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 $\Delta e = \text{energy}(C') - \text{energy}(C)$
        If ($\Delta e < 0$) then $C \leftarrow C'$
        Else with probability $\exp(-\Delta e/T)$ $C \leftarrow C'$
    }
    Until (several energy drops or too many iterations)
    Lower temperature $T \leftarrow \alpha T$
}

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

$$r_k = a_k^T h_k + n_k,$$

where $a_k^T = [a_k, a_{k-1}, \cdots, a_{k-M+1}]$ and $M$ is the channel length indicating how far the past transmitted bits affect the received bit in this time instance. $h_k$ 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 $\sigma^2$.

• For Gauss-Markov channel, $h_k = \alpha h_{k-1} + v_k$ at time $k$, where $v_k$ is a complex white Gaussian with mean $d$ and covariance $C$. 
• Let $H = [h_1, h_2, \cdots, h_N]$ be the matrix of channel coefficient vectors representing by columns, and let $A = [a_1, \cdots, a_N]^T$ represent the matrix of transmitted data representing by rows. Also let $r = [r_1, \cdots, r_N]^T$ represent the received vector, while $n = [n_1, \cdots, n_N]^T$ be the vector of noise from time 1 to time $N$.

• Assuming the mean of $v_k$ is zero and initial channel coefficient $h_0$ is known, we can derive the likelihood function of Gauss-Markov channel as

$$f(r|A) = E[f(r|H, A)] = \int_H f(r|H, A)f(H)dH = \prod_{k=1}^{N} e^{-\frac{|r_k - a_k^T h_k|^2}{\sigma^2}} f(h_1) \prod_{k=2}^{N} f(h_k|h_{k-1})dH.$$
After marginalizing the channel coefficient distribution,

\[
f(r | A) = \prod_{k=1}^{n} e^{q_k^H G_k q_k | G_k|}.\]

and

\[
\begin{align*}
G_1^{-1} &= \frac{a_1^* a_1^T}{\sigma^2} + (1 + |\alpha|^2) C^{-1} \\
q_1 &= \frac{r_1 a_1^*}{\sigma^2} + \alpha C^{-1} h_0.
\end{align*}
\]

\[
\begin{align*}
G_k^{-1} &= \frac{a_k^* a_k^T}{\sigma^2} + C^{-1} + |\alpha|^2 (C^{-1} - C^{-1} G_{k-1} C^{-1}) \\
q_k &= \frac{r_k a_k^*}{\sigma^2} + \alpha C^{-1} G_{k-1} q_{k-1}.
\end{align*}
\]

\[
\begin{align*}
G_N^{-1} &= \frac{a_N^* a_N^T}{\sigma^2} + C^{-1} - |\alpha|^2 C^{-1} G_{N-1} C^{-1} \\
q_N &= \frac{r_N a_N^*}{\sigma^2} + \alpha C^{-1} G_{N-1} q_{N-1}.
\end{align*}
\]
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

$$P_e \triangleq \Pr(\hat{x} \neq x) = 2^{-K} \sum_i \Pr(\hat{x} \neq x(i) | x(i) \text{ is transmitted})$$

$$\leq 2^{-K} \sum_i \sum_{j \neq i} p_{j|i},$$
• The pairwise error probability

\[
p_{j|i} = \Pr \left[ \log \frac{\Pr[r|x(j)]}{\Pr[r|x(i)]} > 0 \right] + \frac{1}{2} \Pr \left[ \log \frac{\Pr[r|x(j)]}{\Pr[r|x(i)]} = 0 \right].
\]

\[
= \Pr \left[ \log \frac{\prod_{k=1}^{N} e^{q_k(j)^H G_k(j) q_k(j)} |G_k(j)|}{\prod_{k=1}^{N} e^{q_k(i)^H G_k(i) q_k(i)} |G_k(i)|} > 0 \right]
\]

\[
+ \frac{1}{2} \Pr \left[ \log \frac{\prod_{k=1}^{N} e^{q_k(j)^H G_k(j) q_k(j)} |G_k(j)|}{\prod_{k=1}^{N} e^{q_k(i)^H G_k(i) q_k(i)} |G_k(i)|} = 0 \right]
\]

\[
= \Pr \left[ \sum_{k=1}^{N} \left( q_k(j)^H G_k(j) q_k(j) - q_k(i)^H G_k(i) q_k(i) \right) > \sum_{k=1}^{N} \left( \log \frac{|G_k(i)|}{|G_k(j)|} \right) \right]
\]

\[
+ \frac{1}{2} \Pr \left[ \sum_{k=1}^{N} \left( q_k(j)^H G_k(j) q_k(j) - q_k(i)^H G_k(i) q_k(i) \right) = \sum_{k=1}^{N} \left( \log \frac{|G_k(i)|}{|G_k(j)|} \right) \right].
\]
Now we reform the q-matrix and G-matrix as

$$
q(j,i) =
\begin{bmatrix}
q_1(j) \\
q_2(j) \\
\vdots \\
q_N(j) \\
q_1(i) \\
q_2(i) \\
\vdots \\
q_N(i)
\end{bmatrix},
G(j,i) =
\begin{bmatrix}
G_1(j) & 0 & \ldots & 0 & 0 & \ldots & 0 \\
0 & G_2(j) & \ldots & 0 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & G_N(j) & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 & -G_1(i) & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \ddots \\
0 & 0 & \ldots & 0 & 0 & \ldots & -G_N(i)
\end{bmatrix},
$$

$$
p_{j|i} = \Pr \left[ q^H(j,i)G(j,i)q(j,i) > \sum_{k=1}^{N} \left( \log \frac{|G_k(i)|}{|G_k(j)|} \right) \right]
+ \frac{1}{2} \Pr \left[ q^H(j,i)G(j,i)q(j,i) = \sum_{k=1}^{N} \left( \log \frac{|G_k(i)|}{|G_k(j)|} \right) \right],
$$

where \( q(j,i) \) 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 $\mu_k$ 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

$$
\mathbf{S}_k = E[(\mathbf{h}_k - \mu_k)(\mathbf{h}_k - \mu_k)^H] \\
= E[\mathbf{h}_k \mathbf{h}_k^H] - \mu_k E[\mathbf{h}_k^H] - E[\mathbf{h}_k] \mu_k^H + \mu_k \mu_k^H \\
= E[(\alpha \mathbf{h}_{k-1} + \mathbf{v}_k)(\alpha \mathbf{h}_{k-1} + \mathbf{v}_k)^H] - |\alpha|^2 \mu_{k-1} E[\mathbf{h}_{k-1}^H] - |\alpha|^2 E[\mathbf{h}_{k-1}] \mu_{k-1}^H \\
+ |\alpha|^2 \mu_{k-1} \mu_{k-1}^H \text{(because } \mathbf{h}_k = \alpha \mathbf{h}_{k-1} + \mathbf{v}_k, \mu_k = \alpha \mu_{k-1}) \\
= |\alpha|^2 E[\mathbf{h}_{k-1}^H \mathbf{h}_{k-1}] + E[\mathbf{v}_k \mathbf{v}_k^H] - |\alpha|^2 \mu_{k-1} E[\mathbf{h}_{k-1}^H] \\
- |\alpha|^2 E[\mathbf{h}_{k-1}] \mu_{k-1}^H + |\alpha|^2 \mu_{k-1} \mu_{k-1}^H \\
= |\alpha|^2 E[(\mathbf{h}_{k-1} - \mu_{k-1})(\mathbf{h}_{k-1} - \mu_{k-1})^H] + \mathbf{C} \\
= (1 + |\alpha|^2 + |\alpha|^4 + \cdots + |\alpha|^{2(k-1)}) \mathbf{C} \\
= \frac{1 - |\alpha|^{2k}}{1 - |\alpha|^2} \mathbf{C}.
$$
• Because $|\alpha| < 1$, the steady-state probability of $h_k$ (i.e. $k \to \infty$) is

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

• Hence the mean vector of $q(j,i)$ is given as

$$m_q(j,i) = E[q(j,i)] = \begin{bmatrix}
\frac{E[r_1]a_1^*(j)}{\sigma^2} + \alpha C^{-1} E[h_0] \\
\frac{E[r_2]a_2^*(j)}{\sigma^2} + \alpha C^{-1} G_1(j) E[q_1(j)] \\
\vdots \\
\frac{E[r_N]a_N^*(j)}{\sigma^2} + \alpha C^{-1} G_{N-1}(j) E[q_{N-1}(j)] \\
\frac{E[r_1]a_1^*(i)}{\sigma^2} + \alpha C^{-1} E[h_0] \\
\frac{E[r_2]a_2^*(i)}{\sigma^2} + \alpha C^{-1} G_1(i) E[q_1(i)] \\
\vdots \\
\frac{E[r_N]a_N^*(i)}{\sigma^2} + \alpha C^{-1} G_{N-1}(i) E[q_{N-1}(i)]
\end{bmatrix} = 0,$$
• Covariance matrix of \( q(j, i) \) is

\[
S_q(j, i) = E [q(j, i)q^H(j, i)],
\]

where the \((k, l)\) element in \( S_q(j, i) \) is given as

\[
E[q_k(n)q_l^H(m)] = E \left[ \left( \frac{r_k a_k^*(n)}{\sigma^2} + \alpha C^{-1} G_{k-1}(n) q_{k-1}(n) \right) \left( \frac{r_l a_l^*(m)}{\sigma^2} + \alpha C^{-1} G_{l-1}(m) \right) \right] q_{l-1}(m))^H \]

\[
= \left( a_k^T(n) E[h_k h_l^H] a_l^*(m) + E[n_k n_l] \right) \frac{a_k^*(n)a_l^T(m)}{\sigma^4} + \frac{\alpha}{\sigma^2} C^{-1} G_{k-1}(n) E[q_{k-1}(n)h_l^H] a_l^*(m)a_l^T(m)
\]

\[
+ \frac{\alpha^*}{\sigma^2} a_k^*(n)a_k^T(n) E[h_k q_{l-1}^H(m)] G_{l-1}(m) C^{-1}
\]

\[
+ |\alpha|^2 C^{-1} G_{k-1}(n) E[q_{k-1}(n)q_{l-1}^H(m)] G_{l-1}(m) C^{-1}.
\]
• Since $S_q(j, i)$ is real and symmetric, it can be represented by two real and symmetric square roots

$$S_q(j, i) = S_q^{1/2}(j, i)S_q^{1/2}(j, i).$$

• And $S_q^{1/2}(j, i)G(j, i)S_q^{1/2}(j, i)$ is also real and symmetric, it can be expressed by its eigenvalues $\lambda_n$ and eigenvectors $k_n$ as

$$S_q^{1/2}(j, i)G(j, i)S_q^{1/2}(j, i) = \sum_{n=1}^{2MN} \lambda_n k_n k_n^T,$$
• So, \( q^H(j, i) G(j, i) q(j, i) \) can be specified as

\[
\left( S_q^{-1/2}(j, i) q(j, i) \right)^H S_q^{1/2}(j, i) G(j, i) S_q^{1/2}(j, i) \left( S_q^{-1/2}(j, i) q(j, i) \right)
= \sum_{n=1}^{2MN} \lambda_n |k_n^T S_q^{-1/2}(j, i) q(j, i)|^2
= \sum_{n=1}^{2MN} \lambda_n |X_n|^2,
\]

where \( X_n = k_n^T S_q^{-1/2}(j, i) q(j, i) \) is complex Gaussian with mean

\[
k_n^T S_q^{-1/2}(j, i) E[q(j, i)] = 0,
\]

variance

\[
k_n^T S_q^{-1/2}(j, i) E[q(j, i) q(j, i)^H] S_q^{-1/2}(j, i) k_n = k_n^T k_n = 1,
\]

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

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

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

where

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

\[(1, u), (2, u), \ldots, (u - 1, u), (u + 1, u), \ldots, (2^k, u) \text{ and} \]

\[(u, 1), (u, 2), \ldots, (u, u - 1), (u, u + 1), \ldots, (u, 2^k).\]
The energy function is now defined as

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

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

\[ \hat{a}(r) = a(i'), \quad \text{where} \quad i' = \arg \max_{i \in 2^K} \left( \sum_{k=1}^{N} (q_k(i))^H G_k(i) q_k(i) + \log |G_k(i)| \right). \]
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 $h_0$ 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 $C=0.001I$.
Table 1: comparison of complexity

<table>
<thead>
<tr>
<th>Criterion for Gauss-Markov channel</th>
<th>Criterion of [7]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>number of</td>
</tr>
<tr>
<td></td>
<td>multiplication</td>
</tr>
<tr>
<td>( Q(j, i) )</td>
<td>24N-16</td>
</tr>
<tr>
<td>( G(j, i) )</td>
<td>54N-40</td>
</tr>
<tr>
<td>( S_q(j, i) )</td>
<td>( \frac{4}{3} N^3 + \frac{187}{5} N^2 )</td>
</tr>
<tr>
<td>( S_q(j, i) \cdot G(j, i) )</td>
<td>( 64N^3 )</td>
</tr>
<tr>
<td></td>
<td>( S_y(i) \cdot Q(j, i) \cdot (N+1)^3 )</td>
</tr>
<tr>
<td>( P_B(i) )</td>
<td>( 2N^2+12N+13 )</td>
</tr>
<tr>
<td>( Q(j, i) )</td>
<td>0</td>
</tr>
<tr>
<td>( S_y(i) )</td>
<td>( 2N^2+9N+7 )</td>
</tr>
</tbody>
</table>

- 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= $10^{-2}$. 