

Encoder–Decoder Design for Feedback Control over the Binary Symmetric Channel

Lei Bao, Mikael Skoglund and Karl Henrik Johansson

School of Electrical Engineering, Royal Institute of Technology, Stockholm, Sweden

Email: lei.bao@ee.kth.se, skoglund@ee.kth.se, kallej@ee.kth.se

Abstract—Encoder–decoder design is considered for a closed-loop scalar control system with feedback transmitted over a binary symmetric channel. We propose an iterative procedure which can jointly optimize adaptive encoder–decoder pairs for a certainty equivalence controller. The goal is to minimize a design criterion, in particular, the linear quadratic (LQ) cost function over a finite horizon. The algorithm leads to a practically feasible design of time-varying non-uniform encoding and decoding. Numerical results demonstrate the promising performance obtained by employing the proposed iterative optimization algorithm.

I. INTRODUCTION

In recent years, the demand for sharing resources efficiently in large networked systems has been continuously increasing. However, in many situations, there is a challenging conflict between the amount of transmitted data and the response time. In particular for emerging control applications, limits imposed on available signaling bandwidth from communication channels, can severely restrict closed-loop performance and even destabilize the system. Networked control based on limited sensor and actuator information has therefore attracted much attention during the past decade, e.g., [1], [2], [3], [4], [5], and the references therein.

Up to now, the existing work on control with limited information has mainly developed theory for error-free feedback systems. Examples of previous works investigating transmission over noisy channels include [5], [6]. The main contribution of the present paper is a practical synthesis technique for joint optimization of the encoder and the decoder for observation feedback over a noisy channel. In particular, we consider the binary symmetric channel (BSC). The corresponding scenario was considered for noiseless channels in our previous work [7]. In most practical situations, especially in a wireless environment, a noiseless assumption is unrealistic, which motivates this extension to feedback signals corrupted by channel errors. We have derived optimal encoder–decoder pairs for a given probability density function (pdf) of the plant initial state and a given cross-over probability ε . We show that this problem is numerically solvable and it leads to channel optimized encoder–decoder pairs that provide better control performance.

The paper is organized as follows. A motivating example is presented in Sec. II. Sec. III defines the control system with encoder, decoder, and communication channel. The problem statement, which concerns a linear quadratic objective over finite horizon, is presented in Sec. IV. The joint encoder–decoder design and the training procedure based on dynamic

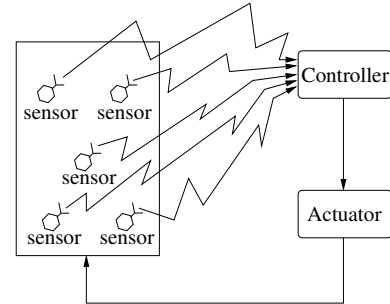


Fig. 1. Block diagram showing a control system that utilizes data from a wireless sensor network. The results in this paper can be applied to optimize the use of the wireless medium.

programming are described in Sec. V. Sec. VI presents numerical results and discusses some implementation issues. Finally the conclusions are given in Sec. VII.

II. MOTIVATING EXAMPLE

Consider the control system in Fig. 1. It consists of a large number of sensor nodes that can connect through a shared wireless medium to a control node. The sensors are spatially distributed over a large area and they measure the state of the control object, which is affected by rarely occurring local disturbances. The control commands for keeping the states around their equilibrium working points are executed through a common actuator, therefore all the sensors report their state measurements to a common control node. The described system is quite representative for many emerging applications with control using wireless sensor networks, in areas such as industrial automation, environmental monitoring, surveillance etc. The actuator is supposed to be able to counteract the disturbances, i.e., keep the states close to zero, by suitable control actions within T time samples. During this period of T time samples, the probability that a new disturbance will occur is negligible, see also the discussion in [7]. In order to efficiently utilize the communication resource, we are interested in the case that each sensor transmits only a few symbols and each symbol consists of a few bits.

Assuming disturbances are possible to detect and an estimate of the probability density function of the disturbance magnitude at each sensor location is known, we can optimize the encoder–decoder pairs for sensor transmission together with the control law. The resulting decentralized control strategy is as follows. Let the control command corresponding to a specific sensor be zero as long as no disturbance is detected.

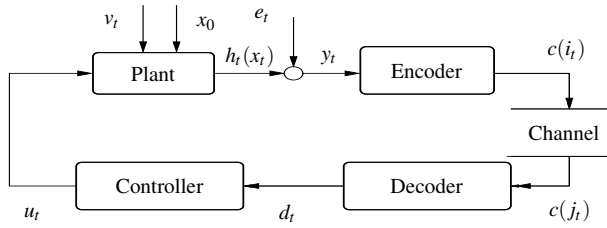


Fig. 2. Feedback control system with channel constraints.

When the sensor detects a disturbance, the sensor reading is encoded and transmitted to the controller node. The message is decoded at the controller node and a control command is derived and actuated to counteract the specific disturbance acting on the transmitting sensor. The controller and the encoder–decoder pairs are designed based on a finite-horizon linear quadratic criterion. A partial solution is therefore to solve the linear quadratic problem with uncertainty both in the initial condition and the information transmission. In this paper, a class of time-varying non-uniform encoder–decoders is derived, which together with a simple state feedback for a control system is suitable to handle this situation.

III. PRELIMINARIES

Consider the control system with a communication channel depicted in Fig. 2. In this section we describe a general version of the system. Let $\mathbf{x}_a^b = \{x_a, \dots, x_b\}$ denote the evolution of a discrete-time signal x_t from $t = a$ to $t = b$. The plant is given by the scalar system

$$\begin{aligned} x_{t+1} &= \theta_t(\mathbf{x}_{t-M_x}^t, u_t) + v_t, \\ y_t &= h_t(x_t) + e_t, \end{aligned} \quad (1)$$

which has a memory of order M_x . $x_t, u_t, y_t \in \mathbb{R}$ are the state, the control, and the measurement, respectively. The process disturbance and the measurement noise are represented by v_t and e_t , respectively.

Let \mathcal{E}_t denote the set of “information” available at the encoder, at time t , that is, the set of variables with values known to the encoder. In particular, we consider an encoder that *causally* utilizes *full information*, in the sense $\mathcal{E}_t = \{\mathbf{y}_0^t, \mathbf{i}_0^{t-1}, \mathbf{u}_0^{t-1}, \mathbf{j}_0^{t-1}\}$. The *encoder* is then a mapping from \mathcal{E}_t to a discrete set of symbols. We take each symbol to be represented by an integer index. At time t , the index is $i_t \in \mathcal{I}_L = \{0, 1, \dots, L-1\}$, where $L = 2^R$ with R denoting the *rate* of the transmission, in bits per state measurement. Hence, the encoder is described by the mapping

$$i_t = f_t(\mathbf{y}_0^t, \mathbf{i}_0^{t-1}, \mathbf{j}_0^{t-1}, \mathbf{u}_0^{t-1}). \quad (2)$$

Let the *discrete channel* have input variable i_t and output j_t , with one channel use defined by

$$j_t = k_t(\mathbf{i}_{t-M_c}^t), \quad (3)$$

where $k_t: \mathcal{I}_L^{M_c+1} \rightarrow \mathcal{I}_L$ is a random mapping, and $M_c \geq 0$ indicates (potential) channel memory.

At the receiver, we denote the information available at the decoder by \mathcal{D}_t ; in particular, we consider a decoder that

causally utilizes full information, i.e., $\mathcal{D}_t = \{\mathbf{j}_0^t, \mathbf{u}_0^{t-1}\}$. The *decoder* is then a deterministic mapping

$$d_t = g_t(\mathbf{j}_0^t, \mathbf{u}_0^{t-1}) \quad (4)$$

from \mathcal{D}_t to \mathbb{R} .

Also, let \mathcal{C}_t , in particular $\mathcal{C}_t = \{\mathbf{d}_0^t, \mathbf{u}_0^{t-1}\}$, denote the (full) information available at the controller. The *controller* is then defined by the mapping

$$u_t = z_t(\mathbf{d}_0^t, \mathbf{u}_0^{t-1}) \quad (5)$$

from \mathcal{C}_t to \mathbb{R} . We also define

$$\hat{x}_{s|t} = \mathbf{E}\{x_s | \mathbf{j}_0^t, \mathbf{u}_0^t\} \quad (6)$$

to be the minimum mean square error estimator of the state x_s , for $s \leq t$, based on \mathbf{j}_0^t and \mathbf{u}_0^t . Finally, note that even if all variables are not available at all nodes in the system, we assume that the *functionalities* of the encoder, decoder, controller and state-estimator are known at all nodes.

IV. PROBLEM STATEMENT

Let us focus on a special instance of the system in Fig. 2, namely a stable scalar linear time-invariant plant, in absence of process noise v_t and observation noise e_t . Perfect state observations are available at the encoder, $y_t = x_t$ and $\mathcal{E}_t = \{\mathbf{x}_0^t, \mathbf{i}_0^{t-1}, \mathbf{u}_0^{t-1}, \mathbf{j}_0^{t-1}\}$. The index i_t is transmitted via a BSC, and the initial state x_0 is unknown, with known pdf $p(x_0)$,

$$x_{t+1} = ax_t + u_t, \quad y_t = x_t, \quad |a| < 1. \quad (7)$$

Let $c(i_t) \in \{0, 1\}^R$ be a binary codeword of length R representing the encoder output $i_t \in \mathcal{I}_L$. The mapping between i_t and $c(i_t)$ is referred to as the *index assignment* (IA), cf. [8]. For example, the natural binary code assigns the natural binary representation of an integer i to the codeword c_i . At time t , the relation between the i_t and $c(i_t)$ is then

$$i_t = k \Leftrightarrow c(i_t) = c_k, \quad k \in \mathcal{I}_L.$$

In a similar way, $c(j_t)$ denotes the received binary codeword, where $j_t \in \mathcal{I}_L$ is the received index.

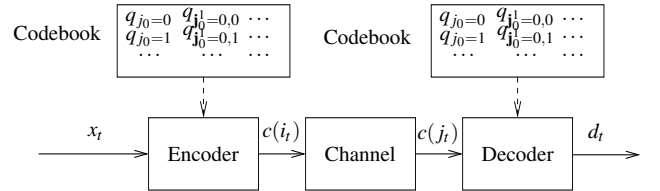


Fig. 3. There is a common codebook at the encoder and the decoder.

A BSC is described by the cross-over probability $\varepsilon = p(0|1) = p(1|0)$. Assuming independent transmission of each bit, the conditional probability of a codeword $p(c_j|c_i)$ is a function of ε ,

$$p(c_j|c_i) = (1 - \varepsilon)^{R - d_H(c_i, c_j)} \varepsilon^{d_H(c_i, c_j)}, \quad (8)$$

where $d_H(c_i, c_j)$ is the Hamming distance between the binary codewords c_i and c_j . Moreover, we assume the encoder and the decoder have access to the same codebook (Fig. 3), which consists of reconstruction points $\{q_{j_0}^t\}_{t=0}^{T-1}$, with j indicating the received index.

A. Performance Measure

The goal of this paper is to solve an integrated encoder–decoder design and optimal control problem for the system in (7). The performance measure is the linear quadratic cost function with horizon T ,

$$J_T = \mathbf{E} \left\{ x_T^2 + \sum_{t=0}^{T-1} [x_t^2 + \rho u_t^2] \right\}, \quad (9)$$

where $\rho \geq 0$ specifies the relative weight of control u_t .

B. Encoder–Decoder Structure

Intuitively, for control applications with strict delay constraints and bandwidth limitations, joint designs can improve upon traditional approaches. In this section, we present a design method, suitable for short codewords, accomplishing source compression and channel protection simultaneously, in a similar fashion to traditional channel optimized source coding techniques, e.g., [8].

The presence of a noisy channel raises an encoder–decoder synchronization problem. In case of noise-free feedback, i.e., $\mathbf{i}_0^t = \mathbf{j}_0^t$, there is no uncertainty in the information the encoder intends to convey to the decoder. That is to say, the encoder and the decoder have identical information about previous events. Such synchronization between the encoder and the decoder is eliminated when the channel produces transmission errors, since the encoder can not make error-free prediction of what index j_t the decoder will receive in advance. Even worse, the decoder has completely lost the ability to perfectly deduce the transmitted indices \mathbf{i}_0^t by examining received values for \mathbf{j}_0^{t-1} . However, since we assume a reliable actuation link, the encoder can in fact calculate j_{t-1} based on knowing x_t . Thus, the sequence \mathbf{j}_0^{t-1} can define common information about previous events, referred here to as *common memory*.

A fundamental problem in system design involving noisy channels is how to handle the impairment caused by channel errors. The two basic strategies are either to prevent in advance or handle errors after they have appeared. In this paper, we choose the first strategy, that is designing the reconstruction points taking the probability of bit errors into account. Our strategy is realized by letting both encoding and decoding be based on the received indices \mathbf{j}_0^{t-1} , cf. Fig. 3. Encoder–decoder design is then essentially equivalent to the problem of constructing an optimal strategy for the encoder to successively inform the decoder about the initial state, x_0 . Note that the encoder knows \mathbf{u}_0^t , based on \mathbf{j}_0^t . Moreover, \mathbf{x}_1^t can be computed from $(x_0, \mathbf{u}_0^{t-1})$. Hence we assume, from now on, that i_t is generated as

$$i_t = f_t(x_0, \mathbf{j}_0^{t-1}), \quad t = 0, \dots, T-1, \quad (10)$$

based on x_0 and the past indices \mathbf{j}_0^{t-1} .

At the receiver, \mathbf{u}_0^t can be computed from \mathbf{j}_0^t , hence in the following the assumed decoder mapping will be

$$d_t = g_t(\mathbf{j}_0^t), \quad t = 0, \dots, T-1. \quad (11)$$

Finally, the optimal state estimator, based on \mathcal{D}_t is

$$\hat{x}_{s|t} = \mathbf{E}\{x_s | \mathbf{j}_0^s\}, \quad t = 0, \dots, T-1, \quad s \leq t. \quad (12)$$

As will be explained later, the decoder mapping will be specialized to $d_t = \hat{x}_{t|t}$, so $\hat{x}_{t|t}$ will be utilized by the controller to produce the control output u_t .

V. ENCODER–DECODER DESIGN

This section presents the main result of the paper. After discussing certainty equivalence and encoder–decoder mappings, we present an overall training algorithm.

A. Certainty Equivalence

In general, optimal performance is achievable only when the coding and the control are designed jointly. There are examples, however, when the separation principle applies to subclasses of the problem considered in this paper. The so-called certainty equivalence controller remains optimal when the estimation errors are independent of previous control commands [9].

To apply the certainty equivalence principle to our problem, in designing the optimal decoder $\{g_t\}_0^{T-1}$ and controller $\{z_t\}_0^{T-1}$, for fixed encoder $\{f_t\}_0^{T-1}$, with $i_t = f_t(x_0, \mathbf{j}_0^{t-1})$, we note that the necessary and sufficient condition in [9] corresponds to the requirement that the average estimation error

$$\mathbf{E}\{(x_0 - \hat{x}_{0|t})^2\}, \quad t = 0, \dots, T-1, \quad (13)$$

is not a function of \mathbf{u}_0^s , for $s = 0, \dots, t-1$, for any sequence \mathbf{u}_0^s of controls. It is quite straightforward to verify that this condition holds true for the proposed scheme, since $i_0 = f_0(x_0)$, $i_1 = f_1(x_0, j_0)$, \dots , $i_t = f_t(x_0, \mathbf{j}_0^{t-1})$. Hence \mathbf{j}_0^t depend only on x_0 and potential channel errors. Consequently, based on a similar argumentation to that in [9], in absence of process noise, we can show that certainty equivalence control is optimal for fixed encoders $\{f_t\}_0^{T-1}$ with the structure $f_t(x_0, \mathbf{j}_0^{t-1})$. The optimal control signal is then given by

$$d_t = \hat{x}_{t|t}, \quad \text{and} \quad u_t = -\ell_t d_t, \quad (14)$$

(i.e., $z_t(\mathbf{d}_0^t) = -\ell_t d_t$) with

$$\ell_t = \frac{ap_{t+1}}{p_{t+1} + \rho}, \quad p_t = 1 + \frac{a^2 p_{t+1} \rho}{p_{t+1} + \rho}, \quad (15)$$

for $t = 0, \dots, T-1$, and p_t is initialized with $p_T = 1$. Note that since certainty equivalence holds, there is no loss in separating the decoder–controller into two separate entities, as done already in Fig. 2.

The remaining parts of the paper are devoted to the encoder design problem, which is the essential part of the joint design.

B. Design Algorithm

The encoder–decoder mapping consists of functions with memory, whose role is to successively provide refinements of information about the initial state x_0 . In this section we propose a training method to obtain the optimal reconstruction points. This iterative approach starts with any given initial set-up, and updates reconstruction points and encoding rules considering not only the past but also the future state evolution.

Similar to traditional quantizer design, the idea is to fix the encoder and update the decoder, then fix the decoder and

Algorithm 1: Encoder–Decoder Design

1. Initialize the encoder and decoder mappings $\{f_t, g_t\}_{t=0}^{T-1}$. Compute the controller parameters $\{\ell_t\}$ by using (15).
 2. For each $t = 0, \dots, T - 1$:
 - 2a. Update the encoder mapping f_t by using (17).
 - 2b. Update the decoder mapping g_t by using (16).
 - 2c. Set $u_t = -\ell_t d_t$.
 3. If J_T has not converged, return to Step 2, otherwise terminate the iteration.
-

update the encoder etc. With $u_t = -\ell_t d_t$, where $d_t = g_t(\mathbf{j}'_0)$, and for fixed $\{f_t\}_0^{T-1}$, the optimal decoder is

$$g_t(\mathbf{j}'_0) = \hat{x}_{t|t} = \mathbf{E}\{x_t | \mathbf{j}'_0\} = a^t q_{\mathbf{j}'_0} + \sum_{s=0}^{t-1} a^{t-1-s} u_s, \quad (16)$$

where $q_{\mathbf{j}'_0} = \mathbf{E}\{x_0 | \mathbf{j}'_0\}$ is the *reconstruction point* at time t , stored in a codebook at the decoder (and encoder, cf. Fig. 3).

The optimal encoder needs to take the impact of future state evolutions into account. Hence, for fixed decoders and controllers, and given the encoder mappings f_t for $t = 0, \dots, t - 1$ and $t + 1, \dots, T - 1$, we get the optimal $f_t = f_t(x_0, \mathbf{j}'_0^{t-1})$ as

$$i_t = \arg \min_{i \in \mathcal{I}_L} \mathbf{E} \left\{ \sum_{k=t}^T x_k^2 + \rho u_k^2 \middle| x_0, \mathbf{j}'_0^{t-1}, i_t = i \right\}, \quad i \in \mathcal{I}_L. \quad (17)$$

Note that the optimal mapping f_t indeed has the form $i_t = f_t(x_0, \mathbf{j}'_0^{t-1})$. According to (17), the encoding rule is updated once the reproduction points are recalculated. Therefore the coder is specified by the set of reconstruction points $\{q_{\mathbf{j}'_0}\}_0^{T-1}$.

Based on (15), (16) and (17), we can formulate an encoder–decoder design algorithm, i.e., Algorithm 1. The algorithm requires that the pdf $p(x_0)$ is known or described by a training set [10]. Given a pdf $p(x_0)$ for x_0 , together with transition probabilities $p(c_j | c_i)$ and \mathbf{j}'_0^{t-1} , the pdf for $\hat{x}_{0|t}$, $t < T$, can be derived. Similarly, given the pdf of the current state and $p(c_j | c_i)$, the pdfs of future estimates can also be estimated. Convergence is monitored based on updating the value of J_T in each step. Unfortunately, as mentioned, the above design cannot guarantee global optimality. Still, the algorithm converges to a local minimum, and in this sense produces a “good” solution. Also, the proposed design assumes a certain value for the cross-over probability ε of the BSC. There will be a loss in performance if the true ε deviates from the design value. However, according to our experience, the performance degradation is modest and underestimating ε is in general more harmful than overestimating it. Thus, robustness against changes in ε can be obtained by designing for a high value in the range over which the true ε is expected to vary.

VI. NUMERICAL RESULTS

In this section we present simulation results from applying the encoder–decoder design of Sec. V. We look at a numerical example with plant dynamics $a = 0.8$. The random magnitude of x_0 is modeled using the generalized Gaussian distribution,

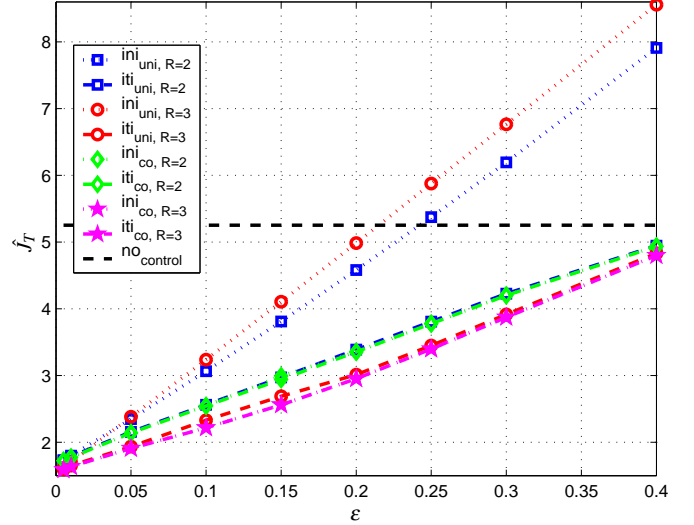


Fig. 4. Performance comparisons among encoder–decoder pairs. **ini_{uni}**: uniform initialization; **ini_{co}**: channel optimized initialization; **iti_{uni}**: iteratively trained with **ini_{uni}** as initialization; **iti_{co}**: iteratively trained with **iti_{co}** as initialization.

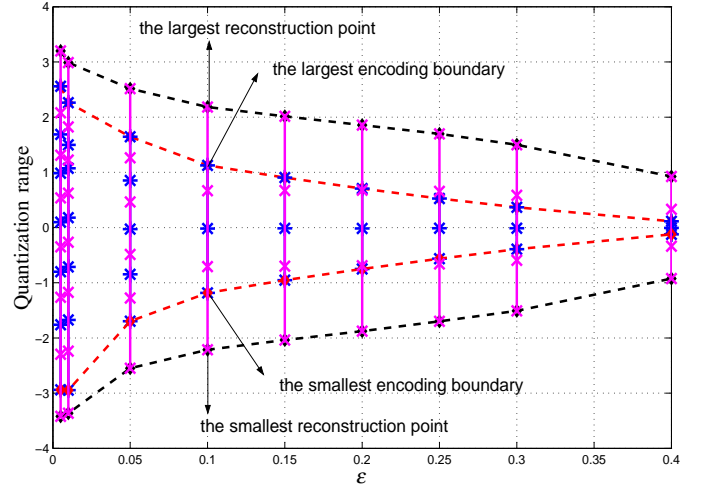


Fig. 5. Reconstruction points $q_{\mathbf{j}'_0}$ and encoding boundary versus ε . *: encoding boundary, \times : reconstruction point.

e.g., [11]. The $GGD(\alpha, \beta)$ provides a wide coverage from narrow-tailed to broad-tailed pdf’s. The parameter α describes the exponential rate of decay, and β is the standard deviation.

Fig. 4 shows simulation results for $p(x_0) = GGD(4, 2)$, rates $R = 2$ and $R = 3$. The performance is evaluated by $J_T = \sum_{t=1}^3 [x_t^2 + u_{t-1}^2]$, i.e., a modified version of (9) with $T = 3$ and $\rho = 1$. Compared to (9), here $\mathbf{E}\{x_0^2\}$ is omitted, since it affects all the results identically. As has been described, the iterative design requires an initial set-up of reconstruction points. Two initialization methods are compared. The first method, referred to as **ini_{uni}**, employs properly scaled uniform quantizers. The second method, referred to as **ini_{co}**, applies channel optimized quantizers [8] designed for each pdf $p(x_0 | \mathbf{j}'_0^{t-1})$, $t = 0, \dots, T - 1$. Starting by designing f_0 for $p(x_0)$, and using the resulting reconstruction points as initial

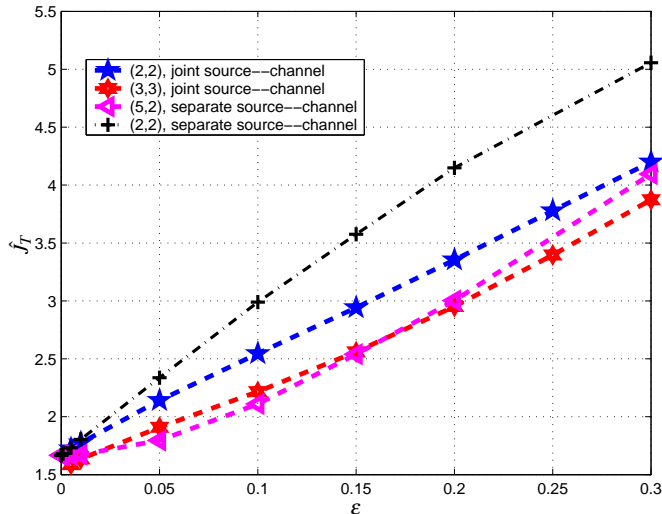


Fig. 6. Performance comparison to source-channel separate approaches. The notation (n, k) denotes that the channel code takes k input bits and provides n output bits.

estimates for $\hat{x}_{0|0}$ as well as ℓ_0 from (15), the conditional pdf's $p(x_0|j_0)$, $j_0 \in \mathcal{S}_L$, can be determined, and so on for increasing t . Fig. 4 illustrates that with iterative improvements, both initializations converge to quite similar final results, \mathbf{ini}_{uni} and \mathbf{ini}_{co} , although the performance of \mathbf{ini}_{uni} is noticeably worse than the performance of \mathbf{ini}_{co} . The computationally demanding initialization \mathbf{ini}_{co} provides fairly good results so that the improvement by training is minor. In the case of a noisy channel, a good IA is essential for the system performance. The training algorithm has the advantage that it produces a good initial IA at the same time as generating the encoder-decoder pairs (cf. [8]).

The horizontal dashed line in Fig. 4 represents the cost \hat{J}_T if no control action is taken. The figure demonstrates that there is still some advantage to feedback state observation on a noisy channel even with 40% bit errors. An interesting problem would be to examine the trade-off between ϵ and establish how coarse the feedback information can be, in order to meet a certain performance requirement.

For stable systems, the contribution $x_t^2 + \rho u_t^2$ to the total cost (9) decreases with time. Hence, the reconstruction points at $t = 0$, $\{q_{j_0=k}\}_{k=0}^{L-1}$, are the most important design parameters. Fig. 5 shows the resulting encoding boundaries and the reconstruction points at $t = 0$ versus ϵ . We can note that the number of reconstruction points chosen by the encoder decreases with increasing ϵ . This phenomenon is well known in quantization for noisy channels and is attributed to the varying abilities of binary codewords in combating channel errors. For very noisy channels, it is beneficial to transmit only the “stronger” codewords [8], providing true redundancy for error protection. Another impact of increasing ϵ is that the ranges of the reconstruction points decrease so the boundaries and the reconstruction points are all moved closer to zero, indicating that only small-valued control actions are allowed.

A simple comparison between optimized joint source-

channel coding and separate source and channel coding is illustrated in Fig. 6. The figure shows $R = 2$ and $R = 3$ joint approaches, designed according to Sec. V-B at the different values for ϵ . The figure furthermore shows the performance of $R = 2$ source coding combined with a $(5, 2)$ shortened Hamming code. The source code is trained according to Sec. V-B, under the assumption $\epsilon = 0$. The channel code has minimum distance 3, so it can correct all single-bit errors and some two-bit errors. The performance of using only the $R = 2$ source code without channel coding is also illustrated. As can be seen, training for the correct ϵ gives a clear gain, and the $R = 3$ joint source-channel approach performs better than the two schemes with $R = 2$ and no channel coding. However, the $R = 2$ source-optimized code combined with the $(5, 2)$ channel code has the best performance in the interval $0.01 < \epsilon < 0.15$ (approximately). Note that the gain is not large and it comes at the price of significantly increased bandwidth (from 3 to 5 bits per sample).

VII. CONCLUSION

This paper has introduced an approach that jointly optimizes the encoder and decoder in closed-loop control of a linear plant with low-rate feedback over a binary symmetric channel. Having argued that a fixed sequence of encoder mappings will result in the certainty equivalence condition being fulfilled, we fixed the receiver structure to implement separate decoding (estimation) and control. Our main contribution is to introduce optimized encoders given this structure on the estimation-control. We have performed various numerical investigations of the proposed optimization algorithm and illustrated quite promising results.

REFERENCES

- [1] V. S. Borkar and S. K. Mitter, “LQG control with communication constraints,” in *Proc. of the 38th Conference on Decision and Control*, Dec. 1999.
- [2] R. W. Brockett and D. Liberzon, “Quantized feedback stabilization of linear systems,” *IEEE Transactions on Automatic Control*, vol. 45, no. 7, July 2000.
- [3] F. Fagnani and S. Zampieri, “Stability analysis and synthesis for scalar linear systems with a quantized feedback,” *IEEE Transactions on Automatic Control*, vol. 48, no. 9, Sept. 2003.
- [4] G. N. Nair and R. J. Evans, “Stabilizability of stochastic linear systems with finite feedback data rates,” *SIAM Jour. Contr. Opt.*, Feb. 2002.
- [5] S. Tatikonda, A. Sahai, and S. Mitter, “Stochastic linear control over a communication channel,” *IEEE Transactions on Automatic Control*, vol. 49, no. 9, Sept. 2004.
- [6] T. Simsek, R. Jain, and P. Varaiya, “Scalar estimation and control with noisy binary observation,” *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1598–1603, Sept. 2004.
- [7] L. Bao, M. Skoglund, and K. H. Johansson, “Encoder-decoder design for event-triggered feedback control over bandwidth limited channels,” *2006 American Control Conference*, to appear.
- [8] N. Farvardin, “A study of vector quantization for noisy channels,” *IEEE Transactions on Information Theory*, vol. 36, no. 4, pp. 799–809, July 1990.
- [9] Y. Bar-Shalom and E. Tse, “Dual effect, certainty equivalence, and separation in stochastic control,” *IEEE Transactions on Automatic Control*, vol. AC-19, no. 5, Oct. 1974.
- [10] A. Gersho and R. M. Gray, *Vector quantization and signal compression*. Kluwer, 1992.
- [11] N. Farvardin and J. W. Modestino, “Optimum quantizer performance for a class of non-Gaussian memoryless sources,” *IEEE Transactions on Information Theory*, vol. 30, no. 3, pp. 485–497, May 1984.