Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 49-22 (2016) 133–138

Optimal Block Length for Data-Rate Minimization in Networked LQG Control

Takashi Tanaka* **Karl Henrik Johansson**** **Mikael Skoglund*****

KTH Royal Institute of Technology, Stockholm, Sweden
* *ttanaka@kth.se* ** *kallej@kth.se* *** *skoglund@kth.se*

Abstract: We consider a discrete-time networked LQG control problem in which state information must be transmitted to the controller over a noiseless binary channel using prefix-free codewords. Quantizer, encoder and controller are jointly designed to minimize average data-rate while satisfying required LQG control performance. We study the effects of selecting large block-lengths (data transmission intervals) from the perspectives of information-theoretic advantage due to coding efficiency and control-theoretic disadvantage due to delay. In particular, we demonstrate that the performance of networked control scheme by Tanaka et al. (2016) can be improved by adjusting the block-length optimally. As a by-product of this study, we also show that the data-rate theorem for mean-square stability similar to Nair and Evans (2004) can be recovered by considering sufficiently large block-lengths.

Keywords: LQG control, Quantization, Sampling rates, Huffman codes, Data compression, Stability.

1. INTRODUCTION

In this paper, we consider a discrete-time state-feedback control system in which state information must be transmitted to the controller through a noiseless binary channel at every time step. As depicted in Figure 1, we assume that the plant to be controlled has a linear time-invariant state space model [1]

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + B\mathbf{u}_t + F\mathbf{w}_t \tag{1}$$

where $\mathbf{x}_t$ is the $\mathbb{R}^n$-valued state process, $\mathbf{u}_t$ is the $\mathbb{R}^m$-valued control input, and $\mathbf{w}_t$ is the $\mathbb{R}^p$-valued white Gaussian noise with unit covariance I_p. Assume that $\det(A) \neq 0$ and (A, B) is a stabilizable pair. For every non-negative integer t, let

$$\mathcal{A}_t \subset \{0,1\}^\star \triangleq \{0, 1, 00, 01, 10, 11, 000, \cdots\}$$

be a set of prefix-free variable-length binary codewords (Cover and Thomas, 1991, Ch.5). There are multiple ways to choose a codebook $\mathcal{A}_t$, and we allow the choice to be time-varying. Designing appropriate codebooks $\mathcal{A}_0^\infty = (\mathcal{A}_0, \mathcal{A}_1, \cdots)$ is part of our design problem. We also design an encoder policy

$$\mathcal{E}_0^\infty \triangleq \{e_t(a_t|x_0^t, a_0^{t-1})\}_{t=0,1,\ldots} \tag{2}$$

which is a sequence of Borel measurable stochastic kernels [2] on $\mathcal{A}_t$ given $\mathcal{X}_0^t \times \mathcal{A}_0^{t-1}$, and a controller policy

$$\mathcal{D}_0^\infty \triangleq \{d_t(u_t|a_0^t, u_0^{t-1})\}_{t=0,1,\ldots} \tag{3}$$

which is a sequence of Borel measurable stochastic kernels on $\mathcal{U}_t$ given $\mathcal{A}_0^t \times \mathcal{U}_0^{t-1}$. Notice that $\mathcal{E}_0^\infty$ and $\mathcal{D}_0^\infty$ are general (possibly non-deterministic) causal decision policies. The length of a codeword $a_t \in \mathcal{A}_t$ transmitted from the encoder to the controller at time step t will be denoted by a random variable $\boldsymbol{\ell}_t$. A triplet $(\mathcal{A}_0^\infty, \mathcal{E}_0^\infty, \mathcal{D}_0^\infty)$ will be called a *design*, and the space of such designs is denoted by Γ.

[1] In this paper, random variables are denoted by lower case bold symbols such as $\mathbf{x}$. Calligraphic symbols such as $\mathcal{X}$ are used to denote sets, and $x \in \mathcal{X}$ is an element. We denote by x_0^t a sequence $x_0, x_1, ..., x_t$, and $\mathbf{x}_0^t$ and $\mathcal{X}_0^t$ are understood similarly.

[2] Foundational discussions on stochastic kernels can be found in Bertsekas and Shreve (1978).

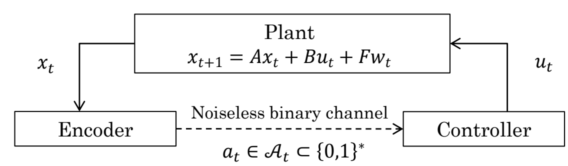

Fig. 1. LQG control over noiseless binary channel.

In this paper, we are interested in a design $(\mathcal{A}_0^\infty, \mathcal{E}_0^\infty, \mathcal{D}_0^\infty)$ that minimizes data rate (i.e., average expected codeword length per time step) while satisfying a required level of LQG control performance γ. Formally, we consider an optimization problem:

$$R(\gamma) \triangleq \min_{\Gamma} \limsup_{T\to+\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}(\boldsymbol{\ell}_t) \tag{4a}$$

$$\text{s.t.} \limsup_{T\to+\infty} \frac{1}{T}\sum_{t=0}^{T-1} \mathbb{E}c(\mathbf{x}_t, \mathbf{u}_t) \leq \gamma. \tag{4b}$$

The cost function $c(\mathbf{x}_t, \mathbf{u}_t)$ is given by a positive definite quadratic form

$$c(\mathbf{x}_t, \mathbf{u}_t) = \begin{bmatrix} \mathbf{x}_t \\ \mathbf{u}_t \end{bmatrix}^\top \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \begin{bmatrix} \mathbf{x}_t \\ \mathbf{u}_t \end{bmatrix}.$$

In the optimization problem (4), we require the *average* data rate to be minimized. However, the number of bits $\boldsymbol{\ell}_t$ to be transmitted in a particular time step t can be arbitrary large. Problem (4) is motivated by a common engineering situation in which a digital communication channel must be shared by multiple control systems. For instance, many real-time applications share a common communication bus in an automobile system (e.g., Johansson et al. (2005)). Even though the communication requirements by individual applications might be small, they could collectively cause a serious packet congestion in the shared bus. It is therefore important for individual control applications to minimize their channel use in order to utilize a shared communication resource efficiently.

10.1016/j.ifacol.2016.10.385

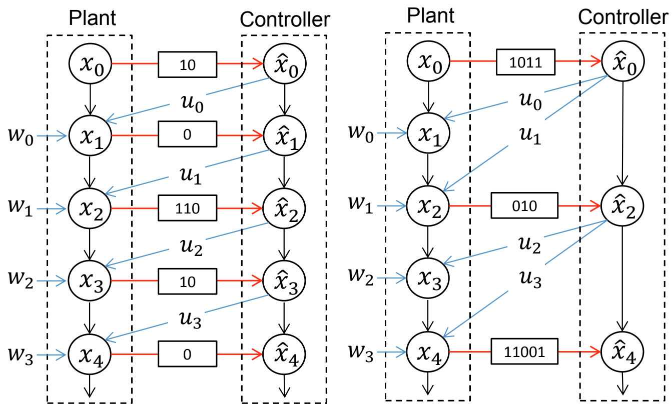

Fig. 2. Standard design (left) and design with block-length two (right).

Obtaining the exact solution to (4) is computationally challenging. However, it is shown by Silva et al. (2011) that a lower bound $DI(\gamma) \leq R(\gamma)$ is obtained by solving a convex optimization problem

$$DI(\gamma) \triangleq \min_{\tilde{\Gamma}} \limsup_{T\to+\infty} \frac{1}{T} I(\mathbf{x}_0^{T-1} \to \mathbf{u}_0^{T-1}) \tag{5a}$$

$$\text{s.t.} \limsup_{T\to+\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}c(\mathbf{x}_t, \mathbf{u}_t) \leq \gamma \tag{5b}$$

where $I(\mathbf{x}_0^{T-1} \to \mathbf{u}_0^{T-1}) \triangleq \sum_{t=0}^{T-1} I(\mathbf{x}_0^t; \mathbf{u}_t | \mathbf{u}_0^{t-1})$ is called *directed information*. Here, we employ a definition of directed information by Massey (1990), while several generalized definitions are proposed in the literature. Minimization in (5a) is over the space $\tilde{\Gamma}$ of the sequences of stochastic kernels $\{p_t(u_t|x_0^t, u_0^{t-1})\}_{t=0,1,\ldots}$.

Silva et al. (2011) also show that $R(\gamma)$ is upper bounded by $DI(\gamma)+c$, where the c is a scalar constant. Their proof is based on the construction of an entropy-coded dithered quantizer (ECDQ) whose rate loss due to the space-filling loss of the lattice quantizers and the loss of entropy coders is bounded by c. This result, although restricted to SISO control systems in the original paper, is recently extended to MIMO control systems by Tanaka et al. (2016). For completeness, we present the synthesis procedure proposed there in Section 3.

Although the quantizer/controller design methods proposed in Silva et al. (2011) and Tanaka et al. (2016) are suboptimal (with bounded performance loss), they are relatively simple to implement, since they are built on uniform quantizers. The optimal quantizer/controller design is much more involved; iterative design algorithms and structural results are studied by Bao et al. (2011) and Yüksel (2014).

The purpose of this paper is to demonstrate that the performance of the design in Silva et al. (2011) and Tanaka et al. (2016) can be further improved by choosing the *block-length*[3] optimally. If the block-length is k, the state information is quantized, encoded, and transmitted to the controller only once in every k time steps. In other time steps, no information is transmitted from the encoder to the controller, i.e., $\ell_t = 0$ for t such that $t \bmod k \neq 0$. This is visualized in Figure 2.

[3] In this paper *block-length* is used as a synonym for *data transmission intervals* or *sampling intervals*.

There are pros and cons of selecting block-lengths larger than one. They can be intuitively seen in the following trade-off:

(A) From a control-theoretic viewpoint, selecting large block-lengths is a disadvantage. Notice that control input $\mathbf{u}_t$ can depend on the state sequence only up to $\mathbf{x}_{t'}$, where t' is the largest integer such that $t' \leq t$ and $t' \bmod k = 0$. In particular, large block-length implies large delay.
(B) From an information-theoretic viewpoint, transmitting high-resolution data once in a while rather than transmitting low-resolution data at every time step is beneficial, since the former allows more efficient data compression. Namely, the effect of space-filling loss due to quantization and the entropy coding loss per time step can be made smaller.

In this paper, we perform a quantitative study on this trade-off using theoretical lower and upper bounds of the best achievable rates. Since the best achievable rates are still difficult to evaluate exactly, it remains difficult to determine the optimal block-length analytically. Hence, we also perform a numerical study to see how the performance of Tanaka et al. (2016) varies with different block-lengths.

2. RELATED WORK

LQG control problems with data-rate constraints have been tackled by many papers from various angles. Joint controller and quantizer design was considered by Bao et al. (2011) and Yüksel (2014). Extensions of the classical separation principle are discussed in Matveev and Savkin (2004), Fu (2009) and You and Xie (2011). Rate-performance trade-off studies can be found in Tatikonda et al. (2004), Huang et al. (2005), Lemmon and Sun (2006) and Freudenberg et al. (2011). In particular, Silva et al. (2011) establishes a connection between (4) and (5). Many important results that cannot be mentioned here can be found in Yüksel and Başar (2013).

Optimal block-length for data-rate minimization in LQG control is considered in Borkar and Mitter (1997). A similar problem for controlled hidden Markov chain is considered in Tan et al. (2004). Although these papers are closest in spirit to this paper, these analyses involve computationally intractable steps (e.g., dynamic programming) for the block-length optimization, which restrict venues for quantitative studies. In this paper, we apply newly obtained theoretical upper and lower bounds (Silva et al. (2011)) and synthesis (Tanaka et al. (2016)) to make the study more quantitative and computationally accessible.

In this paper, we assume time-triggered communications between encoder and controller. However, average data-rate minimization may also be attained by event-triggered communications. Åström and Bernhardsson (2002) compared conventional periodic (Riemann) sampling and event-based (Lebesgue) sampling strategies, and showed that the latter achieves a better performance in a certain context. This observation is extended by Cervin and Henningsson (2008) to the situations in which multiple control loops share a communication medium. Information-theoretic framework for event-triggered control is recently considered by Tallapragada and Cortés (2014).

3. NETWORKED CONTROLLER DESIGN

In this section, we summarize the result obtained by Tanaka et al. (2016). A control architecture proposed there for problem (4) (Figure 1) is shown in Figure 3. Components in Figure 3 can be constructed by the following procedure.

(1) Determine a controller gain by

$$K \triangleq -(B^\top XB + R)^{-1}(B^\top XA + S^\top)$$

where X is the unique positive definite solution to the algebraic Riccati equation

$$A^\top XA - X + Q$$
$$-(A^\top XB + S)(B^\top XB + R)^{-1}(B^\top XA + S^\top) = 0.$$

(2) Solve a semidefinite programming problem [4]

$$DI(\gamma) = \min_{P,\Pi\succ 0} \frac{1}{2}\log\det\Pi^{-1} + \log|\det A| \tag{6}$$
$$\text{s.t. } \operatorname{Tr}(\Theta P) + \operatorname{Tr}(F^\top XF) \le \gamma$$
$$P \preceq APA^\top + FF^\top$$
$$\begin{bmatrix} I-\Pi & F^\top \\ F & APA^\top + FF^\top \end{bmatrix} \succeq 0$$

where $\Theta \triangleq K^\top(B^\top XB + R)K$.

(3) Let $r \ge 0$ be the rank of a positive semidefinite matrix

$$\mathsf{SNR}(\gamma) \triangleq P(\gamma)^{-1} - (AP(\gamma)A^\top + FF^\top)^{-1}$$

where $P(\gamma)$ is the optimal solution to (6) obtained in step (2). Choose a matrix $C \in \mathbb{R}^{r\times n}$ with orthonormal rows and a positive definite diagonal matrix V by the singular value decomposition $C^\top V^{-1}C = \mathsf{SNR}(\gamma)$.

(4) Construct a control architecture shown in Figure 3, where

(a) $Q_\Delta = (Q_{\Delta_1}, ..., Q_{\Delta_r})$ denotes uniform quantizers such that for each $i = 1, ..., r$, $Q_{\Delta_i}(\cdot)$ is defined by

$$Q_{\Delta_i}(x) = j\Delta_i \text{ for } j\Delta_i - \tfrac{\Delta_i}{2} \le x < j\Delta_i + \tfrac{\Delta_i}{2}$$

where Δ_i is the quantizer step size such that

$$\tfrac{\Delta_i^2}{12} = V_{ii} \text{ (the } i\text{-th diagonal entry of } V).$$

(b) $\boldsymbol{\xi}_t$ is an $\mathbb{R}^r$-valued i.i.d. dither signal (an artificial noise) such that its i-th entry $\boldsymbol{\xi}_{t,i}$ has a uniform distribution $\mathcal{U}[-\frac{\Delta_i}{2}, \frac{\Delta_i}{2}]$.

(c) $\mathbf{q}_t = (\mathbf{q}_{t,1}, ..., \mathbf{q}_{t,r})$ with $\mathbf{q}_{t,i} = Q_{\Delta_i}(\boldsymbol{\theta}_{t,i} + \boldsymbol{\xi}_{t,i})$ is a quantized signal taking values in

$$\tilde{\mathbb{Z}}^r \triangleq \{(j_1\Delta_1, ..., j_r\Delta_r) : (j_1, ..., j_r) \in \mathbb{Z}^r\},$$

(d) $(E(\xi_t), D(\xi_t))$ is a Shannon-Fano encoder-decoder pair for a countably infinite alphabet $\tilde{\mathbb{Z}}^r$, adjusted to the conditional probability distribution $p(\mathbf{q}_t|\boldsymbol{\xi}_t = \xi_t)$.

(e) the "Kalman filter" represents a recursive process

$$\hat{\mathbf{x}}_{t|t-1} = A\hat{\mathbf{x}}_{t-1} + B\mathbf{u}_{t-1}$$
$$\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_{t|t-1} + L\mathbf{q}_t$$

with $L = YC^\top(CYC^\top + V)^{-1}$ where Y is the unique positive definite solution to the algebraic Riccati equation

$$AYA^\top - Y + FF^\top - AYC^\top(CYC^\top + V)^{-1}CYA^\top = 0.$$

Theorem 1. For every $\gamma > \operatorname{Tr}(F^\top XF)$, $DI(\gamma)$ defined by (5) can be computed by (6) and provides a lower bound of the achievable rate $R(\gamma)$ defined by (4a). Moreover, the control architecture in Figure 3 designed by steps (1)-(4) achieves the control performance γ and the average data rate $\bar{R}(\gamma)$ strictly smaller than $DI(\gamma) + \frac{r}{2}\log\frac{4\pi e}{12} + 1$, where $r \triangleq \operatorname{rank}(\mathsf{SNR}(\gamma)) \le n$. In particular, we have

$$DI(\gamma) \le R(\gamma) \le \bar{R}(\gamma) < DI(\gamma) + \frac{r}{2}\log\frac{4\pi e}{12} + 1. \tag{7}$$

Proof. See Tanaka et al. (2016). Basic form of this result is attributed to Silva et al. (2011).

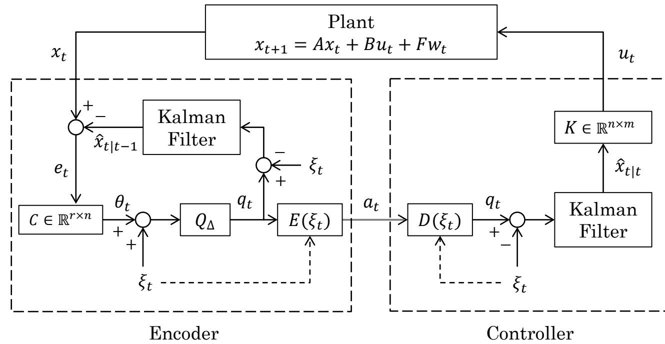

Fig. 3. Proposed architecture for LQG control.

Remark 1. In Figure 3, encoder and controller must share the same realization of the dither signal $\boldsymbol{\xi}_t$. This is a common difficulty in the implementation of subtractive dithered quantizers (see, e.g., Gray and Stockham (1993)). In practice, pseudo-random number generators with shared seeds are used by both parties.

Remark 2. Step (4)-(d) cannot be performed in practice, since this step requires different codebooks for different realizations of $\boldsymbol{\xi}_t$ (there are infinite possibilities). Moreover, for a fixed ξ_t, true distribution $p(\mathbf{q}_t|\boldsymbol{\xi}_t = \xi_t)$ is difficult to obtain and thus Shannon-Fano code cannot be designed optimally. In our simulation study in Section 6, we make a modification to this step. First, we obtain the probability distribution $p(\mathbf{q}_t)$ empirically by means of simulation. Then, we design the Huffman encoder-decoder pair (E, D) adapted to the obtained probability distribution $p(\mathbf{q}_t)$, and the same (E, D) pair is used for all realizations of $\boldsymbol{\xi}_t$. After this modification, the result of Theorem 1 cannot be guaranteed. However, our simulation study shows that Theorem 1 predicts the performance of the controller well even if these modifications are applied.

4. OPTIMAL BLOCK-LENGTH

As in Borkar and Mitter (1997), we design a control strategy with block-length k by simply applying control design scheme presented in Section 3 to the k-lifted state space model. Then we analyze the performance of k-lifted design using Theorem 1.

4.1 k-lifted system

Given a positive integer k, the *k-lifted system* corresponding to the original system (1) has $\mathbf{x}_\tau^{(k)} \triangleq \mathbf{x}_{k\tau}$ as the state vector, where $\tau = 0, 1, 2, ...$ is a new time index. For each $\tau = 0, 1, 2, ...$, new control and noise inputs are defined by

$$\mathbf{u}_\tau^{(k)} \triangleq \begin{bmatrix} \mathbf{u}_{k\tau} \\ \mathbf{u}_{k\tau+1} \\ \vdots \\ \mathbf{u}_{k\tau+k-1} \end{bmatrix}, \mathbf{w}_\tau^{(k)} \triangleq \begin{bmatrix} \mathbf{w}_{k\tau} \\ \mathbf{w}_{k\tau+1} \\ \vdots \\ \mathbf{w}_{k\tau+k-1} \end{bmatrix}.$$

Introducing new matrices $A^{(i)}, B^{(i)}, F^{(i)}$ for $i = 0, ..., k$ and $E^{(i)}$ for $i = 0, ..., k-1$, [5] the state space model of the k-lifted system corresponding to the original system (1) is

$$\mathbf{x}_{\tau+1}^{(k)} = A^{(k)}\mathbf{x}_\tau^{(k)} + B^{(k)}\mathbf{u}_\tau^{(k)} + F^{(k)}\mathbf{w}_\tau^{(k)}. \tag{8}$$

[4] In Tanaka et al. (2016), SDP (6) is written in a different form. It can be shown that the expressions are equivalent if $\det A \neq 0$ and $W \triangleq FF^\top$ has full-rank.

[5] We define $A^{(i)} \triangleq A^i$, $B^{(i)} \triangleq \begin{bmatrix} A^{i-1}B & \cdots & AB & B & 0_{n\times m(k-i)} \end{bmatrix}$, $F^{(i)} \triangleq \begin{bmatrix} A^{i-1}F & \cdots & AF & F & 0_{n\times p(k-i)} \end{bmatrix}$ for each $i = 0, 1, 2, \cdots, k$. We consider $A^{(0)} = I_n$, $B^{(0)} = 0_{n\times mk}$ and $F^{(0)} = 0_{n\times pk}$. We also define $E^{(i)} \triangleq \begin{bmatrix} 0_{m\times mi} & I_{m\times m} & 0_{m\times m(k-i-1)} \end{bmatrix}$ for $i = 0, 1, 2, \cdots, k-1$.

The original state and control inputs are related to $\mathbf{x}_\tau^{(k)}, \mathbf{u}_\tau^{(k)}$ and $\mathbf{w}_\tau^{(k)}$ by

$$\mathbf{x}_{k\tau+i} = A^{(i)}\mathbf{x}_\tau^{(k)} + B^{(i)}\mathbf{u}_\tau^{(k)} + F^{(i)}\mathbf{w}_\tau^{(k)}$$
$$\mathbf{u}_{k\tau+i} = E^{(i)}\mathbf{u}_\tau^{(k)}$$

where $i = 0, 1, ..., k-1$. We also introduce a new cost function

$$c^{(k)}(\mathbf{x}_\tau^{(k)}, \mathbf{u}_\tau^{(k)}) \triangleq \begin{bmatrix}\mathbf{x}_\tau^{(k)} \\ \mathbf{u}_\tau^{(k)}\end{bmatrix}^\top \begin{bmatrix}Q^{(k)} & S^{(k)} \\ S^{(k)\top} & R^{(k)}\end{bmatrix}\begin{bmatrix}\mathbf{x}_\tau^{(k)} \\ \mathbf{u}_\tau^{(k)}\end{bmatrix} + \sum_{i=0}^{k-1}\mathrm{Tr}(F^{(i)\top}QF^{(i)}) \quad (9)$$

where

$$\begin{bmatrix}Q^{(k)} & S^{(k)} \\ S^{(k)\top} & R^{(k)}\end{bmatrix} = \sum_{i=0}^{k-1}\begin{bmatrix}A^{(i)} & B^{(i)} \\ 0 & E^{(i)}\end{bmatrix}^\top \begin{bmatrix}Q & S \\ S^\top & R\end{bmatrix}\begin{bmatrix}A^{(i)} & B^{(i)} \\ 0 & E^{(i)}\end{bmatrix}.$$

We have selected a new cost function so that (9) is equal to the value of the original cost function aggregated over $t = k\tau, k\tau + 1, ..., k\tau + k - 1$. It is easy to verify that

$$\sum_{i=0}^{k-1}\mathbb{E}c(\mathbf{x}_{k\tau+i}, \mathbf{u}_{k\tau+i}) = \mathbb{E}c^{(k)}(\mathbf{x}_\tau^{(k)}, \mathbf{u}_\tau^{(k)})$$

holds under the condition that $(\mathbf{x}_\tau^{(k)}, \mathbf{u}_\tau^{(k)})$ and $\mathbf{w}_\tau^{(k)}$ are mutually independent. This condition is satisfied by all *k-lifted designs*, which will be defined next. [6]

4.2 k-lifted design

Let $\mathcal{A}_\tau^{(k)}$ be the set of prefix-free variable-length binary codewords $a_\tau^{(k)}$ transmitted at time $t = k\tau$. The length of $a_\tau^{(k)}$ is a random variable $\boldsymbol{\ell}_\tau^{(k)}$. Note that we require $\mathcal{A}_t = \{\phi\}$ for all t such that $t \bmod k \neq 0$. Hence the encoder and decoder policies are adequately modified to

$$\{\mathcal{E}^{(k)}\}_0^\infty \triangleq \{e_\tau^{(k)}(a_\tau^{(k)}|\{x^{(k)}\}_0^\tau, \{a^{(k)}\}_0^{\tau-1})\}_{\tau=0,1,...}$$
$$\{\mathcal{D}^{(k)}\}_0^\infty \triangleq \{d_\tau^{(k)}(u_\tau^{(k)}|\{a^{(k)}\}_0^\tau, \{u^{(k)}\}_0^{\tau-1})\}_{\tau=0,1,...}.$$

The space of *k-lifted design* $(\{\mathcal{A}^{(k)}\}_0^\infty, \{\mathcal{E}^{(k)}\}_0^\infty, \{\mathcal{D}^{(k)}\}_0^\infty)$ is denoted by $\Gamma^{(k)}$. The minimum average rate per time step is defined by

$$R^{(k)}(\gamma) \triangleq \min_{\Gamma^{(k)}}\limsup_{T\to+\infty}\frac{1}{kT}\sum_{\tau=0}^{T-1}\mathbb{E}(\boldsymbol{\ell}_\tau^{(k)}) \quad (10)$$
$$\text{s.t.} \limsup_{T\to+\infty}\frac{1}{kT}\sum_{\tau=0}^{T-1}\mathbb{E}c^{(k)}(\mathbf{x}_\tau^{(k)}, \mathbf{u}_\tau^{(k)}) \leq \gamma.$$

Notice that the factor k in the denominator is introduced so that rate and control cost are evaluated per time step in the original time scale. If $k = 1$, (10) is equivalent to (4). Similarly, we define

$$DI^{(k)}(\gamma) \triangleq \min_{\tilde{\Gamma}^{(k)}}\limsup_{T\to+\infty}\frac{1}{kT}I(\{\mathbf{x}^{(k)}\}_0^{T-1} \to \{\mathbf{u}^{(k)}\}_0^{T-1}) \quad (11)$$
$$\text{s.t.} \limsup_{T\to+\infty}\frac{1}{kT}\sum_{\tau=0}^{T-1}\mathbb{E}c^{(k)}(\mathbf{x}_\tau^{(k)}, \mathbf{u}_\tau^{(k)}) \leq \gamma.$$

for the k-lifted state space model (8). Minimization in (11) is over the space $\tilde{\Gamma}^{(k)}$ of the sequences of stochastic kernels

$$\{p_\tau^{(k)}(u_\tau^{(k)}|\{x^{(k)}\}_0^\tau, \{u^{(k)}\}_0^{\tau-1})\}_{\tau=0,1,...}.$$

[6] For instance, in Figure 2 (right), $(\mathbf{x}_2, \mathbf{u}_2, \mathbf{u}_3)$ and $(\mathbf{w}_2, \mathbf{w}_3)$ are mutually independent.

The function $DI^{(k)}(\gamma)$ has a semidefinite representation similar to (6), except that matrices A, B, F, Q, R and S are replaced by $A^{(k)}, B^{(k)}, F^{(k)}, Q^{(k)}, R^{(k)}$ and $S^{(k)}$.

Let $\bar{R}^{(k)}(\gamma)$ be the average rate of the k-lifted design obtained by applying the design procedure in Section 3 to the control problem (10). It follows from Theorem 1 that for every $\gamma > 0$,

$$DI^{(k)}(\gamma) \leq R^{(k)}(\gamma) \leq \bar{R}^{(k)}(\gamma) < DI^{(k)}(\gamma) + \frac{r}{2k}\log\frac{4\pi e}{12} + \frac{1}{k}. \quad (12)$$

4.3 Block-length selection

The inequality (12) has the following implications. First, observe that $DI(\gamma) \leq DI^{(k)}(\gamma)$ holds for every $\gamma > 0$. [7] This means that the lower bound of the best achievable rate increases by choosing $k > 1$, and as a result some rate-performance region becomes theoretically unattainable. This is a disadvantage of choosing $k > 1$, which is related to the item (A) mentioned in Section 1. Second, the effect of space-filling loss $\frac{r}{2k}\log\frac{4\pi e}{12}$ and the entropy coding loss $\frac{1}{k}$ can be suppressed by selecting a large k. This is an advantage corresponds to the item (B) in Section 1. Also, large k helps to tighten the gap between the lower and upper bounds in (12). This means that the rate performance of the proposed control design can be made more predictable by selecting large block-lengths.

Although (12) roughly estimates how $R^{(k)}(\gamma)$ and $\bar{R}^{(k)}(\gamma)$ depend on k, it is still difficult determine the optimal k that minimizes these quantities. Thus, in Section 6, we numerically demonstrate that selecting $k > 1$ can indeed reduce $\bar{R}^{(k)}(\gamma)$ in certain occasions, and that the optimal k also depends on γ.

5. DATA-RATE THEOREM FOR STABILIZATION

Another interesting result is obtained by taking the limit $k \to +\infty$ in (12). This leads to an achievability proof of the data-rate for mean-square stability similar to Nair and Evans (2004). Denote by $\sigma(A)$ the collection of eigenvalues of A counted with multiplicity. Introduce a quantity

$$\mathcal{H}(A) \triangleq \sum_{\lambda_i\in\sigma(A)}\max\{\log|\lambda_i|, 0\}. \quad (13)$$

From the semidefinite representation (6), it can be directly shown (Tanaka et al., 2015, Corollary 1) that

$$\lim_{\gamma\to+\infty}DI(\gamma) = \mathcal{H}(A).$$

Since $DI^{(k)}(\gamma)$ has a semidefinite representation similar to (6), it can be easily shown that

$$\lim_{\gamma\to+\infty}DI^{(k)}(\gamma) = \frac{1}{k}\mathcal{H}(A^k) = \mathcal{H}(A). \quad (14)$$

With this observation, we obtain the next Theorem.

Theorem 2. Define $R^{(k)}(\gamma)$ by (10) and $\mathcal{H}(A)$ by (13). Then

$$\lim_{k\to+\infty}\lim_{\gamma\to+\infty}R^{(k)}(\gamma) = \mathcal{H}(A). \quad (15)$$

Moreover, for every $\epsilon > 0$, there exists a k-lifted design with sufficiently large block-length k such that the average rate is at most $\mathcal{H}(A) + \epsilon$ and the mean-square stability is attained.

[7] To prove $DI(\gamma) \leq DI^{(k)}(\gamma)$, use the fact that for every policy in $\tilde{\Gamma}^{(k)}$, equalities $I(\{\mathbf{x}^{(k)}\}_0^{T-1} \to \{\mathbf{u}^{(k)}\}_0^{T-1}) = I(\mathbf{x}^{kT-1} \to \mathbf{u}^{kT-1})$ and $\sum_{\tau=0}^{T-1}\mathbb{E}c^{(k)}(\mathbf{x}_\tau^{(k)}, \mathbf{u}_\tau^{(k)}) = \sum_{t=0}^{kT-1}\mathbb{E}c(\mathbf{x}_t, \mathbf{u}_t)$ hold. This means that (5) and (11) are equivalent optimization problems when the domain is restricted to $\tilde{\Gamma}^{(k)}$. Next, observe that $\tilde{\Gamma}^{(k)}$ can be considered as a subset of $\tilde{\Gamma}$.

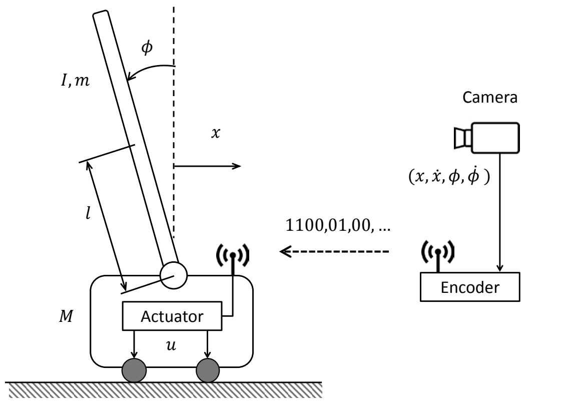

Fig. 4. Networked LQG control of inverted pendulum on cart.

Proof. Due to (12) and (14), for every k, we have

$$\mathcal{H}(A) \leq \lim_{\gamma\to+\infty} R^{(k)}(\gamma) < \mathcal{H}(A) + \frac{n}{2k}\log\frac{4\pi e}{12} + \frac{1}{k}$$

from which (15) follows. Given $\epsilon > 0$, choose sufficiently large k such that

$$\frac{n}{2k}\log\frac{4\pi e}{12} + \frac{1}{k} \leq \frac{\epsilon}{2}.$$

For such k, choose sufficiently large γ such that

$$DI^{(k)}(\gamma) \leq \mathcal{H}(A) + \frac{\epsilon}{2}.$$

Due to (12), the rate of the controller constructed by the procedure presented in this paper with selected k and γ satisfies

$$\bar{R}^{(k)}(\gamma) < DI^{(k)}(\gamma) + \frac{n}{2k}\log\frac{4\pi e}{12} + \frac{1}{k} \leq \mathcal{H}(A) + \epsilon.$$

Since the constructed controller stabilizes the closed-loop system in the mean-square sense, the proof is complete.

In Nair and Evans (2004), a discrete noiseless channel model is considered and the rate is measured by the log-cardinality of the channel alphabet. Theorem 2 shows a slightly different result, as it shows that the same rate is achievable by a noiseless binary channel, even under the prefix-free restriction. However, currently our proof relies on the use of subtractive dither and the optimal Shannon-Fano code (Remarks 1 and 2), and thus the value of Theorem 2 is more theoretical rather than practical.

6. NUMERICAL EXAMPLE

In this section, we consider the optimal block-length k that minimizes the average rate $\bar{R}^{(k)}(\gamma)$ for a given γ by means of numerical simulation. We consider a state space model with

$$A = \begin{bmatrix} 1.000 & 0.0498 & 0.0028 & 0.0001 \\ 0.000 & 0.9913 & 0.1116 & 0.0028 \\ 0.000 & -0.0005 & 1.0327 & 0.0508 \\ 0.000 & -0.0189 & 1.3062 & 1.0327 \end{bmatrix}, B = \begin{bmatrix} .0098 \\ .3908 \\ .0212 \\ .8485 \end{bmatrix}$$

$$F = \begin{bmatrix} .1000 & .0009 & .0001 & .0020 \\ .0009 & .1371 & .0020 & .0805 \\ .0001 & .0020 & .1001 & .0044 \\ .0020 & .0805 & .0044 & .2748 \end{bmatrix}, Q = 2I, R = 1, S = 0.$$

This is a state space model of an inverted pendulum on a cart linearized around its unstable equilibrium (Figure 4).[8] The state space $(x, \dot{x}, \phi, \dot{\phi})$ is four dimensional, where x is the position of the cart and ϕ is the angle of the inverted pendulum. Discrete time model is obtained with sampling frequency 20Hz.

[8] Online resource "Control Tutorials for MATLAB and Simulink (CTMS)" available at `http://ctms.engin.umich.edu/CTMS/index.php?example=InvertedPendulum&section=SystemModeling`.

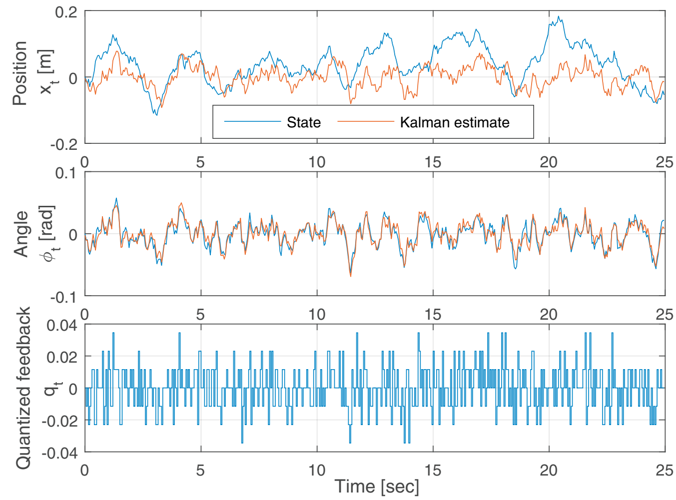

Fig. 5. Simulated state trajectory and quantized feedback signal.

Table 1. Quantized feedback signal $\mathbf{q}_t$, its empirical probability distribution, and Huffman codes.

Source q_t	Empirical Probability	Huffman code
.0460	0.0006	11010100
.0345	0.0106	11011
.0230	0.0748	111
.0115	0.2388	01
0	0.3507	00
-.0115	0.2388	10
-.0230	0.0748	1100
-.0345	0.0106	110100
-.0460	0.0006	1101011

6.1 *Quantized LQG control design*

We first consider a basic design with block-length $k = 1$. Control architecture shown in Figure 3 is constructed by the procedure explained in Section 3. Pseudo-random numbers are used for dither signals, which are shared by encoder and controller (Remark 1). Huffman codes are constructed based on the empirical probability distribution $p(\mathbf{q}_t)$ (Remark 2). Figure 5 shows simulated trajectories of the cart position x_t, angle of the pendulum ϕ_t, and the quantized feedback q_t. Notice that q_t takes discrete values. Table 1 shows codewords assigned to the quantized source q_t.

6.2 *Block-length*

In Figure 6 (left), empirical performances of k-lifted designs are shown for $k = 1, 2, 3, 4$. From this plot, it can be seen that for $\gamma = 15$, the average rate is minimized by block-length $k = 3$. Figure 6 (right) shows theoretical lower and upper bounds of $R^{(k)}(\gamma)$ obtained in (12). It can be seen that better upper bounds are available with larger block-lengths if control performance requirements are not very stringent.

REFERENCES

Åström, K.J. and Bernhardsson, B. (2002). Comparison of riemann and lebesgue sampling for first order stochastic systems. In *Proceedings of the 41st IEEE Conference on Decision and Control*.

Bao, L., Skoglund, M., and Johansson, K.H. (2011). Iterative encoder-controller design for feedback control over noisy channels. *IEEE Transactions on Automatic Control*, 56(2).

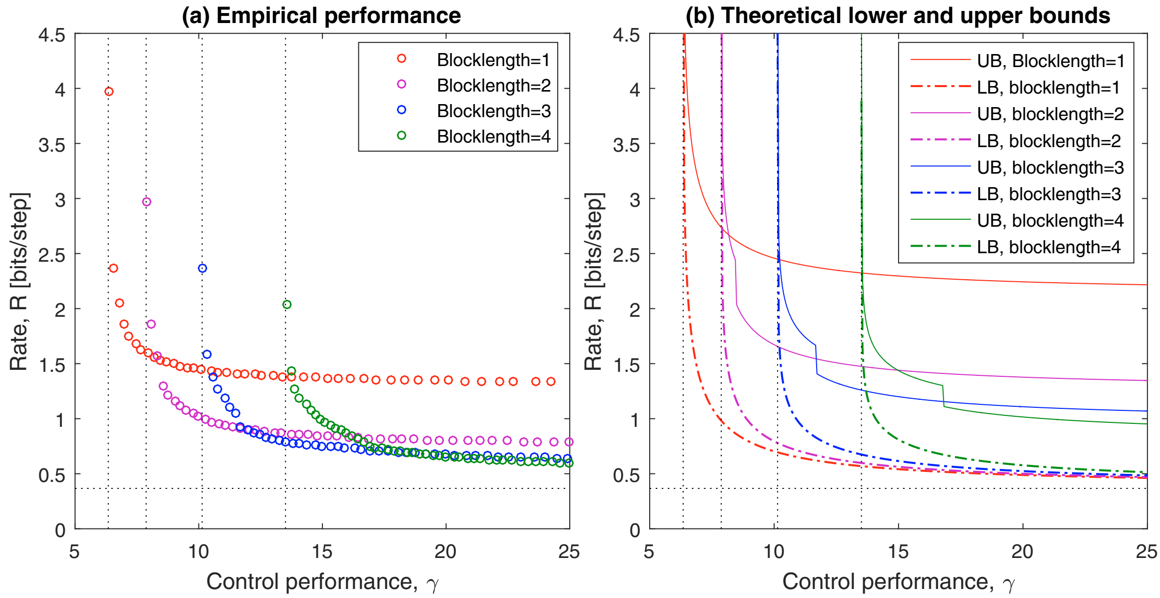

Fig. 6. Left: Average rates and LQG performances of proposed control architecture in Section 3 with various γ and block-lengths k. Rates (average length of Huffman codes) and performances are calculated empirically by simulating a control system for $T = 1,000,000$ time steps. Right: Theoretical lower bounds $DI^{(k)}(\gamma)$ and upper bounds $DI^{(k)}(\gamma)+\frac{r}{2k}\log\frac{4\pi e}{12}+\frac{1}{k}$ obtained in (12). Horizontal asymptote is $\mathcal{H}(A) \triangleq \sum_{\lambda_i\in\sigma(A)} \max\{\log|\lambda_i|, 0\}$. Discontinuity of upper bounds are due to rank-dropping phenomena of $r = \text{rank}(\mathsf{SNR}(\gamma))$.

Bertsekas, D.P. and Shreve, S.E. (1978). *Stochastic optimal control: The discrete time case*, volume 139. Academic Press New York.

Borkar, V.S. and Mitter, S.K. (1997). LQG control with communication constraints. In *Communications, Computation, Control, and Signal Processing*. Springer US.

Cervin, A. and Henningsson, T. (2008). Scheduling of event-triggered controllers on a shared network. In *Proceedings of the 47th IEEE Conference on Decision and Control*.

Cover, T.M. and Thomas, J.A. (1991). *Elements of Information Theory*. Wiley-Interscience, New York, NY, USA.

Freudenberg, J.S., Middleton, R.H., and Braslavsky, J.H. (2011). Minimum variance control over a Gaussian communication channel. *IEEE Transactions on Automatic Control*, 56(8).

Fu, M. (2009). Linear quadratic Gaussian control with quantized feedback. In *Proceedings of the 2009 American Control Conference*.

Gray, R.M. and Stockham, T.G. (1993). Dithered quantizers. *IEEE Transactions on Information Theory*, 39(3).

Huang, M., Nair, G.N., and Evans, R.J. (2005). Finite horizon LQ optimal control and computation with data rate constraints. In *Proceedings of the 44th IEEE Conference on Decision and Control*.

Johansson, K.H., Törngren, M., and Nielsen, L. (2005). Vehicle applications of controller area network. In *Handbook of networked and embedded control systems*, 741–765. Springer.

Lemmon, M.D. and Sun, R. (2006). Performance-rate functions for dynamically quantized feedback systems. In *Proceedings of the 45th IEEE Conference on Decision and Control*.

Massey, J. (1990). Causality, feedback and directed information. In *Proceedings of the 1990 International Symposium on Information Theory and its Applications*.

Matveev, A.S. and Savkin, A.V. (2004). The problem of LQG optimal control via a limited capacity communication channel. *Systems & control letters*, 53(1).

Nair, G.N. and Evans, R.J. (2004). Stabilizability of stochastic linear systems with finite feedback data rates. *SIAM Journal on Control and Optimization*, 43(2).

Silva, E., Derpich, M.S., Ostergaard, J., et al. (2011). A framework for control system design subject to average data-rate constraints. *IEEE Transactions on Automatic Control*, 56(8).

Tallapragada, P. and Cortés, J. (2014). Event-triggered stabilization of linear systems under bounded bit rates.

Tan, X., Xi, W., and Baras, J.S. (2004). Numerical study on joint quantization and control under block-coding. In *Proceedings of the 43rd IEEE Conference on Decision and Control*.

Tanaka, T., Johansson, K.H., Oechtering, T., Sandberg, H., and Skoglund, M. (2016). Rate of prefix-free codes in LQG control systems. In *Proceedings of the 2016 IEEE International Symposium on Information Theory*.

Tanaka, T., Esfahani, P.M., and Mitter, S.K. (2015). LQG control with minimal information: Three-stage separation principle and SDP-based solution synthesis. *arXiv:1510.04214*.

Tatikonda, S., Sahai, A., and Mitter, S. (2004). Stochastic linear control over a communication channel. *IEEE Transactions on Automatic Control*, 49(9).

You, K. and Xie, L. (2011). Linear quadratic gaussian control with quantised innovations kalman filter over a symmetric channel. *Control Theory & Applications, IET*, 5(3), 437–446.

Yüksel, S. (2014). Jointly optimal LQG quantization and control policies for multi-dimensional systems. *IEEE Transactions on Automatic Control*, 59(6).

Yüksel, S. and Başar, T. (2013). *Stochastic networked control systems*, volume 10 of *Systems & Control Foundations & Applications*. Springer, New York, NY.