Code Construction for Gauss-Markov Fading Channel

Prepared by Chin-Hui Cheng

Advisory by Prof. Po-Ning Chen

Institute of Communications Engineering

National Chiao-Tung University

Hsinchu, Taiwan 300, R.O.C.

June 18, 2005

Outline

- Introduction
- Simulated Annealing
- Channel Characteristic
- Design Methodology
- Simulation Results
- Conclusions

Introduction

- Simulated annealing has been applied to code design.
- In 2002, Skoglund, Giese and Parkvall designed a block code and the performance over the block Rayleigh-fading channel was excellent [7].
- Since the channel coefficients are changing during the period of transmission of codeword in practice, we try to design code to adjust well to Gauss-Markov channels.

Simulated Annealing for Code Design

- "Anneal" means to heat and then to cool steel or glass. The cooling stage is slowly so that the potential energy stored in the molecular configuration can be minimized
- A code structure is similar to the molecular constitution in the aspect of the analogy between good distance properties of a code and low potential energy of a molecular configuration.

- Choose initial temperature $T = T_0$, code C

```
Do{
    Do{
        Choose C', a random perturbation of C
        Let Δe=energy(C')-energy(C)
        If (Δe<0) then C←C'
        Else with probability exp(-Δe/T) C←C'
        }
    Until (several energy drops or too many iterations)
    Lower temperature T← αT
    }
Until (stable code configuration is obtained or running time is up )
```

Channel Characteristic

- For a time-varying channel with noise and inter-symbol interference, the received data at time k is

$$r_k = \mathbf{a}_k^T \mathbf{h}_k + n_k,$$

where $\mathbf{a}_k^T = [a_k, a_{k-1}, \cdots, a_{k-M+1}]$ and M is the channel length indicating how far the past transmitted bits affect the received bit in this time instance. $\mathbf{h}_k$ is a complex column vector with M channel impulse response coefficients at time k, and n_k is a Gaussian complex noise with mean zero, variance σ^2.

- For Gauss-Markov channel, $\mathbf{h}_k = \alpha \mathbf{h}_{k-1} + \mathbf{v}_k$ at time k, where $\mathbf{v}_k$ is a complex white Gaussian with mean $\mathbf{d}$ and covariance $\mathbf{C}$.

- Let $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_N]$ be the matrix of channel coefficient vectors representing by columns, and let $\mathbf{A} = [\mathbf{a}_1, \cdots, \mathbf{a}_N]^T$ represent the matrix of transmitted data representing by rows. Also let $\mathbf{r} = [r_1, \cdots, r_N]^T$ represent the received vector, while $\mathbf{n} = [n_1, \cdots, n_N]^T$ be the vector of noise from time 1 to time N.
- Assuming the mean of $\mathbf{v}_k$ is zero and initial channel coefficient $\mathbf{h}_0$ is known, we can derive the likelihood function of Gauss-Markov channel as

$$\begin{aligned} f(\mathbf{r}|\mathbf{A}) &= E[f(\mathbf{r}|\mathbf{H}, \mathbf{A})] \\ &= \int_{\mathbf{H}} f(\mathbf{r}|\mathbf{H}, \mathbf{A}) f(\mathbf{H}) d\mathbf{H} \\ &= \int_{\mathbf{H}} \left(\prod_{k=1}^{N} e^{-\frac{|r_k - \mathbf{a}_k^T \mathbf{h}_k|^2}{\sigma^2}} \right) f(\mathbf{h}_1) \prod_{k=2}^{N} f(\mathbf{h}_k|\mathbf{h}_{k-1}) d\mathbf{H}. \end{aligned}$$

- After marginalizing the channel coefficient distribution,

$$f(\mathbf{r}|\mathbf{A}) = \prod_{k=1}^{n} e^{\mathbf{q}_k^H \mathbf{G}_k \mathbf{q}_k} |\mathbf{G}_k|.$$

and

$$\begin{cases} \mathbf{G}_1^{-1} = \frac{\mathbf{a}_1^* \mathbf{a}_1^T}{\sigma^2} + (1 + |\alpha|^2)\mathbf{C}^{-1} \\ \mathbf{q}_1 = \frac{r_1 \mathbf{a}_1^*}{\sigma^2} + \alpha \mathbf{C}^{-1} \mathbf{h}_0. \end{cases}$$

$$\begin{cases} \mathbf{G}_k^{-1} = \frac{\mathbf{a}_k^* \mathbf{a}_k^T}{\sigma^2} + \mathbf{C}^{-1} + |\alpha|^2 (\mathbf{C}^{-1} - \mathbf{C}^{-1} \mathbf{G}_{k-1} \mathbf{C}^{-1}) \\ \mathbf{q}_k = \frac{r_k \mathbf{a}_k^*}{\sigma^2} + \alpha \mathbf{C}^{-1} \mathbf{G}_{k-1} \mathbf{q}_{k-1}. \end{cases}$$

$$\begin{cases} \mathbf{G}_N^{-1} = \frac{\mathbf{a}_N^* \mathbf{a}_N^T}{\sigma^2} + \mathbf{C}^{-1} - |\alpha|^2 \mathbf{C}^{-1} \mathbf{G}_{N-1} \mathbf{C}^{-1} \\ \mathbf{q}_N = \frac{r_N \mathbf{a}_N^*}{\sigma^2} + \alpha \mathbf{C}^{-1} \mathbf{G}_{N-1} \mathbf{q}_{N-1}. \end{cases}$$

Design Methodology

- Let the code be with length N and rate $R = K/N$. The average block error rate can be upper-bounded by union bound as

$$\begin{aligned} P_e &\triangleq \Pr(\hat{\mathbf{x}} \neq \mathbf{x}) \\ &= 2^{-K} \sum_i \Pr(\hat{\mathbf{x}} \neq \mathbf{x}(i) | \mathbf{x}(i) \text{ is transmitted}) \\ &\leq 2^{-K} \sum_i \sum_{j \neq i} p_{j|i}, \end{aligned}$$

- The pairwise error probability

$$
\begin{aligned}
p_{j|i} &= \Pr\left[\log \frac{\Pr[\mathbf{r}|\mathbf{x}(j)]}{\Pr[\mathbf{r}|\mathbf{x}(i)]} > 0\right] + \frac{1}{2}\Pr\left[\log \frac{\Pr[\mathbf{r}|\mathbf{x}(j)]}{\Pr[\mathbf{r}|\mathbf{x}(i)]} = 0\right]. \\
&= \Pr\left[\log \frac{\prod_{k=1}^{N} e^{\mathbf{q}_k(j)^H \mathbf{G}_k(j)\mathbf{q}_k(j)} |\mathbf{G}_k(j)|}{\prod_{k=1}^{N} e^{\mathbf{q}_k(i)^H \mathbf{G}_k(i)\mathbf{q}_k(i)} |\mathbf{G}_k(i)|} > 0\right] \\
&\quad + \frac{1}{2}\Pr\left[\log \frac{\prod_{k=1}^{N} e^{\mathbf{q}_k(j)^H \mathbf{G}_k(j)\mathbf{q}_k(j)} |\mathbf{G}_k(j)|}{\prod_{k=1}^{N} e^{\mathbf{q}_k(i)^H \mathbf{G}_k(i)\mathbf{q}_k(i)} |\mathbf{G}_k(i)|} = 0\right] \\
&= \Pr\left[\sum_{k=1}^{N} \left(\mathbf{q}_k(j)^H \mathbf{G}_k(j)\mathbf{q}_k(j) - \mathbf{q}_k(i)^H \mathbf{G}_k(i)\mathbf{q}_k(i)\right) > \sum_{k=1}^{N} \left(\log \frac{|\mathbf{G}_k(i)|}{|\mathbf{G}_k(j)|}\right)\right] \\
&\quad + \frac{1}{2}\Pr\left[\sum_{k=1}^{N} \left(\mathbf{q}_k(j)^H \mathbf{G}_k(j)\mathbf{q}_k(j) - \mathbf{q}_k(i)^H \mathbf{G}_k(i)\mathbf{q}_k(i)\right) = \sum_{k=1}^{N} \left(\log \frac{|\mathbf{G}_k(i)|}{|\mathbf{G}_k(j)|}\right)\right].
\end{aligned}
$$

- Now we reform the q-matrix and G-matrix as

$$\mathbf{q}(j,i) = \begin{bmatrix} \mathbf{q}_1(j) \\ \mathbf{q}_2(j) \\ \vdots \\ \mathbf{q}_N(j) \\ \mathbf{q}_1(i) \\ \mathbf{q}_2(i) \\ \vdots \\ \mathbf{q}_N(i) \end{bmatrix}, \mathbf{G}(j,i) = \begin{bmatrix} \mathbf{G}_1(j) & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \mathbf{G}_2(j) & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \dots & 0 \\ 0 & 0 & \dots & \mathbf{G}_N(j) & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & -\mathbf{G}_1(i) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & -\mathbf{G}_N(i) \end{bmatrix}$$

$$\begin{aligned} p_{j|i} &= \Pr\left[\mathbf{q}^H(j,i)\mathbf{G}(j,i)\mathbf{q}(j,i) > \sum_{k=1}^{N}\left(\log\frac{|\mathbf{G}_k(i)|}{|\mathbf{G}_k(j)|}\right)\right] \\ &\quad + \frac{1}{2}\Pr\left[\mathbf{q}^H(j,i)\mathbf{G}(j,i)\mathbf{q}(j,i) = \sum_{k=1}^{N}\left(\log\frac{|\mathbf{G}_k(i)|}{|\mathbf{G}_k(j)|}\right)\right], \end{aligned}$$

where $\mathbf{q}(j,i)$ is complex Gaussian with mean $\mathbf{m}_q(j,i)$ and covariance $\mathbf{S}_q(j,i)$.

- Note that $\mathbf{h}_k$ is complex Gaussian with mean μ_k and covariance $\mathbf{S}_k$,

$$\mu_k = E[\mathbf{h}_k] = E_{\mathbf{h}_{k-1}}[E_{\mathbf{h}_k}[\mathbf{h}_k|\mathbf{h}_{k-1}]] = E_{\mathbf{h}_{k-1}}[\alpha\mathbf{h}_{k-1}] = \alpha\mu_{k-1} = \alpha^k\mathbf{h}_0,$$

and

$$\begin{aligned}
\mathbf{S}_k &= E[(\mathbf{h}_k - \mu_k)(\mathbf{h}_k - \mu_k)^H] \\
&= E[\mathbf{h}_k\mathbf{h}_k^H] - \mu_k E[\mathbf{h}_k^H] - E[\mathbf{h}_k]\mu_k^H + \mu_k\mu_k^H \\
&= E[(\alpha\mathbf{h}_{k-1} + \mathbf{v}_k)(\alpha\mathbf{h}_{k-1} + \mathbf{v}_k)^H] - |\alpha|^2\mu_{k-1}E[\mathbf{h}_{k-1}^H] - |\alpha|^2E[\mathbf{h}_{k-1}]\mu_{k-1}^H \\
&\quad +|\alpha|^2\mu_{k-1}\mu_{k-1}^H(\text{because } \mathbf{h}_k = \alpha\mathbf{h}_{k-1} + \mathbf{v}_k, \mu_k = \alpha\mu_{k-1}) \\
&= |\alpha|^2E[\mathbf{h}_{k-1}\mathbf{h}_{k-1}^H] + E[\mathbf{v}_k\mathbf{v}_k^H] - |\alpha|^2\mu_{k-1}E[\mathbf{h}_{k-1}^H] \\
&\quad -|\alpha|^2E[\mathbf{h}_{k-1}]\mu_{k-1}^H + |\alpha|^2\mu_{k-1}\mu_{k-1}^H \\
&= |\alpha|^2E[(\mathbf{h}_{k-1} - \mu_{k-1})(\mathbf{h}_{k-1} - \mu_{k-1})^H] + \mathbf{C} \\
&= (1 + |\alpha|^2 + |\alpha|^4 + \cdots + |\alpha|^{2(k-1)})\mathbf{C} \\
&= \frac{1 - |\alpha|^{2k}}{1 - |\alpha|^2}\mathbf{C}.
\end{aligned}$$

- Because $|\alpha| < 1$, the steady-state probability of $\mathbf{h}_k$ (i.e. k$\rightarrow \infty$) is

$$f(\mathbf{h}_k) \sim CN\left(0, \frac{1}{1-|\alpha|^2}\mathbf{C}\right).$$

- Hence the mean vector of $\mathbf{q}(j,i)$ is given as

$$\begin{aligned}
\mathbf{m}_q(j,i) &= E\left[\mathbf{q}(j,i)\right] \\
&= \begin{bmatrix}
\frac{E[r_1]\mathbf{a}_1^*(j)}{\sigma^2} + \alpha\mathbf{C}^{-1}E[\mathbf{h}_0] \\
\frac{E[r_2]\mathbf{a}_2^*(j)}{\sigma^2} + \alpha\mathbf{C}^{-1}\mathbf{G}_1(j)E[\mathbf{q}_1(j)] \\
\vdots \\
\frac{E[r_N]\mathbf{a}_N^*(j)}{\sigma^2} + \alpha\mathbf{C}^{-1}\mathbf{G}_{N-1}(j)E[\mathbf{q}_{N-1}(j)] \\
\frac{E[r_1]\mathbf{a}_1^*(i)}{\sigma^2} + \alpha\mathbf{C}^{-1}E[\mathbf{h}_0] \\
\frac{E[r_2]\mathbf{a}_2^*(i)}{\sigma^2} + \alpha\mathbf{C}^{-1}\mathbf{G}_1(i)E[\mathbf{q}_1(i)] \\
\vdots \\
\frac{E[r_N]\mathbf{a}_N^*(i)}{\sigma^2} + \alpha\mathbf{C}^{-1}\mathbf{G}_{N-1}(i)E[\mathbf{q}_{N-1}(i)]
\end{bmatrix} \\
&= \mathbf{0},
\end{aligned}$$

- Covariance matrix of $\mathbf{q}(j,i)$ is

$$\mathbf{S}_q(j,i) = E\left[\mathbf{q}(j,i)\mathbf{q}^H(j,i)\right],$$

where the (k,l) element in $\mathbf{S}_q(j,i)$ is given as

$$\begin{aligned}
& E[\mathbf{q}_k(n)\mathbf{q}_l^H(m)] \\
= \; & E\Bigg[\left(\frac{r_k\mathbf{a}_k^*(n)}{\sigma^2} + \alpha\mathbf{C}^{-1}\mathbf{G}_{k-1}(n)\mathbf{q}_{k-1}(n)\right)\left(\frac{r_l\mathbf{a}_l^*(m)}{\sigma^2} + \alpha\mathbf{C}^{-1}\mathbf{G}_{l-1}(m)\right. \\
& \left.\mathbf{q}_{l-1}(m)\right)^H\Bigg] \\
= \; & \left(\mathbf{a}_k^T(n)E[\mathbf{h}_k\mathbf{h}_l^H]\mathbf{a}_l^*(m) + E[n_k n_l]\right)\frac{\mathbf{a}_k^*(n)\mathbf{a}_l^T(m)}{\sigma^4} \\
& + \frac{\alpha}{\sigma^2}\mathbf{C}^{-1}\mathbf{G}_{k-1}(n)E[\mathbf{q}_{k-1}(n)\mathbf{h}_l^H]\mathbf{a}_l^*(m)\mathbf{a}_l^T(m) \\
& + \frac{\alpha^*}{\sigma^2}\mathbf{a}_k^*(n)\mathbf{a}_k^T(n)E[\mathbf{h}_k\mathbf{q}_{l-1}^H(m)]\mathbf{G}_{l-1}(m)\mathbf{C}^{-1} \\
& + |\alpha|^2\mathbf{C}^{-1}\mathbf{G}_{k-1}(n)E[\mathbf{q}_{k-1}(n)\mathbf{q}_{l-1}^H(m)]\mathbf{G}_{l-1}(m)\mathbf{C}^{-1}.
\end{aligned}$$

- Since $\mathbf{S}_q(j,i)$ is real and symmetric, it can be represented by two real and symmetric square roots

$$\mathbf{S}_q(j,i) = \mathbf{S}_q^{1/2}(j,i)\mathbf{S}_q^{1/2}(j,i).$$

- And $\mathbf{S}_q^{1/2}(j,i)\mathbf{G}(j,i)\mathbf{S}_q^{1/2}(j,i)$ is also real and symmetric, it can be expressed by its eigenvalues λ_n and eigenvectors $\mathbf{k}_n$ as

$$\mathbf{S}_q^{1/2}(j,i)\mathbf{G}(j,i)\mathbf{S}_q^{1/2}(j,i) = \sum_{n=1}^{2MN} \lambda_n \mathbf{k}_n \mathbf{k}_n^T,$$

- So, $\mathbf{q}^H(j,i)\mathbf{G}(j,i)\mathbf{q}(j,i)$ can be specified as

$$
\begin{aligned}
& \left(\mathbf{S}_q^{-1/2}(j,i)\mathbf{q}(j,i)\right)^H \mathbf{S}_q^{1/2}(j,i)\mathbf{G}(j,i)\mathbf{S}_q^{1/2}(j,i)\left(\mathbf{S}_q^{-1/2}(j,i)\mathbf{q}(j,i)\right) \\
= & \sum_{n=1}^{2MN} \lambda_n |\mathbf{k}_n^T \mathbf{S}_q^{-1/2}(j,i)\mathbf{q}(j,i)|^2 \\
= & \sum_{n=1}^{2MN} \lambda_n |X_n|^2,
\end{aligned}
$$

where $X_n = \mathbf{k}_n^T \mathbf{S}_q^{-1/2}(j,i)\mathbf{q}(j,i)$ is complex Gaussian with mean

$$\mathbf{k}_n^T \mathbf{S}_q^{-1/2}(j,i)E[\mathbf{q}(j,i)] = \mathbf{0},$$

variance

$$\mathbf{k}_n^T \mathbf{S}_q^{-1/2}(j,i)E[\mathbf{q}(j,i)\mathbf{q}(j,i)^H]\mathbf{S}_q^{-1/2}(j,i)\mathbf{k}_n = \mathbf{k}_n^T\mathbf{k}_n = 1,$$

and is independent in n.

- $$\begin{aligned} p_{j|i} &= \Pr\left[\mathbf{q}^H(j,i)\mathbf{G}(j,i)\mathbf{q}(j,i) > \sum_{k=1}^{n}\left(\log\frac{|\mathbf{G}_k(i)|}{|\mathbf{G}_k(j)|}\right)\right] \\ &= \Pr\left[\sum_{l=1}^{\bar{L}}\bar{\lambda}_l\chi^2(2\kappa_l) > \sum_{k=1}^{n}\left(\log\frac{|\mathbf{G}_k(i)|}{|\mathbf{G}_k(j)|}\right)\right] \end{aligned}$$

- Assuming the eigenvalues have been ordered as $\bar{\lambda}_1 > \bar{\lambda}_2 > \ldots > \bar{\lambda}_p > 0 > \bar{\lambda}_{p+1} > \bar{\lambda}_{p+2} > \ldots > \bar{\lambda}_{\bar{L}}$, and $n = \sum_{l=1}^{\bar{L}}\kappa_l$, a closed-form expression of $p_{j|i}$ is

$$\begin{aligned} p_{j|i} &= \Pr\left[\mathbf{q}^H(j,i)\mathbf{G}(j,i)\mathbf{q}(j,i) > \sum_{k=1}^{n}\left(\log\frac{|\mathbf{G}_k(i)|}{|\mathbf{G}_k(j)|}\right)\right] \\ &= \sum_{l=1}^{p}\frac{1}{(\kappa_l-1)!}\left[\frac{\partial^{\kappa_l-1}}{\partial x^{\kappa_l-1}}F_l(x)\right]_{x=\bar{\lambda}_l} \end{aligned}$$

where

$$F_l(x) = x^{q-1}\exp\left[\frac{-\sum_{k=1}^{n}\left(\log\frac{|\mathbf{G}_k(i)|}{|\mathbf{G}_k(j)|}\right)}{2x}\right]\prod_{r=1,r\neq l}^{\bar{L}}(x-\bar{\lambda}_r)^{-\kappa r}.$$

- Define the energy function of simulated annealing algorithm as

$$E = 2^{-K} \sum_{i=1}^{2^K} \sum_{j=1, j \neq i}^{2^K} p_{j|i},$$

which results in searching for a code having minimum average block error rate.

- The perturbation of u_{th} codeword will only affect the pairwise error probabilities of those u_{th}-related pairs, i.e. $(1, u), (2, u), \ldots, (u-1, u), (u+1, u), \ldots, (2^k, u)$ and $(u, 1), (u, 2), \ldots, (u, u-1), (u, u+1), \ldots, (u, 2^k)$.

- The energy function is now defined as

$$E(C') = E(C) - 2^{-K} \sum_{i=1\ldots 2^K, i\neq u, (i,u)\in C} (p_{i|u} + p_{u|i}) + 2^{-K} \sum_{i'=1\ldots 2^K, i'\neq u', (i',u')\in C'} (p_{i'|u'} + p_{u'|i'}).$$

- It saves about $1 - \frac{2\cdot(2^K-1)}{2^K\cdot(2^K-1)} = 1 - 2^{-K+1}$ portion of running time, but the cost is that all pairwise error probabilities must be recorded.

Simulation Results

- The codes we found by our criterion and by [7]'s were first encoded as BPSK signals, suffered fading through the Gauss-Markov channel, and then decoded according to the optimal rule

$$\hat{\mathbf{a}}(\mathbf{r}) = \mathbf{a}(i'), \text{ where } i' = \arg\max_{i \in 2^K} \left(\sum_{k=1}^{N} \left(\mathbf{q}_k(i)^H \mathbf{G}_k(i) \mathbf{q}_k(i) + \log |\mathbf{G}_k(i)| \right) \right).$$

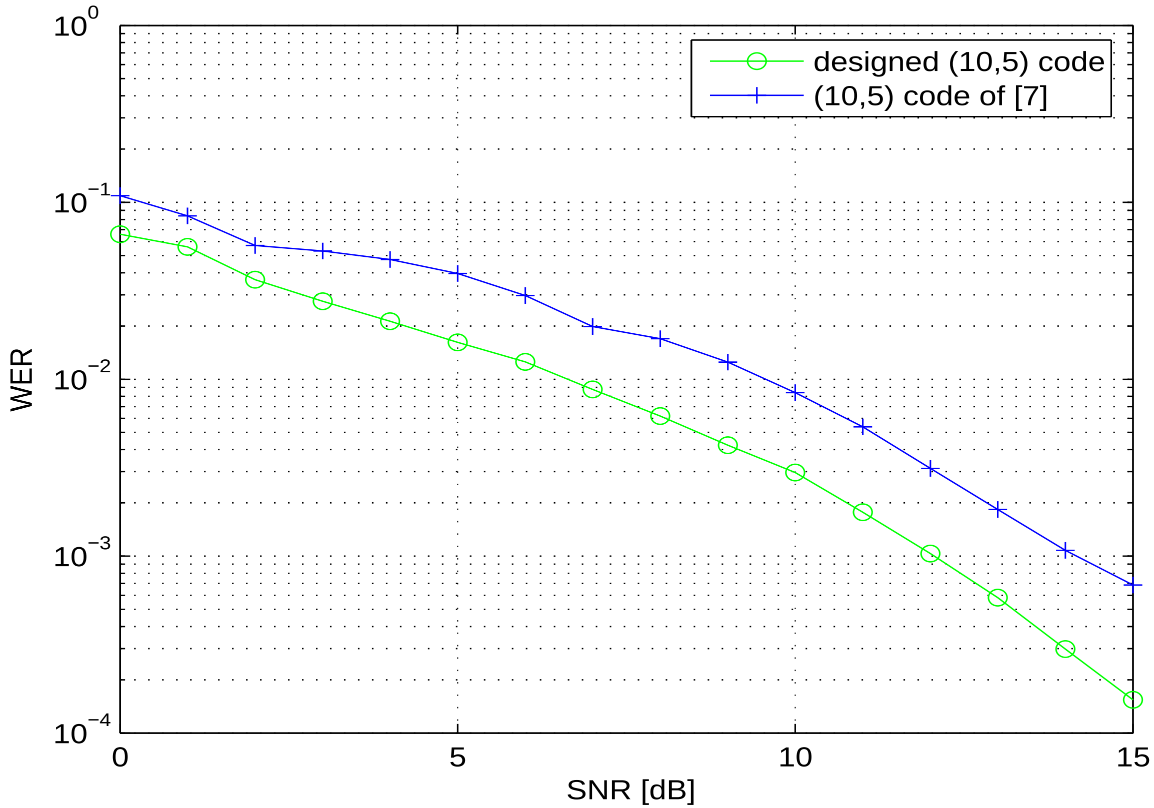

Figure 1: Word error rate for (10,5) code on the Gauss-Markov channel (M=1) with initial channel coefficient h_0 assumed to be known at transmitter and receiver is given as 1, and $h_k = 0.9h_{k-1} + v_k$, where v_k has mean d=0 and covariance $C = 0.001$.

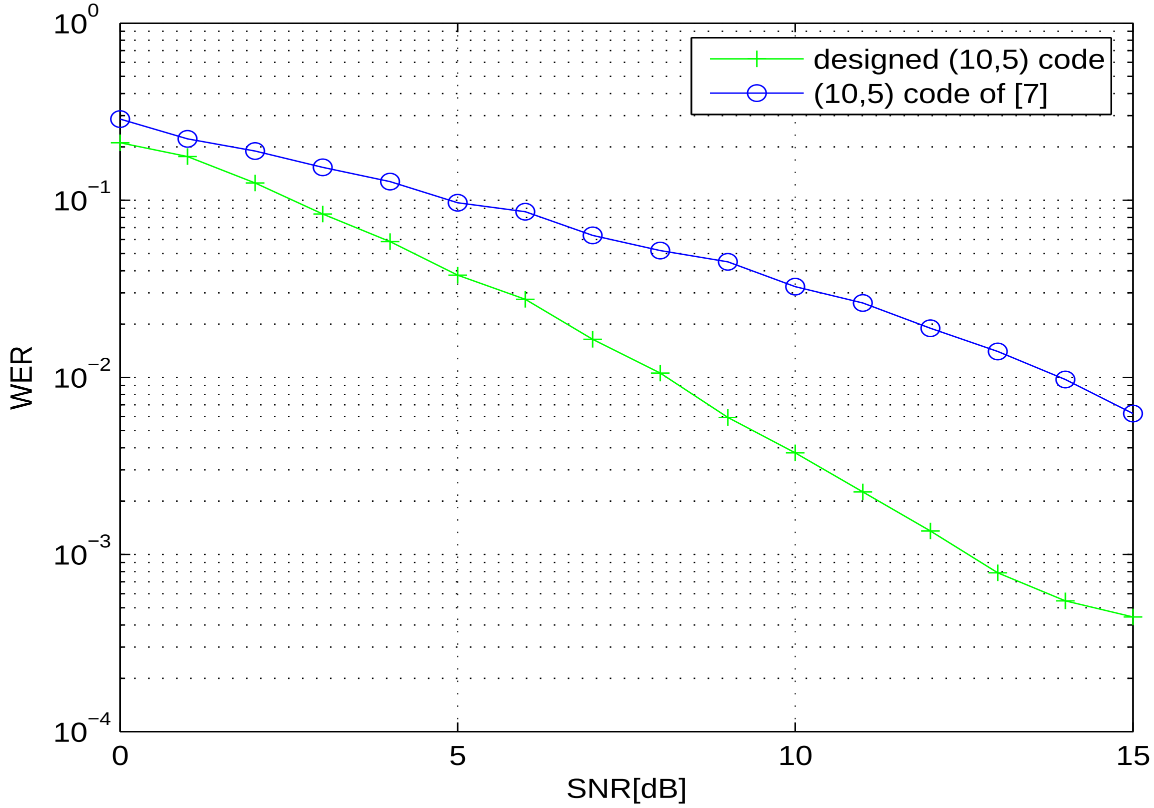

Figure 2: Word error rate for (10,5) code on the Gauss-Markov channel (M=2) with $\mathbf{h}_0$ assumed to be known at transmitter and receiver is given as $\mathbf{1}$, and $\mathbf{h}_k = 0.9\mathbf{h}_{k-1} + \mathbf{v}_k$, where $\mathbf{v}_k$ has mean $\mathbf{d}$=$\mathbf{0}$ and covariance $\mathbf{C}$=0.001$\mathbf{I}$.

Table 1: comparison of complexity

Criterion for Gauss-Markov channel			Criterion of [7]		
	number of multiplication	number of addition		number of multiplication	number of addition
$\mathbf{Q}(j,i)$	24N-16	8N-6	$\mathbf{P}_B(i)$	$2N^2$+12N+13	N^2+8N+4
$\mathbf{G}(j,i)$	54N-40	18N-12	$\mathbf{Q}(j,i)$	0	N^2+2N+1
$\mathbf{S}_q(j,i)$	$\frac{4}{3}N^3+\frac{187}{6}N^2$ $-\frac{241}{6}N+23$	$4N^2$+8N+2	$\mathbf{S}_y(i)$	$2N^2$+9N+7	$2N^2$+6N+4
$\mathbf{S}_q(j,i)\cdot\mathbf{G}(j,i)$	$64N^3$	$64N^3-16N^2$	$\mathbf{S}_y(i)\cdot\mathbf{Q}(j,i)$	$(N+1)^3$	$N(N+1)^2$

- For $N = 10$, the total numbers of multiplications and additions are 68795 and 63124 respectively for our criterion, and 1961 and 1779 respectively for that given in [7].

Conclusion

- We describe the Gauss-Markov channel and its likelihood function which is a recursive form of every prior state.
- Then we derive the pairwise error probability as a function of the criterion for the simulated annealing algorithm.
- Our designed codes provide a coding gain of about 3.5 dB and 6 dB on the Gauss-Markov channels with channel lengths 1 and 2 respectively over those given in [7] at WER$= 10^{-2}$.