Hindawi Publishing Corporation
EURASIP Journal on Applied Signal Processing
Volume 2006, Article ID 35352, Pages 1–16
DOI 10.1155/ASP/2006/35352

Spectrally Efficient Communication over Time-Varying Frequency-Selective Mobile Channels: Variable-Size Burst Construction and Adaptive Modulation

Francis Minhthang Bui and Dimitrios Hatzinakos

The Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, ON, Canada M5S 3G4

Received 1 June 2005; Revised 10 March 2006; Accepted 15 March 2006

Methods for providing good spectral efficiency, without disadvantaging the delivered quality of service (QoS), in time-varying fading channels are presented. The key idea is to allocate system resources according to the encountered channel. Two approaches are examined: variable-size burst construction, and adaptive modulation. The first approach adapts the burst size according to the channel rate of change. In doing so, the available training symbols are efficiently utilized. The second adaptation approach tracks the operating channel quality, so that the most efficient modulation mode can be invoked while guaranteeing a target QoS. It is shown that these two methods can be effectively combined in a common framework for improving system efficiency, while guaranteeing good QoS. The proposed framework is especially applicable to multistate channels, in which at least one state can be considered sufficiently slowly varying. For such environments, the obtained simulation results demonstrate improved system performance and spectral efficiency.

1. INTRODUCTION

Achieving high spectral efficiency is an important goal in communication. However, it is equally important that the quality of service (QoS), quantified by the bit error rate (BER), will not deteriorate as a result of this goal. We propose strategies that allocate resources for improving the spectral efficiency, while maintaining good QoS, for burst-by-burst communication systems. In these systems, data are transmitted in bursts or blocks, possibly with training and other types of symbols to aid data recovery at the receiver. Over any such burst, the channel is assumed to be sufficiently constant or stationary, that is, a single channel environment is approximately experienced by the entire data burst (also known as a quasi-static or block-fading channel). The rationale for employing burst transmission is that since the channel is approximately the same over the entire received burst, it can be estimated, and a single time-invariant equalizer can be used to mitigate interferences for all data symbols within a single burst. In other words, the various data bursts can be independently processed at the receiver, on a burst-by-burst basis.

Unfortunately, with the advent of the systems employing high-frequency carriers and used in high-speed environments, the quasi-static channel assumption is becoming more questionable. Essentially, the channel can be regarded as constant over a burst if the burst duration is less than the channel coherence time T_C. However, the channel coherence time is itself actually a statistical measure, whose precise formula depends on the definition criterion. Loosely speaking, [1, 2],

$$T_C \approx \frac{1}{f_m} \tag{1}$$

or alternatively, defined as the time over which the time correlation function is above 0.5 [1, 2],

$$T_C \approx \frac{9}{16\pi f_m}, \tag{2}$$

where f_m is the maximum Doppler shift given by

$$f_m = \frac{v_m}{\lambda} = \frac{v_m f_c}{c} \tag{3}$$

with v_m being the mobile speed, λ the wavelength, f_c the carrier frequency, and c the speed of light. The relationship with the burst duration can also be viewed using the normalized Doppler shift $f_m T_S$, where T_S is the symbol duration. Then, using (1), a burst is within a coherence time if the number of symbols in the burst, that is, the burst size B_S, is

$$B_S < \frac{1}{f_m T_S}. \tag{4}$$

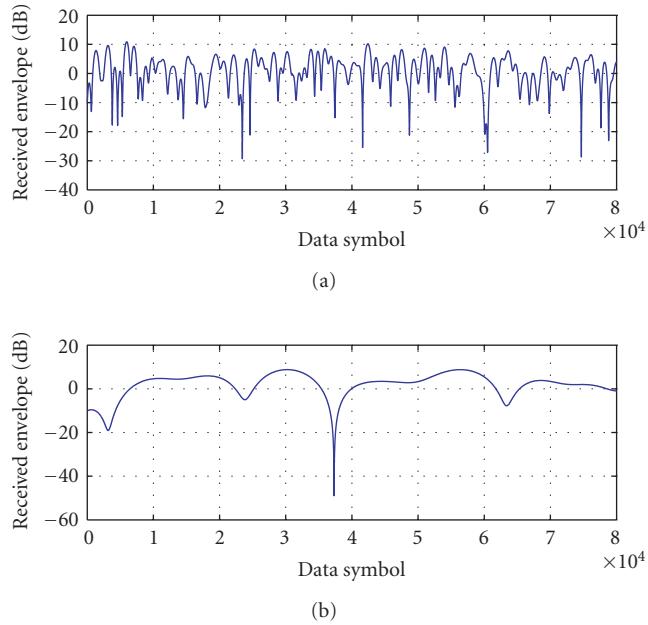


Figure 1: Received envelopes over fading channels at carrier frequency $f_c = 3.5\,\text{GHz}$: (a) mobile speed $v_m = 100\,\text{km/h}$, or normalized maximum Doppler shift $f_m T_S = 5.55 \times 10^{-4}$; (b) $v_m = 10\,\text{km/h}$, or $f_m T_S = 5.55 \times 10^{-5}$.

Regardless of which definition, (1) or (2), is used, the coherence time T_C is inversely proportional to the both carrier frequency f_c and the mobile speed v_m. Hence, with an increase of the carrier frequency f_c in modern systems, T_C tends to become shorter. In practice, the burst duration is chosen to be significantly less than T_C in order to justify the quasi-static assumption. For example, in GSM [1, 2], a burst duration is 0.577 ms, while $T_C \approx 11\,\text{ms}$ (using (1) with $f_c = 960\,\text{MHz}$, $v = 100\,\text{km/h}$).

With an increased carrier frequency, for example, $f_c < 3.5$ GHz in the developing IEEE802.20 standard, the coherence time reduces to $T_C \approx 3.6\,\text{ms}$, and with target bitrates on the order of 1 Mbps, the symbol duration $T_S \approx 2\mu s$ (assuming 2 bits/symbol, e.g., using 4-QAM [2, 3]). Hence, the normalized Doppler shift is $f_m T_S \approx 5.55 \times 10^{-4}$, and a coherence time contains at a maximum $1/(f_m T_S) = 1800$ symbols.

For visualization purposes, Figure 1 shows typical fading envelopes versus the symbol index for the above calculated normalized Doppler shift $f_m T_s \approx 5.55 \times 10^{-4}$, and also for $f_m T_s \approx 5.55 \times 10^{-5}$. Here, the time variations are described by the Jakes power spectral density (see (7)). The smaller normalized Doppler shift corresponds to a more slowly varying channel.

In coping with the reduced coherence time T_C, a number of approaches can be considered. First, the channel invariance assumption can be eliminated, and new receiver structures can be designed. However, suppose that such changes are not permissible, for example, due to existing infrastructure or hardware constraints. Then, the question is whether basic burst-by-burst techniques can still be used in rapidly time-varying channels. We examine techniques for achieving reliable communications in such channels, while still using the same basic burst-by-burst receiver methodology.

Ultimately the goal is to shorten the burst duration in some manner, so that it remains within the coherence duration. Following are example methods that can be considered.

(S1) Reduce the number of data symbols per burst

To reduce the overall burst duration, the symbol duration T_S must not be increased. With this solution, the transmission efficiency, that is, the ratio of useful data symbols over all symbols in a burst, can be severely affected, especially in rapidly varying channels.

(S2) Reduce the burst duration

Alternatively, the same number of symbols in a burst can be maintained, but the symbol duration T_S is reduced. While the transmission efficiency is maintained, if the symbol duration is too short relatively to the channel delay spread, the channel becomes highly frequency selective, with severe intersymbol interference (ISI). The use of a high-complexity equalizer would be needed for acceptable QoS.

(S3) Use a variable-size burst approach

A key bottleneck in the previous two methods is the assumption of a fixed-size burst, chosen to satisfy the worst case scenario. This is inefficient when the encountered channel is slowly changing, for example, when the mobile speed is low. The idea of a variable-size burst [4] is to use a shorter burst when the channel is changing quickly. Conversely, durations

over which the channel is slowly changing will be exploited to use a larger burst. As will be seen in Section 3, this enables a better use of the available training symbols for improved transmission efficiency and QoS. Moreover this construction can be achieved entirely at the receiver.

If the channel quality is further known for each burst, it is also possible to adapt the modulation mode for the data symbols on a burst-by-burst basis. When the channel is benign or of good quality, a higher-order modulation constellation, for example, 16-QAM, can be used for efficiency while still maintaining a good QoS, defined by a target BER. However, when the channel is hostile or of poor quality, a lower-order modulation mode, for example, BPSK, is selected to maintain an acceptable QoS. Known as adaptive modulation [3, 5], this methodology permits an overall improvement in spectral efficiency. Thus, adaptive modulation plays a key role in balancing the system's integrity and efficiency in a time-varying environment.

As will become evident in the remainder of the paper, the overall conclusion of this work is the following: if the underlying time-varying channel can be modeled as multistate, where at least one state is slowly varying, then reliable communication is still possible using conventional burst-by-burst techniques when coupled with a variable-size burst approach. Furthermore, the spectral efficiency can be enhanced with the use of adaptive modulation. When combined together, these two strategies deliver an attractive framework, with minimal modifications of existing systems, for reliable and efficient communication over time-varying channels.

When there is no slow state in the underlying channel, the transmission efficiency is poor since the burst size needs to be very small. By combining variable-size burst construction with basis-expansion modeling (BEM) of the channel [6, 7], the transmission efficiency can be improved. However, in this case, the system complexity is increased due to more complicated estimation and equalization procedures. With some performance loss, the complexity can be reduced significantly using time-varying FIR equalization [8]. But more importantly, even with the addition of basis-expansion modeling, the variable-size burst methodology remains applicable [6]. This is because, under certain conditions, BEM essentially allows a rapidly varying channel to be treated as an equivalent slow fading channel. In fact, at the cost of system complexity, the BEM modification only improves the flexibility of variable-size burst construction, making it applicable to a wider range of time-varying channels [6]. In the interest of brevity and clarity, this work will thus focus on burst construction, and the integration with adaptive modulation, all using conventional channel modeling.

The rest of this paper is organized as follows. After describing a mobile channel model with multistate considerations in Section 2, a variable-size burst structure is presented in Section 3. Channel equalization technique and estimation techniques are then outlined in Section 4. These techniques are subsequently incorporated into a channel-tracking framework for constructing variable-size bursts in Section 5. And to further improve the spectral efficiency, an adaptive modulation method coupled with variable-size burst construction is discussed in Section 6. Next, to demonstrate the performance of the proposed methods, simulation results are obtained in Section 7. Lastly, conclusions are made in Section 8.

2. CHANNEL MODEL

2.1. Mobile fading channels

In this paper, time-varying frequency-selective mobile fading channels are assumed. Under the well-known wide-sense stationary uncorrelated scatterers, (WSSUS) assumptions [2, 9], such channels can be viewed as equivalent time-varying FIR filters, with impulse response

$$h(t,\tau) = \sum_{p=0}^{P-1} \alpha_p(t)\delta(\tau - \tau_p), \tag{5}$$

where P is the number of observable paths, as will as τ_p and $\alpha_p(t)$, respectively, the delay and gain of the pth path.

The time variations, due to the Doppler effect as mentioned in Section 1, are described for each of the P paths by the autocorrelation function [9]:

$$r_p(\tau) = \sigma_p^2 J_0(2\pi f_m \tau) \tag{6}$$

or, equivalently, in the frequency domain, by the Jakes power spectral density:

$$S_p(f) = \begin{cases} \dfrac{\sigma_p^2}{\pi f_m \sqrt{1 - (f/f_m)^2}}, & |f| < f_m, \\ 0, & |f| > f_m, \end{cases} \tag{7}$$

where σ_p^2 is the average power of the pth path, $J_0(\cdot)$ the zero-order Bessel function of the first kind, and f_m the maximum Doppler shift. Note that the coherence time T_C from (2) is defined based on (6).

The channel frequency selectivity is described by specifying the average power for each of the path coefficients $\alpha_p(t)$, resulting in the power-delay profile. For example, a typical urban (TU) COST207-type [3, 9] channel power-delay profile with four observable paths is shown in Figure 2, with parameters summarized in Table 1.

2.2. Multistate extension

While the above mobile channel model is both time and frequency selective, it essentially describes one single *channel state* or environment, where a state is characterized by a particular f_m. From Section 1, f_m is dependent on the mobile velocity v_m for a fixed carrier frequency f_c. Hence, as a user changes his or her mobile activities, the perceived operating environment is also effectively modified. In the context of a variable-size burst, it is beneficial to model such activities explicitly, since the goal is to exploit low-mobility activities for efficiency. To this end, we consider a multistate channel model, where each state is defined by an associated Doppler shift f_m or mobile speed v_m. Evidently, the more states considered, the more accurate is the approximation of the user's mobile activities, at the cost of complexity.

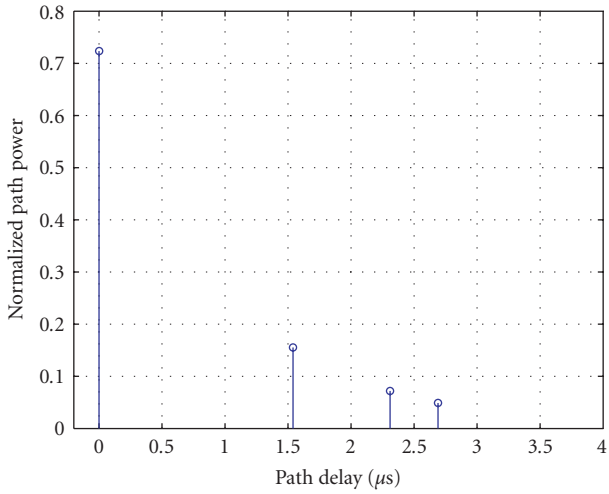


Figure 2: Normalized power-delay profile for a 4-path typical urban (TU) COST207-type channel, with parameters summarized in Table 1.

Table 1: Normalized power-delay profile for a typical urban (TU) COST207-type channel, as depicted in Figure 2.

Delay position (μs)	Path power
0	0.7236
1.54	0.1554
2.31	0.0720
2.69	0.0490

Suppose the user's mobile activities are such that there are κ distinguishable states: $\{k_1, k_2, \ldots, k_\kappa\}$. Denote the probability of the user being in the k_i state as $p(k_i)$, so that

$$\sum_{i=1}^{\kappa} p(k_i) = 1. \tag{8}$$

To fully describe the user's mobile behavior as a function of time, the joint probability mass function (pmf) needs to be specified as a function of the current state, and the past state(s), that is, memory consideration. However, for simplicity, we assume in this paper a memoryless model. Then, the channel states for various time instants can be considered discrete i.i.d random variables, with the pmf specified by

$$p(k_i), \quad i = 1, \ldots, \kappa. \tag{9}$$

Note that when considering a quasi-static channel approximation, the probability of the channel for any burst being in a certain state is specified by (9), that is, on a burst-by-burst basis.

2.3. A Two-state channel example

As an example of a channel with two states, when using a Gauss-Markov approximation to the Jakes model, consider the following composite Gauss-Markov channel, used previously in [4]. Denote the channel taps for the nth time instant as $\mathbf{h}_n$. Let the two states be s-state and f-state. Then the channel changes between time instants as

$$\mathbf{h}_n = \nu(\eta_s \mathbf{h}_{n-1} + \mathbf{u}_s) + (1 - \nu)(\eta_f \mathbf{h}_{n-1} + \mathbf{u}_f), \tag{10}$$

where ν is a Bernoulli random variable, η_s, η_f the correlation coefficients for each state, and $\mathbf{u}_s$, $\mathbf{u}_f$ the noise terms. Hence, by appropriately assigning values to η_s and η_f, the channel can be considered as composing of a slow and a fast state, with state probabilities specified by the Bernoulli rv ν.

For the above composite Gauss-Markov model, each state is specified by parameters relating to the associated Doppler shift f_m, for example, s-state by η_s. In this paper, each channel state is described more generally using (6) and (7).

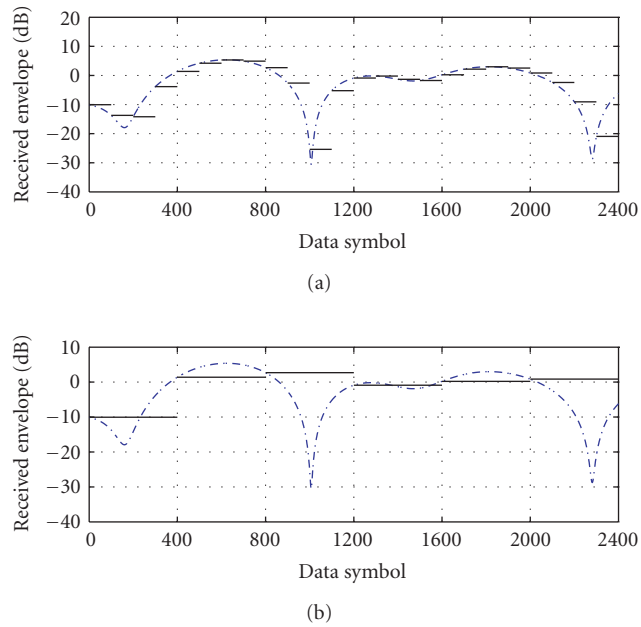


Figure 3: Quasi-static channel approximation for $f_m T_S = 1 \times 10^{-3}$ using: (a) fixed-size bursts of 100 symbols; (b) fixed-size bursts of 400 symbols.

3. VARIABLE-SIZE BURST STRUCTURE

A variable-size burst structure, based on a conventional fixed-size burst, is described in this section.

3.1. Motivation

As mentioned in Section 1, the idea of using a burst transmission system originates from approximating the channel as constant or quasi-static over some interval, which should be less than the coherence time. In the context of a time-varying mobile channel, Figure 3 illustrates this approximation on a channel with normalized Doppler shift $f_m T_S = 1 \times 10^{-3}$ for two different fixed-size bursts: (a) a smaller burst of 100 data symbols; and (b) a larger burst of 400 data symbols. For this scenario, the smaller burst approximates more accurately

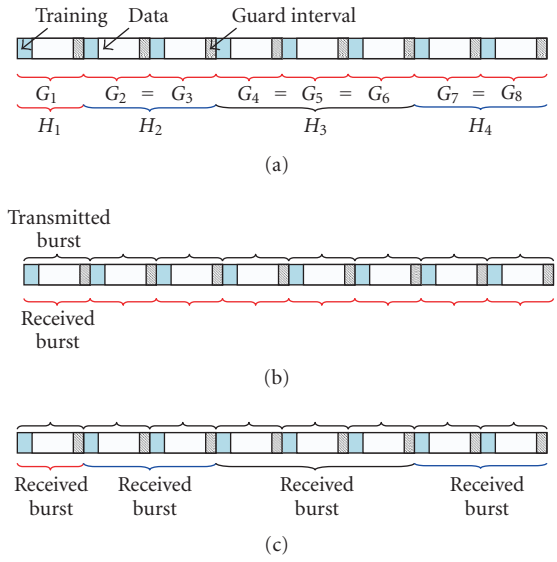


FIGURE 4: Variable-size burst structure with preamble training symbols: (a) quasi-static channel approximations for each burst, where some channels may be the same, for example, $G_2 = G_3 \equiv H_2$; (b) fixed-size burst system, assuming all channels are different; (c) variable-size (received) burst system, exploiting knowledge of channel similarities.

the channel using a total of 24 fixed data bursts. The larger burst approximates the same channel using fewer data bursts, a total of 6 in this case. With a fixed overhead of training symbols per burst, it is more desirable to use the larger burst, since the transmission efficiency (which is proportional to the spectral efficiency) would be higher. However, as illustrated by Figure 3(b), the larger-burst approximation is quite inaccurate at certain times, for example, the deep fade around symbol 1000 is missed entirely. On the other hand, the smaller burst is rather redundant at certain times, for example, over the symbol range 1200–1500, a single-burst approximation suffices. Hence, a compromise between the two different burst sizes, using a variable-size burst, is advantageous in terms of efficiency.

3.2. Accumulated received burst structure

Figure 4 shows a potential variable-size burst structure. The key idea here is to realize the distinction between a transmitted and a received burst: regardless of what the transmitter sends, the receiver ultimately can make a choice on what it considers a received burst (used for further processing, such as channel estimation). Then, the transmitter simply transmits fixed-size *fundamental* bursts. At the receiver, a variable-size burst is constructed by combining consecutive transmitted fundamental bursts appropriately. For this scheme to function, as in a fixed-size burst system, the fundamental bursts need to satisfy the quasi-static channel conditions. The difference is that, by tracking the channel, the receiver can detect a slowly changing duration, and accordingly adapts the burst size by combining the consecutive fundamental bursts within this duration. The result is a larger *accumulated* burst, composed of fundamental bursts, with an enlarged set of training symbols delivering a more accurate channel estimation.

3.3. Example construction

To illustrate the described procedure, Figure 4(a) shows an example scenario, where the channels for eight consecutive fundamental bursts are designated: $G_1, G_2, \ldots, G_8$. A fixed-size burst receiver simply assumes that these channels are all different and constructs received bursts of the same size as the transmitted bursts as shown in Figure 4(b). However, if the underlying channels are not all different, then a variable-size burst can combine appropriate consecutive fundamental bursts to form larger accumulated bursts, while still satisfying the quasi-static assumption. For example, if $G_2 = G_3$, $G_4 = G_5 = G6$, $G7 = G8$ (see Figure 3, e.g., of how this may arise), then the unique channels can be re-designated as H_1, H_2, H_3, H_4, from which there would be four enlarged variable-size accumulated bursts as in Figure 4(c).

3.4. Comparisons to a fixed-size burst

From a transmitter perspective, there is essentially no difference in terms of the burst structure. The fundamental burst size is still specified by the highest-speed f_m. However, in rapidly time-varying channels, the variable-size burst structure is more attractive, because it has the potential to maintain good spectral efficiency.

Indeed, consider using solution (S1), from Section 1, to reduce the number of data symbols per burst. Then, to maintain the same transmission efficiency, the number of training symbols must also be reduced. However, estimation and equalization depend on the raw number of training symbols (and not the transmission efficiency). Hence, a fixed-size burst, which in general has insufficient training symbols in rapidly time-varying channels, will suffer from significant performance degradation due to unsuccessful channel estimation and equalization. By contrast, a variable-size burst has the potential to regain the performance loss by making the best use of the available training symbols.

The effect of training-symbol assignment or placement is not investigated here. While optimal training placement can have a significant impact on the overall performance [10], the present paper has a different perspective: given a training regime (e.g., preamble, midamble, or superimposed), the problem is how to combine the available training symbols from different bursts in an advantageous manner, notably by tracking the channel. This is based on the assumption that more training symbols would yield better overall performance.

4. CHANNEL EQUALIZATION AND ESTIMATION

The proposed variable-size burst scheme requires the receiver to correctly detect the channel changes. Such channel-tracking capability is designed by modifying conventional quasi-static channel equalization and estimation techniques.

First, we will describe the ideal minimum mean-square (MMSE) equalizer, assuming knowledge of the channel. Then, using training symbols, a maximum-likelihood (ML) estimator provides an estimate of the channel. Throughout this section, it is assumed that the accumulated burst is already received under quasi-static channel conditions. In Section 5, the channel estimation and equalization techniques described here will be incorporated in a framework for constructing a quasi-static accumulated burst.

4.1. MMSE equalization

Consider the typical equivalent baseband signal representation

$$y[n] = \sum_{l=0}^{L-1} h[n;l]x[n-l] + v[n], \tag{11}$$

where $x[n]$ is the transmitted symbol at instant n, $y[n]$ the received symbol, $h[n;l]$ the channel impulse response, L the channel length (assumed known), and $v[n]$ the additive white Gaussian noise (AWGN) with variance σ_v^2. When the channel is time invariant as in a burst-by-burst system, the dependence of $h[n;l]$ on n is suppressed:

$$y[n] = \sum_{l=0}^{L-1} h[l]x[n-l] + v[n] = h[n] \star x[n] + v[n], \tag{12}$$

where $\star$ denotes convolution. In this case, a matrix formulation can be obtained. At the instant n, for the potential recovery of the nth symbol $x[n]$, N consecutive received symbols are collected as

$$\mathbf{y}(n) = \mathbf{H}\mathbf{x}(n) + \mathbf{v}(n) \tag{13}$$

with $\mathbf{y}[n] = [y[n],\ldots,y[n-N+1]]^T$, $\mathbf{v}[n] = [v[n],\ldots,v[n-N+1]]^T$, $\mathbf{x}[n] = [x[n],\ldots,x[n-N-L+2]]^T$,

$$\mathbf{H} = \begin{bmatrix} h[0] & \cdots & h[L-1] & \cdots & 0 \\ & & \ddots & & \\ 0 & \cdots & h[0] & \cdots & h[L-1] \end{bmatrix}, \tag{14}$$

where $(\cdot)^T$ denotes matrix transpose, and $\mathbf{H}$ has dimensions $N \times (N+L-1)$.

Using the minimum mean-squared error (MMSE) criterion, a linear equalizer $\mathbf{f} = [f[0], f[1],\ldots, f[N-1]]^T$ is found by minimizing the cost function

$$J_{\mathrm{MSE}}(\mathbf{f}) = E(|\mathbf{f}^H\mathbf{y}(n) - x[n-\delta]|^2), \tag{15}$$

where $E(\cdot)$ denotes the expectation operator, $(\cdot)^H$ the Hermitian transpose, and δ is a delay, with permissible values $\delta = 0,\ldots,N+L-1$ (see (18) and (19) for the effect of δ). The solution to (15) is [11]

$$\mathbf{f} = \mathbf{R}^{-1}\mathbf{p}, \tag{16}$$

where $\mathbf{R} = E(\mathbf{y}(n)\mathbf{y}^H(n))$, $\mathbf{p} = E(x^*[n-\delta]\mathbf{y}(n))$ are known, respectively, as the autocorrelation and cross-correlation. Making the independence assumption of data symbols at different instants, then

$$\mathbf{R} = \sigma_x^2\mathbf{H}\mathbf{H}^H + \sigma_v^2\mathbf{I}_N, \qquad \mathbf{p} = \sigma_x^2\mathbf{H}\mathbf{1}_{\delta+1}, \tag{17}$$

where $\sigma_x^2 = E(|x[n]|^2)$ is the symbol energy, σ_v^2 the noise variance, $\mathbf{I}_N$ the $N \times N$ identity matrix, and $\mathbf{1}_\delta$ an all-zero vector except for the δ element, which is equal to 1 (hence, in (17), $\mathbf{1}_{\delta+1}$ extracts the $(\delta+1)$th column of $\mathbf{H}$).

Given a fixed channel matrix $\mathbf{H}$ [11],

$$\mathrm{MMSE}(\delta-1) = \sigma_x^2(1 - \mathbf{1}_\delta^H\mathbf{H}^H\Delta^{-1}\mathbf{H}\mathbf{1}_\delta), \tag{18}$$

where $\Delta = \mathbf{H}\mathbf{H}^H + \sigma_v^2/\sigma_x^2\mathbf{I}_N$. Hence, the optimal δ can be found by evaluating

$$\Xi = \mathrm{diag}\left(\sigma_x^2(\mathbf{I}_N - \mathbf{H}^H\Delta^{-1}\mathbf{H})\right) \tag{19}$$

from which $(\delta-1)$ corresponds to the row number of Ξ with the minimum value (e.g., if the first row element is the minimum, the delay is $\delta = 0$).

4.2. ML channel estimation

The channel $h[n]$ can be estimated using an ML estimator, with training symbols. This is ultimately where the variable-size burst advantage is realized: a larger accumulated burst provides more training and thus better channel estimate.

Consider the first fundamental burst in an accumulated burst, with M consecutive training symbols located by the index set $I_1 = \{k,\ldots,k+M-1\}$, that is, $x[k],\ldots,x[k+M-1]$ are known symbols. The received signal is

$$\mathbf{y}_{I_1} = \mathbf{x}_{I_1}\mathbf{h} + \mathbf{v}_{I_1}, \tag{20}$$

where $\mathbf{y}_{I_1} = [y[k+L-1],\ldots,y[k+M-1]]^T$, $\mathbf{v}_{I_1} = [v[k+L-1],\ldots,v[k+M-1]]^T$, $\mathbf{h} = [h[0],\ldots,h[L-1]]^T$,

$$\mathbf{x}_{I_1} = \begin{bmatrix} x[k+L-1] & \cdots & x[k] \\ \vdots & & \vdots \\ x[k+M-1] & \cdots & x[k+M-L+1] \end{bmatrix}. \tag{21}$$

Note that when preamble training and zero-padding guard intervals are used (see Figure 4), then the dimensions of the above quantities can be enlarged for better estimation. If $x[k-L+1],\ldots,x[k-1]$ correspond to the guard symbols and are thus known to be all equal to zero, then the received signal can be formed as $\mathbf{y}_{I_1} = [y[k],\ldots,y[k+M-1]]^T$, with appropriate modifications of the related quantities from (20).

Similarly, the second fundamental burst has training symbols with the index set $I_2 = B \oplus I_1$, where $\oplus$ denotes element-wise addition with a scalar B, which is the number of symbols in a fundamental burst. Then, $\mathbf{y}_{I_2} = \mathbf{x}_{I_2}\mathbf{h} + \mathbf{v}_{I_2}$. Thus, if there are μ fundamental bursts in the accumulated

burst,

$$\begin{bmatrix} \mathbf{y}_{I_1} \\ \ddots \\ \mathbf{y}_{I_\mu} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{I_1} \\ \ddots \\ \mathbf{x}_{I_\mu} \end{bmatrix} \mathbf{h} + \begin{bmatrix} \mathbf{v}_{I_1} \\ \ddots \\ \mathbf{v}_{I_\mu} \end{bmatrix} \tag{22}$$

or

$$\mathbf{y}_\Sigma = \mathbf{x}_\Sigma \mathbf{h} + \mathbf{v}_\Sigma. \tag{23}$$

The ML channel estimate is

$$\mathbf{h}_{\mathrm{ML}} = \mathbf{x}_\Sigma^\dagger \mathbf{y}_\Sigma, \tag{24}$$

where $(\cdot)^\dagger$ denotes the Moore-Penrose pseudoinverse [11].

5. CHANNEL TRACKING FOR VARIABLE-SIZE BURST

In this section, the described quasi-static estimation and equalization methods will be incorporated into a threshold-based scheme for detecting channel changes. A receiver procedure for processing variable-size bursts is also presented.

5.1. Threshold-based change detection

The variable-size burst construction problem can be stated iteratively. Suppose that, at the current iteration, the accumulated burst $\mathbf{B}_{\mathrm{current}}$ is composed of μ consecutive fundamental bursts, $\mathbf{B}_{\mathrm{current}} = \{b_k, \ldots, b_{k+\mu-1}\}$, and that the channel is the same over the entire $\mathbf{B}_{\mathrm{current}}$. Then, upon the reception of the candidate fundamental burst $b_{k+\mu}$, the choices are the following.

(H1) Add $b_{k+\mu}$ to the current accumulated burst, forming $\mathbf{B}_{\mathrm{potential}} = \{b_k, \ldots, b_{k+\mu}\}$. Continue with $b_{k+\mu+1}$ as the next candidate.

(H2) Reject $b_{k+\mu}$, terminate $\mathbf{B}_{\mathrm{current}}$, and accept it as the best choice. Reinitialize with $b_{k+\mu}$ as the start of a new accumulated burst.

To decide whether to accept (H1) or (H2), the following procedure is performed.

(1) In (24), estimate the channel using $\mathbf{B}_{\mathrm{current}}$, returning an estimate $\mathbf{h}_C$.

(2) Similarly, estimate the channel using $\mathbf{B}_{\mathrm{potential}}$, returning an estimate $\mathbf{h}_P$.

(3) Compute the squared norm of the estimation difference:

$$\rho_{\mathrm{ed}} = \left| \mathbf{h}_C - \mathbf{h}_P \right|^2. \tag{25}$$

(4) Compare to a threshold ρ_{th} for detection decision:

$$\rho_{\mathrm{ed}} - \rho_{\mathrm{th}} \underset{\mathrm{H1}}{\overset{\mathrm{H2}}{\gtrless}} 0. \tag{26}$$

In the above, ρ_{ed} is a second-order measure of the channel change in the following sense. Suppose that the underlying channel of $\mathbf{B}_{\mathrm{current}}$ is $\mathbf{h}$, and that $\mathbf{h}_C$ is a close estimate of the true channel. Then if $b_{k+\mu}$ experiences the same $\mathbf{h}$, the resulting estimation difference

$$\mathbf{h}_{\mathrm{ed}} = \mathbf{h}_C - \mathbf{h}_P \tag{27}$$

is small (in some norm). But if the channel has changed for the candidate $b_{k+\mu}$, the estimation difference $\mathbf{h}_{\mathrm{ed}}$ is large. In (25), a squared norm is used to quantify this difference. The utility of this choice is made evident by examining (29) and (30), as explained next.

5.2. Threshold function selection

Let the true channel be $\mathbf{h}$, then depending on the detection decision (i) or (ii), the channel estimation error $\mathbf{h}_{\mathrm{ce}}$ is either $\mathbf{h}_{\mathrm{ce},C} = \mathbf{h} - \mathbf{h}_C$ or $\mathbf{h}_{\mathrm{ce},P} = \mathbf{h} - \mathbf{h}_P$. The channel estimation error is unknown, since the true $\mathbf{h}$ is not available. However, an upperbound for its squared norm can be approximated as follows. Noting that $|\mathbf{h}_{\mathrm{ed}}|^2 = |\mathbf{h}_{\mathrm{ce},C} - \mathbf{h}_{\mathrm{ce},P}|^2$ and assuming independence of the estimation errors, so that $E(\mathbf{h}_{\mathrm{ce},C}^* \mathbf{h}_{\mathrm{ce},P}) = E(\mathbf{h}_{\mathrm{ce},C} \mathbf{h}_{\mathrm{ce},P}^*) = 0$,

$$E\left(\left|\mathbf{h}_{\mathrm{ed}}\right|^2\right) \approx E\left(\left|\mathbf{h}_{\mathrm{ce},C}\right|^2\right) + E\left(\left|\mathbf{h}_{\mathrm{ce},P}\right|^2\right) \geq E\left(\left|\mathbf{h}_{\mathrm{ce}}\right|^2\right) \tag{28}$$

which means that by keeping the estimation difference $\mathbf{h}_{\mathrm{ed}}$ small as in (26), the resulting channel estimation error $\mathbf{h}_{\mathrm{ce}}$ should also be statistically small.

Next, consider the effect of a channel estimation error, with impulse response $h_{\mathrm{ce}}[n]$, at the equalizer input. From (12),

$$\begin{aligned} y[n] &= h[n] \star x[n] + v[n] \\ &= (h[n] - h_{\mathrm{ce}}[n]) \star x[n] + \overbrace{h_{\mathrm{ce}}[n] \star x[n]}^{\widetilde{v}[n]} + v[n] \\ &= \widetilde{h}[n] \star x[n] + \widetilde{v}[n] + v[n], \end{aligned} \tag{29}$$

where $\widetilde{h}[n]$ is the estimated channel impulse response (i.e., corresponds to either $\mathbf{h}_C$ or $\mathbf{h}_P$ depending on the detection decision). Hence for an equalizer using the estimated channel $\widetilde{h}[n]$, the second term $\widetilde{v}[n]$, due to the channel estimation error, can be viewed as an additional noise source. For a particular channel realization, this estimation noise error has variance:

$$E\left(\left|h_{\mathrm{ce}}[n] \star x[n]\right|^2\right) = \sigma_x^2 \sum_{l=0}^{L-1} \left|h_{\mathrm{ce}}[l]\right|^2 = \sigma_x^2 \rho_{\mathrm{ce}}, \tag{30}$$

where σ_x^2 is the average symbol energy. From (29), when noise is significant (low SNR), a small estimation error does not necessarily deliver significant performance gain. However, at high SNR, the channel estimation error becomes the bottleneck. In fact, it is well known that channel estimation error can result in an error floor at high SNR [11]. Hence, with a fixed average symbol energy σ_x^2, the channel estimation error variance (30) should be proportional to the channel noise variance σ_v^2 for optimal performance tradeoff.

The above implies that the optimal threshold ρ_{th} in (26) needs to be function of the noise variance. Since the primary goal of this paper is to demonstrate the performance

ρ_{th}: threshold for decision.
N_{total} : total number of fundamental bursts to be processed.
$\text{bsize}_{\max}$: max. number of fundamental bursts in the accumulated burst.
s: fundamental burst defining start of the current accumulated burst.

(I) *Initialization*
  (1) Set $s = 1$
(II) *Iteration*
  for $i = 2, 3, \ldots, N_{\text{total}}$
    if $(i - s + 1 \geq \text{bsize}_{\max})$ or $(i = N_{\text{total}})$,
      (1) Set current accumulated burst = all fundamental bursts from s to i,
      (2) Equalize the current accumulated burst,
      (3) Reset $s = i + 1$,
    else if $(\rho_{\text{ed}} > \rho_{\text{th}})$,
      (1) Set current accumulated burst = all fundamental bursts from s to $i - 1$,
      (2) Equalize the current accumulated burst
      (3) Reset $s = i$,
    end
  end

Algorithm 1: Variable-size burst receiver with channel tracking.

improvement compared to a fixed-size burst in time-varying environments, the effect of threshold optimization will not be explored. Instead, in Section 7, a sensibly predetermined threshold function ρ_{th}, weighted against the noise variance σ_v^2, will be used to assess potential improvement.

5.3. Receiver processing with a variable-size burst

Implicit in the tracking procedure is the requirement of a buffer for computing the intermediate $\mathbf{h}_1$ and $\mathbf{h}_2$, which introduces additional complexity and also latency. To alleviate the incurred penalties, a maximum burst size can be imposed. Fortunately, as evidenced in Section 7, a modest burst size can yield significant performance gain. In fact, when the receiver already has sufficient training to equalize the channel accurately, that is, approaching the MMSE lower-bound, enlarging the accumulated burst does not produce further appreciable improvement. Also, constraining the burst size minimizes the propagation of estimation errors. At low SNR, with inaccurate channel estimates, tracking can erroneously accumulate more fundamental bursts than possible, thus violating the quasi-static requirement.

Accounting for the above factors, Algorithm 1 shows a conceptual receiver procedure for processing variable-size bursts. Essentially, while the accumulated burst has not exceeded the maximum size, the receiver iteratively considers consecutive candidate fundamental bursts for inclusion, using a threshold-based change detection scheme.

5.4. Constrained optimization interpretation

Let the objective $F(\mu) = M\mu$ be the total number of training symbols as a function of M, the number of training symbols in a fundamental burst (see (20)), and μ, the number of fundamental bursts in the accumulated burst (see (22)). Note that M is typically a fixed constant, defined by the training density. Also, let $\mathbf{h}_i$ be the channel associated with the ith fundamental burst in the accumulated burst. Then variable-size burst construction is equivalent to a mixed-integer optimization problem: [12].

Lemma 1. *There exists a unique solution to the following burst construction problem:*

$$\begin{aligned} &\textit{maximize } F(\mu) = M\mu \\ &\textit{subject to } \mu \in \mathbb{Z}\textit{(an integer)}; \quad \mu \leq \text{bsize}_{\max}, \\ &\qquad \mathbf{h}_1 = \mathbf{h}_2 = \cdots = \mathbf{h}_\mu \textit{ (channel invariance).} \end{aligned} \tag{31}$$

Proof. The result follows trivially by noting that $F(\mu)$ is a strict monotonic increasing function of μ. Hence, constrained to a bounded domain, there exists a unique maximum. □

Remarks

If, instead, the objective function is the training density, where the number of training symbols can be adapted per burst, then the optimization problem is not necessarily mixed integer (and M represents essentially a step-size parameter). However, in this case the transceiver design would be more complicated, with some form of feedback required.

Since the existence of a unique solution is guaranteed by Lemma 1, an iterative search for the solution can be implemented. Here, the main difficulty is ensuring that the channel invariance constraint in (31) is maintained. The channels $\mathbf{h}_i$ are not known, and estimates $\widehat{\mathbf{h}}_i$ must be used. Then in the presence of noise and estimation error, with probability one, $\widehat{\mathbf{h}}_1 \neq \widehat{\mathbf{h}}_2 \neq \cdots \neq \widehat{\mathbf{h}}_\mu$, for all μ. Hence, consider instead the equivalent form of the constraint $|\mathbf{h}_{i+1} - \mathbf{h}_i|^2 = 0$, $i = 1, \ldots, \mu - 1$ yielding the squared norm relaxation [12]

$$\left|\mathbf{h}_{i+1} - \mathbf{h}_i\right|^2 < \rho_{\text{th}}, \quad i = 1, \ldots, \mu - 1, \tag{32}$$

where ρ_{th} is a small constant, allowing for some flexibility in accommodating channel estimation error. Essentially, this entails choosing ρ_{th} as in Section 5.2.

Also, at the kth iteration, instead of simply checking $|\widehat{\mathbf{h}_k} - \widehat{\mathbf{h}_{k-1}}|^2$ against the threshold, $|\mathbf{h}_C - \mathbf{h}_P|^2$ as defined by (25) is used to guarantee the constraint. This allows for improved estimation consistency since more training symbols are used for estimation with more iterations.

Algorithm 1 implements the described strategy to iteratively search for μ, which approaches the optimal solution in the squared norm sense.

6. ADAPTIVE MODULATION

The basic scheme of closed-loop burst-by-burst adaptive modulation can be summarized as follows [3, 13].

(1) At the receiver, perform a channel-quality measurement, returning a channel metric.
(2) Relate this channel metric to a suitable modulation mode, which yields the highest throughput while maintaining the required level of QoS.
(3) Signal the selected modulation mode to the transmitter to be used in the next transmission burst.

Note that the average transmitted symbol energy σ_x^2 can be kept the same, regardless of the modulation mode in use. This alleviates the need of power control, which is typical for alternative systems operating in fading channels. The QoS is nonetheless guaranteed, by using the suitable modulation mode for an operating channel quality. In addition, the symbol rate is maintained constant so that the required bandwidth is unchanged, regardless of the selected modulation mode.

6.1. Channel metric

The most accurate metric for quantifying the channel quality is the BER. However, since the BER is often difficult to estimate directly, alternatives are often used instead. For a frequency-non selective or flat-fading channel, the short-term signal-to-noise ratio (SNR) is an appropriate metric [3, 13]. For a frequency-selective channel, the short-term SNR is inadequate, since the influence of ISI must be taken into account. Moreover the BER performance for frequency-selective channel is a complicated function of many factors, including channel length, power-delay profile, and even the form of equalizer used, for example, the number-taps in a linear equalizer, and the value of the equalizer delay. In the following, we outline three possible approaches for computing a channel metric, which can be used to guarantee a target QoS by selecting the appropriate modulation mode.

(1) Exact residual ISI

Given enough side information, the exact probability of error can be computed. Consider the overall equalized channel impulse response:

$$g[n] = f^*[n] \star h[n], \tag{33}$$

where $f[n]$ and $h[n]$ are the impulse responses of the equalizer and the channel, respectively. Following [14], consider the equalizer output at instant n

$$\begin{aligned} z[n] &= f^*[n] \star y[n] \\ &= g[\delta]x[n-\delta] + \sum_{k\neq\delta} g[k]x[n-k] + \sum_{k=0}^{N-1} f^*[k]v[n-k], \end{aligned} \tag{34}$$

where the first term is the desired signal component, the second term the residual ISI, and the last term the equalized noise. Note that $g[n]$ is effectively an FIR filter of length $N+L-1$. Hence, for a particular input sequence $\mathbf{x}_J$ of $N+L-1$ symbols, the corresponding residual ISI term is

$$D_J = \sum_{k\neq\delta} g[k]x_J[n-k]. \tag{35}$$

When using M-PAM, the resulting probability of error is [14]

$$P_M(D_J) = \frac{2(M-1)}{M} Q\left(\sqrt{\frac{(g[\delta]-D_J)^2}{\sigma_n^2}}\right), \tag{36}$$

where σ_n^2 is the variance of the equalized noise

$$\sigma_n^2 = \sigma_v^2 \sum_{n=0}^{N-1} |f[n]|^2. \tag{37}$$

Hence, for a particular channel, input sequence and M, the exact probability of error can be found. A channel metric can then be defined as

$$\Gamma_{\text{ISI}} = D_J, \tag{38}$$

and the appropriate modulation mode, that is, the value of M, can be determined from (35) for a desired QoS. Unfortunately, this exact metric is not practical, since knowledge of $N+L-1$ data symbols surrounding the desired symbol $x[\delta]$ is required (which implies knowledge of the entire sequence of data).

Alternatively, an average and an upper-bound probability of error can be found, respectively, as [14]

$$P_M = \sum_{\mathbf{x}_J} P_M(D_J)P(\mathbf{x}_J), \tag{39}$$

$$P_M(D_J^*), \quad D_J^* = (M-1)\sum_{k\neq\delta} |g[k]|, \tag{40}$$

where (39) is an average over all possible $\mathbf{x}_J$, and (40) is due to the worst-case residual ISI. Unfortunately, the former is computationally expensive, while the latter tends to be rather loose. In addition, for a fading environment, averaging over all fading-channel realizations is required. Thus the exact residual ISI metric is only appropriate for channels with very short length.

(2) Pseudo-SNR

The pseudo-SNR is basically the SNR at the equalizer output:

$$\text{pseudo-SNR} = \frac{\text{wanted signal power}}{\text{residual ISI + noise power}}, \tag{41}$$

and is defined in terms of the coefficients of a decision-feedback equalizer in [3]. Using a linear MMSE equalizer with delay δ,

$$\Gamma_{\text{pSNR}} = \frac{\sigma_x^2 |g[\delta]|^2}{\sigma_x^2 \sum_{k=\delta} |g[k]|^2 + \sigma_n^2} \tag{42}$$

for a particular channel realization, where σ_n^2 is found using

(37). Note that as in [3], a Gaussian approximation of the residual ISI term is made, and independence of the residual ISI and noise is assumed. Then the BER formula in an AWGN channel can be used. For example, the BER for a particular channel realization with 4-QAM:

$$P(\Gamma_{\text{pSNR}}) = P_{\text{4-QAM}}^{(\text{awgn})}(\Gamma_{\text{pSNR}}) = Q\left(\sqrt{\Gamma_{\text{pSNR}}}\right), \tag{43}$$

and more importantly the BER over a mobile fading channel can be found, for a specific m-QAM mode, as

$$P_{m\text{-QAM}}^{(\text{mf})}(\bar{\gamma}) = \int_0^{\infty} P_{m\text{-QAM}}^{(\text{awgn})}(\Gamma_{\text{pSNR}})\, p(\Gamma_{\text{pSNR}}, \bar{\gamma})\, d\Gamma_{\text{pSNR}}, \tag{44}$$

where $\bar{\gamma}$ is the average channel SNR:

$$\bar{\gamma} = \frac{E(|h[n] \star x[n]|^2)}{E(|v[n]|^2)}, \tag{45}$$

$P_{m\text{-QAM}}^{(\text{awgn})}(\cdot)$ the AWGN BER expressions for the m-QAM mode (e.g., can be found in [3, 14]); and $p(\Gamma_{\text{pSNR}}, \bar{\gamma})$ the pdf of the pseudo-SNR Γ_{pSNR} over all fading channel realizations, at a certain average channel SNR $\bar{\gamma}$. In general, the closed-form pdf is not available, and the (discretized) pdf needs to be computed numerically, at each $\bar{\gamma}$ of interest [3]. With Γ_{pSNR} as a channel metric, the appropriate m-QAM mode is selected from (44) for a target QoS.

(3) MSE-based metric

The pseudo-SNR metric requires knowledge of the channel $h[n]$. For methods that find the equalizer $\mathbf{f}$ directly without estimating $h[n]$, a channel metric can be defined based on the MSE computed at the equalizer output [5]. In the sequel, the relationship between the MSE-based metric and the pseudo-SNR is established.

At the equalizer output (34),

$$z[n] = f^*[n] \star y[n] = x[n-\delta] + e[n], \tag{46}$$

where $x[n-\delta]$ is the desired component, and $e[n]$ the overall residual equalization error, which, combines residual ISI, equalized noise, and also scaling. Then, the MSE is the equalization error variance,

$$\sigma_e^2 = E(|e[n]|^2) = E(|x[n-\delta] - z[n]|^2), \tag{47}$$

and can be estimated using training symbols [5]. A corresponding channel metric is

$$\Gamma_{\text{MSE}} = \frac{\sigma_x^2}{\sigma_e^2}. \tag{48}$$

Table 2: Threshold-based switching rules for adaptive modulation.

Switching criterion	Modulation mode
$0 \le \Gamma_C < t_1$	V_1
$t_1 \le \Gamma_C < t_2$	V_2
$\vdots$	$\vdots$
$t_{Q-1} \le \Gamma_C < \infty$	V_Q

Making the assumption of independence between data symbols, residual ISI, and noise,

$$\Gamma_{\text{pSNR}} = \frac{\sigma_x^2 |g[\delta]|^2}{\sigma_e^2 - \sigma_x^2 |g[\delta] - 1|^2}. \tag{49}$$

Comparing (48) and (49), the two metrics are identical when $g[\delta] = 1$, which occurs when the ISI is completely suppressed by the equalizer (at high SNR).

In general, the relationship between the probability of error and MSE is not expressible in a simple closed form. But an upperbound can be obtained [15],

$$P_e(\sigma_e^2) \le \exp\left(-\frac{1 - \sigma_e^2/\sigma_x^2}{\sigma_e^2}\right). \tag{50}$$

Then, the same approach as (44) applies, using the pdf of Γ_{MSE}, which is close to the pdf Γ_{pSNR} at high SNR.

6.2. Threshold-based mode adaptation

Consider a general channel metric Γ_C, for example, $\Gamma_C = \Gamma_{\text{pSNR}}$, which quantifies in some manner the operating channel quality. A threshold-based scheme can be constructed as follows [3, 5]. Designate the choice of available modulation modes by V_q, $q = 1, \ldots, Q$, where Q is the total number of available modulation modes; V_1 is the constellation with the least number of points (most robust); and V_Q the highest (most efficient). Then Table 2 shows the switching rules, based on a set of thresholds $(t_1, \ldots, t_{Q-1})$, where $t_1 < t_2 < \cdots < t_{Q-1}$ are chosen to guarantee some required level of QoS [3].

6.3. Thresholds selection

For a set of thresholds $(t_1, \ldots, t_{Q-1})$, the mean throughput (number of bits per symbol) [3, 16]

$$\begin{aligned} B(\bar{\gamma}) = {} & B_{V_1} \int_0^{t_1} p(\Gamma_C, \bar{\gamma})\, d\Gamma_C \\ & + \sum_{q=2}^{Q-1} B_{V_q} \int_{t_{q-1}}^{t_q} p(\Gamma_C, \bar{\gamma})\, d\Gamma_C \\ & + B_{V_Q} \int_{t_{Q-1}}^{\infty} p(\Gamma_C, \bar{\gamma})\, d\Gamma_C, \end{aligned} \tag{51}$$

where B_{V_q} is the throughput associated with the V_q mode

N_{total} : total number of fundamental bursts to be processed.
s: starting fundamental burst of current accumulated burst.
γ_C: a channel quality metric (e.g., Γ_{pSNR}).

(I) *Initialization*

(1) Set $s = 1$,
(2) Measure channel metric γ_C using sth fundamental burst,
(3) Request QAM-mode(γ_C) to transmitter for the rest of current accumulated bursts.

(II) *Iteration*
for $i = 2, 3, \ldots, N_{\text{total}}$
Track channel starting from sth fundamental burst (using tracking strategy from Section 5, Algorithm 1)
⋮
if (channel change detected at ith fundamental burst)
(1) Set current accumulated burst = all fundamental bursts from s to $i - 1$,
(2) Decode the current accumulated burst,
(3) Reset $s = i$ (i.e., start of new accumulated, burst)
(4) Measure channel metric γ_C using sth fundamental burst,
(5) Request QAM-mode(γ_C) to Tx for the rest of the new accumulated burst.
end
end

Algorithm 2: Adaptive modulation with variable-size burst.

(e.g., throughput of 16-QAM is 4 bps). In a fading channel, the average BER for adaptive modulation

$$P_{\text{AM}}^{(\text{mf})}(\bar{\gamma}) = \frac{1}{B(\bar{\gamma})}\Bigg[B_{V_1}\int_0^{t_1} P_{V_1}^{(\text{awgn})}(\Gamma_C)\,p(\Gamma_C,\bar{\gamma})\,d\Gamma_C + \sum_{q=2}^{Q-1} B_{V_q}\int_{t_{q-1}}^{t_q} P_{V_q}^{(\text{awgn})}(\Gamma_C)\,p(\Gamma_C,\bar{\gamma})\,d\Gamma_C + B_{V_Q}\int_{t_{Q-1}}^{\infty} P_{V_Q}^{(\text{awgn})}(\Gamma_C)\,p(\Gamma_C,\bar{\gamma})\,d\Gamma_C\Bigg]. \tag{52}$$

Hence, with (52), the thresholds can be optimized to produce a desired QoS, for example, using a cost function based on desired BER and average throughput [3, 16].

6.4. Integration with variable-size burst construction

A two-layer strategy is used for adaptation: variable-size burst construction in the first layer, and adaptive modulation method in the second. Feedback is required only in the second layer. A conceptual algorithm for this strategy is summarized in Algorithm 2.

Note that the channel quality is measured once per accumulated burst, that is, the metric obtained with the starting fundamental burst selects the modulation mode for the entire accumulated burst. This is valid because, with channel tracking, the same channel condition, that is, same channel quality, applies to the entire burst.

6.5. Proof of optimality

Let the objective $G(q) = \log_2 q$ be the throughput (number of transmitted bits per symbol) as a function of the modulation mode q. For simplicity, let us assume that there are four modulation modes, that is, $q = 0$ (no transmission), 2 (BPSK), 4 (4-QAM), 16 (16-QAM). Then adaptive modulation with variable-size burst is equivalent to

$$\begin{aligned} &\text{maximize } G(q) = \log_2 q, \\ &\text{subject to } \mu \in \mathbb{Z}(\text{an integer}), \quad \mu \le \text{bsize}_{\max}, \\ &\qquad \mathbf{h}_1 = \mathbf{h}_2 = \cdots = \mathbf{h}_\mu \text{ (channel invariance)}, \\ &\qquad \text{BER}(\mu, q), \le \text{BER}_{\max}, \quad q \in \{0, 2, 4, 16\}, \\ &\qquad \sigma_x^2 = \text{constant}, \end{aligned} \tag{53}$$

where $\text{BER}_{\max}$ specifies the maximum acceptable bit-error rate for a desired QoS, and $\sigma_x^2 = E(|x[n]|^2)$ is the symbol energy.

Proposition 1. *Under the constraints in* (53)*, the given joint optimization problem of burst construction and adaptive modulation has a unique solution. Moreover, the joint optimization is actually separable, that is, burst construction and adaptive modulation can be performed separately in a two-layer strategy.*

Proof. (i) The objective $G(q)$ is a strict monotonic increasing function of q.

(ii) When channel estimation is performed using training symbols, BER is also a function of μ. Under the first three constraints, essentially those from (31), the accumulated burst constructed has more training symbols and also satisfies quasi-static channel requirements. Then, BER is a strict monotonic *decreasing* function of μ.

(iii) Under the last constraint of constant symbol energy, BER is a strict monotonic *increasing* function of q since increasing q decreases the minimum distance between constellation points.

(iv) From (i), (ii), and (iii), a unique solution exists on a bounded domain.

(v) Moreover, to optimally satisfy the fourth BER constraint, μ needs to be as large as possible (for any q). This means that optimization of burst size (which depends on the underlying channel, not on the modulation-mode) can be performed first, followed by the modulation mode search (recall that burst construction deals with channel rate of change, while adaptive modulation addresses the channel quality).

(vi) In other words, a two-layer strategy can be utilized. Once the optimal μ is found as the solution of (31), the

optimal q can then be searched from the given mode choices, producing the largest q that satisfies the BER constraint. □

Remarks

The channel invariance constraint is crucial. Otherwise, if the channel changes between bursts, then increasing the number of training symbols or the modulation mode may or may not improve estimation, depending on the operating channel SNR. In other words, without this constraint, the monotonicity of $\mathrm{BER}(\mu, q)$ may no longer hold. As such, nonunique local maxima may exist on the BER surface over the bounded domain, and the problem would no longer be separable.

Proposition 2. *For each modulation mode q, there is a bijection (one-to-one and onto mapping) between the (pseudo-SNR) channel metric and the BER.*

Proof. This should be quite obvious by construction of any channel metric, because otherwise the constructed metric is not a good metric at all. For the specific case of Γ_{pSNR}, the pseudo-SNR metric, the key is to realize that both Γ_{pSNR} and BER are continuous and strict monotonic decreasing functions of the average channel SNR $\bar{\gamma}$, evident from (42), (44), and (45).

In other words, there exist $\phi, \psi : \mathrm{BER} = \phi(\bar{\gamma}), \Gamma_{\mathrm{pSNR}} = \psi(\bar{\gamma})$, where ϕ, ψ are both bijective (for ϕ, see (44)). Being bijections, ϕ, ψ have bijective inverses: $\bar{\gamma} = \phi^{-1}(\mathrm{BER})$, $\bar{\gamma} = \psi^{-1}(\Gamma_{\mathrm{pSNR}})$. Then, $\Gamma_{\mathrm{pSNR}} = \psi(\phi^{-1}(\mathrm{BER}))$. □

Theoretically, Proposition 2 implies that, when using the channel metric Γ_{pSNR} to maintain the BER constraint in (53), the equivalent condition is $\Gamma_{\mathrm{pSNR}}(\mu, q) \le t_q(\mathrm{BER}_{\max})$, where $t_q(\cdot) = \psi(\phi^{-1}(\cdot))$, for each q. However, note that the above is a purely existential construction, since it is usually difficult to compute the inverses in closed form, for example, computing $\bar{\gamma}$ from BER using (44). Therefore, in practice, the optimal thresholds are usually determined empirically for adaptive modulation [3, 16], as discussed in Section 6.3.

With the above considerations, Algorithm 2 implements a two-layer strategy that iteratively searches for the optimal (μ, q). The switching thresholds (with guaranteed optimal existence by Proposition 2) are empirically approximated and used according to Table 2 for adaptive modulation.

Remarks

Due to the particular forms of the objective and constraints considered here, the optimization can be decoupled as two separate layers. However, this is not always possible. Changing the objective function, for example, addition of delay cost, may necessitate cross-layer optimization. In addition, with more extensive solution spaces (larger $\mathrm{bsize}_{\max}$ and more mode choices), an exhaustive search quickly becomes prohibitively complex due to the combinatorial nature of the mixed-integer problem. For all these cases, suboptimal techniques, such as convexification and relaxation [12], may be applied to reduce complexity.

6.6. Metric errors

It is important to realize that optimality of the above techniques is only guaranteed under ideal situations. In practice, estimation errors lead to constraint violations and therefore suboptimal solutions. In particular, with respect to adaptive modulation, not only can metric errors occur due to insufficient training, delays in transceiver feedback also mean that transmitter mode switching may be too slow.

Algorithm 2 implements closed-loop metric signalling [3], and thus has a minimum latency of one fundamental burst. In other words, even without feedback delay, the metric estimated using the current burst is not used to update the modulation mode until the next transmitted burst, during which time, depending on the Doppler frequency, the channel quality may have changed significantly. In real applications, with feedback delay, the actual latency is even higher. Especially when the channel is changing rapidly, this latency can cause incorrect modes to be invoked by the transmitter receiving outdated metrics.

Under certain conditions, it may be possible to predict the upcoming metrics, thus mitigating the latency effect. Various important considerations in practical implementations of adaptive modulation are surveyed in [3]. In Section 7.5, the effect of latency in the metric estimation will be evaluated by simulation.

7. SIMULATION EXAMPLES

Simulation parameters used are: carrier frequency $f_c = 3\,\mathrm{GHz}$, symbol duration $T_S = 2\,\mu\mathrm{s}$, fundamental burst size = 80 symbols, training density = 10% (i.e., 8 symbols per fundamental burst), normalized data symbols with $\sigma_x^2 = 1$, 4-QAM for fixed-modulation simulations, number of equalizer taps $N = 50$. The power-delay profile is exponential (same shape as Table 1), with delay positions $[0, 4, 6, 7] \times T_S$, so that the channel length $L = 8$.

The maximum accumulated burst size $\mathrm{bsize}_{\max}$ equals 4 fundamental bursts. The threshold function ρ_{th} is defined piece-wise over the SNR-range $\eta \in [0, 40]$ dB:

$$\rho_{\mathrm{th}}(\eta) = \begin{cases} 4\sigma_v^2, & \eta \le 20, \\ 2\sigma_v^2, & 20 < \eta \le 30, \\ \sigma_v^2, & 30 < \eta \le 40, \end{cases} \tag{54}$$

where σ_v^2 is the channel noise variance. This threshold function fulfills the criterion for avoiding potential error floors at high SNR as discussed in Section 5.2: allows larger channel estimation error at low SNR, while forcing smaller estimation error at high SNR.

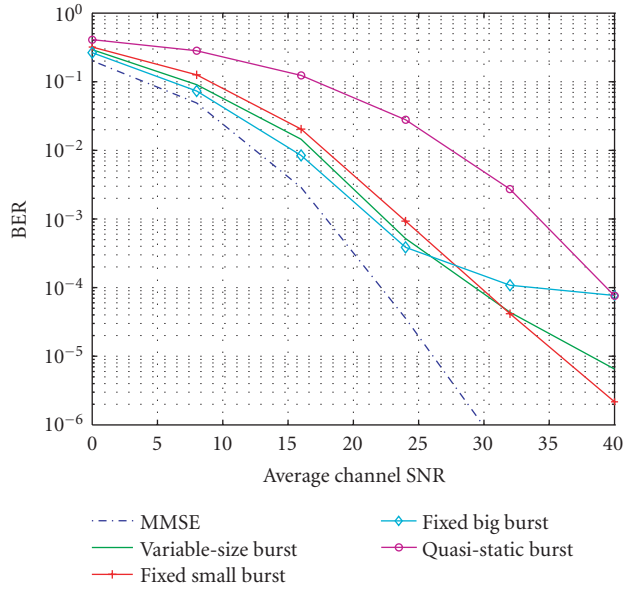


Figure 5: BER performance over fading channel with $f_m T_s = 1 \times 10^{-4}$ or mobile speed $v_m = 18$ km/h.

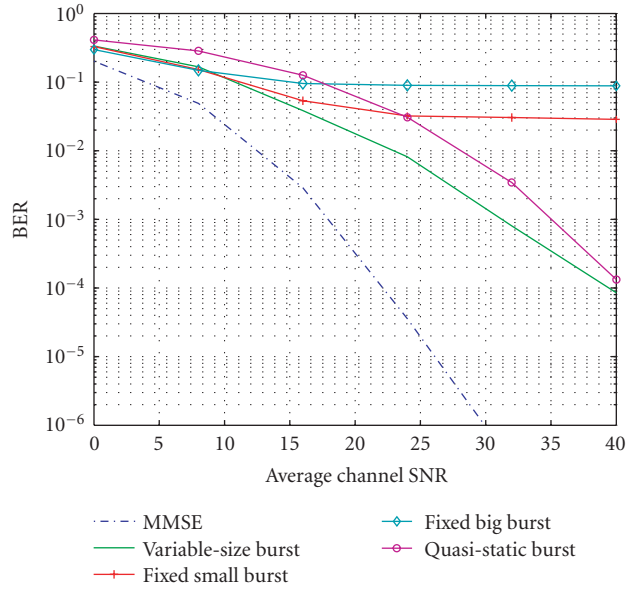


Figure 6: BER performance over fading channel with $f_m T_s = 9 \times 10^{-4}$ or mobile speed $v_m = 162$ km/h.

7.1. Variable-size burst in a slow-fading channel

Here, the channel is characterized by one Doppler state, with $f_m T_s = 1 \times 10^{-4}$ or mobile speed $v_m = 18$ km/h. Figure 5 shows the resulting BER performances for the following schemes.

(1) MMSE

Obtained using a fixed-size burst equal to the fundamental burst, and with *a priori* knowledge of the channel. This is the lower-bound for other cases.

(2) Quasi-static burst

Also obtained using a fixed-size fundamental burst, but with an estimated channel. There is insufficient training for accurate estimation, manifested by a large performance gap from the lower bound.

(3) Fixed small burst

Obtained using a fixed-size burst equal to two fundamental bursts. More training symbols are available compared to the quasi-static burst, resulting in performance improvement.

(4) Fixed big burst

Obtained using a fixed-size burst equal to four fundamental bursts. This scheme approaches the MMSE performance at low SNR, but suffers from an error floor at high SNR due to quasi-static violation being a bottleneck in the absence of noise.

(5) Variable-size burst

Inherits the best characteristics of the previous two fixed-size burst schemes, with good performance at low SNR and no error floor at high SNR.

7.2. Variable-size burst in a fast-fading channel

Here, $f_m T_s = 9 \times 10^{-4}$, corresponding to $v_m = 162$ km/h. Figure 6 shows the resulting performances. Due to construction, the MMSE and quasi-static burst have identical performances as before. In this more rapidly varying scenario, both fixed-size burst schemes suffer from error floors. By contrast, the variable-size burst is able to compensate for the faster channel changes, without being affected by an error floor due to quasi-static violations. Although not as significant as in a slow fading scenario, the variable-size burst still delivers better performance compared to a quasi-static burst.

7.3. Variable-size burst in a two-state fading channel

As described in Section 2.2, the channel here has two Doppler states: a slow state k_1 with $f_m T_s = 1 \times 10^{-4}$, and a fast state k_2 with $f_m T_s = 9 \times 10^{-4}$. In other words, this channel is a combination of the previous two scenarios. The state probabilities are $p(k_1) = 0.8$ and $p(k_2) = 0.2$. This channel is characteristic of a user who spends most of the time in a low-mobility environment, for example, around the $v_m = 18$ km/h range. Figure 7 shows the results.

Although the fast channel state occurs less frequently, it seriously deteriorates the overall performance for the two fixed-size burst schemes, resulting in poor QoS with severe error floors. On the contrary, the variable-size burst delivers

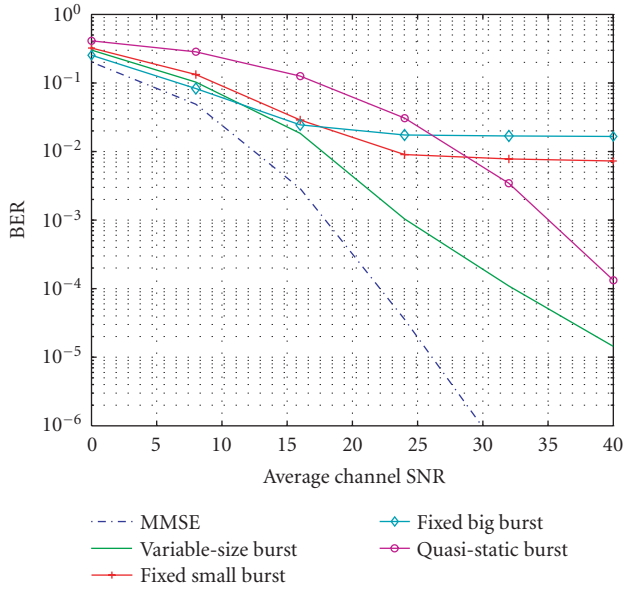


Figure 7: BER performance over fading channel with 2 Doppler states: k_1 with $f_m T_s = 1 \times 10^{-4}$ and k_2 with $f_m T_s = 9 \times 10^{-4}$; the state probabilities are $p(k_1) = 0.8$ and $p(k_2) = 0.2$.

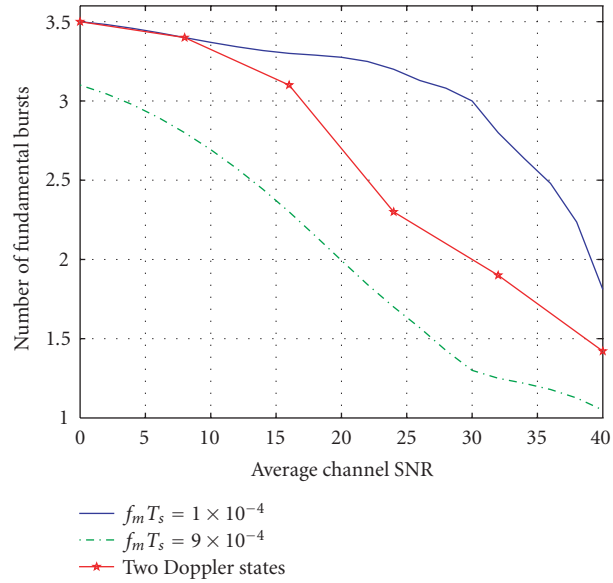


Figure 8: Average burst length (in terms of number of fundamental bursts) of a variable-size burst.

performance gain by exploiting the slower channel state, without being affected by an error floor due to the fast state.

7.4. Average burst length of the variable-size burst

Figure 8 shows the average burst length in the previous channel settings. In a slow fading channel, the burst is closer to the maximum admissible length ($\text{bsize}_{\max} = 4$). But in a fast fading channel, the burst length tends to be shorter in order to satisfy the quasi-static assumption. In a two-state channel, the average burst length is somewhere in between, regulated essentially by the threshold function ρ_{th}.

7.5. Adaptive modulation: BER performance

The previous simulations show that the two fixed-size burst schemes severely fail in a two-state channel, even with fixed modulation. Hence, we will focus on the MMSE, quasi-static and variable-size bursts for adaptive modulation.

The pseudo-SNR metric Γ_{pSNR} is used with thresholds and associated modulation modes summarized in Table 3.

Transmission blocking (no transmission) is invoked for very poor conditions. The highest-throughput mode is 16-QAM, transmitting 4 bits/symbol. To illustrate the effect of metric errors as discussed in Section 6.6, two cases are considered: (i) no feedback delay, resulting in (minimum) latency of 1 burst; (ii) feedback delay of 2 bursts, causing overall latency of 3 bursts. Figure 9 shows the resulting BER performances.

Without feedback delay, the MMSE scheme is able to limit the maximum BER to 10^{-4}, for the SNR range greater than 15 dB. By modifying the thresholds, this range can be changed accordingly, but at the loss of throughput efficiency

Table 3: Switching thresholds for adaptive modulation.

Channel metric (dB)	Modulation mode
$0 \le \Gamma_{\text{pSNR}} < 8$	No transmission
$8 \le \Gamma_{\text{pSNR}} < 12$	BPSK
$12 \le \Gamma_{\text{pSNR}} < 20$	4-QAM
$20 \le \Gamma_{\text{pSNR}} < \infty$	16-QAM

(Figure 10). The obtained results reveal variable-size burst as superior to the fixed-size scheme, guaranteeing a better QoS quantified by the BER.

With delay, the overall QoS is lowered for all cases. This reduction is more noticeable at low SNR since an erroneous metric here implies incorrect invocation of a higher-order mode. By contrast, at high SNR where a higher-order modulation mode is usually already appropriate, an incorrect invocation causes less degradation. And as mentioned in Section 6.6, in certain cases, it may be possible to perform metric prediction to mitigate latency [3].

7.6. Adaptive modulation: throughput performance

A complete comparison of various burst schemes, when using adaptive modulation, also requires examining the corresponding throughputs (number of bits per symbol), depicted in Figure 10.

For throughput, as found in [3], the effect of latency is less significant, with only small performance difference from the ideal case. At low SNR, the MMSE has the lowest throughput. In fact, transmission blocking needs to be the dominant mode here to maintain QoS. Fewer instances of

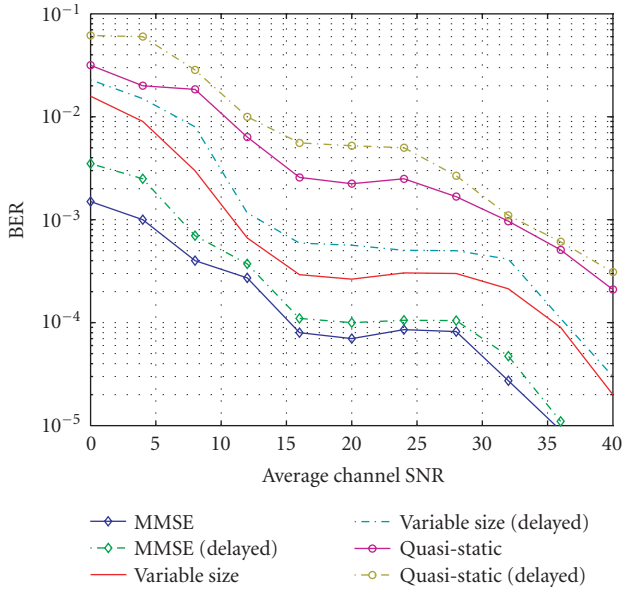


FIGURE 9: Adaptive modulation BER performance over fading channel with 2 Doppler states: k_1 with $f_m T_s = 1 \times 10^{-4}$ and k_2 with $f_m T_s = 9 \times 10^{-4}$; the state probabilities are $p(k_1) = 0.8$ and $p(k_2) = 0.2$.

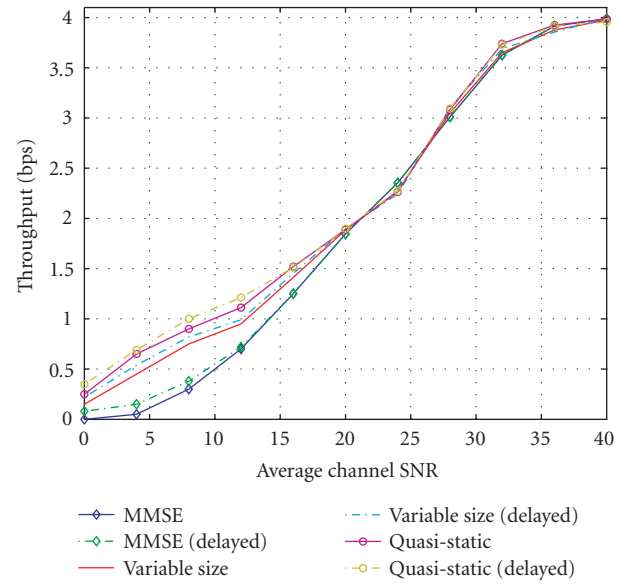


FIGURE 10: Adaptive modulation throughput performance corresponding to Figure 9.

transmission blocking are observed for the variable-size and quasi-static bursts. The reason is that, at low SNR, an accurate channel metric is not available for optimal modulation mode selection. At high SNR, all schemes have nearly identical throughputs, since the estimation of channel metric is more accurate without noise.

The combined BER and throughput performances demonstrate the superiority of a variable-size burst compared to its fixed-size counterpart. It maintains almost identical throughput, but supports much improved QoS.

8. CONCLUDING REMARKS

In this work, two approaches for efficient and reliable communications in time-varying mobile environments are presented: variable-size burst construction and adaptive modulation. It has been shown that, when the underlying time-varying channel is dominated by a slower state, reliable and efficient communication is still possible using a conventional burst-by-burst receiver methodology.

If the channel is dominated by a fast channel state, the variable-size burst performance approaches that of the quasi-static burst, with poor QoS and efficiency. For these scenarios, as mentioned in Section 1, the variable-size burst methodology can be combined with basis-expansion channel models to deliver improved performance at the cost of complexity [6].

ACKNOWLEDGMENTS

The authors thank the anonymous referees whose insightful reviews were instrumental in improving the paper. We are also grateful to the Editor Professor Geert Leus for suggesting important modifications. This work was supported by the Natural Sciences and Engineering Research Council of Canada. Some of the material in this paper was presented at IEEE Globecom 2004, Dallas, Tex.

REFERENCES

[1] T. S. Rappaport, *Wireless Communications: Principles and Practice*, Prentice Hall, Englewood Cliffs, NJ, USA, 1996.

[2] R. Steele and L. Hanzo, *Mobile Radio Communications: Second and Third Generation Cellular and WATM Systems*, John Wiley & Sons, New York, NY, USA, 1999.

[3] L. Hanzo, C. Wong, and M. Yee, *Adaptive Wireless Transceivers: Turbo- Coded, Turbo-Equalized and Space-Time Coded TDMA, CDMA, and OFDM Systems*, John Wiley & Sons, New York, NY, USA, 2002.

[4] F. M. Bui and D. Hatzinakos, "A receiver-based variable-size-burst equalization strategy for spectrally efficient wireless communications," *IEEE Transactions on Signal Processing*, vol. 53, no. 11, pp. 4304–4314, 2005.

[5] F. M. Bui and D. Hatzinakos, "Adaptive modulation using variable-size burst for spectrally efficient interference suppression in wireless communications," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '04)*, vol. 2, pp. 898–902, Dallas, Tex, USA, November-December 2004.

[6] F. M. Bui and D. Hatzinakos, "Identification and tracking of rapidly time-varying mobile channels for improved equalization: a basis-expansion model approach," to appear in *The 5th International Symposium on Communication Systems, Network and Digital Signal Processing (CSNDSP '06)*, Patras, Greece.

[7] G. B. Giannakis and C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 1969–1986, 1998.

[8] I. Barhumi, G. Leus, and M. Moonen, "Time-varying FIR equalization for doubly selective channels," *IEEE Transactions on Wireless Communications*, vol. 4, no. 1, pp. 202–214, 2005.

[9] M. Patzold, *Mobile Fading Channels*, John Wiley & Sons, New York, NY, USA, 1st edition, 2002.

[10] M. Dong and L. Tong, "Optimal design and placement of pilot symbols for channel estimation," *IEEE Transactions on Signal Processing*, vol. 50, no. 12, pp. 3055–3069, 2002.

[11] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, Englewood Cliffs, NJ, USA, 3rd edition, 1996.

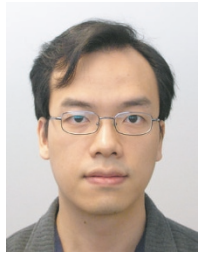
[12] C. A. Floudas, *Nonlinear and Mixed-Integer Optimization: Fundamentals and Applications*, Oxford University Press, New York, NY, USA, 1995.

[13] S. Catreux, V. Erceg, D. Gesbert, and R. W. Heath, "Adaptive modulation and MIMO coding for broadband wireless data networks," *IEEE Communications Magazine*, vol. 40, no. 6, pp. 108–115, 2002.

[14] J. G. Proakis, *Digital Communications*, McGraw Hill, New York, NY, USA, 4th edition, 2001.

[15] P. Balaban and J. Salz, "Optimum diversity combining and equalization in digital data transmission with applications to cellular mobile radio. I: theoretical considerations," *IEEE Transactions on Communications*, vol. 40, no. 5, pp. 885–894, 1992.

[16] C. H. Wong and L. Hanzo, "Upper-bound performance of a wide-band adaptive modem," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 367–369, 2000.



Francis Minhthang Bui received the B.A. degree in French language and the B.S. degree in electrical engineering from the University of Calgary, Alberta, Canada, in 2001. He then completed the M.A.S. degree in electrical engineering in 2003 from the University of Toronto, Ontario, Canada, where he is currently pursuing the Ph.D. degree. His research interests include resource allocation and signal processing methods for power- and spectrum-efficient communications.

Dimitrios Hatzinakos received the Diploma degree from the University of Thessaloniki, Greece, in 1983, the M.A.S. degree from the University of Ottawa, Canada, in 1986, and the Ph.D. degree from Northeastern University, Boston, Mass, in 1990, all in electrical engineering. In September 1990, he joined the Department of Electrical and Computer Engineering, University of Toronto, where now he holds the rank of Professor with tenure. Also, he served as Chair of the Communications Group of the department during the period July 1999 to June 2004. Since November 2004, he is the holder of the Bell Canada Chair in Multimedia at the University of Toronto. His research interests are in the areas of multimedia signal processsing and communications. He is the author/coauthor of more than 150 papers in technical journals and conference proceedings, and he has contributed to 8 books in his areas of interest.