Maximum Likelihood Synchronisation, Equalisation and Sequence Estimation for Unknown, Time-Varying, Frequency-Selective Rician Channels

Brian D. Hart, Member, IEEE
Telecommunications Engineering, Research School of Information Sciences and Engineering,
Australian National University, Canberra, Australia
Brian.Hart@anu.edu.au

Desmond P. Taylor, Fellow, IEEE
Department of Electrical and Electronic Engineering,
University of Canterbury, New Zealand
taylor@elec.canterbury.ac.nz

Abstract

This paper develops a receiver structure to perform jointly ML synchronisation, equalisation and detection of a linearly modulated signal transmitted over a time-varying, frequency-selective, Rician faded channel, corrupted by AWGN. The receiver is particularly suited to a fast fading channel, where other receivers that rely on estimating the channel cannot track it quickly enough. The signal mean and autocovariance are needed, and a scheme is proposed for estimating these quantities adaptively. The receiver processes the specular and diffuse components (corresponding to the signal mean and autocovariance) separately. Processing the known specular component is the classical detection problem. The unknown diffuse component is processed by predictors [11]. We show that the predictors can achieve synchronisation in a novel manner, if synchronisation is required. A union bound on the receiver's BER is derived, and it tightly bounds simulated BERs in fast fading at high SNRs.

I. Introduction

When communicating with fast moving mobile terminals in a multipath channel, the receiver observes a delay and Doppler-spread signal. In the time domain, this Doppler spread is experienced as a time-varying channel. If the Doppler spread is significant compared to the symbol rate, then the channel becomes difficult to track, and most existing receiver structures exhibit an error floor, where an increase in SNR does not improve the BER [1,2]. Several approaches have been considered in the literature to surmount the problem, particularly for frequency-flat channels [3,4,5].

It is instructive to consider receiver structures that are actually optimal for the time-varying, frequency-selective, Rician fading channel model. Different ML sequence estimators have already been derived, for three different assumptions: (*i*) The channel is wholly unknown (e.g. the time-invariant channel [2]). This is the blind ML detection problem. (*ii*) The channel impulse response is unknown, but its mean and autocovariance are known (in this context, "unknown" signifies that the diffuse component is unknown but the specular component is known. This makes most sense when we realise that only zero-mean channels have been considered heretofore) [6,7,8,9,10,11]. (*iii*) The channel is completely known [17,22]. In (*ii*) and (*iii*), the receiver is often described as "genie-aided," since the receiver is assumed to have knowledge that actually cannot be available.

In approach (*i*), the receiver hypothesises all possible transmitted sequences. For each, it makes an ML estimate of the signal mean and autocovariance, from the *entire* received sample sequence. Finally the receiver detects the hypothesised sequence with the maximum probability that the received sample sequence was observed, conditioned on the hypothesised sequence and its estimated mean and autocovariance. This is called "per-sequence-processing," and leads to an intrinsically non-iterative receiver structure.

A near-optimal, practical approximation to the blind MLSE receiver employs per-survivor-processing (PSP) [14]. The signal mean and autocovariance are estimated causally, only a finite number of possible sequences are hypothesised at any one time, and the conditional probability expression is transformed and simplified into an iterative, finite-complexity metric. Approach (*ii*) is also unrealisable, since in practice the signal mean and autocovariance are unknown to the receiver. They must be estimated from past received samples. In fact, a near-optimum, practical approximation to approach (*ii*) is the same as the approximated blind MLSE receiver. Since past samples only are used for estimating the

signal mean and autocovariance, the receiver's performance is poor initially. Accordingly, the receiver's robustness is enhanced when a training sequence is transmitted first, to obtain a reasonable estimate of the signal mean and autocovariance.

In the literature on these receivers, only [7] proposes a way to estimate the signal's autocovariance, for the case of *M*-PSK, rectangular pulses, and Rayleigh fading. Only [7,12] analyse the receiver structure's BER; simulation is used in the other references.

This paper extends these results, and is organised as follows. The signal model is generalised in section II to a time-varying, frequency-selective Rician fading channel, and the receiver's need for synchronisation is explicitly identified. As special cases, this model includes most channels of practical interest. In section III the MLSE receiver structure is derived for a signal distorted by a Rician fading channel and requiring synchronism. The diffuse (random, Rayleigh) component of the receiver signal is processed by MMSE predictors. The receiver derivation in section III assumes perfect knowledge of the signal's mean and autocovariance, whereas they must be estimated from the signal in practice. A scheme for estimating these quantities adaptively is presented in section IV. It employs a minimisation algorithm to search for the signal mean and the predictor tap weights which predict past samples with MMSE. The receiver's BER is evaluated analytically in section V, using a union bound technique. Finally, analytic and simulation results are presented in section VI that illustrate the novel aspects of this paper.

This joint receiver requires *a priori* only (*i*) a stable symbol-rate oscillator; (*ii*) frame timing in TDMA systems; (*iii*) an upper bound on the duration of the received pulses (i.e. the duration of the transmitted pulse plus an upper bound on the delay spread and timing error); (*iv*) an upper bound on the bandwidth that the Doppler spread and shifted signal occupies; and (*v*) for ML performance, perfect knowledge of the signal's mean and autocovariance. In adaptive operation, the mild constraint is made that the channel statistics and signal properties (i.e. the carrier frequency, carrier phase, symbol timing, noise power, and channel mean and autocovariance) change more slowly than receiver's ability to track the changes. Thus, for Rayleigh fading channels, the receiver relies on quasi-stationary *second* order statistics, instead of quasi-stationary *first* order statistics, as is usually the case.

II. System Model

In this section, a mathematical description of the transmitter, channel, and receiver front-end are developed. Figure 1 is a diagram of the communications system.

A. Transmitter

The transmitter maps an M-ary information sequence, $\{\alpha_i\}$, $\alpha_i \in \{0, .., M\text{-}1\}$, to a phasor sequence, $\{\beta_i\}$, taken from an M-ary constellation. The transmitter computes the complex baseband signal,

$$a(t) = \sum_i \beta_i h(t - iT), \tag{1}$$

then translates it to the carrier frequency, f_c. $h(t)$ is the transmitter pulse shape and T is the symbol period.

The lack of an absolute phase reference in the Rayleigh fading channel influences the design of the signal constellation, and the mapping of bits to symbols. We define P such that the constellation has P-ary rotational symmetry and $A = M/P$. Then the constellation consists of P sectors, with A points per sector. Define phase(β_i) $\in \{0, .., P\text{-}1\}$ to label uniquely the sector that β_i is in. Similarly, define amplitude(β_i) $\in \{0, .., A\text{-}1\}$ to label uniquely where β_i is within a sector. This illustrated for 16-QAM in figure 2.

Consider the transmitted and detected sequences, $\{\beta_i\}$ and $\{\hat{\beta}_i\}$. Define $e_{p,i} \in \{0, .., P\text{-}1\}$ such that phase(β_i) = phase($\hat{\beta}_i$) + $e_{p,i}$. Define phase lock as the property that $e_{p,i} = 0$ over a long interval in i (apart from occasional errors). For absolute phase shift keying, a cycle slip or phase slip occurs when $e_{p,i}$ becomes non-zero. For differential phase shift keying, a cycle slip occurs when $e_{p,i}$ has one value for a long interval, then subsequently changes. Similarly, amplitude lock is defined as the property that amplitude(β_i) = amplitude($\hat{\beta}_i$) for long intervals. Define $e_{a,i} \in \{0, .., A\text{-}1\}$ such that amplitude(β_i) = amplitude($\hat{\beta}_i$) + $e_{a,i}$. Amplitude lock is lost when $e_{a,i} \neq 0$ for an interval in i. An amplitude slip occurs when $e_{a,i} = 0$ for an interval, then subsequently changes. The amplitude slip is corrected when $e_{a,i} = 0$ for an interval again.

Receivers for random channels may use predictors [6,9,11]. They use past signal samples as amplitude and phase references for subsequent samples. A cycle slip affects the

phases of subsequent decisions, and phase lock between transmitter and receiver cannot be guaranteed after a deep fade in the absence of channel sounding when the channel is purely Rayleigh. Thus uncoded transmission is effectively a catastrophic "code," in that there are a number of valid sequences whose path metrics are the same but whose hypothesised symbol sequences differ. If the error involves an amplitude slip, then it is ultimately corrected when a symbol is transmitted that reveals the erroneous amplitude reference, if not before[1]. Slips involving amplitude are nearly catastrophic in large constellations, since the anomalous amplitude reference may not be corrected for some time, resulting in long bursts of errors. Accordingly, we conclude that the transmitter constellation and mapping should be designed to be rotationally-invariant and amplitude-slip-tolerant. A radially symmetric constellation is proposed in the following subsection with these properties.

The mapping from $\{\alpha_i\}$ to $\{\beta_i\}$ has two stages. The log_2P phase bits select a sector in a rotationally invariant manner, such as differential encoding; then the log_2A sector bits select a point from that sector, in such a way that fewer bit errors arise from an amplitude slip. Differential amplitude encoding achieves this for the radially symmetric constellation. For *M*-QAM constellations, an effective solution is unclear.

B. Rotationally Symmetric Constellation

A rotationally-invariant and amplitude-slip-tolerant constellation is most easily constructed from geometrically spaced shells, $\{r_{rs}\angle\varphi_{rs}\}$, for $r_{rs} = 1, \rho, \rho^2, \ldots \rho^{A-1}$, $0 < \rho < 1$, and $\varphi_{rs} = 0, 2\pi\frac{1}{P}, \ldots 2\pi\frac{P-1}{P}$. Figure 3 shows one example. To ensure robustness to amplitude and phase slips, each shell is identically differentially Gray-encoded using log_2P bits. The remaining log_2A bits Gray-encode the transitions between shells, with wrap-around if the outer shell is reached. A phase slip introduces one error, and an amplitude slip introduces one error initially, and a second when the amplitude slip is corrected. The proposed constellation is related to the one presented in [25], except that there the shell radii vary arithmetically, whereas here they vary geometrically.

---

[1] Consider 16-QAM, where $\beta_i \in \{\pm\{1,3\}\pm j\{1,3\}\}$. If symbols from $\{\pm1\pm j\}$ are continually transmitted, but the receiver incorrectly detects symbols from $\{\pm3\pm3j\}$ for several symbol periods, then subsequent predictions also approximate $\pm3\pm3j$, and accordingly the receiver detects symbols from $\{\pm3\pm3j\}$, causing continual errors. However, if the transmitter finally sends a symbol from $\{\pm3\pm3j\}$, the receiver predicts $\pm9\pm9j$, and therefore detects a symbol from $\{\pm3\pm3j\}$ as the closest symbol, thus terminating the error event.

C. Channel and Receiver Front-End Processing

The real bandpass transmitted signal is distorted by a time-varying, frequency-selective Rician fading channel. Real bandpass noise is added to this signal, and the composite signal is quadrature demodulated by multiplication with $\exp\left(-j2\pi\left(f_c - f_0\right)t + j\phi\right)$, where $f_c - f_0$ is the receiver's estimate of f_c, and ϕ is the phase offset. The complex near-baseband signal is filtered by a noise-limiting filter with transfer function, $\mathrm{rect}\left(fT_r\right)$, then sampled at the times lT_r-t_0 sec, where t_0 is the timing error. The sampling period, T_r, is chosen as T/r, where the integer $r \geq 1$ is the number of samples per symbol period. The sampling rate must be sufficiently high that the noise-limiting filter does not distort the signal, even allowing for the Doppler spread and the uncertain (but upper bounded) carrier frequency offset due to non-ideal oscillators and a Doppler shift. The signal is not perfectly bandlimited, so strictly the noise-limiting filter's bandwidth is infinite. However, in practice another bandwidth definition (such as the -40dB bandwidth) can be used such that the signal distortion or BER penalty is negligible.

These transformations can be represented simply in complex baseband, as

$$y_l = \int_{-\infty}^{\infty} a\left(lT_r - t_0 - \xi\right) z\left(lT_r - t_0, \xi\right) d\xi \exp\left(j2\pi l f_0 T_r + j\phi\right) + n_l \tag{2}$$

where y_l is the received signal; $z(t,\xi)$ is the fading channel; f_0 is the residual carrier frequency offset; and n_l is the additive noise. Without loss of generality, (2) ignores the timing error in the stationary noise, and lumps the $j2\pi f_0 t_0$ term in the complex exponential with the carrier phase offset. The channel, $z(t,\xi)$, can be visualised as a densely tapped delay line [16], where the taps are indexed by ξ. The input at any delay, ξ_0, is multiplied by the ξ_0-th tap weight, $z(t,\xi_0)$, a time-varying, Rician fading process. For time-invariant channels, $z(t,\xi) = z(0,\xi)$; for frequency-flat channels, $z(t,\xi) = z(t)\delta(\xi)$; for the AWGN channel, $z(t,\xi) = \delta(\xi)$. Define the superscripts *nr* and *r* to denote the non-random (specular) and random (diffuse) signal components, respectively. Thus $z^{nr}(t,\xi) = E\{z(t,\xi)\}$, $z^{r}(t,\xi) = z(t,\xi) - E\{z(t,\xi)\}$, and $z(t,\xi) = z^{nr}(t,\xi) + z^{r}(t,\xi)$.

Substituting (1) into (2) and rearranging, we obtain the familiar notation for linear modulations,

$$y_l = \sum_i \beta_i c_{i,l-ir} + n_l \tag{3}$$

where $c_{i,l\text{-}ir}$ is the complex *received* pulse, accounting for all effects between the original phasor sequence and the received signal[2] (the transmitter pulse shape, carrier frequency offset, carrier phase, symbol timing, and the Rician fading channel). However the received pulse has an extra parameter, *i*, since each received pulse is different, due to the time-varying channel. The received pulse is defined as

$$c_{i,l-ir} = \int_{-\infty}^{\infty} h\big((l-ir)T_r - t_0 - \xi\big) z\big(lT_r - t_0, \xi\big) \exp\big(j2\pi l f_0 T_r + j\phi\big) d\xi \,. \tag{4}$$

It can be split into a specular component, $c^{nr}_{i,l-ir}$, and a diffuse component, $c^{r}_{i,l-ir}$, due to the non-random and random components of the channel. Thus $c_{i,l-ir} = c^{nr}_{i,l-ir} + c^{r}_{i,l-ir}$.

In practice the transmitter pulse is restricted to a finite duration, so that $h(t) = 0$ outside $t \in [0; HT)$, where H is the pulse length in symbol periods; from physical considerations, the maximum total delay spread of the channel is upper bounded by some known τ, so that $z(t,\xi) = 0$, outside $\xi \in [0; \tau]$; and the timing error, t_0, is *a priori* known to be upper bounded by some T_0, so that $t_0 \in [0, T_0]$. Define the length of the received pulse (the channel memory) in symbol periods as[3] $L = \lceil H + (\tau + T_0)/T \rceil$. The *i*th received pulse is fully located within the interval $ir \le l \le (L + ir) - 1$, and (3) is rewritten as $y_l = \sum_{i=-L+1+\lfloor l/r \rfloor}^{\lfloor l/r \rfloor} \beta_i c_{i,l-ir} + n_l$.

From (3), the signal autocovariance obeys

$$R^r_{yy,l,m}(\beta) = E\left\{ y^r_l \overline{y}^r_m \middle| \beta \right\} = \sum_{i=-L+1+\lfloor l/r \rfloor}^{\lfloor l/r \rfloor} \sum_{k=-L+1+\lfloor m/r \rfloor}^{\lfloor m/r \rfloor} \beta_i \overline{\beta}_k \tfrac{1}{2} E\left\{ c^r_{i,l-ir} \overline{c}^r_{k,m-kr} \right\} + \tfrac{1}{2} E\left\{ n_l \overline{n}_m \right\} \tag{5}$$

---

[2] If the noise-limiting filter's passband is too narrow, its impulse response affects the received pulse shape through a further convolution. Low sampling rates have been used [8-10], but the influence of an IF filter has not been addressed. In effect the filter's passband is widened without increasing the sampling rate, and then the additional aliased noise is ignored.

[3] The following notation is used: $\lfloor x \rfloor$ and $\lceil x \rceil$ are the floor and ceiling functions respectively; an overbar, $\overline{x}$, denotes complex conjugation; x^H is the Hermitian transpose of *x*; *x* mod *y* denotes the remainder of *x*/*y*; $x^{(i)}(t)$ is the *i*th derivative of *x*(*t*); and ${}^iC_k = i!/k!(i\text{-}k)!$

where $\frac{1}{2}E\{n_l \bar{n}_m\} = N_0/T_r\,\delta_{lm}$. From (4), the received pulse autocovariance is

$$\begin{aligned}\tfrac{1}{2}E\left\{c^r_{i,l-ir}\bar{c}^r_{k,m-kr}\right\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h\left((l-ir)T_r - t_0 - \xi_1\right)\bar{h}\left((m-kr)T_r - t_0 - \xi_2\right)\times \\ \tfrac{1}{2}E\left\{z^r\left(lT_r - t_0, \xi_1\right)\bar{z}^r\left(mT_r - t_0, \xi_2\right)\right\}\exp\left(j2\pi(l-m)f_0T_r\right)d\xi_1 d\xi_2\end{aligned} \tag{6}$$

where the expectation is implicitly conditioned on the synchronisation parameters. Thus computing the received pulse autocovariance requires either knowledge of the pulse shape, channel autocovariance, and synchronisation parameters, or a time interval over which the channel autocovariance and synchronisation parameters are quasi-stationary long enough for time-averaging or an estimation strategy to converge. Note that (6) does not depend on ϕ.

Often the channel autocovariance satisfies a WSSUS model [16], so that

$$\tfrac{1}{2}E\left\{z^r\left(lT_r, \xi_1\right)\bar{z}^r\left(mT_r, \xi_2\right)\right\} = \delta\left(\xi_1 - \xi_2\right)P^r_{zz}\left((l-m)T_r, \xi_1\right) \tag{7}$$

where $P^r_{zz}\left((l-m)T_r, \xi_1\right)$ is the channel autocovariance function. When the time autocovariance is the same for all values of delay (e.g. when the Doppler spread arises from the mobile's motion and the expectation in (7) includes the multipath arrival angles), this simplifies to

$$\tfrac{1}{2}E\left\{z^r\left(lT_r, \xi_1\right)\bar{z}^r\left(mT_r, \xi_2\right)\right\} = \delta\left(\xi_1 - \xi_2\right)R^r_{tt}\left((l-m)T_r\right)P^r_{\xi\xi}\left(\xi_1\right) \tag{8}$$

$R^r_{tt}\left((l-m)T_r\right)$ is the autocovariance of each tap over time, and $P^r_{\xi\xi}\left(\xi_1\right)$ is the mean tap power. Thus (6) simplifies to

$$\begin{aligned}&\tfrac{1}{2}E\left\{c^r_{i,l-ir}\bar{c}^r_{k,m-kr}\right\} = \\ &\int_{-\infty}^{\infty} h\left((l-ir)T_r - t_0 - \xi\right)\bar{h}\left((m-kr)T_r - t_0 - \xi\right)P_{\xi\xi}\left(\xi\right)d\xi\, R_{tt}\left(l-m\right)\exp\left(j2\pi(l-m)f_0T_r\right)\end{aligned} \tag{9}$$

The average bit energy to noise spectral density is defined as

$$\frac{E_b}{N_0} = \frac{\tfrac{1}{2}E\left\{\int_{-\infty}^{\infty}\left|\int_{-\infty}^{\infty}\beta_i h\left(t-\xi\right)z\left(t,\xi\right)d\xi\right|^2 dt\right\}}{N_0 log_2 M} \tag{10}$$

and the Rice factor, *K*, equals

$$K = \frac{\frac{1}{2} E\left\{ \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} \beta_i h(t_1 - \xi_1) z^{nr}(t_1, \xi_1) d\xi_1 \right|^2 dt_1 \right\}}{\frac{1}{2} E\left\{ \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} \beta_k h(t_2 - \xi_2) z^{r}(t_2, \xi_2) d\xi_2 \right|^2 dt_2 \right\}} \tag{11}$$

D. Notation

It is helpful to have vector and matrix representations of these quantities. Define the length-B vectors, $\mathbf{y}_l^r = \left[y_{-B+l}^r, \ldots, y_{l-1}^r \right]^T$ and $\mathbf{r}_{\mathbf{yy},l}^r(\beta) = \frac{1}{2} E\left\{ \mathbf{y}_{l \bmod r}^r \bar{y}_{l \bmod r}^r \middle| \beta \right\}$; and the $B \times B$ matrix, $\mathbf{R}_{\mathbf{yy},l}^r(\beta) = \frac{1}{2} E\left\{ \mathbf{y}_{l \bmod r}^r \mathbf{y}_{l \bmod r}^{r,H} \middle| \beta \right\}$. We define the length-$(B+r)$ vectors, $\mathbf{y}_{f,ir} = \left[y_{ir-B}, \ldots, y_{(i+1)r-1} \right]^T$ and $\mathbf{n}_{f,ir} = \left[n_{ir-B}, \ldots, n_{(i+1)r-1} \right]^T$; the length-$(B+r)L$ vector, $\mathbf{c}_{f,ir} = \left[\left(c_{-L+1+\lfloor (ir-B)/r \rfloor, ir-B-\left(-L+1+\lfloor (ir-B)/r \rfloor\right) r}, \ldots, c_{\lfloor (ir-B)/r \rfloor, ir-B-\lfloor (ir-B)/r \rfloor r} \right), \ldots, \left(c_{-L+1+\lfloor ((i+1)r-1)/r \rfloor, (i+1)r-1-\left(-L+1+\lfloor ((i+1)r-1)/r \rfloor\right) r}, \ldots, c_{\lfloor ((i+1)r-1)/r \rfloor, (i+1)r-1-\lfloor ((i+1)r-1)/r \rfloor r} \right) \right]^T$; the $(B+r) \times (B+r)$ matrices, $\mathbf{R}_{\mathbf{yy},f}^r(\beta) = \frac{1}{2} E\left\{ \mathbf{y}_{f,0}^r \mathbf{y}_{f,0}^{r,H} \middle| \beta \right\}$, $\mathbf{R}_{\mathbf{nn},f} = \frac{1}{2} E\left\{ \mathbf{n}_{f,ir} \mathbf{n}_{f,ir}^H \right\}$; and the $(B+r) \times (B+r)L$ matrices, $\beta_{f,i} = \mathrm{diag}\left(\left(\beta_{-L+1+\lfloor (ir-B)/r \rfloor}, \ldots, \beta_{\lfloor (ir-B)/r \rfloor} \right), \ldots, \left(\beta_{-L+1+\lfloor ((i+1)r-1)/r \rfloor}, \ldots, \beta_{\lfloor ((i+1)r-1)/r \rfloor} \right) \right)$, and $\mathbf{R}_{\mathbf{cc},f}^r = \frac{1}{2} E\left\{ \mathbf{c}_{f,ir}^r \mathbf{c}_{f,ir}^{r,H} \right\}$. The superscripts *nr* and *r* apply as required.

III. Receiver Derivation

The MLSE receiver searches all allowed symbol sequences in the transmission interval and chooses the one with maximum likelihood. In this section, we derive the sequence metric, then manipulate it into a recursive form suited for on-line detection.

By assuming the synchronisation parameters are unknown but not time-varying, they can be regarded as non-random constants. Hence the signal, y_l, conditioned on the phasor sequence, is still complex Gaussian, since it is a deterministic linear combination of only the complex Gaussian random variables, $c_{i,l-ir}$ and n_l. The analysis of [6] applies, and the sequence log-likelihood can be written as

$$\Lambda(\beta) = \sum_{l} \frac{\left| y_l - y_{l|l-1}(\beta) \right|^2}{\sigma^2_{l|l-1}(\beta)} + \ln\left(\sigma^2_{l|l-1}(\beta)\right) \tag{12}$$

where $y_{l|l-1}(\beta)$ is the expected value of y_l, given the past received samples and a hypothesised symbol sequence; and $\sigma^2_{l|l-1}(\beta)$ is the variance of the prediction. From the principle of orthogonality, the expectation, $y_{l|l-1}(\beta)$, is the MMSE prediction of the Gaussian random variable, y_l, and $\left(y_l - y_{l|l-1}(\beta)\right)\big/\sigma_{l|l-1}(\beta)$ is the Innovations process [11]. $y_{l|l-1}(\beta)$, is computed by an MMSE predictor, as

$$y_{l|l-1}(\beta) = y^{nr}_l(\beta) + \sum_{k=1}^{\infty} b^{ML}_{l,k}(\beta)\left(y_{l-k} - y^{nr}_{l-k}(\beta)\right) \tag{13}$$

where $b^{ML}_{l,k}(\beta)$ is the kth tap for the ML predictor of $y^r_l = y_l - y^{nr}_l$, assuming the transmitted sequence $\{\beta_i\}$; and $y^{nr}_l(\beta)$ is the expected value of y_l, given a hypothesised sequence $\{\beta_i\}$,

$$y^{nr}_l(\beta) = \sum_{\substack{i= \\ -L+1+\lfloor l/r \rfloor}}^{\lfloor l/r \rfloor} \beta_i c^{nr}_{i,l-ir} = \sum_{\substack{i= \\ -L+1+\lfloor l/r \rfloor}}^{\lfloor l/r \rfloor} \beta_i \int_{-\infty}^{\infty} h\left((l-ir)T_r - t_0 - \xi\right) z^{nr}\left(lT_r - t_0, \xi\right) \exp\left\{j2\pi l f_0 T_r + j\phi\right\} d\xi \tag{14}$$

These predictor tap weights depend on the complete history of transmitted symbols. To avoid a tree search, the predictors are restricted to have a fixed number of taps, B, chosen to be large enough that there is a minor BER penalty only. The tap weights are arranged in a vector, $\mathbf{b}_l(\beta) = \left[b_{l,B}(\beta), \ldots, b_{l,1}(\beta)\right]^T$, where $b_{l,k}(\beta)$ is the kth tap for the MMSE predictor of y_l, assuming the transmitted sequence $\{\beta_i\}$. The prediction, the tap weights, and the MMSEs are computed according to

$$y_{l|l-1}(\beta) = y^{nr}_l + \mathbf{b}_l(\beta)^T \mathbf{y}^r_l, \tag{15}$$

$$\mathbf{R}^r_{\mathbf{yy},l}(\beta)\bar{\mathbf{b}}_l(\beta) = \mathbf{r}^r_{\mathbf{yy},l}(\beta), \tag{16}$$

$$\sigma^2_{l|l-1}(\beta) = R^r_{yy,l,l}(\beta) - \mathbf{b}^H_l(\beta)\mathbf{r}^r_{\mathbf{yy},l}(\beta) \tag{17}$$

It is easy to show from (5) that the signal autocovariances in (16) and (17) are different for each combination of the $W = \lceil B/r \rceil + L$ hypothesised symbols in the vicinity of

l, up to the *P*-ary phase ambiguity in $\{\beta_i\}$. These symbols are labelled the hypothesis vector, $\{\text{amplitude}(\beta_{-W+1+\lfloor l/r \rfloor}), \beta_{-W+2+\lfloor l/r \rfloor}, .., \beta_{\lfloor l/r \rfloor}\}$. Using fixed length predictors (and assuming that signal mean and autocovariance are known) has transformed the tree search into a trellis search, where the *i*th symbol's branch metric is

$$\lambda_i(\beta) = \sum_{l=ir}^{(i+1)r-1} \frac{\left| y_l - y_{l|l-1}(\beta) \right|^2}{\sigma^2_{l|l-1}(\beta)} + \ln\left(\sigma^2_{l|l-1}(\beta)\right) \tag{18}$$

There are M^W/P distinct branch metrics, and the receiver has M^{W-1}/P states.

In the purely specular channel, $y_l^r = 0$, $y_{l|l-1}(\beta) = y_l^{nr}(\beta)$, and $\sigma^2_{l|l-1}(\beta) = 2N_0/T_r$ is independent of the hypothesised sequence and can be neglected. Accordingly, (18) reduces to a Euclidean distance, which is related to (5) in [2]: i.e. the conventional MLSE receiver structure for a time-invariant, known channel.

In the purely diffuse case, a Euclidean distance is *not* computed between the signal and a noiseless, hypothesised version of the signal. Instead, a hypothesised sequence's predictors check whether the received sample sequence is internally consistent with that hypothesised sequence. This idea is represented in figure 4. The signal evolves in a correlated manner, according to the non-stationary transmitted signal and the correlated channel. Thus the correlation of the signal is characteristically determined by the transmitted sequence, and it is this property that is checked by the predictors.

We see that this receiver structure achieves synchronisation in a novel manner, since synchronisation, channel estimation and detection are performed jointly by the predictor tap weights. A residual carrier offset causes a rotation of the complex signal around the time axis. The predictor tap weights are computed with this knowledge. In fact their complex tap weights rotate helically around the time axis in the opposite direction to the rotation present in the received samples [18], thereby cancelling it. This is a boon in fast fading channels, since the channel's Doppler spread makes PLL-based carrier acquisition schemes inappropriate [13]. Using predictors makes the carrier phase irrelevant, since the same carrier phase multiplies both the received sample being predicted and the signal's past samples.

Symbol timing is also dealt with by the receiver, since it is fractionally-spaced. By defining a new pulse shape, $g(t) = h(t\text{-}t_0)$, and recognising that the time shift, t_0, in the stationary fading process, $z(t,\xi)$ can be neglected, then the received pulse can be rearranged as

$$c_{i,l-ir} = \int_{-\infty}^{\infty} h\big((l-ir)T_r - t_0 - \xi\big) z^r\left(lT_r - t_0, \xi\right) \exp\big(j2\pi l f_0 T_r + j\phi\big) d\xi$$
$$\sim \int_{-\infty}^{\infty} g\big((l-ir)T_r - \xi\big) z^r\left(lT_r, \xi\right) \exp\big(j2\pi l f_0 T_r + j\phi\big) d\xi \quad . \tag{19}$$

Thus acquiring symbol timing is equivalent to detecting the signal, given the transmitted pulse shape, $g(t)$. However, the receiver is designed for arbitrary pulse shapes, as long as they are restricted to L symbol periods in length. This is satisfied by both $h(t)$ and $g(t)$, from the definition of L. Detection can only proceed when either the signal's autocovariance is known (which requires explicit knowledge of t_0), or when the receiver has sufficient time to estimate the autocovariance.

During detection, the receiver makes rM^W/P complex B-tap predictions, on the zero mean signal, $y_l^r = y_l - y_l^{nr}$. The signal mean, y_l^{nr}, can be precomputed up to its Doppler shift, so the receiver must make $2rM^WB/P$ complex operations per symbol period, and M^{W-1}/P M-way comparisons. The receiver may need to compute the predictor tap weights and MMSEs from the signal autocovariance matrices; this requires approximately $rM^WB^3/6P$ complex operations.

IV. Estimating the Signal Mean and Autocovariance

From (15) and (16), the MLSE receiver requires the signal mean and autocovariance. In this section we describe an effectively optimal "parameter-minimisation" scheme to estimate these quantities. It will become clear that the scheme is impractically complicated, but it does demonstrate how quickly the signal's mean and autocovariance can be estimated. The transmitted symbol sequence is assumed to be known, either through a training sequence, tentative decisions, or because the receiver employs PSP and has conditioned on the transmitted sequence. In the latter case, each survivor has an estimator.

Using the previous definitions,

$$\mathbf{y}_{f,ir}^{nr} = \beta_{f,i} \mathbf{c}_{f,ir}^{nr} \tag{20}$$

$$\mathbf{R}_{\mathbf{yy},f}^{r}(\beta) = \beta_{f,i} \mathbf{R}_{\mathbf{cc},f}^{r} \beta_{f,i}^{H} + \mathbf{R}_{\mathbf{nn},f} \tag{21}$$

The signal mean is calculated from $\mathbf{c}_{f,ir}^{nr}$. The predictor tap weights and MMSEs are

computed from the r (B+1)×(B+1) submatrices of $\mathbf{R}^{r}_{\mathbf{yy},f}(\beta)$, which can in turn be computed from the channel and noise autocovariance matrices, $\mathbf{R}^{r}_{\mathbf{cc},f}$ and $\mathbf{R}_{\mathbf{nn},f}$.

When the signal mean and predictor tap weights are estimated perfectly, the mean square prediction error is at a minimum. Therefore an intuitively satisfying and near-optimal estimation scheme represents $\mathbf{R}^{r}_{\mathbf{cc},f}$, $\mathbf{R}_{\mathbf{nn},f}$, and $\mathbf{c}^{nr}_{f,ir}$ as a vector of parameters, then searches for the parameter vector that minimises the total squared prediction error, $\sum_{l}\left|y_l - y_{l|l-1}(\beta)\right|^2$, over the transmission duration, as shown in figure 5. The parameter vector is initialised to the parameters' *a priori* estimates. From this parameter vector, the signal mean and autocovariance are computed, then predictor tap weights. The signal mean and predictor tap weights are used to predict the received samples of the past signal samples, conditioned on a hypothesised sequence, with some total squared prediction error. The minimisation algorithm iteratively searches for the parameter vector that minimises this error. Assuming the algorithm converges to the global minimum, the final set of predictor tap weights then approximates the minimum mean square error (MMSE) set.

In this way, $\mathbf{R}^{r}_{\mathbf{cc},f}$ and $\mathbf{R}_{\mathbf{nn},f}$ are estimated up to a scaling ambiguity, which must be corrected when computing the MMSEs, $\sigma^2_{l|l-1}(\beta)$. Therefore, $\mathbf{R}^{r}_{\mathbf{cc},f}$ and $\mathbf{R}_{\mathbf{nn},f}$ are scaled so

$$\sum_{l}\left|y_l\right|^2 = \sum_{l}\sum_{\substack{i=\\ -L+1+\lfloor l/r\rfloor}}^{\lfloor l/r\rfloor}\sum_{\substack{k=\\ -L+1+\lfloor l/r\rfloor}}^{\lfloor l/r\rfloor}\beta_i\overline{\beta}_k E\left\{c_{i,l-ir}\overline{c}_{k,l-kr}\right\} + \frac{N_0}{T_r} \tag{22}$$

There is no "best set" of parameters, since *a priori* channel information guides the parameter selection. Some comments can be made. Since $\mathbf{R}_{\mathbf{nn},f} = N_0/T_r\mathbf{I}$, it only requires one parameter, N_0/T_r. In parameterising $\mathbf{R}^{r}_{\mathbf{cc},f}$ when there is no *a priori* information, the entries of $\mathbf{R}^{r}_{\mathbf{cc},f}$ or of its Cholesky decomposition (to guarantee positive definiteness) are appropriate. The number of parameters can be reduced by exploiting the stationarity and Hermitian symmetry of $\mathbf{R}^{r}_{\mathbf{cc},f}$. Normally however there is considerable *a priori* information, such as the transmitted pulse shape and the mathematical structure of the received pulse autocorrelation, (9). Accordingly, a superior set of parameters is f_0T, t_0/T, and parameters for $R^{r}_{tt}(t)$, $P^{r}_{\xi\xi}(\xi)$. These functions can be expanded as the weighted sum of a set of basis

functions, where the weights as used as parameters. The basis functions should be chosen so that $R_{tt}^{r}(t)$ and $P_{\xi\xi}^{r}(\xi)$ are accurately described by a few weighted functions only. A polynomial expansion is convenient. The channel's time autocovariance can be accurately represented by the polynomial expansion,

$$R_{tt}^{r}(t) \approx \sum_{i=0}^{I_t} \lambda_i t^{2i} \tag{23}$$

when the number, I_t+1, of coefficients, $\lambda_0 = 1, \lambda_1, \ldots, \lambda_{I_t}$, is sufficiently large. $z^r(t,\xi)$ is assumed to have independent real and imaginary parts, so that odd terms in the autocovariance expansion can be neglected. Similarly, the channel autocovariance in frequency, $P_{\xi\xi}^{r}(\xi) \leftrightarrow R_{ff}^{r}(f)$, can be parameterised by the I_f+1 coefficients, $\mu_0 = 1, \mu_1, \ldots, \mu_{I_f}$, given the polynomial expansion,

$$R_{ff}^{r}(f) \approx \sum_{i=0}^{I_f} \mu_i (-j2\pi f)^i \leftrightarrow \sum_{i=0}^{I_f} -1^i \mu_i \delta^{(i)}(\xi) \tag{24}$$

and a sufficiently large I_f. With this expansion, the integral in (9) simplifies to

$$\int_{-\infty}^{\infty} h(t_1 - \xi)\bar{h}(t_2 - \xi) P_{\xi\xi}^{r}(\xi) d\xi = \sum_{i=0}^{I_f} \sum_{k=0}^{i} \mu_i \, {}^{i}C_k h^{(k)}(t_1)\bar{h}^{(i-k)}(t_2) \tag{25}$$

When there is no *a priori* information about the signal mean other than L, then the appropriate parameters for the impulse response are the real and imaginary parts of $c_{i,0}^{nr}, \ldots, c_{i,Lr-1}^{nr}$, and the Doppler shift of the specular component relative to the diffuse component. In the mobile radio channel, the signal mean often corresponds to the direct path, so the channel mean's impulse response can be parameterised by the direct path's complex gain only.

Simulations indicate that a simple-minded minimisation algorithm can easily converge to a local minimum. This may be overcome by invoking the minimisation algorithm many times with different initial parameters vectors, or by employing the technique of simulated annealing [19]. When computed from parameters, $\mathbf{R}_{\mathbf{yy},f}^{r}(\beta)$ is not guaranteed positive definite. Calculated MMSEs can be negative, and thus the branch metric, [18], operates incorrectly. This generally occurs when too few parameters are used to describe $R_{tt}^{r}(t)$ and $P_{\xi\xi}^{r}(\xi)$, or when there are insufficient signal samples for the parameters to be

properly estimated. In the rare times this problem arose in simulations, the noise power was repeatedly increased by 20% until all MMSEs were strictly positive.

V. Receiver Analysis

We seek the receiver's BER for a Rician fading, frequency-selective channel in white noise. The rotationally invariant code is assumed to be differential encoding, and the signal mean and autocovariance are ideally known. The same analytic framework applies to coded transmissions also.

First some notation is defined. The actual transmitted sequence is denoted by $\{\beta^{u,v}\}$. Potential error sequences are written as $\{\beta^{u,v,w}\}$. With trellis-based receivers, errors are dependent and appear as error events. The superscript *u* denotes the length of the error event under consideration. The superscript *v* enumerates each distinct transmitted sequence in the vicinity of the length, *u* error event. Each transmitted sequence can be confused with several others, so the error sequences are enumerated by a further index, *w*. When an error occurs, the ML sequence is one of the error sequences, $\{\beta^{u,v,w}\}$.

The probability that the sequence, $\{\beta^{u,v}\}$, is transmitted is labelled by $P\left(\beta^{u,v}\right)$. The probability that an error sequence has a better metric than the transmitted sequence (the pairwise probability of error) is denoted by $P\left(\beta^{u,v} \to \beta^{u,v,w}\right)$. The number of bit errors that arise from the error event is written $e\left(\beta^{u,v} \to \beta^{u,v,w}\right)$.

An upper bound on the BER can be deduced from a union bound over all error events. Since this is an infinite sum, it must be truncated. The truncated bound is a credible upper bound if at least the dominant error events are considered. Thus the BER bound is the union bound of the dominant error events, averaged across the transmitted sequences in the vicinity of the error event,

$$BER < \sum_{u,v,w} \frac{P\left(\beta^{u,v}\right) P\left(\beta^{u,v} \to \beta^{u,v,w}\right) e\left(\beta^{u,v} \to \beta^{u,v,w}\right)}{log_2\, M} \tag{26}$$

The form of an error sequence is $\{\beta^{u,v,w}\} = \left\{\beta^{u,v}\, exp\left(j\theta^{u,v,w}\right) + \varepsilon^{u,v,w}\right\}$, where the sequences $\{\varepsilon^{u,v,w}\}$ and $\{\theta^{u,v,w}\}$ specify the particular error sequence, and are constrained so that $\{\beta^{u,v,w}\}$ is also an allowed sequence. For an error event extending from the *i*th to the

(i+u-1)th symbol period, $\varepsilon_k^{u,v,w} = 0$ for $k < i$ and for $k > i+u-1$. $\theta_k^{u,v,w} = 0$ for $k < i+u$ and is constant for $k \geq i+u$. This remaining phase offset allows the error event to end even when phase lock is lost, since the rotationally invariant code prevents further bit errors. These constraints ensure that $\{\varepsilon^{u,v,w}\}$ and $\{\theta^{u,v,w}\}$ uniquely describe an error event. Without loss of generality, the first symbol error is aligned with time $i = 0$, so that the error sequence can be written as

$$\left\{\ldots \beta_{-2}^{u,v}, \beta_{-1}^{u,v}, \beta_0^{u,v} + \varepsilon_0^{u,v,w}, \ldots, \beta_{u-1}^{u,v} + \varepsilon_{u-1}, \beta_u^{u,v} \exp\left(j\theta_u^{u,v,w}\right), \beta_{u+1}^{u,v} \exp\left(j\theta_{u+1}^{u,v,w}\right), \ldots\right\} \tag{27}$$

The pairwise probability of error depends on the hypothesised and transmitted symbols in the vicinity of the error event, since they determine which predictor tap weights are used. Clearly the erroneous symbols $\left\{\beta_0^{u,v,w}, \ldots \beta_{u-1}^{u,v,w}\right\}$ affect the pairwise error probability. Define $Y = (L+u-1)r-1$. The signal samples, y_0 ... y_Y involve pulse tails from the erroneous symbols. When these samples are predicted, or are used in a prediction, the wrong predictor tap weights are used. Define $\omega = -B$ and $\Omega = (L+u-1)r-1+B$. The signal samples y_ω ... y_Ω, are used with y_0 ... y_Y in predictions, so in fact the symbols that affect the pairwise probability of error are $\left\{\beta_\psi^{u,v}, \ldots \beta_\Psi^{u,v}\right\}$ and $\left\{\beta_\psi^{u,v,w}, \ldots \beta_\Psi^{u,v,w}\right\}$, where $\psi = -W+1$ and $\Psi = W+u-2$. Since the ISI from these symbols is different in each case, pairwise error probabilities must be tediously computed for each ISI combination, up to the P-ary rotational ambiguity. The transmission probability equals, $P\left(\beta^{u,v}\right) = \dfrac{P}{M^{\Psi-\psi+1}}$.

In slow Rayleigh fading channels, the mean fade duration is long, so there are many different yet likely error events. The union bound of (26) is very loose. However, when the fading is sufficiently fast and the SNR high, fade durations are a fraction of a symbol period only. The dominant error event is the cycle slip, and accordingly the union bound is asymptotically tight. A tight union bound also arises when the channel is strongly Rician, since the dominant error events are short there too.

We define normalised predictor tap weights, a bias term and the signal mean as

$$b_{l,k}^{u,v} = \begin{cases} \dfrac{1}{\sigma_{l|l-1}\left(\beta^{u,v}\right)} & k = 0 \\ \dfrac{-b_{l,k}\left(\beta^{u,v}\right)}{\sigma_{l|l-1}\left(\beta^{u,v}\right)} & k = 1..B \end{cases} \quad b_{l,k}^{u,v,w} = \begin{cases} \dfrac{1}{\sigma_{l|l-1}\left(\beta^{u,v,w}\right)} & k = 0 \\ \dfrac{-b_{l,k}\left(\beta^{u,v,w}\right)}{\sigma_{l|l-1}\left(\beta^{u,v,w}\right)} & k = 1..B \end{cases} \tag{28}$$

$$\kappa_{min} = \sum_{l=0}^{\Omega} ln\left(\frac{\sigma_{l|l-1}^2\left(\beta^{u,v,w}\right)}{\sigma_{l|l-1}^2\left(\beta^{u,v}\right)}\right) \tag{29}$$

$$y_l^{nr,u,v} = y_l^{nr}\left(\beta^{u,v}\right) \quad y_l^{nr,u,v,w} = y_l^{nr}\left(\beta^{u,v,w}\right) \tag{30}$$

Then the pairwise probability of error is the probability that an erroneous sequence has a smaller path metric than the transmitted sequence,

$$\begin{aligned} &P\left(\beta^{u,v} \to \beta^{u,v,w}\right) = P\left(\Lambda^{u,v} > \Lambda^{u,v,w}\right) \\ &= P\left(\sum_{l=0}^{\Omega}\sum_{k=0}^{B}\sum_{k'=0}^{B}\left[\begin{array}{l} b_{l,k}^{u,v}\bar{b}_{l,k'}^{u,v}\left(y_{l-k} - y_{l-k}^{nr,u,v}\right)\left(\bar{y}_{l-k'} - \bar{y}_{l-k'}^{nr,u,v}\right) - \\ b_{l,k}^{u,v,w}\bar{b}_{l,k'}^{u,v,w}\left(y_{l-k} - y_{l-k}^{nr,u,v,w}\right)\left(\bar{y}_{l-k'} - \bar{y}_{l-k'}^{nr,u,v,w}\right) \end{array}\right] > \kappa_{min}\right) \end{aligned} \tag{31}$$

from (18). In Rayleigh fading this simplifies to

$$P\left(\beta^{u,v} \to \beta^{u,v,w}\right) = P\left(\sum_{l=0}^{\Omega}\sum_{k=0}^{B}\sum_{k'=0}^{B}\left(b_{l,k}^{u,v}\bar{b}_{l,k'}^{u,v} - b_{l,k}^{u,v,w}\bar{b}_{l,k'}^{u,v,w}\right)y_{l-k}\bar{y}_{l-k'} > \kappa_{min}\right) \tag{32}$$

Define $\kappa^{u,v,w}$ as the left-hand-side of the inequality and the column vector, $\mathbf{y}^{u,v} = \left[y_\omega,\ldots,y_\Omega; y_\omega^{nr,u,v},\ldots,y_\Omega^{nr,u,v}; y_\omega^{nr,u,v,w},\ldots,y_\Omega^{nr,u,v,w}\right]^T$. $\kappa^{u,v,w}$ is a Gaussian quadratic form in the Gaussian random variables, $\mathbf{y}^{u,v}$, and can be written as, $\kappa^{u,v,w} = \mathbf{y}^{u,v,H}\mathbf{Y}^{u,v,w}\mathbf{y}^{u,v}$, where the kernel, $\mathbf{Y}^{u,v,w}$, is a Hermitian symmetric matrix, defined by (31). The entries in the autocovariance matrix, $\mathbf{R}_{\mathbf{yy}}^r = E\left\{\left(\mathbf{y}^{u,v} - E\left\{\mathbf{y}^{u,v}\right\}\right)\left(\mathbf{y}^{u,v} - E\left\{\mathbf{y}^{u,v}\right\}\right)^H\right\}$, of $\mathbf{y}^{u,v}$, are given by (5) and (6). The characteristic function of a Gaussian quadratic form, $\kappa^{u,v,w}$, is given by [1], as

$$\Phi_\xi^{u,v,w}\left(\xi\right) = \frac{\exp\left(j\xi E\left(\mathbf{y}^{u,v}\right)^H\mathbf{Y}^{u,v,w}\left(\mathbf{I} - j2\xi\mathbf{R}_{\mathbf{yy}}^r\mathbf{Y}^{u,v,w}\right)^{-1}E\left(\mathbf{y}^{u,v}\right)\right)}{\det\left|\mathbf{I} - j2\xi\mathbf{R}_{\mathbf{yy}}^r\mathbf{Y}^{u,v,w}\right|} \tag{33}$$

The pairwise probability is calculated by transforming this characteristic function into a pdf, then integrating over the error region, as

$$P\left(\beta^{u,v} \to \beta^{u,v,w}\right) = \int_{\kappa_{\min}}^{\infty} \int_{-\infty}^{\infty} \Phi_{\xi}^{u,v,w}(\xi) \exp\left(j\xi\kappa\right) d\xi d\kappa = \tfrac{1}{2} + \int_{-\infty}^{\infty} \frac{\Phi_{\xi}^{u,v,w}(\xi) \exp\left(-j\xi\kappa_{\min}\right)}{j2\pi\xi} d\xi \quad (34)$$

It is unclear how to easily evaluate this integral for the general Rician case, so numerical integration is appropriate. For Rayleigh fading channels, the vector of Gaussian random variables can be shortened to $\mathbf{y}^{u,v} = \mathbf{y} = \left[y_{\omega}, \ldots, y_{\Omega}\right]^T$, with the matrices $\mathbf{R}_{\mathbf{yy}}^{r}$ and $\mathbf{Y}^{u,v,w}$ cropped also. The characteristic function simplifies to

$$\Phi_{\xi}^{u,v,w}(\xi) = \det\left|\mathbf{I} - j2\xi\mathbf{R}_{\mathbf{yy}}\mathbf{Y}^{u,v,w}\right|^{-1} \quad (35)$$

and from standard residue calculus, the pairwise probability of error equals

$$P\left(\beta^{u,v} \to \beta^{u,v,w}\right) = \begin{cases} 1 - \sum_{i, \Im\{p_i\}>0} \exp\left(-jp_i\kappa_{\min}\right) \prod_{\substack{k=1, \\ k\neq i}}^{I} \frac{1}{\left(1 - p_i / p_k\right)} & \kappa_{\min} \leq 0 \\ + \sum_{i, \Im\{p_i\}<0} \exp\left(-jp_i\kappa_{\min}\right) \prod_{\substack{k=1, \\ k\neq i}}^{I} \frac{1}{\left(1 - p_i / p_k\right)} & \kappa_{\min} \geq 0 \end{cases} \quad (36)$$

where p_i is the ith pole of (35). These poles equal $1/2j$ multiplied by the eigenvalues of $\mathbf{R}_{\mathbf{yy}}\mathbf{Y}^{u,v,w}$, so the poles are found numerically. Equation (36) assumes that the poles of (35) are simple. This holds normally, but the more complicated case can also be dealt with [20].

In the non-fading case, $B = 0$, only the noise is a random variable, so $\mathbf{R}_{\mathbf{yy}}^{r}$ is partly diagonal and partly zero, leading to the usual result in terms of the $Q(.)$ function.

VI. Results

In this section, the novel aspects of the receiver are characterised by simulation, and analysis where possible. The simulated results are generated from a Monte-Carlo simulation of figure 1. Randomly generated data is transmitted and detected until at least 200 bit errors are observed. Unless otherwise stated, the simulation parameters are as follows. The data is differentially encoded BPSK. A root raised cosine pulse is used, with 50% excess bandwidth, and windowed with a Hanning window to $H = 3$ symbol periods. There is no carrier frequency offset. The channel is modelled by $N = 3$ Rayleigh fading taps, spaced equally over $\tau = 0.5T$ seconds, having equal mean power. Each tap is independent from other taps and modelled by passing complex white Gaussian noise through a 192-tap low pass FIR filter,

with impulse response [21], $f_l = J_{\frac{1}{4}}\left(2\pi f_D T_r |l|\right)|l|^{-\frac{1}{4}} \times \text{Hanning}\left(lT_r\right)$, creating a "windowed" $J_0\left(2\pi f_D T_r |l-m|\right)$. The channel is fast fading, $f_D T = 0.1$ The sampling rate for the transmitter, channel and receiver are all the same, with $r = 3$, so the noise-limiting filter is implicit in the discretisation of time. Synchronisation is assumed: $t_0 = 0$, $f_0 = 0$. The MMSE predictor tap weights are all precomputed and not further changed. The predictors have $B = 6$ taps.

The analysis of section V is applied whenever the channel is Rayleigh fading. Only cycle slips and one symbol, nearest neighbour error events are considered. In fast fading BPSK simulations, $f_D T = 0.1$, these two error events are dominant above approximately 20dB; above 30dB only the cycle slips contribute significantly to the BER.

In figure 6, a carrier offset is introduced between transmitter and receiver. In producing this figure only, the transmitter and channel are simulated at a sampling rate of $2r/T$, where $r = 2,.., 4$. Noise is added (equivalent to a final E_b/N_o of 25dB), and this received signal is filtered by a $40r$-tap low pass FIR filter with cutoff frequency, $r/2T$, then 1:2 sub-sampled. The filter is designed in Matlab using a Kaiser window with parameter 12. When the carrier offset is large enough, part of the signal lies in the stopband of the filter, causing signal distortion. This distortion is not accounted for when the predictor tap weights are calculated, so the BER degrades rapidly. This is seen in figure 6. The union bound only applies when the signal is not significantly distorted by the noise-limiting filter. The Doppler spread, windowed, transmitted pulse occupies approximately the -25dB bandwidth, $-0.9/T$ .. $0.9/T$. According, when $r = 2, 3, 4$ sample per symbol are taken, the noise-limiting filter can be neglected for any frequency offset up to approximately $\pm 0.1/T$, $\pm 0.6/T$, $\pm 1.1/T$ (in general the values depend on the SNR). The BER in this region is constant. Beyond this offset, the signal is distorted by the noise-limiting filter, and the BER gets worse as the signal distortion increases. Accordingly, given a sufficiently large r, the receiver can accommodate any carrier offset, albeit at the cost of linearly increasing complexity in r.

In figure 7, a timing offset, t_0, is introduced between transmitter and receiver. E_b/N_o equals 25dB. The predictor tap weights are computed assuming that the timing offset equals $\min(t_0, T_0)$. Thus when the timing offset is properly bounded, $t_0 \le T_0$, the receiver is able to detect the signal reliably. From figure 7, the BER is flat in this region. When t_0 exceeds T_0, there is a mismatch between the predictor tap weights and the received signal, and accordingly the receiver's BER degrades rapidly. The union bound only applies when the

timing mismatch is not significant. The mean variance of the Innovations process, $\frac{1}{N}\sum_{l=1}^{N}\left|y_l - y_{l|l-1}(\beta)\right|^2 \Big/ \sigma^2_{l|l-1}(\beta)$, is approximately unity when the receiver is operating properly. When the timing mismatch increases, it exhibits a cyclic variation, falling to unity as the timing mismatch is an integral number of symbol periods. The decrease at symbol-spaced intervals is due to the receiver reliably detecting adjacent symbols. Accordingly, given a sufficiently large T_0 and a training sequence (or t_0), the receiver can acquire any timing error, albeit at the cost of exponentially increasing complexity in T_0. However, when a training sequence is unavailable, the receiver can only acquire a timing error of up to T seconds, due to the T-second ambiguity in symbol timing.

In figure 8, a simple Rician fading channel is considered. Only the first tap has a non-zero mean, so the channel tends to an AWGN channel as the Rice component gets large, $K \to \infty$. The benefits are seen in figure 8, where the BER curves improve substantially for larger values of K. Recall that the signal mean and autocovariance are assumed known, so for $K \to \infty$, the problem tends to the classical detection problem solved in [17].

In figure 9, the influence of different constellations on the BER is examined. Due to the large number of states, the pulse length is reduced to H = 1.5 symbol periods and not windowed. There is a power penalty in increasing the system's spectral efficiency, and also a substantial increase in complexity. Both 8-PSK and the $A/P/\rho$ = 2/4/0.5 radially-symmetric constellation have the same spectral efficiency, yet the latter has a superior BER. Accordingly it merits further investigation. The union bound is tight at high SNR due to the short mean fade duration. Here the error events are almost exclusively cycle slips.

In figure 10, the acquisition performance of the parameter-minimisation scheme and an RLS-based scheme is studied at E_b/N_0 = 20dB. The receiver is trained for a duration with a random training sequence, then the predictor tap weights are fixed, and detection continues until 200 bit errors are detected. This is repeated 20 times to approximate the ensemble of all training sequences. In this way, the BER as a function of training length is calculated. Note that W = 6 symbols of the training sequence cannot be used since they contain pulse tails from unknown symbols. An unsynchronised, Rayleigh-fading signal is assumed, with $f_0 = 0.2/T$, ϕ = 20°, and $t_0 = 0.2T$.

The parameter-minimisation scheme uses the polynomial expansion, with I_t = 2 and I_f

= 6. The scheme learns the channel autocovariance accurately within 50 symbol periods.

As a comparison, a second scheme is considered where each predictor is directly adapted by its own RLS processor (there is no information pooling between transmitted sub-sequences or sample positions). At least B repetitions of each sub-sequence must be received before the receiver can begin detection, and each predictor can only be updated when its sample and sub-sequence is transmitted. Thus there is a minimum training period of $M^W B/P$ = 192 symbols, and its acquisition time is considerably longer than for the parameter-minimisation scheme. This is seen in figure 10.

In unpublished simulations, the first scheme adapted to Rician fading channels also. However, the carrier offset cannot be estimated perfectly, so the error in estimating $c_{i,l-ir}^{nr} \propto \exp\left(j2\pi l f_0 T_r\right)$ increases in time, l. This problem may be solved either by only using the estimate close to the training sequence, or by continual adaptation. To avoid the problem of local minima, the parameters were initialised to their correct values. The minimisation algorithm (Powell's method [19]) then modified the parameters until they matched the particular training sequence. Despite the favourable initialisation vector, the algorithm still required approximately 2000 iterations to converge. Powell's method does not exploit derivative information, which can potentially improve the rate of convergence. The EM algorithm may also offer faster convergence [27].

VII. Conclusions

The MLSE receiver structure of [11] is generalised to Rician fading and unsynchronised signals. The resulting structure deals with the specular and diffuse components of the received signal separately. The receiver's operation in Rayleigh channels is noteworthy. Given an adequate sampling rate (causing a linear increase in complexity), arbitrarily large carrier frequency offsets can be tolerated. Given an adequate upper bound on the timing error (causing an exponential increase in complexity), arbitrarily large timing offsets can be accommodated. Detection can only commence once accurate estimates of the channel and noise mean and autocovariance are available. In simulation, accurate estimates are available in approximately 50 symbol periods. The receiver's BER is derived in the frequency-selective, fast Rician fading channel, and it can be evaluated for the case of fast Rayleigh fading. The analysis shows good agreement with simulation.

VIII. References

[1] M.Schwartz, W.R.Bennett and S.Stein, “Communication Systems and Techniques,” McGraw-Hill, New York, 1966

[2] K.M.Chugg, A.Polydoros, “MLSE for an Unknown Channel - Part I: Optimality Considerations,” *IEEE Trans. Commun.*, vol. 44, no. 7, pp836-846, July 1996

[3] F. Davarian, “Mobile Digital Communications via Tone Calibration,” *IEEE Trans. Veh. Tech.*, vol. 36, no. 2, pp55-62, May 1987

[4] A.Bateman, “Feedforward Transparent Tone-in-Band: Its Implementations and Applications,” *IEEE Trans. Veh. Tech.*, vol. 39, no. 3, pp235-243, Aug. 1990

[5] J.K.Cavers, “An Analysis of Pilot Symbol Assisted Modulation for Rayleigh Fading Channels,” *IEEE Trans. Veh. Tech.*, vol. 40, no. 4, pp686-693, Nov. 1991

[6] J.H.Lodge and M.J.Moher, “Maximum Likelihood Sequence Estimation of CPM Signals Transmitted over Rayleigh Flat-Fading Channels,” *IEEE Trans. Commun.*, vol. 38, no. 6, pp787-794, June 1990

[7] W.C.Dam, D.P.Taylor, “An Adaptive Maximum Likelihood Receiver for Correlated Rayleigh-Fading Channels”, *IEEE Trans. Commun.*, vol. 42, no. 9, pp2684-2692, Sept. 1994

[8] W.C.Dam, “An Adaptive Maximum Likelihood Receiver for Rayleigh Fading Channels,” M.Eng. thesis, Hamilton Ont. Canada: McMaster University, 1990

[9] G.M.Vitetta, D.P.Taylor, “Maximum Likelihood Decoding of Uncoded and Coded PSK Signal Sequences Transmitted over Rayleigh Flat-Fading Channels,” *IEEE Trans. Comm.*, vol.43, no.11, pp2750-2758, Nov. 1995

[10] Q.Dai, E.Shwedyk, “Detection of Bandlimited Signals over Frequency-Selective Rayleigh Fading Channels,” *IEEE Trans. Commun.*, vol. 42, no. 2-4, pp941-950, 1994

[11] X.Yu and S.Pasupathy, “Innovations-Based MLSE for Rayleigh Fading Channels,” *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp1534-1544, Feb./Mar./Apr. 1995

[12] X.Yu, S.Pasupathy, “Error Performance of Innovations-Based MLSE for Rayleigh Fading Channels,” *IEEE Trans. Veh. Tech.*, vol. 45, no. 4, Nov. 1996, pp631-642

[13] W.C.Jakes, “Microwave Mobile Communications,” IEEE Press (re-issue), 1994

[14] R.Raheli, A.Polydoros, C.-K. Tzou, “Per-Survivor Processing: A General Approach MLSE in Uncertain Environments,” *IEEE Trans. Commun.*, vol. 43, no. 2-4, pp354-364, Feb-Apr 1995

[15] S.Haykin, “Adaptive Filter Theory,” Prentice Hall, Englewood Cliffs, New Jersey, 1991

[16] P.A.Bello, “Characterization of Randomly Time-Variant Linear Channels,” *IEEE Trans. Commun. Sys.*, vol. 11, pp360-393, Dec. 1962

[17] G.D.Forney, “Maximum-Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference,” *IEEE Trans. Inf. Theory*, vol. 18, no. 3, pp363-378, May 1972

[18] P.Y.Kam and C.H.Teh, “Reception of PSK Signals over Fading Channels via Quadrature Amplitude Estimation,” *IEEE Trans. Commun.*, vol. 31, no. 8, pp1024-1027, Aug. 1983

[19] W.H.Press et al, “Numerical Recipes in C,” 2nd Ed., Cambridge University Press, 1992

[20] B.D.Hart, “MLSE Diversity Receiver Structures,” PhD thesis, University of Canterbury, Christchurch, New Zealand, 1996

[21] D.Verdin and T.C.Tozer, “Generating a Fading Process for the Simulation of Land-Mobile Radio Communications,” *Electronics Letters*, vol. 29, no. 23, pp2011-2012, 11 Nov. 1993

[22] B.D.Hart, D.P.Taylor, “Extended MLSE Diversity Receiver for the Time- and Frequency-Selective Channel,” *IEEE Trans. Comm.*, vol. 45, no. 3, pp322-333, Mar. 1997

[23] R.H.Clarke, “A Statistical Theory of Mobile-Radio Reception,” *Bell Syst. Tech. J.*, vol 47, pp957-1000, July-Aug. 1968

[25] W.T.Webb, L.Hanzo, R.Steele, “Bandwidth Efficient QAM Schemes for Rayleigh Fading Channels,” *IEE Proc.*, vol. 138, Pt I, no. 3, Jun. 1991, pp169-175

[27] C.N.Georghiades, J.C.Han, “Sequence Estimation in the Presence of Random Parameters via the EM Algorithm,” *IEEE Trans. Comm.*, vol.45, no. 3, Mar. 1997, pp300-308

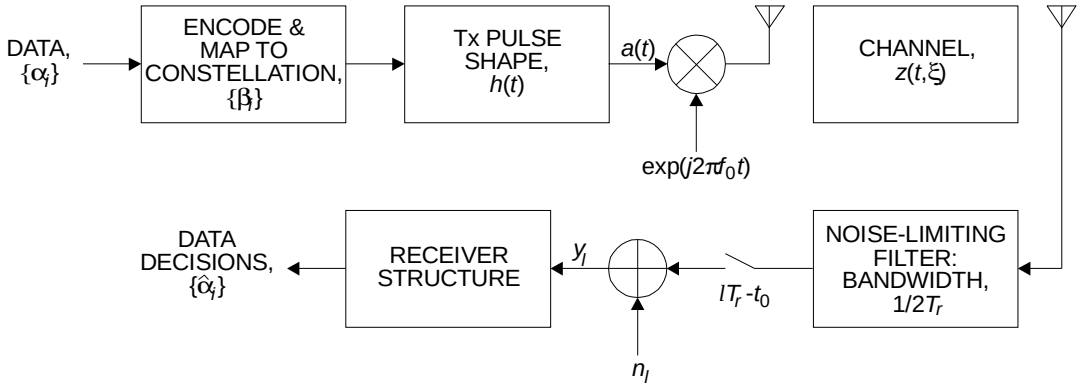

Figure 1: Structural representation of the communications system.

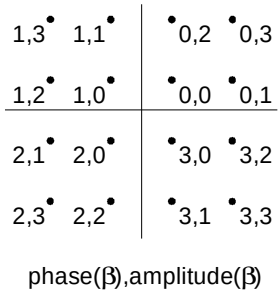

Figure 2: Constellation point labels for 16-QAM, illustrating labelling for differential phase encoding.

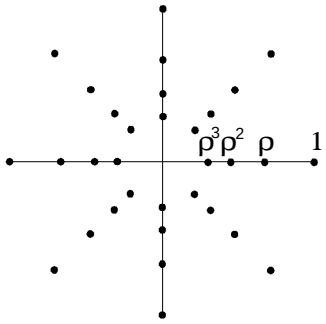

Figure 3: An example of a radially-symmetric constellation. $P = 8$, $A = 4$, $M = 32$.

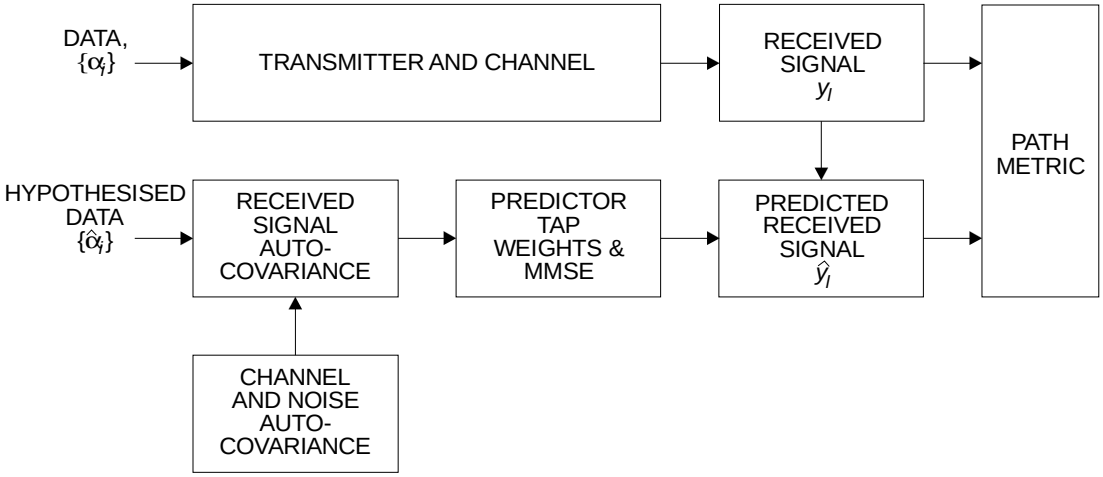

Figure 4: Predictor operation in Rayleigh fading channels

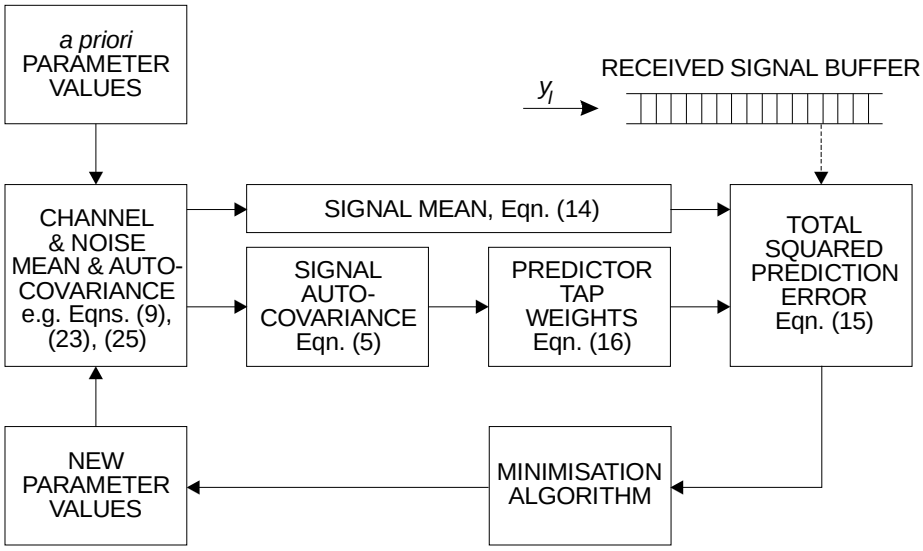

Figure 5: Iterative minimisation algorithm for estimating the channel and noise mean and autocorrelation.

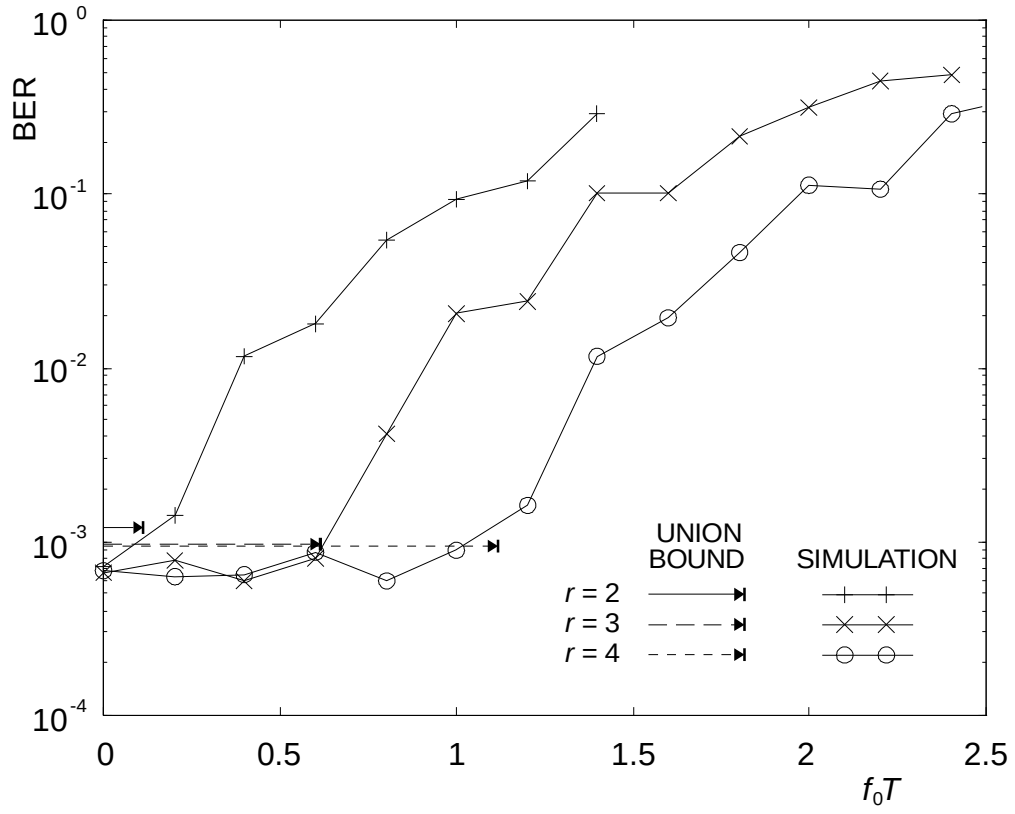

Figure 6: Effect of a carrier offset, f_0. The important simulation parameters are: 2-DPSK; f_DT = 0.1; τ/T = 0.5; E_b/N_0 = 25dB; K = 0; t_0 = 0; B = 2r.

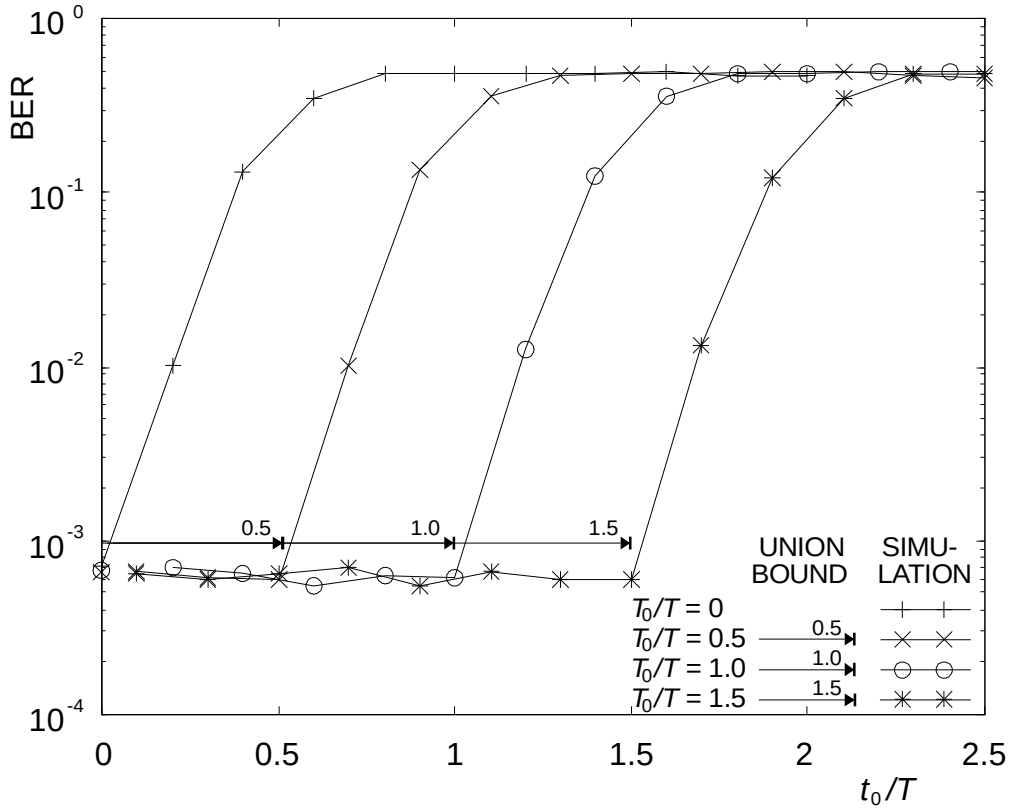

Figure 7: Effect of a timing offset, t_0. The important simulation parameters are: 2-DPSK; $f_D T$ = 0.1; τ/T = 0.5; E_b/N_0 = 25dB; K = 0; f_0 = 0; r = 3; $B = 2r$.

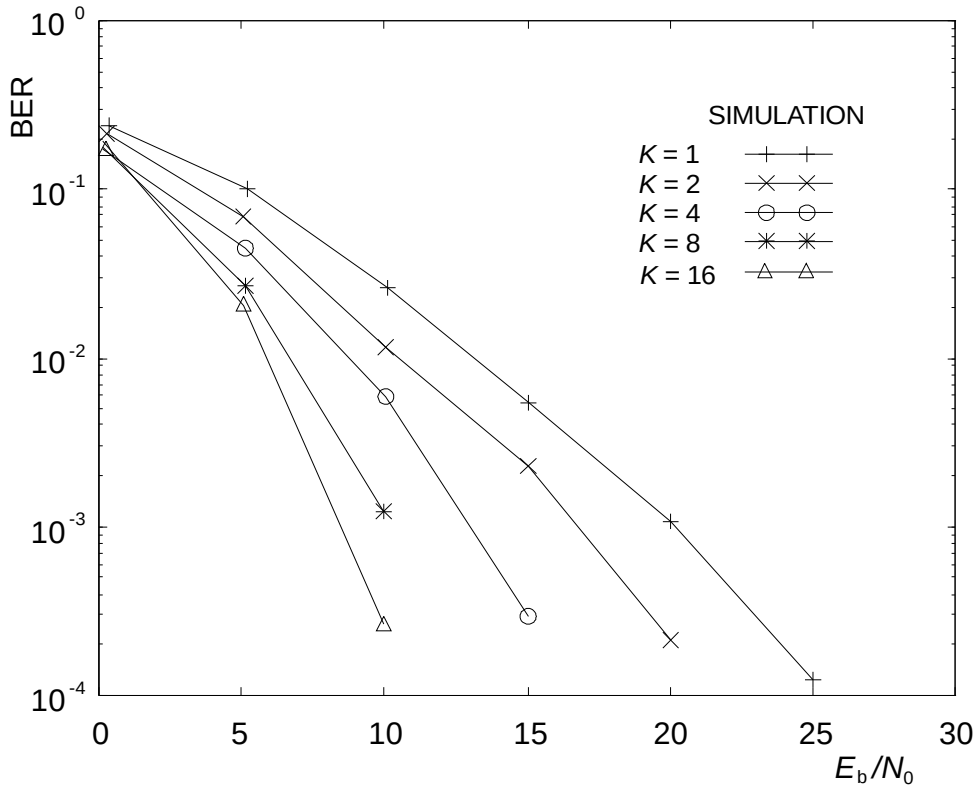

Figure 8: Benefits of a Rician fading channel. The important simulation parameters are: 2-DPSK; $f_D T = 0.1$; $\tau/T = 0.5$; $t_0 = 0$; $f_0 = 0$; $r = 3$; $B = 2r$.

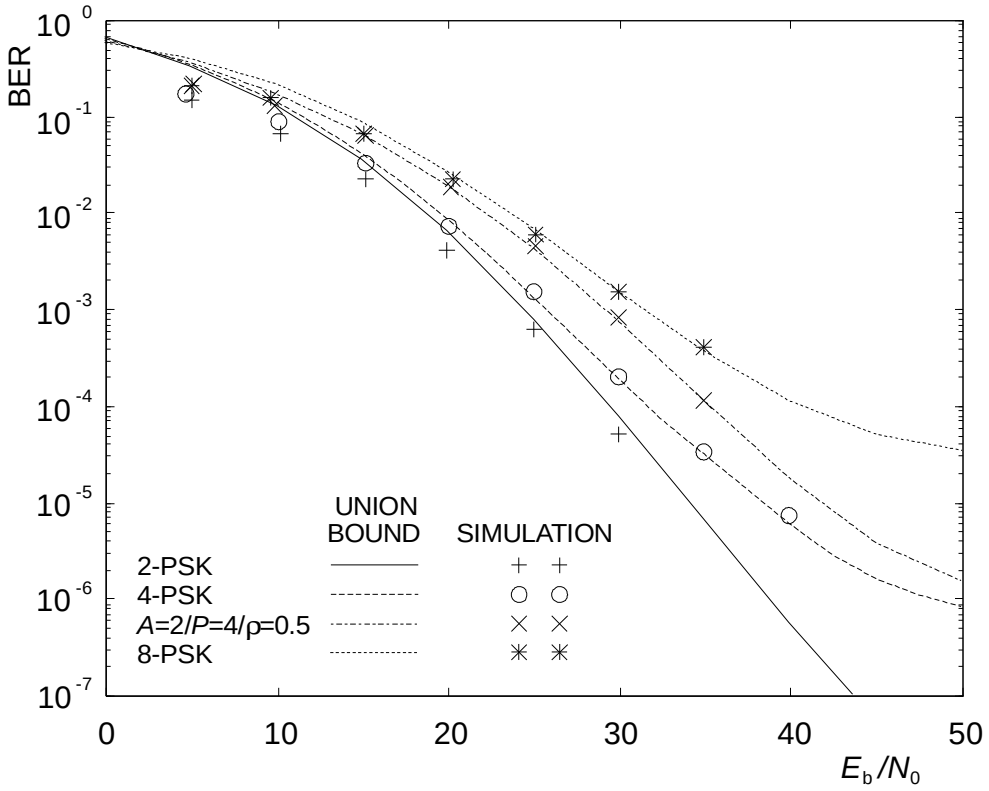

Figure 9: Influence of the constellation size on the BER. The important simulation parameters are: $f_D T = 0.1$; $\tau/T = 0.5$; $K = 0$; $t_0 = 0$; $f_0 = 0$; $r = 3$; $B = 2r$.

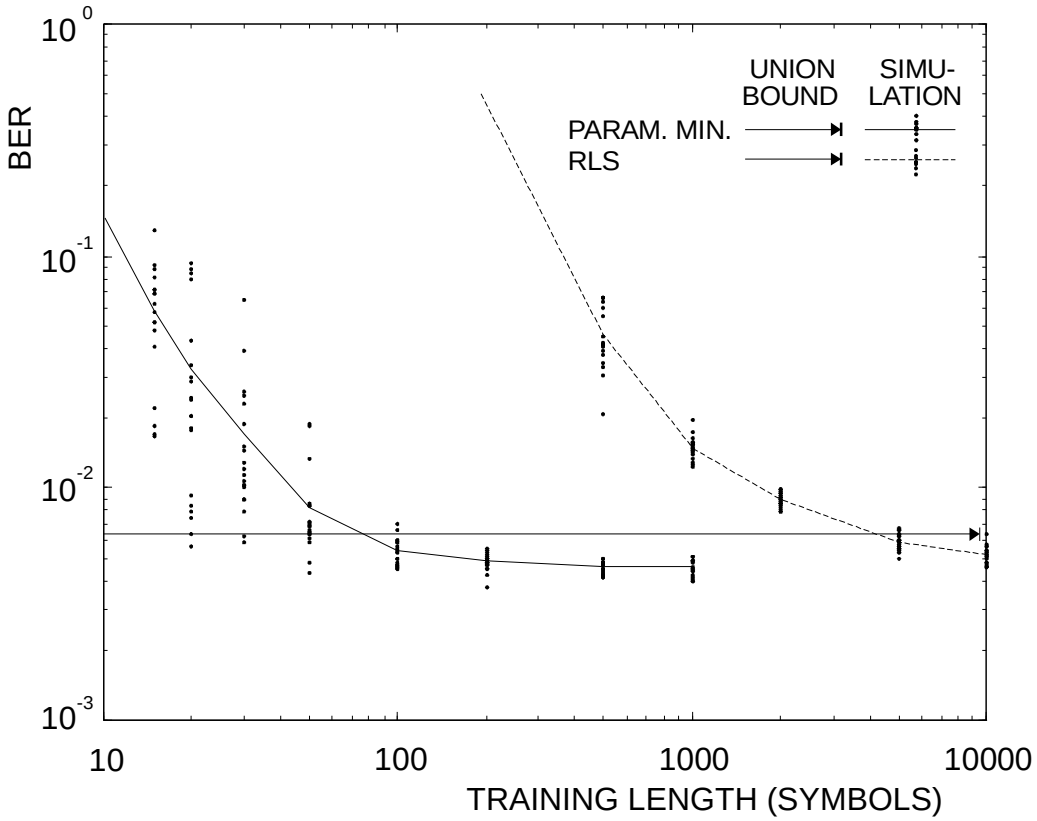

Figure 10: Acquisition performance of the two signal autocovariance estimation schemes. The important simulation parameters are: 2-DPSK; $f_D T = 0.1$; $\tau/T = 0.5$; $K = 0$; $t_0 = 0.2T$; $f_0 T = 0.2$; $r = 3$; $B = 2r$.