Hybrid Digital-Analog Source-Channel Coding with One-to-Three Bandwidth Expansion

Ahmad Abou Saleh, Fady Alajaji, and Wai-Yip Chan

Abstract—**We consider the problem of bandwidth expansion for lossy joint source-channel coding over a memoryless Gaussian channel. A low delay 1:3 bandwidth expansion hybrid digital-analog coding system, which combines a scalar quantizer and a 1:2 nonlinear analog coder, is proposed. It is shown that our proposed system outperforms the 1:3 generalized hybrid scalar quantizer linear coder in terms of signal-to-distortion ratio (SDR). A lower bound on the system SDR is also derived.**

I. INTRODUCTION

The traditional approach for analog source transmission is to use separate source and channel coders. This separation is optimal given infinite delay and complexity in the coders [1]. In practice, joint source-channel coding (JSCC) can achieve better performance when delay and complexity are constrained. A common approach for JSCC design is to jointly optimize the components of a tandem system with respect to the channel and source characteristics. Another approach based on non-linear analog mapping is treated in [2]–[5].

With the increasing popularity of wireless sensors networks, reliable transmission with delay and complexity constraints is more relevant than ever. A sensor node communicates its sensed field information to a fusion center over a noisy wireless channel. In this paper, we focus on low delay and low complexity lossy JSCC by proposing a 1:3 bandwidth expansion scheme based on combining a scalar quantizer with a 1:2 nonlinear analog coder operating on the quantization error. The proposed hybrid digital-analog scheme is shown to be suitable for wireless sensor networks.

The hybrid scalar quantizer linear coder (HSQLC) provides a dimension expansion of rate $R = 2$ [6]. This scheme uses pulse amplitude modulation to send a scalar quantizer output on one channel, and linear uncoded analog transmission of the quantization error on another channel. For rates larger than two ($R > 2$), Coward suggested to repeatedly quantize the error from the previous quantization step and finish the last step with a linear coder [7]. Recently, a similar system was studied in [8], referred to as "generalized HSQLC". For reference, we compare the proposed system to the generalized HSQLC. Similar hybrid systems based on splitting the source into a quantization part and a quantization error part were also proposed in [9]. These schemes, however, use long channel block codes for the quantization part, thus incurring large

delay and complexity and making them not comparable with our proposed low delay scheme. The rest of the paper is organized as follows. In Section II, we describe a Shannon-Kotel'nikov mapping using the 1:2 Archimedes' spiral. Section III describes the system model and optimization. A lower bound on the system signal-to-distortion ratio (SDR) is derived in Section IV. Simulation results are included in Section V. Finally, conclusions are drawn in Section VI.

II. A SHANNON-KOTEL'NIKOV MAPPING

In this section, a 1:2 Archimedes' spiral mapping is described for both uniform and Gaussian memoryless sources. Bandwidth expansion is performed by mapping each source sample $x \in \mathbb{R}$ to a two-dimensional channel symbol, which is a point on the double Archimedes' spiral, given by [5]

$$\mathbf{s}(x) = \begin{bmatrix} z_1(x) \\ z_2(x) \end{bmatrix} = \frac{1}{\pi} \begin{bmatrix} sgn(x)\Delta\varphi(x)\cos\varphi(x) \\ sgn(x)\Delta\varphi(x)\sin\varphi(x) \end{bmatrix} \quad (1)$$

where $sgn(\cdot)$ is the signum function, Δ is the radial distance between any two neighboring spiral arms, and $\varphi(x) = \sqrt{6.25|x|/\Delta}$ is a stretching bijective function. For a given channel signal-to-noise ratio (CSNR) defined as P/σ_N^2, where P and σ_N^2 are the average channel power and noise variance, respectively, the radial distance Δ is optimized to minimize the total distortion by solving the following unconstrained optimization problem

$$\Delta_{\text{opt}} = \arg\min_\Delta [\bar{\varepsilon}_{\text{wn}}^2(\Delta) + \bar{\varepsilon}_{\text{th}}^2(\Delta)] \quad (2)$$

where $\bar{\varepsilon}_{\text{wn}}^2$ and $\bar{\varepsilon}_{\text{th}}^2$ are, respectively, the average weak noise and threshold distortion under maximum likelihood (ML) decoding. For a uniform source on the interval $[-a, a]$, where $a > 0$, the average weak noise distortion is [5]

$$\bar{\varepsilon}_{\text{wn}}^2 = \frac{\sigma_N^2}{\alpha^2} \quad (3)$$

where α is a gain factor related to the average channel power constraint P given by

$$P = \frac{1}{2} \int_{-a}^{a} ||\mathbf{s}(\alpha x)||^2 f_X(x) dx \quad (4)$$

where $f_X(x)$ is the probability density function (pdf) of the source X. The threshold distortion for a uniform source is

$$\bar{\varepsilon}_{\text{th}}^2 = 2a^2 \left(1 - \text{erf}\left(\frac{\Delta}{2\sqrt{2}\sigma_N}\right)\right) \left(\frac{4a}{3} + \frac{\pi^4\eta^2\Delta^2}{\alpha^2 a} + \frac{4\pi^2\eta\Delta}{\alpha}\right) \quad (5)$$

where $\text{erf}(\cdot)$ is the Gaussian error function, and $\eta = 0.16$. Hence, the calculated performance of the 1:2 spiral mapping

under ML decoding is SDR $= \sigma_X^2/(\bar{\varepsilon}_{\text{wn}}^2 + \bar{\varepsilon}_{\text{th}}^2)$, where σ_X^2 is the source variance. For a Gaussian source, the expression for the threshold distortion is derived in [5].

At the receiver side, we use the optimum minimum mean square error (MMSE) decoder instead of the ML decoder used in [5]. MMSE decoding has been shown to achieve a substantial performance improvement over ML decoding at low CSNRs under 2:1 bandwidth reduction [10]. For 1:2 bandwidth expansion, the MMSE decoding rule can be written as follows

$$\hat{x} = E[X|\hat{z}_1, \hat{z}_2] = \int xp(x|\hat{z}_1, \hat{z}_2)dx = \frac{\int xp(\hat{z}_1, \hat{z}_2|x)p(x)dx}{\int p(\hat{z}_1, \hat{z}_2|x)p(x)dx} \tag{6}$$

where $\{\hat{z}_i\}_{i=1,2} = z_i + n_i$ are the received channel outputs. For independent and identically distributed (i.i.d.) Gaussian noise n_i, we have $p(\hat{z}_1, \hat{z}_2|x) = p(\hat{z}_1|x)p(\hat{z}_2|x)$, where

$$p(\hat{z}_i|x) = \frac{1}{\sqrt{2\pi}\sigma_N}e^{-\frac{(\hat{z}_i - z_i(x))^2}{2\sigma_N^2}}, \quad i = 1, 2 \tag{7}$$

and $z_i(x)$ is given by (1). To make the decoder implementation computationally efficient, we devise a decoder based on quantization and table-lookup. This is done by uniform quantization of the output of the channel $\hat{\mathbf{z}} \in \mathbb{R}^2$ and looking up the decoded value $\hat{x}$ for each quantization bin in a table.

As shown in Fig. 1, the performance of the spiral mapping with uniform source is at most 8 dB below the theoretical limit. This is comparable to the case of Gaussian source [5]. It can be noticed that the performance of the 1:2 bandwidth expansion system with MMSE decoder is close to the performance of the block pulse amplitude modulation (BPAM) [11] at low CSNRs (≤ 10 dB), and to the 1:2 bandwidth expansion system with ML decoding at high CSNRs (> 10 dB). Since it is intractable to find an analytical expression for the system performance under MMSE decoding, we use the following approximation to represent it

$$D_{\text{exp}} = \begin{cases} D_{\text{BPAM}} = \left(\frac{\sigma_X \sigma_N}{P + \sigma_N^2}\right)^2 \left(\sigma_N^2 + \frac{P^2}{2}\right), & \text{CSNR} \leq 10\text{dB} \\ D_{\text{ML}} = \bar{\varepsilon}_{\text{wn}}^2 + \bar{\varepsilon}_{\text{th}}^2, & \text{otherwise} \end{cases} \tag{8}$$

where D_{BPAM} and D_{ML} are, respectively, the distortion from the 1:2 BPAM and 1:2 spiral expansion under ML decoding. In Fig. 1, this is given by the dashed curve for CSNR≤ 10dB and by the solid curve (with pluses) for CSNR> 10dB. Similar results are also obtained for Gaussian source.

III. SYSTEM MODEL

A. System Structure

In this section, we assume a Gaussian source X with variance σ_X^2 to be transmitted over a power limited, discrete time, and continuous amplitude channel with additive white Gaussian noise $n \sim \mathcal{N}(0, \sigma_N^2)$. We propose a 1:3 bandwidth expansion system that consists of a scalar quantizer and a 1:2 dimension expansion using Archimedes' spiral, as shown in Fig. 2. The proposed system works as follows. A source symbol X is first quantized using an L-level quantizer Q(.). The quantizer uses a set of decision intervals $D_i = (d_i, d_{i+1})$
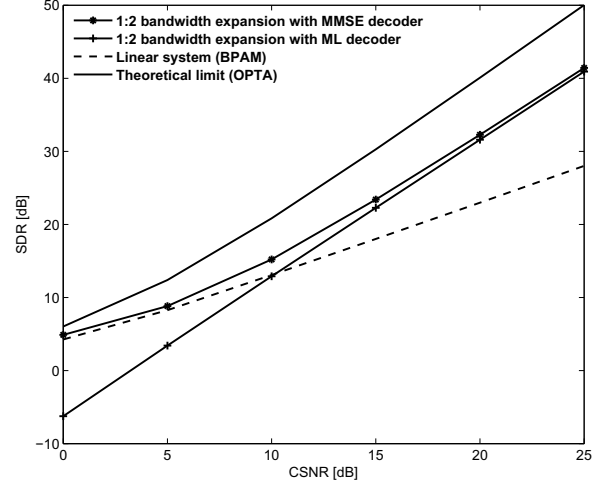

Fig. 1. Performance of the 1:2 bandwidth expansion system for a uniformly distributed source over [-1,1] using the Archimedes' spiral. The optimal performance theoretically attainable (OPTA) SDR$=\sigma_X^2/D$, where D is the total distortion, is plotted for a memoryless Gaussian source which represents a lower bound on OPTA SDR for all other sources. The linear system SDR is also plotted.
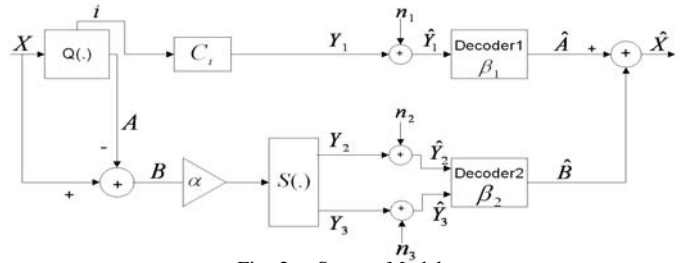

Fig. 2. System Model.

with $d_0 = -\infty$ and $d_L = +\infty$. It returns both the index i and the representation level $A = a_i$, $i = \{0, \ldots, L-1\}$. The index i is represented by the channel input $Y_1 = c_i$. The quantization error $B = X - A$ is mapped to a two-dimensional channel symbol using the 1:2 Archimedes' spiral. The system is optimized to minimize the mean square error (MSE) $E[(X - \hat{X})^2]$ while satisfying the average channel power constraint P

$$\frac{1}{3}\left\{E[Y_1^2] + \int ||\mathbf{s}(\alpha b)||^2 f_B(b)db\right\} \leq P \tag{9}$$

where α is the gain factor, and $f_B(b)$ is the pdf of the quantization error B. At the receiver, we use optimal decoders for both the received quantized symbol and quantization error. The MMSE decoder introduced in Section II is used on the quantization error, and the quantized symbol is decoded as

$$\beta_1(\hat{y}_1) = E[A|\hat{y}_1] = \frac{\sum_{i=0}^{L-1} a_i P(a_i)e^{-\frac{(\hat{y}_1 - c_i)^2}{2\sigma_N^2}}}{\sum_{i=0}^{L-1} P(a_i)e^{-\frac{(\hat{y}_1 - c_i)^2}{2\sigma_N^2}}} \tag{10}$$

where $P(a_i)$ is the probability that $A = a_i$. Note that table-lookup can be used to reduce complexity. We also use a mid-tread uniform quantizer, so that only spacing between levels

need to be specified $\left(i.e., a_i = \left(i - (L-1)/2\right)\delta_a, c_i = Ka_i\right.$ for $i \in 0,...,L-1$, $d_j = \left(j - L/2\right)\delta_d$ for $j \in 1,...,L-1$, and L is an odd number$\big)$.

The distribution of the quantization error B is given by

$$f_B(b) = \sum_{i=0}^{L-1} f_{A,B}(a_i, b) \tag{11}$$

where $f_{A,B}(a_i, b)$ can be expressed as follows

$$f_{A,B}(a_i, b) = \begin{cases} f_X(b + a_i), & b \in (d_i - a_i, d_{i+1} - a_i) \\ 0, & \text{Otherwise} \end{cases}. \tag{12}$$

Using (11), the average channel power constraint in (9) becomes

$$\frac{1}{3}\left\{\sigma_{Y_1}^2 + \frac{\alpha\Delta}{\pi^2\eta}\sum_{i=0}^{L-1}\left(\int_{d_i}^{d_{i+1}} |x - a_i| f_X(x)dx\right)\right\} \leq P \tag{13}$$

where $\sigma_{Y_1}^2$ is the variance of the channel input Y_1.

B. System Optimization

The overall MSE $E[(X - \hat{X})^2]$ can be broken up as follows

$$\text{MSE} = \text{MSE}_q + \text{MSE}_{\text{exp}} \tag{14}$$

where MSE_{exp} is the distortion in decoding B, and MSE_q is the distortion in decoding A, given by

$$\begin{aligned} \text{MSE}_q &= E[(A - \hat{A})^2] = E[(A - \beta_1(C(A) + n))^2] \\ &= \int_{-\infty}^{+\infty}\sum_{i=0}^{L-1}(a_i - \beta_1(c_i + n))^2 P(a_i)f_N(n)dn \end{aligned} \tag{15}$$

where $f_N(n)$ is the pdf of the Gaussian noise.

We minimize numerically the MSE with respect to the quantizer parameters (δ_a, K, δ_d), spiral parameter Δ, and α. For a given amount of power P_1 assigned to Y_1, the quantizer parameters are found by minimizing MSE_q using the line search strategy. The optimization is performed for increasing channel noise levels. As initial condition for the optimization, the result for the previous noise level is used. We also impose additional solution constraints that the parameters are real positive-valued and $\sigma_{Y_1}^2$ is equal to the power assigned to Y_1.

By setting $\delta_a = \delta_d$, the interval $(d_i - a_i, d_{i+1} - a_i) = \left(-\frac{\delta_d}{2}, +\frac{\delta_d}{2}\right)$ in (12) is fixed for $i = 1,\ldots,L-2$. For low CSNRs, one can argue that the pdf $f_B(b)$ can be approximated by a Gaussian distribution, since minimizing MSE_q will make δ_d large. For high CSNRs, B fits well a uniform distribution by the fact that δ_d is relatively small with respect to σ_X and the overload probability is negligible. For the sake of simplicity, the spiral optimization (i.e., finding Δ) is conducted assuming that the quantization error B fits well either a Gaussian or a uniform distribution. Hence, results from Section II can be used; (Δ, α) are optimized using (2) while (13) is satisfied with equality, and MSE_{exp} is calculated in a similar way to (8). The design algorithm is formally stated below.

1) Choose some initial values for (δ_a, K, L).

2) Set $i = 0$, the overall MSE $D^{(0)} = \infty$, $D_{\text{best}} = \infty$, and the power assigned to Y_1, $P_1 = \beta P$, $0 \leq \beta \leq 3$.
3) Let N be the noise variance level for which the system should be optimized. Set N' to correspond to a less noisy channel.
4) Set $i \leftarrow i + 1$.
5) Find the optimal quantizer parameters (δ_a, K) for N' by minimizing (15) under the constraint $\sigma_{Y_1}^2 = P_1$.
6) If $N' \geq N$ go to Step 7. Otherwise set $N' \leftarrow \gamma N'$, $\gamma > 1$, and go to Step 5 using the current parameters as initialization for the new optimization.
7) Optimize (Δ, α) using (2) while (13) is satisfied with equality.
8) Calculate MSE_{exp} using (8), and $D^{(i)}$ using (14).
9) If $D_{\text{best}} > D^{(i)}$, set $D_{\text{best}} \leftarrow D^{(i)}$, and $R \leftarrow (\delta_a, K, \Delta, \alpha, D_{\text{best}}, P_1)$.
10) If $P_1 > \epsilon P$, set $P_1 \leftarrow P_1 - \lambda P$, and go to Step 4.
11) Return R.

In our simulations, we used $L = 35$, $\beta = 1$, the initial noise level N' is set to correspond to a channel with CSNR = 30 dB, $\gamma = 10^{0.5}$ which correspond to a 5 dB decrease in CSNR level, and $\epsilon = \lambda = 0.1$. Note that we also conducted some simulations for $P1 > P$ (i.e., $\beta > 1$).

IV. System Lower Bound on the SDR

In this section, we derive a lower bound on the proposed system SDR following the approach of [8]. For this bound, we assume (suboptimal) ML decoding for recovering both the quantized symbol and quantization error.

For the quantized symbol, the system works as follows. At the encoder side, the quantized symbol is scaled by a gain factor K to satisfy the power constraint and transmitted through the channel. At the receiver, the channel output is rescaled using $1/K$ and ML decoder, which is a minimum distance estimator, is applied on the rescaled signal $\hat{y}_1 = a_i + \tilde{n}$, where $\tilde{n} \sim \mathcal{N}(0, (\frac{\sigma_N}{K})^2)$. The error in decoding the quantized value can be expressed as follows

$$\begin{aligned} E[(A - \hat{A})^2] &= \delta_a^2\sum_{i=0}^{L-1}P(a_i)\sum_{j=0}^{L-1}(j-i)^2 P_{i,j} \\ &\leq \delta_a^2 L\sum_{m=1}^{L-1} m^2 P_m \end{aligned} \tag{16}$$

where $m = |j - i|$, and $P_m = P_{i,j}$ is the probability of receiving a_j when a_i is transmitted, given by

$$P_{i,j} = Q\left(\frac{(|j-i|-1/2)\delta_a}{\frac{\sigma_N}{K}}\right) - Q\left(\frac{(|j-i|+1/2)\delta_a}{\frac{\sigma_N}{K}}\right) \tag{17}$$

where $j = 2,...,L-2$, and $Q(x)$ is the Gaussian Q-function which can be upper bounded for $x \geq 0$ as

$$Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-\tau^2/2}d\tau \leq \frac{1}{2}e^{-x^2/2}. \tag{18}$$

Using (18) and dropping the second term in (17), the transition probability $P_{i,j}$ can be upper bounded by

$$P_{i,j} \leq \frac{1}{2}\exp\left(-K^2\frac{(m-1/2)^2\delta_a^2}{2\sigma_N^2}\right). \tag{19}$$

Thus (16) can be expressed as follows

$$
\begin{aligned}
E[(A - \hat{A})^2] &\leq \frac{\delta_a^2 L}{2} \sum_{m=1}^{L-1} m^2 \exp\left(-\frac{(K\delta_a)^2(m-1/2)^2}{2\sigma_N^2}\right) \\
&\leq \frac{\delta_a^2 L}{2} \exp\left(-\frac{K^2\delta_a^2}{8\sigma_N^2}\right) \\
&\quad + \frac{\delta_a^2 L}{2} \sum_{m=2}^{L-1} m^2 \exp\left(-\frac{(K\delta_a)^2 m}{2\sigma_N^2}\right) \quad (20)
\end{aligned}
$$

where we have used the fact that $(m-1/2)^2 > m$ for $m \geq 2$. The summation can be bounded to obtain an upper bound on the distortion from decoding the quantized symbol

$$
\begin{aligned}
\sum_{m=2}^{\ell} m^2 z^m &\leq \sum_{m=1}^{\ell} m^2 z^m = \frac{z}{(1-z)^3}\big(1 + z - (\ell+1)^2 z^\ell \\
&\quad + (2\ell^2 + 2\ell - 1)z^{\ell+1} - \ell^2 z^{\ell+2}\big) \quad (21)
\end{aligned}
$$

where $z = \exp\left(-\frac{(K\delta_a)^2}{2\sigma_N^2}\right)$, and $\ell = L - 1$.

The distortion from decoding B is bounded by

$$
E[(B - \hat{B})^2] \leq \bar{\varepsilon}_{\text{wn}}^2 + \bar{\varepsilon}_{\text{th}}^2 \quad (22)
$$

where $\bar{\varepsilon}_{\text{wn}}^2$ and $\bar{\varepsilon}_{\text{th}}^2$ are, respectively, the weak noise and threshold distortion given by (3) and (5). Adding (20) and (22) gives an upper bound on the end-to-end distortion, thus yielding a lower bound on the system SDR.

V. NUMERICAL RESULTS

In this section, we assume an i.i.d. Gaussian source X with variance $\sigma_X^2 = 1$. The power allocated to Y_1, P_1, is chosen to maximize the overall MSE in (14). For high CSNRs, it is noticed that $P_1 = P$, while for low CSNRs, assigning lower power to Y_1 gives the best performance. This can be explained by the fact that at low CSNRs, the optimization decreases the number of effective quantization levels (i.e., δ_d is large) in order to minimize the decoding error. This means that the probability of the quantized value being nonzero is low. Two other systems are considered and compared to our proposed system: 1) the optimal linear system (1:3 BPAM) which is the best possible linear solution [11], and 2) a 1:3 generalized HSQLC introduced in [8]. Note that the generalized HSQLC parameters are optimized for the given CSNR level. This is done by searching for the quantization resolution that gives the best achievable performance. As shown in Fig. 3, the proposed system has the highest performance over all CSNR levels and gives a 3~5dB gain over the generalized HSQLC. The lower bound on the proposed system SDR, derived in Section IV, is also plotted. Notice that the lower bound is tighter for CSNR$\geq$ 15 dB and is a few dBs away from the actual performance for CSNR$<$ 15 dB. From other simulations, we noticed that the proposed system gives 2~3 dB gain over the one with ML decoding for CSNR$<$ 10 dB , whereas as CSNR gets larger, both systems give similar performance.

VI. SUMMARY AND CONCLUSION

In this paper, we have presented a low-delay lossy joint source-channel coding scheme which combines a scalar quantizer and a 1:2 nonlinear analog coder. The system is optimized
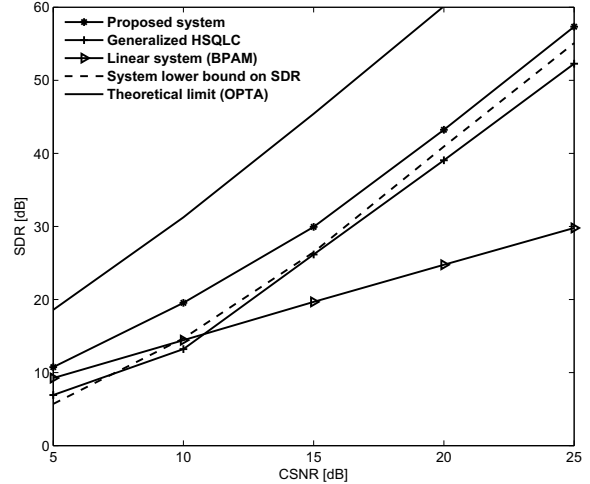

Fig. 3. Performance of the proposed system for $X \sim \mathcal{N}(0, 1)$. This system is also compared to the generalized HSQLC [8], the linear system (BPAM) [11], the system's lower bound, and OPTA.

numerically to minimize the MSE given an average channel power constraint. Simulation results have shown that the proposed system outperforms the generalized HSQLC due to using a nonlinear bandwidth expansion on the quantization error. A lower bound on the system SDR has also been derived and is shown to be a good approximation for moderate to high CSNRs. The proposed hybrid scheme, which has low complexity and low delay, operates at 1:3 bandwidth expansion ratio. In general, to achieve higher expansion ratios, one idea would be to code the source symbol using multiple stages of quantizers and finishing with a 1:2 nonlinear analog encoder at the last stage.

REFERENCES

[1] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.

[2] ——, "Communication in the presence of noise," in *Proc. IRE*, 1949, pp. 10–21.

[3] V. A. Kotel'nikov, *The Theory of Optimum Noise Immunity*. New York: McGraw-Hill, 1959.

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

[5] F. Hekland, P. A. Floor, and T. A. Ramstad, "Shannon-Kotel'nikov mappings in joint source-channel coding," *IEEE Trans. Communications*, vol. 57, no. 1, pp. 94–105, Jan 2009.

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

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

[8] M. Kleiner and B. Rimoldi, "Asymptotically optimal joint source-channel coding with minimal delay," in *Proc. IEEE Global Telecommunications Conference*, Honolulu, HI, 2009.

[9] M. Skoglund, N. Phamdo, and F. Alajaji, "Hybrid digital-analog source-channel coding for bandwidth compression/expansion," *IEEE Trans. Information Theory*, vol. 52, no. 8, pp. 3757–3763, Aug 2006.

[10] Y. Hu, J. Garcia-Frias, and M. Lamarca, "Analog joint source channel coding using space-filling curves and MMSE decoding," in *Proc. Data Compression Conference DCC*, Snowbird, UT, Mar 2009, pp. 103–112.

[11] K. H. Lee and D. P. Petersen, "Optimal linear coding for vector channels," *IEEE Trans. Communications*, vol. COM-24, no. 12, pp. 1283–1290, 1976.