Analysis of Performance of Linear Analog Codes

Yang Liu, Jing Li (Tiffany) Kai Xie
Electrical and Computer Engineering Department,
Lehigh University, Bethlehem, PA 18015
Email: {yal210, jingli, kax205}@ece.lehigh.edu

Abstract—In this paper we carefully study the MSE performance of the linear analog codes. We have derived a lower bound of the MSE performance under Likelihood(ML) and Linear Minimal Mean Square Error(LMMSE) decoding criteria respectively. It is proved in this essay that a kind of linear analog codes called unitary codes can simultaneously achieve both of these two bounds. At the same time, we compare the obtained linear analog codes’ MSE bounds with the performance of some existing nonlinear codes. The results showed that linear analog codes are actually not very satisfying and convinced us that more concerns should be cast onto the nonlinear class in order to find powerful analog codes.

I. INTRODUCTION

Digital signal processing technology has been developed so successfully and its application has entered every corner of our life and changed the world. Since in real life numerous signals we are handling come from analog sources, a typical digital signal processing system converts these signals from the form of analog into digital by sampling and quantization. Sampling theory tells us that the original signal, continuous in time, can be perfectly recovered from the discrete sampled sequence with a sampling frequency higher than Nyquist frequency. However the quantization procedure will introduce permanent loss named as quantization noise, which usually happens in the analog-to-digital(ADC) by rounding. If performing inadequate level quantization, even if there occurs little error in signal processing procedure(for example the powerful error correction codes like turbo or LDPC codes may help us to achieve this), the performance of the system will be dominantly deteriorated by the quantization noise. Although it is possible to suppress quantization noise increasing the number of the quantization levels, this will undoubtedly require a higher transmitting rate, or equivalently an expansion in bandwidth in communication. Thus the tradeoff between the quantization levels and expanded bandwidth is one considering in communication system design. At the same time, random attenuation and noise corruption usually equal-probably happen to each transmitting bits, which carry unequal significance for the signal recovery. Thus the problem of how to evaluate the importance of each bit and how to balance the protection on the most and least important bits can usually be a very hard optimization problem to solve.

In fact, the commonly used procedure of sampling, quantization and then processing is not a native way for the signal processing. Take the conventional digital communication system as an example, the analog signal enters the digital field through ADC, which introduces permanent performance loss; the modulator brings digital signal back into analog world and transmitted it. Since the signal is turned into digital and then turned back into analog, why should not we do it straightly in the analog field? After all this was the way that communication operated in early days. At the same time, the digital error correction coding(DECC) techniques contribute significantly to improving the system performance. Error correction coding is thus a strategy that introduces copies of the information bits in a clever way and enables the system to detect or correct some error patterns. However no constraints have been imposed on this smart technique that it can only be implemented in digital field. Subsequently, a natural idea come up to our mind that can we process the analog signal directly and simultaneously design some error correction code suitable for analog signals? Actually this idea of analog error correction code(AECC), or analog code for short, is not a new idea and can be traced back to some pioneering works in 1980’s. Marshall proposed the concept of real number code [1] and Wolf, independently, used another name analog code in [2].

The encoding procedure of analog code is usually performed by some mapping function which takes information signal as input and outputs “codeword” signal. According to the linearity of the encoding mapping function, analog codes can be classified into two categories: linear analog codes and nonlinear analog codes. Although the research work on AECC started relatively late and until now it has not yet obtained most wide concerns, people still have succeeded in finding some applicable analog codes in both categories.

For the linear class, some codes which have elegant mathematical expression were first found in early years. In reference [1], [2] and [3] discrete Fourier transform(DFT) code was designed, which has its generator matrix composed of a subset of rows of a Fourier transform matrix. With very similar generator matrix structure, discrete cosine transform(DCT) code was proposed by J.L.Wu et. al. in [4]. For a further step, the DCT code was extended to a generalized version called discrete sine transform(DST) code in [7]. The several above linear analog codes all have their generator matrix as Vandermonde matrix, thus are called BCH-like analog codes and can borrow the classical Berlekamp-Massey algorithm or Forney algorithm to implement decoding.

For the nonlinear analog codes, a subclass of the codes named as chaotic analog code(CAC) has been cast special interests, which utilizes chaotic functions to transform the original analog signals. Chaotic functions are named after their unusual property of the sensitivity to the initial states.
Even a small perturbation in the starting point can result in huge difference in the output. This property has been vividly described as “butterfly effect”. Actually in many cases the butterfly effect of the chaotic functions is usually undesirable in practice because it destabilizes the system. However from the viewpoint of extending the codewords’ distance to achieve good error protection ability, this effect is cherishable in error correction code design. This ingenious idea was first proposed in [5] by Chen and Wornell. The constructed a chaotic analog code using the simple chaotic function named tent map code. Though the tent map code’s performance is still far below its digital competitors, it successfully demonstrated feasibility of exploiting the chaotic map to achieve the aim of error correction. Based on that, in [6] Xie et. al. introduced the key thinking of symmetric protection and obtained the chaotic analog turbo(CAT) code. CAT code achieves significant improvement in performance compared to tent map code and for the first time analog code is almost comparable with the conventional digital error correction codes. Also Liu et. al. also applied analog code in the scenario of image transmission and the numerical results showed that under the equal total bandwidth, the analog code can outperform the digital turbo code.

Due to the linearity, linear analog codes are desirable because they are simple for analysis, easy to implement and low in complexity. Though existing nonlinear analog codes usually enjoy favorable performance over the linear ones but their nonlinear mapping function makes it very difficult to develop an efficient decoding algorithm. Moreover, it is already known that for the conventional digital error correction codes, the best codes people have found thus far, the turbo codes and LDPC codes, are both linear. Then, being somewhat empirical, we are very curious about the following problem: is it possible to find some linear analog codes which enjoy satisfying performance? In fact, Xie and Li have explored this problem in their recent paper [7]. They answered this problem from the viewpoint of distance expansion and have got some meaningful results. Here in this paper, we try to study this question in a more rigorous way and give out a much clearer answer.

II. Signal and System Model

In this section, we clarify the signal model and encoding procedure of linear analog coding system, on which the following sections are based.

A. Signal Model

Assume that the original information sequence entering the analog system is discrete in time and continuous in values, which typically can be obtained by sampling an analog signal source. Following the convention of DECC, we still call each element of the input sequence a “bit” and denote this information bit sequence as \( \mathbf{u} \), \( \mathbf{u} \in \mathbb{C}^{K \times 1} \). To make problem clear and easy for analysis, we make the following assumptions on the input signal \( \mathbf{u} \).

A1) Each coordinate \( u_i \) of the original signal sequence \( \mathbf{u} \) follows i.i.d. distribution with distribution density \( p(u) \).

A2) Any coordinate \( u_i \) of the original signal sequence \( \mathbf{u} \) has zero expectation, i.e.

\[
E(u) = \int u p(u) du = 0 \tag{1}
\]

Denote \( D(u_i) = E(u_i^2) - E^2(u_i) = D_u - 0 = D_u \) as the average energy of each coordinate of \( \mathbf{u} \).

Actually the above two assumptions are not very stringent constraints. For example in digital communication we often assume that the input bit has independent distribution, which can be achieved by source coding. On the other hand if the expectation of each input bit is averaged at nonzero constant, we can just subtract this constant from the original signal and make it satisfy the assumption A2).

B. System Model

Since the linear analog codes perform linear mapping function on information bits to generate the codewords, we can represent this procedure in a linear equation mathematically. Denote the linear mapping function as a matrix \( \mathbf{G} \in \mathbb{C}^{K \times N} \) and the codeword as \( \mathbf{v} \in \mathbb{C}^{N \times 1} \). Then analog encoding procedure can be written as:

\[
\mathbf{v} = \mathbf{G}^H \mathbf{u} \tag{2}
\]

Still, following the terminology in conventional coding theory, matrix \( \mathbf{G} \) is called generator matrix and should have \( K \) mutually independent rows, \( K \leq N \). In this paper, we focus on additive white Gaussian noise(AWGN) channel, thus the received \( \mathbf{r} \) can be written as:

\[
\mathbf{r} = \mathbf{v} + \mathbf{n} = \mathbf{G}^H \mathbf{u} + \mathbf{n} \tag{3}
\]

Assume that each coordinate of \( \mathbf{n} \in \mathbb{C}^{N \times 1} \) follows a i.i.d. complex Gaussian distribution, whose covariance \( \mathbf{R}_n \) of \( \mathbf{n} \) is written as:

\[
\mathbf{R}_n = \sigma^2 \mathbf{I}_{N \times N} \tag{4}
\]
Utilize the above assumption $A1)$ and $A2)$ we can easily calculate that:

$$E(u^H v) = E(\sum_{i=1}^{N} |v_i|^2)$$

$$= E(u^H G G^H u)$$

$$= E(\sum_{i=1}^{N} (\sum_{j=1}^{K} g_{ji}^* u_j)(\sum_{j=1}^{K} g_{ji} u_j^*))$$

$$= E(\sum_{i=1}^{N} \sum_{j=1}^{K} |g_{ji}|^2 |u_j|^2 + \sum_{j\neq k}^{K} g_{ji}^* g_{jk} u_j u_k^*)$$

$$= \sum_{i=1}^{N} (\sum_{j=1}^{K} |g_{ji}|^2 E(|u_j|^2)) + 0$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{K} |g_{ji}|^2 D_u$$

$$= \|G\|_F^2 D_u$$

(5)

where the notation $\| \cdot \|_F$ means 2-norm of a matrix, which is also known as Frobenius norm. Though the generator matrix $G$ is up to our choice, we cannot set $G$ to any large for the power constraint in real practice. For a given communication system, the transmission power is usually constant for each information bit to be transmitted, which we denote as $E_b$. Thus the total transmitted energy for the codeword $v$ should have the following constraint:

$$E(v^H v) = \|G\|_F^2 D_u = C = K E_b$$

(6)

We can summarize the statement of system and channel into assumptions as follows:

$A3)$ The average transmitted energy of codeword $v$ is constant as stated in (6).

$A4)$ The noise at any time $n_i$ follows i.i.d. circular symmetric complex Gaussian distribution, $n_i \sim \mathcal{CN}(0, \sigma^2)$, with $\sigma^2/2 = \sigma_x^2 = \sigma_y^2$.

III. OPTIMAL LINEAR ANALOG CODES FOR LMMSE AND ML DECODING SCHEME

For DECC, bit error rate (BER) is used to measure the digital code’s performance. While for signals with continuous values, because the set of potential valid information bits is usually one or multiple continuous regions, mean square error (MSE) is a suitable metric to evaluate the performance of AECC. Mathematically MSE per information bit can be defined as following:

$$MSE = \frac{E\{|\hat{u} - u|^2\}}{K}$$

(7)

where $\hat{u}$ denotes the estimate of original information bits $u$ and $E\{ \cdot \}$ should average over all the support regions of the information bits $u$. There exist different criteria to perform the estimate of information bits (or it should be called decoding criteria in our scenario of analog coding). The maximum likelihood (ML) is a commonly used rule, which is also considered in the existing literatures on analog codes mentioned above. Besides, borrowing from the method of signal estimation, we also consider linear minimal mean square error (LMMSE) criterion which is widely used in MIMO estimation. In the rest of this section, we will discuss the optimal linear analog codes under these two decoding schemes.

A. Optimal Linear Analog Codes for LMMSE Decoding Scheme

LMMSE decoding scheme means that for a given generator matrix $G$, the receiver performs a linear operation to accomplish decoding, which can minimize the MSE per information bit. Assume that linear decoder is $A$, then the estimated $u$ can be expressed as $\hat{u} = Ar$. Substituting the equation (3), (4) and assumption $A2)$ into (7), the mean square error can be calculated as:

$$MSE = \frac{1}{K} E\{|\hat{u} - u|^2\}$$

$$= \frac{1}{K} tr(E\{|\hat{u} - u|^2\})$$

$$= \frac{1}{K} tr(E\{|\hat{u} - u|^2\})$$

$$= \frac{1}{K} D_u \{tr(\{(AG)^H(I)\}) + \sigma^2 tr(\{A^2\})\}$$

(8)

For any given generator matrix $G$, the optimal LMMSE receiver $A_{LMMSE}$ can be determined by reducing $\frac{\partial MSE}{\partial A} = 0$, where $A^*$ is the conjugate transpose of matrix $A$:

$$A_{LMMSE} = (GG^H + \frac{\sigma^2}{D_u}I)^{-1}G$$

(9)

The above equation tells us for any given generator matrix $G$, its corresponding LMMSE decoder is a function of $G$ and can be determined by (9). Consequently the LMMSE optimal estimation of $u_{LMMSE}$ can be determined as

$$\hat{u}_{LMMSE} = A_{LMMSE} r = (GG^H + \frac{\sigma^2}{D_u}I)^{-1}Gr$$

(10)

Based on (10), the minimal mean square error

$$MSE_{LMMSE} = \frac{D_u tr(I + \frac{D_u}{\sigma^2}GG^H)^{-1}}{K}$$

(11)

From the analysis in the previous section we know that matrix $G$ should satisfy the equation (6). At the same time we notice the fact that

$$\|G\|_F^2 = tr(G^H G) = tr(GG^H)$$

(12)

Thus the constraint in (6) can be equivalently expressed as:

$$tr(GG^H) = \frac{KE_b}{D_u}$$

(13)

Then the problem of finding the optimal linear analog code in the meaning of minimal MSE per information bit by LMMSE
decoding can be formulated as the following optimization problem:
\[
\begin{align*}
\min_G & \quad \text{MSE}_{LMMSE} = \frac{D_u}{K} (I + \frac{D_u}{\sigma^2} GG^H)^{-1} \\
\text{s.t.} & \quad \text{tr}(GG^H) = \frac{KE_b}{D_u}
\end{align*}
\] (14)

According to Hadamard’s inequality, whose proof can be found in Appendix I of [9], for any \(N \times N\) positive semidefinite matrix \(B\), we have
\[
\text{tr}(B^{-1}) \geq \sum_{i=1}^{N} \frac{1}{B_{ii}}
\] (16)

where \(B_{ii}\) is the \(i\)-th diagonal element of matrix \(B\). The above inequality (16) achieves equality if and only if \(B\) is diagonal. Notice that \(GG^H\) is positive semidefinite. Then utilize the Hadamard’s inequality, we have
\[
\frac{D_u \text{tr}}{K} (I + \frac{D_u}{\sigma^2} GG^H)^{-1} \geq \frac{D_u \text{tr}}{K} \sum_{i=1}^{K} \frac{1}{1 + \frac{D_u}{\sigma^2} |GG^H|_{ii}}
\] (17)

Equation (17) holds if and only if \(I + \frac{D_u}{\sigma^2} GG^H\) is diagonal matrix, or equivalently \(GG^H\) is diagonal. We denote the diagonal matrix \(Q = GG^H = diag\{q_1, q_2, \ldots, q_K\}\). The problem (13) converts to:
\[
\begin{align*}
\min_{\{q_i\}} & \quad \text{MSE}_{MMSE} = \sum_{i=1}^{K} \frac{D_u}{1 + \frac{D_u}{\sigma^2} q_i} \\
\text{s.t.} & \quad \text{tr}(GG^H) = K \sum_{i=1}^{K} q_i = \frac{KE_b}{D_u} \\
& \quad q_i > 0, \forall i \in \{1, 2, \ldots, K\}
\end{align*}
\] (18)

This is an easy optimization problem and we can figure that the at achieves optimality when
\[
q_1 = q_2 = \cdots = q_K = \frac{KE_b}{KD_u} = \frac{E_b}{D_u}
\] (19)

the objective \(\text{MSE}_{LMMSE}\) achieves the optimal value
\[
\text{MSE}_{LMMSE}^{\text{min}} = \sum_{i=1}^{K} \frac{D_u}{K(1 + \frac{D_u}{\sigma^2} q_i)} = \frac{D_u}{1 + \frac{D_u}{\sigma^2}}
\] (20)

Thus we have obtained the lower bound of the MMSE decoding of linear analog code under AWGN channel as shown in (22).

Now let us check the conditions for \(G\) to achieve this lower bound. Note that the equality (17) stands when \(GG^H\) is diagonal, which means the rows of \(G\) are mutually orthogonal. At the same time the optimal solution in (19) have equal \(q_i\). This means each row of \(G\) has equal energy(norm). We conclude that the analog linear codes achieving lower bound MSE in (22) under AWGN channel have unitary generator matrix \(G\), i.e. each row of \(G\) is drawn from some unitary matrix up to some common scale factor. Thus we can call this optimal code as unitary code.

### B. Optimal Linear Analog Codes for ML Decoding Scheme

For a received vector \(v\), maximum likelihood decoding criterion takes the vector \(u\) which maximizes the likelihood probability as the estimation. Mathematically it can be presented as:
\[
\hat{u} = \arg \max_u P(r|u)
\] (23)

As a result of the i.i.d. complex Gaussian distribution mentioned in assumption A4, the equation (23) can be equivalently reduced to minimizing the squared distance in the exponential part as the following:
\[
\hat{u} = \arg \min_u \|r - GG^H u\|^2
\] (24)

Let us denote the squared distance \(\|r - GG^H u\|^2\) as \(J(u)\) and expand it as following:
\[
J(u) = \|r - GG^H u\|^2 = (r - GG^H u)^H (r - GG^H u) = u^H GG^H u - r^H G^H u - u^H G r + r^H r
\] (25)

By reducing \(\frac{\partial J(u)}{\partial u} = 0\), we can obtain the optimal ML estimator as following:
\[
\hat{u}_{ML} = (GG^H)^{-1} G r
\] (26)

Thus the MSE per information bit of ML decoder \(\text{MSE}_{ML}\) can be calculated as:
\[
\text{MSE}_{ML} = \frac{1}{K} E\{ (\hat{u} - u)(\hat{u} - u)^H \} = \frac{1}{K} \text{tr}(E\{ (\hat{u} - u)(\hat{u} - u)^H \}) = \frac{1}{K} \text{tr}(E\{ ((GG^H)^{-1} G r - u)((GG^H)^{-1} G r - u)^H \}) = \frac{\sigma^2}{K} \text{tr}((GG^H)^{-1})
\] (27)

We can design the generator matrix \(G\) to minimize the MSE per information bit when using ML decoder. This is also a code design problem and can be formulated as an optimization problem as follows:
\[
\begin{align*}
\min_G & \quad \text{MSE}_{ML} = \frac{\sigma^2}{K} \text{tr}((GG^H)^{-1}) \\
\text{s.t.} & \quad \text{tr}(GG^H) = \frac{KE_b}{D_u}
\end{align*}
\] (28)

We can utilize the (16) again
\[
\frac{\sigma^2}{K} \text{tr}((GG^H)^{-1}) \geq \sigma^2 \sum_{i=1}^{K} \frac{1}{|GG^H|_{ii}}
\] (29)

As mentioned, (30) achieves the equality if and only if \(GG^H\) is diagonal.

Thus from here, the analysis for the problem in (28) follows almost the same line for the case of LMMSE in previous
subsection. We omit this part and give out directly the minimal objective $MSE_{ML}$ as:

$$MSE_{ML}^\text{min} = \frac{D_u \sigma^2}{E_b}$$  \tag{31}$$

which is obtained by having equation $[30]$ to hold with equality and the diagonal matrix $Q = GG^H = \text{diag}\{q_1, q_2, \ldots, q_K\}$ to satisfy the solution in $[21]$. 

C. Coincidence of the Optimal Linear Codes by LMMSE and ML Decoding Scheme

Summarizing the results in this section and previous section, we can conclude:

1) Unitary generator matrix $G$ up to some scaler coefficient can simultaneously achieve the lower bound of MSE per information bit of LMMSE decoder and ML decoder for analog linear code on AWGN channel.

2) Comparing the lower bounds in (22) and (31), we can see that LMMSE decoder can achieve lower MSE per information bit.

IV. NUMERICAL RESULTS

In this section, we provide simulation results to confirm the conclusions obtained in previous two sections and make a comparison between the linear and nonlinear analog codes.

A. Numerical Results of Different Analog Linear Codes

Here in our experiments, all the tested analog linear codes are $(60,30)$ codes. The input signal source are complex Gaussian variables following $CN(0,2)$, with $\sigma^2_x = \sigma^2_y = 1$. Thus we can readily calculate that $D_u = 2$ and the average transmitted energy per information bit $E_b = 4$. Noise in channel are also circular symmetric complex Gaussian variables, whose variance is determined by specific $SNR$. The performance is measured by the mean square error per information, illustrated in the figures in a $log_2(\cdot)$ form. In each group, we randomly generate five generator matrices $G$ satisfying some specific designed properties. For the convenience of comparison, in the result of each group, we give out the MSE lower bound for both ML and LMMSE decoding scheme.

1) In the first group, we randomly generate 5 unitary linear analog codes. These $G$s have mutually orthogonal rows and each row has identical norm.

2) In the second group, we randomly generate 5 $G$s, whose rows are mutually orthogonal but unequal in norm.

3) In the third group, we randomly generate 5 $G$s, each of which has linear independent but nonorthogonal rows.

4) In the forth group, we randomly generate 5 DFT codes.

5) In the fifth group, we randomly generate 5 DCT codes.

Obviously the randomly generated unitary analog linear codes are all bound-reaching codes. DFT and DCT codes are also bound reaching and actually they can be regarded as special subclass of the unitary codes. While for non-unitary linear analog codes, its performance typically can not reach the lower bound for both ML and LMMSE decoding schemes.

B. Linear vs Nonlinear

We also make a comparison with the the linear analog code bound obtained above with the performance of some existing nonlinear analog codes, as illustrated in Fig[6] At low SNR, we can see that LMMSE decoding scheme enjoys advantage over the several plotted nonlinear codes. While when SNR is modest or high, the linear codes rarely cherish superiority in performance over their nonlinear competitors. Especially for the relatively powerful Mirrored Baker’s map code with rate $1/8$, we can observe an obvious improvement in MSE when SNR is above $10$dB.

This comparison result is actually both bad and good news for us. On one side, we cannot hope to find powerful linear analog codes, as what has happened in DECC. On the other
hand, in order to improve the analog code’s performance, efforts must focus on nonlinear analog codes.

V. CONCLUSION

In this paper, we have studied the performance of the linear analog codes. For the ML and LMMSE decoding scheme, we have proved that the optimal analog codes in the meaning of minimizing the MSE per information bit is the unitary code, which includes some existing linear analog codes like DFT and DCT codes. We also have figured out the lower bound of MSE performance for linear codes’ ML and LMMSE decoding scheme. By comparing these bounds with the performance of several existing nonlinear analog codes, we conclude that linear analog codes cannot be quite satisfying and subsequently the nonlinear codes will be the real potential for powerful analog codes.

REFERENCES

[1] T. G. Marshall, Jr. “Real number transform and convolutional codes,” Proc. 24th Midwest Symp. Circuits Sys., Editor: S. Kame, Albuquerque, NM, June 29-30, 1981

[2] J. K. Wolf, “Analog codes,” IEEE Intl. Conf. Comm, Boston, MA, USA, June, 1983, pp. 310-312.

[3] J. K. Wolf, “Redundancy, the discrete Fourier transform, and impulse noise cancellation,” IEEE Trans. Comm., Vol. COM-31, No. 3, pp. 458-461, March 1983

[4] J.-L. Wu and J. Shiu, “Discrete cosine transform in error control coding”, IEEE Trans. Comm., pp. 1857-1861, May 1995.

[5] B. Chen and G. W. Wornell, “Analog error-correcting codes based on chaotic dynamical systems,” IEEE Trans. Comm., vol 46, Issue 7, July 1998, pp: 881-890

[6] K. Xie, P. Tan, B. C. Ng, and J. Li (Tiffany), “Analog turbo codes: A chaotic construction,” IEEE Intl. Symp. Inf. Theory, 2009.

[7] K. Xie and J. Li (Tiffany), “Linear analog codes: The good and the bad,” submitted to IEEE Globecom, Houston, TX, 2011.

[8] Y. Liu, T. J. Li (Tiffany) and K. Xie “Efficient Image Transmission Through Analog Error Correction,” IEEE Intl. Workshop on Multimedia Signal Processing Hangzhou, Oct. 2011.

[9] S. Ohno and G. B. Giannakis “Capacity maximizing MMSE-optimal pilots for wireless OFDM over frequency-selective block rayleigh-fading channels,” IEEE Trans. Comm., Vol. 50, pp. 2138f2145, Sep 2004