Block Differential Encoding for Rapidly Fading Channels

Xiaoli Ma, *Member, IEEE*, Georgios B. Giannakis, *Fellow, IEEE*, and Bing Lu

***Abstract*—Rapidly fading channels provide Doppler-induced diversity, but are also challenging to estimate. To bypass channel estimation, we derive two novel block differential (BD) codecs. Relying on a basis expansion model for time-varying channels, our differential designs are easy to implement, and can achieve the maximum possible Doppler diversity. The first design (BD-I) relies on a time-frequency duality, based on which we convert a time-varying channel into multiple frequency-selective channels, and subsequently into multiple flat-fading channels using orthogonal frequency-division multiplexing. Combined with a group partitioning scheme, BD-I offers flexibility to trade off decoding complexity with performance. Our second block differential design (BD-II) improves the bandwidth efficiency of BD-I at the price of increased complexity at the receiver, which relies on decision-feedback decoding. Simulation results corroborate our theoretical analysis, and compare with competing alternatives.**

***Index Terms*—Decision feedback, differential encoding, diversity, fading, interleaving.**

I. Introduction

MODELING temporal channel variations and coping with time-selective fading are important and challenging tasks in mobile communications. Time selectivity arises due to oscillator drifts, phase noise, multipath propagation, and relative motion between the transmitter and the receiver. Especially the latter is difficult to mitigate, but also useful, because it provides Doppler diversity [15]. But in order to collect the Doppler diversity quantified in [15], the time-selective channel must be estimated at the receiver. Periodic insertion of known "pilot" symbols offers an efficient means of estimating time-selective channels with bandlimited variations [2]. Uncoded pilot-symbol-assisted modulation (PSAM) [2], however, does not enable the available Doppler diversity. As an alternative to PSAM, in this paper we develop block differential (BD) schemes that obviate channel estimation, while enabling full diversity gains.

Scalar (as opposed to block) differential phase-shift keying (DPSK) has well-documented merits; see, e.g., [3], [8], [18], [19], and [23] for recent results on improving the decoder's performance. Differential schemes have also been designed to collect space diversity [10], [11], as well as multiplexing gains for multiantenna transmissions over quasi-static flat-fading channels [7], and multipath diversity over frequency-selective channels using orthogonal frequency-division multiplexing (OFDM) [1], [13]. Our work focuses on single-antenna time-selective fading channels. Especially for rapidly fading channels, differential designs are challenging because the channels may change from symbol to symbol, and the differential encoders must be designed properly; otherwise, the power per transmitted block may diverge or even diminish with time. Existing approaches designed to deal with time-selective channels employ either multiple-symbol detection (MSD) [8], [20], or decision-feedback differential detection (DF-DD) [3], [19], [20], [21], [23]. Except for the designs in [4] and [21], these alternatives are not designed to exploit Doppler diversity gains, and may entail high decoding complexity. Repetition codes are used in [4] to collect time diversity by sacrificing bandwidth efficiency. The differential modulation diversity (DMD) in [21] offers diversity without sacrificing bandwidth. However, quantifying the achievable Doppler diversity and designing DMD schemes capable of enabling tunable diversity gains were issues left open in [21].

Based on an existing basis expansion model (BEM) [22], [6], two block-differential (BD) designs are derived here: the first (BD-I) converts a time-selective channel into multiple equivalent frequency-selective channels, and then employs BD modulation for these frequency-selective channels; while the second (BD-II) directly applies BD encoding along with DF-DD, or Viterbi decoding. Note that different from related coherent block transceivers (such as those in [15] and [16] and references therein), our BD approaches here do not require channel information. Our two designs offer appealing features: i) they enable full Doppler gains; ii) they can afford reduced decoding complexity; iii) they obviate channel estimation, while maintaining bandwidth efficiency identical (BD-II), or comparable (BD-I) to, PSAM; and iv) they remain operational even when fading channels vary from symbol to symbol. Our two designs have complementary strengths. BD-I is flexible to trade off performance with complexity, while BD-II can trade off rate with complexity, and exhibits reduced error propagation.

Paper approved by R. Raheli, the Editor for Detection, Equalization, and Coding of the IEEE Communications Society. Manuscript received October 18, 2002; revised June 1, 2003 and September 18, 2003. This work was supported through collaborative participation in the Communications and Networks Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon. This paper was presented in part at the International Conference on Communications, Anchorage, AK, May 11–15, 2003, and in part at the International Conference on Acoustics, Speech and Signal Processing, Hong Kong, April 6–10, 2003.

X. Ma is with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL 36849 USA (e-mail: xiaoli@eng.auburn.edu).

G. B. Giannakis and B. Lu are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: georgios@ece.umn.edu; binglu@ece.umn.edu).

Digital Object Identifier 10.1109/TCOMM.2004.823604

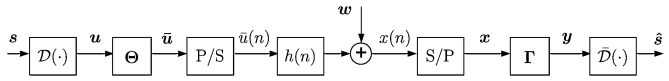


Fig. 1. Discrete-time baseband-equivalent block model of transmitter, channel, and receiver.

The rest of the paper is organized as follows. Section II introduces the channel, and the overall system model. Section III provides a unifying description of our two BD designs. Performance based on diversity-order analysis is presented in Section IV. Simulations in Section V verify our performance analysis. At last, Section VI concludes the paper.

Notation: Upper (lower) boldface letters will be used for matrices (column vectors). Superscript $^{\mathcal{H}}$ will denote Hermitian, * conjugate, and ^{T} transpose. We will reserve $\otimes$ for the Kronecker product, and $\mathrm{E}[\cdot]$ for expectation. We will use $[\boldsymbol{A}]_{k,m}$ to denote the (k,m)th entry of a matrix $\boldsymbol{A}$, and $[\boldsymbol{x}]_m$ to denote the mth entry of the column vector $\boldsymbol{x}$; $\boldsymbol{I}_N$ will denote the $N \times N$ identity matrix, $\mathbf{1}_N$ an $N \times 1$ column vector of all-one entries, and $\boldsymbol{F}_N$ the $N \times N$ normalized (unitary) fast Fourier transformation (FFT) matrix; $\mathrm{diag}[\boldsymbol{x}]$ will stand for a diagonal matrix with $\boldsymbol{x}$ on its main diagonal.

II. System Model

Although extensions to multiantenna systems are possible (see, e.g., independently derived results in [21]), in this paper we focus on single-antenna transmissions over time-varying channels. The information-bearing symbols $s(n)$ are drawn from a finite alphabet $\mathcal{A}_s$ with cardinality 2^R, where R bits/symbol denotes the transmission rate. They are parsed into blocks of size $N \times 1$: $\boldsymbol{s}(k) := [s(kN), \ldots, s((k+1)N-1)]^T$. Each information block $\boldsymbol{s}(k)$ is encoded by a differential encoder $\mathcal{D}(\cdot)$, whose output is (see also Fig. 1): $\boldsymbol{u}(k) := \mathcal{D}(\boldsymbol{s}(k))$. Each block $\boldsymbol{u}(k)$ is linearly precoded by a $P \times N$ matrix[1] $\boldsymbol{\Theta}$ with $P \geq N$, to yield $\bar{\boldsymbol{u}}(k) := \boldsymbol{\Theta}\boldsymbol{u}(k)$. Matrix $\boldsymbol{\Theta}$ comprises our inner encoder, while $\mathcal{D}(\cdot)$ is our outer encoder. After parallel-to-serial (P/S) conversion, pulse shaping, and carrier modulation, the block $\bar{\boldsymbol{u}}(k)$ is transmitted through a time-varying channel, whose delay spread is smaller than the symbol period T_s; hence, no frequency selectivity appears.

The nth sample at the receive-filter output (sampled at the symbol rate $1/T_s$) is

$$x(n) = \sqrt{\rho} h(n) \bar{u}(n) + w(n) \tag{1}$$

where ρ is the signal power per symbol, $h(n)$ is the aggregate time-selective impulse response that includes transmit-receive filters at the nth time slot (notice the channel dependence on n), and $w(n)$ is additive white Gaussian noise (AWGN) with mean zero and variance $N_0/2$. The input–output relationship (1) applies to arbitrarily rapid variations, provided that the channel remains invariant over one sampling period T_s. Note that the discrete-time model in (1) is always an approximation of the underlying continuous-time channel. For the Doppler spectrum of a mobile radio channel, the Jakes' model has been widely used [12]. As the number of sinusoidal terms in the Jakes' model can be prohibitively large, we are motivated to consider the parsimonious BEM of [6] and [22]. In its discrete-time baseband-equivalent form, the BEM describes $h(n)$ as [15]

$$h(n) := \sum_{q=0}^{Q} h_q(k) e^{j\omega_q n} \tag{2}$$

where

$$\omega_q := 2\pi(q - Q/2)/P \quad \text{and} \quad Q := 2\lceil f_{\max} P T_s \rceil \tag{3}$$

with the parameter $f_{\max}$ denoting the maximum Doppler shift, k is the block index which represents the kth realization of h_q's, and $\lceil \cdot \rceil$ stands for the ceiling integer. Because it can be measured experimentally in practice, we assume that $f_{\max}$, and thus Q, are known and bounded (P and T_s are up to the designer's disposal). Time variation in the BEM (2) is captured by the discrete-time complex exponentials, while the coefficients $\{h_q(k)\}_{q=0}^{Q}$ remain invariant over each block containing P symbols. A fresh set of BEM coefficients is considered every PT_s seconds.

At the receiver, the samples $x(n)$ are S/P converted to form the $P \times 1$ blocks with (so-termed polyphase) components $[\boldsymbol{x}(k)]_p := x(kP+p), p \in [0, P-1]$. The matrix-vector counterpart of (1) can then be expressed as

$$\boldsymbol{x}(k) = \sqrt{\rho} \boldsymbol{D}_h(k) \bar{\boldsymbol{u}}(k) + \boldsymbol{w}(k) \tag{4}$$

where $\boldsymbol{D}_h(k) := \mathrm{diag}[h(kP), \ldots, h(kP+P-1)]$, and $\boldsymbol{w}(k)$ is defined similar to $\boldsymbol{x}(k)$. Defining

$$\boldsymbol{\Omega} := \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{j\omega_0} & e^{j\omega_1} & \cdots & e^{j\omega_Q} \\ \vdots & \vdots & & \vdots \\ e^{j\omega_0(P-1)} & e^{j\omega_1(P-1)} & \cdots & e^{j\omega_Q(P-1)} \end{bmatrix}$$
$$\boldsymbol{h}(k) := [h_0(k), h_1(k), \ldots, h_Q(k)]^T \tag{5}$$

we can express $\boldsymbol{D}_h(k)$ alternatively as

$$\boldsymbol{D}_h(k) = \mathrm{diag}[\boldsymbol{\Omega}\boldsymbol{h}(k)]. \tag{6}$$

Notice that $\boldsymbol{h}(k)$ contains the BEM coefficients that remain invariant per block. Each block $\boldsymbol{x}(k)$ is linearly processed by the $N \times P$ matrix $\boldsymbol{\Gamma}$, to obtain $\boldsymbol{y}(k) := \boldsymbol{\Gamma}\boldsymbol{x}(k)$. Corresponding to $\boldsymbol{\Theta}$, matrix $\boldsymbol{\Gamma}$ comprises our inner decoder. The block $\boldsymbol{y}(k)$ is finally decoded by a differential outer decoder $\bar{\mathcal{D}}(\cdot)$ to obtain an estimate of $\boldsymbol{s}(k)$: $\hat{\boldsymbol{s}}(k) := \bar{\mathcal{D}}(\boldsymbol{y}(k))$. Since, in the following, we will work on a block-by-block basis, we will drop the block index k.

III. Block Differential Designs

Since we allow $h(n)$ in (2) to change from symbol to symbol, differential designs for such rapidly time-varying channels are more challenging than existing designs for quasi-static flat [10], [11] and frequency-selective channels [1], [13]. In this section, we will show how to jointly design the differential outer codec $\mathcal{D}(\cdot)$ and $\bar{\mathcal{D}}(\cdot)$, and the inner codec $(\boldsymbol{\Theta}, \boldsymbol{\Gamma})$ for time-selective

[1]Linear precoding has been used in various contexts, and channels differ from the time-selective ones considered here. For a tutorial treatment, we refer the reader to [16] and references therein.

channels to bypass channel estimation, and achieve the maximum possible Doppler diversity.

A. BD Design I (BD-I)

Starting from the diagonal matrix $\boldsymbol{D}_h$ in (4) and based on (6), we can express $\boldsymbol{D}_h$ as ([5, p. 202])

$$\boldsymbol{D}_h = \boldsymbol{F}_P \tilde{\boldsymbol{H}} \boldsymbol{F}_P^{\mathcal{H}} \tag{7}$$

where $\tilde{\boldsymbol{H}}$ is a circulant $P \times P$ matrix with first column $[h_{Q/2} \ \cdots \ h_0 \ 0 \ \cdots \ 0 \ h_Q \ \cdots \ h_{(Q/2)+1}]^T$, and $\boldsymbol{F}_P$ denotes the P-point FFT matrix. From (7), we deduce that (4) could have resulted from an OFDM-based transmission over a frequency-selective channel with $Q+1$ taps, whose channel matrix $\tilde{\boldsymbol{H}}$ is circulant, due to the insertion and removal of the cyclic prefix (CP). We will find this time-frequency duality useful in our ensuing design.

1) Inner Codec: In BD-I, the inner codec will be designed to split the large channel matrix $\tilde{\boldsymbol{H}}$, or $\boldsymbol{D}_h$, into N_b identical submatrices of size $N_{\rm sub}$ each, with guards of size Q between them. To this effect, we select the information block size N to be a multiple of N_b; i.e., $N = N_b N_{\rm sub}$. The inner encoder $\boldsymbol{\Theta}_1$ is given as

$$\boldsymbol{\Theta}_1 = \boldsymbol{F}_P \left(\boldsymbol{I}_{N_b} \otimes \boldsymbol{T}_{\rm cp}\right) \left(\boldsymbol{I}_{N_b} \otimes \boldsymbol{F}_{N_{\rm sub}}^{\mathcal{H}}\right) \tag{8}$$

where $\boldsymbol{T}_{\rm cp} := \sqrt{\rho'/\rho}\,[\boldsymbol{T}_1, \boldsymbol{I}_{N_{\rm sub}}, \boldsymbol{T}_2]^T$ is nothing but the matrix implementation of inserting a CP of length Q per subblock of size $N_{\rm sub}$ using

$$\boldsymbol{T}_1 := \left[\boldsymbol{0}_{(Q/2)\times(N_{\rm sub}-Q/2)}, \boldsymbol{I}_{Q/2}\right]^T$$
$$\boldsymbol{T}_2 := \left[\boldsymbol{I}_{Q/2}, \boldsymbol{0}_{(Q/2)\times(N_{\rm sub}-Q/2)}\right]^T$$

with $\rho' := \rho N_{\rm sub}/(N_{\rm sub}+Q)$ denoting the power per precoded symbol in each transmitted block. From the size of $\boldsymbol{\Theta}_1$ in (8), we can verify that $P = N_b(N_{\rm sub}+Q)$.

At the receiver, we design the inner decoder as

$$\boldsymbol{\Gamma}_1 = \left(\boldsymbol{I}_{N_b} \otimes \boldsymbol{F}_{N_{\rm sub}}\right) \left(\boldsymbol{I}_{N_b} \otimes \boldsymbol{R}_{\rm cp}\right) \boldsymbol{F}_P^{\mathcal{H}} \tag{9}$$

where $\boldsymbol{R}_{\rm cp} := [\boldsymbol{0}_{N_{\rm sub}\times(Q/2)}, \boldsymbol{I}_{N_{\rm sub}}, \boldsymbol{0}_{N_{\rm sub}\times(Q/2)}]$ is the matrix implementation of the CP removal operation per subblock. Note that if $\boldsymbol{w}$ is AWGN, $\boldsymbol{\Gamma}_1\boldsymbol{w}$ is still white, since $\boldsymbol{\Gamma}_1$ is unitary.

Using (7)–(9), it follows by direct substitution that the equivalent aggregate channel becomes

$$\boldsymbol{\Gamma}_1 \boldsymbol{D}_h \boldsymbol{\Theta}_1 = \sqrt{\frac{\rho'}{\rho}} \left(\boldsymbol{I}_{N_b} \otimes \tilde{\boldsymbol{D}}_h\right) \tag{10}$$

where $\tilde{\boldsymbol{D}}_h := \mathrm{diag}[H(0), H(1), \ldots, H(N_{\rm sub}-1)]$, and $H(n) := \sum_{q=0}^{Q} h_q e^{-j2\pi(q-Q/2)n/N_{\rm sub}}$.

The role of the inner codec is to convert the time-selective channel into an equivalent frequency-selective channel (via $\boldsymbol{F}_P$ and $\boldsymbol{F}_P^{\mathcal{H}}$); partition each block containing N information symbols into N_b subblocks; and to apply OFDM per subblock. Note that over the N_b subblocks, the equivalent channel matrix is the same, namely, $\tilde{\boldsymbol{D}}_h$, since the BEM coefficients remain invariant per block of N information symbols. This enables application of BD encoding across the N_b subblocks of the same block, in a fashion that can be thought of as the time-frequency dual of that applied in [13] for frequency-selective channels. What makes this possible is the BEM, which entails quasi-static coefficients per block, while allowing for rapid variations within the block through the bases. In the following, we will introduce $\mathcal{D}(\cdot)$ and $\bar{\mathcal{D}}(\cdot)$.

2) Outer Differential Codec: Substituting $\bar{\boldsymbol{u}} := \boldsymbol{\Theta}_1\boldsymbol{u}$ and inserting (10) into (4), we can rewrite the input–output model as

$$\boldsymbol{y} = \sqrt{\rho'}(\boldsymbol{I}_{N_b} \otimes \tilde{\boldsymbol{D}}_h)\boldsymbol{u} + \boldsymbol{\nu} \tag{11}$$

where $\boldsymbol{\nu} := \boldsymbol{\Gamma}_1\boldsymbol{w}$. In accordance with the inner codec, let us split the vectors $\boldsymbol{s}, \boldsymbol{u}$, and $\boldsymbol{y}$ into N_b equally long subblocks; i.e., $\boldsymbol{s} = [\boldsymbol{s}_0^T, \ldots, \boldsymbol{s}_{N_b-1}^T]^T$, and likewise for $\boldsymbol{u}, \boldsymbol{y}$, and $\boldsymbol{\nu}$. Now, we can write (11) in a per-subblock form as

$$\boldsymbol{y}_b = \sqrt{\rho'}\tilde{\boldsymbol{D}}_h\boldsymbol{u}_b + \boldsymbol{\nu}_b \quad \forall b \in [0, N_b-1]. \tag{12}$$

Before entering the inner encoder, we differentially encode each subblock as follows:

$$\boldsymbol{u}_b = \begin{cases} \boldsymbol{V}_b\boldsymbol{u}_{b-1}, & \text{if } 1 \le b \le N_b-1 \\ \boldsymbol{1}_{N_{\rm sub}} & \text{if } b = 0 \end{cases} \tag{13}$$

where the $N_{\rm sub} \times N_{\rm sub}$ diagonal matrix $\boldsymbol{V}_b \in \mathcal{V}$ conveys the information, and is chosen to correspond one-to-one with the bth subblock $\boldsymbol{s}_b$. Suppose each entry of $\boldsymbol{s}_b$ is chosen from a finite alphabet with cardinality 2^R, i.e., each symbol contains R information bits; therefore, we need to design $\mathcal{V}$ with cardinality $|\mathcal{V}| = 2^{RN_{\rm sub}}$. A simple design comprises a commutative group $\mathcal{V}$ of diagonal matrices with $2^{RN_{\rm sub}}$ elements, so as to make it cyclic [10], as follows:

$$\boldsymbol{V}_b = \boldsymbol{\Phi}^{l_b}, \quad l_b \in [0, 2^{RN_{\rm sub}}-1] \tag{14}$$

where the generator matrix $\boldsymbol{\Phi} := \mathrm{diag}[e^{j2\pi\phi_1 2^{-RN_{\rm sub}}}, \ldots, e^{j2\pi\phi_{N_{\rm sub}} 2^{-RN_{\rm sub}}}], \phi_k \in [0, 2^{RN_{\rm sub}}-1], k \in [1, N_{\rm sub}]$. With this design of $\mathcal{V}$, if we select $N_{\rm sub} \ge Q+1$, the maximum Doppler diversity can be guaranteed [10]. However, when $N_{\rm sub}$ and/or R is large, the drawback of the differential scheme in (13) is twofold: its design entails computationally complex search; and its decoding complexity is high. In the following, we will modify our differential design by partitioning the bases. This partitioning can be thought of as the dual of the subcarrier grouping approach of [13]. The resulting scheme will turn out to enjoy full diversity, and at the same time, it will provide flexibility to trade off decoding complexity for performance.

Suppose that we further split $\boldsymbol{u}_b$ into N_g groups, each with length K; i.e., $N_{\rm sub} = N_g K$. The gth group is given as $[\boldsymbol{u}_b(g)]_k := u_b(g+N_g k), k \in [0, K-1]$. Our idea is to apply the differential encoder (13) per group across all the subblocks, as follows:

$$\boldsymbol{u}_b(g) = \begin{cases} \boldsymbol{V}_{b_g}\boldsymbol{u}_{b-1}(g), & \text{if } 1 \le b \le N_b-1 \\ \boldsymbol{1}_K, & \text{if } b = 0 \end{cases} \tag{15}$$

where $\forall g \in [0, N_g-1]$, the diagonal matrix $\boldsymbol{V}_{b_g} \in \mathcal{V}$ conveys the information from $\boldsymbol{s}_b(g)$. Now, in order to support rate R, we

only need to design a set $\mathcal{V}$ with cardinality $2^{RK} < 2^{RN_{\rm sub}}$. The design of $\mathcal{V}$ in (14) shows that any $\boldsymbol{V}_{b_g}$ is a unitary matrix. Note that when $Q = 0$, i.e., when the channel is time invariant during one block, vectors and matrices in (15) become scalars, and our differential design reduces to the scalar DPSK [18, p. 272].

To derive our inner BD decoder, we use (12) to write the input–output relationship for the gth group of the bth subblock as

$$\boldsymbol{y}_b(g) = \sqrt{\rho'}\tilde{\boldsymbol{D}}_h(g)\boldsymbol{u}_b(g) + \boldsymbol{\nu}_b(g) \tag{16}$$

where $\boldsymbol{y}_b(g)$ and $\boldsymbol{\nu}_b(g)$ are defined as $\boldsymbol{u}_b(g)$, and $\tilde{\boldsymbol{D}}_h(g)$ is the channel submatrix from $\tilde{\boldsymbol{D}}_h$ corresponding to the group g. By interchanging $\tilde{\boldsymbol{D}}_h(g)$ with $\boldsymbol{u}_b(g)$, we have $\tilde{\boldsymbol{D}}_h(g)\boldsymbol{u}_b(g) = \boldsymbol{D}_{u_b}(g)\tilde{\boldsymbol{h}}(g)$, where $\boldsymbol{D}_{u_b}(g) := \mathrm{diag}[\boldsymbol{u}_b(g)]$, and $\tilde{\boldsymbol{h}}(g)$ contains the diagonal elements of $\tilde{\boldsymbol{D}}_h(g)$. Since $\boldsymbol{V}_{b_g}$ is unitary, $\boldsymbol{D}_{u_b}(g)$ is also unitary. Relating $\boldsymbol{y}_b(g)$ with $\boldsymbol{y}_{b-1}(g)$, we find [cf. (16)]

$$\boldsymbol{y}_b(g) = \boldsymbol{D}_{u_b}(g)\boldsymbol{D}^{\mathcal{H}}_{u_{b-1}}(g)[\boldsymbol{y}_{b-1}(g) - \boldsymbol{\nu}_{b-1}(g)] + \boldsymbol{\nu}_b(g). \tag{17}$$

Plugging $\boldsymbol{D}_{u_b}(g) = \boldsymbol{V}_{b_g}\boldsymbol{D}_{u_{b-1}}(g)$ into (17), we obtain

$$\boldsymbol{y}_b(g) = \boldsymbol{V}_{b_g}\boldsymbol{y}_{b-1}(g) + \boldsymbol{\nu}'_b(g) \tag{18}$$

where $\boldsymbol{\nu}'_b(g) := \boldsymbol{\nu}_b(g) - \boldsymbol{V}^{\mathcal{H}}_{b_g}(g)\boldsymbol{\nu}_{b-1}(g)$ is AWGN, because $\boldsymbol{D}_{u_{b-1}}(g)$ is a unitary matrix.

Based on (18), the maximum-likelihood (ML) detector for $\boldsymbol{V}_{b_g}$ is

$$\hat{\boldsymbol{V}}_{b_g} = \arg\min_{\boldsymbol{V}\in\mathcal{V}} \|\boldsymbol{y}_b(g) - \boldsymbol{V}\boldsymbol{y}_{b-1}(g)\| \quad \forall g \in [0, N_g - 1] \tag{19}$$

where $\|\cdot\|$ denotes the Frobenius norm. Using the one-to-one demapping $\boldsymbol{V}_{b_g} \to \boldsymbol{s}_b(g)$, we can obtain the information conveying group $\hat{\boldsymbol{s}}_b(g)$ from $\hat{\boldsymbol{V}}_{b_g}$, with full Doppler diversity order that we will quantify in Section IV.

3) Bandwidth Efficiency: Before we summarize our BD-I design, we will first analyze its bandwidth efficiency, and derive $N_{\rm sub}$ and N_b values that optimize it. Since one subblock is used to initialize the differential recursion, the bandwidth efficiency (defined as the ratio of the number of information-bearing symbols over the block size) is given by

$$\eta_1 := \frac{N_{\rm sub}(N_b - 1)}{P}. \tag{20}$$

Fixing the block size P, η_1 depends on the number of subblocks N_b and their size $N_{\rm sub}$. Since $P = N_b(N_{\rm sub} + Q)$, we deduce that $N_{\rm sub} = P/N_b - Q$. Then (20) can be rewritten as

$$\eta_1 := \frac{\left(\frac{P}{N_b} - Q\right)(N_b - 1)}{P} = 1 - \frac{1}{N_b} - \frac{N_b Q}{P} + \frac{Q}{P}. \tag{21}$$

Temporarily, let us consider N_b as a positive real number, and optimize η_1 by differentiating it with respect to N_b. Based on (21), we obtain

$$\frac{d\eta_1}{dN_b} = \frac{1}{N_b^2} - \frac{Q}{P} = 0 \Rightarrow N_b^{\rm opt} = \sqrt{\frac{P}{Q}} \tag{22}$$

$$N_{\rm sub}^{\rm opt} = \sqrt{PQ} - Q. \tag{23}$$

Since $d^2\eta_1/d^2N_b < 0$, we infer from (20), (22), and (23) that the maximum bandwidth efficiency of our BD-I scheme is

$$\eta_1^{\rm opt} = \left(1 - \sqrt{\frac{Q}{P}}\right)^2. \tag{24}$$

Recalling from (3) that $Q/P \gtrapprox 2f_{\max}T_s$, we infer from (24) that the spectral efficiency of BD-I drops as the Doppler spread (channel variability) increases.

We summarize our codec BD-I in the following steps.

At the transmitter:

T1−1): Given $f_{\max}$ and T_s, select Q, and then decide P based on (3); as we will explain in Section IV, complexity versus performance tradeoffs suggest fixing Q to a value in $[0, 4]$.

T1−2): Choose N_b and $N_{\rm sub}$ from (22) and (23), respectively, so that $P = N_b(N_{\rm sub} + Q)$; then factor $N_{\rm sub}$ such that $N_{\rm sub} = N_g K$.

T1−3): Generate N_b information groups $\boldsymbol{s}_b$, each of length $N_g K$, and use the first subblock to initialize the differential recursion (15), $\forall g \in [0, N_g - 1]$; i.e., $\boldsymbol{s}_0$ does not convey any information.

T1−4): Design a one-to-one mapping from $\{\boldsymbol{s}_b(g)\}$ to $\mathcal{V}$, and from $\boldsymbol{s}_0(g)$ to $\boldsymbol{V}_{0_g} = \boldsymbol{I}_K$.

T1−5): Perform the differential encoding as in (15).

T1−6): Implement the OFDM-based inner coder given in (8).

At the receiver:

R1−1): Perform the OFDM-based inner decoder as in (9).

R1−2): Decode the matrix $\boldsymbol{V}_{b_g}$ by using the ML decoder in (19).

R1−3): Demap $\hat{\boldsymbol{V}}_{b_g}$ to obtain $\hat{\boldsymbol{s}}_b(g)$.

B. BD Design II (BD-II)

On top of the initialization that is inherent to all differential designs, the CP insertion per subblock causes extra bandwidth loss in BD-I. Our desire to improve the bandwidth efficiency of BD-I motivates the BD scheme of this subsection, which avoids the N_b CP guards, and thus operates on blocks of size $P = N$.

1) Inner Codec: With $N = P = M(Q + 1)$, we design the $N \times N$ inner encoder in BD-II as

$$\boldsymbol{\Theta}_2 = \begin{bmatrix} \boldsymbol{I}_{Q+1} \otimes \boldsymbol{e}_1^T \\ \boldsymbol{I}_{Q+1} \otimes \boldsymbol{e}_2^T \\ \vdots \\ \boldsymbol{I}_{Q+1} \otimes \boldsymbol{e}_M^T \end{bmatrix} \tag{25}$$

where $\boldsymbol{e}_m$ is the mth column of $\boldsymbol{I}_M$. Encoding the $N \times 1$ vector $\boldsymbol{u}$ with $\boldsymbol{\Theta}_2$ yields $\bar{\boldsymbol{u}} := \boldsymbol{\Theta}_2\boldsymbol{u} = [\boldsymbol{u}_0^T, \boldsymbol{u}_1^T, \ldots, \boldsymbol{u}_{M-1}^T]^T$ with $\boldsymbol{u}_m := [u(m), u(m+M), \ldots, u(m+QM)]^T, m \in [0, M-1]$. The structure of $\bar{\boldsymbol{u}}$ shows that the inner encoder $\boldsymbol{\Theta}_2$ is a block interleaver with depth M. The interleaver implemented via $\boldsymbol{\Theta}_2$ is at the symbol level, which should not be confused with the bit interleaver advocated in [14] to enable turbo differential decoding. Note that $\boldsymbol{\Theta}_2$ is similar to the block interleaver used in [21]. Different from [21], our BEM-driven design of transmission parameters illuminates rather nicely the relationship among channel variation $(Q+1)$, interleaver depth (M), and the length

of transmitted blocks (P). This relationship also guides the selection of M to satisfy $P = M(Q+1)$.

Accordingly, at the receiver, the inner decoder $\mathbf{\Gamma}_2$ is chosen as $\mathbf{\Gamma}_2 = \mathbf{\Theta}_2$, which deinterleaves the received vector $\boldsymbol{x}$ (see Fig. 1). Interleaving and deinterleaving, based on (25), result in dividing the $P \times (Q+1)$ Vandermonde matrix $\mathbf{\Omega}$ in (5) into M nonoverlapping (polyphase) submatrices $\{\mathbf{\Omega}_m\}_{m=0}^{M-1}$. Each submatrix $\mathbf{\Omega}_m$ is also a $(Q+1)\times(Q+1)$ matrix with $(k+1, q+1)$st entry $[\mathbf{\Omega}_m]_{k+1,q+1} = \exp[j\omega_q(m+kM)]$. Since $\mathbf{\Omega}_m^{\mathcal{H}}\mathbf{\Omega}_m = (Q+1)\boldsymbol{I}_{Q+1}$, it follows that $(1/\sqrt{Q+1})\mathbf{\Omega}_m$ is unitary. The equivalent channel matrix becomes [cf. (6)]

$$\mathbf{\Gamma}_2\boldsymbol{D}_h\mathbf{\Theta}_2 = \operatorname{diag}\left[\boldsymbol{h}_0^T, \boldsymbol{h}_1^T, \ldots, \boldsymbol{h}_{M-1}^T\right] \tag{26}$$

where $\boldsymbol{h}_m := [h(m), h(m+M), \ldots, h(m+QM)]^T = \mathbf{\Omega}_m\boldsymbol{h}\ \forall m \in [0, M-1]$, and $\boldsymbol{h}$ is defined as in (5).

From (4), we can write the input–output relationship as

$$\boldsymbol{y} = \mathbf{\Gamma}_2\boldsymbol{D}_h\mathbf{\Theta}_2\boldsymbol{u} + \boldsymbol{\nu} \tag{27}$$

where $\boldsymbol{\nu} := \mathbf{\Gamma}_2\boldsymbol{w}$. Since $\mathbf{\Theta}_2$ is a permutation matrix, $\boldsymbol{\nu}$ is still AWGN. Based on (26), (27) can be expressed per subblock of size $Q+1$ as

$$\boldsymbol{y}_m = \sqrt{\rho}\boldsymbol{D}_{h_m}\boldsymbol{u}_m + \boldsymbol{\nu}_m, \quad m \in [0, M-1] \tag{28}$$

where $\boldsymbol{y}_m, \boldsymbol{u}_m, \boldsymbol{\nu}_m$ are defined in a polyphase form similar to $\boldsymbol{h}_m$. Note that compared with ρ' in (16), ρ in (28) does not experience signal power loss. Relative to the inner codec $(\mathbf{\Theta}_1, \mathbf{\Gamma}_1)$ of BD-I, we observe that $(\mathbf{\Theta}_2, \mathbf{\Gamma}_2)$ do not include CP, they do not entail IFFT and FFT operations, and they do not exploit the BEM structure; as such, similar to [21], (28) holds true for any time-selective channel model, including Jakes' model.

2) Outer Differential Codec: The outer differential encoder in BD-II starts by splitting the $P\times 1$ vector $\boldsymbol{s}$ to obtain polyphase subblocks $\{\boldsymbol{s}_m\}_{m=0}^{M-1}$ with $[\boldsymbol{s}_m]_{q+1} := s(m(Q+1)+q), q \in [0, Q]$. Mapping $\boldsymbol{s}_m$ to $\boldsymbol{V}_m$, and generating $\boldsymbol{u}_m$ by using the unitary differential modulation of [10] and [11] between consecutive subblocks, we obtain

$$\boldsymbol{u}_m = \begin{cases} \boldsymbol{V}_m\boldsymbol{u}_{m-1}, & \text{if } 1 \le m \le M-1 \\ \boldsymbol{1}_{Q+1}, & \text{if } m = 0 \end{cases} \tag{29}$$

where the $(Q+1)\times(Q+1)$ diagonal matrix $\boldsymbol{V}_m \in \mathcal{V}$ is consistent (except for dimensionality) with the definition in (14).

By interchanging $\boldsymbol{D}_{h_m}$ with $\boldsymbol{u}_m$ in (28), we have $\boldsymbol{D}_{h_m}\boldsymbol{u}_m = \boldsymbol{D}_{u_m}\boldsymbol{h}_m = \boldsymbol{D}_{u_m}\mathbf{\Omega}_m\boldsymbol{h}$. Because the BEM coefficient vector $\boldsymbol{h}$ remains unchanged across subblocks of the same block, and $\boldsymbol{D}_{u_m}, (1/\sqrt{Q+1})\mathbf{\Omega}_m$ are unitary, considering two consecutive subblocks, we obtain [cf. (28)]

$$\boldsymbol{y}_m = \frac{1}{Q+1}\boldsymbol{D}_{u_m}\mathbf{\Omega}_m\mathbf{\Omega}_{m-1}^{\mathcal{H}}\boldsymbol{D}_{u_{m-1}}^{\mathcal{H}}[\boldsymbol{y}_{m-1} - \boldsymbol{\nu}_{m-1}] + \boldsymbol{\nu}_m. \tag{30}$$

Thanks to the inner encoder (25), $\mathbf{\Omega}_{m-1}$ is unitary. Therefore, instead of $\mathbf{\Omega}_{m-1}^{-1}$, we only need $\mathbf{\Omega}_{m-1}^{\mathcal{H}}$, which enables low complexity. Next, we observe that[2]

$$\mathbf{\Upsilon} := \frac{1}{Q+1}\mathbf{\Omega}_m\mathbf{\Omega}_{m-1}^{\mathcal{H}} \tag{31}$$

[2]Note that although $\mathbf{\Omega}_{m-1}^{\mathcal{H}}\mathbf{\Omega}_m = \operatorname{diag}[\boldsymbol{\omega}]$ with $\boldsymbol{\omega} := [1, e^{j2\pi/N}, \ldots, e^{j2\pi Q/N}]^T$, this is not the case for $\mathbf{\Omega}_m\mathbf{\Omega}_{m-1}^{\mathcal{H}}$.

is independent of m, which allows us to rewrite (30) as

$$\boldsymbol{y}_m = \boldsymbol{D}_{u_m}\mathbf{\Upsilon}\boldsymbol{D}_{u_{m-1}}^{\mathcal{H}}\boldsymbol{y}_{m-1} + \boldsymbol{\nu}'_m \quad \forall m \in [1, M-1] \tag{32}$$

where $\boldsymbol{\nu}'_m := \boldsymbol{\nu}_m - \boldsymbol{D}_{u_m}\mathbf{\Upsilon}\boldsymbol{D}_{u_{m-1}}^{\mathcal{H}}\boldsymbol{\nu}_{m-1}$ is AWGN because $\boldsymbol{D}_{u_m}\mathbf{\Upsilon}\boldsymbol{D}_{u_{m-1}}^{\mathcal{H}}$ is a unitary matrix. Note that because $\mathcal{V}$ is a cyclic group, $\boldsymbol{D}_{u_m} = \prod_{l=1}^{m}\boldsymbol{V}_l \in \mathcal{V}$.

Comparing (18) with (18) in BD-I, we notice that the right-hand side of (18) depends on the previous symbol matrix $\boldsymbol{D}_{u_{m-1}}$. Therefore, detecting $\boldsymbol{D}_{u_m}$ in (18) will depend not only on the received vectors $\boldsymbol{y}_m$ and $\boldsymbol{y}_{m-1}$, but also on the symbols transmitted in earlier intervals. This is because in BD-I, the equivalent channels in consecutive subblocks are identical [cf. (12)], while in BD-II, the channels vary across subblocks [cf. (28)]. To handle slow time variations in, e.g., DPSK systems, MSD and DF-DD have well-documented merits (see [3], [8], [19], [20], [21], [23], and references therein). To cope with the multisymbol dependence in (32), we will develop two detectors: one using block DF-DD, and another relying on the Viterbi algorithm (VA).

Letting $\hat{\boldsymbol{D}}_{u_{m-1}}$ denote the estimate of $\boldsymbol{D}_{u_{m-1}}$, we have [cf. (32)]

$$\boldsymbol{y}_m = \boldsymbol{D}_{u_m}\mathbf{\Upsilon}\hat{\boldsymbol{D}}_{u_{m-1}}^{\mathcal{H}}\boldsymbol{y}_{m-1} + \boldsymbol{\nu}''_m \tag{33}$$

where $\boldsymbol{\nu}''_m := \boldsymbol{\nu}'_m + \boldsymbol{D}_{u_m}\mathbf{\Upsilon}(\boldsymbol{D}_{u_{m-1}}^{\mathcal{H}} - \hat{\boldsymbol{D}}_{u_{m-1}}^{\mathcal{H}})\boldsymbol{y}_{m-1}$. Our block DF-DD is defined as

$$\hat{\boldsymbol{D}}_{u_m} = \arg\min_{\tilde{\boldsymbol{D}} \in \mathcal{V}} \left\|\boldsymbol{y}_m - \tilde{\boldsymbol{D}}\mathbf{\Upsilon}\hat{\boldsymbol{D}}_{u_{m-1}}^{\mathcal{H}}\boldsymbol{y}_{m-1}\right\|. \tag{34}$$

Note that the aggregate noise $\boldsymbol{\nu}''_m$ is no longer white, which renders DF-DD suboptimal in general. At the same time, since the decision on $\hat{\boldsymbol{D}}_{u_m}$ depends on $\hat{\boldsymbol{D}}_{u_{m-1}}$, error propagation is unavoidable. Compared with the DF-DD in [21], our detector in (34) also takes into account the dominant channel correlation between two subblocks, because the BEM captures the major variations of the underlying time-selective channel and incorporates the correlation between differentially encoded subblocks.

To mitigate error propagation, one can resort to MSD, which, in our case, detects multiple subblocks, but also increases decoding complexity exponentially with the number of subblocks. Instead, we recommend the VA, which achieves ML optimality with complexity comparable to DF-DD. From (18), we can write the ML sequence detector as

$$\left\{\hat{\boldsymbol{D}}_{u_m}\right\}_{m=1}^{M-1} = \arg\min_{\substack{\forall \tilde{\boldsymbol{D}}_{u_m} \in \mathcal{V} \\ m \in [1, M-1]}} \sum_{m=1}^{M-1}\left\|\boldsymbol{y}_m - \tilde{D}_{u_m}\mathbf{\Upsilon}\tilde{\boldsymbol{D}}_{u_{m-1}}^{\mathcal{H}}\boldsymbol{y}_{m-1}\right\|^2. \tag{35}$$

Because the decision on $\hat{\boldsymbol{D}}_{u_m}$ depends only on a single previous subblock $\boldsymbol{D}_{u_{m-1}}$, we design a VA with $|\mathcal{V}| = 2^{R(Q+1)}$ states. The path metric for the mth stage from state $\boldsymbol{D}_1$ to state $\boldsymbol{D}_2$ is $\|\boldsymbol{y}_m - \boldsymbol{D}_2\mathbf{\Upsilon}\tilde{\boldsymbol{D}}_1^{\mathcal{H}}\boldsymbol{y}_{m-1}\|^2$. Using VA searching through a trellis with $M-1$ stages, we can estimate $\{\boldsymbol{D}_{u_m}\}_{m=1}^{M-1}$. Then, we have $\hat{\boldsymbol{V}}_m = \hat{\boldsymbol{D}}_{u_{m-1}}^{\mathcal{H}}\hat{\boldsymbol{D}}_{u_m}$. The estimate of $\boldsymbol{s}_m$ can be found by demapping $\hat{\boldsymbol{V}}_m$ to $\hat{\boldsymbol{s}}_m$. Notice that the block DF-DD has

complexity of $\mathcal{O}(2^{R(Q+1)})$, while the VA has slightly higher complexity of $\mathcal{O}(2^{2R(Q+1)})$.

3) Bandwidth Efficiency: Since only $Q+1$ symbols are used to initialize (29), the bandwidth efficiency of BD-II is

$$\eta_2 := \frac{P-(Q+1)}{P} = 1 - \frac{Q+1}{P}. \tag{36}$$

For $P > Q+1$, the optimum bandwidth efficiency $\eta_1^{\rm opt}$ of BD-I is smaller than η_2, since BD-I loses bandwidth not only for initialization, but also due to the CP insertion per subblock. Note that the bandwidth efficiencies (η_1 and η_2) of our two schemes are lower than those of scalar DPSK schemes. This is because our transmitter designs take into account the channel time variation, and is the price paid to enable the Doppler diversity. As we have mentioned, an alternative approach to coping with time-varying channels is PSAM. Accounting for the pilot symbols per block, the bandwidth efficiency of PSAM is $\eta_{\rm psam} = 1 - (Q+1)/P$, which is identical to (36).

In summary, BD-II relies on the following steps.

At the transmitter:

T2−1): Given $f_{\max}$ and T_s, select Q, decide P based on (3), and find M as $M := P/(Q+1)$.

T2−2): Generate M information subblocks $\boldsymbol{s}_m$, each with length $Q+1$, and use the first subblock ($m=0$) as reference to initialize the differential recursion (29).

T2−3): Design a one-to-one mapping from $\{\boldsymbol{s}_m\}_{m=1}^{M-1}$ to $\mathcal{V}$, and from $\boldsymbol{s}_0$ to $\boldsymbol{V}_0 = \boldsymbol{I}_{Q+1}$.

T2−4): Perform the differential encoding across m as in (29) to obtain $\{\boldsymbol{u}_m\}_{m=0}^{M-1}$.

T2−5): Implement interleaving of $\{\boldsymbol{u}_m\}$ as in (25) to obtain $\{\bar{\boldsymbol{u}}_m\}$.

At the receiver:

R2−1): Deinterleave the received vectors $\{\boldsymbol{x}_m\}$ to $\{\boldsymbol{y}_m\}$.

R2−2): Decode the matrix $\hat{\boldsymbol{D}}_{u_m}$ by using (34), and then $\hat{\boldsymbol{V}}_m$ by using (29).

R2−3): Demap $\hat{\boldsymbol{V}}_m$ to $\hat{\boldsymbol{s}}_m$.

IV. Performance Analysis

Considering the tradeoffs among diversity advantage, bandwidth efficiency, and decoding complexity, in this section, we want to select the parameters K in BD-I and M (or P) in BD-II, along with the differential group design $\mathcal{V}$ in [10] to provide performance guarantees, at least for high signal-to-noise ratios (SNRs). Our derivations are based on the following operating conditions.

AS1): BEM coefficients $\{h_q\}_{q=0}^{Q}$ are zero-mean, complex Gaussian random variables.

AS2): High SNR is considered only for deriving the diversity order.

If we consider each subchannel in $\tilde{\boldsymbol{h}}(g)$ [cf. (16)] for BD-I as a transmit antenna, our single-antenna setup can be viewed as a system with K transmit and one receive antennae, similar to that considered in, e.g., [10]. Employing ML decoding in (19) at the receiver, we will analyze the pairwise-error probability (PEP) $\Pr[\boldsymbol{V}_{b_g} \to \boldsymbol{V}'_{b_g}]$, that $\boldsymbol{V}_{b_g}$ is incorrectly decoded as $\boldsymbol{V}'_{b_g} \neq \boldsymbol{V}_{b_g}$ [9]–[11]. As usual, the conditional PEP for the model in (18) can be approximated using the Chernoff bound as ([18, p. 53])

$$\Pr\left[\boldsymbol{V}_{b_g} \to \boldsymbol{V}'_{b_g} \middle| \boldsymbol{y}_{b-1}(g)\right] \leq \exp\left[-\frac{\left\|\left(\boldsymbol{V}_{b_g} - \boldsymbol{V}'_{b_g}\right)\boldsymbol{y}_{b-1}(g)\right\|^2}{8N_0}\right]. \tag{37}$$

Based on AS2), we ignore the noise term in (37) and the error-propagation effect (perfect estimation for the previous subblocks), and relate $\boldsymbol{y}_{b-1}(g)$ to $\tilde{\boldsymbol{h}}(g)$ via $\boldsymbol{y}_{b-1}(g) = \sqrt{\rho'}\boldsymbol{D}_{u_{b-1}}(g)\tilde{\boldsymbol{h}}(g)$, to obtain [cf. (37)]

$$\Pr[\boldsymbol{V}_{b_g} \to \boldsymbol{V}'_{b_g} | \tilde{\boldsymbol{h}}(g)] \leq \exp\left[-\frac{\|\sqrt{\rho'}\left(\boldsymbol{V}_{b_g} - \boldsymbol{V}'_{b_g}\right)\boldsymbol{D}_{u_{b-1}}(g)\tilde{\boldsymbol{h}}(g)\|^2}{8N_0}\right]. \tag{38}$$

It is worth mentioning that the variance of $\boldsymbol{\nu}'_b(g)$ in (18) is N_0, which is twice that of $\boldsymbol{\nu}_b(g)$, consistent with the well-known 3-dB loss in SNR that differential detectors exhibit relative to coherent ones.

Although the channel is unknown at the receiver, based on AS1), (38) allows us to average the conditional PEP over the Rayleigh distributed channel parameters. From the definition of $\tilde{\boldsymbol{h}}(g)$ in (6), we find that $\tilde{\boldsymbol{h}}(g) = \boldsymbol{\Omega}_g \boldsymbol{h}$, where $\boldsymbol{\Omega}_g$ denotes the K equispaced rows from the first $Q+1$ columns of the FFT matrix $\boldsymbol{F}_{N_{\rm sub}}$. Defining $\boldsymbol{R}_h := E[\boldsymbol{h}\boldsymbol{h}^{\mathcal{H}}]$, and eigendecomposing $\boldsymbol{R}_h$, we find $\boldsymbol{R}_h = \boldsymbol{U}_h \boldsymbol{\Lambda}_h \boldsymbol{U}_h^{\mathcal{H}}$, where $\boldsymbol{\Lambda}_h := \mathrm{diag}[\lambda_0, \ldots, \lambda_{r_h-1}]$ contains the nonzero eigenvalues of $\boldsymbol{R}_h$, and r_h is the rank of $\boldsymbol{R}_h$. The average PEP $\Pr[\boldsymbol{V}_{b_g} \to \boldsymbol{V}'_{b_g}]$ depends on the rank of the matrix $(\boldsymbol{V}_{b_g} - \boldsymbol{V}'_{b_g})\boldsymbol{D}_{u_{b-1}}(g)\boldsymbol{\Omega}_g \boldsymbol{U}_h \boldsymbol{\Lambda}_h^{1/2}$. Thus, the maximum Doppler diversity is defined as (see, e.g., [10], [11], and [15] for details)

$$G_d = \min_{\forall \boldsymbol{V}_{b_g} \neq \boldsymbol{V}'_{b_g}} \mathrm{rank}\left(\left(\boldsymbol{V}_{b_g} - \boldsymbol{V}'_{b_g}\right)\boldsymbol{D}_{u_{b-1}}(g)\boldsymbol{\Omega}_g \boldsymbol{U}_h \boldsymbol{\Lambda}_h^{1/2}\right). \tag{39}$$

If the correlation matrix $\boldsymbol{R}_h$ has full rank, i.e., $r_h = Q+1$, and $K \geq Q+1$, then the maximum diversity $Q+1$ is enabled by our differential scheme with $\mathcal{V}$ designed as in [10] and [11]. When $K < Q+1$, the achieved diversity is K. Therefore, in general, we have that the diversity gain achieved by BD-I is

$$G_d = \min(K, r_h). \tag{40}$$

Note that the cardinality of $\mathcal{V}$ also depends on the group size K. Hence, the decoding complexity also depends on K. As K increases, the decoding complexity also increases. Since K is a tuning parameter, by adjusting K, the designer has flexibility to trade off affordable decoding complexity for desirable performance.

As for BD-II, we can also achieve maximum diversity order r_h. Assuming there is no error propagation ($\hat{\boldsymbol{D}}_{u_{m-1}} = \boldsymbol{D}_{u_{m-1}}$), based on (33), the conditional PEP for BD-II is

$$\Pr\left[\boldsymbol{D}_{u_m} \to \boldsymbol{D}'_{u_m} \,\middle|\, \boldsymbol{y}_{m-1}\right] \le \exp\left(-\frac{\left\|\left(\boldsymbol{D}_{u_m} - \boldsymbol{D}'_{u_m}\right)\boldsymbol{\Gamma}\hat{\boldsymbol{D}}^{\mathcal{H}}_{u_{m-1}}\boldsymbol{y}_{m-1}\right\|^2}{8N_0}\right). \quad (41)$$

Ignoring the noise at high SNR range, and plugging $\boldsymbol{y}_{m-1} = \boldsymbol{D}_{u_{m-1}}\boldsymbol{\Omega}_{m-1}\boldsymbol{h}$ into (41), we obtain that the conditional PEP is given by

$$\Pr[\boldsymbol{V}_m \to \boldsymbol{V}'_m|\boldsymbol{h}] \le \exp\left[-\frac{\|\sqrt{\rho}(\boldsymbol{V}'_m - \boldsymbol{V}_m)\boldsymbol{D}_{u_{m-1}}\boldsymbol{\Upsilon}\boldsymbol{\Omega}_{m-1}\boldsymbol{h}\|^2}{8N_0}\right]. \quad (42)$$

Unlike BD-I, there is no flexibility in BD-II to adjust the information length per subblock. If the length of each subblock is smaller than $Q+1$, then $\boldsymbol{\Omega}_m$ is not unitary; i.e., $\boldsymbol{\Omega}_m^{\mathcal{H}}\boldsymbol{\Omega}_m \neq (Q+1)\boldsymbol{I}_{Q+1}$, and the differential detectors in (34) and (35) are no longer applicable. When the subblock size is greater than $Q+1$, the decoding complexity increases, while the performance may not increase. Also note that ρ' in (38) is less than ρ in (42) because $\rho' := \rho N_{\rm sub}/(N_{\rm sub}+Q)$, which results in a reduced performance of BD-I, relative to BD-II.

Remark: For a given $f_{\max}$, as the number of symbols per block P, or the symbol period T_s increases, Q increases along with the channel-induced Doppler diversity gain. However, longer block lengths result in longer decoding delays, and require higher decoding complexity (especially for BD-II). These considerations suggest limiting Q to small values, say $Q \in [0,4]$. One could argue that such a choice limits the diversity gain $Q+1$ to ≤ 5. But this is a reasonable compromise, if one takes into account that diversity orders greater than five show up in bit-error rate (BER) curves for very high SNR values (say SNR $>$40 dB) that are well beyond the practical range.

V. Simulated Performance

To illustrate the performance of our differential schemes (BD-I and BD-II) over time-selective channels, we carry out simulated comparisons. The symbol rate is $R = 2$ b/symbol. In Test Cases 2 and 3, we assume that the $Q+1$ BEM coefficients are complex Gaussian distributed with zero mean. In Test Cases 4, 5, and 6, we rely on the Jakes' model containing 200 sinusoidal terms ([12, p. 68]) to generate our time-selective channels. The SNR is defined as E_b/N_0, where E_b is the signal power per information bit. Note that in order to guarantee the validity of (1), we select small Q values.

1) Test Case 1 (Channel-Model Validation): For the Jakes' model in [12, p. 68], we select carrier frequency $f_c = 3.5$ GHz, mobile speed $v = 250$ km/h, sampling period $T_s = 0.5$ μs, and block length $P = 1400$ to generate one realization of the time-varying channel depicted in Fig. 2. To demonstrate the validity of the BEM, we vary the number of bases and find the channel coefficients $\{h_q\}_{q=0}^{Q}$ using the least-squares method in [15]. In Fig. 2, we also depict the approximate channels using the BEM with $Q = 0, 2, 4$ bases. We observe that, as Q increases, the BEM approximates the Jakes' model well.

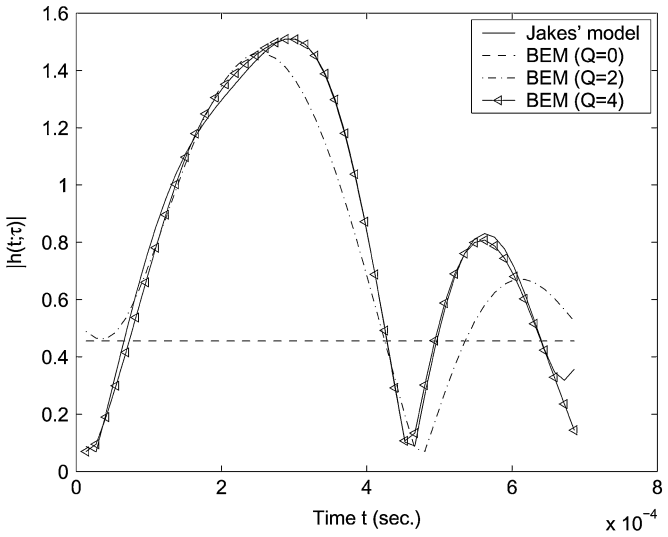


Fig. 2. BEM validation.

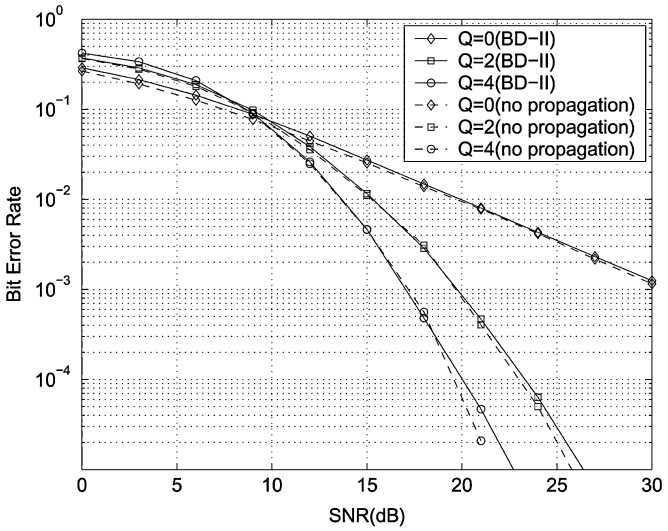


Fig. 3. Performance comparison of BD-II with and without error propagation.

2) Test Case 2 (Error-Propagation Effects): In Section III-B.2, we have claimed that the DF-DD for BD-II causes unavoidable error propagation. To delineate the effects of error propagation, we compare two cases here: one is with error propagation, i.e., the estimate of the previous symbol $\hat{\boldsymbol{D}}_{u_{m-1}}$ is used in (33); and the other is without error propagation, which means that the true value $\boldsymbol{D}_{u_{m-1}}$ is assumed in (33). We select $Q = 0, 2, 4$ corresponding to $f_{\max}T_s = 0, 0.02, 0.04$, and decide the block length from (3) as $P = 48, 48, 50$, respectively. The BER performance is plotted in Fig. 3. We notice that the error-propagation effect on the performance of our DF-DD is small (less than 0.3 dB), or even negligible. This should not be surprising because the initial subblock is known, and the decision of the current subblock only depends on a single previous subblock.

3) Test Case 3 (Performance Comparisons With Coherent Receivers): In this test case, we select $Q = 2, 4$ for a given

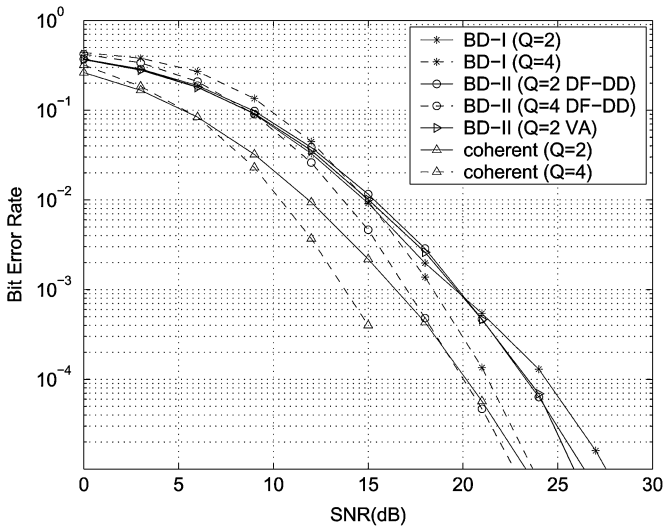


Fig. 4. Performance comparisons with coherent receivers.

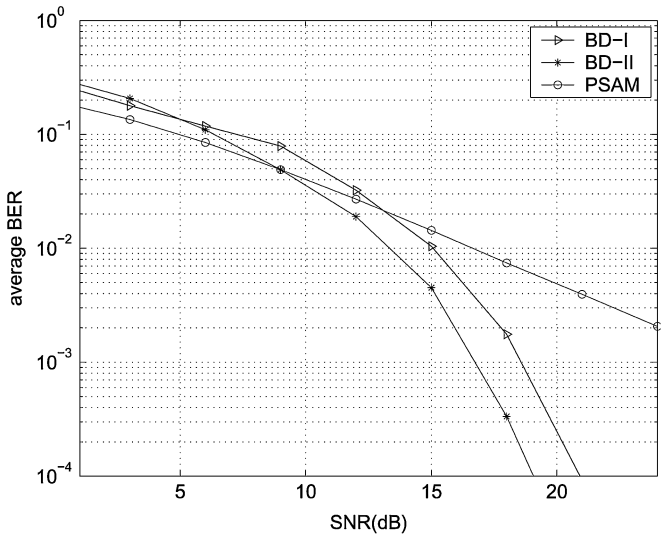


Fig. 5. Performance comparisons of BDs with PSAM (uncoded case).

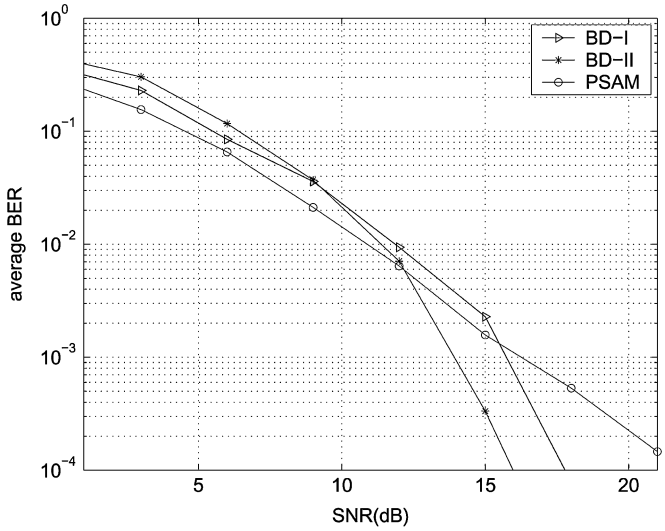


Fig. 6. Performance comparisons of BDs with PSAM (coded case).

$f_{\max}T_s = 0.02$ or 0.04. Accordingly, the block length is chosen from (3) as $P = 48$, or 50. For BD-I, we select N_b and $N_{\rm sub}$ based on (22) and (23), and the group size $K = Q+1$. We compare the performance of BD-I and BD-II in Fig. 4. From the BER slopes, we infer that both BD-I and BD-II (with either DF-DD or VA) achieve full diversity $Q+1$. However, for fixed block length P, as Q increases, the performance of BD-I becomes worse than that of BD-II, because of the power wasted on the CP guards. The bandwidth efficiencies are $\eta_1 = (1-\sqrt{Q/P})^2$ and $\eta_2 = 1-(Q+1)/P$ for BD-I and BD-II, respectively.

In Fig. 4, we also plot the benchmark performance of coherent receivers, assuming that the channel is perfectly known at the receiver. Comparing the performance of our differential schemes with those of coherent receivers, we notice the "3-dB difference" that we expected from Section IV.

4) Test Case 4 (Performance Comparisons With PSAM): Relying on the Jakes' model with $f_c = 3.5$ GHz, $v = 250$ km/h, and sampling period $T_s = 1$ μs, we compare the BER performance of our differential designs (BD-I and BD-II) against the PSAM-based design in [2] for both uncoded and convolutionally coded cases. For BD-I, we select the parameters as $(P, N_b, N_{\rm sub}) = (1176, 21, 54)$. For BD-II and PSAM, we have the block lengths as 1, 203 and 1, 200, respectively.

In Fig. 5, uncoded performance of the three schemes is plotted. We observe that since $Q = 2$, the performance of BD schemes is better than PSAM. This is because fast-varying channels introduce Doppler diversity, and our differential schemes enjoy Doppler diversity gains, although they are not designed based on Jakes' model; whereas for PSAM, the diversity order is always about one. Note that for moderate-to-high SNR values, our designs outperform PSAM, again thanks to the Doppler diversity gains.

We also compare the coded performance of these three schemes in Fig. 6. A rate-$1/2$ convolutional code with generator $[133, 171]$ and a block interleaver with depth 30 are employed before the differential encoder in Fig. 1. After differential decoding, we perform hard Viterbi decoding. Observing Fig. 6, we notice that thanks to the convolutional code, PSAM collects diversity and coding gains. In the low-to-medium SNR range, PSAM outperforms our BD designs, but roles are reversed at medium-high SNR, where diversity benefits become more pronounced.

5) Test Case 5 (Performance Comparisons With [20]): In this test case, we compare our schemes with the one in [20]. The channels are generated using the Jakes' model with $f_c = 2$ GHz, $v = 160$ km/h, $T_s = 0.7$ μs, and $P = 100$. The BER performance is depicted in Fig. 7. Note that our differential schemes demonstrate performance with full diversity. However, the scheme in [20] can only achieve diversity one. Compared with [20], the price we pay here is threefold: our schemes may exhibit longer decoding delays because of the block transmission design; our schemes have relatively higher decoding complexity; and our transmissions have lower bandwidth efficiency. Furthermore, compared with Fig. 7, we notice that the gap between the performance of BD-I and that of BD-II is broadened. This is because theoretically, BD-I performs well if BEM captures the main channel variation. Since the channel here is generated by the Jakes' model, (11) is not exactly true any more.

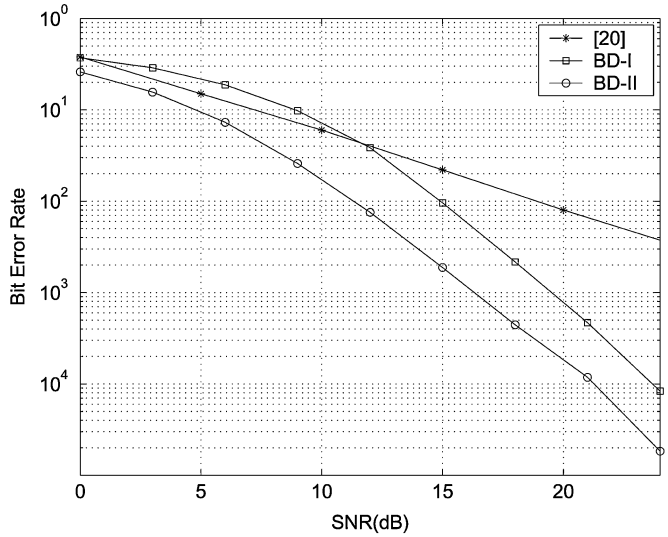


Fig. 7. Performance comparison of BD with [20].

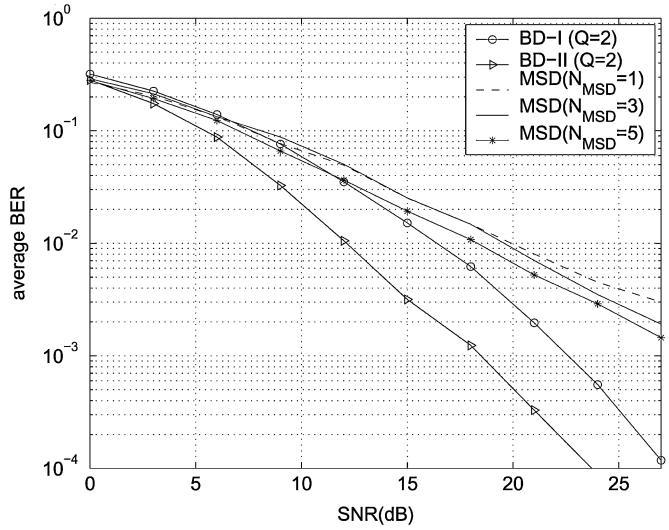


Fig. 8. Performance comparison of BD with MSD [8].

DF-DD can be used to improve the performance of BD-I. However, this goes beyond the scope of this paper.

6) Test Case 6 (Comparisons With [8]): In the literature, MSD has been advocated to deal with channel variations, at the expense of increasing decoding complexity (see, e.g., [8]). The major difference between our schemes and the conventional DPSK with MSD is at the transmitter. Selecting $P = 1\,020$, $f_c = 3.5$ GHz, $v = 250$ km/h, and $T_s = 1\ \mu$s, the number of bases in this case is $Q = 2$. Again, Jakes' model is used to generate the channels. We select the design in [8] as an MSD representative, and test it with different MSD sizes ($N_{\rm MSD} = 1, 3, 5$). Fig. 8 depicts the simulated performance. Thanks to the judicious transmitter design, our schemes achieve higher diversity than [8], even with lower decoding complexity. Note that our receivers have comparable decoding complexity with [8] when $N_{\rm MSD} = 3$. The MSD scheme in [8], though, exhibits diversity order one, even with large $N_{\rm MSD}$ values.

VI. Conclusion

We designed two novel BD schemes that bypass estimation of rapidly varying fading channels. Both encoders take advantage of the maximum Doppler spread at the transmitter. We derived our first differential scheme by converting a time-selective channel to multiple frequency-selective channels, and our second scheme by using block DF-DD or Viterbi decoding. Analysis of the PEP and simulations revealed that both designs enable the maximum available Doppler diversity. Equispaced grouping offered flexibility in BD-I to trade off performance with decoding complexity, while decision feedback allowed BD-II to gain in bandwidth efficiency at the expense of complexity. Numerical simulations indicated that both differential schemes outperform uncoded PSAM, coded PSAM at high SNR, and noncoherent alternatives, especially in rapidly fading channels.[3]

References

[1] K. L. Baum and N. S. Nadgauda, "A comparison of differential and coherent reception for a coded OFDM system in a low C/I environment," in *Proc. Global Communications Conf.*, vol. 1, Phoenix, AZ, Nov. 3–8, 1997, pp. 300–304.

[2] J. K. Cavers, "Pilot-symbol-assisted modulation and differential detection in fading and delay spread," *IEEE Trans. Commun.*, vol. 43, pp. 2206–2212, July 1995.

[3] G. Colavolpe and R. Raheli, "Noncoherent sequence detection in frequency nonselective slowly fading channels," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 2302–2322, Nov. 2000.

[4] Y. E. Dallal and S. Shamai, "Time diversity in DPSK noisy phase channels," *IEEE Trans. Commun.*, vol. 40, pp. 1703–1715, Nov. 1992.

[5] G. H. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1996.

[6] G. B. Giannakis and C. Tepedelenlioglŭ, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proc. IEEE*, vol. 86, pp. 1969–1986, Nov. 1998.

[7] B. Hassibi and B. Hochwald, "Cayley differential unitary space–time codes," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1485–1503, June 2002.

[8] P. Ho and D. Fung, "Error performance of multiple-symbol differential detection of PSK signals transmitted over correlated Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 40, pp. 1566–1569, Oct. 1992.

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

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

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

[12] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

[13] Z. Liu and G. B. Giannakis, "Block differentially encoded OFDM with maximum multipath diversity," *IEEE Trans. Wireless Commun.*, vol. 2, pp. 420–423, May 2003.

[14] L. H.-J. Lampe and R. Schober, "Bit-interleaved coded differential space–time modulation," *IEEE Trans. Commun.*, vol. 50, pp. 1429–1439, Sept. 2002.

[15] X. Ma and G. B. Giannakis, "Maximum-diversity transmissions over doubly selective wireless channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1832–1840, July 2003.

[16] ——, "Complex-field coded MIMO systems: Performance, rate, and tradeoffs," *Wireless Commun. Mobile Comput.*, pp. 693–717, Nov. 2002.

[17] Y. Ma and Q. T. Zhang, "Accurate evaluation for MDPSK with noncoherent diversity," *IEEE Trans. Commun.*, vol. 50, pp. 1189–1200, July 2002.

[18] J. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.

[3]The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government.

[19] R. Schober, W. H. Gerstacker, and J. B. Huber, "Decision-feedback differential detection of MDPSK for flat Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, pp. 1025–1035, July 1999.

[20] R. Schober and L. H.-J. Lampe, "Noncoherent receivers for differential space–time modulation," *IEEE Trans. Commun.*, vol. 50, pp. 768–777, May 2002.

[21] ——, "Differential modulation diversity," *IEEE Trans. Veh. Technol.*, vol. 51, pp. 1431–1444, Nov. 2002.

[22] M. K. Tsatsanis and G. B. Giannakis, "Equalization of rapidly fading channels: Self-recovering methods," *IEEE Trans. Commun.*, vol. 44, pp. 619–630, May 1996.

[23] G. M. Vitetta and D. P. Taylor, "Viterbi decoding of differentially encoded PSK signals transmitted over Rayleigh frequency-flat fading channels," *IEEE Trans. Commun.*, vol. 43, pp. 1256–1259, Feb.-Apr. 1995.

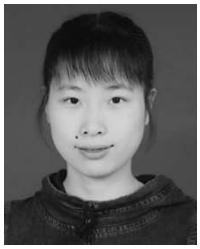
Xiaoli Ma (M'03) received the B.S. degree in automatic control from Tsinghua University, Beijing, P.R. China, in 1998, and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Virginia, Charlottesville, in 1999 and the University of Minnesota, Minneapolis, in 2003, respectively.

Since August 2003, she has been an Assistant Professor with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL. Her research interests include transmitter and receiver diversity techniques for wireless fading channels, communications over time- and frequency-selective channels, complex-field and space–time coding, channel estimation and equalization algorithms, carrier-frequency synchronization for OFDM systems, and wireless sensor networks.

Georgios B. Giannakis (S'84–M'86–SM'91–F'97) received the Diploma in electrical engineering from the National Technical University of Athens, Greece, in 1981. He received the MSc. degree in electrical engineering in 1983, M.Sc. degree in mathematics in 1986, and the Ph.D. degree in electrical engineering in 1986, from the University of Southern California (USC), Los Angeles.

After lecturing for one year at USC, he joined the University of Virginia, Charlottesville, in 1987, where he became a Professor of Electrical Engineering in 1997. Since 1999, he has been with the University of Minnesota, Minneapolis, as a Professor in the Department of Electrical and Computer Engineering, and holds an ADC Chair in Wireless Telecommunications. His general interests span the areas of communications and signal processing, estimation and detection theory, time-series analysis, and system identification, subjects on which he has published more than 180 journal papers, 340 conference papers, and two edited books. Current research focuses on transmitter and receiver diversity techniques for single- and multiuser fading communication channels, comples-field and space–time coding, multicarrier, ultra-wideband wireless communication systems, cross-layer designs, and distributed sensor networks.

Dr. Giannakis is the corecipient of five Best Paper Awards from the IEEE Signal Processing (SP) Society (1992, 1998, 2000, 2001, and 2003). He also received the Society's Technical Achievement Award in 2000. He co-organized three IEEE-SP Workshops, and guest co-edited four special issues. He has served as Editor in Chief for the IEEE SIGNAL PROCESSING L ETTERS, as Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE SIGNAL PROCESSING LETTERS, as secretary of the SP Conference Board, as member of the SP Publications Board, as member and vice-chair of the Statistical Signal and Array Processing Technical Committee, and as chair of the SP for Communications Technical Committee. He is a member of the Editorial Board for the PROCEEDINGS OF THE IEEE, and the steering committee of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He is a member of the IEEE Fellows Election Committee, and the IEEE-SP Society's Board of Governors.



Bing Lu received the B.S. degree in electrical engineering from the University of Science and Technology of China (USTC), Hefei, China, in 2001, and the M.S. degree in electrical engineering from the University of Minnesota, Minneapolis, in 2003.

Her research interest lies in the area of communications and signal processing.