TURBO EQUALIZATION FOR GMSK SIGNALING OVER MULTIPATH CHANNELS

Xiaodong Wang and Zigang Yang

Texas A&M University, Department of Electrical Engineering,
College Station, TX 77843.

ABSTRACT

A novel receiver is derived for Gaussian minimum shift keying (GMSK) signals through a multipath channel. A nonlinear signal model is derived which avoids the linear approximation in the conventional finite impulse response(FIR) system model. A Bayesian equalizer based on the Gibbs sampler, a Markov chain Monte Carlo (MCMC) procedure, is developed for joint channel estimation and symbol detection, and finally, a Turbo equalizer structure is proposed for a coded GMSK system, in which the Bayesian equalizer successively refines its processing based on the information from the decoding stage, and vice versa.

1. INTRODUCTION

Gaussian minimum shift keying(GMSK) modulation is widely used in wireless communication systems due to its low side-lobe and constant modulus properties. A large variety of receiver structures have been proposed for GMSK systems. The conventional method is the Maximum Likelihood Sequence Estimation [1], which employs a finite impulse response(FIR) approximation(linearization) of the system, followed by the Viterbi algorithm. Such an approach is suboptimal due to the following reasons: first, the linear approximation incurs performance loss; second, the separation of channel estimation and data detection(as opposed to joint estimation of both channel and data) also results in performance loss.

In this paper, we propose a Bayesian approach to the problem of joint symbol detection and channel estimation for GMSK systems without linearization. At the transmitter, a precoding method is proposed to transform the system memory to a finite length. By using the band-limited property of the GMSK signal, a tapped-delay model is derived for the channel. The received signal is sampled at twice the symbol rate. With this oversampled nonlinear signal model, we consider the Bayesian inference of all unknown quantities (e. g. , channel and noise parameter, symbol values) from the nonlinearly distorted and noisy observation. A Markov Chain Monte Carlo procedure, called the Gibbs sampler, is employed to calculate the Bayesian estimation. The performance of the proposed Bayesian equalizer is demonstrated via simulations in a near blind way, i. e. , some training symbols are used to resolve the phase and timing ambiguity. Another salient feature of the proposed methods is that being a soft-input soft-output demodulation algorithms, it can be used in conjunction with soft channel decoding algorithm, to accomplish iterative joint equalization and decoding - so-called Turbo equalization.

THIS WORK WAS SUPPORTED BY THE U.S. NATIONAL SCIENCE FOUNDATION UNDER GRANT CAREER CCR–9875314.

2. SYSTEM DESCRIPTION

Assume that the GMSK signal $s(t)$ is transmitted through a multipath channel. The received signal can be written as

$$y(t) = \sum_{l=1}^{p} a_l s(t-\tau_l) + v(t), \tag{1}$$

where p is the number of path, a_l is the fading coefficient for the lth path; τ_l is the propagation delay of the lth path, and $v(t)$ is additive white Gaussian noise. The GMSK signal $s(t)$ can be represented as

$$s(t) = e^{\jmath\theta(t)}, \qquad \text{where } \jmath = \sqrt{-1} \tag{2}$$

$$\text{with} \quad \theta(t) = \sum_i b_i \phi(t-iT), \tag{3}$$

where $b_i \in \{-1, 1\}$ are transmitted data bits; T is symbol duration; the phase pulse-shaping function $\phi(t)$ is given by $\phi(t) = \frac{\pi}{2}\int_{-\infty}^{t} f(u)du$, and $f(t)$ is the GMSK frequency pulse which is characterized by BT(in GSM system, BT is chosen to be 0.3) and defined as

$$f(t) = \mathcal{Q}(\beta(\frac{t-T/2}{T})) - \mathcal{Q}(\beta(\frac{t+T/2}{T})), \tag{4}$$

where $\beta = 2\pi BT/\sqrt{\ln(2)}$.

2.1. Tapped-delay Model

Noticing that the path delays in (1) are difficult to estimate, we next derive a tapped-delay line model for (1). Assuming that signal $s(t)$ with band-width W, is transmitted through a channel $h(t)$, then the output of the channel can be written in the form of a tapped-delay line model [2] with tap interval $1/2W$.

As mentioned before, the GMSK signal has a small side-lobe, and thus provides good frequency efficiency. It has been shown in [3] that 99% of the energy of a GMSK signal with $BT \leq 0.5$ lies within $\pm 1/T$ of the center frequency. Thus, the received signal model becomes

$$y(t) \simeq \sum_{l=p_0}^{p_1} h_l s\left(t - \frac{lT}{2}\right) + v(t), \tag{5}$$

where p_0, p_1 denote the proper truncation from below and above. (The multipath channel will cause infinite number of nonzero tap coefficients, however, we can always truncate them into finite number of items with little loss of energy.)

The received signal is lowpass filtered with bandwidth $1/T$ and then sampled at the Nyquist rate $W = 2/T$. Assuming that the sampling time instants are $t_{k,j} \triangleq kT + j\frac{T}{2}$, then (5) becomes

$$y_{2k+j} = \sum_{l=p_0}^{p_1} h_l S_{2k+j}^{(l)} + v_{2k+j}, \tag{6}$$

where $y_{2k+j} \triangleq y(t_{k,j})$, $v_{2k+j} \triangleq v(t_{k,j})$, $S_{2k+j}^{(l)} \triangleq s(t_{k,j} - \frac{lT}{2})$. With the ideal lowpass filter, the noise samples v_{2k+j} are independent with each other. On the other hand, since over 99% of the GMSK signal energy will pass through the above lowpass filter, the sampled received signal is nearly the sufficient statistic for the transmitted data, and the maximum *a posteriori* probability (MAP) detector based on these sampled signals is a near-optimal receiver.

2.2. Sampled GMSK Signal and Precoding

Assume the duration of the frequency pulse of a GMSK signal is LT, then the phase pulse shape $\phi(t)$ will be 0 when $t \leq -LT$, and $\pi/2$ at $t \geq LT$. Thus,

$$\begin{aligned} S_{2k+j}^{(l)} &= e^{\jmath\theta(kT + \frac{jT}{2} - \frac{lT}{2})} \\ &= \exp\{\jmath\frac{\pi}{2}\langle\sum_{i=0}^{q-L} b_i, 4\rangle\} \prod_{i=-L}^{L-1} \exp\{\jmath b_{q-i}\phi(\rho\frac{T}{2} + iT)\} \end{aligned} \tag{7}$$

where $\langle\cdot,\cdot\rangle$ denote the modulo operation and integer q and $\rho \in \{0, 1\}$ is chosen to satisfy $kT + \frac{jT}{2} - \frac{lT}{2} = qT + \frac{\rho T}{2}$.

To get finite memory to the receive model, we employ a precoding procedure to encode the information bit b_i into I_i as

$$\begin{cases} I_{2k+1} = -b_{2k} b_{2k+1} \\ I_{2k} = b_{2k-1} b_{2k} \end{cases}. \tag{8}$$

At the transmitter, information bits b_i is encoded into I_i and then modulated and transmitted. By introducing the precoding (8), the sampled GMSK signal $S_{2k+j}^{(l)}$ becomes the function of $b_{q-L}, \ldots, b_{q+L-1}, b_{q+L}$ as

$$\begin{aligned} S_{2k+j}^{(l)} &= \jmath^{\langle q-L+1,\, 2\rangle} b_{q-L} \prod_{i=-L}^{L-1} \exp\{\jmath(-1)^{q-i} b_{q-i} \\ &\qquad b_{q-i-1}\phi(\rho\frac{T}{2} + iT)\}. \end{aligned} \tag{9}$$

With the above precoding scheme, given $\{b_i\}$, $S_{2k+j}^{(l)}$ can be obtained as follows: First, compute (q, ρ) from (k, j, l); Then choose one of the four state tables according to the parameters $(q - L)$ and ρ; Finally, look up in this table the state value indexed by $\{b_{q-L}, \ldots, b_{q+L}\}$. This procedure is later used in computing the conditional *a posteriori* distributions (16)-(18).

2.3. Approximation of the phase pulse

In [4], an approximation of phase pulse shape of GMSK signal with $BT = 0.3$ is given. Fig. 1 shows that after this non-significant approximation, L is reduced from 3 to 2, consequently, the state number of the sampled GMSK signal is reduced by a factor of 2. Similar, we can always do this kind of nonlinear approximation, which is non-significant and always better than a linear approximation, to reduce the complexity of the receiver.

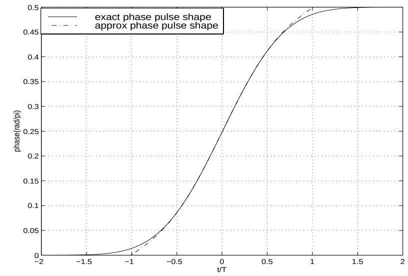

Fig. 1. An approximation for the phase pulse shape (BT $= 0.3$).

2.4. Problem Setup

Now consider the system model (6), in fact channel h_l is assumed constant within a time block. Assuming there are M symbols within a time block. For compactness, define $\boldsymbol{B} \triangleq \{b_0, \ldots, b_{M-1}\}$; $\boldsymbol{Y} \triangleq \{y_0, y_1, \ldots y_N\}$, where $N = 2M + \lceil\max(\tau_k/\frac{T}{2})\rceil$ is the total number of samples at the receiver within a time block. Let's further define $\boldsymbol{h} \triangleq (h_{p_0}, \ldots, h_{p_1})^H$, $\boldsymbol{S}_t \triangleq (S_t^{(p_0)}, \ldots, S_t^{(p_1)})^T$, the system model becomes:

$$y_t = \boldsymbol{h}^H \boldsymbol{S}_t + v_t, \qquad t = 0, 1, \ldots, N-1. \tag{10}$$

In section 3, we consider the problem of estimating the *a posteriori* probabilities of the transmitted symbols

$$P(b_k = \pm 1 | \boldsymbol{Y}), \qquad k = 0, 1, \ldots, M-1. \tag{11}$$

based on the received signals $\boldsymbol{Y}$ and the prior information of $\boldsymbol{B}$, without knowing the channel $\boldsymbol{h}$, and the noise parameters. Notice that $\boldsymbol{S}_t$ in (10) can be computed under the knowledge of $\boldsymbol{B}$ according to the previous discussion.

3. BAYESIAN EQUALIZER

In this section, we consider the problem of computing *a posteriori* symbol probability in (11), under the assumption that the ambient noise distribution is complex Gaussian. That is,

$$v_t \sim \mathcal{N}_c(0, \sigma^2). \tag{12}$$

The problem will be solved under a Bayesian framework: First, the unknown quantities $\theta = \{\boldsymbol{h}, \sigma^2, \boldsymbol{B}\}$ are regarded as independent random variables with some prior distributions. The Gibbs sampler, a Monte Carlo method, is then employed to calculate the maximum *a posteriori* (MAP) estimation of these unknowns.

3.1. Gibbs Sampler

To the problem of joint sequence detection and channel estimation, recent paper [5] has shown that *Gibbs sampler*, a Bayesian approach, is a very powerful Bayesian solution. Let $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_d]^T$ be a vector of unknown parameters, $\boldsymbol{Y}$ be the observed data. Algorithmically, we can describe the Gibbs sampler as follows:

- For $i = 1, \ldots, d$, we draw $\theta_i^{(t+1)}$ from the conditional distribution

$$p(\theta_i^{(n+1)} | \theta_1^{(n+1)}, \ldots, \theta_{i-1}^{(n+1)}, \theta_{i+1}^{(n)}, \ldots, \theta_d^{(n)}, \boldsymbol{Y}).$$

Under regularity conditions,the distribution of $\boldsymbol{\theta}^n$ converges geometrically to $p(\boldsymbol{\theta}|\boldsymbol{Y})$, as $n \rightarrow \infty$, which is the posteriori marginal distribution. Therefore, the maximum a posteriori estimates of each unknown parameter will be relatively easy to compute.

3.2. Prior Distributions

We first specify the prior distributions.

1. For the unknown channel $\boldsymbol{h}$, a complex Gaussian prior distribution is assumed,

$$p[\boldsymbol{h}] \sim \mathcal{N}_c(\boldsymbol{h}_0, \boldsymbol{\Sigma}_0). \tag{13}$$

   Note that large value of $\boldsymbol{\Sigma}_0$ corresponds to less informative prior.

2. For the noise variance σ^2, an inverse chi-square prior distribution is assumed,

$$p[\sigma^2] \sim \chi^{-2}(2\nu_0, \lambda_0). \tag{14}$$

   Small value of $2\nu_0$ corresponds to the less informative priors.

3. The LLR of symbol $\{b_i\}$ can be expressed as

$$\rho_i = \log \frac{P(b_i = +1)}{P(b_i = +1)}. \tag{15}$$

   When there are no prior information for these symbols, ρ_i are set to zero.

3.3. Conditional Posterior Distributions

The following conditional posterior distributions are required by the Bayesian multiuser detector.

1. The conditional distribution of the channel response $\boldsymbol{h}$ given σ^2, $\boldsymbol{B}$, and $\boldsymbol{Y}$ is

$$p(\boldsymbol{h}|\boldsymbol{B}, \sigma^2, \boldsymbol{Y}) \sim \mathcal{N}_c(\boldsymbol{h}_\star, \boldsymbol{\Sigma}_\star), \tag{16}$$

$$\text{with} \quad \boldsymbol{\Sigma}_\star^{-1} \triangleq \boldsymbol{\Sigma}_0^{-1} + \frac{1}{\sigma^2} \sum_{t=0}^{N-1} \boldsymbol{S}_t(\boldsymbol{B}) \boldsymbol{S}_t^H(\boldsymbol{B}),$$

$$\boldsymbol{h}_\star \triangleq \boldsymbol{\Sigma}_\star \left(\boldsymbol{\Sigma}_0^{-1} \boldsymbol{h}_0 + \frac{1}{\sigma^2} \sum_{t=0}^{N-1} \boldsymbol{S}_t(\boldsymbol{B}) y_t^* \right).$$

2. The conditional distribution of the inverse of noise variance σ^2 given $\boldsymbol{h}$, $\boldsymbol{B}$, and $\boldsymbol{Y}$ is

$$p(\sigma^2|\boldsymbol{h}, \boldsymbol{B}, \boldsymbol{Y}) \quad \sim \quad \chi^{-2}\left(2[\nu_0 + N], \frac{\nu_0 \lambda_0 + s^2}{\nu_0 + N} \right), \tag{17}$$

$$\text{with} \quad s^2 \quad \triangleq \quad \sum_{t=0}^{N-1} \left| y_t - \boldsymbol{h}^H \boldsymbol{S}_t(\boldsymbol{B}) \right|^2.$$

3. The conditional distribution of the data bit b_i given $\boldsymbol{h}$, σ^2, $\boldsymbol{B}_i$, and $\boldsymbol{Y}$ can be obtained from [where $\boldsymbol{B}_i \triangleq \boldsymbol{B} \setminus b_i$.]

$$\begin{aligned} & \frac{P(b_i = +1|\boldsymbol{h}, \sigma^2, \boldsymbol{B}_i, \boldsymbol{Y})}{P(b_i = -1|\boldsymbol{h}, \sigma^2, \boldsymbol{B}_i, \boldsymbol{Y})} \\ = \quad & \exp\{\rho_i - \frac{1}{\sigma^2} \sum_{t=t_0}^{t_1} (|y_t - \boldsymbol{h}^H \boldsymbol{S}_t^{i,+}|^2 \\ & -|y_t - \boldsymbol{h}^H \boldsymbol{S}_t^{i,-}|^2)\}, \end{aligned} \tag{18}$$

   where $\boldsymbol{S}_t^{i,+} \triangleq \boldsymbol{S}_t(b_i = +1, \boldsymbol{B}_i)$ and $\boldsymbol{S}_t^{i,-} \triangleq \boldsymbol{S}_t(b_i = -1, \boldsymbol{B}_i)$; $t_0 = 2(i - L) + p_0$ and $t_1 = 2(i + L) + 1 + p_1$.

3.4. The Gibbs Equalizer

Using the above conditional posterior distributions, the Gibbs sampling implementation of the Bayesian equalizer proceeds iteratively as follows. Given the initial values of the unknown quantities $\{\boldsymbol{h}^{(0)}, \sigma^{2(0)}, \boldsymbol{B}^{(0)}\}$ drawn from their prior distributions, and for $n = 1, 2, \ldots$

1. Draw $\boldsymbol{h}^{(n)}$ from $p[\boldsymbol{h}\,|\sigma^{2(n-1)}, \boldsymbol{B}^{(n-1)}, \boldsymbol{Y}]$ given by (16);
2. Draw $\sigma^{2(n)}$ from $p[\sigma^2\,|\boldsymbol{h}^{(n)}, \boldsymbol{B}^{(n-1)}, \boldsymbol{Y}]$ given by (17);
3. For $i = 0, 1, \ldots, M-1$
   Draw $b_i^{(n)}$ from $P[b_i|\boldsymbol{h}^{(n)}, \sigma^{2(n)}, \boldsymbol{B}_i^{(n-1)}, \boldsymbol{Y}]$ given by (18).

To ensure convergence, the above procedure is usually carried out for $(k_0 + K)$ iterations and samples from the last K iterations are used to calculate the Bayesian estimates of the unknown quantities. In particular, the marginal posterior bit probabilities in (11) are calculated as

$$P(b_i = +1|\boldsymbol{Y}) \cong \frac{1}{K} \sum_{k=k_0+1}^{k_0+K} \delta_i^{(k)}, \tag{19}$$

where $\delta_i^{(k)}$ is an indicator such that $\delta_i^{(k)} = 1$, if $b_i^{(k)} = +1$ and $\delta_i^{(k)} = 0$, if $b_i^{(k)} = -1$.

4. TURBO EQUALIZATION

We consider employing iterative equalization and decoding to improve the performance of the Bayesian equalizer in a coded system. Because it utilizes the *a priori* symbol probabilities, and it produces symbol *a posteriori* probabilities, the Bayesian equalizer developed in this paper is well suited for iterative (Turbo) processing. The Turbo receiver consists of two stages: the Bayesian equalizer followed by a soft-input soft-output channel decoder [6]. The two stages are separated by deinterleavers and interleavers. Assume that $\{b_i\}$ is mapped into $\{b_{\pi(i)}\}$ after deinterleaving. An iterative (Turbo) receiver can be implemented as shown in Fig. 2, where $\{\Lambda_1(b_i)\}$ and $\{\Lambda_2(b_{\pi(i)})\}$ are the posterior distribution in terms of LLR at the output of the Bayesian equalizer and the channel decoder. $\lambda_1(b_i)$ and $\lambda_2(b_{\pi(i)})$ are respectively the extrinsic information, which are act as prior information to exchange between the Bayesian equalizer and channel decoder. Note that at the first iteration, the extrinsic information $\{\lambda_1(b_i)\}$ and $\{\lambda_2(b_i)\}$ are statistically independent. But subsequently since they use the same information indirectly, they will become more and more correlated and finally the improvement through the iterations will diminish.

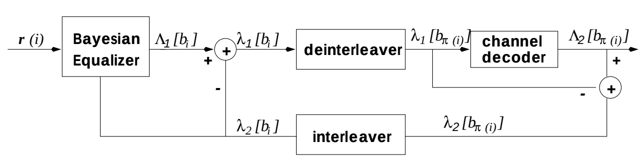

Fig. 2. Turbo equalization.

5. SIMULATION RESULTS

In this section, the GMSK signal with $BT = 0.3$ is chosen to provide simulation examples to illustrate the performance of the Turbo equalizer developed in this paper. We consider the multipath channel shown in Fig. 3 $A\&B$, where $h(t) = \sum_{l=1}^{p} a_l \delta(t - \tau_l)$, with the number of the path $p = 3$. (Note that the channel is normalized to have unit norm, i.e., $|a_1|^2 + |a_2|^2 + |a_3|^2 = 1$). $p_0 = -3$ and $p_1 = 6$ is chosen to make the truncation for the tapped-delay line model (5). The truncated tap coefficients are shown in Fig. 3 $C\&D$.

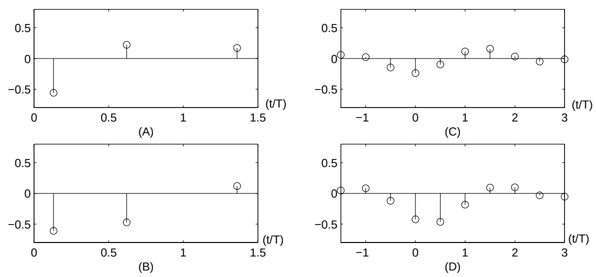

Fig. 3. The simulated multipath channel: (A) The channel delay profile (real part); (B) The channel delay profile (imaginary part); (C) The coefficients of the tapped-delay line model (real part); (D) The coefficients of the tapped-delay line model (imaginary part).

The approximation of the phase pulse-shaping (1) is employed to reduce the number of the states of the sampled GMSK signal. The channel code is a rate of $\frac{1}{2}$ constraint length-5 convolutional code (with generators 23, 35 in octal notation). The interleaver is generated randomly and fixed for all simulations. The block size of the information bits is set to be 128. In order to resolve the phase and shift ambiguities, 25 training bits are added to the 256 interleaved code bits. Finally, 2 tail bits are added to every data block to close the memory of the GMSK signal. In computing the bit probabilities, the Gibbs sampler is iterated 100 runs for each data block, with the first 50 iterations as the "burning-in" period, i.e., $k_0 = K = 50$ in (19).

Fig. 4 illustrates the performance of the Turbo equalizer discussed in Section 5. The code bit error rate at the output of the Bayesian equalizer is plotted for the first three iterations. By constructing a trellis diagram [4] of the uncoded GMSK-modulated signal (the delay channel model (1) is used here), a MAP demodulation algorithm is performed assuming perfect knowledge of the channel, and the simulation result is included in Fig. 4. By combining the MAP decoder with the above MAP demodulator, a conventional iterative (Turbo) equalizer assuming perfect channel knowledge is also implemented, and the code bit error rate at the output of the MAP demodulator at the third iteration is shown in Fig. 4 for the purpose of comparison.

It is seen that in terms of bit error rate, the performance of the proposed Bayesian equalizer is within about 1dB from the uncoded MAP demodulation which assumes the perfect knowledge about channel. It is also seen that by incorporating the extrinsic information provided by the channel decoder, the proposed Bayesian equalizer achieves significant performance improvement by the iterative procedure. For the coded system, the simulation result shows that the code bit error rate of our proposed Turbo equalizer is also within about 1dB from the Turbo equalizer which assumes perfect knowledge about channel.

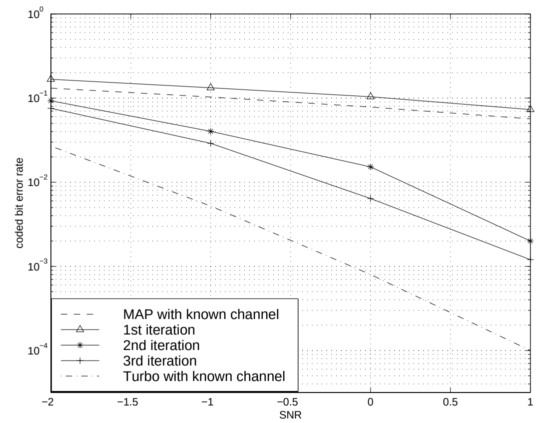

Fig. 4. Performance of Turbo equalizer for the multipath distorted GMSK system with $BT = 0.3$.

6. CONCLUSION

We have considered the problem of signal recovery in GMSK systems with multipath distortions. A nonlinear model is derived for the sampled signal at the receiver due to the nonlinear nature of the GMSK modulation. A novel equalization scheme is developed which is optimal in the sense that it is based on the Bayesian inference of all unknown quantities, and such a Bayesian equalizer can be efficiently implemented using the Gibbs sampler, a Markov Chain Monte Carlo procedure for computing Bayesian estimates. Because of its "soft-input soft output" feature, the Bayesian equalizer is designed to be a part of Turbo equalizer, which refines its processing based on the information from the decoding stage. The effectiveness of the proposed techniques is demonstrated by simulation examples. In fact, this paper provides an example to use the Bayesian equalizer for general nonlinear system.

7. REFERENCES

[1] R. Steele, *Mobile Radio Communications*, IEEE Press, New York, 1992.

[2] John G. Proakis, *Digital Communications*, McGraw-Hill, Inc, third edition, 1995.

[3] Theodore S. Rappaport, *Wireless Communications*, Prentice-Hall, Inc, 1995.

[4] J.-T Chen, A. Paulraj, and U. Reddy, "Multichannel maximum-likelihood sequence estimation(MLSE) equalizer for GSM using a parametric channel model," *IEEE Trans. Commun.*, vol. 47, pp. 53–63, Jan. 1999.

[5] X. Wang and R. Chen, "Adaptive Bayesian multiuser detection for synchronous CDMA with Gaussian and impulsive noise," *IEEE Transactions on Signal Processing*, pp. 2013–2028, July 2000.

[6] L.R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 284–287, Mar. 1974.