An Adaptive MLSD Receiver Employing Noise Correlation

Ting Liu and Saeed Gazor

Abstract

A per-survivor processing (PSP) maximum likelihood sequence detection (MLSD) receiver is developed for a fast time-varying frequency selective Rayleigh fading channel with colored additive noise which follows an auto-regressive (AR) model with unknown parameters. The correlation between noise samples is exploited to considerably enhance the performance of the communications. The maximum likelihood criterion is employed based on unknown noise parameters. This criterion has some desired properties, e.g., it has a unique joint minimum at the true values of the channel and the noise parameters. The new PSP-MLSD algorithm detects the input data and jointly estimates the noise and the channel parameters all together. The proposed structure can be viewed as a traditional PSP-MLSD receiver combined with an adaptive whitening filter. In a colored noise environment, this scheme offers faster tracking property, more accurate estimation of the channel and a substantially lower error probability compared with the traditional PSP-MLSD structure. The Signal-to-Noise-Ratio (SNR) improvement achieved by the proposed receiver, which can be called the Noise Whitening Gain (NWG), is almost equal to the ratio of the energy of the additive noise to the energy of the unpredictable noise component. The squared of the NWG gives also an accurate approximation for the BER improvement ratio obtained by using the proposed algorithm compared with traditional one.

I. INTRODUCTION

It is well known that MLSD is an optimal detection strategy for digital signals transmitted over a dispersive and noisy channel in the sense of minimizing the sequence error probability [1]–[7]. The receiver structure for performing MLSD of a digital signal corrupted by inter-symbol interference (ISI) and white additive Gaussian noise is well known [12], [13]. It may
be divided into two distinct components: a noise-limiting filter and a nonlinear post-processor based on the Viterbi algorithm that searches for the path with minimum cost in the combined ISI and code trellis diagram. This MLSD strategy requires an accurate channel estimate for good performance. When communicating with fast moving mobile terminals in a multipath channel, the receiver observes multiple delayed and Doppler-spread replicas of the signal. In the time-domain, this Doppler spread makes the channel time-varying. The PSP-MLSD, which performs the joint channel parameter estimation and the input data detection, has been introduced to solve this problem, and its tracking schemes have attracted attention in recent years [4]–[8]. With PSP, the estimates of channel parameters are updated independently for each survival trellis path and the decision feedback is retrieved from each individual trellis path with no decision delay. The PSP estimator is very suitable for fast time-varying channels.

In most of signal processing and communications literature, the additive noise is assumed to be white, which is not always true in practice, e.g., in magnetic recording. In such a case, the traditional detection algorithms result in a high error probability. In narrowband communication systems it is reasonable to assume that the additive noise is white. However, the noise is colored in some applications because of the nature of the physical medium such as in magnetic recording, wired communication systems; digital subscriber line (xDSL), power line and cable communication systems. In some other applications the additive noise is colored because of the way the signals are transmitted or received. For instance, in new Ultra-Wide Band (UWB) communication systems [20], the transmitted signal has very wide bandwidth and the additive noise consists of many strong narrow-band interferences and some background noise plus interference from other UWB users. In the communication systems employing multiple receiving antenna or oversampling in the receiver, the noise components are also correlated.

The noise predictive maximum likelihood method is used in the digital magnetic recording by imbedding a noise prediction process into the branch metric computation of the Viterbi detector [16]–[19]. In this method, the channel and the coefficients of the prediction filter are supposed to be known or well estimated during a training period. A simple approach is to first whiten the noise component of the received signal using a whitening filter and then process the filtered signal. We note that most of these whitening methods result in increase in the ISI. Consequently, their complexity is considerably increased by the increase of the number of states to be considered in the MLSD algorithm. The performance improvement using these approach is substantial if
the noise is highly correlated. In the following we will introduce a new low-complexity adaptive algorithm for the case that the noise and the channel models are unknown.

An adaptive whitening approach is discussed in [9], [10], for identification of a fast time-varying channel in a colored noise environment in the general context of adaptive filters. In this paper, this approach is employed in combination with the PSP-MLSD. Two adaptive filters are employed to track the channel and the noise parameters. The adaptive whitening filter is applied jointly to the received and the hypothesized data. The PSP-MLSD will operate based on the whitened signals to jointly detect the input data and estimate the noise and channel parameters to improve the performance. The proposed optimal approach attempts to eliminate the predictable component of the additive noise; therefore considerable improvement is achieved in colored noise environments.

The remainder of the paper is organized as follows. Section II briefly states the structure of the communications system we are dealing with. The new optimal receiver structure with whitening filter is proposed in Section III, and Section IV discusses the adaptive procedures of the estimation of the noise and channel parameters. Some simulations are given in Section V to support our new algorithm. In Section VI, the conclusions of the paper are drawn.

II. THE SIGNAL AND THE NOISE MODELS

The block diagram of the communication system is shown in Figure 1 [3]. The information symbols \( a_k \) are generated by a \( p \)-ary data source with a rate of \( \frac{1}{T} \). After pulse-shaping, the signal \( \sum_{q=1}^{\infty} a_q p(t - qT) \) is transmitted to the channel with transfer function of \( z(t; \tau) \). The received signal \( r(t) = s(t) + n(t) \), is filtered by a noise limiting filter with transfer function of \( g(t) \). The filtered signal \( \bar{r}(t) = r(t) \otimes g(t) = \bar{s}(t) + \bar{n}(t) \), which is sampled at a rate of \( \frac{1}{T_s} \geq \frac{1}{T} \), can be written as [3]:

\[
\bar{r}_l = \bar{r}(lT_s - t_0) = \sum_{q=1}^{\infty} a_q c(lT_s - t_0; lT_s - t_0 - qT) + n_l,
\]

(1)

where \( c(t; \tau) \overset{\Delta}{=} \int_{-\infty}^{\infty} g(\tau - u)z(t; u)du \) represents a time-varying Channel Impulse Response (CIR), and \( n_l = \bar{n}(lT_s - t_0) \) is a zero-mean complex stationary Gaussian random process with variance \( N_0 \). It is not necessary that \( n_l \) is a white noise process. \( t_0 \) is the channel transmission delay. Assuming that the oversampling ratio \( \frac{T}{T_s} = \beta \) is an integer number, the received sequence...
can be written in a folded vector form as:

\[ r_l = [r_l, r_{l-1}, \cdots, r_{l-\beta+1}]^T, \]

where \( r_{l-1}, \cdots, r_{l-\beta+1} \) are new samples received in the time interval of two successive transmitted symbols. The received signal can be written as,

\[ r_{\beta k} = \sum_{q=1}^{\infty} a_q w_{k,k-q}^* + n_{\beta k}, \]

where \( w_{k,p} \) represents the time-varying baseband CIR viewed by the detector that could be expressed in terms of \( c(t; \tau) \) and the noise-limiting filter. The received noise vector is denoted by

\[ n_k = [n_k, n_{k-1}, \cdots, n_{k-\beta+1}]^T. \]

We assume that the support of the CIR, \( c(t; \tau) \), is less than \( T_D \), i.e., \( c(t; \tau) = 0 \) for all \( t - \tau > T_D \), and define,

\[ W_{o,k} = [w_{k,0}, w_{k,1}, \cdots, w_{k,d}]^T, \]

\[ a_k = [a_k, a_{k-1}, \cdots, a_{k-d}]^T, \]

where \( d = \lceil \frac{T_D}{P} \rceil \). Based on this assumption, each \( r_{\beta k} \) only depends on \( d \) previous inputs and (3) can be rewritten as:

\[ r_{\beta k} = W_{o,k}^H a_k + n_{\beta k}. \]

where \((.)^H\) stands for the Hermitian (transpose conjugate) operator. We note that the CIR matrix, \( W_{o,k} \in \mathbb{C}^{d \times \beta} \), is a time-varying random process with correlated elements. The elements of \( W_{o,k} \) are correlated because of two reasons: the first is the signatures of the noise limiting filter and the pulse shaping filter. The second reason is that the channel \( z(t; \tau) \) is supposed to be a Rayleigh fading channel, which means that \( z(t; \tau) \) is a wide sense stationary Gaussian random process with a particular lowpass power spectrum. In this way \( W_{o,k} \) is a time-varying stationary Gaussian random process. The focus of this paper is not how to exploit the spectrum properties of \( W_{o,k} \) to enhance the performance, but to exploit the spectrum properties of the noise.

This paper considers the noise process to be colored with an unknown spectrum and is assumed to follow an autoregressive random process of degree \( L \), \( \text{AR}(L) \), as follows,

\[ n_l = \sum_{q=1}^{L} \theta_{a,q}^* n_{l-q} + v_l = \Theta_a^H n_{L,l} + v_l \]
where $\Theta_o = [\theta_{o,1}, \ldots, \theta_{o,L}]^T$, $n_{L,l} = [n_l, n_{l-1}, \ldots, n_{l-L+1}]^T$, and $\sigma^2 = E[|v_l|^2]$. The vector $\Theta_o$ is an unknown and slowly time-varying noise modeling parameter and $v_l$ is assumed to be a white zero mean Gaussian random process. The subscript “$o$” indicates the true value of the parameter.

One advantage of the AR model for the noise (as opposed to other type of models) is that this model leads to a well behaving optimization criterion. By using this model, we will see that the detection-estimation criterion is a quadratic function of the unknown noise parameters. Therefore, we have a unique global optima. On the other hand, if using other models, e.g., auto-regressive moving average (ARMA), the detection-estimation criterion becomes a multi-modal function of unknown parameters and the model is not stable for all values of the parameters. Therefore, the ARMA model results in very complex optimization problems and expensive adaptive schemes.

Another advantage of the AR model is to deal with the narrow band interferences which are common in communication systems. Such interferences produce unknown peaks in the noise power spectrum. The AR model is known to be capable of efficiently modeling such narrow band components and substantially reduce their impact.

In this paper, the power spectrum properties of the noise are exploited to enhance the performance of the receiver. Using the proposed architecture, the performance of the receiver is anticipated to be improved by an equivalent SNR of $\frac{E[|n_l|^2]}{E[|v_l|^2]}$ in dB. That could be quite considerable.

### III. Metric for AR Colored Noise

In order to perform MLSD, the log–likelihood function must be evaluated for each surviving sequence hypothesis. If we assume that $W_{o,k}$ and $\Theta_o$ are known and $n_l$ is a white Gaussian random process, the optimal metric (referred to as a path metric) is given by [1]–[5]:

$$\Lambda(s) = \frac{\|r_l - \hat{r}_{l-1}\|^2}{\hat{\varsigma}_{l-1}^2} + \beta \ln \varsigma_{l-1}^2, \quad \forall l = k; \beta,$$

where $(\cdot)_{[p]}$ denotes the estimation of $(\cdot)_l$ using the observations up to time $p$. In this way, $\hat{r}_{l-1}(s)$ is the Minimum–Mean–Square–Error (MMSE) prediction of $r_l$ given the past received samples and a hypothesized survivor data sequence labelled by $s$. $\hat{\varsigma}_{l-1}^2(s)$ is the prediction covariance [11]. For simplicity we drop off the state index $s$ from some of the notations. The path metric must be evaluated for each surviving sequence hypothesis to perform MLSD [11]. In contrast, if we assume that the noise is an AR($L$) process as in (8), where $v_l$ is a zero mean white Gaussian...
random process, the optimal path metric is given by:

$$\Gamma(k) = \frac{1}{\sigma^2_{l-1}} \sum_{i=\beta k - \beta + 1}^{\beta k} |v_{i|l-1}|^2 + \beta \ln \sigma^2_{l-1}, \; \forall l = k\beta,$$  \hspace{1cm} (10a)

where $\sigma^2_{l|l}$ is the variance of the whitened noise $v_{i|l}$. For all $i = \beta k - \beta + 1, \ldots, \beta k$ and $l = k\beta$, we have

$$r_{\beta k|l-1} = W^H_k \hat{a}_k,$$  \hspace{1cm} (10b)

$$r_{i|l-1} = [r_{\beta k|l-1}]_{i \mod \beta},$$  \hspace{1cm} (10c)

$$n_{i|l-1} = r_i - r_{i|l-1},$$  \hspace{1cm} (10d)

$$v_{i|l-1} = n_{i|l-1} - \sum_{q=1}^{L} \theta^*_q \hat{n}_{i-q|l-1},$$  \hspace{1cm} (10e)

where $\theta_{q,k}$ and $W_k$ are the estimates of $\theta_{o,q}$ and $W_{o,k}$ at time $l$, respectively, and $\hat{a}_k$ is a hypothesis input vector which is also referred to as the survivor sequence at each trellis state. The likelihood function in (10) is easily obtained by the use of the probability density of the observed signal $r_{\beta k}$ in terms of the iid random variables $\{v_{l}\}_{l=\beta k}^{\beta k + \beta - 1}$. Note that in derivation of (10), we assume the parameters $\{\Theta, W, \sigma^2\}$ are almost constant for a duration of $\beta + L + d$ samples. The MLSD algorithm could be implemented using (10) provided the values of $\Theta, W$ and $\sigma^2$.

In (9) and (10a), the values of $r_{l} - r_{i|l-1}$ and $v_{i|l-1}$ are both the MMSE predictions errors, assuming two different noise models, a white noise model and AR($L$) model, respectively. Using $v$ instead of $n$ in the path metric (10a) removes the redundancy among the samples of the noise. It is like applying an adaptive pre-whitening filter to both the hypothesized input and the received signals before the detection and the adaptation processes [9], [10]. The structure is depicted in Figure 2, where the linear filter $[1, -\Theta_k^T]^{T}$ is the pre-whitening filter. The whitened output $v_{i|l-1}$, which is the estimate of the white noise component, is used for the input signal detection and the filter adaptations. Different from this paper, in [9], [10] the input signal is known and can be colored, and the adaptive pre-filter perform a joint whitening of the input and the noise. Whereas in this paper, the input signal in the algorithm is white and unknown. The adaptive algorithm adapts the estimates of the channel and noise parameters by minimizing the mean square error.

Here, the main objective of the algorithm is detecting the input sequence whereas in [9], [10] the objective is different. In this paper based on a PSP structure, we use the cost in (10a) to adaptively estimate the unknown channel parameters $W_{o,k}$, the noise parameters $g_{o,k}$, as well as
the whitened noise variance. Note that the sampling rate is \( \beta \) times the data rate. We adopt a block processing approach by converting the received signals in vectors of size \( \beta \). In this way the algorithm is executed once for each transmitted symbol.

The detector expressed by (10) requires accurate estimates of the channel response \( W_{o,k} \), the vector of noise parameters \( \Theta_o \), and the variance of the white noise component \( \sigma^2 \). In practice, the random process \( W_{o,k} \) can be assumed to be band-limited, Gaussian and zero mean. For example, if we consider a Rayleigh fading channel with a maximum Doppler frequency \( \omega_D \) in rad/sample, the bandwidth of the CIR is \( f_D \) with auto-correlation matrix that is proportional to \( J_0(\omega_D \tau) \), where \( J_0 \) is the zero-order Bessel function of the first kind [14], [15]. In this case, a recursive state–space estimator can be used (e.g., A minimum–variance unbiased estimate results in a Kalman Filter (KF) or a differential KF, see [3]). However, this is only possible when the additive noise is assumed to be white Gaussian. To obtain an efficient implementation, we suggest to use the very simple recursive PSP algorithm to jointly estimate the required parameters.

### IV. Joint Adaptive Estimation

In order to perform the MLSD, the parameters \( W_{o,k} \) as well as the noise parameters, \( \Theta_o \) and \( \sigma^2 \), are to be estimated recursively for each branch of the Viterbi algorithm, and updated for each state. The parameters are updated based on the state likelihood. The state likelihood is calculated by the branch metrics based on different hypothesis for the input sequence.

In detection theory, when some parameters are unknown (referred to composite hypothesis testing problem) the most popular approach is the generalized likelihood ratio test [21]. In this approach under each hypothesis (here each state) the likelihood function is maximized in terms of unknown parameters in order to obtain an estimate of unknown parameters. Then the resulting likelihood functions are used for performing the hypothesis test. In this paper, we suggest to adaptively maximize the log-likelihood functions under each state, i.e., under each state the cost function in (10) is minimized in order to estimate the unknown parameters (this is to maximize the likelihood functions). Using the gradient based algorithm [10], we propose to use the following recursion for the channel estimation

\[
\begin{align*}
W_k &= W_{k-1} - \mu_W \nabla W_{k-1} \Gamma(k), \\
\Theta_k &= \Theta_{k-1} - \mu_\theta \nabla \Theta_{k-1} \Gamma(k).
\end{align*}
\]
In detail, without loss of generality, the update equations for the case that $\beta = 1$ are:

$$
\begin{align*}
W_k &= W_{k-1} + \frac{\mu_w}{\sigma_{k|k-1}^2 + \epsilon} v^*_k|k-1 (\hat{a}_k - A_{k-1}^H \Theta_{k-1}), \\
\Theta_k &= \Theta_{k-1} + \frac{\mu_\theta}{\sigma_{k|k-1}^2 + \epsilon} v^*_k|k-1 (r_k - R_{k-1}^H W_{k-1}),
\end{align*}
$$

(12)

where $\mu_w$ and $\mu_\theta$ are the step sizes for the adaptations of $W$ and $\Theta$, respectively. $A_{k-1} = [\hat{a}_{k-1} \hat{a}_{k-2} \cdots \hat{a}_{k-L}]^T$ and $R_{k-1} = [r_{k-1} r_{k-1} \cdots r_{k-d}]^T$, where $\hat{a}_{k-1} = [\hat{a}_{k-1} \hat{a}_{k-2} \cdots \hat{a}_{k-d}]^T$ is the survivor sequence for each state in the Viterbi algorithm, and $r_k = [r_k r_{k-1} \cdots r_{k-L+1}]^T$.

The superscript $(.)^*$ stands for the conjugate operator. The small positive constant number $\epsilon$ is introduced to avoid numerical instability. We use $\epsilon = 1$ in our simulations. Since the noise follows the AR($L$) model, we have

$$
v_{k|k-1} = (r_k - \Theta_{k-1}^H r_{k-1}) - W_{k-1}^H (\hat{a}_k - \Theta_{k-1}^H \hat{a}_{k-1}).
$$

(13)

The variance of the whitened noise can be adapted using

$$
\sigma^2_{k|k-1} = \alpha \sigma^2_{k-1|k-2} + (1 - \alpha) |v_{k|k-1}|^2,
$$

where $\alpha$ is a positive constant smaller than but close to 1. The following theorem ensures the existence and uniqueness of the optimum solution. The theorem also shows that the proposed adaptive filter has a similar behavior as expected in the literature of adaptive filter theory [9], [10].

**Theorem 1 ([9], [10]):** The cost function

$$
J = J(W, \Theta) = E[\Gamma(k)|W, \Theta, \sigma^2]],
$$

(14)

has the following properties

- The cost function $J$ for any given value of $\Theta \neq 0$ is a quadratic function in terms of $W$. Furthermore, it has only one unique minimum at the true value of $W$.
- The cost function $J(W, \Theta)$ for any given value of $W$ is a quadratic function of $\Theta$ and has a unique minimum. However, this minimum depends on the value of $W$.
- The cost function $J$ has a unique joint minimum at true values of $W$ and $\Theta = \Theta_o$.
- Except for considering the variance of the innovation noise process this cost function is the mean squared error estimate of the noise source $v_l$ as a function of unknown parameters.

This theorem shows that the proposed adaptive algorithm is able to track the true values for small step-sizes. Note that in this PSP-MLSD algorithm, the use of noise whitening filter does
not change the total number of states. The computational complexity is barely increased because of the extra use of adaptive noise whitening filter. Suppose the number of states is \( M = 2^d \), where \( d \) is the length of the channel. The length of the noise prediction filter \( \Theta \) is \( L \), then for each step, approximately \((d + 2)L\) extra multiplications are required.

The analysis of MLSD when it is combined with adaptive channel estimator (even for the simple case of white noise) is very difficult and is still an open research problem. However, the following discussion gives a rough and asymptotic criterion for the evaluation of the proposed method. It is obvious that not only the additive noise contributes to occupancy of error in the sequence detection but also the channel parameter estimation errors can increase the impact of the noise. If the SNR is high and the channel estimation error is low, from asymptotical analysis of LMS-type adaptive algorithms, the prediction error variance \( E[|υ|_{l-1}|^2] \) is given by \( \sigma^2(1 + M) \), where \( M \) is known as the misadjustment of the channel estimation. This means that the SNR loss of at least \( 10 \log_{10}(1 + M) \) is expected compared with the case of fully known channel. If the proposed whitening process is not performed the effective noise variance will be \( E[|n_l|^2](1 + M) \). We must also note that the MLSD is derived based on the assumption that the successive noise samples are independent. In the case of colored noise process, this assumption makes the traditional MLSD to lose its optimality, i.e., further significant performance loss is expected. Accurate performance analysis and comparison of these algorithms are very difficult (even for the case of fully known channel parameters).

V. SIMULATIONS

The efficiency of the proposed receiver is illustrated by simulations under the following conditions:

- The data source \( a_k \) is generated as a binary independent and identically distributed sequence with equal probability, i.e., \( P(a_k = ±1) = \frac{1}{2} \).
- The fading channel \( W_k \) has 2 independent taps. We used a Rayleigh fading channel following Jake’s model with a Doppler shift frequency of 250Hz and a sampling frequency of 25kHz [15], i.e., \( W_k \) is a complex vector of independent zero mean Gaussian random processes that are generated such that the temporal correlation is given by \( E[W_{k+m}W_H^m] = IJ_0(0.2\pi k) \), where \( I \) is the identity matrix. This represents a relatively fast fading situation, i.e., the ratio of Doppler frequency to the sampling rate is 10%.
• In our simulations, we make a critical sampling, i.e., the oversampling ratio is $\beta = 1$ in (2).
• The vector $\hat{a}_k$ for each surviving processor is formed from surviving memory for each given trellis state. After maximization of the criterion in (14), $\hat{a}_k$ is selected (for each surviving path) and is shifted into the surviving memory to be used in the next time instances.
• According to (8), the noise is generated by an AR model. Specifically in our simulation the noise is generated by $n_k = v_k + \theta n_{k-1}$, where $v_k$ is generated as a white complex Gaussian random process, $\theta$ is a fixed number such that $|\theta| < 1$. In this case, we have $E[|n_l|^2] = E[|v_l|^2] + |\theta|^2 E[|n_l|^2]$ and $n_k = v_k + \theta n_{k-1}$, i.e., the attainable noise variance whitening gain is given by $\frac{E[|n_l|^2]}{E[|v_l|^2]} = \frac{1}{1-|\theta|^2}$. We initialize $n_0$ as a zero mean complex Gaussian random process with variance of $E[|n_l|^2]/E[|v_l|^2]$ in order to have a stationary noise process $n_k$.

The simulated bit error rate (BER) performance for the traditional PSP-MLSD and our proposed algorithm with the whitening approach are given in Figure 3 for $\theta = 0.95$ and 0.99. All the simulations are ensemble averages over 50 independent runs. It’s easy to see that our proposed approach has significant improvement over the traditional one when the noise is colored. Since the additive noise is whitened in the proposed approach, the performance is improved by an equivalent of $E[|n_l|^2]/E[|v_l|^2]$ dB in SNR, which is called the Noise Whitening Gain (NWG), $NGB = E[|n_l|^2]/E[|v_l|^2] = \frac{1}{1-|\theta|^2}$. In Figure 3, compared with the solid curves, the corresponding dashed curves are almost 10dB and 17dB lower, which are the values of $10 \log_{10} NGB = -10 \log_{10}(1 - |\theta|^2)$ in dB when $\theta = 0.95$ and 0.99, respectively.

Figure 4 compares the MSE of the two approaches for SNR= 10 and 30 with $\theta = 0.95$. With almost the same convergence speed, the use of the whitening filter brings down the MSE around 10dB to $-10 \log_{10}(1 - |\theta|^2)$ for both SNR= 10dB and 30dB. The MSE curves for different $\theta$ are plotted in Figure 5 for SNR= 20. The dotted curve with highest MSE is using traditional PSP-MLSD algorithm for $\theta = 0.9$. The middle solid one is for $\theta = 0.9$ and using the proposed approach, which is almost 7dB lower than using the traditional PSP-MLSD algorithm. For the case $\theta = 0.99$, the corresponding MSE using the proposed approach is plotted in the dashed line, which is almost 10dB lower than the middle curve. In this case, only three extra computations (3 multiplications and 3 additions) are required for each state compared with traditional PSP-MLSD algorithm. The SNR improvement in our simulations is approximately given by the NWG.

Figure 6 illustrates the simulated BER behaviors of the two algorithms versus the value of $\theta$, that is, the level of the “correlation” of the additive noise. The bigger the $\theta$ is, that is, the more
colored the noise is, the higher BER we get by using the traditional PSP-MLSD algorithm, since \( \Lambda(s) \) in (9) is not an optimal criterion for the colored noise. The performance of our proposed algorithm is improved by eliminating the predictable components of the additive noise. Therefore, the amount of performance improvement increases with the increase of noise correlation. By using a simple curve fitting in logarithmic scale, we found out that the BER of the proposed algorithm can be approximated by \( (1 - |\theta_o|^2)^2 \times 6 \times 10^{-4} \), in terms of the noise whitening gain, \( \text{NWG}=(1 - |\theta_o|^2) \), and the BER of the traditional algorithm that is \( 6 \times 10^{-4} \). This curve is plotted in dashdot in Figure 6. It shows that the BER of the proposed algorithm is in the order of the squared of the NWG. This also means that the BER of the proposed algorithm is approximately equal to the BER of the traditional algorithm multiplied by the squared of the NWG, i.e.,

\[
\text{BER}_{\text{proposed}} \simeq \text{BER}_{\text{traditional}} \times (\text{NWG})^2.
\]

Therefore, the more colored the noise is, the performance improves more.

VI. CONCLUSIONS

A new PSP-MLSD receiver is introduced for a fast frequency selective Rayleigh fading channel with colored additive AR noise. Besides estimating the channel parameters, the new algorithm whitens the noise as well. The optimal path metric detection criterion is derived based on the noise model. In addition, minimizing this criterion gives the true values of the unknown parameters. When the noise is colored, the performance of the traditional PSP-MLSD algorithm degrades because \( \Lambda(s) \) in (9) is not an optimal criterion. Furthermore, the colored noise causes higher misadjustment in the estimation of the channel. On the contrary, using the proposed approach, the correlation between the samples of the noise in the proposed algorithm is adaptively exploited, which makes the channel estimation more accurate, and reduces the error probability. Simulation shows the proposed algorithm outperforms the traditional PSP-MLSD method with fast time-varying channel and colored additive noise. The ratio of improvement in BER by using the proposed algorithm compared with the traditional algorithm is approximated by the square of NWG.

REFERENCES


Fig. 1. The baseband communication system in the presence of AR(L) colored noise.

Fig. 2. The structure of the adaptive pre-whitening filter corresponding to a survivor path.

Fig. 3. BER performance of the traditional PSP-MLSD algorithm (solid) and our proposed approach (dashed), diamond is for the case that $\theta = 0.99$ and circle is for $\theta = 0.95$. 
Fig. 4. MSE of the traditional and the proposed algorithms in the whitened domain for SNR = 10 and 30.

Fig. 5. MSE of the traditional PSP-MLSD for $\theta = 0.9$ (dotted) and the proposed approach in the whitened domain for $\theta = 0.9$ (solid) and $\theta = 0.99$ (dashed).
Fig. 6. BER performance of the traditional PSP-MLSD algorithm (dashed), the proposed approach (solid) and the curve of 
\((1 - |\theta_o|^2)^2 \times 6 \times 10^{-4}\) obtained by curve fitting on the BER of the proposed algorithm.
**List of Figures**

1. The baseband communication system in the presence of AR($L$) colored noise. . . . 13
2. The structure of the adaptive pre-whitening filter corresponding to a survivor path. 13
3. BER performance of the traditional PSP-MLSD algorithm (solid) and our proposed approach (dashed), diamond is for the case that $\theta = 0.99$ and circle is for $\theta = 0.95$. 13
4. MSE of the traditional and the proposed algorithms in the whitened domain for SNR= 10 and 30. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .