Asymptotically Optimal Joint Source-Channel Coding with Minimal Delay

Marius Kleiner, Bixio Rimoldi
School of Computer and Communication Sciences
Ecole Polytechnique Fédérale de Lausanne
CH-1015 Lausanne, Switzerland
E-mail: firstname.lastname@epfl.ch

***Abstract*—We present and analyze a joint source-channel coding strategy for the transmission of a Gaussian source across a Gaussian channel in n channel uses per source symbol. Among all such strategies, the scheme presented here has the following properties: i) the resulting mean-squared error scales optimally with the signal-to-noise ratio, and ii) the scheme is easy to implement and the incurred delay is minimal, in the sense that a single source symbol is encoded at a time.**

I. Introduction

In this paper we propose and analyze a scheme for the transmission of a discrete-time memoryless Gaussian source across a discrete-time memoryless Gaussian channel, where the channel can be used n times for each source symbol. The parameter n is arbitrary but fixed, given as part of the problem statement.

It is well known that if the source has variance σ_S^2 and the channel noise has variance σ_Z^2 then the average transmit power P and the average mean-squared error D of any communication scheme for this scenario are related by

$$R(D) \leq nC(P), \tag{1}$$

where $R(D) = 0.5\log(\sigma_S^2/D)$ is the rate-distortion function of the source and $C(P) = 0.5\log(1 + P/\sigma_Z^2)$ is the capacity-cost function of the channel (see *e.g.* [1]). Inserting into (1) yields

$$\frac{\sigma_S^2}{D} \leq \left(1 + \frac{P}{\sigma_Z^2}\right)^n,$$

or equivalently

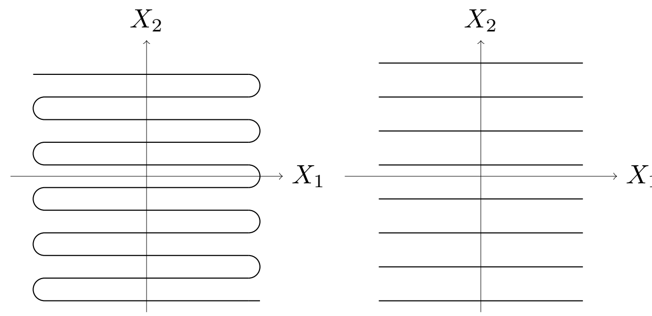
$$\text{SDR} \leq (1 + \text{SNR})^n, \tag{2}$$

where we have defined $\text{SNR} = P/\sigma_Z^2$ and $\text{SDR} = \sigma_S^2/D$. In the limit when SNR goes to infinity,

$$\lim_{\text{SNR}\to\infty} \frac{\log \text{SDR}}{\log \text{SNR}} \leq n. \tag{3}$$

At large SNR, the SDR (signal-to-distortion ratio) behaves thus at best as SNR^n. In this sense n is the best possible scaling exponent that any communication scheme can hope to achieve for a fixed n.

The scheme proposed in this paper achieves this optimal scaling exponent for any fixed n, yet has small complexity and minimal delay in the sense that it operates on a single source symbol at a time. It works by quantizing the source and then repeatedly quantizing the quantization error. The quantized points are sent across the first $n-1$ channel uses and the last quantization error is sent uncoded in the n^{th} channel use.



(a) Shannon's original proposition. (b) Mapping proposed in this paper (for $n = 2$).

Fig. 1. A minimum-delay source-channel code for $n = 2$ can be visualized as a curve in $\mathbb{R}^2$ parametrized by the source. Here we compare the mapping presented in this paper (right) to Shannon's original suggestion (left).

If the quantization resolution is chosen correctly (as a function of the SNR), then the decoding error of the quantization symbols is dominated by that of the uncoded transmission in the last channel use, which is shown to have an asymptotic scaling exponent of n.

Schemes similar to the one proposed here have been considered before. Indeed, one of the first schemes to transmit an analog source across two uses of a Gaussian channel was suggested by Shannon [2]. Notice its resemblance to the constellation studied here, shown in Figure 1.

After Shannon, Wozencraft and Jacobs [3] were among the first to study source-channel mappings as curves in n-dimensional space. Ziv [4] found important theoretical limitations of such mappings. Much of the later work is due to Ramstad and his coauthors (see [5], [6], [7], [8], [9], [10]). A proof that the performance of minimal-delay codes is strictly smaller than that of codes with unrestricted delay when $n > 1$ was given in 2008 by Ingber et al. [11].

For $n = 2$, the presented scheme is almost identical to the HSQLC scheme by Coward [12], which uses a numerically optimized quantizer, transmitter and receiver to minimize the mean-squared error (MSE) for finite values of the SNR. Coward correctly conjectured that the right strategy for $n > 2$

would be to repeatedly quantize the quantization error from the previous step, which is exactly what we do here.

Another closely related communication scheme is the *shift-map* scheme due to Chen and Wornell [13]. Vaishampayan and Costa [14] showed in their analysis that it achieves the scaling exponent $n-\epsilon$ for any $\epsilon > 0$ if the relevant parameters are chosen correctly as a function of the SNR. Up to rotation and a different constellation shaping, the shift-map scheme is in fact virtually identical to the one presented here, a fact that was pointed out recently by Taherzadeh and Khandani [15]. In their own paper they develop a scheme that achieves the optimal scaling exponent exactly and is in addition robust to SNR estimation errors; their scheme, however, is based on rearranging the digits of the binary expansion of the source and is thus quite different from the one presented here.

Shamai, Verdú and Zamir [16] used Wyner-Ziv coding to extend an existing analog system with a digital code when additional bandwidth is available. Mittal and Phamdo [17] (see also the paper by Skoglund, Phamdo and Alajaji [18]) split up the source into a quantized part and a quantization error, much like we do here, but they use a separation-based code (or "tandem" code) to transmit the quantization symbols. Reznic et al. [19] use both quantization and Wyner-Ziv coding, and their scheme includes Shamai et al. and Mittal & Phamdo as extreme cases. All three schemes, however, use long block codes for the digital phase and incur correspondingly large delays, so they are not directly comparable with minimum delay schemes.

While the basic idea of the scheme considered in this paper is not new, the analysis provided is and, to our knowledge, we are the first to give an exact mathematical formulation of the quantization resolution (as a function of the SNR) that leads to the optimal scaling exponent.

II. Proposed Communication Scheme

A. Encoder

To encode a single source letter S into n channel input symbols $X_1, \ldots, X_n$, we proceed as follows. Define $E_0 = S$ and recursively compute the pairs (Q_i, E_i) as

$$Q_i = \frac{1}{\beta}\,\mathrm{int}(\beta E_{i-1})$$
$$E_i = \beta(E_{i-1} - Q_i) \tag{4}$$

for $i = 1, \ldots, n-1$ where $\mathrm{int}(x)$ is the unique integer i such that

$$x \in \left[i - \frac{1}{2}, i + \frac{1}{2}\right)$$

and β is a scaling factor that *grows with the power* P in a way to be determined later. The following result will be useful in the sequel.

Lemma 1: As β goes to infinity, the variance of Q_i converges to that of E_{i-1} for all $i = 1, \ldots, n-1$.

Proof: Intuitively this is so since Q_i is E_{i-1} quantized, and the quantization step becomes smaller as β goes to infinity. A rigorous proof is given in Appendix A. ∎

Proposition 2: The Q_i and E_i satisfy the following properties:

1) The map $S \mapsto (Q_1, \ldots, Q_{n-1}, E_{n-1})$ is one-to-one and
$$S = \sum_{i=1}^{n-1} \frac{1}{\beta^{i-1}} Q_i + \frac{1}{\beta^{n-1}} E_{n-1}. \tag{5}$$
2) The variance of E_0 is σ_S^2 and for all $i = 1, \ldots, n-1$, $E_i \in [-1/2, 1/2)$ and $\mathrm{Var}(E_i) \le 1/4$.
3) For any $\delta > 0$ there exists β_0 such that for $\beta > \beta_0$,
$$\mathrm{Var}(Q_i) \le \begin{cases} \sigma_S^2 + \delta & \text{for } i = 1 \\ 1/4 + \delta & \text{for } i = 2, \ldots, n-1. \end{cases} \tag{6}$$

Proof:

1) From the definition (4) we have
$$E_{i-1} = \frac{1}{\beta} E_i + Q_i. \tag{7}$$
Repeated use of this relationship leads to the given expression for S.
2) First, $\mathrm{Var}(E_0) = \mathrm{Var}(S) = \sigma_S^2$. Next, $E_i \in [-1/2, 1/2)$ follows trivially from the definition of E_i. Furthermore, the variance of any random variable with support in an interval of length 1 is bounded from above by $1/4$.
3) The result follows directly from Lemma 1 and from the bound on the variance of the E_i in point 2 above.

∎

Without loss of generality we assume hereafter that $\sigma_S^2 > 1/4$ so that the first bound of (6) applies to all Q_i.

We determine the channel input symbols X_i from the Q_i and from E_{n-1} according to

$$X_i = \sqrt{\frac{P}{\sigma_S^2 + \delta}}\, Q_i \quad \text{for } i = 1, \ldots, n-1 \text{ and}$$
$$X_n = \sqrt{\frac{P}{\sigma_E^2}}\, E_{n-1},$$

where $\sigma_E^2 = \mathrm{Var}(E_{n-1})$. Following Proposition 2, this ensures that $\mathbb{E}[X_i^2] \le P$ for all i and for $\beta > \beta_0(\delta)$. Since we are interested in the large SNR regime and since we have defined β to grow with P, we can thus assume for the remainder that the power constraint is satisfied.

B. Decoder

The X_i are transmitted across the channel, producing at the channel output the symbols

$$Y_i = X_i + Z_i, \quad i = 1, \ldots, n,$$

where the Z_i are iid Gaussian random variables of variance σ_Z^2. To estimate S from $Y_1, \ldots, Y_n$, the decoder first computes separate estimates $\hat{Q}_1, \ldots, \hat{Q}_{n-1}$ and $\hat{E}_{n-1}$, and then combines them to obtain the final estimate $\hat{S}$. While this strategy is suboptimal in terms of achieving a small MSE, we will see that it is good enough to achieve optimal scaling.

To estimate the Q_i we use a maximum likelihood (ML) decoder, which yields the minimum distance estimate

$$\hat{Q}_i = \frac{1}{\beta} \arg\min_{j \in \mathbb{Z}} \left| \sqrt{\frac{P}{\sigma_S^2 + \delta}} \frac{j}{\beta} - Y_i \right|. \tag{8}$$

To estimate E_{n-1}, we use a linear minimum mean-square error (LMMSE) estimator (see *e.g.* [20, Section 8.3]), which computes

$$\hat{E}_{n-1} = \frac{\mathbb{E}[E_{n-1} Y_n]}{\mathbb{E}[Y_n^2]} Y_n. \tag{9}$$

Finally we use the relationship (5) to obtain

$$\hat{S} = \sum_{i=1}^{n-1} \frac{1}{\beta^{i-1}} \hat{Q}_i + \frac{1}{\beta^{n-1}} \hat{E}_{n-1}. \tag{10}$$

C. Error Analysis

The overall MSE $\mathbb{E}[(S - \hat{S})^2]$ can be broken up into contributions due to the errors in decoding Q_i and E_{n-1} as follows. From (5) and (10), the difference between S and $\hat{S}$ is

$$S - \hat{S} = \sum_{i=1}^{n-1} \frac{1}{\beta^{i-1}} (Q_i - \hat{Q}_i) + \frac{1}{\beta^{n-1}} (E_{n-1} - \hat{E}_{n-1}).$$

The error terms $Q_i - \hat{Q}_i$ depend only on the noise of the respective channel uses and are therefore independent of each other and of $E_{n-1} - \hat{E}_{n-1}$, so we can write the error variance componentwise as

$$\mathbb{E}[(S - \hat{S})^2] = \sum_{i=1}^{n-1} \frac{1}{\beta^{2(i-1)}} \mathcal{E}_{Q,i} + \frac{1}{\beta^{2(n-1)}} \mathcal{E}_E, \tag{11}$$

where $\mathcal{E}_{Q,i} \stackrel{\text{def}}{=} \mathbb{E}[(Q_i - \hat{Q}_i)^2]$ and $\mathcal{E}_E \stackrel{\text{def}}{=} \mathbb{E}[(E_{n-1} - \hat{E}_{n-1})^2]$.

Lemma 3: For each $i = 1, \ldots, n-1$,

$$\mathcal{E}_{Q,i} \in O\left(\exp\{-k\text{SNR}/\beta^2\}\right), \tag{12}$$

where $\text{SNR} = P/\sigma_Z^2$ and $k > 0$ is a constant. (The O-notation is defined in Appendix B.)

Proof: Define the interval

$$\mathcal{I}_j = \left[\frac{(j - \frac{1}{2})\sqrt{P}}{\beta\sqrt{\sigma_S^2 + \delta}}, \frac{(j + \frac{1}{2})\sqrt{P}}{\beta\sqrt{\sigma_S^2 + \delta}} \right).$$

According to the minimum distance decoder (8), $\hat{Q}_i - Q_i = j/\beta$ whenever $Z_i \in \mathcal{I}_j$. The error $\mathcal{E}_{Q,i}$ satisfies thus

$$\begin{aligned} \mathbb{E}[(Q_i - \hat{Q}_i)^2] &= \frac{1}{\beta^2} \sum_{j \in \mathbb{Z}} j^2 \Pr[Z_i \in \mathcal{I}_j] \\ &= \frac{2}{\beta^2} \sum_{j=1}^{\infty} j^2 \Pr[Z_i \in \mathcal{I}_j], \end{aligned} \tag{13}$$

where the second equality follows from the symmetry of the distribution of Z_i. Now,

$$\Pr[Z_i \in \mathcal{I}_j] = Q\left(\frac{(j - \frac{1}{2})\sqrt{\text{SNR}}}{\beta\sqrt{\sigma_S^2 + \delta}} \right) - Q\left(\frac{(j + \frac{1}{2})\sqrt{\text{SNR}}}{\beta\sqrt{\sigma_S^2 + \delta}} \right),$$

where

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\xi^2/2} d\xi,$$

which can be bounded from above for $x \geq 0$ as

$$Q(x) \leq \frac{1}{2} e^{-x^2/2}.$$

For $\beta \geq 1$ we can now bound (13) as

$$\mathcal{E}_{Q,i} \leq \sum_{j=1}^{\infty} j^2 \exp\left\{ -\frac{(j - 1/2)^2 \text{SNR}}{2\beta^2(\sigma_S^2 + \delta)} \right\}.$$

Note that for $j \geq 2$, $(j - 1/2)^2 > j$. Thus

$$\begin{aligned} \mathcal{E}_{Q,i} \quad \leq \quad & \exp\left\{ -\frac{\text{SNR}}{8\beta^2(\sigma_S^2 + \delta)} \right\} \\ & + \sum_{j=2}^{\infty} j^2 \exp\left\{ -\frac{j\text{SNR}}{2\beta^2(\sigma_S^2 + \delta)} \right\}. \end{aligned} \tag{14}$$

To bound the infinite sum we use

$$\sum_{j=2}^{\infty} j^2 p^j \leq \sum_{j=1}^{\infty} j^2 p^j = \frac{p^2 + p}{(1-p)^3} \tag{15}$$

with $p = \exp\{-\text{SNR}/2\beta^2(\sigma_S^2 + \delta)\}$. The first term of (14) thus dominates for large values of SNR/β^2 and

$$\mathcal{E}_{Q,i} \leq c \exp\left\{ -\frac{\text{SNR}}{8\beta^2(\sigma_S^2 + \delta)} \right\}$$

for some $c > 0$, which completes the proof. ■

Lemma 4: $\mathcal{E}_E \in O(\text{SNR}^{-1})$.

Proof: The mean-squared error that results from the LMMSE estimation (9) is

$$\mathcal{E}_E = \sigma_E^2 - \frac{(\mathbb{E}[E_{n-1} Y_n])^2}{\mathbb{E}[Y_n^2]}. \tag{16}$$

Since

$$Y_n = X_n + Z_n = \sqrt{\frac{P}{\sigma_E^2}} E_{n-1} + Z_n,$$

we have $\mathbb{E}[E_{n-1} Y_n] = \sqrt{P\sigma_E^2}$. Moreover, $\mathbb{E}[Y_n^2] = \mathbb{E}[X^2] + \mathbb{E}[Z^2] = P + \sigma_Z^2$. Inserting this into (16) we obtain

$$\begin{aligned} \mathcal{E}_E &= \sigma_E^2 - \frac{P\sigma_E^2}{P + \sigma_Z^2} \\ &= \sigma_E^2 \left(1 - \frac{P}{P + \sigma_Z^2} \right) \\ &= \frac{\sigma_E^2}{1 + \text{SNR}} \\ &< \frac{\sigma_E^2}{\text{SNR}}. \end{aligned}$$

Since σ_E^2 is bounded (cf. Proposition 2), $\mathcal{E}_E \in O(\text{SNR}^{-1})$ as claimed. ■

D. Achieving the Optimal Scaling Exponent

Recall the formula for the overall error

$$\mathbb{E}[(S-\hat{S})^2] = \sum_{i=1}^{n-1} \frac{1}{\beta^{2(i-1)}}\mathcal{E}_{Q,i} + \frac{1}{\beta^{2(n-1)}}\mathcal{E}_E.$$

According to Lemma 3, $\mathcal{E}_{Q,i}$ decreases exponentially when SNR/β^2 goes to infinity. This happens for increasing SNR if we set *e.g.*

$$\beta^2 = \text{SNR}^{1-\epsilon}$$

for some $\epsilon > 0$, in which case $\mathcal{E}_{Q,i} \in O\left(\exp(-k\text{SNR}^{\epsilon})\right)$. From this and Lemma 4, the overall error satisfies

$$\mathbb{E}[(S-\hat{S})^2] \in O(\text{SNR}^{-(n-\epsilon')}), \tag{17}$$

where $\epsilon' = (n-1)\epsilon$ can be made as small as desired. The scaling exponent for a fixed ϵ satisfies therefore

$$\lim_{\text{SNR}\to\infty} \frac{\log \text{SDR}}{\log \text{SNR}} \geq \lim_{\text{SNR}\to\infty} \frac{\log \sigma_S^2 + (n-\epsilon')\log \text{SNR}}{\log \text{SNR}} = n-\epsilon'. \tag{18}$$

Note that the choice of ϵ represents a tradeoff: for small ϵ the error due to the "discrete" part vanishes only slowly, but the scaling exponent in the limit is larger. For larger ϵ, $\mathcal{E}_Q$ vanishes quickly but the resulting exponent is smaller. In the remainder of this section we show how we can choose ϵ as a function of SNR to achieve the optimal scaling.

Let now

$$\epsilon = \epsilon(\text{SNR}) = \frac{\log(n \log \text{SNR}/k)}{\log \text{SNR}}, \tag{19}$$

where k is the constant indicating the decay of $\mathcal{E}_{Q,i}$ in (12). With this choice of ϵ,

$$\begin{aligned}\mathcal{E}_{Q,i} &\in O\left(\exp\left(-k\text{SNR}^{\epsilon}\right)\right)\\ &= O(\text{SNR}^{-n}),\end{aligned}$$

hence the overall error is still dominated as in (17), and (18) still applies. Inserting (19) in (18), we find

$$\begin{aligned}&\lim_{\text{SNR}\to\infty} \frac{\log \sigma_S^2 + (n-(n-1)\epsilon)\log \text{SNR}}{\log \text{SNR}}\\ &= \lim_{\text{SNR}\to\infty} \frac{\log \sigma_S^2 + n\log \text{SNR} - (n-1)\log(n\log \text{SNR}/k)}{\log \text{SNR}}\\ &= n,\end{aligned}$$

which is indeed the optimal scaling exponent.

Remark 1: While the limiting *exponent* above is indeed the optimal one, the SDR scales as $\text{SNR}^n(\log \text{SNR})^{-(n-1)}$ rather than the theoretic optimum SNR^n. This means that the *gap* (in dB) between the theoretically optimal SDR value and our lower bound grows to infinity as $\text{SNR} \to \infty$. According to a result in [15], however, no scheme combining quantization and uncoded transmission as done here can achieve a better SDR scaling than $\text{SNR}^n(\log \text{SNR})^{-(n-1)}$. In the scaling sense, our bound is therefore tight.

III. Conclusions

We have presented and analyzed a joint source-channel communication strategy that achieves the optimal scaling exponent if the channel is to be used n times per source symbol. The given scheme incurs the smallest possible delay and its implementation is straightforward.

While the basic structure of this scheme – separating the source into a quantized part and the associated error – is not new, the simple analysis provided here yields an explicit expression for the quantization resolution in terms of the SNR that leads to the optimal scaling exponent.

Appendix A
Proof of Lemma 1

Since all involved distributions are symmetric, $\mathbb{E}[Q_i] = 0$. Writing Q_i as a function of E_{i-1}, we have

$$\text{Var}(Q_i) = \mathbb{E}[Q_i^2] = \int_{-\infty}^{\infty} Q_i(\xi)^2 f(\xi) d\xi, \tag{20}$$

where $f(\xi)$ is the pdf[1] of E_{i-1}. Now, $Q_i(\xi) = j/\beta$ whenever

$$\xi \in \left[\frac{j-1/2}{\beta}, \frac{j+1/2}{\beta}\right).$$

With this, the integral (20) becomes

$$\begin{aligned}\text{Var}(Q_i) &= \frac{1}{\beta^2}\sum_{j\in\mathbb{Z}} j^2 \int_{\frac{j-1/2}{\beta}}^{\frac{j+1/2}{\beta}} f(\xi)d\xi\\ &= \sum_{j\in\mathbb{Z}} \left(\frac{j}{\beta}\right)^2 \left[F\left(\tfrac{j+1/2}{\beta}\right) - F\left(\tfrac{j-1/2}{\beta}\right)\right],\end{aligned}$$

where $F(\xi)$ is the cdf[2] of E_{i-1}. As β goes to infinity, this sum converges to a Riemann-Stieltjes integral:

$$\text{Var}(Q_i) \longrightarrow \int \xi^2 dF(\xi) = \text{Var}(E_{i-1}) \quad \text{as } \beta \to \infty.$$

■

Appendix B
Big-O Notation

The "Big-O" asymptotic notation used at various points in the paper is defined as follows. Let $f(x)$ and $g(x)$ be two functions defined on $\mathbb{R}$. We write

$$f(x) \in O(g(x))$$

if and only if there exists an x_0 and a constant c such that

$$f(x) \leq cg(x)$$

for all $x > x_0$.

As a simple consequence of this definition, if $f(x) \in O(x^n)$ then

$$\lim_{x\to\infty} \frac{\log f(x)}{\log x} \leq n.$$

[1]probability density function
[2]cumulative distribution function

References

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 1991.

[2] C. E. Shannon, "Communication in the presence of noise," *Proceedings of the I.R.E.*, vol. 37, pp. 10–21, January 1949.

[3] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. John Wiley & sons, 1965.

[4] J. Ziv, "The behavior of analog communication systems," *IEEE Transactions on Information Theory*, vol. 16, pp. 587–594, Sep 1970.

[5] T. A. Ramstad, "Shannon mappings for robust communication," *Telektronikk*, vol. 98, no. 1, pp. 114–128, 2002.

[6] P. Floor and T. Ramstad, "Noise analysis for dimension expanding mappings in source-channel coding," in *2006 IEEE Workshop on Signal Processing Advances in Wireless Communications*, pp. 1–5, July 2006.

[7] H. Coward and T. A. Ramstad, "Bandwidth doubling in combined source-channel coding of memoryless Gaussian sources," in *2000 IEEE International Symposium of Intelligent Signal Processing and Communication Systems*, Honolulu, HI, pp. 571–576, November 2000.

[8] H. Coward and T. Ramstad, "Quantizer optimization in hybrid digital-analog transmission of analog source signals," in *2000 IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 2637–2640, 2000.

[9] N. Wernersson, T. A. Ramstad, and M. Skoglund, "Analog source-channel codes based on orthogonal polynomials," in *Asilomar Conference on Signals, Systems and Computers*, 2007.

[10] F. Hekland, P. Floor, and T. Ramstad, "Shannon-kotel'nikov mappings in joint source-channel coding," *IEEE Transactions on Communications*, vol. 57, pp. 94–105, January 2009.

[11] A. Ingber, I. Leibowitz, R. Zamir, and M. Feder, "Distortion lower bounds for finite dimensional joint source-channel coding," in *2008 IEEE International Symposium on Information Theory*, pp. 1183–1187, 2008.

[12] H. Coward, *Joint Source–Channel Coding: Development of Methods and Utilization in Image Communications*. PhD thesis, Norwegian University of Science and Technology, 2001.

[13] B. Chen and G. Wornell, "Analog error-correcting codes based on chaotic dynamical systems," *IEEE Transactions on Communications*, vol. 46, pp. 881–890, Jul 1998.

[14] V. Vaishampayan and S. Costa, "Curves on a sphere, shift-map dynamics, and error control for continuous alphabet sources," *IEEE Transactions on Information Theory*, vol. 49, pp. 1658–1672, July 2003.

[15] M. Taherzadeh and A. K. Khandani, "Robust joint source-channel coding for delay-limited applications," http://arxiv.org/abs/0805.4023.

[16] S. Shamai, S. Verdu, and R. Zamir, "Systematic lossy source/channel coding," *IEEE Transactions on Information Theory*, vol. 44, pp. 564–579, Mar 1998.

[17] U. Mittal and N. Phamdo, "Hybrid digital-analog (hda) joint source-channel codes for broadcasting and robust communications," *IEEE Transactions on Information Theory*, vol. 48, pp. 1082–1102, May 2002.

[18] M. Skoglund, N. Phamdo, and F. Alajaji, "Design and performance of vq-based hybrid digital-analog joint source-channel codes," *IEEE Transactions on Information Theory*, vol. 48, pp. 708–720, Mar 2002.

[19] Z. Reznic, M. Feder, and R. Zamir, "Distortion bounds for broadcasting with bandwidth expansion," *IEEE Transactions on Information Theory*, vol. 52, pp. 3778–3788, Aug. 2006.

[20] L. Scharf, *Statistical Signal Processing*. Addison-Wesley, 1990.