Article:

Ghogho, M., McLernon, D., Alameda-Hernandez, E. et al. (1 more author) (2005) Channel estimation and symbol detection for block transmission using data-dependent superimposed training. IEEE Signal Processing Letters, 12 (3). pp. 226-229. ISSN 1070-9908

https://doi.org/10.1109/LSP.2004.842283

Channel Estimation and Symbol Detection for Block Transmission Using Data-Dependent Superimposed Training

Mounir Ghogho, Des McLernon, Enrique Alameda-Hernandez, and Ananthram Swami

***Abstract*—We address the problem of frequency-selective channel estimation and symbol detection using superimposed training. The superimposed training consists of the sum of a known sequence and a data-dependent sequence that is unknown to the receiver. The data-dependent sequence cancels the effects of the unknown data on channel estimation. The performance of the proposed approach is shown to significantly outperform existing methods based on superimposed training (ST).**

***Index Terms*—Block transmission, estimation, multipath channels, superimposed training.**

I. Introduction

CHANNEL estimation is a major challenge for reliable wireless transmissions. In most practical systems, this task is achieved by using pilot symbols that are known to the receiver. Often, these pilot symbols are time-division multiplexed (TDM) with the data, i.e., pilots and information-bearing symbols are transmitted on different time slots. Although accurate channel estimates can be obtained if the pilots are judiciously placed [1], this method requires extra bandwidth. An alternative method is the superimposed training (ST) scheme, where pilots are added to the data symbols [2]–[6]. This scheme saves valuable bandwidth at the expense of a reduction in the information signal-to-noise ratio (ISNR), since some of the transmitted energy is allocated to the hidden pilots. ST schemes offer tradeoffs between loss of rate (slots for training) and simplicity of the receiver, channel estimation versus tracking, and, possibly, improved power efficiency.

The idea of using ST to estimate the channel has recently received renewed attention. Recent contributions are based on superimposing a known periodic sequence on the data [3]–[7]. In [7], the deterministic (or sample) mean of each data block was removed prior to transmission, and this approach was shown to reduce the effect of the unknown data on the performance of both channel estimation and equalization. Here, we propose an ST scheme that fully cancels the effects of the unknown data on the performance of the channel estimator. Unlike the conventional ST scheme, the training sequence is the sum of a periodic sequence, which is known to the receiver, and a *data-dependent* sequence which is unknown to the receiver. We show that by judiciously selecting the latter sequence, a very significant improvement in terms of estimation accuracy and symbol error rate can be obtained.

Notation: Superscripts * and ^{T} denote Hermitian and transpose operators. The trace and statistical expectation are denoted by $\mathrm{Tr}\{\cdot\}$ and $E\{\cdot\}$. The discrete Fourier transform (DFT) of a $(N\times 1)$ vector $\boldsymbol{x}$ is denoted by $\tilde{\boldsymbol{x}} = \mathbf{F}_N\boldsymbol{x}$, where $\mathbf{F}_N$ has (m,n) entry $1/\sqrt{N}e^{-j2\pi mn/N}, m,n = 0,\ldots,N-1$. $\mathbf{F}_{P,L}$ will denote the leading $P\times L$ submatrix of $\mathbf{F}_P$. The nth element of a vector $\boldsymbol{z}$ is denoted by $z(n)$, and $\otimes$ will denote the Kronecker product. The $(N\times N)$ identity matrix is denoted by $\mathbf{I}$ and the $(Q\times Q)$ matrix of all ones by $\mathbf{1}_Q$. Finally, $\mathrm{diag}(a_1,\ldots,a_N)$ is the $(N\times N)$ diagonal matrix whose nth diagonal entry is a_n.

II. Proposed ST Scheme

Consider a single-carrier block transmission system operating over a frequency-selective channel. Let N denote the block length. We assume the channel to be time invariant over a single block, but it could vary across blocks. Assume that the memory of the discrete-time channel is upper bounded by $L-1$, which is known. Let $\boldsymbol{h} = [h_0,\ldots,h_{L-1}]^T$ denote the impulse response of the channel. In order to avoid interblock interference, a cyclic prefix (CP) of length $L-1$ is inserted between the blocks. At the receiver, after removing the CP, the signal model for each block can be expressed as

$$\boldsymbol{x} = \mathcal{H}\boldsymbol{s} + \boldsymbol{v} \qquad (1)$$

where $\boldsymbol{s}$ is the $(N\times 1)$ transmitted block, $\mathcal{H}$ is an $(N\times N)$ circulant matrix with first column $[h_0, h_1, \ldots h_{L-1}, 0,\ldots,0]^T$, and $\boldsymbol{v}$ is an additive white noise vector with covariance $\sigma_v^2\mathbf{I}$. Further, $\boldsymbol{s}$ is assumed to be zero mean and independent of $\boldsymbol{v}$. Since we perform block-by-block processing, we do not need a block index in (1).

In a TDM scheme, some of the entries of $\boldsymbol{s}$ are known pilots. In the conventional ST scheme, a known training sequence $\boldsymbol{c}$ is added to the data vector $\boldsymbol{w}$, i.e., $\boldsymbol{s} = \boldsymbol{w} + \boldsymbol{c}$. The data symbols are assumed to be zero mean, independent and identically distributed random variables drawn from a finite alphabet, e.g., phase shift keying or quadrature amplitude modulation; let σ_w^2 denote the data symbol power. The channel coefficients can be

Manuscript received July 21, 2004; revised October 13, 2004. M. Ghogho was supported by the U.K. Royal Academy of Engineering and the European Research Office. E. Alameda-Hernandez was supported by Secretaría de Estado de Educación y Universidades (Spain) and by the European Social Fund. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Lang Tong.

M. Ghogho, D. McLernon, and E. Alameda-Hernandez are with the School of Electronic and Electrical Engineering, University of Leeds, Leeds LS29JT U.K.

A. Swami is with the U.S. Army Research Laboratory, Adelphi, MD 20783 USA.

Digital Object Identifier 10.1109/LSP.2004.842283

consistently estimated using the first-order statistics of the received signal [3]–[5]. In order to simplify channel estimation, $\boldsymbol{c}$ is often chosen to be periodic; let P denote its period, and assume that $Q = N/P$ is an integer. A disadvantage of this method is that the performance of channel estimation is affected by the embedded unknown data, which acts like input noise. In order to better explain this effect and also to motivate the proposed ST scheme, we use the following frequency domain interpretation. Since $\mathcal{H}$ is circulant, the DFT of $\boldsymbol{x}$ can be written as

$$\tilde{\boldsymbol{x}} = \mathbf{H}\tilde{\boldsymbol{c}} + \mathbf{H}\tilde{\boldsymbol{w}} + \tilde{\boldsymbol{v}} \tag{2}$$

where $\mathbf{H} = \text{diag}(H(0), \ldots, H(N-1))$, with $H(k)$ being the frequency response of the channel at frequency k

$$H(k) = \sum_{\ell=0}^{L-1} h_\ell e^{-j2\pi k\ell/N}, \quad k = 0, \ldots, N-1.$$

Since $\boldsymbol{c}$ is periodic with period P, its energy is concentrated only at the P equispaced frequency bins $k = \ell Q, \ell = 0, \ldots, P-1$, which we will refer to as pilot frequencies. In contrast, the energy of the data symbols is spread over all frequency bins. The channel coefficients are then estimated using the pilot frequencies, treating $\mathbf{H}\tilde{\boldsymbol{w}}$ and $\tilde{\boldsymbol{v}}$ as additive noise sequences. Here, we propose to develop a channel estimator that is completely impervious to the unknown data. We propose to *distort* the data vector so that its DFT at the pilot frequencies $k = \ell Q, \ell = 0, \ldots, P-1$ is identically zero. Let the distorted data vector be denoted by $\boldsymbol{w} - \boldsymbol{e}$; it is easy to verify that the corresponding DFT at the pilot frequencies is zero, provided that

$$\sum_{m=0}^{Q-1} w(i+mP) = \sum_{m=0}^{Q-1} e(i+mP), \quad i = 0, \ldots, P-1. \tag{3}$$

A trivial choice is $e(n) = w(n)$, which leads to pure training. For superimposed training, the energy of the distortion sequence $\boldsymbol{e}$ should be minimized for a fixed energy in the data sequence $\boldsymbol{w}$. The solution in this case is found to be

$$e(i+mP) = e(i) = \frac{1}{Q}\sum_{m=0}^{Q-1} w(i+mP), \quad i = 0, \ldots, P-1$$

which is the cyclic mean of the data. In this case, data distortion corresponds to simply removing the cyclic mean of the data. The distortion component can be written as $\mathbf{e} = \mathbf{J}\boldsymbol{w}$, where $\mathbf{J} = (1/Q)\mathbf{1}_Q \otimes \mathbf{I}_P$. The transmitted block is then given by

$$\boldsymbol{s} = (\mathbf{I} - \mathbf{J})\boldsymbol{w} + \boldsymbol{c}. \tag{4}$$

Now, the DFT of $\boldsymbol{s}$ at the pilot frequencies is identical to that of $\boldsymbol{c}$. In other words, channel estimation will only be affected by the additive noise $\boldsymbol{v}$ and not by the unknown data vector $\boldsymbol{w}$. The proposed ST technique can also be seen as a data-dependent superimposed training (DDST) scheme where the training sequence is the sum of a known sequence $\boldsymbol{c}$ and an unknown data-dependent sequence $\boldsymbol{e} := -\mathbf{J}\boldsymbol{w}$.

The proposed ST scheme may seem similar to orthogonally multiplexing pilots and data tones in multicarrier (MC) systems [8]. However, in the proposed method, the symbol energy is spread over the entire bandwidth. Setting P DFT coefficients of the perturbed data $w(n) - e(n)$ to zero affects the symbols equally. Therefore, unlike MC systems, the proposed method for single carrier systems does not entail the waste of bandwidth. Indeed, for MC systems, only $N - P$ information-bearing symbols would be transmitted in each block. Both schemes involve a length $L - 1$ CP.

III. Channel Estimation and Training Design

At the receiver, channel estimation can be carried out as in the conventional ST scheme. A time-domain estimator based on synchronized averaging of the received signal was developed in [3], [5], and [6]. The same estimator can be obtained using the frequency domain [4], which will be used here since equalization will be carried out in this domain.

Since the DFT of the distorted data is identically zero at the pilot frequencies, we have $\tilde{x}(kQ) = \hat{H}(kQ)\tilde{c}(kQ), k = 0, \ldots, P-1$. To ensure consistent channel estimates based on this, at least L of the P pilot cycles of $\boldsymbol{c}$ should be nonzero. We estimate the frequency response of the channel at the pilot frequencies via

$$\hat{H}(kQ) = \tilde{x}(kQ)/\tilde{c}(kQ), \quad k = 0, \ldots, P-1.$$

The channel coefficient vector is estimated as

$$\hat{\boldsymbol{h}} = \frac{1}{\sqrt{P}}\mathbf{F}_{P,L}^{*}\hat{\boldsymbol{d}} \tag{5}$$

where $\hat{\boldsymbol{d}} = [\hat{H}(0), \hat{H}(Q), \ldots, \hat{H}((P-1)Q)]^T$.

The channel estimate in (5) is unbiased, and its mean square error (MSE) is given by

$$\mathsf{mse}(\hat{\boldsymbol{h}}) := E\left\{\sum_{l=0}^{L-1}|\hat{h}_l - h_l|^2\right\} = \frac{\sigma_v^2}{P}\text{Tr}\{\mathbf{F}_{P,L}^{*}\mathbf{C}^{-1}\mathbf{F}_{P,L}\}$$

where $\mathbf{C} = \text{diag}(|\tilde{c}(0)|^2, |\tilde{c}(Q)|^2, \ldots, |\tilde{c}((P-1)Q)|^2)$. Under the constraint of constant training power $(1/N)\sum_{n=0}^{N-1}|c(n)|^2 = \sigma_c^2$, the above MSE is minimized when $|\tilde{c}(kQ)|^2 = \sigma_c^2(N/P), k = 0, \ldots, P-1$. The MSE expression becomes

$$\mathsf{mse}(\hat{\boldsymbol{h}}) = \frac{L\sigma_v^2}{N\sigma_c^2}. \tag{6}$$

Further, since the above MSE is independent of P, and data distortion increases with P, P should be as small as possible, i.e., $P = L$. Note that the MSE on channel estimation is now independent of the unknown data, unlike that in existing ST-based methods [5], [6].

Since there are infinitely many periodic sequences for which the cycles are all equal in magnitude, sequences with minimum peak-to-average power ratio (PAR) are desirable. In fact, "ideal" sequences, i.e., optimal and constant envelope (i.e., unit PAR) sequences, exist for all values of P. One of these sequences is the chirp sequence $c(n) = \sigma_c \exp(j2\pi n(n+i)/P)$ with $i = 0$ when P is even, and $i = 1$ if P is odd [5].

IV. SYMBOL DETECTION

After the channel is estimated, we can remove the contribution of $\boldsymbol{c}$ from $\boldsymbol{x}$ by simply computing

$$z = (\mathbf{I} - \mathbf{J})\boldsymbol{x}.$$

In the frequency domain, this is equivalent to setting the DFT of $\boldsymbol{x}$ at the pilot frequencies to zero. Since both $\mathcal{H}$ and $\mathbf{J}$ are circulant and using the commutativity of circulant matrices, $\boldsymbol{z}$ can be expressed as

$$\boldsymbol{z} = \mathcal{H}(\mathbf{I} - \mathbf{J})\boldsymbol{w} + (\mathbf{I} - \mathbf{J})\boldsymbol{v}$$

where we have used the fact that $(\mathbf{I}-\mathbf{J})^2 = (\mathbf{I}-\mathbf{J})$. The additive noise $\breve{\boldsymbol{v}} = (\mathbf{I} - \mathbf{J})\boldsymbol{v}$ is now slightly colored. However, this color will fade away when Q is large and will therefore be ignored in what follows. Further, it is straightforward to show that the power of $\breve{\boldsymbol{v}}$ is $\breve{\sigma}_v^2 = \sigma_v^2(1 - 1/Q)$.

Since $\mathcal{H}$ is circulant, equalization can be carried out in the frequency domain, i.e., the equalized signal is given by

$$\boldsymbol{u} = \mathbf{F}_N^* \mathbf{G} \tilde{\boldsymbol{z}} \tag{7}$$

where $\tilde{\boldsymbol{z}}$ is obtained by setting the DFT of $\boldsymbol{x}$ at the pilot frequencies to zero, and $\mathbf{G}$ is an $(N \times N)$ diagonal matrix whose kth entry $G(k)$ is $G(k) = 1/\hat{H}(k)$ in the case of zero-forcing equalization and $G(k) = \hat{H}^*(k)/(|\hat{H}(k)|^2 + \breve{\sigma}_v^2)$ in the case of MMSE equalization.

Due to data distortion in the transmission, $\boldsymbol{u} \neq \boldsymbol{w}$, even in the absence of channel estimation error and noise. Indeed, in this ideal scenario, $\boldsymbol{u} = (\mathbf{I} - \mathbf{J})\boldsymbol{w}$. Since $(\mathbf{I} - \mathbf{J})$ is singular, $\boldsymbol{w}$ cannot be recovered *linearly*. However, using the fact that the data symbols are drawn from a finite alphabet and that $\mathbf{J}\boldsymbol{w}$ is small compared to $\boldsymbol{w}$, symbol detection can be accomplished by finding the vector of constellation points $\boldsymbol{w}$ that minimizes the Euclidian distance between $\boldsymbol{u}$ and $(\mathbf{I} - \mathbf{J})\boldsymbol{w}$. However, this sequence detection scheme is computationally cumbersome and will, therefore, not be considered here. Instead, we propose the following iterative symbol-by-symbol detection scheme.

The symbol-by-symbol detection algorithm is initialized by treating $\mathbf{J}\boldsymbol{w}$ as an extra additive noise and considering $\boldsymbol{u}$ in (7) as a soft detector of $\boldsymbol{w}$; the initial hard detector of $\boldsymbol{w}$ is given by

$$\bar{\boldsymbol{w}}^{(0)} = \lfloor \boldsymbol{u} \rfloor$$

where $\lfloor \boldsymbol{u} \rfloor$ denotes the vector of constellation points that are closest to the vector $\boldsymbol{u}$. The detected symbols are used to estimate $\mathbf{J}\boldsymbol{w}$ to be used in the next iteration. The detected symbols at the ith iteration are given by

$$\bar{\boldsymbol{w}}^{(i)} = \left\lfloor \boldsymbol{u} + \mathbf{J}\bar{\boldsymbol{w}}^{(i-1)} \right\rfloor.$$

As we will see in the next section, most of the gain in symbol detection performance over existing ST-based methods is obtained in the very first iteration.

V. SIMULATION RESULTS

We compare the proposed DDST scheme with both the conventional ST and TDM schemes in terms of channel estimation performance and symbol error rate (SER). The length of the data block is set to $N = 240$. The channel is randomly generated at each Monte Carlo run and is assumed to be Rayleigh with length $L = 6$; the coefficients are uncorrelated, and their powers are given by the exponential delay profile $E\{|h_\ell|^2\} = \exp(-0.2\ell)$. The periodic sequence is chosen to be optimal according to Section III, and its power is set to be 10% of the total transmitted power, σ_s^2; its period is set to $P = L$. The data symbols are drawn from quadrature phase shift keying constellations.

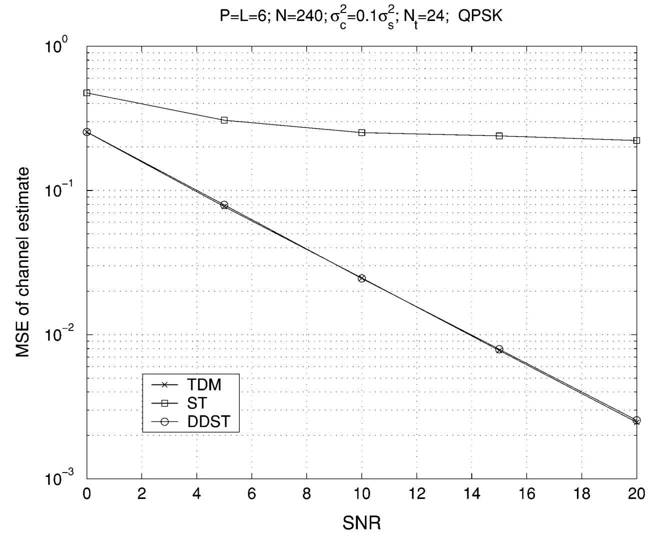

Fig. 1. Empirical Mean square error of channel estimates.

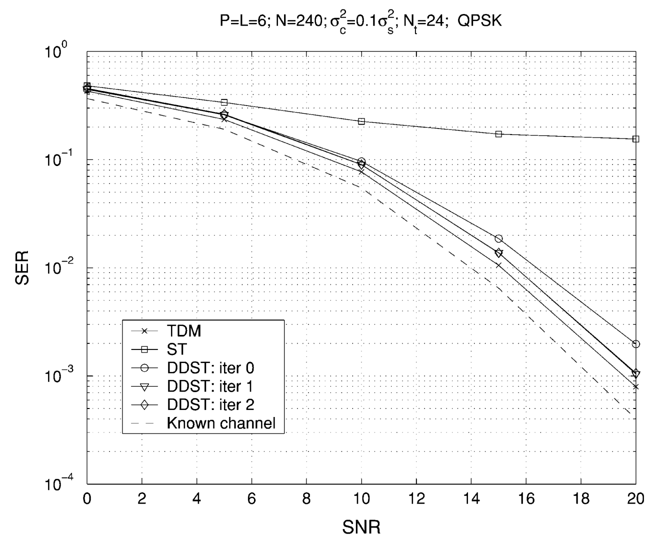

Fig. 2. Symbol error rate for different PAT schemes.

The channel estimation performance of the proposed DDST scheme is the same as that of the TDM scheme if the percentage of power used in the DDST scheme is the same as that of pilot symbols used in the TDM scheme, as is verified in Fig. 1, where the number of training symbols in the TDM scheme is $N_t = 24$. Fig. 2 displays the SER for the three schemes as well as that when the channel is exactly known at the receiver, and full power is assigned to the data symbols. It is seen that the SER for the DDST scheme is close to that for the TDM scheme. The latter, however, consumes 10% of the bandwidth. Full investigation of the effects of N, constellation size, and

σ_c^2 on performance will be carried out in a longer version of this paper.

VI. Conclusion

In this letter, we have presented a new pilot-assisted transmission scheme to estimate frequency-selective channels. The scheme consists of setting a few points of the DFT of the data to known values. This operation can be easily implemented in the time domain when these DFT points are equispaced. The channel is estimated using the DFT of the received signal at these pilot frequencies. Detection of the distorted symbols is carried out using an iterative scheme. The proposed method was shown to outperform existing methods based on superimposed training and compares well with the time-division multiplexed training scheme, in terms of both the performance of the channel estimator and the symbol error rate. Unlike the TDM scheme, the proposed method does not require bandwidth for training.

References

[1] S. Adireddy, L. Tong, and H. Viswanathan, "Optimal placement of training for frequency-slective block-fading channels," *IEEE Trans. Inf. Theory*, vol. 48, no. 8, pp. 2338–2352, Aug. 2002.

[2] B. Farhang-Boroujeny, "Experimental study of semi-blind channel identification/equalization through pilot signals," in *Proc. ICSP*, Beijing, China, 1996, pp. 618–621.

[3] G. T. Zhou, M. Viberg, and T. McKelvey, "First-order statistical method for channel estimation," *IEEE Signal Process. Lett.*, vol. 10, no. 3, pp. 57–60, Mar. 2003.

[4] J. Tugnait and W. Luo, "On channel estimation using superimposed training and first-order statistics," *IEEE Commun. Lett.*, vol. 7, no. 9, pp. 413–415, Sep. 2003.

[5] A. G. Orozco-Lugo, M. M. Lara, and D. C. McLernon, "Channel estimation using implicit training," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 240–254, Jan. 2004.

[6] M. Ghogho and A. Swami, "Improved channel estimation using superimposed training," in *Proc. SPAWC*, Lisbon, Portugal, Jul. 2004.

[7] D. McLernon, E. Alameda-Hernandez, and A. G. Orozco-Lugo, "Implicitly-trained channel estimation and equalization with zero mean input data packets" (in ssity), in *Proc. IEEE ISSPIT*, Rome, Italy, Dec. 2004.

[8] S. Ohno and G. B. Giannakis, "Optimal training and redundant precoding for block transmissions with application to wirelesss OFDM," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 2113–2123, Dec. 2002.