Communicating Over Nonstationary Nonflat Wireless Channels

Karin Sigloch, *Member, IEEE*, Michael R. Andrews, Partha P. Mitra, and David J. Thomson, *Fellow, IEEE*

***Abstract*—We develop the concept of joint time-frequency estimation of wireless channels. The motivation is to optimize channel usage by increasing the signal-to-noise ratio (SNR) after demodulation while keeping training overhead at a moderate level. This issue is important for single-input single-output (SISO) and multiple-input multiple-output (MIMO) systems but particularly so for the latter. Linear operators offer a general mathematical framework for symbol modulation in channels that vary both temporally and spectrally within the duration and bandwidth of one symbol. In particular, we present a channel model that assumes first-order temporal and spectral fluctuations within one symbol or symbol block. Discrete prolate spheroidal sequences (Slepian sequences) are used as pulse-shaping functions. The channel operator in the Slepian basis is almost tridiagonal, and the simple intersymbol interference pattern can be exploited for efficient and fast decoding using Viterbi's algorithm. To prove the concept, we use the acoustic channel as a meaningful physical analogy to the radio channel. In acoustic 2×2 MIMO experiments, our method produced estimation results that are superior to first-order time-only, frequency-only, and zeroth-order models by 7.0, 9.4, and 11.6 db. In computer simulations of cellular wireless channels with realistic temporal and spectral fluctuations, time-frequency estimation gains us 12 to 18 db over constant-only estimation in terms of received SNR when signal-to-receiver-noise is 10 to 20 db. The bit error rate (BER) decreases by a factor of two for a binary constellation.**

***Index Terms*—Channel estimation, discrete prolate spheroidal sequences, MIMO, modulation, nonflat, nonstationary, Slepian, time-frequency.**

I. Motivation

We propose a new modulation scheme for wireless channels that vary both temporally and spectrally. It is a computationally feasible method that can estimate temporal and spectral variations within one symbol while using only a modest number of extra training sequences. Generally, one can expect to raise the signal-to-noise ratio (SNR) after demodulation by applying more realistic modeling to the physical layer. We suggest that systems that vary over both time and frequency (such as wireless channels) can be approximated most efficiently with models that allow for variations in both of those dimensions. By most efficient, we mean that the overall number of parameters (training sequences or "pilots" in wireless) needed to achieve a certain estimation accuracy or SNR is smaller than if one uses models that allow for variation of either only time, only frequency, or neither. This concept is related to linear regression when a general, continuous time series is to be approximated locally by polynomials of low order. Successively higher order polynomials use more parameters per segment but can compensate by approximating the time series more accurately over longer segments. Fig. 1 illustrates the principle for the two simplest functions: zeroth- and first-order polynomials. The overall number of parameters is the same in both cases: In the lower plot, we estimate one constant coefficient per interval Δ; in the top, a constant and a time slope coefficient need to be estimated every 2Δ. Since $h(t)$ varies smoothly over time, the piecewise linear model takes time variation into account and yields a smaller total error. In wireless communications, the polynomial function for a given segment corresponds to the channel model, the accuracy corresponds to errors in fitting the channel, and the number of parameters to be estimated is directly proportional to the number of pilot symbols. The curve in Fig. 1 can be thought of as a temporally varying wireless channel. Since wireless channels also vary in frequency, we should add a frequency axis perpendicular to the two existing axes. Channel variation would

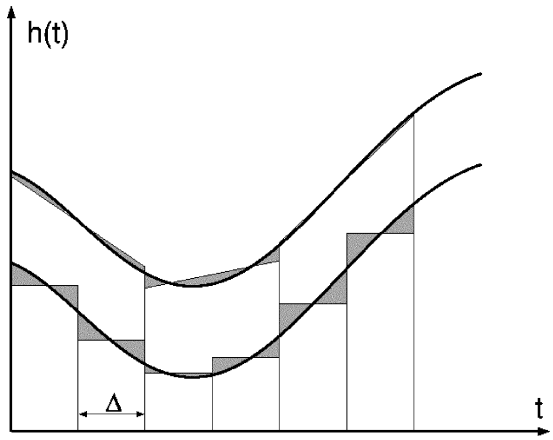

Fig. 1. Local regression of a general, continuous time series using (bottom) zeroth-order polynomials or (top) first-order polynomials. Approximation errors are shaded. Note that globally, the same number of parameters are needed in each case: one per interval Δ for constant segment approximation and two per 2Δ (namely, constant part and slope) for first-order polynomials. The second method results in smaller errors because it takes into account the existing temporal fluctuations.

Manuscript received July 25, 2003; revised May 6, 2004. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Tulay Adali.

K. Sigloch was with the Bell Laboratories, Lucent Technologies, Murray Hill, NJ. She is now with Princeton University, Princeton, NJ 08840 USA (e-mail: sigloch@princeton.edu).

M. R. Andrews was with the Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974-0636 USA. He is now with the Flarion Technologies, Bedminster, NJ 07921 USA (e-mail: mikea@xoba.com).

P. P. Mitra was with the Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974-0636 USA. He is now with the Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724 USA (e-mail: mitra@cshl.edu).

D. J. Thomson was with the Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974-0636 USA. He is now with the Queen's University, Kingston, ON, K7L 3N6 Canada (e-mail: djt@mast.queensu.ca).

Digital Object Identifier 10.1109/TSP.2005.847849

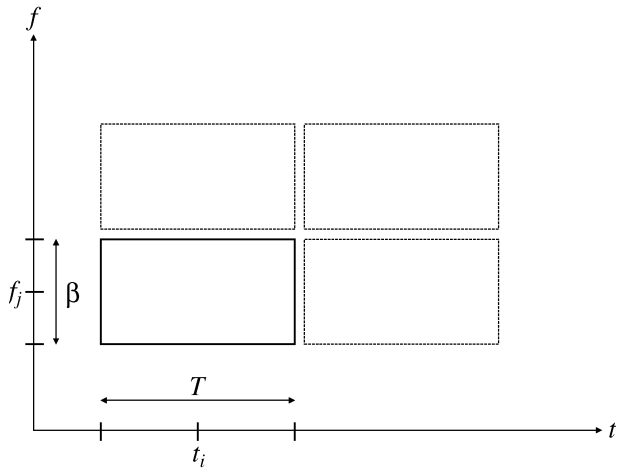

Fig. 2. Dividing up the global time-bandwidth resource into blocks of duration T and bandwidth B; channel model parameters need to be estimated for every block. If the channel varies both temporally and spectrally, then both kinds of fluctuations should be included in the modeling to maximize the SNR after demodulation and to keep the training overhead low.

be drawn as a hyperplane $h(t, f)$ over the time-frequency plane and would be approximated locally by hyperplane segments of duration T and bandwidth B. These blocks of duration T and bandwidth B are tiled to cover the entire time-frequency plane, as shown in Fig. 2, and model function parameters are re-estimated for each block.

In current systems, channel models are either piecewise constant in a block (orthogonal frequency division multiplexing (OFDM) or code division multiple access (CDMA) without equalizing) or "frequency-only" models, i.e., time-invariant transfer function approaches (CDMA or time division multiple acccess (TDMA) with tap equalizing). The more general class of pertinent channel models (namely, temporally *and* spectrally varying functions) has not been exploited so far, even though it would be desirable from the viewpoint of resource efficiency. Current simplistic channel models require an unduly high proportion of symbols to be sacrificed for pilots since time-bandwidth blocks need to be chosen very small for the flat or stationary channel assumption to hold. With true time-frequency models, one could use larger blocks and a smaller overall number of pilots because the model captures more of the actual channel. In other words, it is desirable to account for distortions by explicitly modeling them at the lowest possible level: the level of the pulse-shaping functions. If one allows distortions to propagate up to the decoding level—as simplistic models do—then more redundancy in form of pilots needs to be added in the first place. A recent suggestion to overcome the limitations of flat and stationary models is given by Kuroda [9], where the channel is allowed to have both temporal and spectral modulation over a larger time-bandwidth region. However, this is achieved by interpolating between smaller subblocks over which the channel is flat and stationary. The full efficiency benefits of time-frequency modulation cannot be realized in this way.

To our knowledge, the present study is the first approach that allows for simultaneous temporal and spectral channel variation *within the smallest, or elementary, time-frequency block* relevant to the scheme. We explore a model that allows for first-order variations both in time and bandwidth. In other words, we estimate a constant part, a "time slope coefficient" and a "frequency slope coefficient" for every signaling block. Pilot and data symbols are carried by mutually orthogonal pulse-shaping functions, all of which are spread over the entire block duration and bandwidth.

In any modulation scheme, it is crucial to limit inter-symbol interference (ISI) to a computationally acceptable level. We show that time-frequency channel modeling need not imply a vastly increased complexity in decoding. Our choice of Slepian sequences as pulse-shaping functions limits ISI to only three neighboring symbols while minimizing out-of-block leakage.

The mathematical background of channel parameterization with linear matrix operators is discussed in Section II. We wanted to test our scheme on a real physical context and chose to do acoustic experiments. Section III explains why the inexpensive and easy-to-handle acoustic testbed provides a physically meaningful analogy to the wireless radio channel that justifies its use for the proof of concept that we attempt. Our idea can be used for both single-input single-output (SISO) and multiple-input multiple-output (MIMO) channels. We did both kinds of experiments but choose to present 2×2 MIMO results since space does not allow us to present both. We think that time-frequency estimation may turn out to be most useful in MIMO contexts because of the higher pilot requirements when estimating transmission matrices. Section IV presents computer simulations of cellular wireless channels that vary over time and frequency. This paper is exploratory. We do not mean to provide a conclusive study about decoding algorithms and bit error rates (BERs) for a particular communication protocol. Rather, this is a new and generic idea for improving SNR on the physical layer of wireless systems, enabling more sophisticated coding to build on top of that.

II. Mathematical Background

A. Channel Parameterization

The general theory of propagation in a cellular environment has been given in detail by Jakes [8], Proakis [12], and elsewhere. The *time-variant impulse response* $h(t,\tau)$ as the most general linear channel model for the single-antenna case and in baseband notation is given by the relation between the transmitted waveform $x(t)$, the received waveform $y(t)$, and receiver noise $\eta(t)$

$$y_r(t) = y(t) + \eta(t) = \int_{\tau=0}^{\infty} h(t,\tau)x(t-\tau)d\tau + \eta(t). \quad (1)$$

This model considers the channel to consist of a continuum of multipath components. τ are the delays of an infinite number of echoes at the receiver ("continuous scattering"). Their attenuation $h(t,\tau)$ varies over time t due to Doppler effects, and it also varies with τ, due to wavelength-dependent interference patterns at the receiver.

Since overall channel fluctuation as a function of absolute time t and delay τ appears stochastic, the channel is approximated by local linear fitting of a relatively simple model func-

tion to each of the time-bandwidth blocks in Fig. 2. Since we are working with signals sampled at a rate of $S = 1/\Delta$, we need to discretize our channel model. Let times and delays be discrete ($t_j = j\Delta$ and $(t-\tau)_i = i\Delta$) in order to obtain transmitted and received time series $y_j = y(j\Delta)$ and $x_i = x(i\Delta)$, where $i, j = 1, \ldots, n$, and $n = TS$. Then, channel (1) for one block becomes

$$y_j = \sum_{i=1}^{n} \Delta h\left(j\Delta, (j-i)\Delta\right) x_i$$

or in matrix notation

$$\mathbf{y} = \mathbf{H}\mathbf{x} \tag{2}$$

where $\mathbf{x} = [x_i]$, $\mathbf{y} = [y_j]$, and the $n \times n$ elements of the channel matrix $\mathbf{H}$ are $H_{ji} = \Delta \cdot h(j\Delta, (j-i)\Delta)$. Note that $\mathbf{H}$ is a Toeplitz matrix for the special case of time-translational invariant ("frequency-only") channels $h(t,\tau) = h(t-\tau)$ and is diagonal for flat-fading channels $h(t,\tau) = h_0 \cdot \delta(t-\tau)$.

To open the field of time-frequency modeling, we propose to work with a model function that features temporal and spectral variations of first order. Proceeding from conventional practice, one may begin by ignoring temporal and spectral variations and assuming the channel to be constant during short periods of time T and flat over small bandwidths B. Thus, the channel model in that case would be a constant complex attenuation H^c over time: $y_j^c = H^c x_j$, or, in matrix notation

$$\mathbf{y^c} = H^c \mathbf{C}\mathbf{x}.$$

$\mathbf{C}$ is simply the $n \times n$ identity matrix, and we call it the "constant operator" because it describes how a constant channel operates on the transmitted $\mathbf{x}$.

A natural extension to this simplistic model is to assume that the channel varied linearly over time, and thus, one would use a model with both constant *and* linear parts as functions of time: $y_j = y_j^c + y_j^t = H^c x_j + H^t j\Delta x_j$. In matrix notation, the temporally varying part is

$$\mathbf{y^t} = H^t \mathbf{T}\mathbf{x}$$

where H^t is the complex constant "time slope coefficient" and $\mathbf{T} = [T_{ji}] = [j\Delta \cdot \delta_{ji}]$ is called "time operator" because this diagonal $n \times n$ matrix multiplies $\mathbf{x}$ by the linearly sampled time $j\Delta = -T/2, \ldots, T/2$ of a block.

The next extension would be to proceed similarly for variations in bandwidth and assume a linear variation in the channel over the band; a linear variation over frequency is simply handled in the Fourier domain, just as a linear variation over time is handled in the temporal domain. The $n \times n$ Discrete Fourier Transform (DFT) matrix $\mathbf{D}$ and its inverse (IDFT) matrix $\mathbf{D^\dagger}$ are

$$\begin{aligned} \mathbf{D} &= [D_{kj}] = \left[\frac{1}{\sqrt{n}} \exp\left(-2\pi i \frac{jk}{n}\right)\right] \\ \mathbf{D^\dagger} &= [D_{jk}] = \left[\sqrt{n} \exp\left(+2\pi i \frac{jk}{n}\right)\right]. \end{aligned} \tag{3}$$

Therefore, $\mathbf{X} = \mathbf{D}\mathbf{x}$ and $\mathbf{Y^f} = \mathbf{D}\mathbf{y^f}$ are the discrete Fourier transforms of $\mathbf{x}$ and $\mathbf{y^f}$, the latter being the part of the received signal that is due to spectral channel variation. The relation of the elements of $\mathbf{X} = [X_k]$ and $\mathbf{Y^f} = [Y_k^f]$ is defined as being $Y_k^f = H^f \cdot k\nu \cdot X_k$. In matrix notation

$$\mathbf{D}\mathbf{y^f} = H^f \boldsymbol{F}\mathbf{D}\mathbf{x}$$

H^f is the complex constant "frequency slope coefficient", and $\nu = B/n$ is the spectral resolution of the block. The diagonal matrix $F = [F_{kl}] = [k\nu \cdot \delta_{kl}]$ is called the "frequency operator in the Fourier basis" since it linearly multiplies X_k by the discrete frequencies $k\nu = -B/2, \ldots, B/2$ of the band. In the time domain, the received signal due to first-order spectral variation is then

$$\mathbf{y^f} = \mathbf{F}\mathbf{x} = H^f \mathbf{D^\dagger}\boldsymbol{F}\mathbf{D}\mathbf{x}.$$

The right- and left-multiplication of F with the DFT/IDFT matrices describes the rotation of the frequency operator from the frequency basis to the time basis. $\mathbf{F} = \mathbf{D^\dagger}\boldsymbol{F}\mathbf{D}$ is therefore the frequency operator in the time basis. Pulling everything together, our first-order time-frequency channel model is

$$\begin{aligned} \mathbf{y} &= \mathbf{H}\mathbf{x} \\ &= \mathbf{y^c} + \mathbf{y^t} + \mathbf{y^f} \\ &= H^c\mathbf{C}\mathbf{x} + H^t\mathbf{T}\mathbf{x} + H^f\mathbf{F}\mathbf{x} \end{aligned} \tag{4}$$

which is essentially a first-order Taylor expansion of the channel in both time and frequency and, thus, valid over small T and B. The frequency operator $\mathbf{F}$ and, thus, the whole channel operator $\mathbf{H}$, do not have a simple structure in the time basis. Elements of considerable magnitude far off the diagonals indicate energy spilling (ISI) over a wide range of the transmitted time series.

The discrete SISO channel (4) can be easily generalized to the MIMO case with N transmit and N receive antennas. We introduce the notation here and carry it through the rest of this paper. However, the two new antenna indices (α, β) make the notation somewhat cumbersome. *Everything that follows can be read without loss of understanding by considering the SISO case, which means ignoring all indices α, β and letting $N = 1$.* We suggest to do so for a first reading. In the MIMO case, (4) becomes

$$\begin{aligned} \mathbf{y}_\beta &= \sum_{\alpha=1}^{N} \mathbf{H}_{\beta\alpha}\mathbf{x}_{\beta\alpha} \\ &= \sum_{\alpha=1}^{N} H^c_{\alpha\beta}\mathbf{C}\mathbf{x}_\alpha + H^t_{\alpha\beta}\mathbf{T}\mathbf{x}_\alpha + H^f_{\alpha\beta}\mathbf{F}\mathbf{x}_\alpha \end{aligned} \tag{5}$$

$\alpha, \beta = 1, \ldots, N$ are the transmit and receive antenna indices, and the channel $\mathbf{H}_{\beta\alpha}$ is a rank 4 tensor $(n \times n \times N \times N)$ because we need to describe $N \times N$ transmission paths for every instant in time. Constant, time, and frequency operators are the same as in the SISO case. The constant attenuation $H^c_{\beta\alpha}$, time slope coefficients $H^t_{\beta\alpha}$, and frequency slope coefficients $H^f_{\beta\alpha}$ are now elements of three $N \times N$ matrices, with three complex numbers characterizing each propagation path $\beta\alpha$.

B. Slepian Basis

Our goal is to devise a computationally feasible decoding algorithm. This is the motivation behind rotating (5) to a more parsimonious basis: the basis of a carefully chosen set of pulse-shaping functions. For the first-order time-frequency model, the

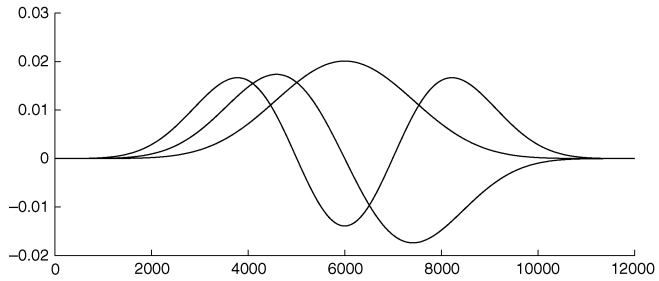

Fig. 3. First three discrete prolate spheroidal sequences, also called Slepian sequences ($T = 2.4$ s, $B = 5$ Hz, $K = 11, n = 12\,000$).

Slepian basis is a suitable basis. Its basis functions are the *discrete prolate spheroidal sequences* (*dpss*) or *Slepian sequences*, in honor of their main proponent in signal processing: D. Slepian [14], [15]. The dpss are a set of orthonormal sequences of length $n = TS$, with the defining property that they are the set of orthonormal sequences that is maximally concentrated in energy within the chosen bandwidth B (S is the sampling rate). For details, see [14] and [15]; a method for efficient computation of the dpss is given in [17, App.]. Fig. 3 shows a plot of the first three dpss for parameters that are typical of our acoustic experiments. We can compute n orthogonal Slepian functions, but after the first $K = TB - 1$ functions, there is a sharp drop in the relative amount of energy concentrated in the band B. The first K functions together account for more than 95% of the energy contained in B. Loosely speaking, K is the number of orthonormal functions that can be accommodated by a time-frequency block of duration T and bandwidth B, which is expected from the uncertainty principle of time and frequency (which is derived, for example, in [11]).[1]

C. Channel Operators in the Slepian Basis

The time operator $\mathbf{T}$ and the frequency operator $\mathbf{F}$ both show the same kind of sparseness in the Slepian basis; only the upper and the lower diagonals contain nonzero elements. It is this time-frequency symmetry of the dpss that we will be exploiting for quick decoding. In the context of spectral analysis, the particular shape of the two operators was originally developed for and applied to the analysis of nonstationary time series (see [18] and [17]; see also [19] for a summary). They give estimates for the time and frequency derivatives of a spectrum. One way of understanding the origin of the sparseness and the symmetry is to realize that the dpss are closely related to the eigenfunctions of the harmonic oscillator in quantum mechanics, which is a problem that has been studied extensively (e.g. [3]). These eigenfunctions describe the location probability of a particle trapped in a parabolic potential well. It is intuitively plausible that the solutions to this kind of concentration problem would be similar to the solutions of maximally concentrating orthogonal functions in time. The difference is that the dpss are *strictly*

[1]It is somewhat arbitrary to choose $K = TB - 1$ instead of $K = TB$. The (TB)th sequence is the last one that still has most of its energy within B, but because of its considerable energy spilling of about 30%, we choose to exclude it. Another manifestation of the uncertainty principle is Shannon's sampling theorem, according to which a signal of bandwidth $B = 2W$ (positive and negative frequencies) needs to be sampled $2W$ times per second to be completely specified: The signal has $2WT$ degrees of freedom.

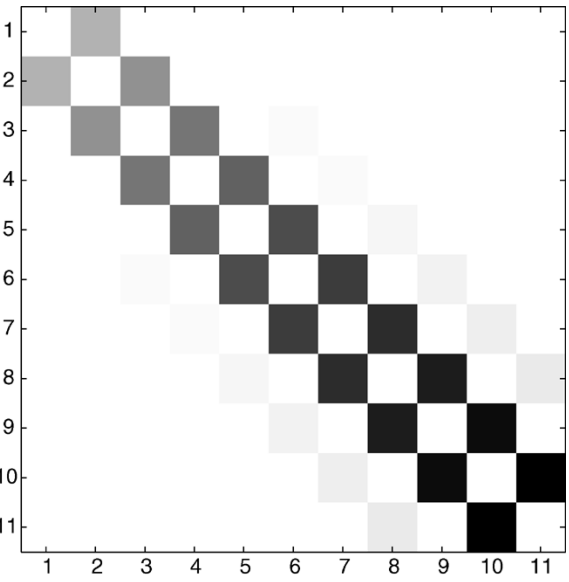

Fig. 4. Magnitude of elements of the 11×11 time operator $\mathbf{T}$ in the Slepian basis. A plot of the frequency operator $\mathbf{F}$ looks almost identical. Both matrices are practically bidiagonal, but the time operator is real and symmetric, whereas the frequency operator is purely imaginary and skew-symmetric. This difference allows us to distinguish between time variation and frequency variation in the channel.

concentrated between $-T/2$ and $T/2$, whereas the harmonic oscillator eigenfunctions extend to $\pm\infty$, although they are exponentially damped. The oscillator eigenfunctions are products of Hermite polynomials with symmetric Gaussian functions. Hermite polynomials $u_k(x)$ can be computed using a simple recursion relation

$$u_{k+1}(x) = 2xu_k(x) - 2ku_{k-1}(x)$$

and therefore, the $(k+1)$th eigenfunction can be calculated if the $(k-1)$th and kth ones are known. This recursion relation means that the position operator $\mathbf{x}$ is bidiagonal, i.e., all nonzero elements are found on the first upper and lower diagonals of the operator matrix, which is easily seen by rewriting the equation as

$$xu_k(x) = \frac{1}{2}u_{k+1}(x) + ku_{k-1}(x).$$

Since the position operator $\mathbf{x}$ directly corresponds to the time operator $\mathbf{T}$ in our time-frequency problem, it would be interesting to know if $\mathbf{T}$ is also bidiagonal in the Slepian basis. The answer is "almost." Since the dpss are truncated versions of the oscillator eigenfunctions, the above recursion relation does not carry over exactly, and the time operator has nonzero elements off the two bidiagonals. However, if only the first K dpss are chosen as basis functions, the magnitudes of the elements off the bidiagonals are orders of magnitude smaller than those on the bidiagonals. For all practical purposes, the $K \times K$ time operator in the Slepian basis can therefore be considered to be bidiagonal. A time operator we computed for our acoustic experiments $(K = 11)$ is shown in Fig. 4. Elements off the two bidiagonals are just barely visible or are not visible at all; they account for only about 10^{-4} of the total energy.

The symmetry is due to another interesting property of the harmonic oscillator eigenfunctions: They are their own Fourier transforms (with appropriate scale changes). The same holds

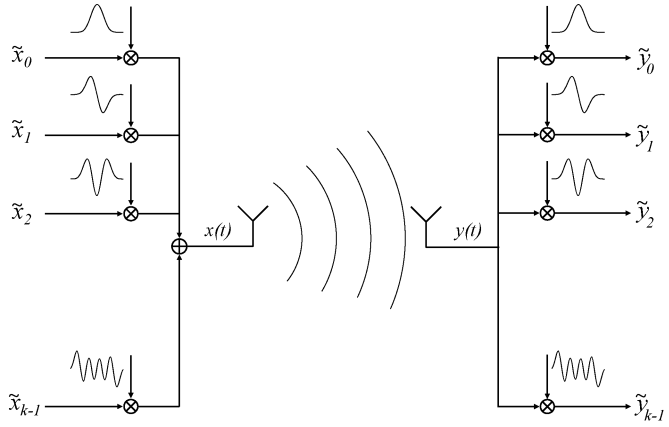

Fig. 5. Communication scheme for the single-antenna case. K symbols are modulated by K Slepian pulse-shaping functions, transmitted over the temporally and spectrally varying channel, and demodulated by a bank of K matched filters.

true for their time-sampled relatives; the dpss and their discrete Fourier transforms are identical, except for appropriate scale factors (asymptotically the Slepians become Hermite functions [15]). Therefore, the frequency operator in the Slepian basis is (almost) bidiagonal as well. While a plot of the frequency operator (the magnitude of its complex elements) looks so much like the time operator in Fig. 4 that it is not featured separately here, there is a critical difference between the two matrices that allows us to distinguish between temporal and spectral channel variation. The time operator is real-valued and symmetric, whereas the frequency operator is purely imaginary and skew symmetric. Thus, $\mathbf{C}$, $\mathbf{T}$, and $\mathbf{F}$ are jointly trace orthogonal.[2]

D. Channel Modulation With Slepian Sequences

Fig. 5 shows the modulation/demodulation process for the single-antenna case. The first K dpss sequences are chosen as pulse-shaping functions and modulated by the K block symbols to transmit. The sum of these waveforms is transmitted, distorted by the channel, and received (carrier modulation and other technical details are omitted in this picture). The vector of K demodulated symbols $\tilde{\mathbf{y}}_\beta$ is obtained after passing the received time series $\mathbf{y}_\beta$ through a bank of K matched filters.[3] For the $N \times N$ MIMO case, imagine N transmitters and receivers on either side. This procedure relates to the operators as follows.

[2]Insight into why Hermite/Slepian functions form a basis in which time and frequency have this symmetric relationship can be obtained by studying the algebra of raising and lowering operators in the quantum mechanics of simple harmonic oscillators. Recall that the Hermite functions $u_k(x)$ are eigenfunctions of the Hamiltonian $\mathbf{u} = (1/2)(\mathbf{x}^2 + \mathbf{p}^2)$ and that the raising/lowering operators $\mathbf{a} = (\mathbf{x} + i\mathbf{p}/\sqrt{2})$ and $\mathbf{a}^\dagger = (\mathbf{x} - i\mathbf{p}/\sqrt{2})$ have the property that they are lower/upper diagonal in the Hermite basis, i.e.,

$$\mathbf{a}^\dagger u_k(x) \propto u_{k+1}(x), \quad \text{and:} \quad \mathbf{a}u_k(x) \propto u_{k-1}(x).$$

Thus, $\mathbf{x} = (\mathbf{a} + \mathbf{a}^\dagger/\sqrt{2})$ and $\mathbf{p} = (\mathbf{a} - \mathbf{a}^\dagger/i\sqrt{2})$ are bidiagonal in the Hermite basis. Making the identification $x \leftrightarrow t$ and $p \leftrightarrow f$ completes the analogy.

[3]The Slepian sequences would be stored in a look-up table. If that were inconvenient for some reason, another set of possible pulse-shaping functions would be the sine tapers $s_k(i) = \sqrt{2/(n+1)}\sin(\pi ki/(n+1))$, where $i = 1, \ldots, n$, and $k = 1, \ldots, K$. These are orthonormal on $[1, n]$ and have the advantage that with the proper choice of n, *all* the coefficients can be computed simultaneously with a fast Fourier transform (FFT), that is, in $\mathcal{O}(n \log n)$ operations. Their energy concentration properties are poorer than those of the Slepians but adequate for cellular applications.

Let $\mathbf{\Psi}$ denote the $n \times K$ matrix that has the first K dpss as its columns.

- Modulation: The length n time series $\mathbf{x}_\alpha$ transmitted by transmitter α is a linear combination of K dpss
$$\mathbf{x}_\alpha = \mathbf{\Psi}\tilde{\mathbf{x}}_\alpha$$
where $\tilde{\mathbf{x}}_\alpha$ are the K digital symbols (pilots and data) to be transmitted.
- Demodulation: The length n received waveform $\mathbf{y}_\beta$ is demodulated by K matched filters $\mathbf{\Psi}^\dagger$:
$$\tilde{\mathbf{y}}_\beta = \mathbf{\Psi}^\dagger \mathbf{y}_\beta$$
and $\tilde{y}_\beta$ are the K symbols demodulated at receiver β.

The whole transmission equation in the Slepian basis is

$$\begin{aligned}\tilde{\mathbf{y}}_\beta &= \mathbf{\Psi}^\dagger \mathbf{y}_\beta \\ &= \mathbf{\Psi}^\dagger \sum_\alpha \left(H^c_{\beta\alpha}\mathbf{C} + H^t_{\beta\alpha}\mathbf{T} + H^f_{\beta\alpha}\mathbf{F} \right) \mathbf{\Psi}\tilde{\mathbf{x}}_\alpha \\ &= \sum_\alpha \left(H^c_{\beta\alpha}\mathbf{\Psi}^\dagger\mathbf{C}\mathbf{\Psi} + H^t_{\beta\alpha}\mathbf{\Psi}^\dagger\mathbf{T}\mathbf{\Psi} + H^f_{\beta\alpha}\mathbf{\Psi}^\dagger\mathbf{F}\mathbf{\Psi} \right) \tilde{\mathbf{x}}_\alpha \\ \tilde{\mathbf{y}}_\beta &= \sum_\alpha \left(H^c_{\beta\alpha}\tilde{\mathbf{C}} + H^t_{\beta\alpha}\tilde{\mathbf{T}} + H^f_{\beta\alpha}\tilde{\mathbf{F}} \right) \tilde{\mathbf{x}}_\alpha \qquad (6) \\ \tilde{\mathbf{y}}_\beta &= \sum_\alpha \tilde{\mathbf{H}}_{\beta\alpha}\tilde{\mathbf{x}}_\alpha. \qquad (7)\end{aligned}$$

Equation (7) expresses a direct relation between K transmitted and received *symbols* and not n instants in time, as in (5) (usually $K \ll n$). Since each Slepian function spans the entire signaling block (length T, bandwidth B), each symbol will be affected by all of the temporal and spectral variation occurring in one signaling block. The constant operator in the Slepian basis $\tilde{\mathbf{C}}$ is the $K \times K$ identity matrix ($\mathbf{\Psi}^\dagger\mathbf{\Psi}$ is the identity matrix since the dpss are orthogonal). $\tilde{\mathbf{T}}$ and $\tilde{\mathbf{F}}$ are bidiagonal for all practical purposes, as we saw earlier. As a result, the overall channel operator $\tilde{\mathbf{H}}_{\beta\alpha}$ has significantly nonzero elements only on its diagonal (due to constant transmission) and its two bidiagonals (due to first-order fluctuations). Physically, this means that the value of the kth received symbol $y_{\beta,k}$ depends only on $3N$ transmitted symbols, namely, $x_{\alpha,k-1}$, $x_{\alpha,k}$, and $x_{\alpha,k+1}$, where $\alpha = 1 \ldots N$. *Every symbol is distorted only by energy spilling from the two next neighbors of the symbol's own dpss*. This is the benefit of choosing the dpss as pulse-shaping functions. We exploit this simple ISI pattern for computationally efficient decoding, using dynamic programming (Viterbi's algorithm) of search depth three.

E. Estimation of the Model Parameters

To determine all channel parameters $H^c_{\beta\alpha}$, $H^t_{\beta\alpha}$, and $H^f_{\beta\alpha}$ ($3N^2$ complex coefficients), we transmit p pilot symbol matrices of size $N \times N$, where $p \geq 3$. This means that pN^2 out of a total of NTB block symbols are used as pilots: Unfortunately, the ratio of pilot symbols to total symbols grows with N in MIMO systems. Each antenna transmits pN pilot symbols per block. We modulate the pilot symbols $\tilde{\mathbf{x}}^{(\mathbf{pil})}_\alpha$ on the last pN dpss since the higher order sequences show stronger ISI (darker regions to the lower right of Fig. 4), which improves estimation accuracy. Between the pilot symbols and the data symbols, we

TABLE I
COMPARISON OF PARAMETERS FOR THE ACOUSTIC INDOOR CHANNELS OF OUR EXPERIMENTS, CELLULAR WIRELESS INDOOR CHANNELS OF THE SAME GEOMETRY, AND CELLULAR WIRELESS CHANNELS IN CARS. REFER TO SECTION III FOR DETAILS

parameter	acoustics	cellular wireless		
	indoors	indoors	at 50 km/h	at 150 km/h
wave velocity	340 m/s	$3 \cdot 10^8$ m/s	$3 \cdot 10^8$ m/s	$3 \cdot 10^8$ m/s
carrier frequency	1–2 kHz	1 GHz	1GHz	1GHz
wavelength	34 cm	30 cm	30 cm	30 cm
fading time	1–4 s	0.1–1 s	20–30 ms	6–10 ms
fading bandwidth	5–20 Hz	1–10 MHz	30–50 kHz	30–50 kHz
coherence number	$\approx 10^1$	$\approx 10^6$	$\approx 10^3$	$\approx 10^2$

leave one dpss unused. This "guard dpss" prevents spilling of energy from data into pilot symbols.[4] The pN received symbols $\tilde{\mathbf{y}}_\beta^{(\mathbf{pil})}$ at each antenna that correspond to the pilots are used to estimate the $3N^2$ channel coefficients by minimizing

$$\left(\tilde{\mathbf{y}}_\beta^{(\mathbf{pil})} - \sum_\alpha \tilde{\mathbf{H}}_{\beta\alpha}^{(\mathbf{pil})} \tilde{\mathbf{x}}_\alpha^{(\mathbf{pil})}\right)^2. \tag{8}$$

$\tilde{\mathbf{H}}_{\beta\alpha}^{(\mathbf{pil})}$ is the $p \times p$ submatrix of the channel operator $\tilde{\mathbf{H}}_{\beta\alpha}$ that is associated with the pilot-carrying dpss. Since the parameters $H_{\beta\alpha}^c$, $H_{\beta\alpha}^t$, and $H_{\beta\alpha}^f$ are multiplied linearly by the channel operator's three components

$$\tilde{\mathbf{H}}_{\beta\alpha} = \left(H_{\beta\alpha}^c \tilde{\mathbf{C}} + H_{\beta\alpha}^t \tilde{\mathbf{T}} + H_{\beta\alpha}^f \tilde{\mathbf{F}}\right) \tag{9}$$

we can fit them in a least mean squares sense ($3N^2$ unknowns and pN^2 measurements) by solving (8). The numerical values of the pilot symbols need to be optimized such that the normal matrix resulting from this least squares problem has a reasonably low condition number (around 4 in our experiments).

F. Quick Decoding

An adapted version of Viterbi's algorithm [20], [21] can be used for decoding. We saw that according to the model, every received symbol is determined by $3N$ transmitted symbols. The channel operator $\tilde{\mathbf{H}}_{\beta\alpha}$ is tridiagonal; if the column index is $k = 1, \ldots, K$, then its only nonzero elements are $h_{k,k-1}^{\beta\alpha}$, $h_{k,k}^{\beta\alpha}$, $h_{k,k+1}^{\beta\alpha}$. For all possible combinations of values that the $3N$ transmitted symbols $x_{\alpha,k-1}$, $x_{\alpha,k}$, $x_{\alpha,k+1}$ could take, we compute a penalty $P(\beta,k)$ as a measure of how well a particular combination explains the received symbol $y_{\beta,k}$. We chose the squared distance between received and predicted triplets

$$P(\beta,k) = \left(y_{\beta,k} - \sum_\alpha \left(h_{k,k-1}^{\beta\alpha} x_{\alpha,k-1} + h_{k,k}^{\beta\alpha} x_{\alpha,k} + h_{k,k+1}^{\beta\alpha} x_{\alpha,k+1}\right)\right)^2$$

as the penalty function $P(\beta,k)$. Every sent symbol $x_{\alpha,k}$ influences $3N$ received symbols, and a serious mismatch in explaining any of them would add to the penalty associated with that guess for $x_{\alpha,k}$. In the forward pass of Viterbi's algorithm, we move through the Slepian index $k = 1, \ldots, K$, calculating the penalties associated with all possible combinations of $3N$ sent symbols. We obtain a minimum penalty and the combination of the best fitting $K \times N$ symbols associated with it, which is our best guess for what was sent. For a constellation of size M, a data block length of K, and a $N \times N$ MIMO channel, the computational complexity of the least mean squares algorithm for a full channel matrix would be M^{KN}, whereas it is M^{3N} for the tridiagonal channel matrix described above. Thus, for small N (e.g., 2×2, 3×3 channels), the method presented in this paper provides a tractable scheme for treating both temporal and spectral variations in the channel.

[4]Note that we are not making use of space-time coding [4], [5], [16] for now (each symbol is transmitted by only one antenna). This or other more sophisticated coding and training techniques could be implemented on top of the scheme that we propose.

III. ACOUSTIC EXPERIMENTS

A. Why Acoustics?

As a proof of concept, we tested the scheme on an actual physical system. For experiments, we chose the acoustic rather than the wireless channel for two reasons. First, it requires far less expensive equipment and is easy to handle due to the greatly reduced bandwidths. Our platform featured two stereo loudspeakers and two microphones from Radio Shack. Second, it has a much lower (spectral) coherence than the radio channel: In an indoor environment, the time-bandwidth product of a block is not much higher than 10. This makes it a particularly difficult test case for channel estimation and provides an excellent opportunity to observe the benefits of time-frequency estimation. We argue that the acoustic testbed provides a physically meaningful analogy because the wave propagation aspects we are interested in are basically the same as in wireless. Polarization plays no role in our scheme. For the discussion that follows, refer to Table I for parameter conversion between the two channel types. The values given for the acoustic channel in Table I were measured in the experiments we conducted in various indoor environments. For the indoor cellular wireless channel, the same room geometries were assumed, and values for temporal and spectral coherence were estimated from first

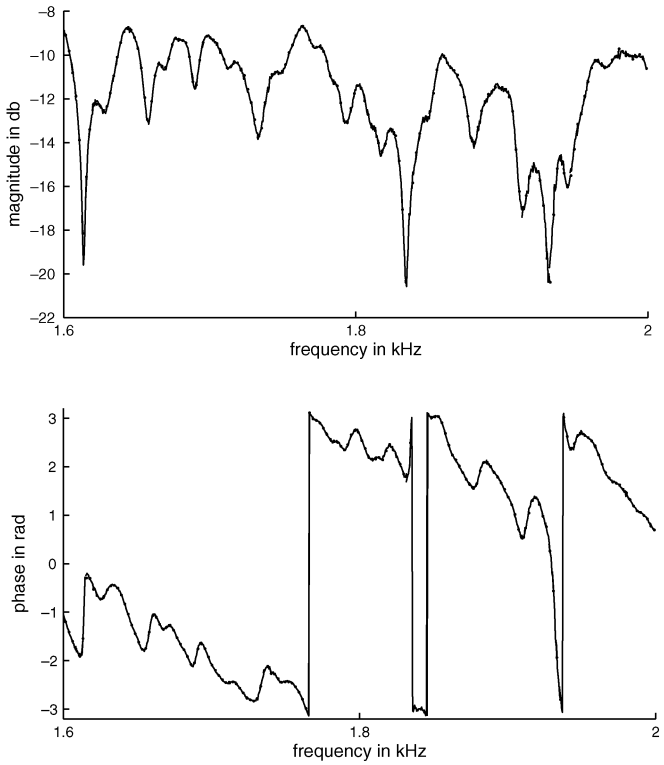

Fig. 6. Magnitude and phase of an acoustic channel transfer function, as measured in the quiet and empty Bell Labs cafeteria. Fades are spaced by about 20 Hz.

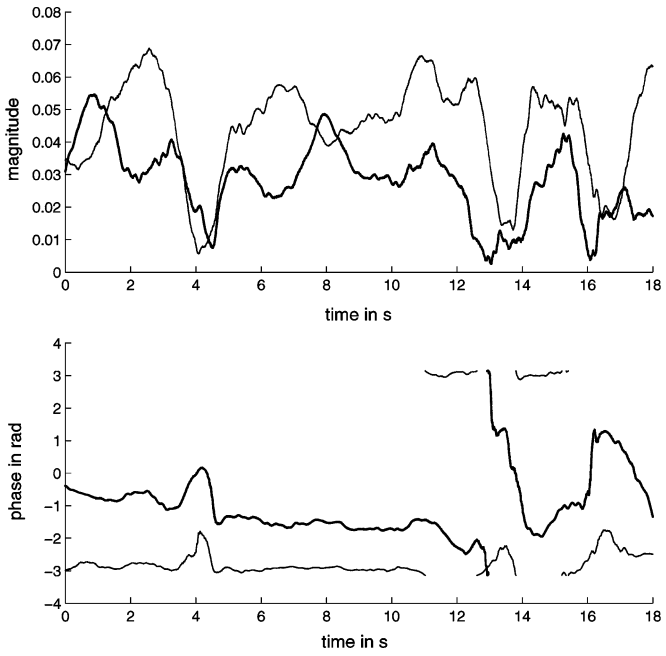

Fig. 7. Temporally varying amplitudes of two acoustic tones at 1500 and 1520 Hz, measured in the Bell Labs cafeteria during lunch time. The time scale of amplitude fluctuations is on the order of 1 sec; amplitudes vary over roughly one order of magnitude.

principles (see below). The values for cellular communications in cars at 50 km/s and 150 km/s are empirical rule-of-thumb values.

Environment geometry is the determining factor for both temporal and spectral channel coherence. Spectral coherence (spacing of spectral channel fades) is usually determined by the largest dimension of the room in which one is working. Temporal fading is due to both Doppler effects and movement relative to standing wave interference patterns in the environment (the second effect tends to be dominant in cellular wireless). We chose our acoustic carrier frequency (1–2 kHz) such that the wavelengths were the same as typical cellular wavelengths (30 cm). For identical environment geometry, the scaling of the two problems is therefore the same. Fig. 6 shows a typical indoor transfer function: in this case in the Bell Labs cafeteria. Its overall appearance with deep fades of 10 dB and more, and with relatively regular spacing, is very similar to the fading radio channel. In this experiment and others, spectral fades were spaced by 5–20 Hz, mainly depending on the size of the room. If we assumed reflection and attenuation coefficients to be the same in acoustics and wireless, we would expect the spectral coherence of the wireless channel to be six orders of magnitude higher than that of the acoustic channel. The reason is that in a given geometry, acoustic waves at a frequency f_0 would generate the same interference pattern as radio waves at $(c_{em}/c_{ac}) \times f_0$. The first-principle calculation of 5 MHz spectral coherence for the geometry of the Bell Labs cafeteria is consistent with these expectations. (We do not conclude that the propagation coefficients actually are the same since there is no reason why they should be. We do conclude that order of magnitude estimation seems to work quite well.) Temporal coherence (spacing of temporal fades) of acoustic and radio channels can be expected to be roughly the same. In the case of Doppler effects, it is basically the inverse of the absolute Doppler shift, which depends on the wavelength and the speed at which the environment changes. When one is moving relative to standing wave patterns, temporal coherence is directly proportional to one's speed. Fig. 7 shows a measurement of time variation over 18 sec of the acoustic channel in the Bell Labs cafeteria during lunch hour. The characteristic speed in this environment is that of people strolling to their lunch tables. Temporal fading in Fig. 7 is on the order of 2 sec, which is reasonably consistent with the first principle estimate of 0.1–1 sec for this room.

The important thing to realize is that the "coherence number" (product of temporal and spectral coherence) is about six orders of magnitude less in acoustics than in wireless. The number of symbols that we can accommodate in one signaling block $(K = TB - 1)$ will be a certain fraction of this coherence number. For a given SNR or number of pilots, this fraction will be greater if we can model both temporal and spectral fluctuations within a block. With a coherence number on the order of only 10, the acoustic channel is a very nonflat and nonstationary channel to communicate over, which provides an excellent opportunity to observe the benefits of time-frequency modeling. Wireless radio channels impose less severe constraints since their coherence number is several orders of magnitude higher. However, rapid temporal variation causes coherence to deteriorate significantly, for example, when cell phones are used in cars (last column of Table I). Pilots in current cellular systems still use up a substantial fraction of the total signaling resources, typically 5–10% in terms of time-bandwidth usage (and even more

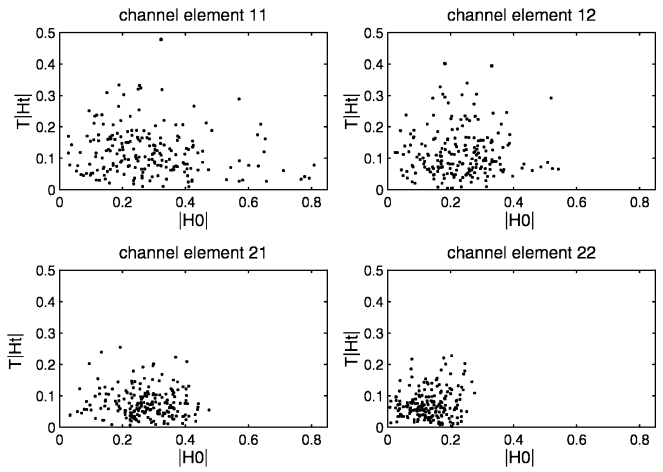

Fig. 8. Estimated magnitudes of first-order temporal variation $T|H^t|$ versus constant channel attenuation $|H^c|$ for the 200 acoustic 2×2 MIMO channels measured in the acoustic experiment.

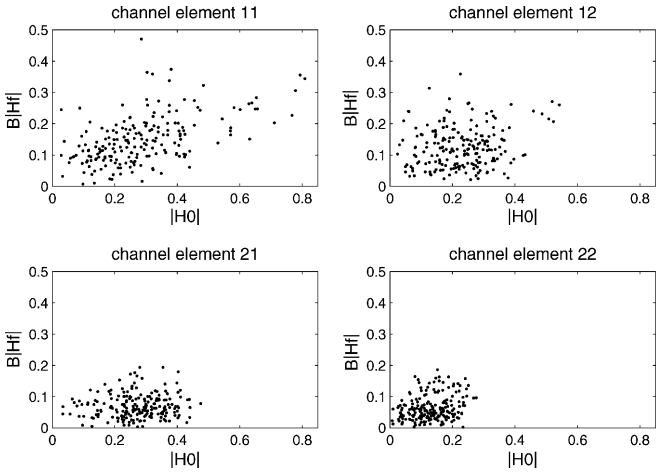

Fig. 9. Estimated magnitudes of first-order spectral variation $B|H^f|$ versus constant channel attenuation $|H^c|$ for the 200 acoustic 2×2 MIMO channels measured in our experiment.

in terms of power but we do not consider the power aspect here). Pilot efficiency is a serious concern in real-world wireless.

B. Results of 2 × 2 MIMO Experiments

Ideally, the block duration T and the block bandwidth B should be chosen such that the channel varies linearly in time as well as in frequency. To learn the characteristic temporal and spectral scales of our channels, we took global measurements like Figs. 6 and 7, as well as measurements that probed many blocks with the highest possible frequency resolution (K sinusoidal tones per block) or the highest time resolution (K pulses per block). The curvature of these partial channel variations plotted in the complex plane provides a convenient measure for the amount of higher order temporal or spectral fluctuations present in the system. We found that purely first-order frequency variation was a reasonable assumption for $B = 4$–8 Hz, depending on the room in which we were working. Purely first-order time variation could be assumed for $T = 1$–3 sec, depending on how fast people were moving around.

We conducted SISO and 2×2 MIMO experiments. Since space does not allow us to present all results, we choose to present MIMO results here for two reasons. First, we want to show that the harder MIMO case actually works, even under the highly nonflat nonstationary conditions in acoustics. Second, we think that our time-frequency approach could turn out to be most useful in contexts such as MIMO, where a disproportionally high pilot requirement mandates particularly efficient usage of the channel. The results below were obtained in an office room of approximately $4 \times 5 \times 3\ \mathrm{m}^3$. One person was asked to move around constantly. Block duration was chosen to be $T = 2$ sec, block bandwidth $B = 6$ Hz, number of symbols $KN = 11 \times 2$, number of pilots $pN^2 = 4 \times 2 \times 2$, carrier frequency 1.5 kHz, and sampling rate $S = 20$ kHz. Estimates of temporal and spectral fluctuations are shown in Figs. 8 and 9 for 200 channel realizations of 2×2 MIMO channels. We plot the magnitudes of temporal fluctuation $T|H^t|$ and the magnitude of spectral fluctuation $B|H^f|$, both versus the magnitude of the constant channel component $|H_0|$. Note that the mean values of $T|H^t|$ and $B|H^f|$ are of the same order of magnitude; both kinds of fluctuations cause comparable amounts of distortion. For the chosen block size, we compare four different channel estimation methods:

- first-order time-frequency model
$$\tilde{\mathbf{y}}_\beta = H^c_{\alpha\beta}\tilde{\mathbf{C}}\tilde{\mathbf{x}}_\alpha + H^t_{\alpha\beta}\tilde{\mathbf{T}}\tilde{\mathbf{x}}_\alpha + H^f_{\alpha\beta}\tilde{\mathbf{F}}\tilde{\mathbf{x}}_\alpha;$$
- first-order time variation only estimation
$$\tilde{\mathbf{y}}_\beta = H^c_{\alpha\beta}\tilde{\mathbf{C}}\tilde{\mathbf{x}}_\alpha + H^t_{\alpha\beta}\tilde{\mathbf{T}}\tilde{\mathbf{x}}_\alpha;$$
- first-order frequency variation only estimation
$$\tilde{\mathbf{y}}_\beta = H^c_{\alpha\beta}\tilde{\mathbf{C}}\tilde{\mathbf{x}}_\alpha + H^f_{\alpha\beta}\tilde{\mathbf{F}}\tilde{\mathbf{x}}_\alpha;$$
- zeroth-order (constant-only) estimation
$$\tilde{\mathbf{y}}_\beta = H^c_{\alpha\beta}\tilde{\mathbf{C}}\tilde{\mathbf{x}}_\alpha.$$

We are interested in the time- and frequency-only schemes because they constitute the "two halves" of the time-frequency model. For all four models, parameters were estimated from the same set of pN^2 symbols $\tilde{\mathbf{y}}^{(\mathbf{pil})}_\beta$, $\beta = 1, 2$, by fitting to the data in a least mean squares sense. Performance criterion was the signal-to-noise power ratio (SNR) in decibels, summed over all $N \times K$ block symbols

$$\begin{aligned}\mathrm{SNR}_m &= 10\log\left(\frac{\sum_{\beta=1}^{N}\left\|\tilde{\mathbf{y}}^{(\mathbf{m})}_\beta\right\|^2}{\sum_{\beta=1}^{N}\left\|\tilde{\mathbf{y}}_\beta - \tilde{\mathbf{y}}^{(\mathbf{m})}_\beta\right\|^2}\right)\\ &= 10\log\left(\frac{\mathrm{var}\left(\tilde{\mathbf{H}}^{(\mathbf{m})}\tilde{\mathbf{x}}\right)}{\mathrm{var}\left(\tilde{\mathbf{H}}_{\mathbf{tr}}\tilde{\mathbf{x}} + \tilde{\eta} - \tilde{\mathbf{H}}^{(\mathbf{m})}\tilde{\mathbf{x}}\right)}\right)\end{aligned} \quad (10)$$

where $\tilde{\mathbf{y}}^{(\mathbf{m})}_\beta$ are the K received symbols predicted by model m, and $\tilde{\mathbf{y}}_\beta$ are the received symbols. $\tilde{\mathbf{H}}^{(\mathbf{m})}$ and $\tilde{\mathbf{H}}_{\mathbf{tr}}$ are the estimated and the true channel in the Slepian basis, and $\tilde{\eta}$ is the receiver noise. For every estimation method, we thus divide the total symbol power by the total noise power due to model misfit

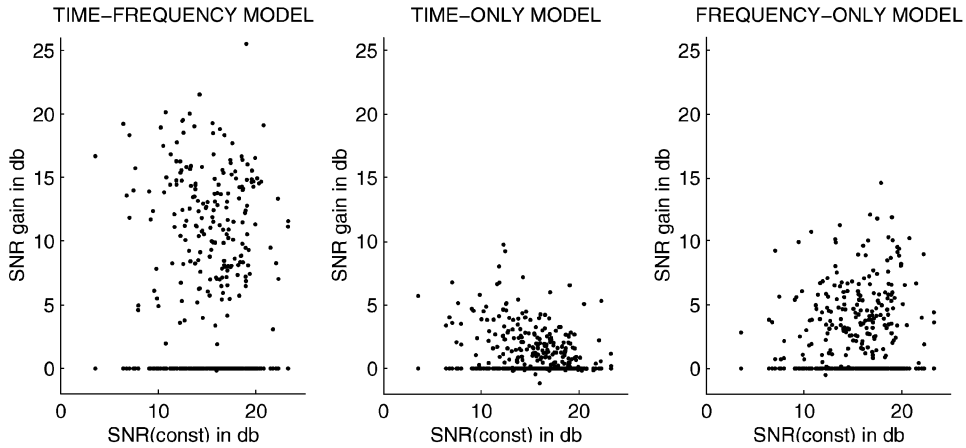

Fig. 10. Relative estimation gains of the three first-order schemes over the zeroth-order scheme, as found in the 2×2 MIMO experiment. $\mathrm{SNR}_{tf} - \mathrm{SNR}_c$, $\mathrm{SNR}_t - \mathrm{SNR}_c$, and $\mathrm{SNR}_f - \mathrm{SNR}_c$ in decibels, plotted over SNR_c, the SNR of the zeroth-order scheme. Note that SNR_c is also plotted in every diagram (the 0 db baseline) for easier visual comparison. Clearly, the joint time-frequency model achieves the best estimation results.

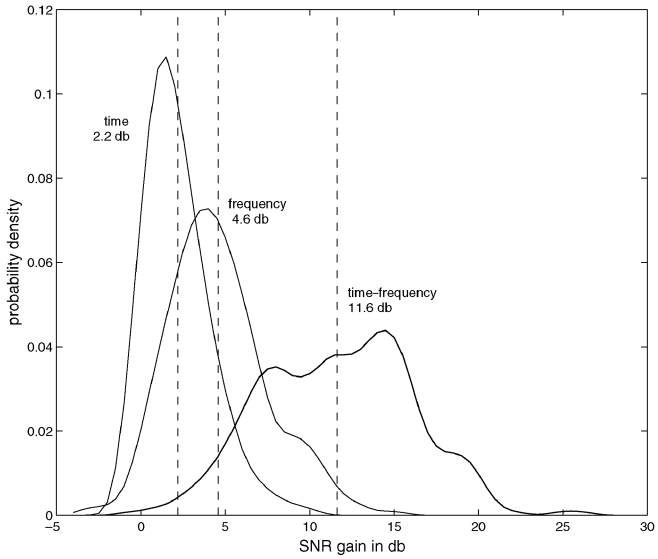

Fig. 11. Histogram of relative method gains in decibels of the three first-order schemes over the zeroth-order scheme, i.e., $\mathrm{SNR}_{tf} - \mathrm{SNR}_c$, $\mathrm{SNR}_t - \mathrm{SNR}_c$, and $\mathrm{SNR}_f - \mathrm{SNR}_c$.

or misestimation. The SNR performance of the three first-order models versus the zeroth-order model is shown in Fig. 10 for each of the 200 channel realizations. Plotted on the x-axes is the absolute SNR_c of the constant-only scheme, whereas the y-axis shows the relative gains $\mathrm{SNR}_{(m)} - \mathrm{SNR}_c$ of the three first-order schemes over the zeroth-order scheme (for easier visual comparison, SNR_c is also plotted as the 0 dB baseline). Clearly, the SNR_{tf} scatterplot is shifted toward higher gains over the baseline than the time- and frequency-only schemes. Time-frequency always did better than time-only or frequency only. In this particular experiment, spectral fluctuations $B|H^f|$ were slightly higher on average than temporal variations $T|H^t|$.

The scatterplots are summed up in Fig. 11, which is a histogram of relative gains of the three first-order schemes over the constant-only model (i.e., $\mathrm{SNR}_{tf} - \mathrm{SNR}_c$, $\mathrm{SNR}_t - \mathrm{SNR}_c$, and $\mathrm{SNR}_f - \mathrm{SNR}_c$). The time-frequency model shows the highest SNR gains in 2×2 acoustic MIMO experiments. Its average gain over the constant-only method is of 11.6 dB, indicated by the dashed line. The peaks of the time- and frequency-only gains are at lower SNRs; their average values are 2.2 and 4.6 dB, respectively. Note that the sum of these two values is much smaller than the average SNR of the time-frequency scheme: *Joint* time-frequency estimation clearly is more than just the performance of the two partial schemes added up. Single antenna experiments in the same environment yielded results very similar to Figs. 10 and 11. In the SISO case, and relative SNR gains of time-frequency estimation were about 3 dB higher.

IV. Simulations

We ran SISO simulations to investigate how the four different schemes compare for typical cellular wireless parameters. To make the results more tangible, we chose $T = 3.3$ ms and $B = 30$ kHz, which are the block duration and bandwidth used by the TDMA air interface (25 kHz for GSM). We aim at simulating cell phone usage while driving: an everyday situation where moderate to severe channel fluctuations put current wireless systems under stress, and where improvements would have a significant impact. Thirty kHz is comparable to the fading bandwidth in an outdoor cellular environment (30–50 kHz, see Table I); therefore, we expect substantial spectral fluctuations (phase variations on the order of π) in a block. For fading time for cars moving on highways (6–10 ms), 3.3 ms is of the same order of magnitude. Time variations are expected to be somewhat less severe within a block than spectral variations for this particular choice of parameters but still on the same order of magnitude. If T and B could be chosen freely in a wireless system, joint time-frequency estimation would perform most effectively if the magnitudes of time and frequency variation were exactly the same. We mimicked nonstationary multipath propagation to compute scalar channel realizations. Two or three echoes with randomized gains, propagation delays, time slopes, and fractional Doppler shifts were added at the receiver. Magnitudes and phases of temporal and spectral variation are uncorrelated. For each of the four methods, $L = 5$ pilots were used to estimate the channel.

Due to our choice of parameters, the ratio between B and fading bandwidth (around 0.8) is roughly two times larger than the ratio between T and fading time (around 0.4). To model this difference, we generated channels that, on average, varied

roughly twice as strongly over frequency than over time. To characterize the magnitude of first-order variation, we use the ratios $R_T = T|H^t|/|H^c|$ and $R_B = B|H^f|/|H^c|$ as estimated for a noise-free channel. We generated 10 000 channels with 30% to 60% first-order time variation $(0.3 \leq R_T \leq 0.6)$ and 70% to 100% first-order spectral variation $(0.7 \leq R_B \leq 1.0)$. For illustration, consider a channel with $H^c = (1 + 0i)$, whose spectral component starts out at $(0.75 - 0.25i)$ at $f = f_0 - B/2$ and ends up at $(1.25 + 0.25i)$ at $f = f_0 + B/2$; let its temporal component be $(0.8 + 0i)$ at $t = 0$ and $(1.2 + 0i)$ at $t = T$. Such a channel would satisfy our specifications since $R_B = |(1.25 + 0.25i) - (0.75 - 0.25i)|/|1 + 0i| = 0.707$, and $R_T = |(1.2 + 0i) - (0.8 + 0i)|/|1 + 0i| = 0.4$. The channels had only small to moderate fluctuations of order higher than one, which was judged by their curvature in the complex plane when probed at highest temporal or spectral resolution.

Each channel realization was run at different receiver noise levels. On the x-axes in Fig. 12, we plot signal-to-receiver-noise ratio (srnr), which is defined as the signal power after passing through the channel, divided by the power of the white Gaussian noise subsequently added at the receiver. The y-axis of the upper plot shows the received SNR values of (10) after demodulation for all four methods (constant-only, time-only, frequency-only,and time-frequency). The middle plot shows SNR gains of the three first-order methods, relative to the zeroth-order (constant-only) method. It shows the same data as the top plot, except that the performance of the constant-only scheme is taken as the 0-dB baseline. The lower plot shows bit error rates (BERs) for a two-symbol constellation using the Viterbi-style decoder described in Section II-F. Cellular systems are designed to operate in the srnr regime between 0 and 20 dB, with occasional "bad" channel realizations falling into the region of rapid bit error increase below 0 dB. For low receiver noise (srnr $>$ 20 dB), all methods are operating at their maximum performance, as indicated by the four flat curves in the top part of Fig. 12. The middle plot compares the relative performance of the four schemes: For srnr $= 20$ dB, the SNR of time-frequency estimation is 18 db above that of constant only; time-only and frequency-only are still performing about 2.5 and 7.5 dB above constant-only. Clearly, the performance of joint time-frequency modeling is far higher than the two partial first-order schemes added up. The BERs in this regime are 0.5% for time-frequency, 0.7% for time-only, 0.6% for frequency-only, and 1.0% for constant-only. At srnr $= 10$ dB, the average relative SNR gains are still 12, 2, and 6 dB, respectively.

In constrast, at high receiver noise around srnr $= -10$ dB, all four methods break down. The SNR versus srnr slopes in the top of Fig. 12 are all close to 1, indicating that the random receiver noise we put in simply reappears as estimation noise after demodulation. Differences between the schemes become apparent as srnr increases. The time-frequency curve keeps rising at a slope of one, which means that all temporal and spectral channel fluctuations get accounted for, and only the added receiver noise stays unexplained. The slopes of constant-only, time-only, and frequency-only level off and approach zero: The partial methods interpret the unmodeled spectral and/or temporal variations as noise. Time-frequency only levels off around srnr $= 20$ dB when all first-order variations are accounted for and higher order

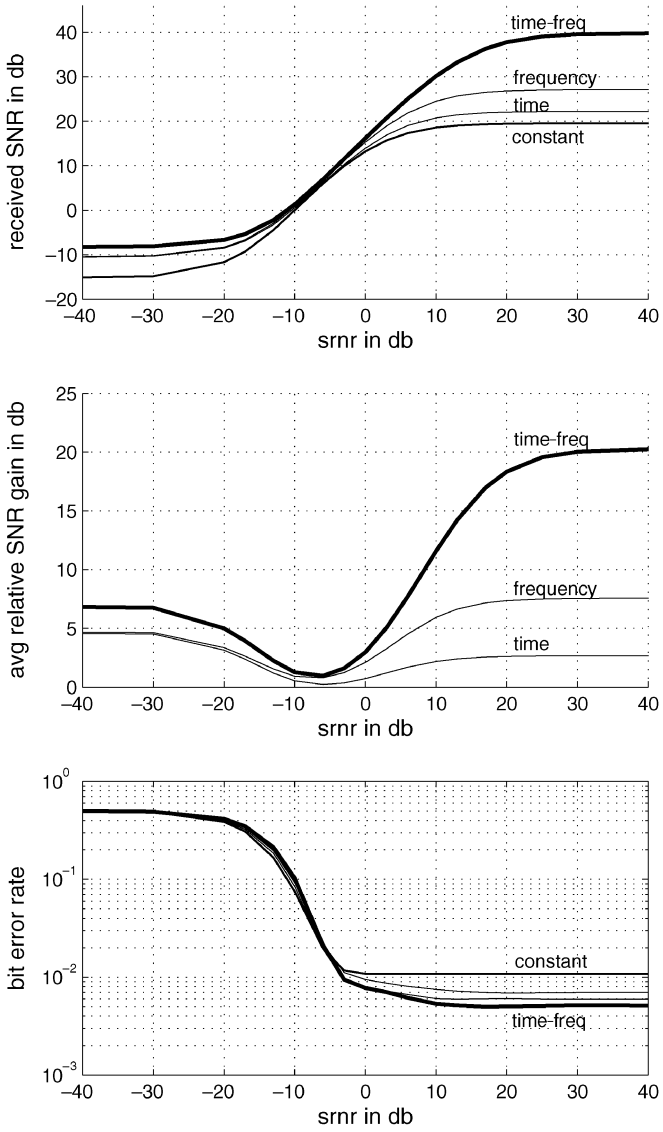

Fig. 12. Estimation results for wireless simulations, as a function of the signal-to-added-noise level (srnr) at the receiver. The channels had temporal and spectral fluctuations of the magnitudes expected for cellular communications in cars (refer to the text for parameters and definitions). (Top) Average received SNR for all four estimation methods. (Middle) Average SNR gains of the three first-order methods over the constant-only method. (Bottom) BERs for a binary constellation.

variations, which are weak but nevertheless present due to the way we generate channel realizations, become the dominant source of estimation noise. Time-frequency also outperforms the other methods in terms of BER, but the differences become smaller as the environment deteriorates. Below -3 db, BER rises exponentially, and all schemes are doing equally badly.

The regime of srnr < -10 dB and BER $> 10\%$ is of no practical interest. However, its seemingly counterintuitive behavior deserves a brief discussion. One might expect that receiver SNR would continue to deteriorate linearly below srnr $= -10$ dB. Instead, all four curves level off toward values between -9 and -15 dB. This is a consequence of the definition of SNR in (10). As the added receiver noise $\tilde{\eta}$ becomes overwhelming, the pilots become so noisy that the magnitude of the resulting channel estimate $\tilde{\mathbf{H}}^{(\mathbf{m})}$ is determined by the magnitude of $\tilde{\eta}$ rather than

staying close to the true channel $\tilde{\mathbf{H}}_{\mathbf{tr}}$. $\tilde{\mathbf{H}}_{\mathbf{tr}}\tilde{\mathbf{x}}$ becomes negligible in (10), and both $\tilde{\eta}$ and $\tilde{\mathbf{H}}^{(\mathbf{m})}\tilde{\mathbf{x}}$ play the role of noise components. Since numerator and denominator now are highly dependent, the SNR values saturate. Time-frequency estimation still achieves the highest relative SNR because it fits the noise with three instead of two or one parameters. Time-frequency BER is slightly worse than for the other schemes: Since its estimate is closer to the noise, it is further from the (drowned) channel $\tilde{\mathbf{H}}_{\mathbf{tr}}$.

We saw that in the region where real-world wireless systems attempt to operate (srnr between 0 and 20 dB), BER decreases by a factor of two (from 1% to 0.5%) when first-order time-frequency estimation is used instead of constant-only. This result was obtained for a binary constellation. Given its high relative SNR gains, the time-frequency scheme could support constellations with more symbols and still achieve the same BER as constant-only does on the binary constellation. For example, an SNR gain of 12 dB (at srnr $= 10$ dB) corresponds to a factor of $10^{12/20} \approx 4$ by which constellation symbols could be spaced more closely.

V. Implications for Pilot Efficiency

In order to give numbers for asymptotic SNR and pilot efficiency gains for the first-order time-frequency channel, we would need to know its distribution. The problem is studied in statistics under "calibration" or "inverse prediction"; a description of the bi-variate problem (i.e. first-order time-only or frequency-only) is given in [13, Sec. 7.2.6], where the distribution is shown to be t_{n-2}. To our knowledge, rigorous treatment of the multivariate complex case still needs to be done but is beyond the scope of this paper. However, we get a feel for the method's potential if we look at the (smaller) benefits of the first-order *time-only* model, as compared to constant-only. Jakes [8] discusses the classical Rayleigh fading model, where the covariance of fading is $J_0(2\pi f_d \tau)$, and f_d is the Doppler frequency. A realistic scenario for a SISO channel could be this: A cell phone system designed for a maximum speed of 40 m/s (144 km/h) and operating at a carrier frequency of 1.8 GHz would have $f_d \approx 240$ Hz. If one plots the eigenvalues of the Karhunen–Loeve expansion of the channel over the product of block duration T and Doppler frequency f_d, a reasonable choice of T would be the point where the first ignored term in the channel is -20 dB down in the eigenvalue plot. Let the channel bandwidth $B = 30$ kHz. For the constant-only model, we get $T \approx 300$ μs, and $K \approx 10$ symbols per block. According to [7], the optimum number of pilots is $p = 3$; therefore, 30% of channel resources are used for pilots. For the time-only model (meaning constant part plus time slope), we get $T \approx 1700$ μs and $K \approx 51$ symbols. We need twice as many pilots ($p = 6$) but still get a lower relative pilot usage, namely 6/51 $\approx$ 12%. The same numbers for 2.4 GHz, which is another standard carrier frequency, work out to be 43% pilot overhead for the constant-only model and 16% for the first-order time-only model. For the first-order *time-frequency* model, we expect another significant gain in channel usage, given the model's superior performance relative to time-only in Figs. 11 and 12. Our hope is that members of the community will take up this interesting line of research.

VI. Conclusions

We have proved the concept of time-frequency channel modeling in wireless communications. Experiments and simulations show that for our first-order scheme, received SNR increases significantly compared to models that do not take into account both temporal and spectral variations. In acoustic 2×2 MIMO experiments, our first-order time-frequency modulation scheme yielded estimation results superior to constant-only modulation by 11.6 dB. In cellular wireless simulations (SISO case) with moderate to severe channel fluctuations, we gain 12 to 18 dB in terms of received SNR when signal-to-receiver-noise is 10 to 20 dB. The BER decreases by a factor of two for a binary constellation. These results were achieved for equal block, signaling power, and number of pilots. By comparison, the benefits of first-order time-only or frequency-only estimation are small if both kinds of fluctuations are actually present. The increase in estimation overhead (three parameters versus one for constant-only or two for time-only and frequency-only) is small or moderate compared with the SNR gain. In terms of computation, Slepian sequences (dpss) are ideal pulse-shaping functions for the first-order time-frequency channel because they limit ISI to three symbols, thus allowing for quick decoding using Viterbi's algorithm. Other time-frequency channel parameterizations could be treated with the formalism of linear operators presented here. MIMO communications require a larger fraction of the channel resource to be allocated to pilots. We think, therefore, that simultaneous treatment of temporal and spectral variations of the channel, as discussed in this paper, may prove critical in optimal usage of MIMO channels.

References

[1] M. R. Andrews, P. P. Mitra, and R. deCarvalho, "Tripling the capacity of wireless communications using electromagnetic polarization," *Nature*, vol. 409, pp. 316–318, Jan. 18, 2001.

[2] M. R. Andrews, P. P. Mitra, and D. J. Thomson, "Communication Employing Triply-Polarized Transmissions," US Patent 6 317 098, Nov. 2001.

[3] C. Flammer, *Spheroidal Wave Functions*. Stanford, CA: Stanford Univ. Press, 1957.

[4] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Techn. J.*, vol. 1, no. 2, pp. 41–59, Autumn 1996.

[5] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, p. 311, Mar. 1998.

[6] J. D. Gibson, *The Mobile Communications Handbook*, Second ed. Boca Raton, FL: CRC/IEEE Press, 1999, ch. 18.

[7] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–964, Apr. 2003.

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

[9] T. Kuroda and T. Matsumoto, "Multicarrier signal detection and parameter estimation in frequency selective Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 46, no. 4, pp. 882–890, Nov. 1997.

[10] C. Loader, *Local Regression and Likelihood*. New York: Springer, 1999.

[11] D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications*. Cambridge, U.K.: Cambridge Univ. Press, 1993.

[12] J. G. Proakis, *Digital Communications*, Third ed. New York: McGraw-Hill, 1995, p. 764.

[13] G. A. F. Seber, *Linear Regression Analysis*. New York: Wiley, 1977.

[14] D. Slepian and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty—I," *Bell Syst. Tech. J.*, vol. 40, pp. 43–63, 1961.

[15] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis and uncertainty—V.: the discrete case," *Bell Syst. Tech. J.*, vol. 57, pp. 1371–1430, 1978.
[16] E. Teletar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, 1999.
[17] D. J. Thomson, "Quadratic-inverse spectrum estimates; applications to paleoclimatology," *Phil. Trans. R. Soc. Lond.*, ser. A, vol. 332, pp. 539–597, 1990.
[18] ——, "Non-stationary fluctuations in 'stationary' time series," in *Proc. SPIE*, vol. 2027, 1993, pp. 236–244.
[19] ——, "Multitaper analysis of nonstationary and nonlinear time series data," in *Nonlinear and Nonstationary Signal Processing*, W. Fitzgerald, R. Smith, A. Walden, and P. Young, Eds. Cambridge, U.K.: Cambridge Univ. Press, 2001, pp. 317–394.
[20] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inf. Theory*, vol. IT-13, pp. 260–269, Apr. 1967.
[21] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. New York: McGraw-Hill, 1979.

Karin Sigloch (M'04) received the Master's degrees in electrical engineering from the University of Karlsruhe, Karlsruhe, Germany, and the Institut Polytechnique de Grenoble (ENSIEG), Grenoble, France, in 2002. Her thesis research was done at Bell Labs, Lucent Technologies, Murray Hill, NJ in 2001 and 2002. She has since been working toward the Ph.D. degree in geophysics/seismology with the Geosciences Department, Princeton University, Princeton, NJ.

Her current research explores the measurement and inversion of seismic body wave amplitudes for global tomography.

Michael R. Andrews received the B.S. degree in physics and mathematics from the State University of New York at Stony Brook, Stony Brook, NY, in 1992 and the Ph.D. degree in physics from the Massachusetts Institute of Technology, Cambridge, in 1998.

From 1998 to 2002, he was a Principal Investigator at Bell Laboratories, Lucent Technologies, Murray Hill, NJ, and since 2002, he has been with Flarion Technologies, Bedminster, NJ.

Partha P. Mitra received the B.Sc. (Hons.) degree in physics from Presidency College, Calcutta University, Calcutta, India, in 1989 and the Ph.D. degree in physics from Harvard University, Cambridge, MA, in 1993.

He joined Bell Laboratories, Murray Hill, NJ, as a postdoctoral fellow in 1993 and became a Member of Technical Staff in the department of theoretical physics in 1995. His research interests at Bell ranged from theoretical physics to neuroscience and theoretical engineering including wireless and optical fiber communications. He spent 1996 at the California Institute of Technology, Pasadena, as an assistant professor of theoretical physics. He returned to Bell in 1997. In 2003, he left Bell Laboratories to join Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, where he is a Professor of neuroscience. His current research includes applying signal processing and data-mining techniques to biological data, as well as understanding the functioning of biological systems from the perspective of engineering principles.

David J. Thomson (F'91) was born in Saint John, NB, Canada in 1942, graduated from Acadia University, Wolfville, NS, in mathematics and physics. He received the M.S. and Ph.D. degrees in electrical engineering from Polytechnic Institute, Brooklyn, NY, in 1967 and 1971, respectively.

He became a Member of Technical Staff at Bell Telephone Laboratories, Murray Hill, NJ, in 1965. He worked on the WT4 Millimeter Waveguide System and the Advanced Mobile Phone Service project, where he was responsible for the circuit design of and software for a microprocessor-controlled modem for Rayleigh fading channels. In 1983, he was a Green Scholar at Scripps Institution of Oceanography, San Diego, CA, and became a Distinguished Member of Technical Staff in the Communications Analysis Research Department. In addition to spectrum estimation, his current research interests are analysis of global climate data and space physics.

Dr. Thomson is a member of the American Geophysical Union, the Statistical Society of Canada, and the American Statistical Association. He is a Chartered Statistician and a Fellow of the Royal Statistical Society. He was a member of the Panel on Sensors and Electron Devices of the Army Research Laboratory Technical Assessment Board, Chairman of Commission C of USNC-URSI, and an associate editor for *Radio Science*. He was associate editor for Communications Theory and for Detection and Estimation of the IEEE TRANSACTIONS ON INFORMATION THEORY. He has taught courses at Princeton University, Princeton, NJ, and Stanford University, Stanford, CA, gave the Houghton lectures at Massachusetts Institute of Technology, Cambridge, and was a participant at the Isaac Newton Institute at the University of Cambridge, Cambridge, U.K. In 2002, he became a Canada Research Chair in Statistics and Signal Processing with the Department of Mathematics and Statistics, Queen's University, Kingston, ON, Canada.