WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
Wirel. Commun. Mob. Comput. 2001; **1**:201–219 (DOI: 10.1002/wcm.14)

REVIEW

Channel estimation in narrowband wireless communication systems

Hüseyin Arslan*,† and Gregory E. Bottomley

Ericsson Inc.
P.O. Box 13969
Research Triangle Park
NC 27709
U.S.A.

Summary

Channel estimation is an integral part of standard adaptive receiver designs used in narrowband, digital wireless communication systems. In this tutorial paper, commonly used approaches to channel estimation are reviewed. Both time-invariant and time-varying channels are considered. For time-varying channels, both pilot symbol interpolation and data-directed channel tracking are considered. Applications include the Global System for Mobile communications, the Enhanced Data rates for Global Evolution system, and another Time-Division Multiple-Access system known as Telecommunications Industry Association/Electronics Industry Association/Interim Standard—136 (TIA/EIA/IS-136 or IS-136).

KEY WORDS

channel estimation
channel tracking
channel identification
parameter estimation
adaptive estimation
time-varying channnels
dispersive channels
time-division multiple access

1. Introduction

In digital wireless communication systems, information is transmitted through a radio channel. For conventional, coherent receivers, the effect of the channel on the transmitted signal must be estimated to recover the transmitted information. For example, with binary phase shift keying (BPSK), binary information is represented as $+1$ and -1 symbol values. The radio channel can apply a phase shift to the transmitted symbols, possibly inverting the symbol values. As long as the receiver can estimate what the channel did to the transmitted signal, it can accurately recover the information sent.

*Correspondence to: H. Arslan, Ericsson Inc., P.O. Box 13969, Research Triangle Park, North Carolina 27709, U.S.A.
†E-mail: eushura@rtp.ericsson.se

Channel estimation is a challenging problem in wireless communications. Transmitted signals are typically reflected and scattered, arriving at the receiver along multiple paths. When these paths have similar delays, they add either constructively or destructively, giving rise to fading. When these paths have very different delays, they appear as signal echoes. Due to the mobility of the transmitter, the receiver, or the scattering objects, the channel changes over time.

This paper provides a tutorial overview of channel estimation approaches commonly applied to narrowband wireless communication systems. Various approaches are described, along with representative references. The channel impulse response is typically modeled as a set of delays and channel coefficients. Only channel coefficient estimation is addressed in this paper. Also, it is assumed that some part of the transmitted signal is known at the receiver. For example, the transmitter can provide a training sequence or periodic pilot symbols that can be used by the receiver as part of channel estimation. Blind channel estimation is not included in this paper, though the interested reader is referred to References [1] and [2]. Channel estimation can be viewed as an application of system identification, which has a long and rich history in the signal processing and control theory communities [3, 4]. A thorough treatment of this subject area is beyond the scope of this paper. Instead, the focus is on channel estimation approaches applied to narrowband digital cellular communication systems. Earlier related work for High Frequency (HF) modem design can be found in Reference [5] and references cited therein.

Three examples are used throughout. The first example corresponds to the Global System for Mobile communications (GSM) [6, 7] and the Enhanced Data rates for Global Evolution (EDGE) system [8]. In these Time-Division Multiple-Access (TDMA) systems, information is transmitted in time slots, as illustrated in Figure 1. The time slot is short enough that, for reasonable vehicle speeds, the channel is approximately constant for the duration of the slot. In the middle of the slot, there is a known 26-symbol training sequence, which can be used for channel estimation.

The second example corresponds to one of the downlink slot structures of the Telecommunications Industry Association/Electronics Industry Association/Interim Standard—136 (TIA/EIA/IS-136 or simply IS-136) system [6, 9, 10]. In this TDMA system, as shown in Figure 2, the slot duration is much longer, so that the channel can change significantly during the slot. However, clusters of known pilot symbols are placed throughout the slot. The channel can be estimated locally at the pilot symbol clusters, then interpolated over the data portions of the slot.

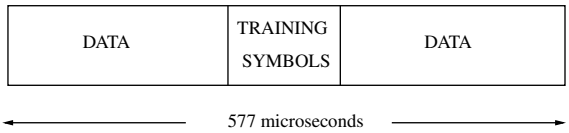

Fig. 1. GSM slot structure.

The third example, shown in Figure 3, corresponds to another of the downlink slot structures of the IS-136 system. The slot duration is long, like the previous example, but there is only a single training sequence. While an initial channel estimate can be obtained using the training sequence, the channel must be tracked over the data portion of the slot using the data symbols.

The paper is organized as follows. In Section 2, baseband equivalent channel models are given for both time-invariant and time-varying channels. For the first example (GSM and EDGE), a time-invariant channel model can be used. Channel estimation for time-invariant channels using a known training sequence is examined in Section 3. In the second and third examples (IS-136), a time-varying channel model is needed. In Section 4, channel estimation approaches based on pilot symbols (second example) are described. For the third example, continuous tracking of the channel is considered. In Section 5, tracking approaches are developed assuming the transmitted symbols are known. The fact that the transmitted symbols are not known is addressed in Section 6, concentrating on techniques developed for decision feedback equalization (DFE) and maximum likelihood sequence estimation (MLSE). Section 7 concludes the paper.

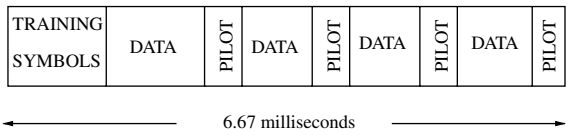

Fig. 2. IS-136 slot structure with pilot symbols.

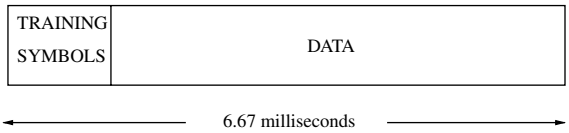

Fig. 3. IS-136 slot structure without pilot symbols.

2. System Model

The complex, baseband equivalent system model given in Figure 4 is used throughout. Both time-invariant and time-varying channels are considered.

2.1. Time-invariant channel

At the transmitter, digital symbols are transmitted as

$$x(t) = \sum_k b_k f(t - kT) \tag{1}$$

where b_k corresponds to a sequence of symbols, $f(\tau)$ is the pulse shape (transmit filter impulse response) as a function of delay τ, and T is the symbol period. The transmitted signal passes through a radio channel which, from Nyquist arguments [11], can be modeled using discrete filter taps. The resulting signal is received in the presence of noise, giving,

$$y(t) = \sum_{\ell=0}^{L-1} c(\ell)x(t - \tau(\ell)) + w(t) \tag{2}$$

where L is the number of channel taps; and $c(\ell)$ and $\tau(\ell)$ are the ℓth complex channel coefficient and delay, respectively. The delays can be assumed to be equally spaced, i.e., $\tau(\ell) = \ell T/M$ where M is an integer, and the spacing (T/M) for accurate modeling depends on the bandwidth of the system. Typically M is 1 for symbol-spaced channel modeling or 2 for fractionally spaced channel modeling. The noise term $w(t)$ is assumed to be white, complex Gaussian noise. For simplicity, symbol-spaced channel modeling is assumed, though extension to fractionally spaced channel modeling is straightforward.

At the receiver, the received signal $y(t)$ is filtered by a filter, matched to the pulse shape, and sampled with sampling period T_s, giving

$$r_k = \int f^*(\tau)y(\tau + kT_s)d\tau \tag{3}$$

where superscript '*' denotes complex conjugation. Substituting Equation (2) into Equation (3), the received samples can be expressed as

$$\boxed{r_k = \sum_{j=0}^{J-1} h(j)b_{k-j} + z_k} \tag{4}$$

where $h(j)$ is the jth composite channel coefficient, reflecting the influence of the transmit filter, the radio channel, and the receive filter, i.e.,

$$h(j) = \sum_{\ell=0}^{L-1} c(\ell)R_{ff}(jT - \ell T) \tag{5}$$

where $R_{ff}(\tau)$ is the pulse shape autocorrelation function. Note that z_k corresponds to a sequence of complex Gaussian noise samples, which may be correlated depending on the pulse shape autocorrelation function. For the special case of Nyquist filtering, $h(j) = c(j)$ and the noise samples are uncorrelated.

Using this model of the received samples, the operation of a simple coherent receiver is as follows. Assuming a one-tap channel model, Equation (4) simplifies to

$$r_k = h(0)b_k + z_k \tag{6}$$

Assume the information symbol b_k is either $+1$ or -1 (BPSK). Assuming the channel coefficient $h(0)$ is known, the information can be recovered using

$$\hat{b}_k = \text{sign}(\text{Re}\{h^*(0)r_k\}) \tag{7}$$

where $\text{Re}\{\cdot\}$ denotes the real part of a complex number. Observe that multiplying the received value by the conjugate of the channel coefficient removes the phase rotation introduced by the channel. In practice, the channel coefficient is estimated.

To explain many of the channel estimation approaches, it is convenient to use vector and matrix forms. Assuming K received samples total, the following column vectors are defined: $\mathbf{r} = [r_0\, r_1 \ldots r_{K-1}]^T$, $\mathbf{h} = [h(0)\, h(1) \ldots h(J-1)]^T$, and $\mathbf{z} = [z_0\, z_1 \ldots z_{K-1}]^T$. Then, Equation (4) can be expressed as

$$\boxed{\mathbf{r} = \mathbf{B}\mathbf{h} + \mathbf{z}} \tag{8}$$

where $\mathbf{B}$ is a $K \times J$ matrix whose rows correspond to different shifts of the transmitted sequence of symbols.

It is common to model the channel coefficients as uncorrelated, zero-mean complex Gaussian random variables [12, 13]. This corresponds to 'Rayleigh' fading, in that channel tap magnitudes (amplitudes) are Rayleigh distributed. This model is used throughout.

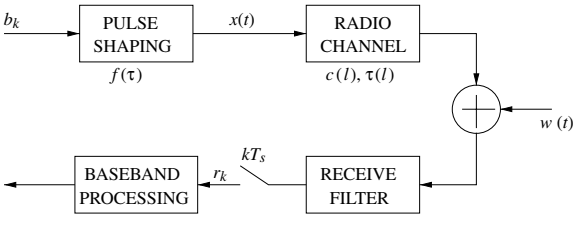

Fig. 4. System model.

2.2. Time-varying channel

With a time-varying channel, the time evolution of the channel must be characterized in addition to the distribution of the channel coefficients. Assuming the mobile receiver is surrounded by a ring of uniformly distributed scattering objects, then each channel coefficient can be modeled as a complex Gaussian random variable with autocorrelation function [14, 15]

$$R(m) = \mathbf{E}\{c_k^*(j)c_{k+m}(j)\} = \sigma_c^2(j)J_0(2\pi f_D m T_s) \tag{9}$$

where mT_s is the autocorrelation delay, $\sigma_c^2(j)$ is the mean square value of the jth channel coefficient, f_D is the Doppler spread which is proportional to the vehicle speed and the carrier frequency, and $J_0(\cdot)$ is the zeroth order Bessel function of the first kind. The corresponding power spectrum of the fading process is shown in Figure 5. As with the time-invariant case, different coefficients are assumed to be uncorrelated with one another.

How fast the channel changes depends on the carrier frequency and vehicle speed. For the IS-136 examples, the carrier frequency can be either in the cellular band or the Personal Communication Systems (PCS) band. The largest carrier frequencies for these bands are 893.97 MHz and 1989.95 MHz, respectively. At high vehicle speeds, e.g. 100 km h^{-1}, these carrier frequencies correspond to Doppler spread values of 83 Hz and 184 Hz, respectively.

It is assumed that the fading is slow relative to the symbol rate, so that the fading coefficients are approximately constant over the pulse shape. As a result, for symbol-spaced channel modeling, the received samples can be approximated by

$$r_k \approx \sum_{j=0}^{J-1} h_k(j)b_{k-j} + z_k \tag{10}$$

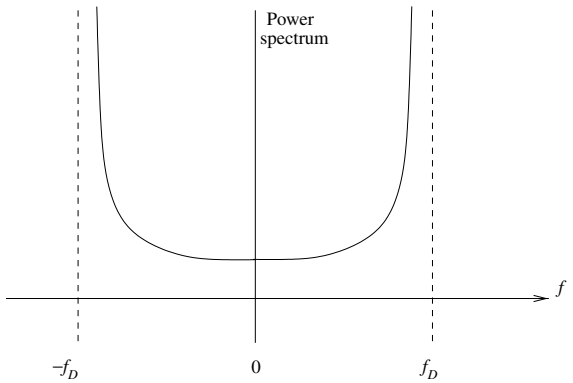

Fig. 5. Spectrum of fading process.

where subscript k in $h_k(j)$ denotes time dependency. For the IS-136 examples, this is reasonable, as the maximum Doppler spread is 184 Hz whereas the symbol rate is 24.3 kbaud.

When developing channel tracking approaches, it is common to formulate the tracker in terms of the conjugate of the channel coefficients, i.e., $g_k = h_k^*$. This gives the equivalent model

$$\boxed{r_k = \mathbf{g}_k^H \mathbf{b}_k + z_k} \tag{11}$$

where superscript H denotes Hermitian (conjugate) transpose, $\mathbf{g}_k = [g_k(0)\ g_k(1) \ldots g_k(J-1)]^T$ and $\mathbf{b}_k = [b_k, \ldots, b_{k-J+1}]^T$.

3. Estimation for Time-Invariant Channels

In this section, channel estimation approaches for the first example, GSM and EDGE, are examined. In this example, the slot duration is short (577 μs), so that the channel response over the time slot is approximately constant for reasonable vehicle speeds [16, 17]. The accuracy of this approximation is explored in References [18] and [19]. It is assumed that the channel is estimated using the known training sequence.

3.1. Correlation

One of the simplest ways of estimating channel coefficients is to correlate the received signal to a subsequence of the known training sequence [20–27]. The resulting channel coefficient estimate is given by

$$\hat{h}_{\mathrm{COR}}(j) = \frac{1}{N_s} \sum_{\ell=\ell_0}^{\ell_0+N_s-1} r_{\ell+j} b_\ell^* \tag{12}$$

where N_s is the length of the subsequence used in correlation and ℓ_0 is the index of the first symbol used in correlation.

Substituting Equation (4) into Equation (12), the correlation approach channel estimate can be expressed as

$$\hat{h}_{\mathrm{COR}}(j) = h(j) + \frac{1}{N_s} \sum_{\substack{m=0 \\ m\neq j}}^{J-1} h(m)R_{j-m} + \tilde{z}_j \tag{13}$$

where $\tilde{z}_j$ results from the noise and R_ℓ is the deterministic correlation between the portion of the training sequence used for channel estimation and the ℓth shift

of the full training sequence. This form of autocorrelation, sometimes called the ambiguity function, is given by

$$R_\ell = \sum_{n=\ell_0}^{\ell_0+N_s-1} b_n^* b_{n+\ell} \tag{14}$$

From Equation (13), the accuracy of channel estimation is affected by noise and by interference between taps, which depends on the tap coefficient values as well as the training sequence autocorrelation function. This interference between taps can be suppressed using an iterative form of successive cancellation [20]. Nonideal autocorrelation can also be addressed by refining the correlation estimate values using the Least Mean Squares (LMS) algorithm [21].

Interference between taps can be eliminated by careful design of the training sequence. In GSM and EDGE, each training sequence of 26 symbols contains a portion of $N_s = 16$ symbols such that $R_\ell = 0$ for nonzero lags between -5 and 5 [22, 26]. Thus, assuming a composite channel of six taps or less, interference between taps is eliminated. In general, to avoid correlation with unknown data symbols, $N_s \leq N - 2(J_{\max} - 1)$, where N is the length of the training sequence and $J_{\max}$ is the maximum number of composite channel coefficients. While design of 'ideal' training sequences is beyond the scope of this paper, the interested reader is referred to References [27–31].

When the training sequence autocorrelation is ideal, the estimation error variance on each coefficient is σ_z^2/N_s. Thus, the estimation error variance is proportional to the noise variance and inversely proportional to the length of the correlation sequence.

One way of quantifying channel estimation performance is by the mean square error (MSE) matrix, also referred to as the error covariance [32]. The MSE matrix is defined as

$$\text{MSE} = \text{E}\{(\mathbf{h} - \hat{\mathbf{h}})(\mathbf{h} - \hat{\mathbf{h}})^H\} \tag{15}$$

where $\text{E}\{\cdot\}$ denotes expected value. The diagonal elements of this matrix give the MSE or estimation error variance on the different coefficient estimates. The off-diagonal elements indicate how much the errors on different channel coefficients are correlated.

For the correlation approach, assuming ideal autocorrelation properties and white noise values, the MSE matrix is given by

$$\text{MSE}_{\text{COR}} = \left(\frac{\sigma_z^2}{N_s}\right)(\mathbf{I} + \mathbf{O}) \tag{16}$$

where $\mathbf{I}$ is the identity matrix, and $\mathbf{O}$ has only nonzero off-diagonal elements that depend on the correlation of the subsequence to shifts of itself.

3.2. Least-squares estimation

Least squares (LS) channel estimation is also commonly used when a training sequence is available [18, 31, 33, 34]. The LS channel coefficient estimate is defined by

$$\hat{\mathbf{h}}_{\text{LS}} = \arg\min_{\mathbf{h}} (\mathbf{r} - \mathbf{B}\mathbf{h})^H (\mathbf{r} - \mathbf{B}\mathbf{h}) \tag{17}$$

where $K = N - (J - 1)$ samples are used to form $\mathbf{r}$ (see Figure 6). By differentiating with respect to each channel coefficient and setting the result to zero, the following closed form expression for the LS channel estimate can be obtained:

$$\hat{\mathbf{h}}_{\text{LS}} = (\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H\mathbf{r} \tag{18}$$

Substituting Equation (8) into Equation (18), the LS channel estimate can be expressed as

$$\hat{\mathbf{h}}_{\text{LS}} = \mathbf{h} + (\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H\mathbf{z} \tag{19}$$

Observe that, unlike the correlation approach, the interference between taps is eliminated, independent of any training sequence correlation properties.

From Equation (19), the MSE matrix assuming white noise is given by $(\mathbf{B}^H\mathbf{B})^{-1}\sigma_z^2$. Thus, the estimation error variance depends on the noise power, the length of the training sequence, the number of channel taps, and the correlation properties of training sequences [31].

The estimation error variance is minimized when $\mathbf{B}^H\mathbf{B} = (N - (J - 1))\mathbf{I}$ [18]. This property defines

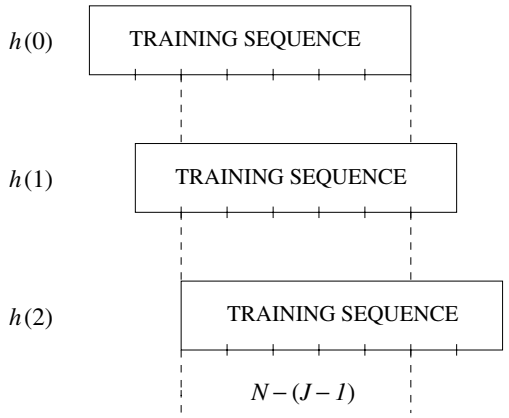

Fig. 6. Samples used for LS, ML, MAP, and MMSE channel estimation, $J = 3$, $N = 7$.

'ideal' autocorrelation properties for a training sequence used with LS channel estimation, as the off-diagonal elements of $\mathbf{B}^H\mathbf{B}$ are correlations between different portions of the training sequence. With ideal autocorrelation properties, the estimation error variance simplifies to $\sigma_z^2/(N-(J-1))$ and the MSE matrix becomes

$$\mathrm{MSE}_{\mathrm{LS}} = \left(\frac{\sigma_z^2}{N-(J-1)}\right)\mathbf{I} \tag{20}$$

Observe that the error on different coefficient estimates is uncorrelated. Also, the estimation variance is inversely proportional to the $N-(J-1)$, not N_s, as is the case for the correlation approach. As N_s must be less than $N-(J-1)$, the LS approach gives a lower error variance under the assumptions given by an improvement factor $(N-(J-1))/N_s$. For the GSM and EDGE example, $N_s = 16$. If a six-tap channel model is assumed, the improvement factor is 21/16 or 1.2 dB. If a five-tap channel model is assumed, which is more common [22], then the improvement factor is 20/16 or 1.0 dB. Thus, the LS approach performs better than the correlation approach, though the complexity of the LS approach is higher.

3.3. Maximum likelihood estimation

With maximum likelihood (ML) channel estimation, the likelihood of the received data samples given the channel coefficients is maximized [13, 32]. The ML solution can be expressed as

$$\hat{\mathbf{h}}_{\mathrm{ML}} = \arg\max_{\mathbf{h}} p(\mathbf{r}|\mathbf{h}) \tag{21}$$

where $p(a|b)$ denotes the conditional probability density function of a given b. The ML solution can be obtained by maximizing a quantity related to the log-likelihood, which depends on the distribution of the noise. For complex Gaussian noise, the ML solution becomes

$$\hat{\mathbf{h}}_{\mathrm{ML}} = \arg\max_{\mathbf{h}} \left[-(\mathbf{r}-\mathbf{B}\mathbf{h})^H\mathbf{R}_z^{-1}(\mathbf{r}-\mathbf{B}\mathbf{h})\right] \tag{22}$$

where $\mathbf{R}_z$ is the covariance matrix of the noise. A simpler form of the solution is obtained by setting the derivative with respect to $\mathbf{h}$ equal to zero, giving

$$\hat{\mathbf{h}}_{\mathrm{ML}} = (\mathbf{B}^H\mathbf{R}_z^{-1}\mathbf{B})^{-1}(\mathbf{B}^H\mathbf{R}_z^{-1})\mathbf{r} \tag{23}$$

The corresponding MSE matrix is $(\mathbf{B}^H\mathbf{R}_z^{-1}\mathbf{B})^{-1}$. Assuming white noise, Equation (23) simplifies to the LS solution given in Equation (18) and, assuming ideal autocorrelation properties, the MSE matrix is given by Equation (20).

Thus, assuming white noise, the ML approach is equivalent to the LS approach. When the noise is not white and the noise covariance matrix is known, then the ML approach allows that knowledge to be used to improve channel estimation [32].

3.4. Maximum a posteriori estimation

With maximum *a posteriori* (MAP) channel estimation [32], the likelihood of the channel coefficients given the received data is maximized. Such an approach treats the channel coefficients as random variables. The MAP solution can be expressed as

$$\hat{\mathbf{h}}_{\mathrm{MAP}} = \arg\max_{\mathbf{h}} p(\mathbf{h}|\mathbf{r}) \tag{24}$$

Using the properties of conditional probability, the MAP estimate can be expressed as

$$\hat{\mathbf{h}}_{\mathrm{MAP}} = \arg\max_{\mathbf{h}} \frac{p(\mathbf{r}|\mathbf{h})p(\mathbf{h})}{p(\mathbf{r})} \tag{25}$$

The $p(\mathbf{r})$ term in the denominator can be removed, so that the MAP estimate is equivalent to maximizing $p(\mathbf{r}|\mathbf{h})p(\mathbf{h})$. Observe that this is similar to the ML criterion, except that *a priori* statistical knowledge of the channel coefficient, represented by $p(\mathbf{h})$, has been included. Thus, the MAP solution depends on both the probability distribution of the noise and the probability distribution of the channel coefficients.

As discussed in Section 2, the channel coefficients can be modeled as zero-mean, complex Gaussian random variables. Assuming the noise and channel coefficients are independent, the MAP solution becomes [32]

$$\hat{\mathbf{h}}_{\mathrm{MAP}} = \mathbf{R}_h\mathbf{B}^H(\mathbf{B}\mathbf{R}_h\mathbf{B}^H + \mathbf{R}_z)^{-1}\mathbf{r} \tag{26}$$

where $\mathbf{R}_h$ is the channel coefficient covariance matrix.

The general expression for the MSE matrix is $\mathbf{R}_h - \mathbf{R}_h\mathbf{B}^H(\mathbf{B}\mathbf{R}_h\mathbf{B}^H + \mathbf{R}_z)^{-1}\mathbf{B}^H\mathbf{R}_h$ [32], which averages over both the channel coefficient and noise distributions. To compare with previous results, the following assumptions are made: (a) the channel coefficients are uncorrelated and have equal strength, so that $\mathbf{R}_h = \sigma_h^2\mathbf{I}$; (b) the training sequence has LS ideal autocorrelation properties and (c) the noise is white. With these assumptions, the MAP MSE matrix simplifies to

$$\mathrm{MSE}_{\mathrm{MAP}} = a\left(\frac{\sigma_z^2}{N-J+1}\right)\mathbf{I} = a\cdot\mathrm{MSE}_{\mathrm{LS}} \tag{27}$$

where a is given by

$$a = \left(\frac{\sigma_h^2}{\sigma_h^2 + \sigma_z^2/(N - J + 1)} \right) \tag{28}$$

Observe that the MSE matrix for MAP is similar to that of LS, except the result is scaled by a term less than 1. Thus, under the assumptions given, the MAP approach provides a more accurate estimate of the channel coefficient than the LS approach, particularly when σ_h^2/σ_z^2 is low. However, the MAP approach requires prior knowledge of the channel coefficient covariance matrix. When the noise is not white, the MAP approach also requires knowledge of the noise covariance.

It is also possible to determine the MSE matrix for a particular realization of the channel coefficients. Under assumptions (b) and (c) earlier, the resulting *conditional* MSE matrix is given by

$$\mathrm{MSE}_{\mathrm{MAP}}(\mathbf{h}) = (1 - a)^2 \mathbf{h}\mathbf{h}^H + a^2 \left(\frac{\sigma_z^2}{N - J + 1} \right) \mathbf{I} \tag{29}$$

Observe that even with certain autocorrelation properties, the MSE matrix of MAP depends on the particular channel realization. It can be shown that taking expected value of Equation (29) with respect to $\mathbf{h}$ gives Equation (27).

3.5. Minimum mean square error estimation

Another channel estimation approach is minimum mean square error (MMSE) estimation [32]. Like MAP, it treats the channel coefficients as random variables. Like LS, this approach is independent of the distribution of the channel coefficients, though it does depend on the moments of the channel coefficients. The MMSE channel estimate is defined as

$$\hat{\mathbf{h}}_{\mathrm{MMSE}} = \arg \min_{\mathbf{h}} [E\{(\mathbf{h} - \hat{\mathbf{h}})^H (\mathbf{h} - \hat{\mathbf{h}})\}] \tag{30}$$

Assuming $\mathbf{h}$ has zero mean and covariance $\mathbf{R}_h$, Equation (30) can be simplified to [32]

$$\hat{\mathbf{h}}_{\mathrm{MMSE}} = \mathbf{R}_h \mathbf{B}^H (\mathbf{B}\mathbf{R}_h \mathbf{B}^H + \mathbf{R}_z)^{-1} \mathbf{r} \tag{31}$$

Observe from Equations (26) and (31) that the MMSE estimate is equivalent to the MAP estimate under the assumption that the channel coefficients are Gaussian [32]. Thus, the MSE matrix is given by Equation (27). Like MAP, the MMSE approach requires prior information. In certain applications, the MMSE leads to optimal soft information when forward error correction (FEC) decoding is applied after demodulation [35, 36].

3.6. Parametric approaches

So far, the focus has been on estimating the composite channel response $\mathbf{h}$. However, according to the system model, the composite response can be expressed in terms of the pulse shape autocorrelation function, a known quantity, and the radio channel response $\mathbf{c}$, an unknown quantity. Channel estimation can be improved by taking advantage of this structure.

In [32, 37, 38], LS channel estimation using a parametric approach is developed. From Equation (5), the composite response can be expressed as $\mathbf{h} = \mathbf{A}\mathbf{c}$, where $\mathbf{h}$ is $J \times 1$, $\mathbf{c}$ is $L \times 1$, and the $J \times L$ matrix $\mathbf{A}$ depends on the pulse shape autocorrelation function. The LS estimate for $\mathbf{c}$ is then given by

$$\hat{\mathbf{c}} = (\mathbf{A}^{\mathbf{H}} \mathbf{B}^{\mathbf{H}} \mathbf{B}\mathbf{A})^{-1} \mathbf{A}^{\mathbf{H}} \mathbf{B}^{\mathbf{H}} \mathbf{r} \tag{32}$$

Observe that Equation (32) is similar to Equation (18), except that $\mathbf{B}$ has been replaced with $\mathbf{BA}$. In essence, the known pulse shape autocorrelation has been folded into the known symbols, which together represent the signal component after transmit and receive filtering. The composite response is then estimated by

$$\hat{\mathbf{h}} = \mathbf{A}\hat{\mathbf{c}} = \mathbf{A}(\mathbf{A}^{\mathbf{H}} \mathbf{B}^{\mathbf{H}} \mathbf{B}\mathbf{A})^{-1} \mathbf{A}^{\mathbf{H}} \mathbf{B}^{\mathbf{H}} \mathbf{r} \tag{33}$$

The parametric approach is particularly useful when the number of radio channel taps L is less than the number of composite channel taps J, leading to a reduction in estimation error variance. The parametric approach can also be used when tracking a time-varying channel [39–42].

4. Estimation Using Pilot Symbols

For the second example, IS-136 with pilot symbols, the long slot duration requires a time-varying channel model. In this section, approaches are discussed which 'measure' the channel at the pilot locations and use these measurements to estimate the channel at the data locations. To simplify the discussion, the example of a one-tap channel model ($J = 1$) is used.

4.1. Linear interpolation

One of the simplest forms of channel estimation using pilot symbols is linear interpolation [43–45]. With linear interpolation, the channel estimate at a certain

time period is a linear combination of the two 'nearest' channel measurements. For example, suppose there are measurements from pilot symbols as times $k = 0$ and $k = M$, denoted $\tilde{g}_0$ and $\tilde{g}_1$, respectively. Then, the channel estimate at time k, $0 < k < M$, is given by

$$\hat{g}_k = a_k(0)\tilde{g}_0 + a_k(1)\tilde{g}_1 \tag{34}$$

where $a_k(0)$ and $a_k(1)$ are given as

$$a_k(0) = \frac{M-k}{M}, \quad a_k(1) = \frac{k}{M} \tag{35}$$

Linear interpolation can be viewed as applying a filter with symbol-spaced taps to the channel measurements, which contain zeros at the unknown data symbol points, as illustrated in Figure 7. This suggests that other filtering approaches may be applied.

4.2. Wiener filtering

To obtain MMSE channel estimates, Wiener filtering can be used [46–50]. In general, the channel estimate as a function of the channel measurements is given by

$$\hat{g}_k = \mathbf{a}_k^T \tilde{\mathbf{g}} \tag{36}$$

where $\mathbf{a}_k$ is a $N_m \times 1$ vector of interpolation coefficients associated with time k and $\tilde{\mathbf{g}}$ is a $N_m \times 1$ vector of channel coefficient measurements corresponding to N_m pilot clusters. For Wiener filtering, the vector of filter coefficients are given by

$$\mathbf{a}_k = \mathbf{R}_{\tilde{g}}^{-1}\mathbf{p}_k \tag{37}$$

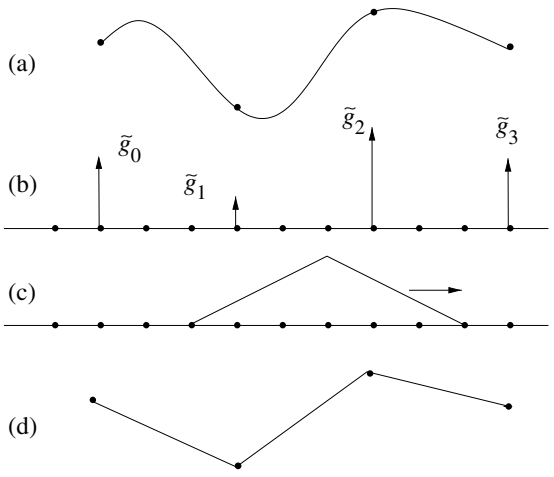

Fig. 7. Linear interpolation filtering using pilot symbols: (a) fading channel; (b) measurements at pilot locations; (c) interpolation filter; (d) estimated channel.

where $\mathbf{R}_{\tilde{g}}$ is the autocorrelation matrix of the channel at the pilot clusters and $\mathbf{p}_k$ is the cross-correlation vector between the channel at the pilot clusters and the channel at time k. From the system model given in Section 2, the elements of $\mathbf{R}_{\tilde{g}}$ and $\mathbf{p}_k$ can be expressed in terms of the channel coefficient mean-square values and a Bessel function [see Equation (9)].

The Wiener interpolation coefficients depend on the Doppler spread. When this information is not available, the interpolation filter coefficients can be selected based on the worst case expected Doppler spread value [47, 48].

4.3. Other filtering

In addition to the above interpolation techniques, other types of filters have been proposed. In Reference [51] , Gaussian interpolation filters are used. In References [52–54], a truncated version of Nyquist interpolation is used. Other interpolation filters can be found in Reference [55].

Interpolation may also be based on a deterministic polynomial model of the channel (see Section 5). In Reference [56], measurements at the pilot locations are used to estimate parameters of the polynomial model.

4.4. Channel measurements

The interpolation approach requires channel measurements at the pilot cluster locations. The pilot clusters are usually short so that the channel is approximately constant over these symbols. In case of very long pilot clusters, the cluster can be divided into smaller segments [46]. With the constant channel assumption, the channel measurements at each pilot cluster can be determined using the LS approach [46, 52, 54]. A channel tracker run over the pilot cluster only can also be used to obtain a channel measurement for interpolation [45].

In general, channel estimation performance (MSE) will depend on the slot design, fading channel characteristics, noise level, and the interpolation approach. Slot design is a trade-off between minimizing overhead and optimizing channel estimation performance. For example, the more rapid the fading, the more pilot clusters are needed.

5. Channel Tracking

Like the second example, the third example (IS-136 without pilot symbols) also requires a time-varying

channel model. With the third example, there is only a single training sequence. Thus, the channel must be tracked over the data portion of the slot using the data symbols.

In this section, channel tracking approaches are developed assuming that the symbols are known. Channel tracking approaches can be related to an underlying model of how the channel changes in time. These models for channel estimation are discussed first. Then, commonly used tracking approaches are given.

5.1. Models for channel estimation

Channel tracking is often based on a model of the channel. To obtain reasonable complexity, the model is usually an approximation to the model given in Section 2. Both stochastic and deterministic models are used.

5.1.1. Stochastic models

With stochastic models, the channel coefficients are modeled as stochastic random processes. The most commonly used models can be described using the Kalman filtering state space model [57, 58] given in Figure 8. In this model, the $N_s \times 1$ state vector $\mathbf{s}_k$ includes the channel coefficients. Each updated state $\mathbf{s}_{k+1}$ depends on the previous value of the state through a $N_s \times N_s$ state transition matrix $\mathbf{F}$ as well as the plant noise $\mathbf{u}_k$ through a $N_s \times N_s$ gain matrix $\mathbf{G}$. The plant noise is assumed to be a zero-mean, complex white Gaussian process with covariance $\mathbf{I}$.

The state is mapped to the output (observation) through a $N_s \times 1$ measurement matrix $\mathbf{H}_k$, which includes the symbol values. The output is observed in the presence of measurement noise z_k with mean square value σ_z^2. The observation is defined as the conjugate of the received samples (r_k^*) so that standard expressions can be used. Mathematically, the system is described by the following *process* and *measurement* equations:

$$\mathbf{s}_{k+1} = \mathbf{F}\mathbf{s}_k + \mathbf{G}\mathbf{u}_k \tag{38}$$

$$r_k^* = \mathbf{H}_k^H \mathbf{s}_k + z_k^* \tag{39}$$

In general, these equations can model an autoregressive moving average (ARMA) process [59].

One of the simplest channel models is the *random walk* model [57, 60]. With this model the state vector is the channel coefficient vector ($\mathbf{s}_k = \mathbf{g}_k$), the measurement matrix is the symbol vector ($\mathbf{H}_k = \mathbf{b}_k$), the state transition matrix is identity ($\mathbf{F} = \mathbf{I}$), and $\mathbf{G} = \sigma_g \mathbf{I}$. Therefore, the *process* and *measurement* equations become

$$\mathbf{g}_{k+1} = \mathbf{g}_k + \mathbf{u}_k \tag{40}$$

$$r_k^* = \mathbf{b}_k^H \mathbf{g}_k + z_k^* \tag{41}$$

Another commonly used channel model is a first-order autoregressive (AR) process [18, 61–63]. This process is similar to the random walk process, except that the transition matrix $\mathbf{F}$ is set to $\beta\mathbf{I}$ (assuming independently fading taps), where β indicates the correlation between two subsequent sample intervals. As a result, the *process* equation can be written as

$$\mathbf{g}_{k+1} = \beta\mathbf{g}_k + \mathbf{u}_k \tag{42}$$

In References [61] and [62], β is given as a function of the Doppler spread, $\beta = e^{-2\pi f_m T_s}$.

Second-order autoregressive (AR-2) channel modeling [17, 18, 64] has also been used, particularly when the channel changes rapidly. Such an approach can more accurately model the two spectral peaks that occur at plus and minus the Doppler spread (see Figure 5). For the simple case of a one-tap channel model ($J = 1$), the model quantities become

$$\mathbf{s}_k = \begin{pmatrix} g_k \\ -a_2 g_{k-1} \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} -a_1 & 1 \\ -a_2 & 0 \end{pmatrix},$$
$$\mathbf{G} = \begin{pmatrix} G_1 \\ 0 \end{pmatrix}, \mathbf{H} = \begin{pmatrix} b_k \\ 0 \end{pmatrix} \tag{43}$$

In References [65] and [66], higher order AR modeling is used, including adaptive estimation of the AR parameters.

Higher order modeling can also be obtained by tracking different order derivatives of the channel coefficients. For example, the channel tap and its derivative are tracked in the second-order *integrated*

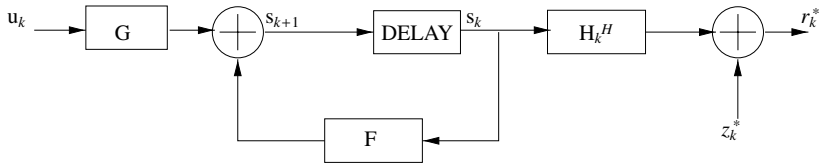

Fig. 8. Kalman signal model.

random walk (IRW) channel model [64, 67]. For the simple case of a one-tap channel model ($J = 1$), the model quantities become

$$\mathbf{s}_k = \begin{pmatrix} g_k \\ \dot{g}_k \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$
$$\mathbf{G} = \begin{pmatrix} 0 \\ G_2 \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} b_k \\ 0 \end{pmatrix} \tag{44}$$

where $\dot{g}_k$ denotes the time derivative of g_k. This approach can be generalized to include higher order derivatives [68].

5.1.2. Deterministic models

The time evolution of channel coefficients can also be expressed as a deterministic parametric function. Once the function's parameters have been estimated, the channel coefficient waveform is determined.

One popular deterministic model is the complex sinusoidal model [69–73]. Consider the simple case of a one-tap channel. With this model, the tap is modeled as

$$g_k = \sum_{n=1}^{N_e} A_n e^{j(2\pi f_n k T_s + \phi_n)} \tag{45}$$

where N_e is the number of exponentials, A_n, f_n and ϕ_n are the amplitude, the frequency and the phase of the nth exponential, respectively. Note that although g_k may change rapidly, the parameters A_n, f_n, ϕ_n are assumed to not change significantly during any given block of data.

In addition to sinusoidal modeling, the fading channel can also be modeled over a short interval using a small set of polynomial basis functions [60, 74–77]. For example, in Reference [75], a time-varying model of the channel with a linear or quadratic variation of channel with time is suggested. Quadratic variation of the channel can be represented as

$$g_k = d_0 + d_1 k + d_2 k^2 \tag{46}$$

To obtain a linear variation, d_2 is set to zero. In Reference [77], optimal polynomial model is discussed, and the optimal order is given as a function of fading rate, signal-to-noise ratio (SNR), and the duration of the received data.

5.2. Tracking approaches

In this subsection, commonly used channel tracking approaches are presented. Work comparing certain approaches can be found in Reference [57].

5.2.1. Least mean square

The least-mean-square (LMS) algorithm [13, 57, 78–80] is one of the simplest approaches to channel tracking. LMS channel tracking is performed according to

$$\hat{\mathbf{g}}_{k+1} = \hat{\mathbf{g}}_k + \mu \mathbf{b}_k e_k^* \tag{47}$$

where

$$e_k = r_k - \mathbf{g}_k^H \mathbf{b}_k \tag{48}$$

In essence, at symbol period k, an error e_k between what is received and what is modeled is formed. This error is used to update the channel coefficient estimate for the next symbol period.

For a time-invariant channel, LMS channel tracking can be interpreted as an iterative, stochastic gradient approach for finding the channel coefficients that minimize the mean square error (MSE) between the *received samples* and the model of the received samples [78]. For two non-complex (real) channel coefficients, the MSE as a function for the channel coefficients has a bowl shape [78]. At symbol period k, the LMS algorithm forms a noisy estimate of the slope or gradient at the place on the bowl corresponding to the current channel coefficient estimates. It then updates the taps in the direction of the negative gradient, so as to find the bottom of the bowl, where the MSE is minimized. Selection of the step size μ is a trade-off between rate of convergence and how noisy the model is at convergence (misadjustment noise).

For a time-varying channel, the step size μ trades tracking ability for misadjustment noise. The LMS algorithm can be derived from the Kalman filter (see later section), assuming a random walk model for the channel coefficients [60, 81].

LMS channel tracking was proposed early on for adaptive receivers [82, 83], and it continues to be a popular form of tracking [53, 84–101]. Different step sizes can be used for different taps depending on the average tap power [61, 102, 103]. In References [61] and [103], this form is derived from the Kalman filter using a first-order AR model of the channel. Also, the step size can be adaptive in time [41, 104].

5.2.2. Recursive least squares

Recursive least squares (RLS) [13, 57, 78–80] has also been applied to channel tracking, due to its rapid convergence properties. Conventional RLS channel

tracking is performed according to the following

$$\hat{\mathbf{g}}_{k+1} = \hat{\mathbf{g}}_k + \left(\frac{\mathbf{A}_k}{\lambda + \mathbf{b}_k^H \mathbf{A}_k \mathbf{b}_k} \right) \mathbf{b}_k e_k^* \quad (49)$$

$$\mathbf{A}_{k+1} = \frac{1}{\lambda} \left(\mathbf{A}_k - \frac{\mathbf{A}_k \mathbf{b}_k \mathbf{b}_k^H \mathbf{A}_k^H}{\lambda + \mathbf{b}_k^H \mathbf{A}_k \mathbf{b}_k} \right) \quad (50)$$

where e_k is given in Equation (48). Compared to LMS, RLS updates a second quantity, the $J \times J$ matrix $\mathbf{A}_k$, which corresponds to the inverse of an exponentially smoothed estimate of the symbol vector ($\mathbf{b}_k$) correlation matrix. This quantity is typically initialized to a diagonal matrix with large diagonal entries. The term λ is the 'forgetting factor' and, like the LMS step size, determines convergence rate and misadjustment properties. Typically, λ is chosen to be slightly less than 1.

RLS channel tracking can be viewed as finding the set of channel coefficients that minimizes a deterministic weighted squared error between the received samples and the modeled samples, i.e., [78]

$$E = |r_k - \mathbf{g}_{k+1}^H \mathbf{b}_k|^2 + \lambda |r_{k-1} - \mathbf{g}_{k+1}^H \mathbf{b}_{k-1}|^2 + \lambda^2 |r_{k-2} - \mathbf{g}_{k+1}^H \mathbf{b}_{k-2}|^2 + \cdots \quad (51)$$

As errors in modeling past received samples are weighted less, the RLS approach tracks the channel by trying to accurately model the most recent data. Because of the exponentional weighting, this form of RLS algorithm is also referred to as exponentially windowed RLS (EW-RLS) [105, 106]. RLS channel tracking can also be related to Kalman filtering, based on either a random walk model with a special form of the measurement noise covariance [60, 107] or a first order AR process with zero plant noise [57, 108]. Based on the latter relation, improved tracking approaches are developed by considering nonzero plant noise and a time-varying state transition matrix [109].

Conventional EW-RLS has been used to track wireless channels [32, 98–101, 110–112]. It has also been extended to track the channel tap and its derivative [75, 113].

Alternatively, a sliding window RLS (SW-RLS) approach can be used [77, 106]. In Reference [106] an optimized weighting function is applied to the window, based on statistical knowledge of the fading channel and SNR. In Reference [77], SW-RLS is used to track parameters for a polynomial model of the channel.

5.2.3. Kalman filtering

Kalman filtering [57, 59] provides a general, flexible framework for channel tracking. For the general model for channel evolution given in the previous subsection, the corresponding one-step prediction Kalman filter is given by

$$\hat{\mathbf{s}}_{k+1} = \mathbf{F}\hat{\mathbf{s}}_k + \mathbf{K}_k e_k^* \quad (52)$$

$$\mathbf{P}_{k+1} = \mathbf{F} \left(\mathbf{P}_k - \frac{\mathbf{P}_k \mathbf{H}_k \mathbf{H}_k^H \mathbf{P}_k}{\mathbf{H}_k^H \mathbf{P}_k \mathbf{H}_k + \sigma_z^2} \right) \mathbf{F}^H + \mathbf{G}\mathbf{G}^H \quad (53)$$

where

$$\mathbf{K}_k = \frac{\mathbf{F}\mathbf{P}_k \mathbf{H}_k}{\mathbf{H}_k^H \mathbf{P}_k \mathbf{H}_k + \sigma_z^2} \quad (54)$$

and e_k is given by Equation (48). Note that $\mathbf{K}_k$ is a $N_s \times 1$ vector and $\mathbf{P}_k$ is a $N_s \times N_s$ matrix.

The Kalman filter is an MMSE channel estimation approach for time-varying channels. Note that the Kalman filter requires knowledge of the measurement noise variance, as well as model parameters ($\mathbf{F}$, $\mathbf{G}$, $\mathbf{H}_k$).

Kalman filtering has been applied to channel tracking using a variety of channel models. It has been used with the most simple model, the random walk model [81, 114, 115]. Using Equations (40) and (41), expressions (52) through (54) simplify to [114, 115]

$$\hat{\mathbf{g}}_{k+1} = \hat{\mathbf{g}}_k + \mathbf{K}_k \mathbf{e}_k^* \quad (55)$$

$$\mathbf{P}_{k+1} = \left(\mathbf{P}_k - \frac{\mathbf{P}_k \mathbf{b}_k \mathbf{b}_k^H \mathbf{P}_k}{\mathbf{b}_k^H \mathbf{P}_k \mathbf{b}_k + \sigma_z^2} \right) + \mathbf{G}\mathbf{G}^H \quad (56)$$

$$\mathbf{K}_k = \frac{\mathbf{P}_k \mathbf{b}_k}{\mathbf{b}_k^H \mathbf{P}_k \mathbf{b}_k + \sigma_z^2} \quad (57)$$

Kalman filtering has been used with other models as well. In References [62, 63, 66, 90], AR-1 channel modeling is used along with Kalman-based tracking. In Reference [65], higher order AR modeling is used, and ARMA modeling is used in Reference [58].

5.2.4. Kalman LMS

In References [60] and [64], a series of approximations are used to the Kalman filter to obtain a lower complexity tracking approach similar to LMS. The key approximation is to average out the effect of the time-varying symbol vector $\mathbf{b}_k$, which causes the Kalman gain $\mathbf{K}_k$ to vary with time. The resulting 'Kalman LMS' (KLMS) approach was applied to two second-order models: the integrated random walk (IRW) model and the AR-2 model. More recently, the

KLMS IRW form was developed using a different approximation of the Kalman filter [67].

For a single channel tap, the KLMS tracking expressions for these two models are given by

$$\hat{\mathbf{s}}_{k+1} = \mathbf{F}\hat{\mathbf{s}}_k + \boldsymbol{\mu} b_k e_k^* \tag{58}$$

where $\boldsymbol{\mu}$ is a 2×1 vector of step sizes, and $\hat{\mathbf{s}}_k$ and $\mathbf{F}$ are defined earlier for the two models. The KLMS AR2 form is particularly useful for rapidly fading channels [17].

The IRW Kalman LMS form can alternatively be developed from degree-1 least squares prediction [68]. This approach was extended to track acceleration (second derivative) as well. In Reference [113], an RLS cost function is used to track the coefficient and its derivative.

5.2.5. Linear filtering based approaches

A simple tracking approach, often applied to one-tap channels, is linear prediction [116–118]. The approach is similar to interpolation with pilot symbols, except that the known symbols are the previous N_m symbols. Using the one-tap model in Equation (6), channel measurements at symbol times $k-1$ through $k-(N_m-1)$ can be obtained by multiplying r_k by b_k^* and normalizing by $|b_k|^2$ (typically 1). As in Equation (36), a weighted sum of these past measurements can be used to estimate the channel for symbol period $k+1$. Note that the measurements, not past estimates, are filtered. The weights can be learned adaptively [117]. The approach can also be used to predict the channel at some future iteration $k+D$ (see Section 6). Recently, low-complexity tracking approaches have been developed based on Wiener filtering [119].

6. Data-Directed Tracking

With the third example (IS-136 without pilot symbols), the channel is tracked over the data portion of the slot using the data symbols. In this section, approaches are discussed which detect symbol values and use these detected values with the tracking approaches presented in the previous section.

When the channel has significant dispersion (multiple taps), some form of equalization is needed at the receiver. Only equalization approaches that rely on channel estimates are discussed. Adaptation of other equalizer parameters, such as filter coefficients in linear and decision feedback equalizers, is not addressed.

Many of the data-directed approaches to channel estimation are designed for equalization based on either decision feedback equalization (DFE) or maximum likelihood sequence estimation (MLSE). These approaches are preferred over linear equalization, which is sensitive to spectral nulls in the channel response [13]. First, the basic principles for these equalizers are reviewed. Initial channel estimation using a training sequence is then discussed. Various data-directed tracking approaches are developed, showing the trade-offs between tracking delay, symbol value accuracy, and complexity. Finally, results for other forms of demodulation are discussed.

6.1. Equalization

Adaptive equalization has an extensive history [120, 121]. In narrowband wireless communication systems, DFE and MLSE are two commonly used forms of equalization. To understand the basic principles of these approaches, consider the simple case of binary modulation and a two-tap channel model, so that

$$r_k = g_k^*(0) b_k + g_k^*(1) b_{k-1} + z_k \tag{59}$$

The DFE, shown in Figure 9, consists of a feedforward filter, a feedback filter, and a decision device. Conceptually, the feedforward filter tries to collect all signal energy for b_k (which appears in both r_k and r_{k+1}) while suppressing intersymbol interference (ISI) from subsequent symbols (e.g., b_{k+1}). The feedback filter removes ISI from previous symbols (e.g., b_{k-1}). Notice that the feedforward filter introduces delay between when the first image of a symbol arrives and when that symbol is decided.

With MLSE, the likelihood of the received data samples, conditioned on the symbol values, is maximized. Two classic forms have been developed, assuming the channel is known [122, 123]. Here a form related to the Forney form [94] is used.

The conditional log-likelihood of the kth data sample, assuming z_k is Gaussian and white, is related to

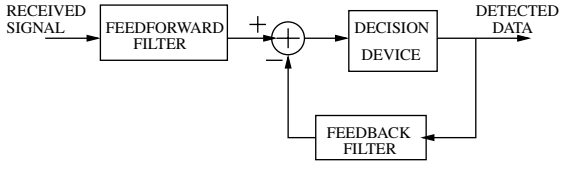

Fig. 9. DFE receiver.

the following metric or cost function:

$$M(b_k, b_{k-1}) = -|r_k - (g_k^*(0)b_k + g_k^*(1)b_{k-1})|^2 \quad (60)$$

This metric is accumulated for different symbol sequence hypotheses. The sequence corresponding to the largest metric determines the detected sequence. More intuitively, the negative of the metric, which should be minimized, indicates how well the model of the received data fits the actual received data.

A brute force search of all possible sequences would require high computational complexity. However, it is possible to determine the largest metric sequence through a process of hypothesis pruning known as the Viterbi algorithm [124].

The Viterbi algorithm is best explained using the trellis in Figure 10. Let $b_0 = -1$ be a known training symbol and $k > 0$ correspond to unknown data. For $k = 1$, there are two possible values for $M(b_1, b_0)$ corresponding to the two possible values for b_1. These two metrics correspond to the two branches shown in Figure (10) between $k = 0$ and $k = 1$. These two branches end in two 'states' corresponding to the possible values for b_1. At $k = 2$, there are four possible branches, two from each state, corresponding to the possible values for the pair (b_1, b_2). Observe that two branches end at the state corresponding to $b_2 = +1$ and two end at the state corresponding to $b_2 = -1$. At each state, only the 'best' sequence of the two is kept; the other will be 'pruned'. The best is determined by candidate 'path' metrics.

For example, at the state corresponding to $b_2 = +1$, the two sequences considered are $\{b_1 = +1,\ b_2 = +1\}$ and $\{b_1 = -1,\ b_2 = +1\}$. The corresponding candidate metrics are $M(b_2 = +1,\ b_1 = +1) + M(b_1 = +1,\ b_0 = -1)$ and $M(b_2 = +1,\ b_1 = -1) + M(b_1 = -1,\ b_0 = -1)$. Whichever candidate metric is largest determines which 'path' is kept and which 'path metric' is accumulated. Thus, each state has a 'path history' that determines previous symbol values.

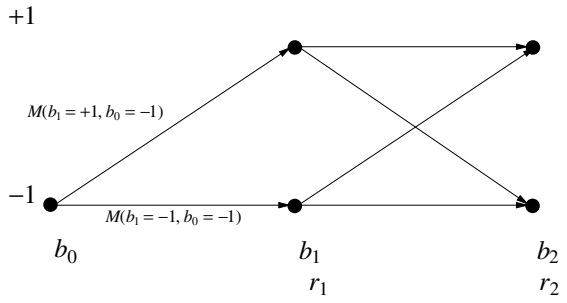

Fig. 10. Trellis for Viterbi algorithm.

Ideally, the process continues until all the data samples have been processed. Then, the state with the largest metric determines the detected sequence. In practice, a decision can be made after some delay D. The path history corresponding to the state with the best path metric after processing the kth sample can be used to determine the $(k - D)$th symbol value.

6.2. Initialization

The channel is initially estimated using the training sequence. An initial estimate can be obtained using the time-invariant approaches, such as the correlation [17, 94, 99, 125, 126], LS [111] or ML [110] approaches. This initial estimate can be refined by performing tracking over the training sequence [17, 94, 95]. Alternatively, tracking only may be used to obtain an initial estimate [41, 84, 93, 99, 127, 128]. Such an approach can also be used for the time-invariant case (example 1) [129].

6.3. Decision feedback

Early work on MLSE with data-directed tracking used decision feedback [82, 83, 123]. With this popular approach [84, 99, 104, 125, 130], symbol values are decided with delay D, and these detected symbols are used to update channel estimates. The final detected symbol values can be determined using a longer delay [94, 100]. In References [68] and [113], the concept of using tentative decisions for tracking is proposed for an arbitrary detector.

To obtain reliable symbol decisions, a certain delay is needed, which can seriously impact performance when the channel is varying rapidly. Part of this problem can be overcome by using a channel prediction approach [17, 92, 110, 127]. Channel prediction can be performed using either a form of linear prediction [92, 110, 127] or using a Kalman-based approach [17]. However, channel prediction becomes less reliable the larger D is. Thus, there is still a trade-off between accuracy of symbol values and tracking delay.

There are several other issues related to decision feedback tracking and MLSE. First, the 'best' path can suddenly change, so that the detected symbol sequence does not correspond to any one path history. Regularization can be used to improve channel tracking in this situation [42]. Second, due to symbol errors that occur in deep fades, the tracker can lock

onto the wrong phase [50, 130]. In other words, the tracker actually tracks a phase rotated version of the channel, consequently causing the detected symbols to be rotated. For example, when the channel estimate is phase rotated by 180°, the demodulator produces the opposite symbols (i.e., $g_k s_k = (-g_k)(-s_k)$). In Reference [50], a form of bi-directional channel tracking is proposed to resolve this problem.

Overmodelling the finite impulse response (FIR) channel can cause equalizer timing divergence after fading dips in fast fading channels [17]. Figure 11 shows an example where a two-tap model is used when only one tap is needed. During recovery from a deep fade, the position of the nonzero tap might change (false lock or time slip). When this happens, the channel is, in fact, being tracked, but with the wrong delay. A solution is given in Reference [17], where different tracking step sizes are used for different taps. Bi-directional tracking [131] can also be used to resolve this problem.

Related to decision feedback, there are also methods that try to account for or limit the effect of incorrect decisions. In Reference [60], total least squares (TLS) is suggested as a way to account for decision errors. Also, to limit the effects of incorrect decisions, limiting or clipping is applied to the update quantity.

Decision feedback for channel estimation can also be used with DFE [56, 62, 101, 102, 111, 132]. Various approaches for mapping channel estimates to DFE filter coefficients have been developed [133]. Similar to MLSE, delay can be addressed with prediction either using linear prediction [102] or a Kalman-based approach [62].

Decision feedback has also been applied to linear equalization. In Reference [134], transversal equalizer weights are determined by estimating the data autocorrelation and the cross-correlation between the received data and the symbol values. Decision feedback is used to estimate the cross-correlation term, which can be interpreted as a channel coefficient estimate. Finally, decision feedback can also be used without equalization [118].

A related form of decision feedback is multipass demodulation. In References [56] and [116], the data is demodulated a second time, using channel estimation based on decisions from the first pass.

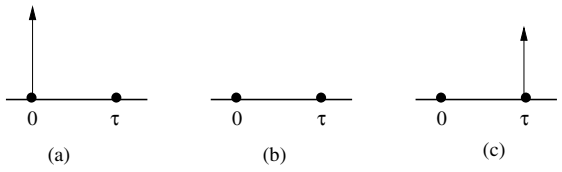

Fig. 11. Time slip problem in decision feedback: (a) before fade; (b) during fade; (c) after fade.

6.4. Per survivor processing

For MLSE, tracking delay can be eliminated by keeping different channel models for different states [86, 88]. This approach has been used in several receiver designs [63, 93]. Each state corresponds to a symbol sequence and a channel estimate associated with it. Thus, a channel tracker corresponding to that state can be updated without a decision delay. When forming candidate metrics for new states, the branch metrics are formed using the corresponding channel estimate. In the pruning process, the path that is kept also determines which channel model is updated for the new state. Such an approach is a form of per survivor processing (PSP) [88]. Note that it is possible to update the channel estimate before pruning [135]. Also, even for a one-tap channel, an MLSE approach with channel models for different states can be used for channel tracking purposes [49, 76, 116, 136].

To reduce complexity, the number of channel models can be less than the number of states. In Reference [89], channel models are kept for the best K states. In Reference [97], the states are partitioned and a channel estimate is maintained for each partition. In Reference [112], PSP or decision feedback tracking is adaptively selected based on the instantaneous signal power estimate.

Alternatively, there could be multiple channel models per state. For example, the List Viterbi [137, 138] algorithm maintains $K \geq 1$ best data sequences into each state. Therefore, each surviving path keeps its own channel estimate [85, 95, 139, 140].

When the channel is unknown, then ideally there would be a channel model for every possible symbol sequence [49, 91, 141]. The trellis in Figure 10 would then become an ever-expanding tree. To limit complexity, some form of tree-pruning is needed, leading to generalized-PSP approaches [91, 95], where associated with each surviving path of any tree-search algorithm is an estimate of the channel. In Reference [140], generalized PSP is used with the M-algorithm [142], which keeps the best M paths at each stage.

6.5. Data-directed tracking including pilot symbols

So far, it is assumed that there is a single training sequence to initialize data-directed channel tracking.

However, there may be other portions of the received data that may be treated as known. In Reference [17], a portion of the IS-136 slot is treated as known and used to retrain the channel estimates. In Reference [116], known portions are used in the Viterbi algorithm to constrain the trellis, which indirectly affects channel tracking. Pilot symbols may be used to resolve the phase ambiguity between the channel coefficient and the symbol values [76, 136].

It is also possible to use data-directed tracking in conjunction with pilot symbol interpolation, particularly when the interpolated estimates are not accurate enough. In Reference [44], interpolated estimates are used in conjunction with LMS updating to track the channel.

6.6. Other detectors

In References [90] and [143], Kalman type channel estimation is performed in conjugation with fixed-lag maximum *a posteriori* (MAP) symbol estimation. Alternatively, Kalman filtering can be used as a form of equalization for estimating the symbol sequence (Kalman equalization) [96, 144]. In practice, Kalman equalization can be performed in conjunction with channel tracking [96]. The Expectation–Maximization (EM) algorithm can be used to alternate between estimating the channel and estimating the symbol values [135, 145, 146]. Other approaches that alternate between channel estimation and data recovery have been developed [147–149].

Ultimately, channel estimation is an intermediate step in determining the symbol values. Approaches have been developed which directly detect the symbol values, folding in statistical information regarding the channel [150–154].

7. Conclusion

Channel estimation is an important part of receiver design in digital wireless communication systems. For a time-invariant channel, it was shown how channel estimation performance can be improved through careful design of training sequences as well as through use of prior information regarding the channel and the noise. For a time-varying channel, pilot symbols can be used to estimate the varying channel using interpolation. When pilot symbols are unavailable, data-directed tracking can be used. Tracking approaches are based on a model of how the channel changes in time. First-order models are often used for slowly varying channels, whereas higher order models are used for fast varying channels. These tracking approaches require knowing the data symbol values. When detected symbol values are used, there is a trade-off between the need for reliable symbol decisions and tracking delay. The tracking delay can be minimized by keeping multiple channel models, corresponding to different hypotheses of the data symbol values.

Currently, the next generation of digital wireless communication systems is being developed. As modulation formats for these systems require coherent reception, channel estimation will continue to be a key element in receiver design.

Acknowledgment

The authors would like to thank A. Duel-Hallen, A. Khayrallah, L. Lindbom, and K. Molnar for reviewing a draft of this paper. The authors gratefully acknowledge L. Lindbom for referring us to References [4] and [119], and K. Molnar for referring us to Reference [80]. Finally, the authors wish to thank their colleagues for many helpful discussions on channel estimation.

References

1. Tugnait JK, Tong L, Ding Z. Single-user channel estimation and equalization. *IEEE Sig. Proc. Mag.* 2000; **17**: 17–28.
2. Giannakis G. Highlights of statistical signal and array processing: Channel estimation and equalization. *IEEE Sig. Proc. Mag.* 1999; **15**: 37–40.
3. Ljung L. *System Identification*. Prentice Hall: Englewood Cliffs, NJ, 1987.
4. Söderström T, Stoica P. *System Identification*. Prentice Hall: Englewood Cliffs, NJ, 1989.
5. Clark AP. *Adaptive Detectors for Digital Modems*. Pentech: London, U.K., 1989.
6. Goodman DJ. Second generation wireless information networks. *IEEE Trans. Veh. Technol.* 1991; **40**: 366–374.
7. ETSI. GSM 05.02 *Digital Cellular Telecommunications Systems (Phase 2+): Multiplexing and multiple access on the radio path*, 1996/1997.
8. Furuskar A, Mazur S, Muller F, Olofson H. EDGE: Enhanced data rates for GSM and TDMA/136 evolution. *IEEE Pers. Commun. Mag.* 1999; **6**: 56–67.
9. Austin M, Buckley A, Coursey C, Hartman P, Kobylinski R, Majmundar M, Raith K, Seymour J. Service and system enhancements for TDMA digital cellular systems. *IEEE Personal Communications* 1999; **6**: 20–33.
10. TIA/EIA 136-131-B: *TDMA Third generation wireless: Digital Traffic Channel Layer 1*, 2000.
11. Van Trees HL. *Detection, Estimation, and Modulation Theory Part 3*. Krieger Publishing Co: Malabar, FL, 1992.
12. Bello PA. Characterization of randomly time-variant linear channels. *IEEE Trans. Commun. Sys.* 1963; **CS-11**: 360–393.
13. Proakis J. *Digital Communications* (3rd edn). McGraw-Hill: New York, U.S.A., 1995.

14. Clarke RH. A statistical theory of mobile radio reception. *Bell Sys. Tech. J.* 1968; **47**: 957–1000.
15. Jakes W. *Microwave Mobile Communications*. (1st edn). IEEE Press: Piscataway, NJ, U.S.A., 1993.
16. Liu Y-J, Wallace M, Ketchum JW. A soft-output bidirectional decision feedback equalization technique for TDMA cellular radio. *IEEE Journal on Sel. Areas in Commun.* 1993; **11**: 1034–1045.
17. Jamal K, Brismark G, Gudmundson B. Adaptive MLSE performance on the D-AMPS 1900 channel. *IEEE Trans. Veh. Technol.* 1997; **46**: 634–641.
18. Ziegler RA, Cioffi JM. Estimation of time-varying digital radio channels. *IEEE Trans. Veh. Technol.* 1992; **41**: 134–151.
19. Wautier A, Dany J-C, Mourot C, Kumar V. A new method for predicting the channel estimate influence on performance of TDMA mobile radio systems. *IEEE Trans. Commun.* 1995; **44**: 594–602.
20. Montemayor CA, Flikkema PG. Near-optimum iterative estimation of dispersive multipath channels. *Proc. IEEE Veh. Technol. Conf.*, Ottawa, Ontario, Canada, 1998; 2246–2250.
21. Song H-K. A channel estimation using sliding window approach and tuning algorithm for MLSE. *IEEE Commun. Letters* 1999; **3**: 211–213.
22. Benelli G, Garzelli A, Salvi F. Simplified Viterbi processors for the GSM Pan-European cellular communication system. *IEEE Trans. Veh. Technol.* 1994; **43**: 870–877.
23. Steele R. *Mobile Radio Communications.* IEEE Press: New York, U.S.A., 1992.
24. Baier A, Heinrich G, Wellens U. Bit synchronization and timing sensitivity in adaptive Viterbi equalizers for narrowband-TDMA digital mobile radio systems. *Proc. IEEE Veh. Technol. Conf.*, 1988; 377–384.
25. Ariyavisitakul S, Greenstein LJ. Reduced-complexity equalization techniques for broadband wireless channels. *IEEE Journal on Sel. Areas in Commun.* 1997; **15**: 5–15.
26. Park JI, Wicker WB, Owen HL. Trellis-based soft-output adaptive equalization for TDMA cellular systems. *IEEE Trans. Veh. Technol.* 2000; **49**: 83–94.
27. Ng JCL, Letaief KB, Murch RD. Complex optimal sequences with constant magnitude for fast channel estimation initialization. *IEEE Trans. Commun.* 1998; **46**: 305–308.
28. Milewski A. Periodic sequences with optimal properties for channel estimation and fast start-up equalization. *IBM J. Res. Develop.* 1983; **27**: 426–431.
29. Tellambura C, Guo YJ, Shepherd SJ, Barton SK. Optimal sequences for channel estimation using discreet fourier transform techniques. *IEEE Trans. Commun.* 1999; **47**: 230–238.
30. Fan MDP. *Sequence Design for Communications Applications*. John Wiles & Sons: New York, U.S.A., 1996.
31. Crozier SN, Falconer DD, Mahmoud SA. Least sum of squared errors (LSSE) channel estimation. *IEEE Proceedings-F* 1991; **138**: 371–378.
32. Lee H-N, Pottie GJ. Fast adaptive equalization/diversity combining for time-varying dispersive channels. *IEEE Trans. Commun.* 1998; **46**: 1146–1162.
33. Gorokhov A. On the performance of the Viterbi equalizer in the presence of channel estimation errors. *IEEE Sig. Proc. Letters* 1998; **5**: 321–324.
34. Dzung D. Error probability of MLSE equalization using imperfect channel estimation. *Proc. IEEE Int. Conf. Commun.* Denver, CO, 1991; 19.4.1–19.4.5.
35. Guey J-C, Fitz MP, Bell MR, Kuo W-Y. Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels. *IEEE Trans. Commun.* 1999; **47**: 527–537.
36. Ling F. Optimal reception, performance bound, and cutoff rate analysis of reference-assisted coherent CDMA communications with applications. *IEEE Trans. Commun.* 1999; **47**: 1583–1592.
37. Khayrallah AS, Ramesh R, Bottomley GE, Koilpillai D. Improved channel estimation with side information. *Proc. IEEE Veh. Technol. Conf.*, Phoenix, AZ, 1997; 1049–1053.
38. Ng BC, Cedervall M, Paulraj A. A structured channel estimator for maximum likelihood sequence detection. *IEEE Commun. Letters* 1997; **1**: 52–55.
39. Molnar KJ, Bottomley GE, Ramesh R. A novel fractionally spaced MLSE receiver and channel tracking with side information. *Proc. IEEE Veh. Technol. Conf.*, Ottawa, Ontario, Canada, 1998; 2251–2255.
40. Hafeez A, Molnar KJ, Bottomley GE. Co-channel interference cancellation for D-AMPS handsets. *Proc. IEEE Veh. Technol. Conf.*, Houston, TX, 1999; 1026–1030.
41. Denno S, Saito Y. Orthogonal-transformed variable-gain least mean squares (OVLMS) algorithm for fractional tap-spaced adaptive MLSE equalizers. *IEEE Trans. Commun.* 1999; **47**: 1151–1160.
42. Martone M. Optimally regularized channel tracking techniques for sequence estimation based on cross-validated subspace signal processing. *IEEE Trans. Commun.* 2000; **48**: 95–105.
43. Kamio Y, Sampei S. Performance of reduced complexity DFE using bidirectional equalization in land mobile communication. *Proc. IEEE Veh. Technol. Conf.*, Denver, CO, 1992; 372–375.
44. Davidson GW, Falconer DD, Sheikh AUH, An investigation of block adaptive decision feedback equalization for frequency selective fading channels. *Proc. IEEE Int. Conf. Commun.* Atlanta, GA, 1998; 360–365.
45. Sampei S. Computation reduction of decision feedback equalizer using interpolation for land mobile communications. *Proc. IEEE Globecom Conf.*, Phoenix, AZ, 1991; 16.1.1–16.1.5.
46. Fechtel SA, Meyr H. Optimal parametric feedforward estimation of frequency-selective fading radio channels. *IEEE Trans. Commun.* 1994; **42**: 1639–1650.
47. Cavers JK. An analysis of pilot symbol assisted modulation for Rayleigh fading channels. *IEEE Trans. Veh. Technol.* 1991; **40**: 686–693.
48. Kuo W-Y, Fitz MP. Designs for pilot-symbol-assisted burst-mode communications with fading and frequency uncertainty. *Int. Journal of Wireless Inf. Networks* 1994; **1**: 239–252.
49. Meyr SFH, Moeneclaey M. *Digital Communication Receivers: Synchronization, Channel Estimation and Signal Processing*. John Wiley and Sons: New York, USA, 1998.
50. Arslan H, Ramésh R, Mostafa A. Interpolation and channel tracking based receivers for coherent Mary-PSK modulations. *Proc. IEEE Veh. Technol. Conf.*, Houston, TX, 1999; 2194–2199.
51. Sampei S, Sunaga T. Rayleigh fading compensation for QAM in land mobile radio communications. *IEEE Trans. Veh. Technol.* 1993; **42**: 137–147.
52. Lo NWK, Falconer DD, Sheikh AUH. Adaptive equalization and diversity combining for mobile radio using interpolated channel estimates. *IEEE Trans. Veh. Technol.* 1991; **40**: 636–645.
53. Koilpillai RD, Chennakeshu S, Toy RT. Low complexity equalizers for U.S. digital cellular system. *Proc. IEEE Veh. Technol. Conf.*, Denver, CO, 1992; 744–747.
54. Eyceöz T, Duel-Hallen A. Simplified block adaptive diversity equalizer for cellular mobile radio. *IEEE Commun. Letters* 1997; **1**: 15–18.
55. Sampei S. *Applications of Digital Wireless Technologies to Global Wireless Communications*. Prentice Hall: New Jersey, U.S.A., 1997.
56. Borah DK, Hart BD. Receiver structures for time-varying frequency-selective fading channels. *IEEE Journal on Sel. Areas in Commun.* 1999; **17**: 1863–1875.
57. Haykin S. *Adaptive Filter Theory (third edition)*. Prentice Hall: New Jersey, U.S.A., 1996.

58. Dai Q, Shwedyk E. Detection of bandlimited signals over frequency selective Raleigh fading channel. *IEEE Trans. Commun.* 1994; **42**: 941–950.
59. Anderson BDO, Moore JB. *Optimal Filtering*. Prentice–Hall: Englewood Cliffs, NJ, 1979.
60. Lindbom L. Adaptive equalization for fading mobile radio channels. *Licentiate Thesis*, Technol. Dept., Uppsala Univ., Uppsala, Sweden, 1992.
61. Liu W. Performance of joint data and channel estimation using tap variable step size LMS for multipath fast fading channel. *Proc. IEEE Globecom Conf.*, San Francisco, CA, 1994; 973–978.
62. Stojanovic M, Proakis JG, Catipovic JA. Analysis of the impact of channel estimation errors on the performance of a decision-feedback equalizer in fading multipath channels. *IEEE Trans. Commun.* 1995; **43**: 877–885.
63. Rollins ME, Simmons SJ. Simplified per-survivor Kalman processing in fast frequency-selective fading channels. *IEEE Trans. Commun.* 1997; **45**: 544–552.
64. Lindbom L. Simplified Kalman estimation of fading mobile radio channels: High performance at LMS computational load. *Proc. IEEE Int. Conf. Acoust., Speech, Sig. Proc.*, 1993; 352–355.
65. Zamiri-Jafarian H, Pasupathy S. Adaptive MLSD receiver with identification of flat fading channels. *Proc. IEEE Veh. Technol. Conf.*, Phoenix, AZ, 1997; 695–699.
66. Davis LM, Collings IB, Evans RJ. Coupled estimators for equalization of fast fading mobile channels. *IEEE Trans. Commun.* 1998; **46**: 1262–1265.
67. Gazor S. Prediction in LMS-type adaptive algorithms for smoothly time varying environments. *IEEE Trans. Sig. Proc.*, 1999; **47**: 1735–1739.
68. Clark AP. Adaptive channel estimator for an HF radio link. *IEEE Trans. Commun.* 1989; **37**: 918–926.
69. Duel-Hallen A, Hu S, Hallen H. Long-range prediction of fading signals. *IEEE Sig. Proc. Mag.* 2000; **17**: 62–75.
70. Eyceoz T, Duel-Hallen A, Hallen H. Deterministic channel modeling and long range prediction of fast fading mobile radio channels. *IEEE Commun. Letters* 1998; **2**: 254–256.
71. Andersen JB, Jensen J, Jensen SH, Frederiksen F. Prediction of future fading based on past measurements. *Proc. IEEE Veh. Technol. Conf.*, Amsterdam, Netherland, 1999; 151–155.
72. Hwang JK, Winters JH. Sinusoidal modeling and prediction of fast fading process. *Proc. IEEE Globecom Conf.*, 1998; 892–897.
73. Ekman T, Kubin G. Nonlinear prediction of mobile radio channels: Measurements and Mars model design. *Proc. IEEE Int. Conf. Acoust., Speech, Sig. Proc.*, Phoenix, AZ, 1999; 2667–2670.
74. Rummler WD, Coutts RP, Liniger M. Multipath fading channel models for microwave digital radio. *IEEE Commun. Mag.* 1986; **24**: 30–42.
75. Zhou N, Holte N. Least squares channel estimation for a channel with fast time variations. *Proc. IEEE Int. Conf. Acoust., Speech, Sig. Proc.*, 1992; 165–168.
76. Vitetta GM, Taylor DP. Multisampling receivers for uncoded and coded PSK signal sequences transmitted over Rayleigh frequency-flat fading channels. *IEEE Trans. Commun.* 1996; **44**: 130–133.
77. Borah DK, Hart BD. Frequency-selective fading channel estimation with a polynomial time-varying channel model. *IEEE Trans. Commun.* 1999; **47**: 862–873.
78. Alexander ST. *Adaptive Signal Processing*. Springer-Verlag: New York, U.S.A., 1986.
79. Rappaport T. *Wireless Communications*. Prentice–Hall: New York, U.S.A., 1996.
80. Glentis G-O, Berberidis K, Theodoridis S. Efficient least squares adaptive algorithms for FIR transversal filtering. *IEEE Sig. Proc. Mag.* 1999; **16**: 13–41.
81. Ljung L, Gunnarsson S. Adaptation and tracking in system identification: a survey. *Automatica* 1990; **26**: 7–21.
82. Magee FR, Proakis JG. Adaptive maximum-likelihood sequence estimation for digital signaling in the presence of intersymbol interference. *IEEE Trans. Info. Theory* 1973; **18**: 120–124.
83. Qureshi SUH, Newhall EE. An adaptive receiver for data transmission over time-dispersive channels. *IEEE Trans. Info. Theory* 1973; **IT-19**: 448–457.
84. Wan Y, Liu Q, Sendyk A. A fractionally spaced maximum likelihood sequence estimation receiver in multipath fading environment. *Proc. IEEE Int. Conf. Acoust., Speech, Sig. Proc.*, 1992; 689–692.
85. Seshadri N. Joint data and channel estimation using blind trellis search techniques. *IEEE Trans. Commun.* 1994; **42**: 1000–1011.
86. Kubo H, Murakami K, Fujino T. An adaptive maximum-likelihood sequence estimator for fast time-varying intersymbol interference channels. *IEEE Trans. Commun.* 1994; **42**: 1872–1880.
87. Kubo H, Murakami K, Fujino T. Adaptive maximum-likelihood sequence estimation by means of combined equalization and decoding in fading environments. *IEEE Journal on Sel. Areas in Commun.* 1995; **13**: 102–109.
88. Raheli R, Polydoros A, Tzou C-K. Per-survivor processing: A general approach to MLSE in uncertain environments. *IEEE Trans. Commun.* 1995; **43**: 354–364.
89. Raheli R, Marino G, Castoldi P. Per-survivor processing and tentative decisions: what is in between?. *IEEE Trans. Commun.* 1996; **44**: 127–129.
90. Giridhar K, Shynk JJ, Iltis RA, Mathur A. Adaptive MAPSD algorithms for symbol and timing recovery of mobile radio TDMA signals. *IEEE Trans. Commun.* 1996; **44**: 976–987.
91. Chugg KM, Polydoros A. MLSE for an unknown channel—part 2: tracking performance. *IEEE Trans. Commun.* 1996; **44**: 949–958.
92. Chiu M-C, Chao C-C. Analysis of LMS-adaptive MLSE equalization on multipath fading channels. *IEEE Trans. Commun.* 1996; **44**: 1684–1692.
93. Hamied KA, Stüber GL. An adaptive truncated MLSE receiver for Japanese personal digital cellular. *IEEE Trans. Veh. Technol.* 1996; **45**: 41–50.
94. Bottomley GE, Chennakeshu S. Unification of MLSE receivers and extension to time-varying channels. *IEEE Trans. Commun.* 1998; **46**: 464–472.
95. Bontu CS, Falconer DD, Strawczynski L. Diversity transmission and adaptive MLSE for digital cellular radio. *IEEE Trans. Veh. Technol.* 1999; **48**: 1488–1502.
96. Marcos S. A network of adaptive Kalman filters for data channel equalization. *IEEE Trans. Sig. Proc.* 2000; **48**: 2620–2627.
97. Bradley MJ, Mars P. Application of multiple channel estimators in MLSE detectors for fast time-varying and frequency selective channels. *Electronics Letters* 1996; **32**: 620–621.
98. Chugg KM, Polydoros A. MLSE for an unknown channel—part 1: Optimality considerations. *IEEE Trans. Commun.* 1996; **44**: 836–846.
99. Newson P, Mulgrew B. Adaptive channel identification and equalization for GSM European digital mobile radio. *Proc. IEEE Int. Conf. Commun.*, Denver, CO, 1991; 1.5.1–1.5.5.
100. Castellini G, Conti F, Del Re E, Pierucci L. A continuously adaptive MLSE receiver for mobile communications: algorithm and performance. *IEEE Trans. Commun.* 1997; **45**: 80–89.
101. Shukla PK, Turner LF. Channel-estimation-based adaptive DFE for fading multipath radio channels. *IEEE Proc. I*, 1991; **138**: 525–543.

102. Fechtel SA, Meyr H. An investigation of channel estimation and equalization techniques for moderately rapid HF-channels. *Proc. IEEE Int. Conf. Commun.*, Denver, CO, 1991; 25.2.1–25.2.5.
103. Liu W, Romanens H. Adaptive channel equalization for high-speed train. *Proc. IEEE Veh. Technol. Conf.*, 1994; 225–229.
104. Shiino H, Yamaguchi N, Shoji Y. Performance of an adaptive maximum-likelihood receiver for fast fading multipath channel. *Proc. IEEE Veh. Technol. Conf.*, Denver, CO, 1992; 380–383.
105. Eleftheriou E, Falconer DD. Tracking properties and steady state performance of RLS adaptive filter algorithms. *IEEE Trans. ASSP* 1986; **ASSP-34**: 1097–1110.
106. Lin J, Proakis JG, Ling F, Lev-ari H. Optimal tracking of time-varying channels: A frequency domain approach for known and new algorithms. *IEEE Journal on Sel. Areas in Commun.* 1995; **13**: 142–154.
107. Haykin S. *Adaptive Filter Theory (second edition)*. Prentice–Hall: Englewood Cliffs, NJ, 1991.
108. Sayed AH, Kailath T. A state-space approach to adaptive RLS filtering. *IEEE Sig. Proc. Mag.* 1994; **11**: 18–60.
109. Haykin S, Sayed AH, Zeidler JR, Yee P, Wei PC. Adaptive tracking of linear time-variant systems by extended RLS algorithms. *IEEE Trans. Sig. Proc.* 1997; **45**: 1118–1128.
110. Wu J, Aghvami H. A new adaptive equalizer with channel estimator for mobile radio communications. *IEEE Trans. Veh. Technol.* 1996; **45**: 467–474.
111. Ziegler R, Al-Dhahir N, Cioffi J. Nonrecursive adaptive decision-feedback equalization from channel estimates. *Proc. IEEE Veh. Technol. Conf.*, Denver, CO, 1992; 600–603.
112. Suk J, Lee YH. An adaptive MLSE receiver for TDMA systems: A hybrid of per-survivor processing and tentative decision MLSE. *Proc. IEEE Veh. Technol. Conf.*, Amsterdam, Netherlands, 1999; 712–716.
113. Clark AP, Hariharan S. Efficient estimators for an HF radio link. *IEEE Trans. Commun.* 1990; **38**: 1173–1180.
114. Bottomley GE, Molnar KJ. Adaptive channel estimation for multichannel MLSE receivers. *IEEE Commun. Letters* 1999; **3**: 40–42.
115. Bottomley GE, Molnar KJ. Interference cancellation for improved channel estimation in array processing MLSE receivers. *Proc. IEEE Veh. Technol. Conf.*, Phoenix, AZ, 1997; 140–144.
116. D'Andrea AN, Diglio A, Mengali U. Symbol-aided channel estimation with non-selective Rayleigh fading channels. *IEEE Trans. Veh. Technol.* 1995; **44**: 41–49.
117. Young RJ, Lodge JH. Detection of CPM signals in fast Rayleigh flat-fading using adaptive channel estimation. *IEEE Trans. Veh. Technol.* 1995; **44**: 338–347.
118. Bin L, Ho P. Data-aided linear prediction receiver for coherent DPSK and CPM transmitted over Rayleigh flat-fading channels. *IEEE Trans. Veh. Technol.* 1999; **48**: 1229–1236.
119. Ahlen A, Lindbom L, Sternad M. Channel tracking with WLMS algorithms: high performance at LMS computational load. *Proc. IEEE Veh. Technol. Conf.*, Tokyo, Japan, 2000; 16–20.
120. Lucky RW. A survey of the communication theory literature: 1968–1973. *IEEE Trans. Info. Theory* 1973; **19**: 725–739.
121. Proakis JG. Adaptive equalization for TDMA digital mobile radio. *IEEE Trans. Veh. Technol.* 1991; **40**: 333–341.
122. Forney GD. Maximum-likelihood sequence estimation of digital sequences in presence of ISI. *IEEE Trans. on Inf. Theory* 1972; **IT-18**: 363–378.
123. Ungerboeck G. Adaptive maximum-likelihood receiver for carrier-modulated data-transmission systems. *IEEE Trans. Commun.* 1974; **22**: 624–636.
124. Forney GD. The Viterbi algorithm. *Proc. of the IEEE* 1973; **61**: 268–277.
125. D'Avella R, Moreno L, Sant'Agostino M. An adaptive MLSE receiver for TDMA digital mobile radio. *IEEE Journal on Sel. Areas in Commun.* 1989; **7**: 122–129.
126. Larsson G, Gudmundson B, Raith K. Receiver performance of the North American digital cellular system. *Proc. IEEE Veh. Technol. Conf.*, St. Louis, MI, 1991; 1–6.
127. Dahlman E. New adaptive Viterbi detector for fast-fading mobile-radio channels. *Electronics Letters* 1990; **26**: 1572–1573.
128. Lin J, Ling F, Proakis J. Fading channel tracking properties of several adaptive algorithms for the Northern American digital cellular systems. *Proc. IEEE Veh. Technol. Conf.*, Secaucus, NJ, 1993; 273–276.
129. Cheung JC, Steele R. Soft-decision feedback equalizer for continuous phase modulated signals in wideband mobile radio channels. *IEEE Trans. Commun.* 1994; 1628–1638.
130. Shukla PK, Turner LF. Examination of an adaptive DFE and MLSE/near-MLSE for fading multipath radio channels. *IEE Proc.-I* 1992; **129**: 418–428.
131. Liu Y-J, Wallace M, Ketchum JW. A soft-output bidirectional decision feedback equalization technique for TDMA cellular radio. *IEEE Journal on Sel. Areas in Commun.* 1993; **11**: 1034–1045.
132. Farhang-Boroujeny B. Channel equalization via channel identification: algorithms and simulation results for rapidly fading HF channels. *IEEE Trans. Commun.* 1996; **44**: 1409–1412.
133. Al-Dhahir N, Cioffi JM. Fast computation of channel-estimate based equalizers in packet data transmission. *IEEE Trans. Sig. Proc.* 1995; **43**: 2462–2473.
134. Butler P, Cantoni A. Noniterative automatic equalization. *IEEE Trans. Commun.* 1975; **COM-23**: 621–633.
135. Zamiri-Jafarian H, Pasupathy S. Adaptive MLSDE using the EM algorithm. *IEEE Trans. Commun.* 1999; **47**: 1181–1193.
136. Vitetta GM, Taylor DP. Maximum likelihood decoding of uncoded and coded PSK signal sequences transmitted over Rayleigh flat-fading channels. *IEEE Trans. Commun.* 1995; **43**: 2750–2758.
137. Hashimoto T. A list-type reduced-constraint generalization of the Viterbi algorithm. *IEEE Trans. Inf. Theory* 1987; **IT-33**: 866–876.
138. Seshadri N, Sundberg C-EW. List Viterbi decoding algorithms with applications. *IEEE Trans. Commun.* 1994; **42**: 313–323.
139. Lin J, Ling F, Proakis J. Joint data and channel estimation for TDMA mobile channels. *Proc. IEEE PIMRC Conf.*, Boston, MA, 1992; 235–239.
140. Castoldi P, Raheli R, Marino G. Efficient trellis search algorithms for adaptive MLSE on fast Raleigh fading channels. *Proc. CTMC-Globecom*, San Francisco, CA, 1994; 196–200.
141. Chugg KM. The condition for the applicability of the Viterbi algorithm with implications for fading channel MLSD. *IEEE Trans. Commun.* 1998; **46**: 1112–1116.
142. Anderson JB, Mohan S. Sequential coding algorithms: A survey and cost analysis. *IEEE Trans. Commun.* 1984; **COM-32**: 169–176.
143. Baccarelli E, Cusani R, Galli S. A novel adaptive receiver with enhanced channel tracking capability for TDMA-based mobile radio communications. *IEEE Journal on Sel. Areas in Commun.* 1998; **16**: 1630–1639.
144. Thielecke J. A soft-decision state-space equalizer for FIR channels. *IEEE Trans. Commun.* 1997; **45**: 1208–1217.
145. Georghiades N, Han JC. Sequence estimation in the presence of random parameters via the EM algorithm. *IEEE Trans. Commun.* 1997; **45**: 300–308.
146. Zamiri-Jafarian H, Pasupathy S. EM-based recursive estimation of channel parameters. *IEEE Trans. Commun.* 1999; **47**: 1297–1302.

147. Chang K-H, Georghiades CN. Iterative joint sequence and channel estimation for fast time-varying intersymbol interference channels. *Proc. IEEE Int. Conf. Commun.* Seattle, WA, 1995.
148. Yuan WS, Gheorghiades CN. An application of the generalized likelihood function to sequence estimation for frequency nonselective Rayleigh fading channels. *29th Annual Conf. on Info. Sciences and Systems (CISS)*, Baltimore, MD, 1995; 408–413.
149. Kong H, Schwedyk E. Sequence detection and channel state estimation over finite state Markov channels. *IEEE Trans. Veh. Technol.* 1999; **48**: 833–839.
150. Yu X, Pasupathy S. Innovations-based MLSE for Rayleigh fading channels. *IEEE Trans. Commun.* 1995; **43**: 1534–1544.
151. Seymour JP, Fitz MP. Near-optimal symbol-by-symbol detection schemes for flat Rayleigh fading. *IEEE Trans. Commun.* 1995; **43**: 1525–1533.
152. Matolak DW, Wilson SG. Detection for a statistically known, time-varying dispersive channel. *IEEE Trans. Commun.* 1996; **44**: 1673–1983.
153. Hart BD, Taylor DP. Maximum likelihood synchronization, equalization, and sequence estimation for unknown time-varying frequency-selective Rician channels. *IEEE Trans. Commun.* 1998; **46**: 211–221.
154. Hart BD. Maximum likelihood sequence detection using a pilot tone. *IEEE Trans. Veh. Technol.* 2000; **49**: 550–560.

Authors' Biographies

Hüseyin Arslan (eushura@rtp.ericsson.se) was born in Nazilli, Turkey, in 1968. He received the B.S. degree from Middle East Technical University, Ankara, Turkey, and M.S. and Ph.D. degrees from Southern Methodist University, Dallas, Texas, USA, in 1992, 1994, and 1998, respectively, all in electrical engineering. Since January 1998, he has been with Advanced Development and Research group of Ericsson Inc., RTP, North Carolina, USA. His research interests are in baseband signal processing for mobile communications, including interference cancellation, channel estimation, and equalization. Dr. Arslan is a Member of The Institute of Electrical and Electronics Engineers, Inc. (IEEE).

Gregory E. Bottomley received the B. S. and M. S. degrees from Virginia Polytechnic Institute and State University, Blacksburg, USA in 1983 and 1985, respectively, and the Ph. D. degree from North Carolina State University, Raleigh, USA in 1989, all in electrical engineering.

From 1985 to 1987 he was with AT&T Bell Laboratories, Whippany, NJ, USA, working in the area of sonar signal processing. In 1990, he was a Visiting Lecturer at North Carolina State University, Raleigh, USA. Since 1991, he has been with Ericsson Inc., Research Triangle Park, NC, USA, where he is currently a Senior Consulting Engineer in the Advanced Development and Research Department. His research interests are in baseband signal processing for mobile communications, including equalization, RAKE reception, and interference cancellation.

Dr. Bottomley is an associate member of Sigma Xi and a Senior Member of The Institute of Electrical and Electronics Engineers, Inc. (IEEE). He served as Associate Editor and is currently Editor for the *IEEE Transactions On Vehicular Technology*.