

OVERLAID OR TIME-MULTIPLEXED PILOTS FOR CHANNEL ESTIMATION IN ITERATIVE MIMO RECEIVERS

Mohammad Ali Khalighi[†], Lamia Berriche[‡], Jean-François Hélard[†]

[†] Institut d'Électronique et de Télécommunications de Rennes (IETR), 20 av. des Buttes de Coësmes, 35043 Rennes, France

[‡] UMR 5141 CNRS - GET / Télécom Paris (ENST), 37-39 rue Dareau, 75014 Paris, France

Email: Ali.Khalighi@ieee.org, Lamia.Berriche@enst.fr, Jean-Francois-Helard@insa-rennes.fr

ABSTRACT

We consider the use of overlaid pilots for multiple input multiple output (MIMO) channel estimation in the context of single-user and single-carrier modulation. Imposing no loss in spectral efficiency, overlaid pilots are potentially of special interest, notably in relatively fast fading channels. However, when channel is to be estimated over short sequences, estimation errors can be important, due to the non-zero cross-correlation of data and pilot sequences. We show how these estimation errors limit the performance of the receiver, when iterative data detection and channel estimation is performed. Also, for the case of iterative detection at receiver, we compare the performances of the overlaid and the time-multiplexed pilot schemes.

1. INTRODUCTION

Conventional time-multiplexed data and pilot (training) symbols can cause an important loss in spectral efficiency. This is especially the case for MIMO systems when a large number of antennas is employed at transmitter and/or when the communication channel undergoes fast variations [1]. An alternative solution may be to use overlaid (also called superimposed or embedded) training sequences for channel estimation [2, 3] that impose no loss in spectral efficiency.

We consider in this paper MIMO channel estimation based on overlaid pilots (OP). Block fading channel model is considered where the channel is assumed almost constant over a block of symbols. Assuming uncorrelated data and pilot sequences, we estimate channel coefficients by calculating the cross-correlation between the received sequences on each antenna and the transmitted pilot sequences, known to receiver. However, although data and pilot sequences are *statistically* uncorrelated, the cross-correlation is calculated over a block of symbols of limited length, over which the channel coefficients are supposed to remain unchanged. The smaller the block length (i.e., the faster the channel fading), the more important this cross-correlation is. This can result in an error floor in the receiver bit-error-rate (BER) [4] and make the OP scheme lose its interest.

In this work, our aim is to investigate whether or not iterative data detection and channel estimation can be a solution to the problem of error floor. In fact, by iterative detection, while profiting from the channel coding gain,

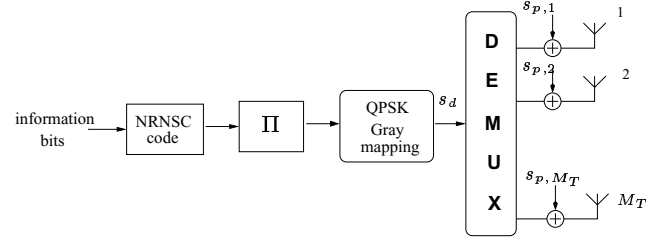


Fig. 1. Block diagram of transmitter with overlaid pilots

we can estimate the transmitted data symbols through iterations, and try to reduce their effect on the estimated channel coefficients. Moreover, we want to see if overlaid pilots are preferable to time-multiplexed pilots (MxP) for such an iterative scheme.

2. SYSTEM MODEL

Let us denote by M_T and M_R the number of antennas at transmitter and receiver, respectively. Frequency non-selective Rayleigh fading conditions are considered. The block diagram of the transmitter is shown in Fig.1. Bit interleaved coded modulation is performed using a non recursive non systematic convolutional (NRNSC) code. The interleaver Π is of random type and QPSK modulation is considered. Pilot sequences $s_{p,i}$, $i=1, \dots, M_T$ are assumed to be mutually orthogonal. The vector of transmitted symbols from M_T antennas \mathbf{s} at a given time sample is the superposition of the data symbols vector \mathbf{s}_d and the pilot symbols vector \mathbf{s}_p : $\mathbf{s} = \mathbf{s}_d + \mathbf{s}_p$. We denote by σ_d^2 and σ_p^2 the variance (power) of data and pilot symbols, respectively. The received vector \mathbf{y} of dimension M_R at a given time sample is

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n} = \mathbf{H}(\mathbf{s}_p + \mathbf{s}_d) + \mathbf{n}. \quad (1)$$

Channel matrix \mathbf{H} is of dimension $(M_R \times M_T)$ and \mathbf{n} is the vector of additive complex white Gaussian noise of zero mean and variance σ_n^2 . We assume that σ_n^2 is known at receiver, and focus on the estimation of channel coefficients.

3. ITERATIVE DATA DETECTION AND CHANNEL ESTIMATION

The block diagram of the receiver is shown in Fig.2. We perform soft-input soft-output (SISO) log-MAP symbol detection, as well as SISO log-MAP channel decoding.

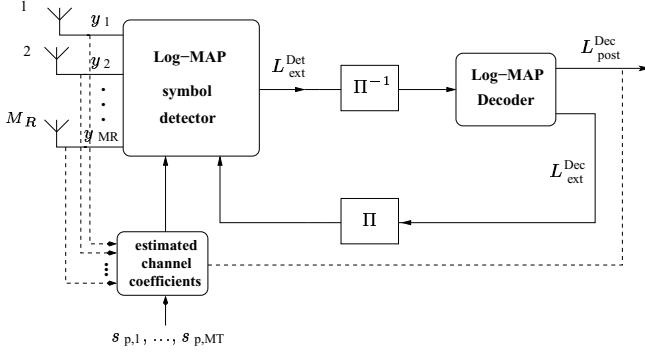


Fig. 2. Block diagram of the receiver

Soft-values are in the form of logarithmic likelihood ratio (LLR). In Fig.2, L^{Det}_{ext} denotes extrinsic LLR at the symbol detector output, and L^{Dec}_{ext} and L^{Dec}_{post} denote extrinsic and a posteriori LLR at the decoder output, respectively.

Consider, for example, the estimation of the channel coefficient \mathbf{H}_{ij} . At the first iteration, this estimate is obtained by calculating the cross-correlation Γ between the sequences \mathbf{y}_i and $\mathbf{s}_{p,j}$:

$$\Gamma(\mathbf{s}_{p,j}, \mathbf{y}_i) = \frac{1}{N_s} \sum_{n=1}^{N_s} \mathbf{s}_{p,j}^*[n] \cdot \mathbf{y}_i[n]. \quad (2)$$

N_s is the number of channel uses, i.e., the frame length, and n is the channel use time reference. From (1) we have:

$$\mathbf{y}_i[n] = \sum_{t=1}^{M_T} (\mathbf{s}_{p,t}[n] + \mathbf{s}_{d,t}[n]) \mathbf{H}_{it} + \mathbf{n}_i[n]. \quad (3)$$

Assuming mutually orthogonal pilot sequences we obtain:

$$\begin{aligned} \Gamma(\mathbf{s}_{p,j}, \mathbf{y}_i) &= \mathbf{H}_{ij} + \frac{1}{N_s} \sum_{n=1}^{N_s} \sum_{t=1}^{M_T} \mathbf{s}_{d,t}[n] \mathbf{s}_{p,j}^*[n] \mathbf{H}_{it} \\ &\quad + \frac{1}{N_s} \sum_{n=1}^{N_s} \mathbf{n}_i[n] \mathbf{s}_{p,j}^*[n]. \end{aligned} \quad (4)$$

At the first iteration, we take $\hat{\mathbf{H}}_{ij}^{(1)} \approx \Gamma(\mathbf{s}_{p,j}, \mathbf{y}_i)$. We notice that there are two interference terms on the estimated $\hat{\mathbf{H}}_{ij}$: one term from the non-zero cross-correlation of data and pilot sequences and the other from the cross-correlation of noise samples with pilots. The more important one is obviously the former, especially at high SNR where the channel estimation errors become important and result in an error floor in the receiver performance. In the succeeding iterations, we calculate soft estimates of the transmitted data symbols $\tilde{\mathbf{s}}_d$ by using a posteriori LLRs at the SISO decoder output and cancel their effect in $\Gamma(\mathbf{s}_{p,j}, \mathbf{y}_i)$. In this way, we try to reduce the first interference term in (4). For expressions for the soft estimated symbols see [5]. At iteration $m > 1$, we calculate first the data removed vectors $\tilde{\mathbf{y}}_i^{(m)}$, and then the new estimate of \mathbf{H}_{ij} , as follows.

$$\tilde{\mathbf{y}}_i^{(m)}[n] = \mathbf{y}_i[n] - \sum_{t=1}^{M_T} \tilde{\mathbf{s}}_{d,t}[n] \hat{\mathbf{H}}_{it}^{(m-1)} \quad (5)$$

$$\hat{\mathbf{H}}_{ij}^{(m)} = \Gamma(\mathbf{s}_{p,j}, \tilde{\mathbf{y}}_i^{(m)}) = \frac{1}{N_s} \sum_{n=1}^{N_s} \mathbf{s}_{p,j}^*[n] \cdot \tilde{\mathbf{y}}_i^{(m)}[n] \quad (6)$$

4. CRAMÉR-RAO BOUNDS

In our performance analysis, we use the Cramér-Rao bound (CRB), which is the lower bound to the mean square of estimation errors in the set of unbiased estimates. Let \mathbf{h} be the vector of parameters to be estimated, in which the columns of \mathbf{H} are stacked. Also, let $p_{\mathbf{h}}(\mathbf{h})$ be the pdf of \mathbf{h} . We consider the *Bayesian* CRB that assumes Gaussian distributed symbols and whose expression for the OP scheme is [6]:

$$\text{CRB} = \frac{\sigma_n^2}{\sigma_n^2 \rho_h^2 + N_s \sigma_p^2}, \quad \rho_h^2 = \mathbb{E} \left[\left| \frac{\partial \ln p_{\mathbf{h}}(\mathbf{h})}{\partial \mathbf{h}^*} \right|^2 \right]. \quad (7)$$

This corresponds to the case of non-coded data and channel estimation based on pilots only. To modify (7) for the case of coded data, we assume that after channel coding and interleaving, data symbols remain almost uncorrelated. So, for a code rate R , we use in (7) the modified power $\sigma_d'^2 = \sigma_d^2/R$ [7]. In other words, the redundancy insertion in time is transposed into power increase.

Note that we could alternatively use the more realistic stochastic CRB for the given constellation of symbols. However, the numerical calculation of the stochastic CRB is too complex. We will later see in Section 5.2 that the Bayesian CRB is useful enough for our analyses.

5. NUMERICAL RESULTS

To study the performance of the proposed estimation scheme we consider the channel code $(5, 7)_8$ of rate $R = 1/2$. We denote by α the ratio of power of pilot symbols to the total transmit power at a symbol time, i.e., $\alpha = \sigma_p^2/(\sigma_p^2 + \sigma_d^2)$. SNR is considered in the form of E_b/N_0 with E_b the average received energy per information data bit, and N_0 the unilateral noise psd. We include in E_b/N_0 the receive array gain M_R ; however, it returns to data symbols only, and does not take into account pilot symbols. Here, for the OP case, the total (from pilot and data) average received SNR (dB) can be obtained by adding the factor $10 \log_{10} \frac{1}{1-\alpha}$ to E_b/N_0 (dB).

5.1. OP scheme, performance versus α

Figures 3 and 4 show curves of BER versus E_b/N_0 for different values of α , and for the data frame length of $N_d = 200$ bits, corresponding to 50 channel uses. BER curves are shown for the first and fifth iterations where almost full convergence is attained. We have $M_T = M_R = 2$ in Fig.3 and $M_T = 2$ and $M_R = 4$ in Fig.4, that we denote by 2×2 and 2×4 , respectively. No constraint is imposed on the total transmit power as we want to highlight the effect of α . It is seen that for the 2×2 case, although iterative processing helps improve considerably the performance for $\alpha > 5\%$, we have important error floors even for values of α as large as 50%. Remember that this error floor is due to non-zero cross-correlation between data and pilots. When more receive diversity is available at receiver (as in Fig.4), data detection can be performed more efficiently, and the cross-correlation of data and symbols can be suppressed more efficiently too.

The interfering cross-correlations are inversely proportional

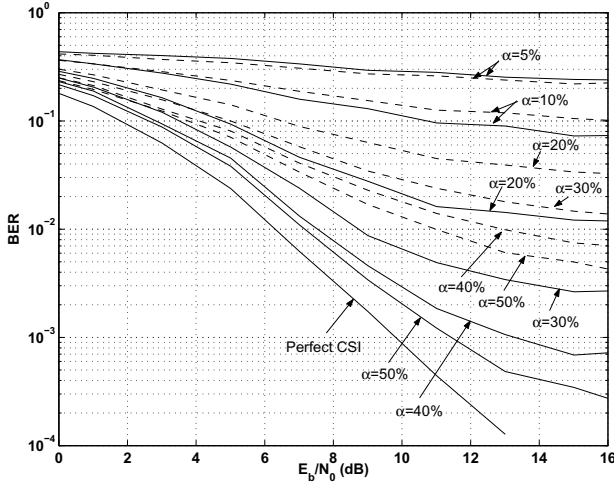


Fig. 3. (2×2) MIMO channel, code $(5, 7)_8$, $N_d = 200$; dashed line: first iteration; solid line: fifth iteration

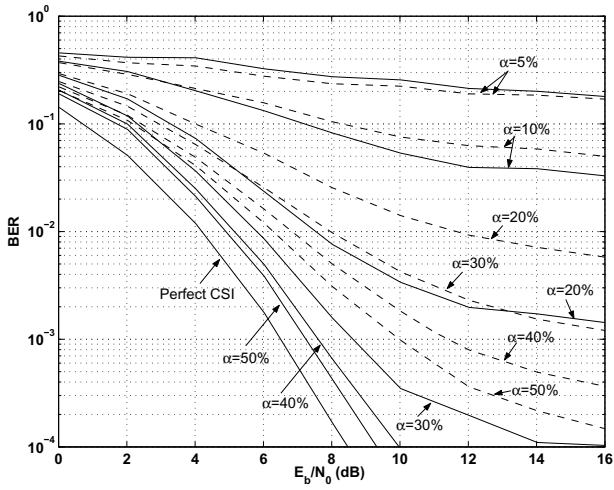


Fig. 4. (2×4) MIMO channel, code $(5, 7)_8$, $N_d = 200$; dashed line: first iteration; solid line: fifth iteration

to the frame length (see [4] for analytical calculations). If we consider ten times longer blocks with $N_d = 2000$, the performances are much better (Fig.5). Here, for $\alpha \geq 4\%$, we are very close to the case of perfect channel knowledge. Notice that this case can be regarded as with channel time variations ten times slower than those in Fig.3,4.

5.2. Channel estimation errors

To see how the mean square of the estimation errors (MSE) improves through iterations, curves of MSE versus E_b/N_0 are shown in Figures 6 and 7 for 2×4 system with $N_d = 2000$ and $\alpha = 2\%, 5\%$. MSE is the sum of mean square error on the $M_T M_R$ sub-channels' coefficients. We see that when α is not large enough, the interference originating from the cross-correlation of data and pilots can not be reduced sufficiently, although it is reduced through the four first iterations. For $\alpha = 5\%$, however, MSE is efficiently reduced and after five iterations, MSE (dB) decreases almost linearly with SNR (dB); it is also closer to the CRB.

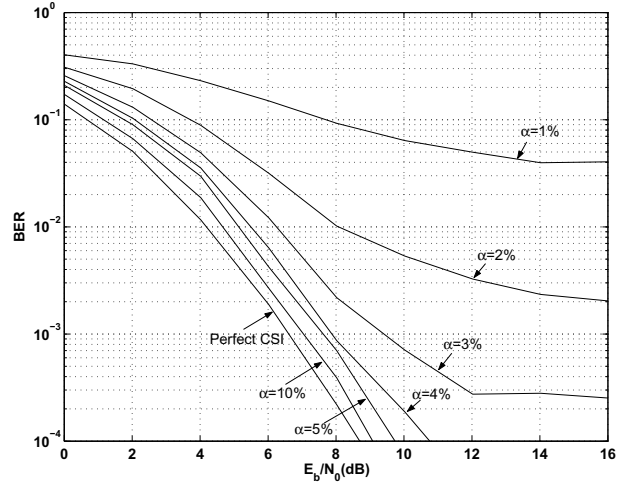


Fig. 5. (2×4) channel, code $(5, 7)_8$, $N_d = 2000$

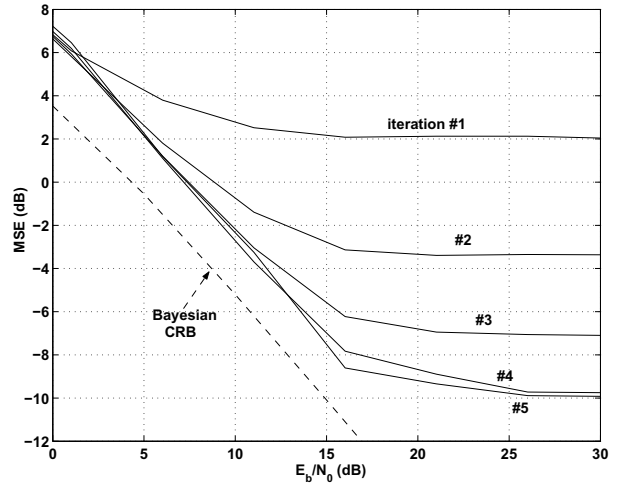


Fig. 6. (2×4) channel, code $(5, 7)_8$, $N_d = 2000$, $\alpha = 2\%$

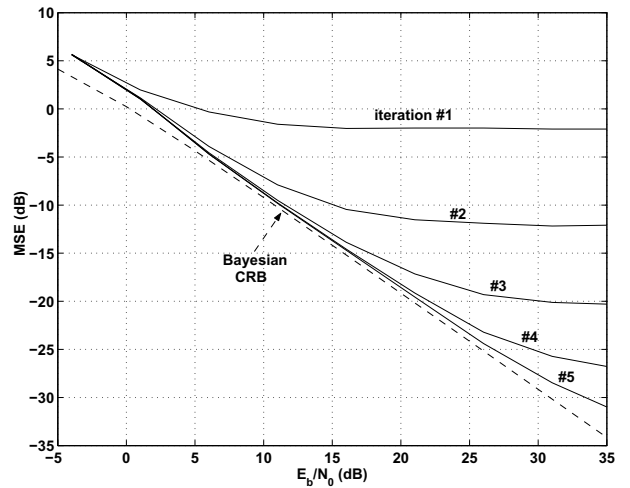


Fig. 7. (2×4) channel, code $(5, 7)_8$, $N_d = 2000$, $\alpha = 5\%$

5.3. Comparison with MxP scheme

Consider the MxP scheme with N_p and N_d , the number of pilot and data symbols in a frame, respectively. In order to compare the receiver performance with OP and MxP

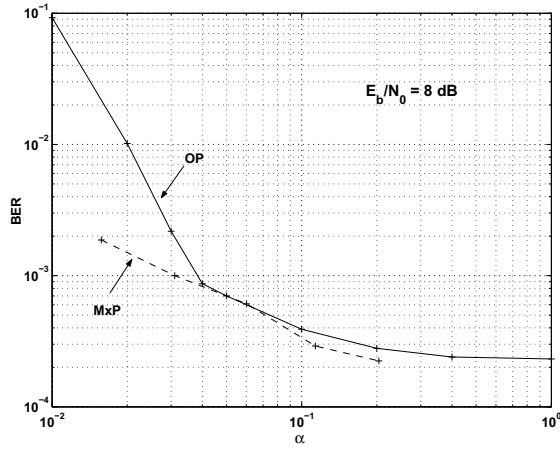


Fig. 8. (2×4) channel, code (5, 7)₈, $N_d = 2000$

schemes, we impose the constraint of equal transmit energy for the transmission of N_d data bits. We attribute the same power σ_d^2 to both data and pilot symbols. By taking $N_p/(N_p + N_d) = \alpha$ and considering the same E_b/N_0 for data symbols in both OP and MxP cases, the same energy will be used for the transmission of the ensemble of data and pilot symbols. Meanwhile, The MxP scheme has a spectral efficiency loss of α , compared to OP. Consider first the case of 2×4 system with $N_d = 2000$. We saw in Subsection 5.1 that a relatively small α can result in a good convergence of the detector for the OP scheme. To compare the two estimation schemes, let us first fix E_b/N_0 and see which scheme works better for a given α . Fig.8 shows BER curves of the fifth iteration for $E_b/N_0 = 8$ dB. It is seen that the two schemes provide almost equivalent performances for $\alpha \geq 4\%$. For relatively large α , OP is preferred since it imposes no loss in spectral efficiency. For smaller α values, MxP scheme works better.

Let us now consider a desired performance, for example, the MSE of the estimation errors, and control the SNR or α . Curves of α versus E_b/N_0 are shown in Fig.9 for two MSE values of -10 dB and -20 dB and the case of short frames. We can see how to set α for an available SNR to obtain the desired MSE. Note that high SNR values are more interesting since they correspond to low BER. We see for high SNRs how power (for OP) should be traded off with spectral efficiency (for MxP). For example, consider the criterion of attaining $\text{MSE} = -10$ dB. For a given $E_b/N_0 = 15$ dB, either we should use the MxP scheme, which imposes a loss of 10% in the data rate, or use the OP scheme, which imposes an increase of about 20% in the transmit energy. On the other hand, for a given $E_b/N_0 = 5$ dB, the OP scheme is quite preferable to MxP, since it requires less transmit energy and imposes no loss in spectral efficiency.

6. CONCLUSIONS

We studied the OP-based MIMO channel estimation in an iterative detection scheme. Channel estimation was based on pilots only. We showed that for long enough frames, the OP scheme is very interesting. However, for too short frames, i.e., too fast fading channels, even by it-

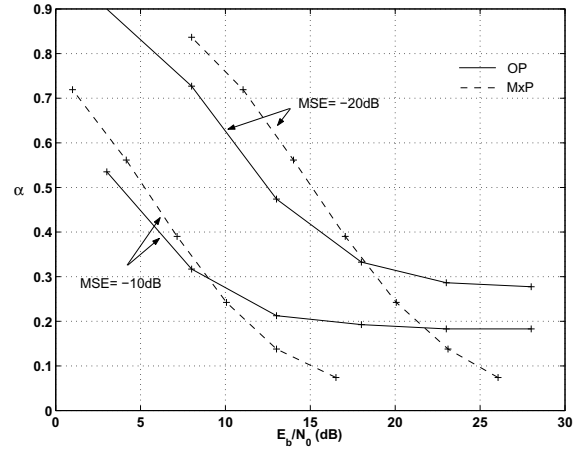


Fig. 9. (2×2) channel, code (5, 7)₈, $N_d = 200$

erative processing the resulting performance may not be satisfying unless relatively high power is dedicated to pilot sequences. We showed that in such a case, time multiplexed pilots can still be preferred to overlaid pilots. In some cases, however, we can trade off the transmit energy for OP scheme with spectral efficiency for MxP scheme. Anyway, for fast fading channels, without iterative processing, or in the absence of coding, OP is almost of no interest, compared to MxP. At last, we should point out that the OP scheme is especially interesting when used with spread spectrum techniques, where a distinct code can be dedicated to each one of the M_T pilot sequences.

7. REFERENCES

- [1] B. Hassibi and B.M. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. IT*, vol. 49, no. 4, pp. 951–963, Apr. 2003.
- [2] H. Zhu, B. Farhang-Boroujeny, and C. Schlegel, "Pilot embedding for joint channel estimation and data detection in MIMO communication systems," *IEEE Comm. Lett.*, vol. 7, no. 1, pp. 30–32, Jan. 2003.
- [3] P. Hoeher and F. Tufvesson, "Channel estimation with superimposed pilot sequence," *GLOBECOM*, vol. 4, pp. 2162–2166, 1999, Rio de Janeiro, Brazil.
- [4] V. Jungnickel, T. Haustein, E. Jorswieck, V. Pohl, and C. von Helmolt, "Performance of a MIMO system with overlay pilots," *GLOBECOM*, vol. 1, pp. 594–598, Nov. 2001, San Antonio, TX.
- [5] M.A. Khalighi and J. Boutros, "Channel estimation in Turbo-BLAST detectors using EM algorithm," *ICASSP*, Mar. 2005, Philadelphia, PA.
- [6] L. Berriche, K. Abed-Meraim, and J.C. Belfiore, "Cramer-Rao bounds for MIMO channel estimation," *ICASSP*, vol. 4, pp. 397–400, May 2004.
- [7] H. A. Cirpan and M. K. Tsatsanis, "Maximum-likelihood estimation of FIR channels excited by convolutionally encoded inputs," *IEEE Trans. COM*, vol. 49, no. 7, pp. 1125–1128, July 2001.