Fractionally Spaced Coherent Detection with DF MAP Estimation of Statistically Known Time-Varying Rayleigh Fading Channel

Huaping Liu† Zoran Siveski† Nirwan Ansari

Department of Electrical and Computer Engineering
New Jersey Institute of Technology, Newark, NJ 07102 USA

Abstract

Coherent symbol by symbol detection using fractionally spaced samples in a time-varying frequency flat Rayleigh fading channel is presented. The proposed detector incorporates a decision feedback (DF) maximum a posteriori (MAP) channel estimator that can be applied to a time-varying flat Rayleigh fading channel with an arbitrary but known spaced-time correlation function. The best achievable bit error rate performance for this detector is obtained by assuming correct bit decisions while performing channel estimation. Differential encoding is used in order to prevent the phase reversal phenomenon during the estimator's operation in a decision directed mode. Numerical results suggest that with only a few samples per symbol and with a short memory estimator depth, most of the gain due to oversampling can be achieved.

1. Introduction

Radio channels in mobile communication systems exhibit time varying fading, the rapidity of which is quantified by the Doppler spread. Differentially coherent PSK schemes employed in such channels will–whenever the assumption of a nearly constant channel gain over two adjacent symbol intervals does not hold–exhibit an error floor [1]. If more power-efficient coherent PSK signaling is employed instead, channel estimation (i.e., carrier recovery) is needed. A systematic approach to carrier recovery is given in [2], while a minimum mean square error (MMSE) channel estimator with decision feedback (DF), in particular, is used in a number of references on coherent detection (e.g., [3], [4]). Perfect estimates, however, can be obtained only if the channel is noiseless and time invariant, whereas channel estimation errors in coherent detection will cause an error floor. For fading channels modeled by a discrete time autoregressive (AR) process, the Kalman filter is the optimum MMSE DF channel estimator [2]. In [5] it was shown that the error floor is the same for both coherent and differentially coherent detection when the Kalman filter is used for performing first-order AR fading channel model estimation for the former.

Fractionally spaced sampling (multi-sampling) receivers in time-varying fading channels have the potential to alleviate the error floor. Differentially coherent detection of PSK signals in a time-varying flat Rayleigh fading channel is presented in [6], in which K samples were obtained in each symbol interval. Maximum likelihood detection was then applied on the $2K$-sample snapshot over two consecutive symbol intervals, resulting in a significant reduction of the error floor. In [7] and [8], the same multi-sampling technique was applied for sequence estimation of a PSK signal, and significant performance improvement over the symbol spaced sampling approach was observed. Fractional sampling was also used in [9] for a time-varying dispersive statistically known channel for improved error performance.

This paper proposes a fractionally spaced coherent symbol by symbol detector with MAP channel estimator that can be implemented in a time-varying flat Rayleigh fading channel with an arbitrary but known spaced-time correlation function. A lower bound on the error probability is derived analytically and differential encoding is employed, resulting in a more robust coherent detector. Numerical examples for different fading rates and the detector's parameters are presented.

2. Detection

The received baseband equivalent complex signal $r(t)$ in a single path Rayleigh fading channel is expressed as

$$y(t) = \sum_i \sqrt{E} b(i) s(t - iT) c(t) + n_w(t), \qquad (1)$$

where E, $b(i) \in \{-1, 1\}$, $s(t)$, and $n_w(t)$ denote bit energy, the information bit, unit energy pulse of duration T, and an additive, zero-mean, complex white Gaussian noise with the one-sided power spectral density N_0, respectively. The fading channel process $c(t)$ is modeled as a normalized, zero-mean complex-valued wide-sense stationary Gaussian process.

† Now with Lucent Technologies, Whippany, NJ 07981.

We obtain K samples at the matched filter output at time instants $(i-1)T+\tau_1, \ldots, (i-1)T+\tau_K$ over the i^{th} bit interval[1]. Such a K-sample snapshot of $y(t)$ is expressed in a vector form as

$$\begin{aligned} \boldsymbol{y}(i) &= [y^{(1)}(i), \ldots, y^{(K)}(i)]^T \\ &= \sqrt{E}\mathcal{E}\boldsymbol{C}(i)b(i) + \boldsymbol{\xi}(i), \end{aligned} \tag{2}$$

where

$$\begin{aligned} \mathcal{E} &= \text{diag}[\sqrt{\epsilon^{(1)}}, \ldots, \sqrt{\epsilon^{(K)}}], \\ \epsilon^{(j)} &= \int_{\tau_{j-1}}^{\tau_j} s^2(t)dt, \quad j = 1, \ldots, K, \\ \boldsymbol{C}(i) &= [c^{(1)}(i), \ldots, c^{(K)}(i)]^T \\ &= [c(t)|_{t=(i-1)T+\tau_1}, \ldots, c(t)|_{t=(i-1)T+\tau_K}]^T, \\ \boldsymbol{\xi}(i) &= [\xi^{(1)}(i), \ldots, \xi^{(K)}(i)]^T. \end{aligned}$$

In (2) we adopted a discrete-time fading channel formulation in deriving the detector structure. Specifically, the channel is assumed to be constant during one sample interval, but it changes from one sample to another. The covariance matrix of $\boldsymbol{\xi}$ is

$$\boldsymbol{R}_{\xi(i)} = E\{\boldsymbol{\xi}(i)\boldsymbol{\xi}(i)^H\} = N_0\boldsymbol{I}, \tag{3}$$

where $\boldsymbol{I}$ is a $K \times K$ identity matrix. Given $\widehat{\boldsymbol{C}}(i)$, the estimate of the channel vector $\boldsymbol{C}(i)$, the decision variable for $b(i)$ can be obtained by combining the K elements in (2) according to the maximal ratio combining (MRC) rule[2] as

$$d(i) = Re\{\widehat{\boldsymbol{C}}(i)^H\boldsymbol{y}(i)\}. \tag{4}$$

Hence, to coherently detect the information bit $b(i)$ from the received vector in (2), we need to provide a procedure for channel estimation.

3. Channel Estimation

Define $MK \times 1$ vector

$$\mathcal{Y}_{i-1} = [\widetilde{\boldsymbol{y}}(i-M)^T \ \widetilde{\boldsymbol{y}}(i-M+1)^T \ \ldots \ \widetilde{\boldsymbol{y}}(i-1)^T]^T,$$

where $\widetilde{\boldsymbol{y}}(i-m) = \boldsymbol{y}(i-m)\widehat{b}(i-m)$, $m = 1, \ldots, M$, and $\widehat{b}(i-m)$ is the decision on the information bit $b(i-m)$. By defining the $(M+1)K \times 1$ vector

$$\mathcal{Y} = [\mathcal{Y}_{i-1}^T \ \boldsymbol{C}(i)^T]^T, \tag{5}$$

the covariance matrix of $\mathcal{Y}$ can be written as

$$\boldsymbol{R}_{\mathcal{Y}} = E\{\mathcal{Y}\mathcal{Y}^H\} = \begin{bmatrix} \boldsymbol{R}_{\mathcal{Y}11} & \boldsymbol{R}_{\mathcal{Y}12} \\ \boldsymbol{R}_{\mathcal{Y}21} & \boldsymbol{R}_{\mathcal{Y}22} \end{bmatrix}, \tag{6}$$

where

$$\begin{aligned} \boldsymbol{R}_{\mathcal{Y}11} &= E\{\mathcal{Y}_{i-1}\ \mathcal{Y}_{i-1}^H\}, \qquad (MK \times MK) & (7) \\ \boldsymbol{R}_{\mathcal{Y}12} &= E\{\mathcal{Y}_{i-1}\ \boldsymbol{w}(i)^H\}, \qquad (MK \times K) & (8) \\ \boldsymbol{R}_{\mathcal{Y}21} &= E\{\boldsymbol{w}(i)\ \mathcal{Y}_{i-1}^H\}, \qquad (K \times MK) & (9) \\ \boldsymbol{R}_{\mathcal{Y}22} &= E\{\boldsymbol{w}(i)\ \boldsymbol{w}(i)^H\}. \qquad (K \times K) & (10) \end{aligned}$$

The covariance matrices (7) - (9) depend not only on the channel spaced-time correlation function, but also on the previous bit decisions. From the practical point of view, for a reliable communication the decision error is usually small enough, so these matrices can be calculated by using channel correlation information only; that is, previous bit decisions can be assumed as being correct for computing $\boldsymbol{R}_{\mathcal{Y}11}$, $\boldsymbol{R}_{\mathcal{Y}12}$ and $\boldsymbol{R}_{\mathcal{Y}21}$, although the matrices depend on $\mathcal{Y}_{i-1}$, which is data-decision dependent. Given this approximation, the matrices are assumed decision independent and can be calculated *a priori*.

The MAP channel estimates of $\boldsymbol{C}(i)$ can be obtained by maximizing $p\{\boldsymbol{C}(i)|\mathcal{Y}_{i-1}\}$; i.e.,

$$\max_{\boldsymbol{C}(i)} \ p\{\boldsymbol{C}(i)|\mathcal{Y}_{i-1}\}. \tag{11}$$

The estimation procedure is illustrated in Fig. 1.

According to (5), because both $\mathcal{Y}_{i-1}$ and $\boldsymbol{C}(i)$ are complex Gaussian random vectors, so is $\mathcal{Y}$. The conditional probability density function in (11) can be written as

$$\begin{aligned} p\{\boldsymbol{C}(i)|\mathcal{Y}_{i-1}\} &= \frac{p\{\mathcal{Y}_{i-1}, \boldsymbol{C}(i)\}}{p\{\mathcal{Y}_{i-1}\}} \\ &= \frac{\frac{1}{\pi^{(M+1)K}|\boldsymbol{R}_{\mathcal{Y}}|}e^{-\mathcal{Y}^H\boldsymbol{R}_{\mathcal{Y}}^{-1}\mathcal{Y}}}{\frac{1}{\pi^{MK}|\boldsymbol{R}_{\mathcal{Y}11}|}e^{-\mathcal{Y}_{i-1}^H\boldsymbol{R}_{\mathcal{Y}11}^{-1}\mathcal{Y}_{i-1}}}. \end{aligned} \tag{12}$$

By using the partitioned matrix inversion from [11], $\boldsymbol{R}_{\mathcal{Y}}^{-1}$ can be expressed as

$$\boldsymbol{R}_{\mathcal{Y}}^{-1} = \begin{bmatrix} \boldsymbol{R}_{\mathcal{Y}11} & \boldsymbol{R}_{\mathcal{Y}12} \\ \boldsymbol{R}_{\mathcal{Y}21} & \boldsymbol{R}_{\mathcal{Y}22} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{R}_{\mathcal{Y}i11} & \boldsymbol{R}_{\mathcal{Y}i12} \\ \boldsymbol{R}_{\mathcal{Y}i21} & \boldsymbol{R}_{\mathcal{Y}i22} \end{bmatrix}, \tag{13}$$

where

$$\begin{aligned} \boldsymbol{R}_{\mathcal{Y}i11} &= (\boldsymbol{R}_{\mathcal{Y}11} - \boldsymbol{R}_{\mathcal{Y}12}\boldsymbol{R}_{\mathcal{Y}22}^{-1}\boldsymbol{R}_{\mathcal{Y}21})^{-1} \\ \boldsymbol{R}_{\mathcal{Y}i22} &= (\boldsymbol{R}_{\mathcal{Y}22} - \boldsymbol{R}_{\mathcal{Y}21}\boldsymbol{R}_{\mathcal{Y}11}^{-1}\boldsymbol{R}_{\mathcal{Y}12})^{-1} \\ \boldsymbol{R}_{\mathcal{Y}i12} &= -\boldsymbol{R}_{\mathcal{Y}i11}\boldsymbol{R}_{\mathcal{Y}12}\boldsymbol{R}_{\mathcal{Y}22}^{-1} \\ \boldsymbol{R}_{\mathcal{Y}i21} &= -\boldsymbol{R}_{\mathcal{Y}i22}\boldsymbol{R}_{\mathcal{Y}21}\boldsymbol{R}_{\mathcal{Y}11}^{-1}. \end{aligned}$$

Maximizing the conditional probability density function is equivalent to minimizing the following quadratic function:

$$l = \mathcal{Y}^H\boldsymbol{R}_{\mathcal{Y}}^{-1}\mathcal{Y} - \mathcal{Y}_{i-1}^H\boldsymbol{R}_{\mathcal{Y}11}^{-1}\mathcal{Y}_{i-1}. \tag{14}$$

Forcing the derivative of l with respect to $\boldsymbol{C}(i)$ to a zero column vector gives

$$\begin{aligned} \widehat{\boldsymbol{C}}(i) &= -\boldsymbol{R}_{\mathcal{Y}i22}^{-1}\boldsymbol{R}_{\mathcal{Y}i21}\mathcal{Y}_{i-1} \\ &= \boldsymbol{R}_{\mathcal{Y}21}\boldsymbol{R}_{\mathcal{Y}11}^{-1}\mathcal{Y}_{i-1}, \end{aligned} \tag{15}$$

[1] It is not necessary for the time instants to be uniformly distributed over the i^{th} bit interval, but it is required that $0 = \tau_0 < \tau_1 < \ldots < \tau_K = T$.

[2] Due to the assumption that the channel changes from one sample to another, the elements in (2) will not be fully correlated, and the MRC should provide performance improvement.

where, as mentioned before, $\widehat{\boldsymbol{C}}(i)$ denotes the estimate of $\boldsymbol{C}(i)$.

4. Error Performance

As pointed out in the previous section, we will derive the error probability lower bound for this detector by assuming that the previous bit decisions are correct, and, without loss of generality, that the transmitted bit is 1. The analysis is similar to that of the error performance of quadratic receivers in [13] and [14]. For simplicity, the time index i will be omitted whenever possible.

Introducing

$$\boldsymbol{v} = \begin{bmatrix} \widehat{\boldsymbol{C}} \\ \boldsymbol{y} \end{bmatrix},$$

the decision variable in (4) can be reformulated as

$$\begin{aligned} d(i) &= \begin{bmatrix} \widehat{\boldsymbol{C}} \\ \boldsymbol{y} \end{bmatrix}^H \begin{bmatrix} \mathbf{o} & 0.5\boldsymbol{I} \\ 0.5\boldsymbol{I} & \mathbf{o} \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{C}} \\ \boldsymbol{y} \end{bmatrix} \\ &= \boldsymbol{v}^H \boldsymbol{Q} \boldsymbol{v}, \end{aligned} \tag{16}$$

where $\mathbf{o}$ is a $K \times K$ matrix with all elements equal to zero. Defining matrix

$$\boldsymbol{V} = E\{\boldsymbol{v}\boldsymbol{v}^H\} = E\left\{ \begin{bmatrix} \widehat{\boldsymbol{C}} \\ \boldsymbol{y} \end{bmatrix} [\widehat{\boldsymbol{C}}^H \boldsymbol{y}^H] \right\},$$

from the results and definitions of the previous section, we have the following expressions:

$$\begin{aligned} E\{\widehat{\boldsymbol{C}}\widehat{\boldsymbol{C}}^H\} &= \boldsymbol{R}_{\mathcal{Y}21}\boldsymbol{R}_{\mathcal{Y}11}^{-1}\boldsymbol{R}_{\mathcal{Y}12}, \\ E\{\boldsymbol{y}\boldsymbol{y}^H\} &= a_1\boldsymbol{R}_{\mathcal{Y}22} + N_0\boldsymbol{I}, \\ E\{\widehat{\boldsymbol{C}}\boldsymbol{y}^H\} &= \sqrt{a_1}\boldsymbol{R}_{\mathcal{Y}21}\boldsymbol{R}_{\mathcal{Y}11}^{-1}\boldsymbol{R}_{\mathcal{Y}12}, \\ E\{\boldsymbol{y}\widehat{\boldsymbol{C}}^H\} &= \sqrt{a_1}\boldsymbol{R}_{\mathcal{Y}21}\boldsymbol{R}_{\mathcal{Y}11}^{-1}\boldsymbol{R}_{\mathcal{Y}12}. \end{aligned}$$

Therefore, matrix $\boldsymbol{V}$, defined above, can be expressed as

$$\begin{aligned} \boldsymbol{V} &= E\left\{ \begin{matrix} \widehat{\boldsymbol{C}}\widehat{\boldsymbol{C}}^H & \widehat{\boldsymbol{C}}\boldsymbol{y}^H \\ \boldsymbol{y}\widehat{\boldsymbol{C}}^H & \boldsymbol{y}\boldsymbol{y}^H \end{matrix} \right\} \\ &= \begin{bmatrix} \boldsymbol{R}_{\mathcal{Y}21}\boldsymbol{R}_{\mathcal{Y}11}^{-1}\boldsymbol{R}_{\mathcal{Y}12} & \sqrt{a_1}\boldsymbol{R}_{\mathcal{Y}21}\boldsymbol{R}_{\mathcal{Y}11}^{-1}\boldsymbol{R}_{\mathcal{Y}12} \\ \sqrt{a_1}\boldsymbol{R}_{\mathcal{Y}21}\boldsymbol{R}_{\mathcal{Y}11}^{-1}\boldsymbol{R}_{\mathcal{Y}12} & a_1\boldsymbol{R}_{\mathcal{Y}22} + N_0\boldsymbol{I} \end{bmatrix}, \end{aligned}$$

and matrix $\boldsymbol{V}\boldsymbol{Q}$ can be expressed as

$$\boldsymbol{V}\boldsymbol{Q} = 0.5 \begin{bmatrix} \sqrt{a_1}\boldsymbol{R}_{\mathcal{Y}21}\boldsymbol{R}_{\mathcal{Y}11}^{-1}\boldsymbol{R}_{\mathcal{Y}12} & \boldsymbol{R}_{\mathcal{Y}21}\boldsymbol{R}_{\mathcal{Y}11}^{-1}\boldsymbol{R}_{\mathcal{Y}12} \\ a_1\boldsymbol{R}_{\mathcal{Y}22} + N_0\boldsymbol{I} & \sqrt{a_1}\boldsymbol{R}_{\mathcal{Y}21}\boldsymbol{R}_{\mathcal{Y}11}^{-1}\boldsymbol{R}_{\mathcal{Y}12} \end{bmatrix}.$$

The probability of error of the quadratic decision variable is the one given in [13]:

$$P_e = P\{d(i) < 0\} = \sum_{\gamma_l < 0} \prod_{\substack{m=1 \\ m \neq l}}^{2K} \frac{\gamma_l}{\gamma_l - \gamma_m}, \tag{17}$$

in which $\gamma_1, \ldots, \gamma_{2K}$ are the eigenvalues of the matrix $\boldsymbol{V}\boldsymbol{Q}$.

5. Numerical Examples and Discussion

In the numerical examples, we consider the Jakes channel model [10]. Fading dynamics measured by the normalized Doppler frequency f_dT is the most important factor that affects the channel estimation error. The number of samples per bit K and the channel memory depth M are the parameters that affect the channel estimation accuracy. In the numerical examples, we will illustrate their effects on the error performance of the detector. The lower bound on the bit error rate can be evaluated by assuming correct previous decisions while performing channel estimation. The phase reversal phenomenon observed with the Kalman filter estimator in [12] could also occur in the proposed decision feedback MAP channel estimator. For the purpose of analytical calculation of the error performance lower bound, we assume that the previous decisions are correct in (15). In the actual channel estimation, differential encoding can be employed to mitigate the effects of phase reversals. Fig. 2 shows the curves obtained both by numerical evaluation of equation (17) and by simulation, under the assumption of correct previous bit decisions during channel estimation, and by simulation with differential encoding employed. Two samples per bit interval ($K = 2$), the memory depth $M = 4$, and normalized Doppler spread $f_dT = 0.04$ scenario was considered whereupon differential encoding demonstrated robust behavior. Results for a faster fading channel with $f_dT = 0.08$, the same number of samples per bit, and the same memory depth are shown in Fig. 3, with very small performance degradation. For comparison, curves for conventional symbol spaced differentially coherent detection performance in the same channel are shown in Figs. 2 and 3. The error floor for the conventional DPSK was observed while the proposed detector ($K = 2$) provided much better performance. When the number of samples per bit is increased, the error performance improves somewhat, as shown in Figs. 4 and 5, in which the fading rate is $f_dT = 0.04$ and $f_dT = 0.08$, respectively.

The detector's error performance with differential encoding is very close to the lower bound, which assumes that the previous decisions are correct during the channel estimation. Under the idealized assumption of perfect channel estimates in the fully interleaved channel, coherent detection with differential encoding is approximately 3 dB worse than that without differential encoding. However, in our more realistic examples, that difference is smaller due to clustering of errors. Also, even with 2 samples ($K = 2$) per bit the detector's performance is already very close to that of $K = 4$, suggesting that a large number of samples per bit is not necessary for reasonably good performance. For both $K = 2$ and $K = 4$, the error performance of the fractionally spaced sampling detector in the numerical examples considered was almost invariant to the fading rate, and no error floor was found for both fading rates in the SNR regions of interest.

References

[1] P. K. Kam, "Bit error probabilities of MDPSK over the nonselective Rayleigh fading channel with diversity reception," *IEEE Trans. on Communications*, vol. 39, pp. 220–224, Feb. 1991.

[2] R. Haeb and H. Meyr, "A systematic approach to carrier recovery and detection of digitally phase modulated signals on fading channels," *IEEE Trans. on Communications*, vol. 37, pp. 748–754, July 1989.

[3] P. K. Varshney and A. H. Haddad, "A receiver with memory for fading channels," *IEEE Trans. on Communications*, vol. COM-32, pp. 278–283, Feb. 1978.

[4] P. K. Kam and C. H. Teh, "Reception of PSK signals over fading channels via quadrature amplitude estimation," *IEEE Trans. on Communications*, vol. COM-31, pp. 1024–1027, Aug. 1983.

[5] Z. Zvonar and M. Stojanovic, "Performance of multiuser diversity reception in nonselective Rayleigh fading CDMA channels," *Proc. Third Communication Theory Mini-Conf. (CTMC'94)*, pp. 171–175, San Francisco, CA, Nov. 1994.

[6] W. Dam and D. Taylor, "An adaptive maximum likelihood receiver for correlated Rayleigh-fading channels," *IEEE Trans. on Communications*, vol. 42, pp. 2684–2692, Sep. 1994.

[7] G. M. Vitetta and D. P. Taylor, "Maximum likelihood decoding of uncoded and coded PSK signal sequences transmitted over Rayleigh flat-fading channels," *IEEE Trans. on Communications*, vol. 43, pp. 2750–2758, Nov. 1995.

[8] G. M. Vitetta and D. P. Taylor, "Multisampling receivers for uncoded and coded PSK signal sequences transmitted over Rayleigh frequency-flat fading channels," *IEEE Trans. on Communications*, vol. 44, pp. 130–133, Feb. 1996.

[9] D. W. Matolak and S. G. Wilson, "Detection for a statistically known, time-varying dispersive channel," *IEEE Trans. on Communications*, vol. 44, pp. 1673–1683, Dec. 1996.

[10] W. C. Jakes, *Microwave Mobile Communications*. New York, NY: John Wiley and Sons, Inc., 1974.

[11] S. M. Kay, *Modern Spectral Estimation*. Englewood Cliffs, NJ: Prentice Hall, 1988.

[12] H. Y. Wu and A. Duel-Hallen, "Multiuser detection with differentially encoded data for mismatched flat Rayleigh fading CDMA channels," *Proc. 1996 Conference on Information Sciences and Systems*, pp. 332–337, Princeton, NJ, Mar. 1996.

[13] M. Barrett, "Error probability for optimal and suboptimal quadratic receivers in rapid Rayleigh fading channels," *IEEE Journal on Selected Areas in Communications*, vol. SAC-5, pp. 302–304, Feb. 1987.

[14] M. Schwartz, W. R. Bennett, and S. Stein, *Communication Systems and Techniques, Chapter 10*. New York, NY: McGraw-Hill, 1966.

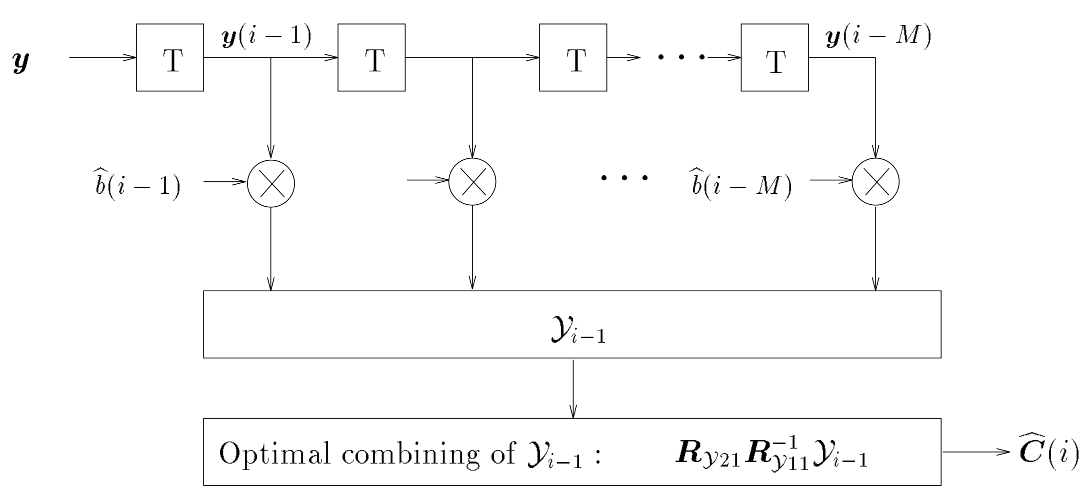

Figure 1: The single-stage DF MAP channel estimator

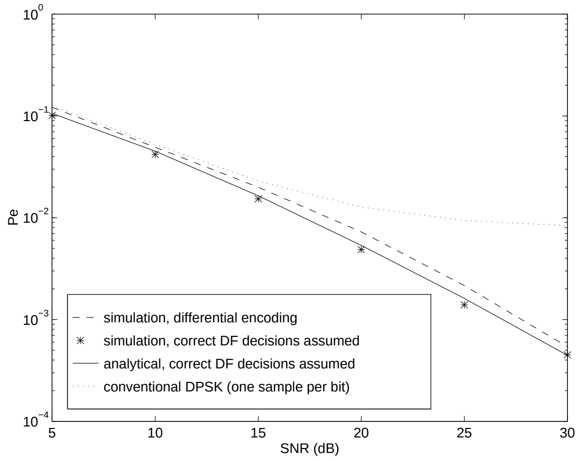

Figure 2: Error performance ($K = 2, M = 4,\ f_d T = 0.04$)

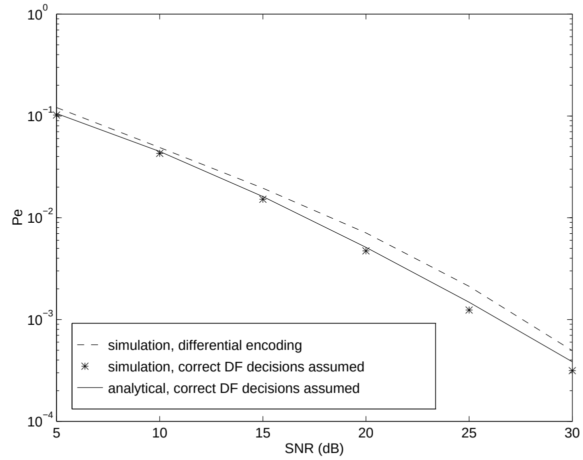

Figure 4: Error performance ($K = 4, M = 4,\ f_d T = 0.04$)

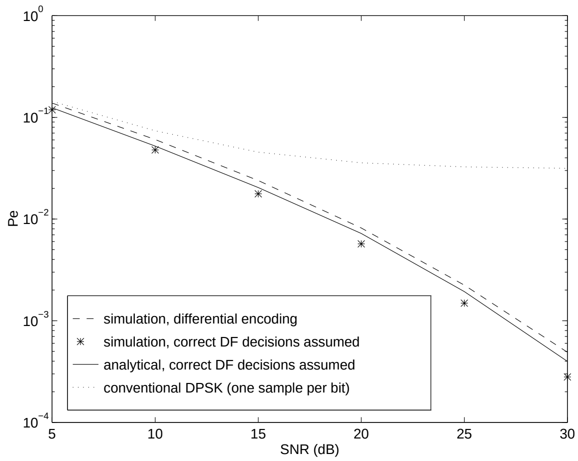

Figure 3: Error performance ($K = 2, M = 4,\ f_d T = 0.08$)

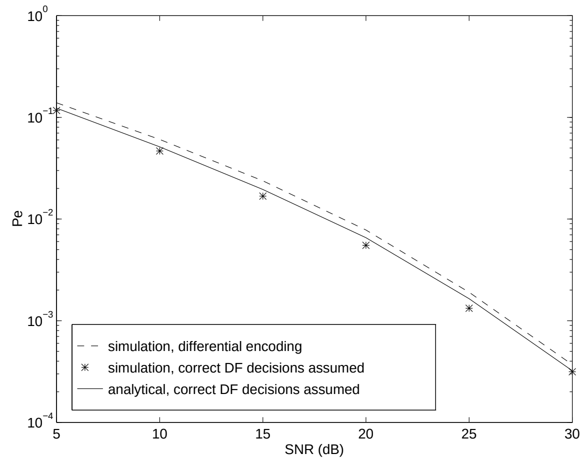

Figure 5: Error performance ($K = 4, M = 4,\ f_d T = 0.08$)