPC) codes, whose length is typically longer than $10000$ bits to approach capacity. In addition, a single-antenna system cannot exploit spatial diversity techniques [2]–[5] to deal with channel fading.

High-dimensional modulation (HDM) schemes [6]–[8] have been proposed to provide reliable communications for short-message M2M communications. For example, a greedy sphere packing has been proposed in [6] to design best-possible short block constellations having a larger Euclidean distance by using densest known lattices. In [7] and [8], alternative HDM methods have been proposed based on linear block codes and nonlinear block codes, respectively. It was shown that the HDM based on the Nordstrom–Robinson (NR) nonlinear code [8]–[10] provides a larger coding gain compared to the best known linear code. Although the HDM can offer larger coding gain, the HDM itself cannot achieve a full diversity for fading channels.

One potential method to overcome fading for reliable M2M communications has been reported in [11], where pseudo-random phase precoding (PRPP) was proposed. The PRPP uses randomly rotated constellations to achieve a higher diversity gain, similar to the rotated-constellation precoding for space-time diversity [4]. By exploiting an analogy between single-antenna block transmission and multiple-input multiple-output (MIMO) transmission, an MIMO detection scheme based on likelihood ascent search (LAS) [12] was employed. This scheme performs well in practice when the block size is at least $400$ symbols. However, our goal is to reduce the block size without losing much of the performance.

Maximum-likelihood detection (MLD) is optimal for symbol detection. However, the complexity of the MLD grows exponentially with the block size. Therefore, a variety of suboptimal detection algorithms with a reduced complexity have been developed, e.g., [12]–[16]. For example, a symbol ordering method was proposed in [17] for probabilistic data association (PDA) detection [13]. In [18], an improved PDA algorithm was proposed for short block sizes. However, if the block size is very short, the MLD is still feasible. Therefore, we use the MLD for simplicity of analysis, and focus on how to design the precoding at the transmitter side in order to achieve a higher diversity gain.

The existing precoder based on PRPP [11] has two major drawbacks: i) there is a non-zero probability that the precoding matrix becomes singular to lose the diversity because of the randomness, and ii) the precoding matrix is almost surely non-unitary and, thus, the transmission signal power can unpredictably vary. Those issues can be solved by using unitary transforms and diagonal phase rotation as in [4] and [5]. The discrete Fourier transform (DFT) and the discrete Walsh–Hadamard transform (DWT) have been investigated in conjunction with diagonal phase rotation for two-stage precoding in [4] and [5], respectively. However, there is few literature considering more than two stages for precoding design. In addition, there is no comparative study considering different unitary transforms. A Grassmannian precoding design was also discussed in [2], where a limited feedback channel is considered for precoder adaptation. However, such an adaptive precoding requires a complicated feedback protocol. For some applications, it may be preferred to avoid complicated adaptation and to keep using the same precoding matrix regardless of the channel conditions. Hence, we focus on non-adaptive precoding.

In this paper, we propose a new precoding method by extending the existing precoders in several directions. Compared to existing works, some novel contributions of this paper are summarized below.

- **Monomial Phase Rotation**: We propose to use a monomial function parameterized by two factors to generate deterministic phase rotations. This parametric function enables us to optimize the precoding matrix to achieve a higher diversity gain.
- **Multi-Stage Precoding**: We propose multiple stages of fast unitary transforms to distribute diagonal phase rotations across different symbols. In particular, the four-stage precoding achieves a full diversity and shows significant robustness against channel selectivity in time and frequency.
- **Fast Unitary Transform**: We make performance comparison for various different unitary transforms [19], which are fast transformable, including the discrete cosine transform (DCT), discrete slant transform (DST), and discrete Haar transform (DHT) as well as the DFT and DWT. We show that the four-stage precoding based on the DHT has a great advantage in terms of computational complexity while maintaining its achievable full diversity.
- **High-Dimensional Modulation**: We incorporate the HDM with precoding to improve the diversity gain. A joint precoding and HDM design based on the NR nonlinear block code [8]–[10] provides an additional diversity gain.

## II. Multi-Stage Precoding

### A. System Model

We assume M2M communications from a transmitter to a receiver, each of which is equipped with a single antenna, over Rayleigh fading channels. Fig. 1 shows a schematic of the M2M communications system. The transmitter uses a short coded modulation format to generate an $N$-symbol block $\boldsymbol{x} = [x_0, x_1, \ldots, x_{N-1}]^{\mathrm{T}}$, where $x_n \in \mathbb{C}$ is the $n$th constellation ($\mathbb{C}$ is a complex-valued set) and the superscript $[\cdot]^{\mathrm{T}}$ denotes a transpose. We assume that the energy of modulated signals is normalized as $\frac{1}{N}\mathbb{E}\big[\|\boldsymbol{x}\|^2\big] = 1$, where $\mathbb{E}[\cdot]$ and $\|\cdot\|$ denote the expectation and the Frobenius norm, respectively. From a low latency requirement, we suppose that any long error correction codes such as capacity-approaching LDPC codes are not available.

Before transmitting $\boldsymbol{x}$, the signals are precoded by a precoding matrix $\boldsymbol{P} \in \mathbb{C}^{N\times N}$ to deal with deep fading. Through wireless fading channels, the receiver obtains the received signal block $\boldsymbol{y} = [y_0, y_1, \ldots, y_{N-1}]^{\mathrm{T}} \in \mathbb{C}^{N\times 1}$, which is modeled as

$$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{P}\boldsymbol{x} + \boldsymbol{z}, \tag{1}$$

where $\boldsymbol{G} \in \mathbb{C}^{N\times N}$ is a channel matrix and $\boldsymbol{z} = [z_0, z_1, \ldots, z_{N-1}]^{\mathrm{T}} \in \mathbb{C}^{N\times 1}$ is an additive noise vector. The noise vector $\boldsymbol{z}$ comes from a zero-mean circular-symmetric complex Gaussian distribution of $\mathcal{CN}(\boldsymbol{0}, \sigma^2\boldsymbol{I}_N)$ having a covariance matrix of $\sigma^2\boldsymbol{I}_N$, where $\boldsymbol{I}_N$ is an identity matrix of size $N \times N$.

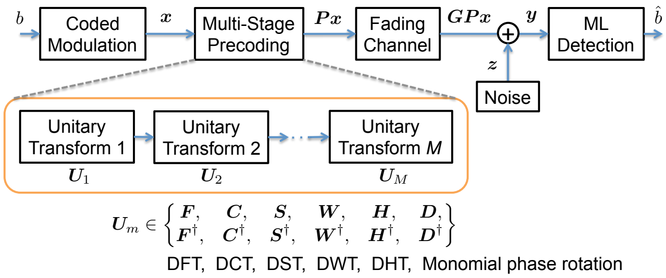


Fig. 1. Schematic of M2M communications systems employing coded modulation and $M$-stage precoding based on fast unitary transforms including DFT, DCT, DST, DWT, DHT, and monomial phase rotation.

For time-selective fading channels, we model the channel matrix $\boldsymbol{G}$ as follows:

$$\boldsymbol{G} = \mathrm{diag}\big[g_0, g_1, \ldots, g_{N-1}\big], \tag{2}$$

where $g_n \in \mathbb{C}$ is the $n$th channel gain following a zero-mean complex Gaussian distribution according to Rayleigh fading, and $\mathrm{diag}[\cdot]$ denotes a diagonal matrix, whose diagonal entry is formed by an argument vector. The channel gain $g_n$ is correlated each other. More specifically, the correlation between $g_n$ and $g_m$ is well determined by a Jakes model as $\mathbb{E}[g_n g_m^*] = J_0(2\pi|n-m|f_{\mathrm{D}}T_{\mathrm{s}})$, where $[\cdot]^*$ denotes a complex conjugate, $J_0(\cdot)$ is the zero-th order Bessel function of the first kind, $f_{\mathrm{D}}$ is the maximum Doppler frequency, and $T_{\mathrm{s}}$ is the symbol duration.

If the fading channel is frequency-selective, we can alternatively model the channel matrix $\boldsymbol{G}$ as follows:

$$\boldsymbol{G} = \boldsymbol{F}^{\dagger}\mathrm{diag}\big[g_0, g_1, \ldots, g_{N-1}\big]\boldsymbol{F}, \tag{3}$$

where $\boldsymbol{F}$ is the DFT matrix of size $N \times N$, and a superscript $[\cdot]^{\dagger}$ denotes the Hermitian transpose. Here, we assume a cyclic prefix to prevent inter-block interference. The DFT matrix is defined as $[\boldsymbol{F}]_{m,n} = \frac{1}{\sqrt{N}}\exp(-\jmath\frac{2\pi mn}{N})$, where $[\cdot]_{m,n}$ denotes the $(m,n)$th entry of an argument matrix and $\jmath = \sqrt{-1}$ is an imaginary unit. In this case, $g_n$ represents the $n$th frequency-domain channel gain, and the correlation among $g_n$ depends on the power delay profile of the channel. The correlation is typically characterized by an exponential function.

### B. High-Dimensional Modulation (HDM)

For most M2M applications, high-rate communications are not required, whereas high reliability and low latency are stringent. Hence, low-order modulation schemes such as binary/quadrature phase-shift keying (BPSK/QPSK) have often been used for signals $\boldsymbol{x}$. To improve the reliability, high-dimensional modulation (HDM) schemes [6]–[8] have been studied as an advanced modulation. For example, a sphere cutting method has been proposed in [6] to design lattice-packed HDM for short-message reliable M2M communications, by using densest known lattices [20], e.g., hexagonal $A_2$ lattice, checker-board $D_4$ lattice, diamond $E_8$ lattice, Coxeter-Todd $K_{12}$ lattice, Barnes-Wall $\Lambda_{16}$ lattice, and Leech $\Lambda_{24}$ lattice. In [7], alternative HDM schemes based on short linear block codes have been proposed to increase the minimum Euclidean distance. For example, a 24-dimensional modulation using the extended Golay code achieves 6 dB gain in the squared minimum Euclidean distance compared to BPSK.

In [8], a 16-dimensional HDM scheme based on the NR nonlinear block code [9] has been investigated. The NR code improves the squared minimum Euclidean distance by $0.8\,$dB compared to the best known linear code. We use a systematic NR code, which is a binary image of an octacode with a generator matrix defined as

$$\boldsymbol{G}_{\mathrm{NR}} = \begin{bmatrix} 1 & 3 & 1 & 2 & 1 & 0 & 0 & 0 \\ 1 & 0 & 3 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 3 & 1 & 2 & 1 & 0 \\ 1 & 0 & 0 & 0 & 3 & 1 & 2 & 1 \end{bmatrix}. \tag{4}$$

As shown in [10], the NR codeword is derived by multiplying the generator matrix $\boldsymbol{G}_{\mathrm{NR}}$ over the $\mathbb{Z}_4$ Lie ring, followed by a Gray mapping.

The benefit of HDM is three-fold as follows:

- We can improve a coding gain because of the increased Euclidean distance.
- We can improve the diversity gain because of the interdependent constellations.
- We can prevent severe diversity loss when channel selectivity is changed after moving to a different environment.

The achievable coding gain and higher diversity order are important for reliable M2M communications because long and strong error correction codes are not available. NR-coded 16-dimensional QPSK modulation itself provides a diversity order of four. This may facilitate the precoding design since some of the diversity gain can come from the HDM and we can relax the requirement that the precoder alone achieves a full diversity.

### C. Multi-Stage Precoding

In this paper, we propose a multi-stage precoding, whose precoding matrix $\boldsymbol{P}$ is a cascaded product of $M$ unitary transforms as follows:

$$\boldsymbol{P} = \boldsymbol{U}_M \boldsymbol{U}_{M-1} \cdots \boldsymbol{U}_1, \tag{5}$$

where $M$ is the number of stages, and $\boldsymbol{U}_m$ is the $m$th unitary transform, which is either the DFT, DCT, DST, DWT, DHT, diagonal phase rotation, or inverse transform (i.e., conjugate transpose) thereof. Although there are many different combinations to generate the precoding matrix, we are particularly interested in the following precoders for comparison:

- 1-stage precoding of the form $\boldsymbol{P} = \boldsymbol{U}$,
- 2-stage precoding of the form $\boldsymbol{P} = \boldsymbol{U}\boldsymbol{D}$,
- 3-stage precoding of the form $\boldsymbol{P} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^\dagger$, and
- 4-stage precoding of the form $\boldsymbol{P} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^\dagger\boldsymbol{D}^\dagger$,

for a given unitary transform $\boldsymbol{U}$ and a diagonal phase rotation $\boldsymbol{D}$. Note that most existing precoding methods [4], [5] consider at most two stages. As we will see later, the four-stage precoding has a significant advantage.

### D. Fast Unitary Transforms

In order to keep the computational complexity low, we consider precoding based on fast-transformable unitary matrices [19]: the DFT matrix $\boldsymbol{F}$, DCT matrix $\boldsymbol{C}$, DST matrix $\boldsymbol{S}$, DWT matrix $\boldsymbol{W}$, and DHT matrix $\boldsymbol{H}$ as briefly defined below. Note that the above transforms are listed in order of decreasing computational complexity. The DHT is the most attractive transform because of the lowest complexity with a linear order of $\mathcal{O}[N]$.

*1) **Discrete Fourier Transform (DFT):*** The DFT unitary transform $\boldsymbol{F}$ has been widely used in a number of applications. In particular for radio communications, the DFT has often been used to convert a frequency-selective channel into parallel frequency-flat subchannels as in orthogonal frequency-division multiplexing (OFDM). The DFT is fast-transformable in a Cooley-Tukey architecture with a log-linear complexity order of $\mathcal{O}[N\log_2 N]$. However, the DFT requires complex-valued multiplications unlike the other unitary transforms below.

*2) **Discrete Cosine Transform (DCT):*** The DCT matrix $\boldsymbol{C}$ of type II is expressed as

$$[\boldsymbol{C}]_{m,n} = c_m \frac{2}{\sqrt{N}} \cos \frac{\pi m(2n+1)}{2N}, \tag{6}$$

where $c_0 = 1/\sqrt{2}$ and $c_m = 1$ for $m \neq 0$. The DCT has also a Cooley-Tukey-type architecture with a log-linear complexity order of $\mathcal{O}[N\log_2 N]$. However, the DCT is simpler than the DFT because the DCT is real-valued.

*3) **Discrete Slant Transform (DST):*** The DST matrix $\boldsymbol{S}$ is defined by the following recursion:

$$\boldsymbol{S}_k = \frac{1}{\sqrt{2}} \begin{bmatrix} \begin{matrix} 1 & 0 \\ a_k & b_k \end{matrix} & \boldsymbol{0} & \begin{matrix} 1 & 0 \\ -a_k & b_k \end{matrix} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{2^{k-1}-2} & \boldsymbol{0} & \boldsymbol{I}_{2^{k-1}-2} \\ \begin{matrix} 0 & 1 \\ -b_k & a_k \end{matrix} & \boldsymbol{0} & \begin{matrix} 0 & -1 \\ b_k & a_k \end{matrix} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{2^{k-1}-2} & \boldsymbol{0} & -\boldsymbol{I}_{2^{k-1}-2} \end{bmatrix} \boldsymbol{S}_{k-1},$$

$$\boldsymbol{S}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \; a_k = \sqrt{\frac{3 \cdot 4^{k-1}}{4^k - 1}}, \; b_k = \sqrt{\frac{4^{k-1} - 1}{4^k - 1}}.$$

The computational complexity of the DST is also log-linear of $\mathcal{O}[N\log_2 N]$ with a Cooley-Tukey-type algorithm. Since the DST has piecewise linear basis functions, it is slightly simpler than the DCT.

*4) **Discrete Walsh–Hadamard Transform (DWT):*** The DWT matrix $\boldsymbol{W}$ is also generated recursively as follows

$$\boldsymbol{W}_k = \boldsymbol{W}_1 \otimes \boldsymbol{W}_{k-1}, \quad \boldsymbol{W}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \tag{7}$$

where $\otimes$ denotes a Kronecker product. The DWT is a symmetric unitary transform, whose entry consists of $\pm 1$ excluding a normalization term of $\sqrt{N}$. Although the computational complexity of the DWT is log-linear $\mathcal{O}[N\log_2 N]$, no multiplications are needed except for the normalization. Consequently, the DWT is even simpler than the DST.

*5) **Discrete Haar Transform (DHT):*** The DHT based on a rectangular mother wavelet has a unitary matrix $\boldsymbol{H}$, defined recursively as follows

$$\boldsymbol{H}_k = \frac{1}{\sqrt{2}} \begin{bmatrix} \boldsymbol{H}_{k-1} \otimes [1 \quad 1] \\ \boldsymbol{I}_{2^{k-1}} \otimes [1 \quad -1] \end{bmatrix}, \quad \boldsymbol{H}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{8}$$

The DHT is the lowest complexity transform among the unitary transforms considered in this paper because the basis functions contain a lot of zero entries while non-zero entries are $\pm 1$ without normalization like the DWT. Although the DHT has also a Cooley-Tukey-type structure, its computational complexity is significantly reduced to a linear order $\mathcal{O}[N]$ not a log-linear order $\mathcal{O}[N \log_2 N]$ because of zero entries. This is a great advantage against some existing precoders based on the DFT [4] or the DWT [5].

### E. Monomial Phase Rotation

As an additional unitary transform, we also consider a diagonal phase rotation, whose matrix is expressed as

$$\boldsymbol{D} = \mathrm{diag}\big[\exp(\jmath\theta_0), \exp(\jmath\theta_1), \ldots, \exp(\jmath\theta_{N-1})\big], \quad (9)$$

where the $n$th phase $\theta_n$ is pre-determined. Rather than using a pseudo-random number generator like the PRPP [11], we propose the use of a monomial function parameterized by a scale factor $\alpha$ and an exponent factor $\beta$ as follows:

$$\theta_n = \pi\, \alpha\, n^{\beta}. \quad (10)$$

More precisely, the monomial function requires only integers for $\beta$, and it becomes a power function otherwise. However, we call it the monomial function regardless of $\beta$ for convenience.

As we will evaluate later, this parametric phase rotation based on the monomial function is simple but effective to design the precoding matrix $\boldsymbol{P}$. Note that the computational complexity of the diagonal phase rotation is also linear order of $\mathcal{O}[N]$ as the DHT.

### F. Bit-Error Rate (BER) Union Bound

We aim at optimizing the precoding matrix in terms of a minimum union bound of bit-error rate (BER) performance. The union bound is expressed as

$$p_{\mathrm{union}} = \frac{1}{B} \sum_{b,b'} d_{\mathrm{H}}(b, b')\, p(b, b'), \quad (11)$$

where $B$ is the number of bits in a coded modulation, $d_{\mathrm{H}}(b, b')$ denotes the Hamming distance between bit indices $b$ and $b'$ (for $b, b' \in \mathbb{Z}_B$), and $p(b, b')$ is a pairwise error probability (PEP). The PEP is expressed as

$$p(b, b') = \mathop{\mathbb{E}}_{\boldsymbol{G}} \left[ \frac{1}{2} \mathrm{erfc} \sqrt{\frac{1}{\sigma^2} \big\| \boldsymbol{G}\boldsymbol{P}\big(\boldsymbol{x}(b) - \boldsymbol{x}(b')\big) \big\|^2} \right], \quad (12)$$

where $\mathrm{erfc}(\cdot)$ is a complementary error function, and $\boldsymbol{x}(b)$ is a coded modulation constellation for the bit index $b$.

In *i.i.d.* Rayleigh fading channels for $g_n$, we obtain the PEP approximated by

$$p(b, b') \simeq \frac{1}{12} \prod_n \frac{1}{1 + d_n^2(b, b')\rho} + \frac{1}{4} \prod_n \frac{1}{1 + \frac{4}{3} d_n^2(b, b')\rho}, \quad (13)$$

where we used an approximation of $\mathrm{erfc}(x) \simeq \frac{1}{6}\exp(-x^2) + \frac{1}{2}\exp(-\frac{4}{3}x^2)$. Here, $\rho$ is an average signal-to-noise ratio (SNR), and $d_n(b, b')$ is defined as

$$d_n^2(b, b') = \big| \boldsymbol{e}_n^{\mathrm{T}} \boldsymbol{P}\big(\boldsymbol{x}(b) - \boldsymbol{x}(b')\big) \big|^2, \quad (14)$$

where $\boldsymbol{e}_n$ is a standard unit vector, whose elements are all zeros but one at the $n$th entry.

**Algorithm 1** Universal precoding design for different channel selectivity

**Require:** Modulation format $\{\boldsymbol{x}(b)\}$, selected unitary transforms $\{\boldsymbol{U}_m\}$ for $m \in \{1, 2, \ldots, M\}$, average SNR $\rho$, channel selectivity set $\boldsymbol{G} \in \mathcal{G}$
  **for all** monomial exponent factor $\beta$ **do**
    **for all** monomial scale factor $\alpha$ **do**
      generate monomial phase rotations $\boldsymbol{D}$ in (9) and (10)
      construct precoding matrix $\boldsymbol{P}$ in (5)
      **for all** channel selectivity set $\mathcal{G}$ **do**
        calculate union bound $p_{\mathrm{union}}$ in (11) and (13)
      **end for**
      **print** the worst-case union bound in a loop of $\mathcal{G}$
    **end for**
  **end for**
  **return** the best pair of $\alpha$ and $\beta$ minimizing the worst-case union bound

### G. Universal Precoding Design

We optimize the multi-stage precoding with monomial phase rotations $\boldsymbol{D}$ to minimize the union bound $p_{\mathrm{union}}$. Because the monomial function in (10) has two factors $\alpha$ and $\beta$, we just need to sweep the values across a two-dimensional grid to search for the best pair of the factors. We consider the scan ranges for $0 \leq \alpha \leq 1$ with a grid step of 0.001, and for $0.25 \leq \beta \leq 4$ with a grid step of 0.25 as an example. Given modulation format $\boldsymbol{x}(b)$ and average SNR $\rho$, the union bound $p_{\mathrm{union}}$ in (11) is calculated with PEP expression in (13) for each pair of monomial factors $\alpha$ and $\beta$ to find the optimal values minimizing the union bound.

It was shown in [18] that a precoder based on the DFT offers diversity gain in time-selective fading channels whereas the diversity gain is seriously lost in frequency-selective fading channels. It suggests that the precoding matrix should be changed according to the channel selectivity. However, such an adaptive use of different precoders may have a drawback for some applications, where the channel selectivity is unknown at the time of system deployment.

In order to be robust against different conditions of channel selectivity, we propose a universal precoding design method, whose pseudo-code is described in Algorithm 1. This method searches for the best pair of monomial factors (i.e., $\alpha$ and $\beta$) to minimize the worst-case union bound for a different channel conditions set $\mathcal{G}$. In this paper, we focus on two specific channel conditions; *i.i.d.* time-selective fading as in (2) and *i.i.d.* frequency-selective fading as in (3). As shown in (2) and (3), the frequency-selective channel model effectively adds up an additional unitary transform of $\boldsymbol{F}$ for a given precoding matrix $\boldsymbol{P}$ as $\boldsymbol{F}\boldsymbol{P}$. Correspondingly, monomial factors are optimized by considering the worst-case union bound for both an original precoder $\boldsymbol{P}$ and an effective precoder $\boldsymbol{F}\boldsymbol{P}$ in the universal precoding design.

## III. Performance Evaluations

### A. Block Size

We first evaluate the impact of block size $N$ in Fig. 2, where BER performance as a function of the average SNR is plotted under *i.i.d.* time-selective Rayleigh fading channels for

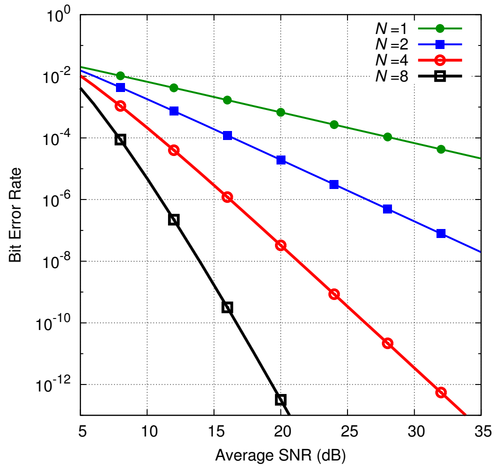


Fig. 2. BER performance of BPSK with 2-stage DFT precoding for different block size of $N = 1, 2, 4, 8$ in Rayleigh fading channels ($\alpha = 1/N$ and $\beta = 1$).

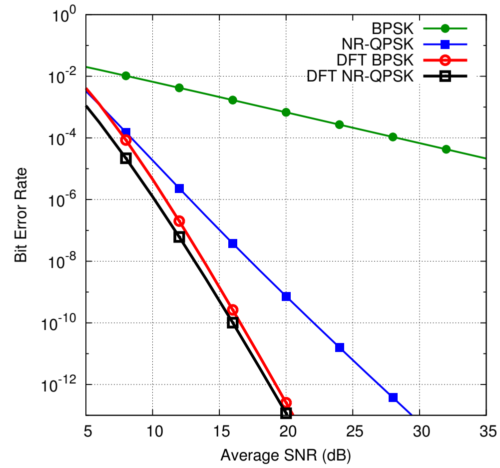


Fig. 4. BER performance for 2-stage DFT precoding with BPSK or NR-coded QPSK in Rayleigh fading channels.

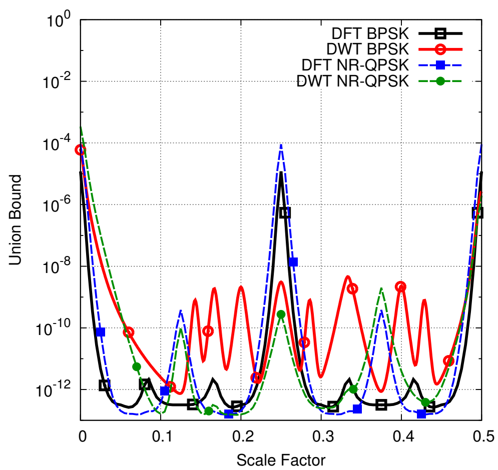


Fig. 3. Union bound vs. monomial scale factor $\alpha$ for 2-stage precoding based on DFT or DWT with BPSK or NR-coded QPSK in Rayleigh fading channels ($\beta = 1$ and $N = 8$).

a block size of $N = 1, 2, 4, 8$ using two-stage precoding based on the DFT. Here, we use suboptimal monomial factors of $\alpha = 1/N$ and $\beta = 1$ for BPSK modulations. It is observed in Fig. 2 that a higher diversity gain is difficult to be achieved with precoding for shorter block sizes. This result shows difficulty in satisfying low latency and high reliability at the same time for M2M communications. We consider a block size of $N = 8$ in the following evaluations.

### B. Monomial Factor and HDM

We next show the impact of the monomial factor with/without HDM for two-stage precoding based on the DFT or the DWT in Fig. 3, where a union bound as a function of a scale factor $\alpha$ is present. Here, we consider an exponent factor of $\beta = 1$ as an example. We assume BPSK for the case without HDM, and NR-coded QPSK as HDM. Note that spectral efficiency is identical to 1 bps/Hz for both cases. Some key results in Fig. 3 are summarized as follows:

- The diagonal phase rotations with properly chosen monomial factors can significantly improve the union bound compared to the schemes without rotation (i.e., 1-stage precoding) at $\alpha = 0$.
- If the phase rotation is randomly chosen like PRPP [11], the union bound can be seriously degraded by several magnitudes in a certain probability.
- The DFT precoding slightly outperforms the DWT precoding for BPSK when we use a linearly increasing phase rotation with $\beta = 1$.
- Although the benefit of the HDM is marginal, the minimum union bound of the DWT precoding becomes comparable to that of the DFT precoding with HDM.

In Fig. 4, the BER performance curves of BPSK and NR-coded QPSK are shown with and without the two-stage DFT precoding. For the case with precoding, we use optimized monomial factors of $\alpha = 0.435, \beta = 1.5$ for BPSK, and $\alpha = 0.619, \beta = 3.25$ for NR-coded QPSK. From Fig. 4, we can observe the following results.

- The HDM based on the NR code itself achieves a higher diversity gain (4-order diversity) than BPSK when precoding is not employed.
- The two-stage precoding with optimized monomial factors provides a significant improvement in BER achieving a full diversity order of 8 for both BPSK and HDM.
- The NR-coded QPSK has a very marginal gain compared to BPSK when precoding is employed.

### C. Unitary Transforms and Universal Diversity

In Table I, we list the union bound achieved by different unitary transforms (i.e., DFT, DCT, DST, DWT, and DHT) for two-stage precoding, in which the monomial factors are optimized for time-selective fading channels at an average SNR of 20 dB, and the HDM based on the NR code is used. It is seen that a good performance of a union bound below $10^{-12}$ is obtained for each unitary transform for time-selective fading channels. However, if we keep using the same precoding for different channel conditions, the performance can be considerably degraded for some unitary transforms.

TABLE I. UNION BOUND WITH 2-STAGE PRECODING OPTIMIZED FOR TIME-SELECTIVE FADING CHANNELS

| 2-stage precoding $\boldsymbol{P} = \boldsymbol{U}\boldsymbol{D}$ | | | | | |
|---|---|---|---|---|---|
| | DFT | DCT | DST | DWT | DHT |
| Monomial factor $(\alpha, \beta)$ | $0.619, 3.25$ | $0.87, 3.5$ | $0.142, 2.5$ | $0.086, 2$ | $0.375, 3$ |
| Union bound (time selective) | $1.2 \times 10^{-13}$ | $1.3 \times 10^{-13}$ | $1.4 \times 10^{-13}$ | $8.7 \times 10^{-14}$ | $6.5 \times 10^{-13}$ |
| Union bound (freq. selective) | $7.3 \times 10^{-10}$ | $1.9 \times 10^{-12}$ | $4.8 \times 10^{-13}$ | $1.2 \times 10^{-10}$ | $3.4 \times 10^{-13}$ |

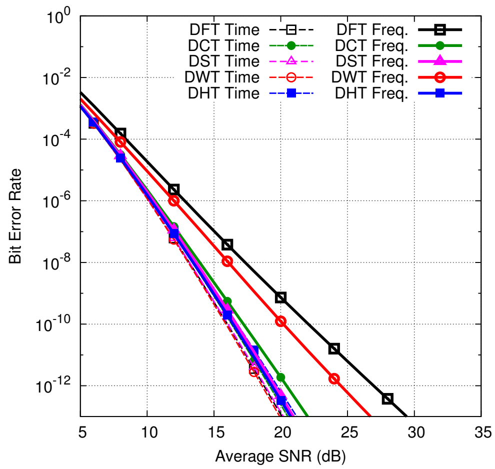


Fig. 5. BER performance for 2-stage precoding with different unitary transforms for NR-coded QPSK in time-selective or frequency-selective Rayleigh fading channels.

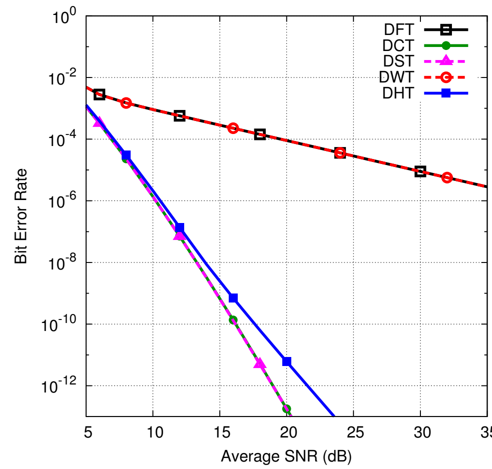


Fig. 6. BER performance for 3-stage precoding with different unitary transforms for NR-coded QPSK in worst case between time-selective and frequency-selective Rayleigh fading channels.

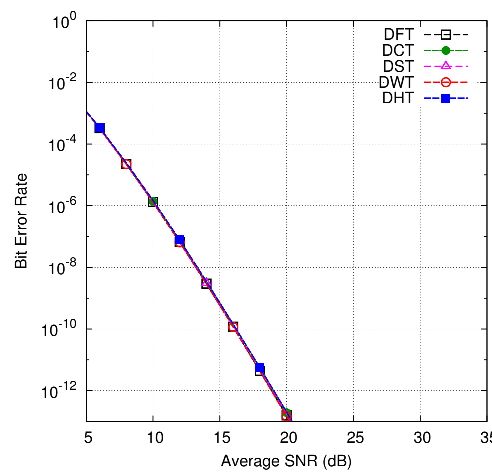


Fig. 7. BER performance for 4-stage precoding with different unitary transforms for NR-coded QPSK in worst case between time-selective and frequency-selective Rayleigh fading channels.

For example, the precoding based on the conventional DFT or DWT degrades the union bound to higher than $10^{-10}$ for frequency-selective fading channels.

This is more clearly shown in Fig. 5, where the BER performance of two-stage precoding with different unitary transforms is present for time-selective fading channels or frequency-selective fading channels. Here, we use monomial factors listed in Table I. Some key observations in Fig. 5 are described as follows.

- The two-stage precoding with optimized monomial factors offers full-diversity performance for time-selective fading channels regardless of unitary transforms.
- The precoding based on the DFT, DWT, or DCT can suffer from diversity loss at frequency-selective fading channels.
- The precoding based on the DST or the DHT shows robustness against difference of channel selectivity.

We now investigate the performance benefit of the universal multi-stage precoding design, proposed in this paper to achieve diversity gain for different channel selectivity. In Table II, we show a special example of the universal precoding design, which optimized the monomial factors ($\alpha$ and $\beta$) for different unitary transforms to minimize the worst-case union bound between *i.i.d.* time-selective Rayleigh fading channels and *i.i.d.* frequency-selective Rayleigh fading channels. The performance curves for 3-stage and 4-stage precoding with the optimized monomial factors are plotted, respectively, in Figs. 6 and 7, which show the worst-case BER performance between time-selective and frequency-selective fading channels. We can see the following results.

- The precoding with the DFT or DWT cannot universally achieve a high diversity gain for less-than four stages.
- The precoding based on the DST exhibits universal diversity for $M \geq 2$ stages.
- The 4-stage precoding provides low union bound below $2 \times 10^{-13}$, and universally achieves full diversity no matter which unitary transforms are employed.
- In particular, the 4-stage precoding based on the DHT can be an attractive candidate for short-message M2M communications since full diversity, low latency, and low complexity are simultaneously achieved.

The universal diversity gain can be explained by the fact that the modulated symbols are uniformly well distributed via

TABLE II. UNIVERSAL PRECODING DESIGN RESULTS FOR WORST-CASE UNION BOUND BETWEEN TIME- AND FREQUENCY-SELECTIVE CHANNELS

| 1-stage precoding $\boldsymbol{P}=\boldsymbol{U}$ | | | | | |
|---|---|---|---|---|---|
| | DFT | DCT | DST | DWT | DHT |
| Worst-case union bound | $9.0\times10^{-5}$ | $9.0\times10^{-5}$ | $1.4\times10^{-4}$ | $3.4\times10^{-4}$ | $9.0\times10^{-5}$ |
| 2-stage precoding $\boldsymbol{P}=\boldsymbol{U}\boldsymbol{D}$ | | | | | |
| | DFT | DCT | DST | DWT | DHT |
| Monomial factor $(\alpha,\beta)$ | $0.928, 1$ | $0.495, 3$ | $0.932, 3.75$ | $0.972, 2.5$ | $0.375, 3$ |
| Worst-case union bound | $7.3\times10^{-10}$ | $4.1\times10^{-13}$ | $2.6\times10^{-13}$ | $3.0\times10^{-11}$ | $6.5\times10^{-13}$ |
| 3-stage precoding $\boldsymbol{P}=\boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^\dagger$ | | | | | |
| | DFT | DCT | DST | DWT | DHT |
| Monomial factor $(\alpha,\beta)$ | $0.37, 0.75$ | $0.97, 1.75$ | $0.955, 2.25$ | $0.835, 3$ | $0.25, 4$ |
| Worst-case union bound | $9.0\times10^{-5}$ | $1.8\times10^{-13}$ | $1.8\times10^{-13}$ | $9.0\times10^{-5}$ | $6.1\times10^{-12}$ |
| 4-stage precoding $\boldsymbol{P}=\boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^\dagger\boldsymbol{D}^\dagger$ | | | | | |
| | DFT | DCT | DST | DWT | DHT |
| Monomial factor $(\alpha,\beta)$ | $0.385, 2.5$ | $0.769, 4$ | $0.88, 3.75$ | $0.802, 2.5$ | $0.78, 1.25$ |
| Worst-case union bound | $1.6\times10^{-13}$ | $1.9\times10^{-13}$ | $1.7\times10^{-13}$ | $1.5\times10^{-13}$ | $2.0\times10^{-13}$ |

many-stage precoding onto all possible bases to be robust against various fading realizations. Although we assumed MLD for analysis under idealistic *i.i.d.* Rayleigh fading channels, it is expected that a significant gain is still possible for realistic channel environments with low-complexity PDA detection as we have shown in [18].

## IV. CONCLUSIONS

We have proposed multiple stages of fast unitary transforms in the precoding design to achieve a higher diversity gain. For low-latency and highly reliable M2M communications, we have used a short block-coded modulation based on the Nordstrom–Robinson nonlinear code. By introducing a monomial phase rotation, we can optimize the precoding matrix, which universally achieves a full diversity gain for both time-selective and frequency-selective fading channels without changing the precoding matrix. It has been verified that the four-stage precoding offers such a universal diversity gain irrespective of the considered unitary transforms. For low-complexity applications, the precoding based on the discrete Haar transform has a great potential compared to the one using the other transforms including the discrete Fourier transform because the computational complexity becomes linear order. More practical performance evaluations in doubly-selective fading channels remain as future works.

## ACKNOWLEDGMENT

The authors would like to thank Dr. Kieran Parsons at MERL for valuable advice.

## REFERENCES

[1] G. Wu, S. Talwar, K. Johnsson, N. Himayat, and K. Johnson, “M2M: From mobile to embedded internet,” *IEEE Commun. Magazine*, vol. 49, no. 4, pp. 36–43, 2011.

[2] D. J. Love and R. W. Heath, “Limited feedback unitary precoding for spatial multiplexing systems,” *IEEE Trans. Inform. Theory*, vol. 51, no. 8, pp. 2967–2976, Aug. 2005.

[3] R. W. Heath and A. J. Paulraj, “Linear dispersion codes for MIMO systems based on frame theory,” *IEEE Trans. Signal Proccess.*, vol. 50, no. 10, pp. 2429–2441, Oct. 2002.

[4] Y. Xin, Z. Wang, and G. B. Giannakis, “Space-time diversity systems based on linear constellation precoding,” *IEEE Trans. Wireless Commun.*, vol. 2, no. 2, pp. 294–304, Mar. 2003.

[5] M. O. Damen, K. Abed-Meraim, and J. C. Belfiore, “Transmit diversity using rotated constellations with Hadamard transform,” *Adaptive Systems for SP, Com., and Control Conf.*, Lake Louise, Alberta, Canada, Oct. 2000.

[6] T. Koike-Akino and V. Tarokh, “Sphere packing optimization and EXIT chart analysis for multidimensional QAM signaling,” *IEEE Int. Conf. Commun. (ICC)*, June 2009.

[7] D. S. Millar, T. Koike-Akino, S. Ö. Arik, K. Kojima, K. Parsons, T. Yoshida, and T. Sugihara, “High-dimensional modulation for coherent optical communications systems,” *Optics Express*, vol. 22, no. 7, pp. 8798–8812, Apr. 2014.

[8] T. Koike-Akino, D. S. Millar, K. Kojima, K. Parsons, K. Sugihara, Y. Miyata, and T. Yoshida, “LDPC-coded 16-dimensional modulation based on the Nordstrom-Robinson nonlinear block code,” *OSA Conf. Laser Electro-Optics (CLEO)*, JTh2A.24, May 2015.

[9] A. W. Nordstrom and J. P. Robinson, “An optimum nonlinear code,” *Information and Control*, vol. 11, no. 5, pp. 613–616, Nov.-Dec. 1967.

[10] G. D. Forney Jr., N. J. A. Sloane, and M. D. Trott, “The Nordstrom-Robinson code is the binary image of the octacode,” *Coding and Quantization: DIMACS/IEEE Workshop*, pp. 19–26, Oct. 1992.

[11] R. Annavajjala and P. Orlik, “Achieving near-exponential diversity on uncoded low-dimensional MIMO, multi-user and multi-carrier systems without transmitter CSI,” *IEEE Information Theory and Applications Workshop (ITA)*, pp. 1–11, 2011.

[12] B. S. Mohammed, A. Chockalingam, and B. S. Rajan, “Low-complexity detection and performance in multi-gigabit high spectral efficiency wireless systems,” *IEEE Personal Indoor and Mobile Radio Communications (PIMRC)*, pp. 1–5, 2008.

[13] D. Pham, K. Pattipati, P. Willett, and J. Luo, “A generalized probabilistic data association detector for multiple antenna systems,” *IEEE Commun. Lett.*, vol. 8, no. 4, pp. 205–207, 2004.

[14] K. J. Kim, Y. Yue, R. A. Iltis, and J. D. Gibson, “A QRD-M/Kalman filter-based detection and channel estimation algorithm for MIMO-OFDM systems,” *IEEE Trans. Wireless Commun.*, vol. 4, pp. 710–721, Mar. 2005.

[15] B. Rajan, S. Mohammed, A. Chockalingam, and N. Srinidhi, “Low-complexity near-ML decoding of large non-orthogonal STBCs using reactive TABU search,” *IEEE Int. Symp. Information Theory (ISIT)*, pp. 1993–1997, 2009.

[16] P. Som, T. Datta, N. Srinidhi, A. Chockalingam, and B. Rajan, “Low-complexity detection in large-dimension MIMO-ISI channels using graphical models,” *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 8, pp. 1497–1511, 2011.

[17] A. Yellepeddi, K. J. Kim, C. Duan, and P. Orlik, “On probabilistic data association for achieving near-exponential diversity over fading channels,” *IEEE Int. Conf. Commun. (ICC)*, pp. 1–6, June 2013.

[18] M. Pajovic, K. J. Kim, T. Koike-Akino, and P. Orlik, “Modified probabilistic data association algorithms,” *IEEE Int. Conf. Commun. (ICC)*, June 2014.

[19] S. Minasyan, “Parametric transforms based adaptive methods in signal processing and logic design,” *Tampere Univ. Technol.*, pp. 1–74, June 2010.

[20] J. H. Conway and N. J. A. Sloane, *Shere Packings, Lattices and Groups*, Third Ed., *Springer*, Dec. 1998.