Adaptive Synchronization and Channel Parameter Estimation Using an Extended Kalman Filter

ABBAS AGHAMOHAMMADI, STUDENT MEMBER, IEEE, HEINRICH MEYR, FELLOW, IEEE, AND GERD ASCHEID, MEMBER, IEEE

Abstract—**Carrier-phase synchronization can be approached in a general manner by estimating the multiplicative distortion (MD) to which a baseband received signal in an RF or coherent optical transmission system is subjected. This paper presents a unified modeling and estimation of the MD in finite-alphabet digital communication systems. A simple form of MD is the carrier phase exp ($j\theta$) which has to be estimated and compensated for in a coherent receiver. A more general case with fading must, however, allow for amplitude as well as phase variations of the MD.**

We assume a state-variable model for the MD and generally obtain a nonlinear estimation problem with additional randomly-varying system parameters such as received signal power, frequency offset, and Doppler spread. An extended Kalman filter is then applied as a near-optimal solution to the adaptive MD and channel parameter estimation problem. Examples are given to show the use and some advantages of this scheme.

I. Introduction

The power efficiency provided by coherent detection in digital communication systems is only possible when the receiver is supplemented by a carrier phase synchronization, or said generally, a multiplicative distortion (MD) estimation unit. Such a unit, which may or may not have access to modulation-free sections of the carrier, can optimally estimate the MD from the received signal by modeling the dynamics generating the MD. These models have in general additional unknown or randomly varying parameters and a good estimate of them might be essential for successful synchronization. For example, in a UHF mobile communication system the MD estimation is mostly concerned with signals which strongly depend on parameters such as the vehicle speed and received signal power (Section IV).

The problems of identification, state estimation with tracking, and adaptive control of systems with unknown parameters have been studied mainly in the fields of control and aerospace for over 20 years ([1]–[4], for example). In the area of signal processing similar ideas are known as "adaptive algorithms" whereas in statistics the methods are usually called "sequential parameter estimation." A coherent picture and analysis of recursive identification composed of the many approaches used in each of these disciplines is presented in [5].

The major contribution of this paper is a formulation of the extended Kalman filter [6] which can be applied in a unifying manner to many MD estimation problems in digital communication systems with additional unknown or randomly varying parameters. This opens the way for novel MD estimation schemes and the improvement of some current carrier synchronization techniques. Similar nonlinear estimation methods [10]–[14] which are extensions of [7]–[9] are suggested in the context of quasioptimum angle demodulation. All these methods are in principle related to the example considered in Section IV-A. The demodulation structure for fading channels [15] seems to have similarities to our example, Section IV-B. However, a main difference is that we face an easy-to-eliminate *digital* message and concentrate on the synchronization problem while the above work has to jointly estimate the fading distortion and the *analog* FM message.

A popular approach for carrier recovery [16] is the maximum likelihood (ML) method where an appropriate likelihood function of the carrier phase taken as unknown and essentially constant over an observation interval is maximized and no assumption about the *a priori* probability densities (except for the additive noise) is made. For a simple additive white gaussian noise (AWGN) channel, other approaches (e.g., Bayesian) which use some *a priori* knowledge about the carrier phase distribution yield similar results because the module-2π nature of phase effectively produces a uniform distribution. This can, for example, be seen in Fig. 4 where the phase of y_k has nearly a uniform distribution over any interval of 2π even though the phase θ_k is actually gaussian. However, in the case where the MD is a two-dimensional (complex, baseband) signal as in Fig. 6, we are not working with periodic quantities. Then, by relying on certain *a priori* statistical information on both the MD and the additive noise we can aim to deliver an optimum, in the sense of minimum mean-square error (MMSE), estimate of the MD recursively. Such statistical information can be provided in many cases of interest (e.g., Section IV-B) and results in improvements over the ML method.[1] Other disadvantages of the ML method are that an extensive search may be required to find the optimal estimate when the density function has several peaks and that the favorable properties, the unbiasedness and consistency of the estimates, do not necessarily hold if the unknown quantity is actually time varying.

Section II defines a general model used to represent digital communications in baseband. The resulting complex-valued, noisy received signal and its random parameters are the subjects of a joint estimation algorithm in Section III using an extended Kalman filter. Section IV presents the examples of MD estimation for AWGN channels, mobile communication systems, and receivers with randomly varying frequency offset and received signal power.

II. Signal Model

Consider the complex-valued baseband linearly modulated received signal of the form $r(t) = \Sigma_k u_k y(t) h_c(t - kT) + v_r(t)$ generated by using a fixed local carrier. Here, u_k represents the finite-alphabet digital modulation with $E[|u_k|^2] = 1$, $y(t)$ the MD with $E[|y(t)|^2] = 1$, $h_c(t)$ the convolution of the signaling pulse and the channel impulse-response, and $v_r(t)$ the AWGN with one-sided power spectral density (PSD) of N_0. The goal of a general baseband "carrier synchronizer" is to provide a good estimate of the MD $y(t)$ which may

Paper approved by the Editor for Synchronization Systems and Techniques of the IEEE Communications Society. Manuscript received March 11, 1988.

The authors are with Aachen University of Technology (RWTH), Lehrstuhl fur elektrische Regelungstechnik 5240, 5100 Aachen, West Germany.

IEEE Log Number 8930928.

[1] It can be proved that on a channel with memory the detection process is optimized under the minimization of probability-of-error criterion by employing a MMSE estimator.

0090-6778/89/1100-1212$01.00

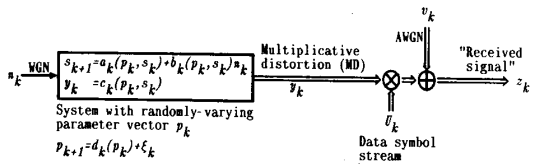


Fig. 1. Baseband discrete-time model of a digital communications system. (Perfect symbol timing and absence of intersymbol interference assumed.)

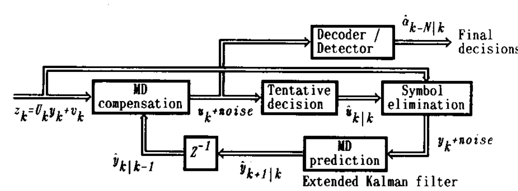


Fig. 2. Full-digital baseband decision-aided MD estimation and compensation.

include phase/frequency errors of the mixers, time-varying amplitude and phase introduced by a frequency-flat fading channel, etc. For the synchronizer, however, the modulation u_k plays the role of an MD and has to be somehow eliminated.

We pass $r(t)$ through a matched filter with impulse response $h_c^*(-t)/E$ and assume that the combination of $h_c(t)$ and the matched filter forms a Nyquist pulse for intersymbol-interference-free reception [17], i.e., $\int h_c^*(-\tau)h_c(kT - \tau)d\tau = E\delta(k)$ where E is the received energy per channel symbol. Then sampling the result at the channel symbol rate $1/T$ with correct symbol timing and assuming that $y(t)$ changes little over a time period equal to the duration of $h_c(t)$ yields the sufficient statistic representation $z_k = u_k y_k + v_k$ where v_k can be shown to be a discrete-time AWGN with average power N_0/E, $y_k := y(kT)$ and $E[|u_k y_k|^2] = 1$. For the rest of the paper, we refer to z_k as the received signal.[2]

The discrete-time model shown in Fig. 1 can describe a large class of structures in digital communications represented in baseband. For the sake of synchronization, this model emphasizes the MD in opposed to the modulation. Using real-vector notation hereafter, the input zero mean white Gaussian noise (WGN) vector n_k has an arbitrary dimension whereas y_k, v_k, and z_k are two-dimensional vectors formed from the cartesian coordinates (in/quadrature-phase (I/Q) components) of the corresponding complex-valued baseband signals. The I/Q components of the data symbol u_k form the following 2×2 matrix or 2×1 vector:

$$U_k = \begin{bmatrix} u_k^I & -u_k^Q \\ u_k^Q & u_k^I \end{bmatrix}; \quad u_k = \begin{bmatrix} u_k^I \\ u_k^Q \end{bmatrix}.$$

The MD y_k with the generalized role of the "carrier phase" is generated by a nonlinear dynamical system. This system has the explicit state variable s_k and an implicit dependence on the parameter vector p_k which itself has a nonlinear state-variable description and is assumed to be varying more slowly than the state s_k. In the next section s_k and p_k will be put together in a joint state vector.

The received signal, written as

$$z_k = U_k y_k + v_k \tag{1}$$

must be manipulated for the sake of the detection of the transmitted data symbol stream (see Fig. 2). Specifically, an estimate $\hat{y}_{k|k-1}$ of y_k, given $z_0 \cdots z_{k-1}$ and (estimates of) $u_0 \cdots u_{k-1}$, is used to eliminate the effect of the MD y_k and leaves us with u_k + *noise* which is the decision variable yielding the tentative detected symbol $\hat{u}_{k|k}$. This result is now used in (1) to build the signal y_k + *noise* which can be processed to give the new prediction $\hat{y}_{k+1|k}$ of the MD. The recursion then goes on.

Assuming that the receiver is working in a region of reasonable error rate the process of eliminating the data modulation by using the detected symbols in (1) is quite satisfactory. This implies a decision-aided synchronization scheme but there are other alternatives like the introduction of a suitable nonlinearity for eliminating the data modulation effects. In any case, the tentative symbol elimination needed here is for the sake of MD estimation and does not necessarily coincide with the actual detection process of the receiver which may have a long decoding/detection delay when coding is used. So, we can simplify the arguments for the estimation of the MD by equivalently assuming continuous wave (CW) transmission, i.e.,

$$z_k = y_k + v_k \quad (\text{for } U_k = I_2, \text{ the } 2\times 2 \text{ identity matrix}). \tag{2}$$

The receiver of this CW transmission must now concentrate on estimating y_k along with the parameter p_k. However, one must be aware that parameter estimation is only possible if the model structure is selected properly, resulting in an "identifiable" parameterization. This is described roughly in [20] (with further references for more specific results named there) as a set of conditions on the system modes affected by a parameter. The conditions, rephrased to suit our model in Fig. 1, are that at least one of these modes be 1) observable, 2) excited by the initial conditions or controllable from the inputs, and 3) such that $a_k(\cdot, s)$, $b_k(\cdot, s)$, *and* $c_k(\cdot, s)$ do not assume identical values for different parameter values. The third condition can apparently be relaxed for periodic functions of a parameter, the identification of which is acceptable over any period. This will be the case in Section IV-C.

Besides the identifiability of parameters, it can be important from a computational viewpoint to perform a sensitivity analysis [20], [21] for distinguishing the parameters that are crucial to estimate, rather than estimating all uncertain ones.

III. Joint Estimation: The Extended Kalman Filter

Looking at Fig. 1 and omitting U_k, the structure exactly fits the signal model for which an extended Kalman filter is a "near optimum" estimator [6]. The governing equations are

$$s_{k+1} = a_k(p_k, s_k) + b_k(p_k, s_k)n_k \tag{3}$$

$$y_k = c_k(p_k, s_k) \tag{4}$$

$$z_k = y_k + v_k \tag{5}$$

$$p_{k+1} = d_k(p_k) + \xi_k \tag{6}$$

where a_k, c_k, and d_k are in general vector-valued, differentiable, nonlinear functions and b_k is a matrix-valued linear function. These equations are quite general and would be tailored to suit specific cases of interest. For example, $d_k(p_k) = p_k$ in the case of an unknown, constant parameter vector. Or, b_k is usually independent of p and s.

The dependence of the system model on the parameter vector p_k is handled by augmenting the state vector with the vector p_k, namely, by defining a new state vector

$$x_k := \begin{bmatrix} s_k \\ p_k \end{bmatrix}. \tag{7}$$

Then we have the general model

$$\begin{cases} x_{k+1} = f_k(x_k) + g_k(x_k)w_k & (8) \\ z_k \;\; = h_k(x_k) + v_k & (9) \end{cases}$$

[2] For unknown channels, a fractionally spaced untrained [18] or a trained cyclic [19] equalizer could be thought to precede the symbol-rate sampler.

with

$$f_k(x_k) = \begin{bmatrix} a_k(p_k, s_k) \\ d_k(p_k) \end{bmatrix} \quad (10)$$

$$g_k(x_k) = \begin{bmatrix} b_k(p_k, s_k) & 0 \\ 0 & I \end{bmatrix} \quad (11)$$

$$h_k(x_k) = c_k(p_k, s_k) = y_k \quad (12)$$

$$w_k = \begin{bmatrix} n_k \\ \xi_k \end{bmatrix}. \quad (13)$$

We further assume that $\{v_k\}$ and $\{w_k\}$ are zero mean, white Gaussian processes and x_0 is a Gaussian random variable, all being mutually independent with

$$E[v_k v_k^T] = R_k \quad (14)$$

$$E[w_k w_k^T] = Q_k \quad (15)$$

$$E[x_0] = \bar{x}_0;\ E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] = P_0. \quad (16)$$

Now, we introduce the matrices

$$F_k = \frac{\partial f_k(x)}{\partial x}\bigg|_{x=\hat{x}_{k|k}} = \begin{bmatrix} \partial a_k/\partial s & \partial a_k/\partial p \\ 0 & dd_k/dp \end{bmatrix}_{x=\hat{x}_{k|k}} \quad (17)$$

$$H_k^T = \frac{\partial h_k(x)}{\partial x}\bigg|_{x=\hat{x}_{k|k-1}} = [\partial c_k/\partial s\ \ \partial c_k/\partial p]_{x=\hat{x}_{k|k-1}} \quad (18)$$

$$G_k = g_k(\hat{x}_{k|k}) = \begin{bmatrix} b_k(\hat{p}_{k|k}, \hat{s}_{k|k}) & 0 \\ 0 & I \end{bmatrix} \quad (19)$$

and follow the work in [6] to derive an approximate linear signal model and the corresponding Kalman filter. The functions $f_k(\cdot)$ and $h_k(\cdot)$, if sufficiently smooth, can be expanded in Taylor series about the conditional means $\hat{x}_{k|k}$ and $\hat{x}_{k|k-1}$, respectively. Neglecting the second and higher order terms enables us to approximate (8)–(9) by the linear model

$$\begin{cases} x_{k+1} = F_k x_k + G_k w_k + q_{k,x} & (20) \\ z_k = H_k^T x_k + v_k + q_{k,z} & (21) \end{cases}$$

with known, external insertions

$$q_{k,x} = f_k(\hat{x}_{k|k}) - F_k \hat{x}_{k|k} \quad (22)$$

$$q_{k,z} = h_k(\hat{x}_{k|k-1}) - H_k^T \hat{x}_{k|k-1}. \quad (23)$$

Note that the linearization of f_k and h_k is performed at each time step. The Kalman filter for this approximate linear signal model is defined to be the extended Kalman filter for the model (8)–(9) and is shown in Fig. 3.

The Kalman gain L_k is computed online using (10)–(19) and the relations

$$\left.\begin{aligned} \hat{x}_{0|-1} &= E[x_0] = \bar{x}_0 & (24) \\ \Sigma_{0|-1} &= E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] = P_0 & (25) \end{aligned}\right\} \text{Initialization}$$

$$\Omega_k = H_k^T \Sigma_{k|k-1} H_k + R_k \quad (26)$$

$$L_k = \Sigma_{k|k-1} H_k \Omega_k^{-1} \quad (27)$$

$$\left.\begin{aligned} \hat{x}_{k|k} &= \hat{x}_{k|k-1} + L_k(z_k - h_k(\hat{x}_{k|k-1})) \\ \hat{x}_{k+1|k} &= f_k(\hat{x}_{k|k}) \end{aligned}\right\} \text{Fig. 3} \quad (28)$$

$$\Sigma_{k|k} = (I - L_k H_k^T)\Sigma_{k|k-1}(I - L_k H_k^T)^T + L_k R_k L_k^T \quad (29)$$

$$\Sigma_{k+1|k} = F_k \Sigma_{k|k} F_k^T + G_k Q_k G_k^T. \quad (30)$$

While being an exact Kalman filter for the signal mode (20)–(21), this extended Kalman filter is no longer linear or optimal when applied to the general model (8)–(9). The smaller the errors $\|x_k - \hat{x}_{k|k}\|$ and $\|x_k - \hat{x}_{k|k-1}\|$, the less is lost by neglecting the higher order terms in Taylor series expansion of $f_k(\cdot)$ and $h_k(\cdot)$. Some indications of the extended Kalman filter quality are how small the trace($\Sigma_{k|k}$) and trace($\Sigma_{k|k-1}$) and how white the pseudoinnovations are.

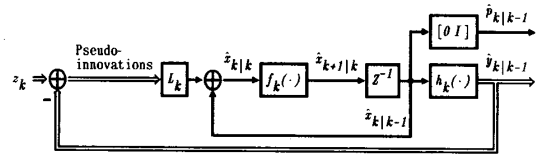


Fig. 3. Extended Kalman filter for the model of (8)–(9).

There are other algorithms [6] for gain and covariance recursions which are algebraically equivalent to (26)–(30) but have some improved properties such as better computational aspects or more efficiency. Most of these alternatives like square root or information square root Kalman filtering can be used with little extra effort in the extended version too.

We see that in (26)–(30) the computation of L and Σ are coupled to $\hat{x}$ through (17)–(19) and hence cannot be carried out offline. But for some applications it might be possible to find approximations which result in decoupling of these computations.

It is possible to improve the estimation accuracy by iteration techniques which improve the reference trajectory or by including more terms in the Taylor series expansion of f_k and h_k [22]. On the other hand, one can use other expansion methods like the "statistical approximation" technique [23] which does not require the differentiability of system nonlinearities and has potential performance advantages.

Unfortunately, no general performance or stability results are known for the extended Kalman filter. The algorithm may give biased estimates and may sometimes diverge. But there is room for improvements through modifications which are, for example, explained in [23]–[25]. In specific cases, when the nonlinearities f_k and h_k are cone bounded, similar estimators can be obtained by taking the structure of Fig. 3 but using some other procedure for the gain calculation which is decoupled from the state estimate and yields a filter with obtainable performance bound. Moreover, attempting to minimize the bound on the error covariance (instead of the error covariance itself) a "bound optimal" filter with precomputable bound and gain can be found [6].

The MD estimation unit of Fig. 3, besides being nearly optimal, has the general advantage that no VCO or NCO is involved in this unit and that its structure well suits into signal processor and VLSI realizations. Additionally, some specific advantages are pointed out as we go through the examples in the next section.

IV. Applications

Any MD estimation problem which can be formulated by some choice of the functions a_k, b_k, c_k, and d_k of Fig. 1 has the (near) optimum solution offered in Section III. We present a few examples by first finding a phase synchronizer for AWGN channels which is possibly only of academic significance. Then, we proceed to other channel models where actual practical implications are at stake.

A. Carrier Phase Estimation on AWGN Channels

The sampled baseband received signal in a digital communication receiver on a pure AWGN channel is

$$z_k = U_k \begin{bmatrix} \cos\theta_k \\ \sin\theta_k \end{bmatrix} + v_k. \quad (31)$$

A coherent receiver must somehow estimate the carrier phase θ_k and eliminate its effect before attempting to detect the data symbol u_k. Here we approach phase synchronization (in the

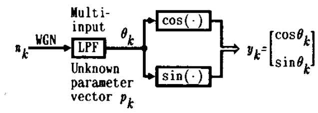


Fig. 4. Model for the carrier phase process on AWGN channels.

context of the structure of Fig. 1) by the assumption that the phase process is generated using the system of Fig. 4.

Using the notation of the previous section we have

$$s_{k+1}=a_k(p_k, s_k)+b_k(p_k, s_k)n_k=A(p_k)s_k+B(p_k)n_k \quad (32)$$

$$y_k=c_k(p_k, s_k)=\begin{bmatrix}\cos\,[C^T(p_k)s_k]\\ \sin\,[C^T(p_k)s_k]\end{bmatrix} \quad (33)$$

where the carrier phase is recognized as

$$C^T(p_k)s_k =: \theta_k. \quad (34)$$

The matrices $A(\cdot)$, $B(\cdot)$, and $C^T(\cdot)$ are set by the low-pass filter (LPF) which determines the PSD of the carrier phase process θ_k. Similar models to that of Fig. 4 have been used [13] in the context of near-optimum demodulation of analog FM with θ_k representing the message process. In our case here, the carrier phase process θ_k is mostly the result of the flicker noise of the oscillators used in the mixers and its PSD has components of the form $|\omega|^\alpha$ for α some nonpositive number [26]. For example,

$$S_\theta(\omega)=\sum_{\alpha=-4}^{0} h_\alpha|\omega|^\alpha; \quad 0<\omega_l\le|\omega|\le\omega_h<\infty \quad (35)$$

where the coefficients h_α depend on the oscillator type and ω_l and ω_h represent limitations on the measurements of PSD. For practical reasons, in this example we let $h_{-3} = h_{-1} = h_0 = 0$. The two remaining terms correspond to random phase and frequency walk which can represent most oscillator inaccuracies and instabilities of interest.[3] Now assuming white inputs with unity average power the LPF with two inputs and one output has the transfer function matrix

$$F(s)=[\sqrt{h_{-4}}/s^2 \;\sqrt{h_{-2}}/s] \quad (36)$$

with the unknown, constant parameter vector

$$p=[\sqrt{h_{-4}}\;\sqrt{h_{-2}}]^T \quad (37)$$

A discrete-time filter $F_d(z)$ can be found from $F(s)$ by using the impulse-invariant transformation [27]. Then the relation $F_d(z) = C^T(zI - A)^{-1}B$ poses a realization problem for determining the matrices A, B, and C. Finally, using $\xi_k = 0$ and $d_k(p) = p$ in (6) the extended Kalman filter of Fig. 3 performs the carrier phase synchronization by delivering the estimate $\hat{y}_{k|k-1} = [\cos\hat{\theta}_{k|k-1}\;\sin\hat{\theta}_{k|k-1}]^T$. Additionally, $\hat{p}_{k|k-1}$ gives information about the unknown h_α coefficients in the model.

It is interesting to study the similarities of the above extended Kalman filter and the phase locked loop (PLL).[4] In the extremely simplified case with known parameter $p = [0\;1]^T$, i.e., for $F_d(z) = [0\;1/(z-1)]$, and with no AWGN the extended Kalman filter, after using some algebra, turns into a first-order PLL (in baseband) transformed to discrete time (Fig. 5). The box $1/(z - 1)$ in the estimator plays the role of a voltage controlled oscillator (VCO) with transfer function $1/s$. Furthermore, using $p = [1\;1]^T$, i.e., a LPF with transfer

[3] In coherent optical communications, the noise process is mainly determined through nonzero h_0, h_{-2}, and h_{-4} terms which represent the shot noise, the laser phase noise, and the frequency mismatch, respectively.

[4] Such a comparison has been made by many authors for the case of near-optimum angle demodulation, too.

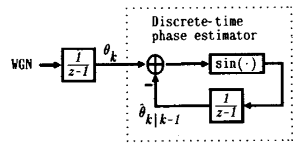


Fig. 5. Noiseless observations of a random phase walk process and the equivalent model of an extended Kalman filter (first-order "discrete-time PLL" in baseband) as the estimator.

function matrix $F_d(z) = [1/(z-1)^2\;1/(z-1)]$, results in an extended Kalman filter which after reaching steady state has the same form as a second-order PLL, the most popular one in application. In both of these cases, including the AWGN $v_k = (v_k^I v_k^Q)^T$ in the signal model results in a term $v'(k, \hat{\theta}) := -v_k^I \sin\hat{\theta} + v_k^Q \cos\hat{\theta}$ to be added to the output of the sin $(\cdot)$ of the estimator. A PLL with noisy input experiences an identical modification of its noiseless-input baseband equivalent model.

In a realistic problem, we might have significant amounts of additive noise in addition to oscillator instability and inaccuracy. Then, optimization of the tracking performance for a PLL requires quite laborious arguments for finding loop parameters which minimize the variance of the phase error $\theta - \hat{\theta}$. These arguments are based on assuming a set of values for the h_α coefficients of (35) and a fixed known signal-to-noise ratio (SNR) and have to be repeated if the SNR or the h_α coefficients change. On the other hand, approaching the same problem with an extended Kalman filter, the Kalman gain L_k which effectively contains all the loop parameters is easily set in the "best possible" manner automatically and adaptively. (For SNR adaptation, see Section IV-D below.) In any case, the synchronizers of the form in Fig. 5 are optimum during the tracking mode, as the sin$(\cdot)$ nonlinearity can be ignored for small phase errors.

Rapid acquisition of phase is also essential in many applications. In general, any synchronizer resulting from the model of Fig. 4 (i.e., some form of a PLL) is not reliable for fast acquisition of phase or frequency by itself. For example, in Fig. 5, a situation with an initial estimation error $\theta - \hat{\theta} = \pi$ might be left uncorrected for a long time ("Hangup", [28].) A possible solution to this problem is known as "acquisition-aided PLL" [16, paper by H. Meyr and L. Popken]. Utilizing the ideas in this paper, a more systematic solution is being currently investigated. The idea is to stay away from phase and frequency as explicit state variables in an attempt to avoid the acquisition problems of the PLL.

B. Independent-I/Q Models: Adaptive MD Estimation for Mobile Communications

Starting with a PLL to achieve phase-coherent detection of digital data or voice transmitted over many fading channels results in receivers with unacceptable behavior. This is a direct reflection of the poor response of the PLL during and immediately after deep fades in channel amplitude, generally leading to the conclusions that coherent detection is not suitable for such fading channels and noncoherent detection must be resorted to, [29]. Here we apply our method of near optimum MD estimation which does not have the weaknesses of the PLL and results in a coherent receiver that outperforms the noncoherent one.

We modify the last example by assuming that the two components of y_k are generated independently (Fig. 6). This allows amplitude as well as phase variations of the MD y_k and results in the Rayleigh flat fading observed in many mobile communication systems.[5] Working on the quadrature components of y_k (and not its phase directly) has interesting properties which will be explained at the end of this example.

[5] The signal model for Rice fading includes both Figs. 4 and 6 with their output vectors summed together after being properly scaled.

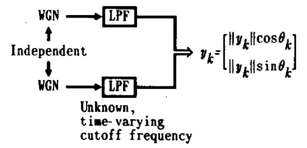


Fig. 6. Model with independent I/Q components for generation of MD.

The received signal is

$$z_k = U_k y_k + v_k. \quad (38)$$

In comparison to (31) the estimation of the MD y_k can be interpreted as the "carrier estimation" in the receiver. Again, we ignore U_k and concentrate on the estimation of y_k by using specific information about the dynamics of the LPF's in Fig. 6.

In mobile communications, under the assumption of uniform angular distribution of the received electrical waves, a transmitted CW signal at frequency f_c has the following baseband I- or Q-component PSD at the receiver [30]

$$S_{yI/Q}(f) = (1-(f/f_D)^2)^{-1/2}; \quad |f| < f_D \quad (39)$$

with

$$f_D = f_c v/c \quad \text{(half of the Doppler spread)}$$
$$v = \text{vehicle speed}$$
$$c = \text{speed of light}. \quad (40)$$

For a carrier frequency of 1 GHz and maximum vehicle speed of 260 km/h (e.g., for trains in Europe), the Doppler spread [17] runs up to about 500 Hz which is a considerable amount.

Assuming a white input, the magnitude of the frequency response of a LPF with output PSD of (39) is

$$|F(f)| = (1-(f/f_D)^2)^{-0.25}; \quad |f| < f_D \quad (41)$$

which is plotted in Fig. 7. Having a fixed carrier frequency f_c in (40), the filter cutoff frequency f_D depends on the vehicle speed and is an unknown, time-varying parameter.

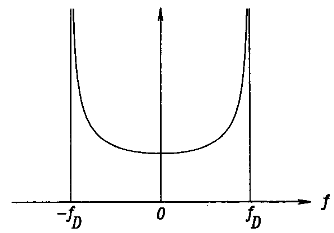


Fig. 7. Magnitude of the frequency response of the LPF's of Fig. 6 for land-mobile communications. f_D depends on vehicle speed and is unknown.

In our example, we assume the following stochastic model for the parameter $\omega = 2\pi T f_D$ which is the normalized cutoff frequency in radians:

$$\delta_{k+1} = \delta_k + \xi_k^\delta \quad (42)$$

$$\omega_{k+1} = \omega_k + \delta_k + \xi_k^\omega \quad (43)$$

where $\{\xi_k^\delta\}$ and $\{\xi_k^\omega\}$ are zero mean WGN processes. Relating this to the notation of Section II we get the random parameter p_k and its dynamics

$$p_k = \begin{bmatrix} \delta_k \\ \omega_k \end{bmatrix} \quad (44)$$

$$p_{k+1} = d_k(p_k) + \xi_k = Dp_k + \begin{bmatrix} \xi_k^\delta \\ \xi_k^\omega \end{bmatrix} \quad (45)$$

$$D = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}. \quad (46)$$

It is observed in practice that such random-parameter models yield good filter designs even if the parameters actually have deterministic behavior.

An LPF with very similar characteristics to (41) has been realized [31] by an eighth-order elliptic filter which we used to generate the multiplicative Rayleigh distortion. But in the extended Kalman filter we simplified matters further by assuming a second-order filter with unknown, time-varying cutoff frequency ω. The performance loss due to this simplification was found to be minimal. So, the MD y_k is assumed by the Kalman filter to be generated as

$$s_{k+1} = a_k(p_k, s_k) + b_k(p_k, s_k)n_k = A(p_k)s_k + Bn_k \quad (47)$$

$$y_k = c_k(p_k, s_k) = C^T(p_k)s_k \quad (48)$$

with

$$A = \begin{bmatrix} A' & 0 \\ 0 & A' \end{bmatrix}; \; B = \begin{bmatrix} B' & 0 \\ 0 & B' \end{bmatrix}; \; C^T = \begin{bmatrix} C'^T & 0 \\ 0 & C'^T \end{bmatrix}. \quad (49)$$

The primed matrices represent a second-order, discrete-time filter with cutoff frequency ω and damping ratio ζ:

$$A' = \begin{bmatrix} 0 & 1 \\ -\exp(-2\zeta\omega) & 2\exp(-\zeta\omega)\cos(\omega\sqrt{1-\zeta^2}) \end{bmatrix} \quad (50)$$

$$B' = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (51)$$

$$C' = \begin{bmatrix} 0 \\ \sqrt{\dfrac{1-\exp(-2\zeta\omega)}{1+\exp(-2\zeta\omega)}\{1-2\exp(-2\zeta\omega)\cos[2\omega(1-\zeta^2)^{1/2}]+\exp(-4\zeta\omega)\}} \end{bmatrix}. \quad (52)$$

The complex expression in (52) is the result of the requirement that, with white input, the output average power of the filter $\{A', B', C'\}$ be independent of the time-varying cutoff frequency ω. This is implied by the assumption that the fading average power $E[\|y_k\|^2]$ is independent of the vehicle speed. The parameter ζ takes on a fixed value, for example, 0.2. Alternatively, it could also have been assumed unknown and taken into account by the parameter vector p_k.

Now, with (44)–(52) and (7)–(19) the extended Kalman filter of Fig. 3 is completely determined. We used the setup in Fig. 8 to simulate the above ideas. The time-varying cutoff frequency ω_k and its estimate are displayed in Fig. 9.

We also ran this example for 4-phase, differentially-coherent phase-shift-keying (4-DC-PSK) modulation with decision-aided synchronization. Fig. 10 shows the error-rate performance of this receiver (using the "tentative decisions"

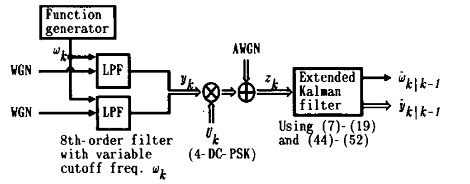


Fig. 8. MD estimation by joint estimation of a noisy, complex-valued signal and its randomly varying bandwidth.

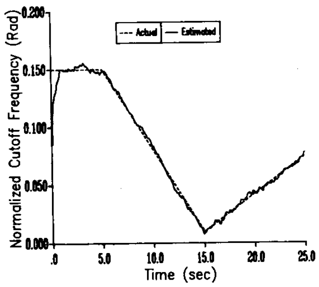


Fig. 9. The time-varying cutoff frequency and its estimate after removing the bias due to modeling differences.

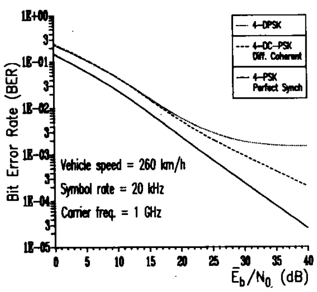


Fig. 10. Probability of error versus SNR for a mobile communication system.

of Fig. 2) and that of a noncoherent one with differential detection (4-DPSK) where the superiority of the coherent detection can be clearly seen. The theoretical result for coherent 4-PSK with perfect synchronization [17] is also plotted in that figure. In the presence of diversity (or coding, in general,) similar ordering of performances are observed.

The MD estimation scheme outlined above has several more interesting properties. It is free of hangup, as the phase variable θ is not approached directly. Also, the magnitude $\|y_k\|$ can be used as the channel quality information in diversity reception for maximal ratio combining or in a soft-decision decoder when coded transmission is present.

Before leaving this example, we point out that (in opposed to Fig. 4) the model of Fig. 6 has no explicit nonlinearity and if it were not for the sake of random parameters in the LPF's, we would have a purely linear model. Such a linear model for the MD is assumed in [32] and used to discuss optimal detection and synchronization.

C. Adaptive MD Estimation with Frequency Offset

A problem faced in receivers is a frequency offset after the received signal is mixed down to baseband. Depending on the modulation scheme in use, a reasonable amount of left over frequency offset can be handled by our baseband synchronizer.

In example A this corresponds to the term h_{-4}/ω^4 in (35) which can represent a random frequency walk along with frequency offset. The latter is taken care of by merely using a larger value for the first element of P_0 in (25) which corresponds to the frequency variable in this example. Such an easy modification for frequency offset is always possible when one of the system states is actually the frequency variable.

In general, the matter can be solved by assuming a frequency offset $\mu(t)$ in Hz, resulting in a phase change $\varphi_k = 2\pi\int_0^{kT}\mu(\tau)d\tau$. Then the signal $U_k y_k$ must make a phase rotation of φ_k by premultiplying it with the time-varying matrix

$$\Phi_k = \begin{bmatrix} \cos\varphi_k & -\sin\varphi_k \\ \sin\varphi_k & \cos\varphi_k \end{bmatrix}. \tag{53}$$

So, for Section IV-B, the third part of (49) takes the form

$$C^T = \Phi_k \begin{bmatrix} C'^T & 0 \\ 0 & C'^T \end{bmatrix}. \tag{54}$$

Note that C^T in the presence of frequency offset is no longer block diagonal and hence a cross-coupling between the two components of y_k exists which was not otherwise present in Fig. 6.

It remains now to extend the random parameter p_k and the equation describing its dynamics. We can, for example, assume that $\{\varphi_k\}$ is the result of random phase and frequency walk processes

$$\mu_{k+1} = \mu_k + \xi_k^\mu$$
$$\varphi_{k+1} = \varphi_k + \mu_k + \xi_k^\varphi \tag{55}$$

where all inputs are zero mean WGN, as usual. Actually, in addition to frequency offset this model takes care of instability of the mixer oscillators. Now, (44)–(46) have the form

$$p_k = [\delta_k \ \omega_k \ \mu_k \ \varphi_k]^T \tag{56}$$

$$p_{k+1} = d_k(p_k) + \xi_k = Dp_k + [\xi_k^\delta \ \xi_k^\omega \ \xi_k^\mu \ \xi_k^\varphi]^T \tag{57}$$

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}. \tag{58}$$

Taking all the above system and parameter changes into account Section IV-B can now be tried again, resulting in an extended Kalman Filter which adaptively estimates an MD with randomly varying bandwidth and in the presence of frequency offset and oscillator instability, the estimates of which are also delivered.

We modified Fig. 8 by applying a frequency offset to z_k. A safe upper limit of the normalized frequency offset with CW transmission is about $\pi/2$ rad and with 4-PSK modulation $\pi/8$ rad, which for symbol rate of $1/T$ = 20 kHz corresponds to $(\pi/8)/(2\pi T)$ = 1.25 kHz. Fig. 11 shows the estimation of the frequency offset by the extended Kalman filter explained above. The simulated bit error rate showed negligible performance loss compared to Fig. 10, as tracking a frequency offset using the above method is very satisfactory. However, the initial acquisition behavior requires a closer look (cf. end of Section IV-A) which is out of the scope of this paper.

D. Adaptive MD and Received Signal Power Estimation

The received signal power (RSP) is a parameter which may be needed in many receiver configurations. If the assumption holds that the PSD of the AWGN is fixed and approximately known, the estimations of the RSP and the signal-to-noise ratio (SNR) are equivalent. We define RSP as an average power, where time-averaging is performed only long enough to leave out the "short-term" effects but to include the "long-term" ones. For a standard AWGN channel, there are no long-term effects and the RSP might be an unknown, constant parameter.

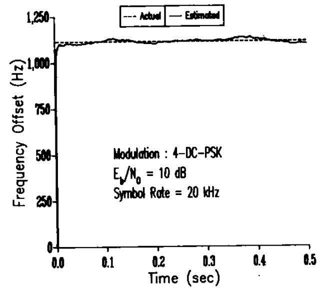


Fig. 11. Estimation of frequency offset by the extended Kalman filter.

On the other hand, the RSP is a slow, randomly varying parameter in mobile communications with Rayleigh fading (short-term) and log-normal shadowing (long-term) [30].

Here, we only outline the simple procedure when adaptive RSP estimation is necessary in the receiver. The schemes in the examples A, B, and C can be extended by appending the additional parameter η_k to (37), (44), and (56) and multiplying the right-hand sides of (33) and (52) by η_k. The square of $\hat{\eta}_{k|k-1}$ is a measure for RSP. Therefore, we have adaptive MD estimation schemes which provide estimates of the RSP in addition to other time-varying parameters of interest. The estimate $\hat{\eta}_{k|k-1}$ can be used for AGC amplifier setting in analog and/or digital parts of the receiver.

V. Conclusions and Further Remarks

Near-optimal MD estimation for digital communications involves, in general, nonlinear models with additional unknown or randomly varying parameters. The systematic approach offered in this paper can be regarded as a new tool for "carrier synchronization" where an extended Kalman filter is applied for estimating a noisy signal with randomly varying parameters governed by nonlinear dynamical equations. Under simplifying conditions on AWGN channels, the solutions found here reconcile with the classical PLL, showing the "asymptotic optimality" of the PLL in the sense of MMSE. Novel algorithms with significant advantages have been derived here for other channels with more involved models containing fading, shadowing, and frequency offset.

In general, the extended Kalman filter should be tried out for specific cases, resulting in possible refinements. For example, in the event of nonsatisfactory performance the same filter structure with modified algorithms [6], [24] or more sophisticated filters [6], [14], [25] can be used. If too complicated to be practical, the extended Kalman filter could be taken as a starting point with desirable properties. It could then be watered down to suit the practical limitations.

References

[1] H. Cox, "On the estimation of state variables and parameters for noisy dynamic systems," *IEEE Trans. Automat. Contr.*, vol. AC-9, pp. 5–12, Jan. 1964.

[2] C. G. Hilborn, Jr., and D. G. Lainiotis, "Optimal estimation in the presence of unknown parameters," *IEEE Trans. Syst. Sci. Cybern.*, vol. SSC-5, pp. 38–43, Jan. 1969.

[3] J. J. Kappl, "Nonlinear estimation via Kalman filtering," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-7, pp. 79–84, Jan. 1971.

[4] D. G. Lainiotis, "Optimal adaptive estimation: Structure and parameter adaptation," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 160–170, Apr. 1971.

[5] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification.* Cambridge, MA: The M.I.T. Press, 1983.

[6] B. D. O. Anderson and J. B. Moore, *Optimal Filtering.* Englewood Cliffs, NJ: Prentice-Hall, 1979.

[7] D. L. Snyder, "The state variable approach to analog communication theory," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 94–110, Jan. 1968.

[8] A. H. Jazwinski, "Filtering for nonlinear dynamical systems," *IEEE Trans. Automat. Contr.*, vol. AC-11, pp. 765–766, Oct. 1966.

[9] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part II.* New York: Wiley, 1971.

[10] G. Pasternack and R. L. Whalin, "Analysis and synthesis of a digital phase locked loop for FM demodulation," *Bell Syst. Tech. J.*, pp. 2207–2237, Dec. 1968.

[11] C. N. Kelly and S. C. Gupta, "The digital phase locked loop as a near optimum FM demodulator," *IEEE Trans. Commun. Technol.*, vol. COM-20, pp. 406–411, June 1972.

[12] D. R. Polk and S. C. Gupta, "Quasi-optimum digital phase locked loop," *IEEE Trans. Commun. Technol.*, vol. COM-21, pp. 75–82, Jan. 1973.

[13] A. L. McBride, "On optimum sampled-data FM demodulation," *IEEE Trans. Commun. Technol.*, vol. COM-21, pp. 40–50, Jan. 1973.

[14] P. K. S. Tam and J. B. Moore, "A Gaussian sum approach to phase and frequency estimation," *IEEE Trans. Commun.*, vol. COM-25, pp. 935–942, Sept. 1977.

[15] M. T. Dharamsi and S. C. Gupta, "Discrete-time demodulation of angle modulated analog signals in fading channels," *Inform. Sci.*, vol. 9, pp. 99–132, Jun. 1975.

[16] Special Issue on Synchronization, *IEEE Trans. Commun.*, vol. COM-28, Aug. 1980.

[17] J. G. Proakis, *Digital Communications.* New York: McGraw-Hill, 1983.

[18] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. COM-28, pp. 1867–1875, Nov. 1980.

[19] S. U. H. Qureshi, "Adaptive equalization," *Proc. IEEE*, vol. 73, pp. 1349–1387, Sept. 1985.

[20] P. S. Maybeck, *Stochastic Models, Estimation, and Control, Vol. 2.* New York: Academic, 1982.

[21] P. M. Frank, *Introduction to System Sensitivity Theory.* New York: Academic, 1978.

[22] A. P. Sage and J. L. Melsa, *Estimation Theory with Applications to Communications and Control.* New York: McGraw-Hill, 1971.

[23] A. Gelb, Ed., *Applied Optimal Estimation.* Cambridge, MA: The M.I.T. Press, 1974.

[24] G. C. Goodwin and K. S. Sin, *Adaptive Filtering Prediction and Control.* Englewood Cliffs, NJ: Prentice-Hall, 1984.

[25] H. W. Sorenson, "On the developement of practical nonlinear filters," *Inform. Sci.*, vol. 7, pp. 253–270, 1974.

[26] V. F. Kroupa, "Noise properties of PLL systems," *IEEE Trans. Commun.*, vol. COM-30, pp. 2224–2252, Oct. 1982.

[27] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1975.

[28] F. M. Gardner, "Hangup in phase locked loops," *IEEE Trans. Commun.*, vol. COM-25, pp. 1210–1214, Oct. 1977.

[29] S. Stein, "Fading channel issues in system engineering," *IEEE J. Select. Areas Commun.*, vol. SAC-5, pp. 68–89, Feb. 1987.

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

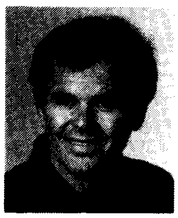
[31] H. Brehm, W. Stammler, and M. Werner, "Design of a highly flexible digital simulator for narrowband fading channels," *EUSIPCO-86 Proc. Signal Processing III: Theories and Applications*, I. T. Young, Ed. Elsevier Science Publishes B. V. (North Holland), pp. 1113–1116, 1986.

[32] R. Haeb and H. Meyr, "A systematic approach to carrier recovery and detection of digitally phase modulated signals on fading channels," *IEEE Trans. Commun.*, accepted for publication.

Abbas Aghamohammadi (S'88) was born in Tehran, Iran, in 1957. He received the B.S. and M.S. degrees in electrical engineering and computer science from Massachusetts Institute of Technology, Cambridge, in 1979 and 1981, respectively. He is presently working towards the Ph.D. degree in electrical engineering at Aachen University of Technology (RWTH), Aachen, West Germany.

From 1979 to 1984, he was partly a Research or a Teaching Assistant at M.I.T. and at the University of California, Santa Barbara, and since 1985, at Aachen University of Technology. His research interests include synchronization in digital communications with emphasis on digital receiver techniques.

★

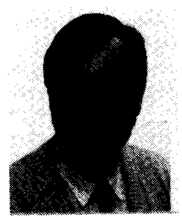


Heinrich Meyr (M'75–SM'83–F'86) received the Dipl.-Ing. and the Ph.D. degrees from the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, in 1967 and 1973, respectively.

From 1968 to 1970 he held research positions at Brown Boveri Corporation, Zurich, Switzerland and the Swiss Federal Institute for Reactor Research. From 1970 to the Summer of 1977 he was with Hasler Research Laboratory, Bern, Switzerland, doing research in fields of digital facsimile encoding and tracking systems. His last position at Hasler was Manager of the Research Department. During 1974 he was a Visiting Assistant Professor with the Department of Electrical Engineering, University of Southern California, Los Angeles. Since the Summer of 1977 he has been a Professor of Electrical Engineering at the Aachen University of Technology (RWTH Aachen), West Germany. His research includes synchronization, estimation, and digital signal processing. He is a frequent consultant to the industry in his field of specialization.

Dr. Meyr has published work in various fields and journals and holds over a dozen patents. He served as a Vice Chairman for the 1978 IEEE Zurich Seminar and as an International Chairman for the 1980 National Telecommunications Conference, Houston, TX. In 1984 and 1985, he served as the Vice President, International Affairs of the IEEE Communications Society.

★



Gerd Ascheid (M'84) was born in Cologne, Germany, in 1951. He received the Dipl.-Ing. degree in electrical engineering at Aachen University of Technology (RWTH), Aachen, West Germany, in 1978 and the Doctorate degree in 1984.

From 1979 to 1984 he was a Research Assistant and from 1985 to 1988 a Senior Engineer at the Department of Electrical Engineering at Aachen University of Technology. He now is Managing Director of CADIS GmbH, Aachen. His main interests are in synchronization, in particular for bandwidth efficient transmission, and in digital signal processing in communications.

Dr. Ascheid is coauthor of the forthcoming book, *Synchronization in Digital Communications*.