Estimation of Continuous Flat Fading MIMO Channels

Qinfang Sun, *Student, IEEE*, Donald C. Cox, *Fellow, IEEE*, Howard C. Huang, *Member, IEEE*, and Angel Lozano, *Senior Member, IEEE*

***Abstract*—Multiple-input–multiple-output (MIMO) systems can provide high data rate wireless services in a rich scattering environment. In this paper, we study one of the proposals for MIMO systems, the Bell Labs Layered Space-Time (BLAST) architecture. Channel estimation using training sequences is required for coherent detection in BLAST. We apply the maximum-likelihood channel estimator and the optimal training sequences for block flat fading channels to continuous flat fading channels and analyze the estimation error. The optimal training length and training interval that maximize the throughput for a given target bit error-rate are found by computer simulations as functions of the Doppler frequency and the number of antennas.**

***Index Terms*—Bell Labs layered space-time, channel estimation, continuous fading, multiple-input–multiple-output, optimal training parameters, training.**

I. Introduction

MULTIPLE-INPUT–multiple-output (MIMO) systems, e.g., the Bell Laboratories layered space-time (BLAST) wireless communication system invented by Foschini [1], use multielement antenna arrays at both the transmitter and receiver to achieve high spectral efficiency (see Fig. 1). Independent data streams sharing both frequency bands and time slots are transmitted from multiple antennas and jointly detected at the receiver. It has been shown that the theoretical capacity increases linearly with the number of antennas in rich scattering environments [2].

While noncoherent MIMO techniques have been proposed [3], the enormous capacity gains predicted in [1] assume knowledge of the channels at the receiver. Similar gains in the capacity limits are shown in [4] for noncoherent cases, but the result only applies to block fading channels. Different coherent detection algorithms have been proposed for BLAST systems [5], [6]. All of them require knowledge of the channel at the receiver, which is obtained by estimation using training sequences.

The optimal training sequences for the maximum-likelihood (ML) channel estimator for block flat fading MIMO channels are orthogonal sequences [5]. An optimal channel estimator for continuous fading channels should account for the structure of the channel variation using some form of Kalman filtering [7]. However, it is nontrivial to obtain the information on the channel variation and the complexity of Kalman filtering is significantly higher. Furthermore, high-speed data services such as BLAST generally target low mobility users, where the channels are slow fading and resemble block fading channels. Therefore, we apply the simple ML estimator for block fading channels with orthogonal training sequences to continuous flat fading MIMO channels in this paper. The objective is to investigate the performance of this admittedly suboptimal but practically appealing solution.

Manuscript received October 30, 2000; revised July 30. 2001 and December 2, 2001; accepted December 2, 2001. The editor coordinating the review of this paper and approving it for publication is A. F. Molisch.

Q. Sun is with the Atheros Communications, Sunnyvale, CA 94085 USA (e-mail: qfsun@wireless.stanford.edu).

D. C. Cox is with the Department of Electrical Engineering, Stanford University, CA 94305 USA (e-mail: dcox@spark.stanford.edu).

H. C. Huang and A. Lozano are with the Lucent Bell Labs, Crawford Hill, NJ 07733 USA (e-mail: hchuang@lucent.com; aloz@lucent.com).

Digital Object Identifier 10.1109/TWC.2002.804178

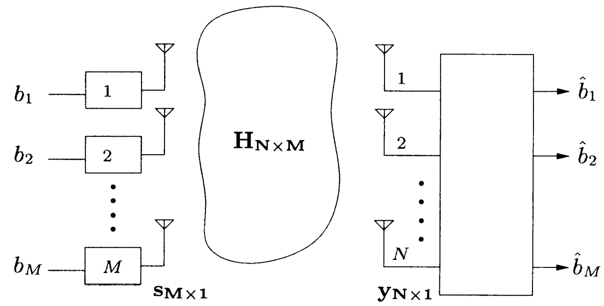

Fig. 1. Block diagram of BLAST system.

We show that the channel estimation error for continuous fading channels is caused by noise as well as the temporal variation of the channel. There are tradeoffs between minimizing the channel estimation error and minimizing the training overhead. We write a vertical BLAST (V-BLAST) simulator to obtain numerical results of the optimal training interval and training length that maximize the throughput within a specific bit error-rate (BER) constraint. V-BLAST is a simplified version of BLAST, where no coding is implemented across different transmitting antennas [6].

This paper is organized as follows. In Section II, we derive the channel estimation error for the continuous flat Rayleigh fading MIMO channels. In Section III, we consider optimal training parameters of the orthogonal training scheme for V-BLAST. We conclude with Section IV.

II. Channel Estimation Error

In this section, we study the training scheme for BLAST and derive the correlation properties of the channel estimation error. We assume narrowband channels in our analysis. Optimal training sequences for broadband channels are discussed in [8].

A. System Model

Consider a communication system with M transmitting antennas and N receiving antennas (see Fig. 1). We use a matrix of complex coefficients to represent the flat fading channel in a baseband equivalent model. The $N \times 1$ signal vector received at time i can be denoted as

$$\mathbf{y}_i = \sqrt{\frac{\rho}{M}}\mathbf{H}_i\mathbf{s}_i + \mathbf{w}_i \tag{1}$$

where $\mathbf{H}_i$ is the $N \times M$ channel matrix, $\mathbf{s}_i$ is the $M \times 1$ transmitted signal vector and $\mathbf{w}_i$ is the $N \times 1$ vector of complex additive white Gaussian noise with zero-mean and unit variance at time instant i. The average power of the components in $\mathbf{H}_i$ and $\mathbf{s}_i$ are normalized to unity, so the average signal-to-noise ratio (SNR) at each receiving antenna is ρ, independent of the number of transmitting antennas (for fair comparison).

During the training phase, training sequences of L_t symbols long are transmitted from all the transmitting antennas. An estimate of the channel, $\hat{\mathbf{H}}$, is obtained at the end. During the payload phase, data sequences of L_d symbols long are transmitted and jointly detected. We define L_t as the *training length* and $L = L_t + L_d$ as the *training interval*. The *duty cycle factor*, $\eta = 1 - L_t/L$, is the fraction of time spent in data transmission.

Since the channel is continuously fading, the actual channel will deviate progressively from the channel estimate obtained at time $i = L_t$. The BER performance will be dominated by the worst channel estimation error. Therefore, we consider $\hat{\mathbf{H}} - \mathbf{H}_L$ as a measure of the channel estimation error, where $\mathbf{H}_L$ is the channel at the end of the training period.[1] Define the difference between the channel at time i and at time L as

$$\Delta\mathbf{H}_i = \mathbf{H}_i - \mathbf{H}_L \tag{2}$$

then, we can rewrite (1) as follows:

$$\mathbf{y}_i = \sqrt{\frac{\rho}{M}}\mathbf{H}_L\mathbf{s}_i + \sqrt{\frac{\rho}{M}}\Delta\mathbf{H}_i\mathbf{s}_i + \mathbf{w}_i. \tag{3}$$

Let $\mathbf{S}$ be the matrix of training symbols, $\mathbf{S} = [\mathbf{s}_1 \quad \mathbf{s}_2 \quad \cdots \quad \mathbf{s}_{L_t}]$, where $\mathbf{s}_i$ for $1 \le i \le L_t$ is the $M \times 1$ training symbol vector at time i. Let the matrix of received signals be $\mathbf{Y} = [\mathbf{y}_1 \quad \mathbf{y}_2 \quad \cdots \quad \mathbf{y}_{L_t}]$ and the matrix of noise be $\mathbf{W} = [\mathbf{w}_1 \quad \mathbf{w}_2 \quad \cdots \quad \mathbf{w}_{L_t}]$. Then

$$\mathbf{Y} = \sqrt{\frac{\rho}{M}}\mathbf{H}_L\mathbf{S} + \mathbf{W} + \sqrt{\frac{\rho}{M}}[\Delta\mathbf{H}_1\mathbf{s}_1 \quad \Delta\mathbf{H}_2\mathbf{s}_2 \quad \cdots \quad \Delta\mathbf{H}_{L_t}\mathbf{s}_{L_t}]. \tag{4}$$

B. Channel Estimation Error

For *block fading* channels where the channel realization remains constant within a block of certain length and then changes to an independent realization for the next block, it has been shown [5] that the ML channel estimator is ($*$ stands for Hermitian transpose)

$$\hat{\mathbf{H}} = \sqrt{\frac{M}{\rho}}\mathbf{Y} \cdot \mathbf{S}^* \cdot (\mathbf{S}\mathbf{S}^*)^{-1} \tag{5}$$

and the optimal training sequences which minimize the mean square estimation error are orthogonal across all transmitting antennas, i.e.,

$$\mathbf{S}\mathbf{S}^* = L_t\mathbf{I}_M \tag{6}$$

where $\mathbf{I}_M$ is the $M \times M$ identity matrix. A necessary condition for the matrix inversion $(\mathbf{S}\mathbf{S}^*)^{-1}$ to exist is $L_t \ge M$.

[1] The result can be applied to any arbitrary reference point with minor changes.

As discussed in Section I, (5), and (6) are suboptimal but practically appealing for continuous fading channels. Apply these to (4), we obtain

$$\hat{\mathbf{H}} = \mathbf{H}_L + \Delta\mathbf{H}_{\text{noise}} + \Delta\mathbf{H}_{\text{Doppler}} \tag{7}$$

where

$$\Delta\mathbf{H}_{\text{noise}} = \frac{1}{L_t}\sqrt{\frac{M}{\rho}}\mathbf{W}\mathbf{S}^* \tag{8}$$

is the estimation error due to noise and

$$\Delta\mathbf{H}_{\text{Doppler}} = \frac{1}{L_t}\sum_{i=1}^{L_t}\Delta\mathbf{H}_i \cdot (\mathbf{s}_i\mathbf{s}_i^*) \tag{9}$$

is the estimation error due to the temporal variation of the channel. It is easy to show that $\Delta\mathbf{H}_{\text{noise}}$ has i.i.d. complex Gaussian entries of zero-mean and variance of $M/(\rho L_t)$. We assume the components of $\mathbf{H}_i$ are uncorrelated with each other (rich scattering) and Rayleigh fading with respect to i. Let $\Delta\mathbf{h}_n^T$ represent the n-th row of $\Delta\mathbf{H}_{\text{Doppler}}$. Using the properties of Rayleigh fading [9], we can show

$$E\left\{\left(\Delta\mathbf{h}_{n_1}^T\right)^* \Delta\mathbf{h}_{n_2}^T\right\} = \delta_{n_1n_2} \cdot \frac{1}{L_t^2}\sum_{i_1=1}^{L_t}\sum_{i_2=1}^{L_t}\mathbf{s}_{i_1}\mathbf{s}_{i_1}^* \cdot [\xi(i_1 - i_2) - \xi(i_1 - L) - \xi(i_2 - L) + 1]\,\mathbf{s}_{i_2}\mathbf{s}_{i_2}^* \tag{10}$$

where δ_{jk} is the discrete Dirac delta function. $\xi(x) = J_0(2\pi f_{d\max}T \cdot x)$, where $J_0(\cdot)$ is the zeroth order Bessel function of the first kind, $f_{d\max}$ is the maximum Doppler frequency and T is the symbol period. For channel estimation without tracking, it is reasonable to assume that the phase change during one training period is small, i.e., $2\pi f_{d\max}TL \ll 1$. Then (10) can be simplified as

$$E\left\{\left(\Delta\mathbf{h}_{n_1}^T\right)^* \Delta\mathbf{h}_{n_2}^T\right\} = \delta_{n_1n_2} \cdot 2\left(\frac{\pi f_{d\max}T}{L_t}\right)^2 \cdot \left(\sum_{i=1}^{L_t}(L - i)\mathbf{s}_i\mathbf{s}_i^*\right)^2 \tag{11}$$

using $J_0(x) \approx 1 - x^2/4$ for small x. The result indicates that the estimation error due to temporal variation increase quadratically with the Doppler frequency. The error also depends on the training length L_t, the training interval L and the training sequences $\mathbf{s}_i$.

For the simple case where there are only one transmitting antenna and one receiving antenna, i.e., $M = N = 1$, the variance of the estimation error in (11) can be computed directly

$$\sigma_{\text{Doppler}}^2 = 2\left[\pi f_{d\max}T\left(L - \frac{L_t + 1}{2}\right)\right]^2. \tag{12}$$

We observe the following.

- If L is fixed and L_t increases, the error decreases.
- If L_t is fixed and L increases, the error increases.
- If both L_t and L increase at a fixed ratio L_t/L, the error increases.

When $M, N > 1$, the expression of the mean square estimation error is generally very complicated and it depends on the

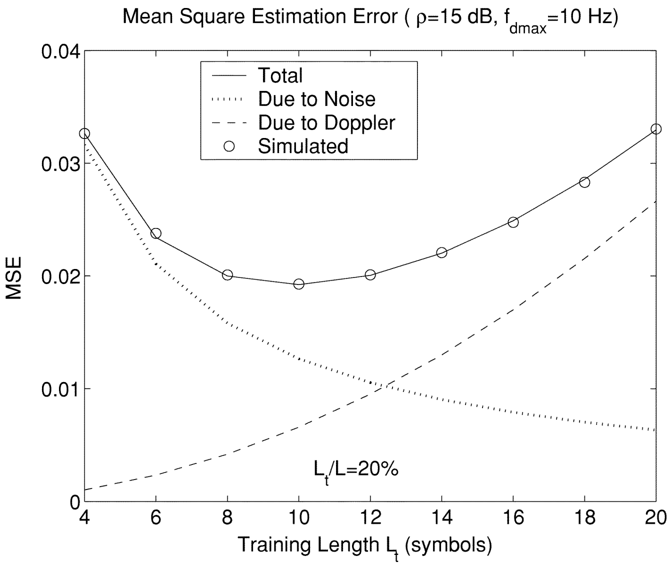

Fig. 2. Channel estimation MSE versus the training length L_t. Four transmitting antennas and four receiving antennas.

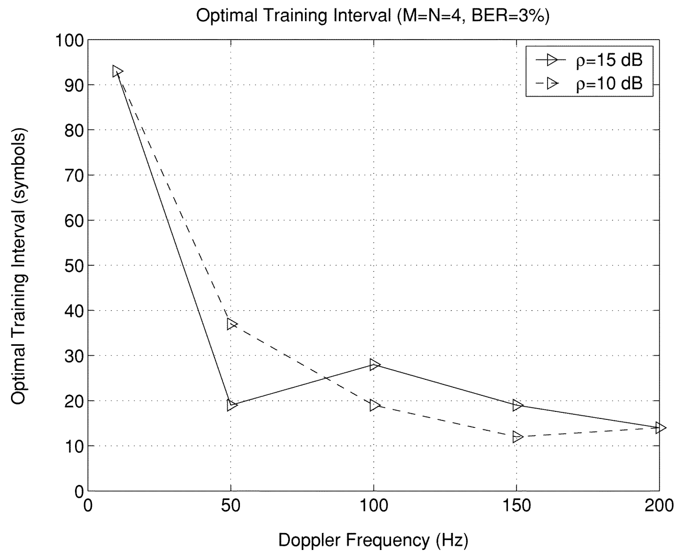

Fig. 3. Optimal training interval versus Doppler frequency.

exact training sequences. A good choice of orthogonal training sequences is the FFT matrix, i.e.

$$\mathbf{S}_{m,i} = e^{-j2\pi(m-1)(i-1)/L_t} \tag{13}$$

where $\mathbf{S}_{m,i}$ is the (m,i)-th component of the training matrix $\mathbf{S}$, $1 \le m \le M$, $1 \le i \le L_t$. It can be shown that with such training sequences, the leading component of the variance is the same as (12). Therefore, the earlier observations also apply to multiple antenna systems.

C. An Example

We use $M = 4$ transmitting antennas and $N = 4$ receiving antennas. The training sequences are the fast Fourier transform (FFT) sequences in (13). The average receiving SNR is $\rho =$ 15 dB. The carrier frequency is $f_c = 2$ GHz and the maximum Doppler frequency is $f_{d\max} = 10$ Hz, which corresponds to a pedestrian speed. The symbol period is $T = 41\ \mu$s, corresponding to the IS-136 standard [10]. The channel coefficients are generated using the Jakes model [11] and continuously fading. Fig. 2 shows the mean square error (MSE) of the channel estimation as a function of the training length L_t. Both L_t and the training interval L increase at a fixed ratio, $L_t/L =$20%. The MSE due to noise decreases with L_t but the MSE due to temporal variation increases with L_t and L. As a result, the overall MSE first decreases and then increases.

We note here that as long as the flat fading model holds, the above results will apply to systems with different symbol period T, if we scale the maximum Doppler frequency $f_{d\max}$ appropriately. This is valid because the estimation error due to temporal variation depends only on the product $f_{d\max}T$ from (10) and (11).

III. Optimal Training Parameters

From the analysis in Section II, when the training interval L is fixed and the training length L_t increases, the channel estimation errors decrease but the duty cycle factor also decreases. When L_t is fixed and L increases, the duty cycle factor increases but the channel estimation error also increases. When both L_t and L increase at a fixed ratio, there is a tradeoff between the two channel estimation error terms. To maximize the data throughput of such a system for a specific target BER, we need to optimize the training interval and training length to balance the channel estimation error and the training overhead.

Analysis of the optimal training length for block fading channels can be found in [5], [12]. The optimal solution greatly depends on the problem definition. When a continuous Rayleigh fading channel is considered, theoretical analysis of optimal training interval and training length is very difficult, as it is difficult to define a meaningful yet workable metric. Therefore, we simulate a system using V-BLAST training and detection algorithms to obtain numerical results of the optimal training parameters.

The simulator generates a continuous Rayleigh fading MIMO channel and performs the channel estimation and data detection repeatedly using orthogonal training sequences and the MMSE V-BLAST detection algorithm [6], [13]. Quadrature amplitude modulation (QAM) is assumed, where the number of bits per symbol is denoted as b. For a given set of system parameters ($f_{d\max}$, M, N, ρ and target BER), we did an exhaustive search for the training interval (L), training length (L_t) and modulation levels (b) that maximize the throughput $R = Mb(1-L_t/L)$. No coding is used and we consider only raw BER. The same carrier frequency, symbol period and channel model as in Section II are used.

We study the optimal training interval, training length and the resulting maximum throughput as functions of the Doppler frequency and the number of antennas. Fluctuations in the results are due to statistical noise, finite step sizes of the exhaustive search and the discrete levels of the QAM modulation.

A. Dependency on Doppler Frequency

Fig. 3 shows the optimal training interval as a function of the Doppler frequency for $M = 4$, $N = 4$, target raw BER of 3% and average SNR levels $\rho =$15 and 10 dB. The target BER is consistent with IS-136 specifications [10]. A Doppler frequency of 10 Hz corresponds to pedestrian speed and 100 Hz corresponds to about 35 mph for the parameters chosen (carrier frequency is 2 GHz). We can see that the optimal training interval

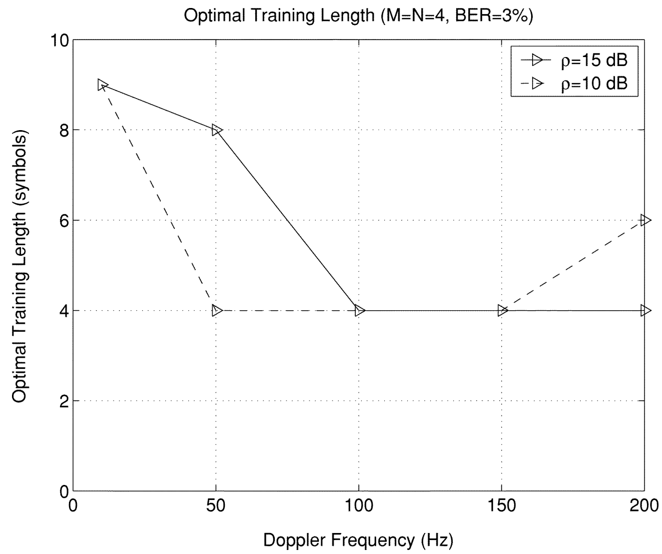

Fig. 4. Optimal training length versus Doppler frequency.

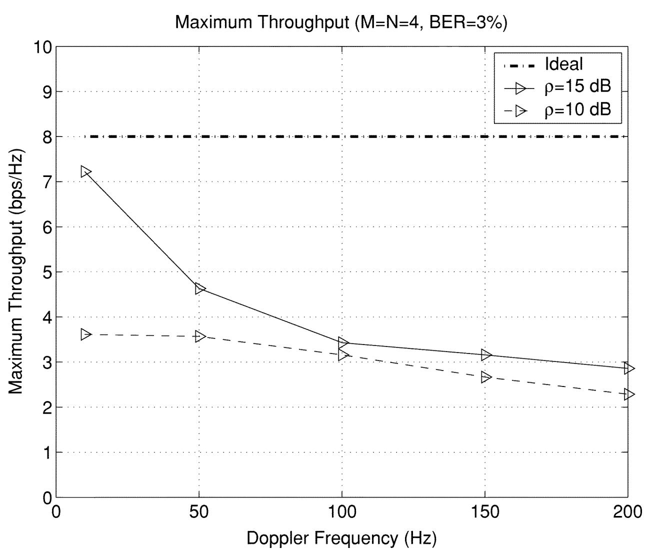

Fig. 5. Maximum throughput versus Doppler frequency resulting from the optimal training interval in Fig. 3 and the optimal training length in Fig. 4. "Ideal" indicates the maximum throughput with ideal channel estimation.

decreases as the Doppler frequency increases. This agrees with intuition that when the channel changes rapidly, training should be done more often. The two curves cross over since there is a modulation level change from quadrature phase shift keying (QPSK) to binary phase-shift keying (BPSK) for $\rho = 15$ dB from 50 to 100 Hz.

Fig. 4 shows the corresponding optimal training length. As the Doppler frequency increases, the optimal training length converges to the number of transmitting antennas, the minimum requirement for orthogonal training. This agrees with the results in Fig. 3. Since the training interval decreases with Doppler frequency, the training length should decrease as well to maintain a reasonable duty cycle factor. There is an increase in the training length at 200 Hz on the 10-dB curve in Fig. 4. This results from the increase at the same location in Fig. 3 caused by the finite step size in the exhaustive search. Since the training interval is slightly increased there, a longer training length can be afforded

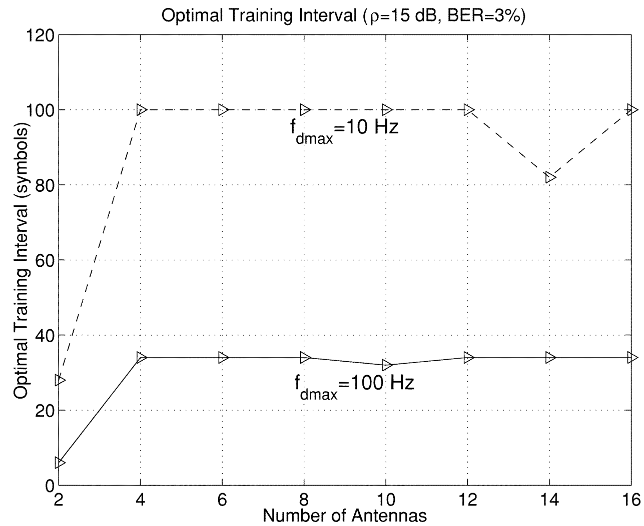

Fig. 6. Optimal training interval versus the number of antennas.

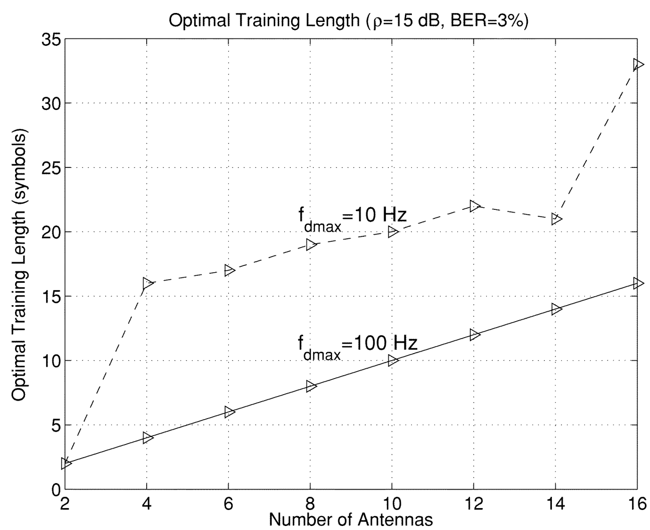

Fig. 7. Optimal training length versus the number of antennas.

to reduce the channel estimation error with little loss in the duty cycle factor.

Fig. 5 shows the maximum throughput esulting from the optimal training interval and training length shown in Figs. 3 and 4. The throughput decreases when the Doppler frequency increases. This suggests that tracking may be needed for communications with high-speed mobiles. The gap between the two curves is roughly the same except for the first two Doppler frequency points, where the achievable modulation levels are different (QPSK versus BPSK). As a comparison, we also include the maximum throughput for ideal channel estimation where the channel is perfectly known at the receiver. The ideal maximum throughput is the same for $\rho = 15$ dB and $\rho = 10$ dB due to the discrete levels of QAM modulation. The performance degrades significantly as a result of training overhead and channel estimation error.

B. Dependency on Number of Antennas

We also study the optimal training interval and training length as functions of the number of antennas. For simplicity, we let $M = N$. Figs. 6–8 show the results for a system with

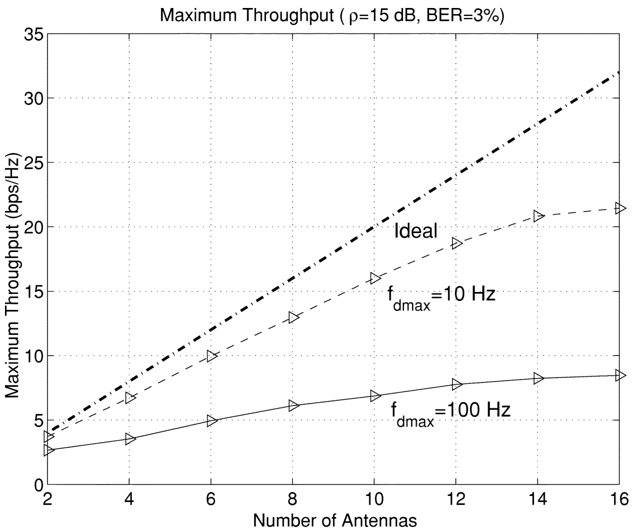

Fig. 8. Maximum throughput versus the number of antennas resulting from the optimal training interval in Fig. 6 and the optimal training length in Fig. 7. "Ideal" indicates the throughput with ideal channel estimation.

ρ =15 dB and 3% raw target BER. Two Doppler frequencies, 10 and 100 Hz are compared. It can be seen that as the number of antennas increases, the optimal training interval increases first and then saturates. The optimal training length is proportional to the number of antennas in both cases and equals the number of antennas for the 100 Hz case. A dip in the 10-Hz curve at $M = 14$ in both Figs. 6 and 7 is caused by statistical error and the finite step sizes of the exhaustive search. The throughput saturates as the number of antennas increases, upper-limited by the Doppler frequency. Without tracking, the throughput is limited to about 21.4 b/s/Hz and 8.5 bp/s/Hz for Doppler frequencies of 10 and 100 Hz, respectively, with 16 antennas on both sides and a 15-dB average SNR. The modulation schemes used are only up to QPSK. The maximum throughput with ideal channel estimation is also included for comparison. Again, the actual performance degrades as a result of training overhead and channel estimation error.

In summary, the simulations show that the required training length is proportional and converges to the number of antennas when the maximum Doppler frequency increases. The optimal training interval and maximum throughput both decrease when the maximum Doppler frequency increases. They saturate as the number of antennas increases for a given maximum Doppler frequency.

IV. Conclusion

Estimation of continuous Rayleigh fading MIMO channels using orthogonal training sequences is subject to errors due to two effects: noise and temporal variation of the channel. The error due to noise is inversely proportional to the number of training symbols, while the error due to temporal variation is proportional to the square of the maximum Doppler frequency and depends on the training length and training interval.

Previous work [5], [12] on optimal training interval and length are based on estimation error only due to noise. Numerical results obtained from our V-BLAST simulations suggest that the required training length is proportional to the number of transmitting antennas; when the maximum Doppler frequency increases, the optimal training length converges to the number of transmitting antennas. The optimal training interval decreases as the maximum Doppler frequency increases; for a given maximum Doppler frequency, it first increases then saturates as the number of antennas increases. The maximum throughput decreases as the maximum Doppler frequency increases; for a given maximum Doppler frequency, it first increases and then saturates as the number of antennas increases.

The throughput is significantly reduced from that with ideal channel estimation as a result of channel estimation errors, especially for high Doppler frequencies. This suggests that advanced channel estimation and tracking algorithms are needed to provide satisfactory performance for high-speed mobiles.

Acknowledgment

The authors would like to thank the anonymous reviewers for the comments and suggestions that helped to improve the quality of this paper.

References

[1] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Tech. J.*, vol. 1, Autumn 1996.
[2] I. E. Telatar, "Capacity of multiantenna Gaussian channels," *Euro. Trans. Telecommun.*, vol. 10, no. 6, Nov./Dec. 1999.
[3] B. M. Hochwald and T. L. Marzetta, "Space-time modulation for unknown fading," in *Proc. SPIE AeroSense Conf.*, Orlando, FL, Apr. 1999.
[4] L. Zheng and D. N. C. Tse, "Information theoretic limits for non coherent multiantenna communications," in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, vol. 1, Sept. 2000.
[5] T. L. Marzetta, "BLAST training: Estimating channel characteristics for high capacity space-time wireless," in *Proc. 37th Annual Allerton Conf. Communication, Control and Computing*, Monticello, IL, 1999.
[6] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architectures," *Electron. Lett.*, vol. 35, pp. 14–15, Jan. 1999.
[7] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Englewood Cliffs, NJ: Prentice-Hall, 2000.
[8] J. Balakrishnan, M. Rupp, and H. Viswanathan, "Optimal channel training for multiple antenna systems," in *Proc. Conf. Multiple Access, Mobility and Teletraffic for Wireless Communications (MMT)*, Duck Key, FL, Dec. 2000.
[9] G. L. Stüber, *Principles of Mobile Communication*, 2nd ed. Norwell, MA: Kluwer , 2001.
[10] "TDMA Cellular PCS," TIA/EIA-136, TR-45.3, 1999.
[11] W. C. Jakes, *Microwave Mobile Communications*. Piscataway, NJ: IEEE Press Reprints, 1994.
[12] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?," Bell Laboratories, Lucent Technologies, Apr. 2000.
[13] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.