GMSK FOR MOBILE COMMUNICATION

M. Kent
PCSI
San Diego, CA
92121

D.D. Sworder
Dept. of ECE
UC San Diego
La Jolla, CA 92093-0407
sworder@ece.ucsd.edu

R.J. Elliott
Dept. of Math. Sciences
University of Alberta
Edmonton, Alberta
CANADA

Abstract

A new estimation protocol is proposed for GMSK. Current receivers sample the received signal when the "eye-opening" is maximal. Clock drifts and intersymbol interference degrade performance when the out-of-channel power spectral density is reduced. This paper shows that a nonlinear, multisample receiver can be used to improve detection performance. Further, the multisample approach uses more of the transmitted energy in the filtered pulse to improve detection probability.

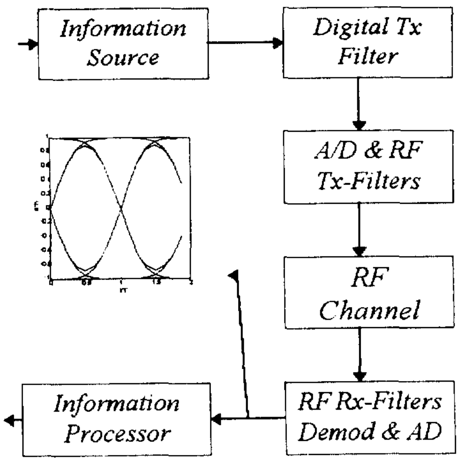

Figure 1: Signal flow in a wireless communication system.

1. Introduction

Figure 1 is a block diagram portraying a mobile communication link. A binary symbol sequence, $\{x_k\}$, is recast by a transmit filter to create the baseband information signal. This signal is modulated, transmitted and demodulated to yield the baseband signal at the receiver. This latter is sampled to generate an observation sequence $\{y_k\}$ from which $\{x_k\}$ can be reconstructed. The mapping from $\{y_k\}$ to the decoder output is the detection protocol.

An important advantage accruing to a properly chosen transmit filter is that the RF spectrum required for the link can be significantly reduced. If the input is replaced with a bivalued simple process with period T, (called $\{x_t\}$), and if the transmit filter has impulse response $\{g_t\}$, the baseband signal, $\{f_t\}$, can be represented as $\{f_t\} = \{g_t *(2x_t-1)\}$ where "*" is orthodox convolution. Specifically, consider Gaussian minimum shift keying (GMSK), a popular algorithm in mobile communications.[1], [2] The (noncausal) transmit filter for GMSK is given by the parametric family: $g_t= (\alpha/\sqrt{\pi})\exp(-\alpha t^2)$; $t\in(-\infty,\infty)$ (or better in terms of β, the bandwidth of the filter relative to the bit rate: $\alpha=\pi(2\ln2)\beta T$). As a notational convention, let **1** be a vector of "ones," and denote by I_A the indicator of a set A; e.g., $I_{[0,T]}$ is the indicator of [0,T] (for convenience $I_{[kT,(k+1)T]}=I_k$, and where it is easier to read I_A=I(A). Figure 2 shows the $\{h_t\}=\{g_t*I_0\}$, sightly translated, for β=.5 along with I_0 ($\beta=\infty$). The (normalized) power spectral density of $\{f_t\}$ with GMSK is contrasted with that without it in Figure 3 (again for β=.5). The reduction in spectral spill over is apparent.

There are various ways in which the data sequence could be reconstructed, the simplest being: sample the received baseband signal at the midpoint of the current temporal bin, and if the signal exceeds a threshold, x_k=1 is declared. This decoder uses the maximum value of $\{f_t\}$, and the effects of additive transmission noise should be ameliorated. Unfortunately, the temporal filtering that reduces the transmission bandwidth, creates intersymbol interference (ISI) at the receiver. The support of $\{g_t*I_0\}$ is not confined to the coincident temporal bin, but is readily discernible in the contiguous neighbors; the transmit filter has an infinite tail (both pre- and postsentient), but when β=0.5, the essential support of $\{h_t\}$ is confined to just these three bins. Figure 4 displays the set of all 8 possible $\{f_t\}$ signals as (x_{-1}, x_0, x_1) varies over their range: the "eye" chart.

Neglecting exogenous effects, the three bit sequence (x_{-1} x_0 x_1) could be unambiguously retrieved by multiple sampling of $\{f_t\}$ on just the single interval [0,T]. Detection is made complicated by anomalous effects in signal generation and transmission; e.g., channel noise, ISI, drifts in synchronization, fading, etc. In this paper, focus will be

1058-6393/95 $4.00 © 1995 IEEE

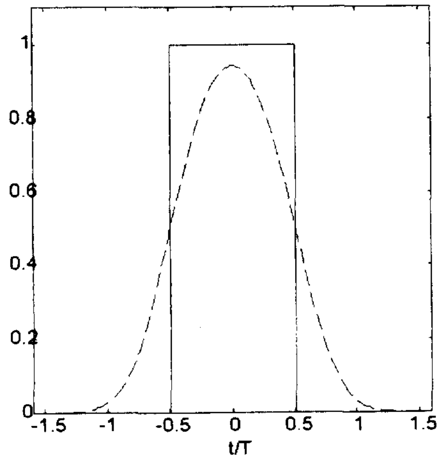

Figure 2: Response of the transmit filter in GMSK: β=0.5.

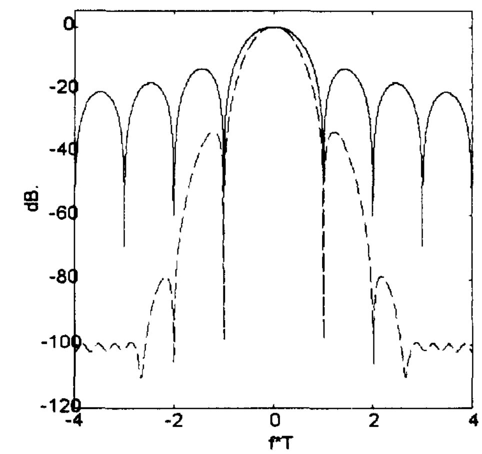

Figure 3: Spectrum of the transmitted signal: β=.5 and β=∞.

on channel noise. To be more precise, suppose the data sequence, $\{x_k;\ k\in Z^+\}$, is a Markov process, and suppose the decoder samples $\{f_t\}$ at the rate λ samples/symbol, generating a data process $\{y_\ell;\ \ell\in Z^+\}$ thereby. Denote the filtration generated by $\{y_n\}$ by $\boldsymbol{Y}_n$. The decoder is a $\boldsymbol{Y}_n$-adapted mapping to {0,1}; i.e. the decoder approximates x_k with $\mathbf{x}_k$ where $\mathbf{x}_k$ is $\boldsymbol{Y}_n$-adapted. (If $n\geq\lambda k$, the decoder is noncausal.) For example, in the simplest GMSK decoder, $\lambda=1$ and $\mathbf{x}_k=1$ if $y_k>0$ with $\mathbf{x}_k=0$ otherwise.

More sophisticated approaches to this problem involve decoders which are based upon the $\boldsymbol{Y}_n$-conditional probability distribution of $\{x_k\}$; or equivalently, since x_k is an indicator function, $\{\hat{x}_k\}$ where $\hat{x}_k=E\{x_k|\boldsymbol{Y}_n\}$. The natural analogue of the simple decoder described above would be: $\mathbf{x}_k=1$ if $\hat{x}_k>.5$ with $\mathbf{x}_k=0$ otherwise. To derive the dynamic equation for $\{\hat{x}_k\}$, careful modeling of both the data dynamics and the observation is required. This paper uses a descriptive model to represent both the data sequence and ISI to derive a nonlinear receiver algorithm which provides an accurate estimate of the binary sequence. By using ISI for prediction and correction, it is possible to achieve decoding results superior to those attainable using a single sample per temporal bin.

2. A nonlinear processing algorithm

The data sequence $\{x_k\}$ ($x_k\in\{0,1\}$) is well described by a Markov model: $x_{k+1}=Px_k+m_k$ where $\{m_k\}$ is a martingale difference sequence. The output of the transmit filter attributable to the single symbol $x_0=1$ is $\{h_t\}$. For a data sequence on Z^+, the baseband process can be found by superposition; $f_t=\sum_{i\in Z+}(2x_i-1)h_{t-iT}$.

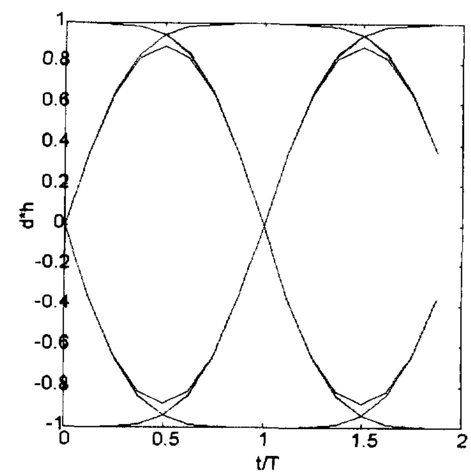

Figure 4: The "eye" chart showing ISI for β=0.5.

The information processing problem is that of recovering $\{x_k\}$ from noisy measurements of $\{f_t\}$. Denote the observations at the information processor input by $\{y_n\}$. These measurements will be made at a rate λ samples/symbol, and at the times $\{\tau_1,\ldots,\tau_\lambda\}$ after the start of a temporal bin. As a notational convenience, when the observation is the rth within interval of the kth symbol a more detailed labeling convention will be used: $y_{kr}=y_n$ where $n=k\lambda+r$; $1\leq r\leq\lambda$. This distinction between absolute and symbol-relative time will be made in what follows as appropriate.

It is evident from Fig. 2 that the baseband signal, $\{f_t\}$, is well approximated as a combination of contributions from the current symbol and its two temporal nearest neighbors: if $t \in I_k$, $f_t \cong \sum_{i\in(-1,0,1)} (2x_{k+i}-1)h_{t-(k+i)T}$. The baseband signal can be written more concisely by modifying the state space. Use the digits of the binary expansion of the numbers 0 through 7 to create an 8×3-matrix $\mathbf{N}=[(a_1, a_2, a_3)]$. Define $\mathbf{u}_t=\mathbf{N}(h_{t-T}, h_t, h_{t+T})'I_0$. Then $(2\mathbf{u}_t-1)I_0$ is a vector process, and its eight components form the eye chart displayed in Fig. 4.

To create the comprehensive state space of the signal, define an 8-dimensional random sequence $\{\phi_k\}$ as follows: $\phi_n=\mathbf{e}_i$ if $n=k\lambda+r$ and the binary number $x_{k-1}x_k x_{k+1}+1=i$. Then, ϕ_n is a unit vector in $\Re^8$ with transitions at symbol transitions. Further, since $\{x_k\}$ is a Markov process, $\{\phi_n\}$ is also Markovian: $\phi_n= Q_n'\phi_{n-1}+M_n$ where $Q_n=I$ and M_n is zero if $n\neq k\lambda$; while if $n=k\lambda$, $Q_n=Q$, a matrix easily constructed from P. The sequence $\{M_n\}$ is a martingale difference; e.g., if $\{\phi_n\}$ makes an $i\rightarrow j$ transition, then $(M_n)_\ell = 1-Q_{i\ell}$ if $\ell=j$; $-Q_{i\ell}$ otherwise. With these conventions, $\{f_t\}$ can be written: if $t\in I_k$, $f_t=\phi_k'\mathbf{u}_{t-kT}$. Note that the signal space is of higher dimension than the symbol space to account for the existence of ISI.

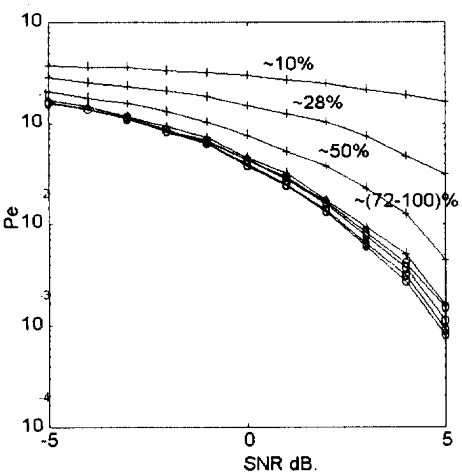

Figure 5: Bit-error rate for the MSKEP as a function of number of samples per symbol.

The decoder generates an estimate of $\{x_k\}$ from $\{y_n\}$ and this estimate need not be causal. The proper estimate will be determined from the application, but most reasonable estimates involve the Y_n-conditional expectation of $\{x_k\}$. Suppose the impact of the transmitter-channel-demodulator can be represented by an iid noise process $\{b_n\}$: $y_{kr} =\phi_k'\mathbf{u}_{\tau r}+b_{kr}$ (or $y_n =(\phi'\mathbf{u})_n+b_n$ for short) where the b_n have common density φ. The conditional distribution of the symbols can be written $\mathcal{P}(x_k=1|Y_n\}=\sum_{\phi\in\{.I.\}} E\{\phi_k|Y_n\}$ (with $\mathcal{P}(x_k=0|Y_n\}= 1-\mathcal{P}(x_k=1|Y_n\})$. The information processor requires the computation of $\hat{\phi}_k=E\{\phi_k|Y_n\}$, the dynamic equation for which is derived in the appendix. Define an 8-vector $\boldsymbol{\varphi}_n =[\varphi(y_n - \mathbf{e}_i'\mathbf{u}_n)]/\varphi(y_n)$, and let "★" be the Hadamard product. The MSK estimation protocol (MSKEP) is:

$$\hat{\phi}_n= q_n/\mathbf{1}'q_n \text{ subject to: } q_n =\boldsymbol{\varphi}_n\star Q_n'q_{n-1}.$$

Multisampling has significant advantages. Figure 5 shows the bit-error-rate (BER) of the MSKEP as a function of the number and location of the samples (see Figure 6), and the signal-to-noise ratio SNR of the channel. Four samples/symbol gives performance essentially indistinguishable from more rapid sampling and will be the normative selection in what follows.

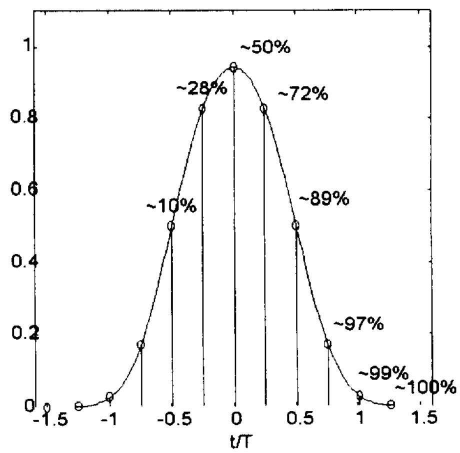

Figure 6: Sample location an total energy for different sample rates of MSKEP.

3. Performance of the algorithm

The MSK estimation protocol is presented in the previous section. It uses a simple linear recursion with observation dependent "gain." In this section, the performance of a simple decoder based upon the MSK estimation protocol is compared to two well known alternatives: the matched filter-maximum likelihood sequence estimator (MLSE) [3]; and a decoder based upon a single sample at the symbol midpoint. Consider the former first. Suppose φ is $N(0,\sigma)$, and the sample rate $\lambda=4$. The MLSE first processes $\{y_k\}$ using a filter matched to $\{h_t\}$. This smoothing of the noise reduces its intensity, but creates a nonwhite noise spectrum and an increase in ISI. Indeed, it can be shown that the matched filter reduces the effective β at the decoder by a factor of 2: $\beta_{eff}=0.25$. To minimize the correlation in the processed noise, the output of the matched filter is decimated (by a factor of 4 in what follows). Note that the spectral reduction in the channel is still that associated with $\beta=0.5$. For a system in temporal synchronism, define the SNR by average peak in the eye chart to the postfilter noise variance. Matched filtering reduces both, but it can be shown that SNR is improved by 2.7 dB over that which would obtain from a single sample algorithm.

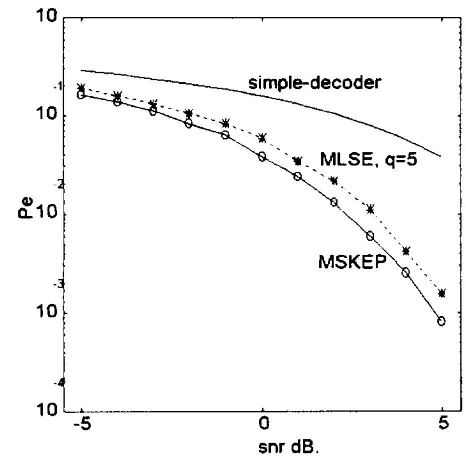

Figure 7: Performance of the MSKEP-based decoder along with the simple-decoder and the VA with one sample/symbol.

The decimated output sequence can be used in the Viterbi algorithm (VA).[4] The VA is a model-based procedure similar in concept to the MSKEP. The MSKEP is noncausal, computing a $Y_{(k+1)\lambda}$-predictable version of $\hat{x}_k$. A maximum *a posteriori* logic is then used for decoding. The VA carries this relaxation of causality further. Let $N >> k\lambda$. Then a Y_N-predictable version of $\hat{x}_k$ would necessarily be more accurate than that used in MSKEP. In principle, the VA is based upon maximizing $\wp(\{x_\ell\};\ \ell\in\{1,...,N/\lambda\}\cap\{y_\ell\};\ell\in\{1,...,N\})$ over $\{x_\ell\};\ell\in\{1,...,N/\lambda\}$. Denote this maximizing $\{x_\ell\};\ \ell\in\{1,...,N/\lambda\}$ by $\{\hat{x}_\ell\}$. The kth element of the sequence, $\hat{x}_k$, is the decoded signal at time k. It is a conditional expectation as is the MSKEP estimate, though with respect to a different filtration. The VA uses the same decoding logic, but bases it on a more accurate estimate of x_k--if N is large.

Unfortunately, the VA can not be implemented as indicated because $\{\hat{x}_k\}$ is too difficult to compute. The smoothed mean of the symbol sequence $\{x_k\}$ involves both forward and backward filtering. The VA uses a forward dynamic programming formalism to approximate the maximum likelihood estimate from a limited set of observations spanning the symbol in question. Further, to reduce the set of symbol sequences that must be kept current, severe pruning in required. To algorithmic complexity comparable with the MSKEP while balancing these exigencies, the VA was implemented using 11 temporal bins, and because of decimation, one sample/symbol.

The single sample threshold decoder (simple-decoder) uses a single sample/symbol and uses a threshold of zero. The results of a comparison of the simple-decoder, the one sample/symbol VA and the four sample/symbol MSKEP are shown in Figure 7 as a function of channel SNR. Probability of error is largest for the simple-decoder for all SNR, and the decay as a function of SNR is slow. The MSKEP is superior to the MLSE at all SNR even with a shorter delay (one symbol vs. five symbols). This along with its implementation simplicity suggest that the MSKEP is an algorithm that warrants consideration in this application.

4. Conclusions

In this paper, a novel decoding algorithm has been presented for a channel using GMSK. A simple recurrence formula has been used to produce the unnormalized conditional distribution of the symbol sequence. From this, the message can be decoded, and a measure of confidence in the ostensible message created. The algorithm is simple enough that multiple samples/symbol can be used with simple processors.

Only decoding temporally synchronous processes without fading have been considered here. Current work involves relaxing this hypothesis and including the influence of more general channel distortions. The procedures used here extend naturally to this more general class of estimation problems.

Appendix

In this appendix, a concise development of the MSKEP is provided. This development is based upon a procedure presented in considerably more detail in [5]. Begin with the probability space $(\Omega,\mathcal{A},\wp_1)$ and two vector random processes $\{\phi_n\}$ and $\{y_n\}$. Let $\{\boldsymbol{F}_n\}$ be the filtration generated by $\{\phi_n, y_{n-1}\}$ and $\{\boldsymbol{Y}_n\}$ be the filtration generated by $\{y_n\}$. Suppose that $\{\phi_n\}$ is an $\boldsymbol{F}_n$-Markov process and $\{y_n\}$ is an $\boldsymbol{F}_n$-iid sequence with density φ. Then the dynamic equation for $\{\phi_n\}$ is of the form: $\phi_n = Q_n'\phi_{n-1} + M_n$ where M_n is a $\boldsymbol{F}_n$-martingale difference.

Create three random sequences $\{b_n\}$, $\{\breve{g}_n\}$ and $\{\breve{G}_n\}$: $b_n = y_n - (\phi'\mathbf{u})_n$; $\breve{g}_n = \varphi(b_n)/\varphi(y_n)$; $\breve{G}_n = \prod_{\ell=[1,n]} \breve{g}_\ell$. Define a new probability measure $\wp$ on $(\Omega,\mathcal{A})$ by setting the $\boldsymbol{F}_n$-restriction of $d\wp/d\wp_1$ equal to $\breve{G}_n$. Consider the behavior of $\{b_n\}$ under $\wp$: $\wp\{b_n \le t|\boldsymbol{F}_n\} = E\{I(b_n\le t)|\boldsymbol{F}_n\} = E_1\{\breve{G}_n I(b_n\le t)|\boldsymbol{F}_n\}/E_1\{\breve{G}_n|\boldsymbol{F}_n\}$. But $E\{\breve{g}_n|\boldsymbol{F}_n\} = \int_{(-\infty,\infty)} [\varphi(b_n)/\varphi(y_n)]\varphi(y_n)dy_n = \int_{(-\infty,\infty)} \varphi(y_n - \phi_n'\mathbf{u}_n)\, dy_n = 1$, and $E\{\breve{g}_n I(b_n\le t)|\boldsymbol{F}_n\} = \int_{(-\infty,\infty)} \varphi(b_n)\, I(b_n\le t)\, db_n = \int_{(-\infty,t)} \varphi(b_n)db_n = \wp_1\{b_n\le t\}$. Hence, $\{b_n\}$ is a $\wp$-independent sequence.

Let the s-vector sequence $\{q_n\}$ be given by $[E_1\{\check{G}_n\phi_n |\mathcal{Y}_n\}]$. Then $\hat{\phi}_n=E\{\phi_n|\mathcal{Y}_t\}=q_n/1'q_n$. The recursive equation for q_n can be written: $E_1\{\check{G}_n\phi_n|\mathcal{Y}_n\}= E_1\{\check{G}_{n-1}\phi_n[\varphi(b_n)/\varphi(y_n)] |\mathcal{Y}_n\}= (\varphi(y_n))^{-1} E_1\{\check{G}_{n-1}\phi_n\varphi(b_n)|\mathcal{Y}_t\}$. The ith component of q_n is $q_n(i) =E_1\{\check{G}_n I(\phi_n=\mathbf{e}_i)|\mathcal{Y}_n\} =(\varphi(y_n))^{-1} E_1\{\check{G}_{n-1}I(\phi_n=\mathbf{e}_i)\varphi(b_n) |\mathcal{Y}_n\} =(\varphi(y_n))^{-1} \varphi(y_n -\mathbf{e}_i'\mathbf{u}_n) E_1\{\check{G}_{n-1}I(\phi_n=\mathbf{e}_i)|\mathcal{Y}_n\} =(\varphi(y_n))^{-1} \varphi(y_n -\mathbf{e}_i'\mathbf{u}_n) \sum_j E_1\{\check{G}_{n-1}I(\phi_n=\mathbf{e}_i,\phi_{n-1}=\mathbf{e}_j)|\mathcal{Y}_t)\}$. But under $\wp_1$, $\{y_k\}$ is a sequence of independent random variables. So $q_n(i)=(\varphi(y_n))^{-1} \varphi(y_n -\mathbf{e}_i'\mathbf{u}_n) \sum_j E_1\{\check{G}_{n-1}I(\phi_{n-1}=\mathbf{e}_j)|\mathcal{Y}_n\}(Q_n)_{ji} =(\varphi(y_n))^{-1} \varphi(y_n -\mathbf{e}_i'\mathbf{u}_n) \sum_j q_{n-1}(j)(Q_n)_{ji}$.

References

1.M.K. Simon and C.C. Wang, "Differential Detection of Gaussian MSK in a Mobile Radio Environment," *IEEE Trans. on Vehicular Technology*, Vol VT-33, No. 4, Nov. 1984, pp. 307-320.

2. K. Murota, and K.
Hirade, "GMSK Modulation for Digital Mobile Radio Telephony," *IEEE Trans. on Communications Technology*, Vol COM-29, July 1981, pp. 1044-1050.

3. G.D. Forney, "Maximum-Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference," *IEEE Trans. on Information Theory*, Vol IT-19, May 1972, pp. 363-378.

4. G.D. Forney, "The Viterbi Algorithm," *Proceedings of the IEEE*, Vol 61, No. 3, March 1973, pp. 268-277.

5. R.J. Elliott, "A General Recursive Discrete Time Filter," *J. of Applied Probability*, Vol. 30, No. 3, Sept. 1993, 575-588.