Fig. 9. RNMF convergence: Relative error versus Iteration number.

V. CONCLUSION

In this paper, we proposed a robust approach to nonnegative matrix factorization of a spectral library. The proposed method is formulated as an energy minimization problem whose solution is achieved by updating alternatively two equations. Unlike the K–L method, the robust nonnegative matrix factorization technique is based on a robust cost function, resistant to outliers, and generates nonnegative basis functions which balances the logical attractiveness of measurement functions against their physical feasibility. We have successfully tested the robust nonnegative factorization algorithm on a library of reflectance spectra, and the experimental results clearly show that the proposed technique outperforms the current reconstruction methods including PCA, NNSC, NMFSC, NMF, and cNMF.

REFERENCES

[1] J. M. Chalmers and P. R. Griffiths, *Handbook of Vibrational Spectroscopy*. New York: Wiley, 2002.
[2] I. T. Jolliffe, *Principal Component Analysis*. New York: Springer-Verlag, 1986.
[3] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate Analysis*. London, U.K.: Academic, 1979.
[4] D. Lee and H. Seung, "Learning the parts of objects by nonnegative matrix factorization," *Nature*, vol. 401, pp. 788–791, 1999.
[5] ——, "Algorithms for nonnegative matrix factorization," *Adv. Neural Inf. Proc. Syst.*, vol. 13, pp. 556–562, 2000.
[6] F. Shahnaz, M. Berry, V. P. Pauca, and R. Plemmons, "Document clustering using nonnegative matrix factorization," *Information Processing and Management*, vol. 42, no. 2, pp. 556–562, Mar. 2005.
[7] J. Piper, V. P. Pauca, R. Plemmons, and M. Giffin, "Object characterization from spectral data using nonnegative factorization and information theory," presented at the Amos Technical Conf., Maui, Sep. 2004.
[8] P. O. Hoyer, "Nonnegative sparse coding," in *Proc. IEEE Workshop Neural Networks Signal Processing*, Sep. 2002, pp. 557–565.
[9] ——, "Non-negative Matrix Factorization with sparseness constraints," *J. Mach. Learn. Res.*, no. 5, pp. 1457–1469, 2004.
[10] C. Samson, L. Blanc-Féraud, G. Aubert, and J. Zerubia, "A variational model for image classification and restoration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 5, pp. 460–472, 2000.
[11] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*. Philadelphia, PA: SIAM, 1995, vol. 16, Frontiers in Applied Mathematics.
[12] P. Huber, *Robust Statistics*. New York: Wiley, 1981.
[13] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust Statistics: The Approach Based on Influence Functions*. New York: Wiley, 1986.
[14] M. J. Black and A. Rangarajan, "On the unification of line processes, outlier rejection, and robust statistics with applications in early vision," *Int. J. Comput. Vis.*, vol. 19, pp. 57–92, Jul. 1996.
[15] M. J. Black, G. Sapiro, D. H. Marimont, and D. Heeger, "Robust anisotropic diffusion," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 421–432, Mar. 1998.
[16] H. Krim and I. C. Schick, "Minimax description length for signal denoising and optimized representation," *IEEE Trans. Inf. Theory*, vol. 45, no. 3, pp. 898–908, Apr. 1999.
[17] A. B. Hamza and H. Krim, "Image denoising: a nonlinear robust statistical approach," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3045–3054, Dec. 2001.
[18] M. Turk and A. Pentland, "Face recognition using eigenfaces," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1991, pp. 586–591.
[19] P. Sajda, S. Du, T. R. Brown, R. Stoyanova, D. C. Shungu, M. Xiangling, and L. C. Parra, "Nonnegative matrix factorization for rapid recovery of constituent spectra in magnetic resonance chemical shift imaging of the brain," *IEEE Trans. Med. Imag.*, vol. 23, no. 12, pp. 1453–1465, Dec. 2004.

Low-Complexity Equalization of Time-Varying Channels With Precoding

Zijian Tang, *Student Member, IEEE*, and
Geert Leus, *Senior Member, IEEE*

Abstract—This correspondence deals with time-varying (TV) single-input multiple-output (SIMO) channels, which are both frequency selective (due to high data rate) and time selective (due to mobility). A complex exponential basis expansion model (CE-BEM) is used to model the channel. We consider a block transmission system, where on the transmit side a precoder is employed to enable the maximum available diversity for a CE-BEM channel. After direct decoding on the receive side, the resulting channel resembles a finite-impulse-response (FIR) filter on both block and symbol level. We therefore propose an equalizer that bears a structure analogous to the effective channel. In comparison with a standard block minimum mean-square error decision-feedback equalizer (BMMSE-DFE) that involves the inversion of a large-size matrix, the proposed parametric equalizer renders a similar performance but at a lower computational cost if there are multiple outputs present. Another contribution of this correspondence is a semiblind algorithm to estimate this equalizer when the channel state information is not available: the equalizer taps and the information symbol estimates are refined recursively by means of normalized least-mean-squares (NLMS) adaptation.

Index Terms—Basis expansion model (BEM), diversity, time-varying (TV) channel.

I. INTRODUCTION

In high-data-rate mobile communication systems, the relative velocity between the transmitter and the receiver gives rise to a Doppler

spread and makes the channel time varying (TV). Jakes' model is often adopted if there exist a large number of scatterers in the vicinity of the mobile [1]. For simplicity, we adopt a parsimonious model called complex exponential basis expansion model [(CE-BEM), see [2] among others] to approximate the channel's time variation in a discrete form. This model utilizes truncated complex Fourier series as expansion basis functions, i.e., the lth tap at the nth time interval of a TV channel $h[n; l]$ is approximated as

$$h[n; l] \approx \sum_{q=0}^{Q} h_{q,l} e^{j2\pi qn/P} \qquad (1)$$

where Q indicates the order of the CE-BEM and P defines the size of the observation window, within which the CE-BEM coefficients $h_{q,l}$ are assumed to remain constant. In comparison with other parsimonious models, e.g., the discrete prolate spheroidal BEM (DPS-BEM [4]), the CE-BEM inflicts a larger modeling mismatch, especially on the edges of the observation window. However, the average modeling mismatch can be made arbitrarily small by increasing Q [6] at the expense of system complexity.

The CE-BEM expression has the unique property that it has a finite-impulse-response (FIR) structure in both time domain and frequency domain [9]. Benefiting from this time–frequency duality, the authors in [8] give a general analytical formulation of the delay-Doppler diversity for a precoded scheme. However, the precoder in [8] destroys the FIR structure of the channel and makes a standard maximum-likelihood (ML) equalizer extremely expensive. Suboptimal equalizers such as a sphere decoder (SD) [14], or a block minimum mean-square error decision-feedback equalizer (BMMSE-DFE) [13] have a complexity that is exponential or polynomial in P, which could still be too high. [27, Sec. IV-A] presents a decoder for one special case of [8] that makes smart use of the commutability between the channel and the (de-)precoder. The resulting effective channel, which associates the decoded samples directly with the transmitted data symbols, can then be characterized by a two-level (2-L) FIR filter incorporating certain frequency shifts.

In this correspondence, we present a low-complexity DFE to equalize such a channel. Motivated by the fact that a DFE with time-invariant FIR feedforward and feedback filters can accurately equalize a time-invariant FIR channel at a low complexity cost, we here also consider a DFE with feedforward and feedback filters that both have the same structure as the effective channel, i.e., a 2-L FIR incorporating certain frequency shifts. We will show how to acquire the DFE coefficients in two cases. In the first case, we need the channel (CE-BEM) knowledge; in the second case, we estimate the DFE coefficients semiblindly following an adaptive approach.

Notation: We use upper (lower) boldface letters to denote matrices (column vectors). $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^\dagger$ represent conjugate, transpose, complex conjugate transpose (Hermitian), and pseudoinverse, respectively. $\mathcal{E}\{\cdot\}$ stands for the expected value. $\otimes$ represents the Kronecker product. $\lceil x \rceil$ represents the smallest integer that is greater or equal to x. We denote the $N \times N$ identity matrix as $\mathbf{I}_N$, and the $M \times N$ all zero matrix as $\mathbf{0}_{M \times N}$.

II. SYSTEM MODEL

Let us consider a block transmission system with one input and A outputs [single-input multiple-output (SIMO)]. Let us further assume that all the channels are FIR with a maximum order of L. If $\mathbf{s}$ represents an $(M - L)(N - K) \times 1$ data symbol vector and $\mathbf{y}^{(a)}$ an $MN \times 1$ received sample vector, the corresponding I/O relationship for the ath channel with precoding is described by [for details, see [8], in particular (16)]

$$\mathbf{y}^{(a)} = \mathbf{H}^{(a)} \mathbf{\Theta} \mathbf{s} + \mathbf{w}^{(a)} \qquad (2)$$

where $\mathbf{w}^{(a)}$ denotes the additive noise. The $MN \times MN$ matrix $\mathbf{H}^{(a)}$ stands for the ath channel matrix, given by $[\mathbf{H}^{(a)}]_{k,m} = h^{(a)}[k, k - m]$. If we assume that all the channels can be modeled by a CE-BEM as defined in (1) within an observation window of length $P = MN$, we have $[\mathbf{H}^{(a)}]_{k,m} := \sum_{q=0}^{Q} h_{q,k-m}^{(a)} e^{j2\pi qk/(MN)}$. The $MN \times (M - L)(N - K)$ precoder $\mathbf{\Theta}$ is defined as in [8] [27, Sec. IV-A]

$$\mathbf{\Theta} := \left(\mathbf{F}_N^H \mathbf{T}_1 \right) \otimes \mathbf{T}_2 \qquad (3)$$

where $\mathbf{F}_N$ is a unitary N-point DFT matrix with entries $[\mathbf{F}_N]_{p,q} := (1/\sqrt{N})e^{-j(2\pi/N)(p-1)(q-1)}$, $\mathbf{T}_1 := [\mathbf{I}_{N-K}, \mathbf{0}_{(N-K)\times K}]^T$, and $\mathbf{T}_2 := [\mathbf{I}_{M-L}, \mathbf{0}_{(M-L)\times L}]^T$. It is proved in [8] that if $Q \geq K$, then $\mathbf{\Theta}$ enables a diversity order up to $(L + 1)(K + 1)$ under the assumption of a CE-BEM channel (1). For such a precoder, [27, Sec. IV-A] proposes a decoder $\mathbf{\Psi} := \mathbf{F}_N \otimes \mathbf{I}_M$. As a result, the ath channel output after decoding becomes

$$\bar{\mathbf{y}}^{(a)} = \mathbf{\Psi} \mathbf{y}^{(a)} = \mathbf{\Psi} \mathbf{H}^{(a)} \mathbf{\Theta} \mathbf{s} + \bar{\mathbf{w}}^{(a)} \qquad (4)$$

with $\bar{\mathbf{w}}^{(a)} := \mathbf{\Psi} \mathbf{w}^{(a)}$. Next, we partition the data symbol sequence $\mathbf{s}$ into $N - K$ blocks, i.e., the ith data block $\mathbf{s}(i)$ contains the symbols $[\mathbf{s}]_{i(M-L)+1}$ to $[\mathbf{s}]_{(i+1)(M-L)}$. Likewise, the decoded sequence $\bar{\mathbf{y}}^{(a)}$ is partitioned into N blocks with the jth block $\bar{\mathbf{y}}^{(a)}(j)$ containing the symbols $[\bar{\mathbf{y}}^{(a)}]_{jM+1}$ to $[\bar{\mathbf{y}}^{(a)}]_{(j+1)M}$. Accordingly, we are able to split the expression in (4) up into smaller blocks[1]

$$\bar{\mathbf{y}}^{(a)}(i) = \sum_{q=0}^{Q} \bar{\mathbf{H}}_q^{(a)} \mathbf{s}(i - q) + \bar{\mathbf{w}}^{(a)}(i) \qquad (5)$$

where $\bar{\mathbf{w}}^{(a)}(i)$ is similarly defined as $\bar{\mathbf{y}}^{(a)}(i)$, and $\bar{\mathbf{H}}_q^{(a)} := \mathbf{\Delta}_M^q \mathbf{H}_q^{(a)}$ with $\mathbf{H}_q^{(a)}$ being an $M \times (M - L)$ Toeplitz matrix $[\mathbf{H}_q^{(a)}]_{k,m} := h_{q,k-m}^{(a)}$ and $\mathbf{\Delta}_\kappa^q$ being a diagonal frequency shift matrix $\mathbf{\Delta}_\kappa^q := \text{diag}([1, e^{j(2\pi q/MN)}, \cdots, e^{j(2\pi q/MN)(\kappa-1)}])$. Note that in (5), we assume $\bar{\mathbf{y}}^{(a)}(i) = \bar{\mathbf{w}}^{(a)}(i) = \mathbf{0}_{M \times 1}$ for $i < 0$ and $i \geq N$, and $\mathbf{s}(i) = \mathbf{s}(N + i)$ for $-Q \leq i < 0$ and $\mathbf{s}(i) = \mathbf{0}_{(M-L)\times 1}$ for $i \geq N - K$.

III. PARAMETRIC DFE

For the standard SIMO DFE, which contains a set of A linear feedforward filters (FFs), and a feedback filter (BF), we can readily apply a classical BMMSE-DFE on (4), with a design complexity of $\mathcal{O}((M - L)^3(N - K)^3)$ and an implementation complexity of $AMN(M - L)(N - K) + (M - L)^2(N - K)^2/2$ complex multiply/add (MA) operations per data symbol vector. Note that since we are dealing with TV channels, both design and implementation steps have to be carried out in every observation window.

However, it is easy to observe from (5) that the effective channel takes on a 2-L FIR structure incorporating certain frequency shifts, as is illustrated in Fig. 1: on the block level, the channel can be viewed as an FIR filter with $\bar{\mathbf{H}}_q^{(a)}$ being its block-level taps; on the symbol level, except for the frequency shifts captured in $\mathbf{\Delta}_M^q$, there is again an FIR filter with $h_{q,l}^{(a)}$ being its symbol-level taps. The fact that a time-invariant FIR channel can be accurately equalized by a DFE with time-invariant FIR FFs and a BF at a low cost motivates the consideration to similarly structure the DFE in the present context using a 2-L FIR incorporating certain frequency shifts for both the FFs and BF. An important observation is that the induced design and implementation complexity will generally decrease, at the cost of a slight BER degradation, as we will show later on.

---

[1] The proof can be found in [27, Sec. IV-A] with the difference that we deal in this correspondence with the more general case $Q \geq K$.
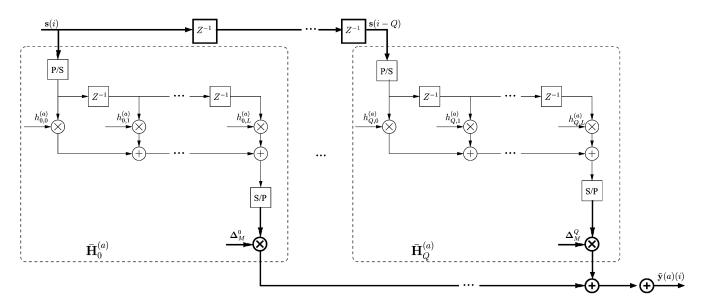
Fig. 1. Block diagram of the ath effective channel.

A. Structure of the DFE

In line with the aforementioned ideas, we design the FFs and the BF using a 2-L FIR incorporating certain frequency shifts, i.e., we estimate $\mathbf{s}(i)$ as

$$\hat{\mathbf{s}}(i) = \sum_{a=1}^{A} \sum_{q_e=-Q_e}^{0} \bar{\mathbf{F}}_{q_e}^{(a)T} \bar{\mathbf{y}}^{(a)}(i-d_Q-q_e) - \sum_{q_b=0}^{Q_b} \bar{\mathbf{B}}_{q_b} \check{\mathbf{s}}(i-q_b) \quad (6)$$

with $\check{\mathbf{s}}(i)$ standing for the quantized estimate $\check{\mathbf{s}}(i) = \mathcal{Q}(\hat{\mathbf{s}}(i))$. In (6), the first summation on the right-hand side represents the operation of the FFs and the second summation the operation of the BF.

We notice that each FF is equipped with Q_e+1 block-level taps $\bar{\mathbf{F}}_{q_e}^{(a)}$ of size $M \times (M-L)$. Like in the time-invariant case where a delay is often incorporated in the equalizer, we introduce here a parameter d_Q to denote the block-level delay for the FF, which must satisfy $-Q \leq d_Q \leq Q_e$ to ensure that $\mathbf{s}(i)$ is present in the equalizer input. To stick to the 2-L FIR structure, we construct each block-level tap $\bar{\mathbf{F}}_{q_e}^{(a)}$ to be the product of a frequency-shift matrix and a Toeplitz matrix, i.e.,

$$\bar{\mathbf{F}}_{q_e}^{(a)} := \mathbf{\Delta}_M^{q_e+d_Q} \mathbf{F}_{q_e}^{(a)} \quad (7)$$

where $[\mathbf{F}_{q_e}^{(a)}]_{k,m} = [\mathbf{f}_{q_e}^{(a)}]_{k-m+d_L+1}$, with $\mathbf{f}_{q_e}^{(a)}$ collecting all the L_e+1 symbol-level taps of the q_e^{th} FIR filter for the ath channel output. d_L is the symbol-level delay and should satisfy $-L \leq d_L \leq L_e$ to avoid all-zero columns in $\mathbf{F}_{q_e}^{(a)}$.

A similar structure is imposed upon the BF: it is assigned Q_b+1 block-level taps $\bar{\mathbf{B}}_{q_b}$ of size $(M-L) \times (M-L)$, each of which is again the product of a frequency-shift matrix and a Toeplitz matrix, i.e.,

$$\bar{\mathbf{B}}_{q_b} := \mathbf{\Delta}_{M-L}^{q_b} \mathbf{B}_{q_b} \quad (8)$$

where $[\mathbf{B}_{q_b}]_{k,m} = [\mathbf{b}_{q_b}]_{k-m+d_b+1}$, with $\mathbf{b}_{q_b}$ collecting all the L_b+1 symbol-level taps of the q_b^{th} FIR filter. d_b is the delay parameter intrinsic to the BF and should satisfy $0 \leq d_b \leq L_b$ to avoid all-zero rows in $\mathbf{B}_{q_b}$. Because the BF must take the channel causality into account, we require $[\mathbf{b}_0]_i = 0$ for $i = 1, \cdots, d_b+1$.

The proposed DFE is therefore characterized by $A(L_e+1)(Q_e+1) + Q_b(L_b+1) + L_b - d_b$ coefficients. From the description above,

we understand that the implementation complexity is about $AM(N-K)(Q_e+1)(L_e+2) + (M-L)(N-K)(Q_b+1)(L_b+2)$ MA operations per data vector. In general, it is smaller than that of the classical BMMSE-DFE that entails $AMN(M-L)(N-K) + (M-L)^2(N-K)^2/2$ MA operations per data vector.

B. Assuming Channel (CE-BEM) Knowledge

If the past decisions are correct, it is straightforward to transform (6) into

$$\hat{\mathbf{s}}(i)^T = \sum_{a=1}^{A} \sum_{q_e=-Q_e}^{0} \sum_{q=0}^{Q} \left(\mathbf{\Delta}_M^{q'} \mathbf{H}_q^{(a)} \mathbf{s}(i-q') \right)^T \mathbf{F}_{q_e}^{(a)}$$
$$+ \sum_{a=1}^{A} \sum_{q_e=-Q_e}^{0} \left(\mathbf{\Delta}_M^{q''} \bar{\mathbf{w}}^{(a)}(i-q'') \right)^T \mathbf{F}_{q_e}^{(a)}$$
$$- \sum_{q_b=0}^{Q_b} \mathbf{s}(i-q_b)^T \mathbf{B}_{q_b}^T \mathbf{\Delta}_{M-L}^{q_b} \quad (9)$$

with $q' := d_Q + q + q_e$ and $q'' := d_Q + q_e$. We need to express the above equation as an explicit function of $\mathbf{f}_{q_e}^{(a)}$ and $\mathbf{b}_{q_b}$. First, it can be derived after some mathematical manipulation that

$$\left(\mathbf{\Delta}_M^{q'} \mathbf{H}_q^{(a)} \mathbf{s}(i-q') \right)^T \mathbf{F}_{q_e}^{(a)}$$
$$= e^{-j\omega_{q'}} \mathbf{f}_{q_e}^{(a)T} \mathbf{\Delta}_{L_e+1}^{q'} \mathcal{H}_q^{(a)} \mathcal{S}(i-q') \mathbf{\Delta}_{M-L}^{q'} \quad (10)$$

where $\omega_q := 2\pi q d_L/P$. $\mathcal{H}_q^{(a)}$ is a Toeplitz matrix with $[h_{q,L+1}^{(a)}, \mathbf{0}_{1 \times L_e}]^T$ as its first column and $[h_{q,L+1}^{(a)}, \cdots, h_{q,1}^{(a)}, \mathbf{0}_{1 \times L_e}]$ as its first row, and $\mathcal{S}(i)$ denotes an $(L_e + L + 1) \times (M-L)$ Hankel matrix $[\mathcal{S}(i)]_{k,m} = [\mathbf{s}(i)]_{k+m-L-d_L-1}$.

Likewise, we can show that

$$\left(\mathbf{\Delta}_M^{q''} \bar{\mathbf{w}}^{(a)}(i-q'') \right)^T \mathbf{F}_{q_e}^{(a)}$$
$$= e^{-j\omega_{q''}} \mathbf{f}_{q_e}^{(a)T} \left(\mathbf{\Delta}_{L_e+1}^{q''} \mathcal{W}^{(a)}(i-q'') \mathbf{\Delta}_{M-L}^{q''} \right) \quad (11)$$

where $\mathcal{W}^{(a)}(i)$ is an $(L_e + 1) \times (M-L)$ Hankel matrix with $[\mathcal{W}^{(a)}(i)]_{k,m} := [\bar{\mathbf{w}}^{(a)}(i)]_{k+m-d_L-1}$. Finally, it holds that

$$\mathbf{s}(i-q_b)^T \mathbf{B}_{q_b}^T \mathbf{\Delta}_{M-L}^{q_b} = \mathbf{b}_{q_b}^T \mathbf{P} \mathcal{S}(i-q_b) \mathbf{\Delta}_{M-L}^{q_b} \quad (12)$$

with $\mathbf{P} := [\mathbf{0}_{(L_b+1)\times(L-L_b+d_L+d_b)}, \mathbf{I}_{L_b+1}, \mathbf{0}_{(L_b+1)\times(L_e-d_L-d_b)}]$.

Substituting (10), (11), and (12) in (9), we obtain

$$\hat{\mathbf{s}}^T(i) = \sum_{a=1}^{A}\sum_{q_e=-Q_e}^{0}\sum_{q=0}^{Q}\mathbf{f}_{q_e}^{(a)T}$$
$$\times\left(e^{-j\omega_{q'}}\Delta_{L_e+1}^{q'}\mathcal{H}_q^{(a)}\mathcal{S}(i-q')\Delta_{M-L}^{q'}\right)$$
$$+\sum_{a=1}^{A}\sum_{q_e=-Q_e}^{0}\mathbf{f}_{q_e}^{(a)T}$$
$$\times\left(e^{-j\omega_{q''}}\Delta_{L_e+1}^{q''}\mathcal{W}^{(a)}(i-q'')\Delta_{M-L}^{q''}\right)$$
$$-\sum_{q_b=0}^{Q_b}\mathbf{b}_{q_b}^T\mathbf{P}\mathcal{S}(i-q_b)\Delta_{M-L}^{q_b}. \qquad (13)$$

To avoid the multiple summations, we define $\mathbf{f}^{(a)} := [\mathbf{f}_0^{(a)T}, \cdots, \mathbf{f}_{-Q_e}^{(a)T}]^T$, $\mathbf{b} := [\mathbf{b}_0^T, \cdots, \mathbf{b}_{Q_b}^T]^T$, and

$$\boldsymbol{\Theta}$$
$$:= \operatorname{diag}\left\{[e^{-j\omega_{d_Q+Q}}, \cdots, e^{-j\omega_{d_Q-Q_e}}]\right\}\otimes\mathbf{I}_{L_e+L+1}$$

$$\tilde{\mathcal{H}}^{(a)}$$
$$:=\begin{bmatrix}\Delta_{L_e+1}^{d_Q+Q}\mathcal{H}_Q^{(a)} & \cdots & \Delta_{L_e+1}^{d_Q}\mathcal{H}_0^{(a)} \\ & \ddots & & \ddots \\ & & \Delta_{L_e+1}^{d_Q+Q-Q_e}\mathcal{H}_Q^{(a)} & \cdots & \Delta_{L_e+1}^{d_Q-Q_e}\mathcal{H}_0^{(a)}\end{bmatrix}$$

$$\tilde{\mathcal{S}}(i)$$
$$:=\begin{bmatrix}\mathcal{S}(i-d_Q-Q)\Delta_{M-L}^{d_Q+Q} \\ \vdots \\ \mathcal{S}(i-d_Q+Q_e)\Delta_{M-L}^{d_Q-Q_e}\end{bmatrix}$$

$$\tilde{\mathcal{W}}^{(a)}(i)$$
$$:=\begin{bmatrix}e^{-j\omega_{d_Q}}\Delta_{L_e+1}^{d_Q}\mathcal{W}^{(a)}(i-d_Q)\Delta_{M-L}^{d_Q} \\ \vdots \\ e^{-j\omega_{d_Q-Q_e}}\Delta_{L_e+1}^{d_Q-Q_e}\mathcal{W}^{(a)}(i-d_Q+Q_e)\Delta_{M-L}^{d_Q-Q_e}\end{bmatrix}$$

such that (13) can be written in a compact form as

$$\hat{\mathbf{s}}^T(i) = \sum_{a=1}^{A}\mathbf{f}^{(a)T}\left(\tilde{\mathcal{H}}^{(a)}\boldsymbol{\Theta}\tilde{\mathcal{S}}(i)+\tilde{\mathcal{W}}^{(a)}(i)\right) - \mathbf{b}^T\tilde{\mathbf{P}}\tilde{\mathcal{S}}(i)$$
$$= \mathbf{f}^T\left(\tilde{\mathcal{H}}\boldsymbol{\Theta}\tilde{\mathcal{S}}(i)+\tilde{\mathcal{W}}(i)\right) - \mathbf{b}^T\tilde{\mathbf{P}}\tilde{\mathcal{S}}(i) \qquad (14)$$

with $\mathbf{f} := [\mathbf{f}^{(1)T}, \cdots, \mathbf{f}^{(A)T}]^T$; $\tilde{\mathcal{H}} := [\tilde{\mathcal{H}}^{(1)T}, \cdots, \tilde{\mathcal{H}}^{(A)T}]^T$; $\tilde{\mathcal{W}}(i) := [\tilde{\mathcal{W}}^{(1)T}(i), \cdots, \tilde{\mathcal{W}}^{(A)T}(i)]^T$; and $\tilde{\mathbf{P}} := [\mathbf{0}_{(L_b+1)(Q_b+1)\times(d_Q+Q-Q_b)(L_e+1)}, \mathbf{I}_{Q_b+1}\otimes \mathbf{P}, \mathbf{0}_{(L_b+1)(Q_b+1)\times(Q_e-d_Q)(L_e+L+1)}]$.

The MMSE solution for the DFE coefficients is found by minimizing the mean-square error (MSE) $\mathcal{J} := \mathcal{E}\{|\hat{\mathbf{s}}(i)-\mathbf{s}(i)|^2\}$. With the observation that $\mathbf{s}^T(i) = \mathbf{e}^T\tilde{\mathcal{S}}(i)$, where $\mathbf{e}$ is an $(L_e+L+1)(Q+Q_e+1)\times 1$ unit vector with a one in the position $(d_Q+Q)(L_e+L+1)+L+d_L+1$, we can derive the MMSE solution in a similar way as done in [13], as follows:

$$\mathbf{b}_{\mathrm{MMSE}} = \tilde{\mathbf{P}}\left(\frac{\mathbf{R}_{\mathrm{MMSE}}\mathbf{e}}{\mathbf{e}^T\mathbf{R}_{\mathrm{MMSE}}\mathbf{e}}\right)^* - \tilde{\mathbf{P}}\mathbf{e}$$
$$\mathbf{R}_{\mathrm{MMSE}} = \boldsymbol{\Theta}^H\tilde{\mathcal{H}}^H\tilde{\mathbf{R}}_{\mathcal{W}}^{-1}\tilde{\mathcal{H}}\boldsymbol{\Theta}+\tilde{\mathbf{R}}_{\mathcal{S}}^{-1}$$
$$\mathbf{f}_{\mathrm{MMSE}}^T = \left(\mathbf{b}_{\mathrm{MMSE}}^T\tilde{\mathbf{P}}+\mathbf{e}^T\right)\mathbf{R}_{\mathrm{MMSE}}^{\dagger}\boldsymbol{\Theta}^H\tilde{\mathcal{H}}^H\tilde{\mathbf{R}}_{\mathcal{W}}^{-1} \qquad (15)$$

where $\tilde{\mathbf{R}}_{\mathcal{S}}$ denotes the data covariance matrix $\tilde{\mathbf{R}}_{\mathcal{S}} := \mathcal{E}\{\tilde{\mathcal{S}}(i)\tilde{\mathcal{S}}^H(i)\}$ and $\tilde{\mathbf{R}}_{\mathcal{W}}$ denotes the noise covariance matrix $\tilde{\mathbf{R}}_{\mathcal{W}} := \mathcal{E}\{\tilde{\mathcal{W}}(i)\tilde{\mathcal{W}}^H(i)\}$, both of which are assumed to be known at the receiver. It is easy to derive the expressions for $\tilde{\mathbf{R}}_{\mathcal{S}}$ and $\tilde{\mathbf{R}}_{\mathcal{W}}$, but we have to omit them here due to space considerations.

From the expressions in (15), it can be seen that most computational effort is invested in inverting the matrix $\mathbf{R}_{\mathrm{MMSE}}$ with a design complexity of $\mathcal{O}((L+L_e+1)^3(Q+Q_e+1)^3)$. Again, this is cheaper than the classical BMMSE-DFE, which requires $\mathcal{O}((M-L)^3(N-K)^3)$ per data vector.

C. Direct Semiblind Equalization

In the previous section, we have seen how to obtain the DFE coefficients based on channel (CE-BEM) knowledge, which is often not available in practice. In that case, direct equalizer estimation can be appealing because it skips the intermediate channel estimation step [23]. In this correspondence, we will follow this approach using an adaptive algorithm [28].

To estimate $\mathbf{s}(i)$, let us use $\boldsymbol{\mathcal{U}}(i)$ to denote the corresponding FF input in (6). Now that the channel is transparent to the receiver, we have to express $\boldsymbol{\mathcal{U}}(i)$ as a function of $\bar{\mathbf{y}}^{(a)}(i)$

$$\boldsymbol{\mathcal{U}}(i) = \begin{bmatrix}e^{-j\omega_{d_Q}}\Delta_{L_e+1}^{d_Q}\boldsymbol{\mathcal{Y}}^{(1)}(i-d_Q)\Delta_{M-L}^{d_Q} \\ \vdots \\ e^{-j\omega_{d_Q-Q_e}}\Delta_{L_e+1}^{d_Q-Q_e}\boldsymbol{\mathcal{Y}}^{(1)}(i-d_Q+Q_e)\Delta_{M-L}^{d_Q-Q_e} \\ \vdots \\ e^{-j\omega_{d_Q}}\Delta_{L_e+1}^{d_Q}\boldsymbol{\mathcal{Y}}^{(A)}(i-d_Q)\Delta_{M-L}^{d_Q} \\ \vdots \\ e^{-j\omega_{d_Q-Q_e}}\Delta_{L_e+1}^{d_Q-Q_e}\boldsymbol{\mathcal{Y}}^{(A)}(i-d_Q+Q_e)\Delta_{M-L}^{d_Q-Q_e}\end{bmatrix}. \qquad (16)$$

Here, $\boldsymbol{\mathcal{Y}}^{(a)}(i)$ denotes an $(L_e+1)\times(M-L)$ Hankel matrix with $[\boldsymbol{\mathcal{Y}}^{(a)}(i)]_{k,m} := [\bar{\mathbf{y}}^{(a)}(i)]_{k+m-d_L-1}$. Including the corresponding BF input $\tilde{\mathbf{P}}\mathcal{S}(i)$, we obtain the following estimate prior to quantization:

$$\hat{\mathbf{s}}(i) = \tilde{\boldsymbol{\mathcal{U}}}(i)\mathbf{v} \qquad (17)$$

where $\tilde{\boldsymbol{\mathcal{U}}}(i) = [(\boldsymbol{\mathcal{U}}(i))^T, (-\tilde{\mathbf{P}}\tilde{\mathcal{S}}(i))^T]^T$, and $\mathbf{v}$ contains both the FF and BF coefficients $\mathbf{v} := [\mathbf{f}^T, \mathbf{b}^T]^T$. In particular, we have $[\hat{\mathbf{s}}(i)]_j = \tilde{\mathbf{u}}_{i,j}^T\mathbf{v}$ with $\tilde{\mathbf{u}}_{i,j}^T$ denoting the jth row of $\tilde{\boldsymbol{\mathcal{U}}}(i)$.

With the above notations, we are in a position to use the normalized least-mean squares (NLMS) algorithm [28] to minimize the square error between the input and the output of the quantizer. Suppose $\mathbf{v}^{(\kappa)}$ contains the DFE coefficients at the κth iteration. Then, it is refined at the next iteration as

$$\mathbf{v}^{(\kappa+1)} = \mathbf{v}^{(\kappa)} + \frac{\mu}{\|\tilde{\mathbf{u}}_{i,j}\|^2}\left([\check{\mathbf{s}}(i)]_j - \tilde{\mathbf{u}}_{i,j}^T\mathbf{v}^{(\kappa)}\right)\tilde{\mathbf{u}}_{i,j}^* \qquad (18)$$

where $[\check{\mathbf{s}}(i)]_j$ is the output of the quantizer, $\|\cdot\|$ denotes the Frobenius norm, and μ is the step size, which satisfies the convergence requirement. It is noteworthy that the iterations operate on the symbol level and need to go through the whole data symbol sequence for multiple loops until convergence is reached. Therefore, the iteration index κ must be associated with the block index i and symbol index j as

$$\operatorname{mod}(\kappa, (M-L)(N-K)) = i(M-L)+j \qquad (19)$$

with $\operatorname{mod}(p,q)$ standing for the residue of p divided by q. To ensure proper convergence, training symbols will be inserted at the head of the data symbol sequence.

Direct equalizer estimation has a design complexity that is linear in the length of the data symbol vector. It does not rely on channel

Fig. 2. With channel (BEM $Q = 2$) knowledge.

Fig. 3. With channel (BEM $Q = 4$) knowledge.
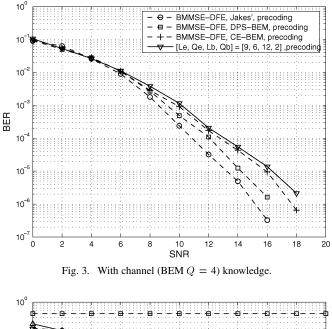

Fig. 4. Semiblind performance.

(CE-BEM) knowledge; the CE-BEM channel assumption is basically not relevant here. Perhaps more important, we can apply this approach without assuming any specific channel model, in which sense the channel modeling error is no longer a key concern.

IV. SIMULATION RESULTS

For the simulations, we generate TV channels using a Jakes' simulator as given in [31]. We assume the channels have an order $L = 3$ and a maximal normalized Doppler spread $\nu_{\max} = 0.0025$. In order to explore the full Doppler diversity with the precoder in (3) for an observation window of length $P = 400$, we choose $K = 2$, satisfying the Nyquist criterion $K = \lceil 2\nu_{\max}P \rceil$.

We assume the transmitted data symbol sequence is white and zero-mean and consists of $N - K = 18$ blocks, with each block containing $M - L = 17$ quadrature phase-shift keying (QPSK) symbols (note that $N = M = 20$ in order to reach $P = MN = 400$). We employ one transmit antenna and two receive antennas that are corrupted by additive white Gaussian noise $\mathcal{N}(0, \sigma_n^2)$. All the antennas are equipped with a rectangular filter. At the receive side, the received signals are further oversampled by a factor of two. Hence, we deal with a SIMO system with $A = 4$ channel outputs. Note that due to the oversampling, the noise from each channel output is not necessarily uncorrelated with each other, but we will still use this assumption in the equalizer design. The obtained bit error rate (BER) is averaged over 5000 Monte Carlo runs across an SNR range from 0 to 20 dB. Here, the SNR is defined as $(M - L)(N - K)/P\sigma_n^2$, taking the precoder-induced redundancy into account.

Test Case 1 (Equalizers Based on Channel (BEM) Knowledge): In Fig. 2, we first list the performance of four BMMSE-DFEs, the first three of which are based on a precoded system and are constructed using the knowledge of the true Jakes' channel, the best DPS-BEM fit with $Q = 2$, and the best CE-BEM fit with $Q = 2$, respectively. The last BMMSE-DFE also utilizes the channel knowledge of the best CE-BEM fit with $Q = 2$, but the transmission system is only zero-stuffed, e.g., $\mathbf{F}_N^H = \mathbf{I}_N$ in (3) (referred to as "no precoding" in the figure). For those BMMSE-DFEs, the design complexity is $\mathcal{O}(316^3)$ and the implementation complexity is about 536 418 MA operations per data vector. In contrast, the proposed DFE entails a much lower cost: for the DFE scaled by $[L_e, Q_e, L_b, Q_b] = [3, 2, 0, 0]$ (thus no decision feedback), the design complexity is $\mathcal{O}(35^3)$ and the implementation complexity is 22 212 MA operations per data vector. Affording a bit more design complexity of $\mathcal{O}c(143^3)$ and implementation
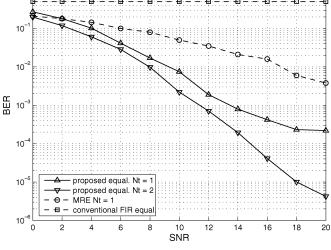
complexity of 155 412 MA operations per data vector, the DFE with $[L_e, Q_e, L_b, Q_b] = [9, 8, 12, 2]$ yields a performance close to that of the BMMSE-DFE based on the same CE-BEM channel knowledge.

It can be observed that due to the channel modeling error, all the BEM-based equalizers suffer from a performance gap with respect to the one that utilizes the true Jakes' channel knowledge. However, this performance gap can be mitigated if we adopt a larger Q for both the CE-BEM and DPS-BEM. As revealed in Fig. 3, the DFE with $[L_e, Q_e, L_b, Q_b] = [9, 6, 12, 2]$ that is based on the best CE-BEM fit with $Q = 4$ exhibits a 2-dB improvement at a BER of 10^{-5} than its counterpart in the $Q = 2$ case. In addition, the induced implementation complexity is even reduced to 123 732 MA operations per data vector, while the same design complexity is maintained. Actually, the performance of all the BEM-based equalizers is improved.

Test Case 2 (Semiblind Equalizers): For the semiblind equalization, we implement a moderate DFE with $[L_e, Q_e, L_b, Q_b] = [3, 2, 2, 2]$ to reduce the number of unknowns in the estimation. If we assume the first N_t data symbol blocks are pilots, the overall bandwidth efficiency becomes $(M - L)(N - K - N_t)/P$. We follow the NLMS in (18) using $\mu = 0.3$ and plot the performance in Fig. 4. For the case $N_t = 1$, equivalent to a bandwidth efficiency of 72%, the performance of the proposed DFE suffers from a BER floor at high SNR. This is alleviated
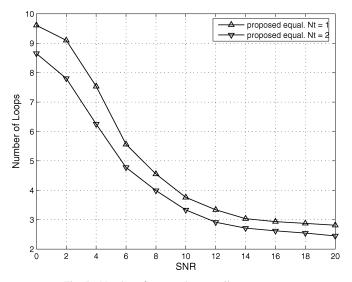
Fig. 5.   Number of average loops until convergence.

if the system can afford more overhead, e.g., $N_t = 2$ equivalent to a bandwidth efficiency of 68%. For comparison, we list the performance of the equalizer given in [23] (referred to as "MRE") and a classical time-invariant FIR DFE. The MRE is less accurate since it is unable to explore the finite-alphabet property of the data symbols. The classical DFE is obviously not capable of tracking fast-fading channels.

The proposed DFE also has the virtue of fast convergence. As revealed in Fig. 5, the average number of loops until convergence decreases to around 3 for an SNR higher than 12 dB. In such cases, the resulting design complexity is $\mathcal{O}(316 \times 3)$ and the implementation complexity is about 75 816 MA operations per data vector, which are lower than the DFE based on channel (CE-BEM) knowledge.

V. CONCLUSION

In this correspondence, we have proposed an equalization/decoding scheme for a precoded transmission system. The precoder enables the maximum available diversity for a CE-BEM channel, but makes most existing equalizers ineffective or very expensive. By adopting a CE-BEM to approximate the TV channel and commuting the CE-BEM with the (de-)precoder, we can apply a parametric DFE after the decoder, which is computationally attractive for a channel with moderate L and Q. Two approaches are proposed to construct the equalizer: 1) utilizing the channel (CE-BEM) knowledge and 2) in a semiblind adaptive fashion.

The first approach affords higher bandwidth efficiency and yields better performance but relies on the CE-BEM assumption of the TV channel. Consequently, the equalizer is penalized by the channel modeling error. From the simulations, we observe that the precoder combined with a larger CE-BEM helps to mitigate the influence of the channel modeling error, without increasing the equalization complexity. Of course, this could pose more pressure on any possible channel estimator, which has not been treated in this correspondence.

REFERENCES

[1] W. C. Jakes, *Microwave Mobile Channels*.   New York: Wiley, 1974.
[2] M. K. Tsatsanis and G. B. Giannakis, "Modeling and equalization of rapidly fading channels," *Int. J. Adapt. Control. Signal Process.*, vol. 10, pp. 159–176, Mar. 1996.
[3] D. K. Borah and B. D. Hart, "Frequency-selective fading channel estimation with a polynomial time-varying channel model," *IEEE Trans. Commun.*, vol. 47, no. 6, pp. 862–873, Jun. 1999.

[4] T. Zemen and C. F. Mecklenbräuker, "Time-variant channel estimation using discrete prolate spheroidal sequences," *IEEE Trans. Signal Process.*, vol. 53, no. 9, pp. 3597–3607, Sep. 2005.
[5] D. Slepian, "Prolate spheroidal wave functions, fourier analysis and uncertainty—V: The discrete case," *Int. J. Adapt. Control Signal Process.*, vol. 57, pp. 1371–1430, May 1978.
[6] C. Lanczos, *Applied Analysis*.   Englewood Cliffs, NJ: Prentice-Hall, 1956.
[7] G. Leus, "On the estimation of rapidly time-varying channels," in *Proc. Eur. Signal Processing Conf. (EUSIPCO)*, Vienna, Austria, Sep. 2004, pp. 2227–2230.
[8] X. Ma and G. B. Giannakis, "Maximum-diveristy transmissions over doubly selective wireless channels," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1832–1840, Jul. 2003.
[9] A. M. Sayeed and B. Aazhang, "Joint multipath-Doppler diversity in mobile wireless communications," *IEEE Trans. Commun.*, vol. 47, no. 1, pp. 123–132, Jan. 1999.
[10] H.-N. Lee and G. J. Pottie, "Fast adaptive equalization/diversity combining for time-varying dispersive channels," *IEEE Trans. Commun.*, vol. 46, no. 9, pp. 1146–1162, Sep. 1998.
[11] S. Benedetto and E. Biglieri, *Principles of Digital Transmission With Wireless Applications*.   Norwell, MA: Kluwer Academic/Plenum Publishers, 1999.
[12] Z. Wang, S. Zhou, and G. B. Giannakis, "Joine coding-precoding with low-complexity turdo-decoding," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 832–842, May 2004.
[13] N. Al-Dhahir and J. M. Cioffi, "MMSE decision-feedback equalizers: Finite-length results," *IEEE Trans. Inf. Theory*, vol. 41, no. 4, pp. 961–975, Jul. 1995.
[14] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1639–1642, Jul. 1999.
[15] G. Leus and M. Moonen, *Signal Processing for Mobile Communications Handbook*.   Boca Raton, FL: CRC Press, 2004, ch. 16.
[16] X. Ma, G. Giannakis, and S. Ohno, "Optimal training for block transmissions over doubly-selective fading channels," *IEEE Trans. Signal Process.*, vol. 51, no. 5, pp. 1351–1366, May 2003.
[17] M. Dong, L. Tong, and B. M. Sadler, "Optimal insertion of pilot symbols for transmissions over time-varying flat fading channels," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1403–1418, May 2004.
[18] H. A. Cirpan and M. K. Tsatsanis, "Maximum likelihood blind channel estimation in the presence of Doppler shifts," *IEEE Trans. Signal Process.*, vol. 47, no. 6, pp. 1559–1569, Jun. 1999.
[19] S. Haykin, A. H. Sayed, J. R. Zeidler, P. Yee, and P. C. Wei, "Adaptive tracking of linear time-variant systems by extended RLS algorithms," *IEEE Trans. Signal Process.*, vol. 45, no. 5, pp. 1118–1128, May 1997.
[20] M.-A. R. Baissas and A. M. Sayeed, "Pilot-based estimation of time-varying multipath channels for coherent CDMA receivers," *IEEE Trans. Signal Process.*, vol. 50, no. 8, pp. 2037–2049, Aug. 2002.
[21] G. Leus and M. Moonen, "Deterministic subspace based blind channel estimation for doubly-selective channels," in *Proc. IEEE Signal Processing Workshop Signal Processing Advances Wireless Communications (SPAWC)*, Jun. 2003, pp. 210–214.
[22] J. K. Tugnait and W. Luo, "Linear prediction error method for blind identification of periodically time-varying channel," *IEEE Trans. Signal Process.*, vol. 50, no. 12, pp. 3070–3082, Dec. 2002.
[23] G. Leus, I. Barhumi, O. Rousseaux, and M. Moonen, "Direct semiblind design of serial linear equalizers for doubly-selective channels," in *Proc. IEEE Int. Conf. Communications (ICC)*, Jun. 2004, pp. 2626–2630.
[24] G. B. Giannakis and C. Tepedelenlioŭglu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proc. IEEE*, vol. 86, pp. 1969–1986, Oct. 1998.
[25] H. Liu and G. B. Giannakis, "Deterministic approaches for blind equalization of time-varying channels with antenna arrays," *IEEE Trans. Signal Process.*, vol. 46, no. 11, pp. 3099–3104, Nov. 1998.
[26] D. Gesbert, P. Duhamel, and S. Mayrargue, "On-line blind multichannel equalization based on mutually referenced filters," *IEEE Trans. Signal Process.*, vol. 45, no. 9, pp. 2307–2317, Sep. 1997.
[27] C. Tepedelenlioŭglu and G. B. Giannakis, "Transmitter redundancy for blind estimation and equalization of time- and frequency-selective channel," *IEEE Trans. Signal Process.*, vol. 48, no. 7, pp. 2029–2043, Jul. 2000.

[28] S. Haykin, *Adaptive Filter Theory.* Englewood Cliffs, NJ: Prentice-Hall, 1996.

[29] Y.-S. Choi, P. J. Voltz, and F. A. Cassara, "On channel estimation and detection for multicarrier signals in fast and selective Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 49, no. 8, pp. 1375–1387, Aug. 2001.

[30] P. Schniter, "Low-complexity equalization of OFDM in doubly-selective channels," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 1002–1011, Apr. 2004.

[31] Y. R. Zheng and C. Xiao, "Simulation models with correct statistical properties for rayleigh fading channels," *IEEE Trans. Commun.*, vol. 51, no. 6, pp. 920–928, Jun. 2003.

[32] R. A. Horn and C. R. Johnson, *Matrix Analysis.* Cambridge, U.K.: Cambridge Univ. Press, 1999.

Narrowband Interference Suppression Using Undecimated Wavelet Packets in Direct-Sequence Spread-Spectrum Receivers

Emilia Pardo, Miguel A. Rodríguez-Hernández, and Juan J. Pérez-Solano

Abstract—A new algorithm for narrowband interference suppression in direct-sequence spread-spectrum (DS-SS) communications is presented. The algorithm combines the undecimated wavelet packet transform (UWPT) with frequency shifts to confine the interference energy in a subband that is subsequently eliminated. Computer simulation shows a robust performance that appears to be independent of the interference frequency.

Index Terms—Direct-sequence spread spectrum (DS-SS), interference suppression, tree structuring algorithm (TSA), undecimated wavelet packets.

I. INTRODUCTION

Spread-spectrum (SS) communication systems have certain interference rejection capacities that make them suitable for transmitting information in congested or noisy channels. Nevertheless, the performance of these systems is strongly degraded by high-power interferences. In this context, interference suppression improves the immunity of the system without increasing the bandwidth. Two classes of interference rejection schemes have been extensively used: time-domain adaptive filtering [1] and transform-domain suppression [2]. Time-domain adaptive filtering can eliminate completely sine-wave interferences but needs a convergence time to reach the optimal solution. On the other hand, transform-domain techniques present

quick tracking of changing interferences, but their performance depends on the ability of the transform to confine the interference in a few bins [1].

This correspondence is focused on narrowband interference suppression by transform-domain excision. The received signal is represented in the transform domain, and the frequency bands that exceed a certain energy threshold are eliminated. In this way, most of the interference energy is removed with low distortion of the desired broadband signal. The inverse transformation is then applied to reconstruct the interference-suppressed SS signal.

A fundamental aspect in the preceding process is the choice of the transform. The objective is to obtain a set of basis functions that generate a compact representation of the interference. Previous approaches were based on the Fourier transform [1], [3], [4] but their performance is limited due to the windowing effect, which spreads the interference energy. Subsequently, more general multichannel filter bank structures were proposed, of which the Fourier transform is a particular case [5]. These transforms offer greater design freedom. Different types of perfect reconstruction multichannel filter banks have been applied to the interference suppression problem, some examples of which are modulated filter banks [6], paraunitary filter banks [7], and lapped transforms [2].

The wavelet transform [8] has also been employed to eliminate interferences in DS-SS communications [9]. The multiresolution time–frequency decomposition implemented by this transform performs relatively well with various types of interference, but the spectral partition provided is fixed, in the same way as those obtained with multichannel filter banks, and presents similar limitations, such as interband spectral leakage and fixed time–frequency resolution.

An evolution of the wavelet transform is the wavelet packet transform [8]. It provides greater flexibility than previous transforms, offering a variety of unequal bandwidth spectral decompositions. Wavelet packet transform allows the implementation of adaptive algorithms in the sense that they select the best decomposition tree according to the interference frequency location for each processed frame. In this group, the algorithm proposed in [10] deserves special attention. It is called adaptive time–frequency algorithm because interference suppression can be accomplished in the time or the frequency domain. The switch between time and frequency methods depends on the kind of interference to be eliminated. Pulsed interference is eliminated better in the time domain, because its energy is localized in time and expanded in frequency. On the other hand, narrowband interference suppression is easier in the frequency domain because the energy is localized in frequency and expanded in time. The dual treatment in time or in frequency provides significant performance improvements. When transformation from time to frequency is required, it is carried out by the tree structuring algorithm (TSA) [10]. The TSA employs adaptive subband tree decompositions or wavelet packets. This method has demonstrated superior performance over conventional fixed transforms for narrowband interference suppression [2], [10], but the results are slightly dependent on the interference frequency.

A new algorithm based on wavelet packets for narrowband interference suppression in spread spectrum communications is described here. Its novelty lies in the application of frequency shifts to the SS signal in order to concentrate the interference in a subband and thus eliminate it completely. The method has been successfully applied to frequency-hopping and direct-sequence SS (DS-SS) systems [11], [12]. In this paper, a detailed description of the scheme for interference suppression in DS-SS systems is presented, and experimental results confirming the improvements are provided.