Quantized stabilization of linear systems: complexity versus performance

Fabio Fagnani
Dipartimento di Matematica
Politecnico di Torino
C.so Duca degli Abruzzi, 24
10129 Torino, Italy
fagnani@calvino.polito.it

Sandro Zampieri
Dipartimento di Elettronica ed Informatica
Università di Padova
via Gradenigo, 6/a
35131 Padova, Italy
zampi@dei.unipd.it

October 13, 2007

Abstract

Quantized feedback control has been receiving much attention in the control community in the past few years. Quantization is indeed a natural way to take into consideration in the control design the complexity constraints of the controller as well as the communication constraints in the information exchange between the controller and the plant. In this paper we analyze the stabilization problem for discrete time linear systems with multidimensional state and one-dimensional input using quantized feedbacks with a memory structure, focusing on the trade off between complexity and performance. A quantized controller with memory is a dynamical system with a state space, a state updating map and an output map.

The quantized controller complexity is modelled by means of three indices. The first index L coincides with the number of the controller states. The second index is the number M of the possible values that the state updating map of the controller can take at each time. The third index is the number N of the possible values that the output map of the controller can take at each time. The index N corresponds also to the number of the possible control values that the controller can choose at each time.

In this paper the performance index is chosen to be the time T needed to shrink the state of the plant from a starting set to a target set. Finally, the contraction rate C, namely the ratio between the volumes of the starting and target sets, is introduced. We evaluate the relations between these parameters for various quantized stabilizers, with and without memory, and we make some comparisons. Then we prove a number of results showing the intrinsic limitations of the quantized control. In particular we show that, in order to obtain a control strategy which yields arbitrarily small values of $T/\ln C$ (requirement which can be interpreted as a weak form of the pole assignability property), we need to have that $LN/\ln C$ is big enough.

Keywords: Stabilization, communication constraints, dynamic quantizers, quantized feedback.

1   Introduction

The stabilization problem by quantized feedback has been widely studied in the last few years: see [1, 2, 4, 5, 7, 13, 16, 18, 19] and the reference therein. Quantization can not be avoided in the digital control setting and it is indeed a natural way to insert into the control design complexity constraints of the controller and communication constraints of the channels which connect the controller and the plant.

In this paper we consider the stabilization problem for general discrete time linear systems with one-dimensional input and full state observation. The quantized feedbacks considered here possess a memory structure. In this paper we try to extend some of the results obtained in [7, 8] under the assumptions that the state is one dimensional and the feedback is memoryless.

The main focus of this paper is on the trade off between controller complexity and the closed loop stability performance. The controller complexity will be described by three integer parameters. The number L of discrete states of the controller will measure the computational complexity of the control algorithm. The number N of quantization subsets of the controller output map and the number M of quantization subsets of the controller state updating map are related to the information flow which is needed for the data transmission between the plant and the controller. The mean time T needed to shrink the state of the plant from a starting set to a target set will instead measure the controller performance.

We can expect a trade-off between the complexity and the performance indices, namely, to obtain small times T we need controllers with high complexity indices L, N, M. In order to quantify this trade-off we introduce another parameter C, called the contraction rate, which coincides with the ratio between the volumes of the starting and target sets.

This framework constitutes a common general setting in which various quantized control strategies already appeared in the literature (including the quantized controller with memory proposed in [2, 18, 14, 15]) can be analyzed and compared on the basis of these indices. We then prove a number of results showing intrinsic limitations of the quantized control. The only hypothesis we need on the plant is that its state matrix A possesses a real eigenvalue. As far as the quantized controller instead we need to impose some geometric characteristics on the quantization subsets of the state updating map. In this way, we can establish inequality constraints between the performance parameter T and the complexity parameters L, N, M, and the contraction C. The geometric hypotheses imposed on the quantization subsets of the state updating map make the dependence on M of these inequalities quite involved. For this reason we prefer to consider it fixed and to use these bounds to study the relations of the parameters T, L, N, M and C only.

In particular we show that, under these assumptions, in order to obtain a controller yielding the value of the ratio $T/\ln C$ arbitrarily small, requirement which can be interpreted as a weak form of the pole assignability property in the classical linear feedback theory, we need to have that the controller complexity is such that $LN/\ln C$ is big enough. On the other hand, the various examples presented show that this is also a sufficient condition: a logarithmic growth of L or of N with respect to C insures arbitrarily small logarithmic time rate. These constraint results are obtained through the use of symbolic dynamics representations of the closed loop maps (which turn out to be piecewise affine maps) and combinatorial results established in [8] for the one-dimensional case. We believe that this kind of analysis is missing in the literature on quantized control. The only complexity parameter considered in the literature is N and the results proposed consists in lower bounds on N, depending only on the system, which guarantee the existence of a quantized asymptotic stabilizer. This can be obtained at a price however of an unlimited number of discrete states, namely, $L = +\infty$. In [8] the above result on logarithmic growth was proven in the special case of one-dimensional state systems and

memoryless quantized feedbacks $L = 1$.

In the remaining part of the introduction we specify the problem, we introduce all assumptions, definitions and notations used in the paper and we provide an interpretation of the control under communication constraints in out context. Sections 2 and 3 are devoted to a careful analysis of some examples of quantized feedback stabilization techniques for which we evaluate the various complexity and performance indices. We first consider a memoryless uniform quantizer like the one considered in [16]. Then we show that, by nesting scaled version of this quantizer, we obtain a quantized feedback map with performance similar to the logarithmic quantizer proposed in [5]. Then we consider dynamic quantized stabilizers illustrating the zooming in/zooming out procedure proposed in [2, 18, 15]. In Section 4 we introduce the symbolic dynamics formalism, we recall some results established in [8] and we prove some inequality constraints. Section 5 contains some concluding remarks.

1.1 Problem statement

Consider a linear discrete time system

$$\begin{cases} x_{t+1} = Ax_t + Bu_t \\ y_t = Gx_t \ , \end{cases} \tag{1}$$

where $x_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^m$ and $y_t \in \mathbb{R}^p$ and where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$ and $G \in \mathbb{R}^{p \times n}$. A controller in our set up is a system

$$\begin{cases} s_{t+1} = f(s_t, y_t) \\ u_t = k(s_t, y_t) \ , \end{cases} \tag{2}$$

where $s_t \in S$ and where $f : S \times \mathbb{R}^p \to S$ and $k : S \times \mathbb{R}^p \to \mathbb{R}^m$. A controller is called quantized if the set S is finite or denumerable and if for each $s \in S$ the map $k(s, \cdot) : \mathbb{R}^p \to \mathbb{R}^m$ is quantized, namely, there exists a finite or denumerable family $\mathcal{K}_s = \{K_s^1, K_s^2, \ldots, K_s^{N_s}\}$ of disjoint subsets of $\mathbb{R}^p$ such that

$$\bigcup_{j=1}^{N_s} K_s^j = \mathbb{R}^p$$

and such that the map $k(s, \cdot)$ is constant on each K_s^j. Notice that, since S is finite or denumerable, then also the map $f(s, \cdot) : \mathbb{R}^p \to S$ will be quantized, since for each $s \in S$ there will exist a finite or denumerable family $\mathcal{F}_s = \{F_s^1, F_s^2, \ldots, F_s^{M_s}\}$ of disjoint subsets of $\mathbb{R}^p$ such that

$$\bigcup_{i=1}^{M_s} F_s^i = \mathbb{R}^p$$

and such that the map $f(s, \cdot)$ is constant on each F_s^i. The subsets K_s^j will be called *input quantization subsets* while subsets F_s^i will be called *state quantization subsets* of the quantized controller. Notice that, for every fixed $s \in S$, the map $x \mapsto (f(s, x), k(s, x))$ is constant on the intersections $F_s^i \cap K_s^j$ and so also this map can be considered quantized with at most $N_s M_s$ quantization subsets. We will assume that the initial condition s_0 of the controller state is fixed and it will be denoted by the symbol 0. Notice that a quantized controller is completely defined by the quadruple $(S, f, k, 0)$.

The interconnection of the system and the controller yields the closed loop system

$$\begin{cases} x_{t+1} = \Gamma(s_t, x_t) \\ s_{t+1} = \Lambda(s_t, x_t) \ , \end{cases} \tag{3}$$

where $\Gamma(s,x) = Ax + Bk(s, Gx)$ and $\Lambda(s,x) = f(s, Gx)$. Observe that for each $s \in S$ the map $\Gamma_s := \Gamma(s, \cdot)$ is a piecewise affine map, namely,

$$\Gamma_s(x) = Ax + Bu_s^j \quad \text{if } x \in G^{-1}(K_s^j).$$

Observe moreover that the dynamical system (3) is a hybrid system since its dynamic behavior is the result of the combination between a system with a continuous state and a system with discrete state. In fact, we can define a map

$$(\Lambda, \Gamma) : S \times \mathbb{R}^n \to S \times \mathbb{R}^n$$

$$(\Lambda, \Gamma)(s, x) = (\Lambda(s,x), \Gamma(s,x)) \ .$$

Since the initial discrete state is fixed, to any $x \in \mathbb{R}^n$ we can associate, through (Λ, Γ), the double orbit $(s_t, x_t) = (\Lambda, \Gamma)^t(0, x)$. We denote by Π_x the canonical projection from $S \times \mathbb{R}^n$ to $\mathbb{R}^n$ so that we have $x_t = \Pi_x(\Lambda, \Gamma)^t(0, x)$.

The problem considered in this paper can be formulated as follows:

Problem: Given three subsets $0 \in V \subseteq W \subseteq Z \subseteq \mathbb{R}^n$, find a quantized controller $(S, f, k, 0)$ such that the closed loop system satisfies the following properties:

1. For any $x \in W$, the evolution $\Pi_x(\Lambda, \Gamma)^t(0, x) \in Z$ for every $t \in \mathbb{N}$.

2. For any $x \in W$, there exists $t_0 \in \mathbb{N}$ such that $\Pi_x(\Lambda, \Gamma)^t(0, x) \in V$ for every $t \geq t_0$.

A controller satisfying the above properties is called (Z, W, V)-stabilizing and the corresponding closed loop (Λ, Γ) is said to be (Z, W, V)-stable.

In the case when $V = W = Z$ we are simply requiring to remain inside V if we start from V. The case $V = W$ has been considered in [19], where it is called containability. The general case, beyond containability, also requires an attraction towards the smallest target subset V. In case when $Z = W$ we will simply talk of (W, V)-stability. In this paper we assume that W and V are convex subsets of finite non-zero n-dimensional Lebesgue measure.

There are some important complexity and performance parameters to be considered in the above problem. On the one hand, the number L of the states of the controller state space S gives a measure of the complexity of the control algorithm and so of its computational demand (of course we consider only that part of S really used by the controller in driving the system from W to V). On the other hand, $N = \sup_s N_s$ and $M = \sup_s M_s$ provide instead a measure of the required information flow between the system and the controller. Of course we would like to have all these integer parameters as small as possible. This will clearly depend on the contraction rate $C = \lambda[W]/\lambda[V]$, where λ is the Lebesgue measure in $\mathbb{R}^n$, which describes how small is the target set with respect to the starting set. It is important however to consider another index which should give an idea of the performance of the transient behavior of the closed loop system in its convergence from W to V. Many choices are possible. Here we will consider the entrance time map $T_{(W,V)} : W \to \mathbb{N} \cup \{+\infty\}$ defined as

$$T_{(W,V)}(x) = \min\{t \in \mathbb{N} \mid \Pi(\Lambda, \Gamma)^{t+n}(0, x) \in V \ \forall n \geq 0\} \ .$$

There are various ways to obtain a performance index from this map. Here we will consider the expected value $T = \mathbb{E}[T_{(W,V)}]$ of this map with respect to the uniform probability density on W. Motivated by the results obtained in [7, 8] in the one-dimensional case, we expect a trade-off between T and the complexity parameters L, M and N as functions of C, illustrating

4

a general feature of the control problem under computation and communication constraints: there is a strict link between the amount of information transmitted and elaborated in the control process and the level of performance that can be reached.

If (Λ, Γ) is (Z, W, V)-stable, then we can introduce two (Λ, Γ)-invariant subsets of $S \times \mathbb{R}^n$ which are naturally linked to the definition of (Z, W, V)-stability. Indeed, consider

$$\bar{W} := \bigcup_{t \in \mathbb{N}} (\Lambda, \Gamma)^t (\{0\} \times W) \supseteq \{0\} \times W \tag{4}$$

$$\bar{V} := \{(\Lambda, \Gamma)^t(0, x) \ : \ (0, x) \in S \times W, \ (\Lambda, \Gamma)^{t+n}(0, x) \in S \times V, \ \forall n \geq 0\} \subseteq S \times V \tag{5}$$

Notice that (Λ, Γ) is (Z, W, V)-stable if and only if $\Pi_x(\bar{W}) \subseteq Z$ and for any $(s, x) \in \bar{W}$, there exists $t_0 \in \mathbb{N}$ such that $(\Lambda, \Gamma)^{t_0}(s_0, x_0) \in \bar{V}$. The previously defined entrance time map can be reformulated more compactly in terms of $\bar{V}$ as follows

$$T_{(W, \bar{V})} : W \to \mathbb{N}$$

$$T_{(W, \bar{V})}(x) = \inf\{t \in \mathbb{N} \mid (\Lambda, \Gamma)^t(0, x) \in \bar{V}\}.$$

We then also have $T := \mathbb{E}[T_{(W, \bar{V})}]$.

In this paper we assume that the input space is one-dimensional $m = 1$ and that we have full state observation $p = n$ and $G = I$. Also we assume (A, B) to be a reachable pair. Under these assumptions it is well known that, if we forget quantization issues, we can find a memoryless linear controller $u = Kx$ such that the closed loop map $x_{t+1} = (A + BK)x_t$ is asymptotically stable. Actually, reachability permits to assign arbitrarily the eigenvalues of $A + BK$. If all eigenvalues are equal to 0 we obtain a dead-beat controller: in this case any initial state is driven to zero in at most n steps. If instead all eigenvalues are located in the open ball $B(0, \delta)$, with $\delta < 1$, we can obtain bounds like $\|x_t\| \leq \text{const } \delta^t$ on the state evolution. Consequently, in this case entrance time T will depend on the contraction C in such a way that $T/\ln C \sim \delta$. In other terms all the values of the ratio $T/\ln C$ can be achieved by allocating the eigenvalues in a suitable way. According to the previous interpretation, if a function $f(C)$ is such that $f(C)/\ln C \to \delta$, then we will call δ the *logarithmic time rate* of f.

A natural question to be posed in the quantized control context is for which values of the complexity parameters L, M and N (as functions of C) we can obtain the same type of stabilization results which could be obtained with linear feedback controllers. In particular, we are interested in evaluating for which values of L, M and N we can obtain the dead-beat controller, or arbitrarily small logarithmic time rate. These parameters will be evaluated for the various examples presented in Sections 2 and 3. General bounds will be established in Section 4.

1.2 Control under communication constraints

In this section we will give an interpretation of the control under communication constraints in the context of the quantized control proposed in this paper. The example below is helpful for understanding this interpretation.

Example Assume we have a number of vehicles moving on the plane according to the dynamic equations

$$\begin{aligned} x_{t+1} &= x_t + u_t \\ y_{t+1} &= y_t + v_t \ , \end{aligned} \tag{6}$$

where x_t and y_t are the two coordinates of one of these vehicles and u_t, v_t are the control inputs. Assume that this vehicle can not measure its position, but that it receives through

a communication channel an approximation of its position from a remotely positioned sensor, such as a camera positioned remotely, for instance on a satellite. Because of the great number of vehicles the sensor has to deal with, a very limited information rate is allowed from the sensor to the vehicle. Choose the following control strategy. Define the quantized map $k : \mathbb{R} \to \mathbb{R}$ as

$$k(z) := \begin{cases} 1/2 & \text{if } -1 < x < 0 \\ -1/2 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

and

$$k(s, x, y) := (\delta^s k(\delta^{-s} x), \delta^s k(\delta^{-s} y)) \qquad \forall s \in \mathbb{Z} \; .$$

Notice that $k(s, x, y)$ is the scaling of $k(0, x, y)$ and that it divides $\mathbb{R}^2$ into 5 quantization subsets. Notice moreover that, if $|x| \leq \delta^s$ and $|y| \leq \delta^s$ and $(x', y') = \Gamma(s, x, y)$, then $|x'| \leq \delta^{s+1}$ and $|y'| \leq \delta^{s+1}$ whenever we choose $\delta \geq 1/2$. We finally introduce the state updating map

$$f(s, x, y) := \begin{cases} s + 1 & \text{if } |x| \leq \delta^s \text{ and } |y| \leq \delta^s \\ s - 1 & \text{otherwise} \end{cases}$$

It is clear that, under the previous assumptions, the control law

$$\begin{cases} s_{t+1} = f(s_t, x_t, y_t) \\ u_t = k(s_t, x_t, y_t) \end{cases}$$

yields the exponential convergence of the closed loop system since the state (x_t, y_t) converges to zero as δ^t. Therefore the best convergence rate is given by $\delta = 1/2$, while $\delta > 1/2$ guarantees some robustness. Notice that in this case $M = 2$ and that the combined map (f, k) has 5 quantization subsets. This is an example of the zooming strategy proposed in [2] and illustrated in Section 3.

Notice that in this example the sensor has to send to each vehicle only one of the five possible scaled quantization subsets the vehicle belongs to. The vehicle needs to know the scaling factor s and to this aim it needs to know only the updating map and the initial scaling s_0. This works only in case we assume that there is no transmission error. In fact, a difference between the sensor scaling factor and the vehicle scaling factor will cause instability. So it is necessary an absolute reliability when the sensor informs the vehicle about its being inside the set $[-\delta^{s_t}, \delta^{s_t}]^2$ in order to maintain the synchronization between the encoder and the decoder state. The transmission of both the state and of the quantization subset prevents this problem. However, it requires the same transmission rate needed by memoryless quantized feedback strategies.

Notice moreover that, from a more technological point of view, the communication complexity indices M and N appear to be much more critical design parameters than the index L describing the memory requirement of the control algorithm. It is our opinion that, under noisy communication, the presence of many states can make the state synchronization problem more critical and so we expect that L should be more relevant in this case.

The idea illustrated in the previous example can be generalized in the context of general quantized controllers as follows. Define the map

$$q : S \times \mathbb{R}^n \to \mathbb{N} \times \mathbb{N}$$

such that $q(s, x) = (i, j)$ if and only if $x \in F_s^i \cap K_s^j$. Notice that there exist maps $\bar{f} : \mathbb{N} \times \mathbb{N} \to S$ and $\bar{k} : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ such that $f = \bar{f} \circ q$ and $k = \bar{k} \circ q$. Control under communication constraints

can be formalized as illustrated in the following picture by letting the sensor/encoder to be

$$\text{ENCODER} \left\{ \begin{array}{l} s_{t+1} = f(s_t, x_t) \\ (i_t, j_t) = q(s_t, x_t) \end{array} \right.$$

by letting the controller/decoder to be

$$\text{DECODER} \left\{ \begin{array}{l} s_{t+1} = \bar{f}(s_t, i_t, j_t) \\ u_t = \bar{k}(s_t, i_t, j_t) \end{array} \right.$$
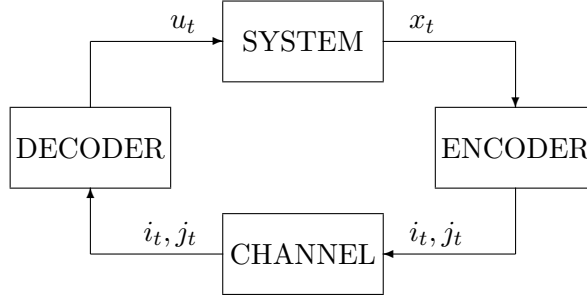
2 Memoryless feedback quantizers

In this section we will consider the class of quantized controllers which have no memory, namely, $S = \{0\}$, $L = 1$. These controllers, called memoryless quantized feedback, are determined by a single quantized map

$$u_t = k(x_t).$$

2.1 Uniform quantized feedback

The simplest way to obtain quantized feedback maps is by quantizing linear feedback maps uniformly or logarithmically. More precisely, define a uniform quantizer to be any quantized map $q : \mathbb{R} \to \mathbb{R}$ such that $|z - q(z)| \leq 1/2$. We can take for instance

$$q(z) = k + 1/2 \quad \text{for all } z \text{ such that } k < z < k + 1. \tag{7}$$

Define moreover the scaled version of $q(z)$ as

$$q_\Delta(z) := \Delta q\left(\frac{z}{\Delta}\right).$$

It is easy to see that this quantizer has quantization intervals of length Δ and that $|z - q_\Delta(z)| \leq \Delta/2$. We know that there exists a linear state feedback

$$u_t = kx_t$$

yielding a closed loop system which control to zero any initial state in n steps. The first strategy is simply to choose the quantized feedback

$$k(x) = q_\Delta(kx).$$

Since the system is reachable, we can assume with no loss of generality that (A, B) is in the controller canonical form, namely

$$
A = \begin{pmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 1 \\
-a_0 & -a_1 & \cdots & -a_{n-2} & -a_{n1}
\end{pmatrix}
\qquad
B = \begin{pmatrix}
0 \\ 0 \\ \vdots \\ 0 \\ 1
\end{pmatrix}.
$$

In this case the dead-beat controller is given by the feedback matrix

$$
k = \begin{bmatrix} a_0 & a_1 & \cdots & a_{n-1} \end{bmatrix}.
$$

If $x = (x_1, \ldots, x_{n-1}, x_n) \in Q_\Delta := [-\Delta/2, \Delta/2]^n \subseteq \mathbb{R}^n$ and if $x' = (x'_1, \ldots, x'_{n-1}, x'_n) = Ax + Bq_\Delta(kx)$, then it is clear that

$$
x'_1, \ldots, x'_{n-1} \in \left[-\frac{\Delta}{2}, \frac{\Delta}{2} \right]
$$

Moreover, notice that

$$
|x'_n| = \left| -\sum_{i=1}^n a_{i-1} x_i + q_\Delta \left(\sum_{i=1}^n a_{i-1} x_i \right) \right| \leq \frac{\Delta}{2}
$$

and hence we have that $x' \in Q_\Delta$. In this way we showed that the hypercube Q_Δ is invariant with respect to the closed loop map $\Gamma(x) := Ax + Bk(x)$.

Notice that this quantized feedback yields a quantization partition inside Q_Δ which results by cutting Q_Δ by equidistant hyperplanes which are orthogonal to the vector $(a_0, \ldots, a_{n-1}) \in \mathbb{R}^n$. Notice moreover that, if $x \in Q_\Delta$, then $|kx| \leq a\Delta/2$, where

$$
a := \sum_{i=0}^{n-1} |a_i|. \tag{8}
$$

This implies that only $2 \left\lceil \frac{a}{2} \right\rceil$ levels of the quantized map $k(x) = q_\Delta(kx)$ are active and so this quantized feedback requires

$$
N = 2 \left\lceil \frac{a}{2} \right\rceil
$$

quantization subsets inside Q_Δ. With a little more effort we can obtain the same goal with $N = \lceil a \rceil$ quantization subsets.

The technique just presented allows to obtain a quantized feedback making a hypercube Q_Δ invariant. A simple modification of the above construction leads to a class of (Q_Δ, Q_ϵ)-stabilizing feedbacks, as shown in the following proposition (see [16]).

Proposition 1 *[16] For every $\epsilon < \Delta$, there exists a (Q_Δ, Q_ϵ)-stabilizing feedback with*

$$
N = \left\lceil a \frac{\Delta}{\epsilon} \right\rceil
$$

quantization subsets in Q_Δ and entrance time $T_{(Q_\Delta, Q_\epsilon)}$ such that

$$
\mathbb{E}[T_{(Q_\Delta, Q_\epsilon)}] = n - \left(\frac{\epsilon}{\Delta} \right)^n \frac{(\Delta/\epsilon)^n - 1}{\Delta/\epsilon - 1}
$$

Proof Consider the quantized feedback

$$k(x) = q_\epsilon(kx).$$

The same arguments presented above ensures that in at most n steps this controller drives the any state in $\mathbb{R}^n$ into Q_ϵ and moreover we have that inside Q_Δ this feedback has $\lceil a\Delta/\epsilon \rceil$ quantization subsets. The only thing that needs to be proved is the expected entrance time. Notice that

$$\mathbb{E}[T_{(Q_\Delta,Q_\epsilon)}] = \sum_{k=0}^{\infty} \mathbb{P}[T_{(Q_\Delta,Q_\epsilon)} \geq k+1] = \sum_{k=0}^{n-1} \mathbb{P}[T_{(Q_\Delta,Q_\epsilon)} \geq k+1] = \sum_{k=0}^{n-1}(1 - \mathbb{P}[T_{(Q_\Delta,Q_\epsilon)} \leq k])$$

Since

$$\{x \in Q_\Delta : T_{(Q_\Delta,Q_\epsilon)}(x) \leq k\} = \left[-\frac{\Delta}{2}, \frac{\Delta}{2}\right]^k \times \left[-\frac{\epsilon}{2}, \frac{\epsilon}{2}\right]^{n-k},$$

then

$$\mathbb{E}[T_{(Q_\Delta,Q_\epsilon)}] = n - \Delta^{-n} \sum_{k=0}^{n-1} \Delta^k \epsilon^{n-k}$$

which yields the result.

∎

Notice that the contraction rate C in the above case is $C = (\Delta/\epsilon)^n$. Hence we have, for $C \to +\infty$

$$
\begin{aligned}
N &= \left\lceil aC^{1/n} \right\rceil \sim aC^{1/n} \\
T &= \mathbb{E}[T_{(Q_\Delta,Q_\epsilon)}] = n - C^{-1/n}\frac{1 - C^{-1}}{1 - C^{-1/n}} \sim n
\end{aligned}
$$

where the last approximation, which holds when C is big enough, shows that in this case the expected entrance time tends to be independent of the contraction C. Notice that these estimates are in agreement with the results given in [7, 8].

If instead of considering uniform quantized approximations of linear state feedback, we consider logarithmic quantized approximations we obtain what has been proposed in [5]. A logarithmic quantizer is any quantized map $q : \mathbb{R} \to \mathbb{R}$ such that $|z - q(z)| \leq \delta|z|$ where $\delta > 0$. This technique has the advantage to be less demanding in terms of the number quantization subsets. This is paid by a bigger expected entrance time, which grows at least logarithmically as a function of the contraction rate. We prefer not to investigate in detail the properties and the performance of this technique. Instead, we will present an alternative method which provides similar results but which is simpler to be analyzed.

2.2 Nested quantizers

Different stabilization performance can be obtained by nesting quantized feedbacks. This method is inspired by the zooming techniques proposed in [2]. Assume we have three subsets $W_1 \supseteq W_2 \supseteq W_3$ and two feedback quantizers: k_1 which is (W_1, W_2)-stabilizing and k_2 which is (W_2, W_3)-stabilizing. Define the nested quantized feedback $k(x)$ defined by

$$k(x) = \begin{cases} k_1(x) & x \in W_1 \setminus W_2 \\ k_2(x) & x \in W_2 \end{cases}$$

9

where $W_1 \setminus W_2 := \{x \in W_1, \ x \notin W_2\}$. It is clear that $k(x)$ is (W_1, W_3)-stabilizing. Moreover, if N_i is the number of quantization subsets of k_i inside W_i, we have that the number N of quantization subsets used by k in W_1 is bounded by $N_1 + N_2$. The evaluation of the expected entrance time is more difficult and we will return to this question shortly.

This construction can be repeated as many times as we want. Let us see what we obtain by considering a nesting of the stabilizing feedbacks constructed in the previous section. Assume again for simplicity that the system (1) is already in the controller canonical form. Fix $\Delta > 0$ and $0 < \delta < 1$ and consider a quantized feedback $\bar{k}$ such that

$$\bar{\Gamma}^n(Q_\Delta) \subseteq Q_{\delta\Delta}$$

where $\bar{\Gamma}(x) := Ax + B\bar{k}(x)$. Notice that this quantized feedback requires $\lceil a/\delta \rceil$ quantization subsets inside Q_Δ, where a is defined in (8). Consider the scaled quantized feedback $k_i(x) := \delta^i \bar{k}(\delta^{-i}x)$. It is easy to verify that the corresponding closed loop map, which is $\Gamma_i(x) := \delta^i \bar{\Gamma}(\delta^{-i}x)$, is such that

$$\Gamma_i^n(Q_{\delta^i\Delta}) \subseteq Q_{\delta^{i+1}\Delta}.$$

The nested quantized feedback defined as

$$k(x) := k_i(x) \quad \text{if} \quad x \in Q_{\delta^i\Delta} \setminus Q_{\delta^{i+1}\Delta} \tag{9}$$

will be $(Q_\Delta, Q_{\delta^r\Delta})$-stabilizing with contraction rate δ^{-rn} and $r\lceil a/\delta \rceil$ quantization subsets. As far as the mean entrance time $\mathbb{E}[T_{(Q_\Delta, Q_{\delta^r\Delta})}]$ is concerned, we clearly have that $\mathbb{E}[T_{(Q_\Delta, Q_{\delta^r\Delta})}] \leq rn$, but we would like to have better estimates.

Consider the map

$$\Psi : Q_\Delta \to Q_\Delta : x \mapsto \delta^{-1}\bar{\Gamma}^{n(x)}(x),$$

where $n(x) := T_{(Q_\Delta, Q_{\delta\Delta})}(x) \leq n$ is the first entrance time function for $\bar{k}$. In this way we obtain that the first entrance time for k is

$$T_{(Q_\Delta, Q_{\delta^r\Delta})}(x) = \sum_{i=0}^{r-1} T_{(Q_\Delta, Q_{\delta\Delta})}(\Psi^i(x)) \ .$$

The map Ψ is a piecewise affine map and it induces an operator

$$\mathcal{P} : L^1(Q_\Delta) \to L^1(Q_\Delta)$$

mapping probability densities on Q_Δ into themselves. This is called the Perron-Frobenius operator associated with Ψ and it satisfies the following duality relation [9]

$$\int_{Q_\Delta} (g \circ \Psi)(x)f(x)dx = \int_{Q_\Delta} g(x)(\mathcal{P}f)(x)dx \tag{10}$$

for all $g \in L^\infty(Q_\Delta), f \in L^1(Q_\Delta)$. Consequently, we can write

$$T_r := \mathbb{E}[T_{(Q_\Delta, Q_{\delta^r\Delta})}] = \sum_{k=0}^{r-1} \mathbb{E}_{g_k}[T_{(Q_\Delta, Q_{\delta\Delta})}]$$

where $g_k = \mathcal{P}^k g$ and g is the uniform probability density in Q_Δ and where the symbol $\mathbb{E}_{g_k}$ means that the expected value of with respect to the density g_k. We have the following result.

Theorem 1 *Assume that the initial state is uniformly distributed on Q_Δ. Then there exists a $\mathcal{P}$-invariant probability density $\bar{g}$ such that*

$$T_r = r\bar{T} + a_r \qquad \forall r \in \mathbb{N}$$

where $\{a_r\}$ is a bounded sequence and where

$$\bar{T} := \mathbb{E}_{\bar{g}}[T_{(Q_\Delta, Q_{\delta\Delta})}]$$

is the expected entrance time from Q_Δ to $Q_{\delta\Delta}$ with respect to the probability density $\bar{g}$.

We do not furnish here a proof of this result but we make a few comments. In the one-dimensional case $n = 1$ it is proven in [8]. The proof is based on the well-known spectral properties of the Perron Frobenius operator of one-dimensional piecewise affine maps on bounded variation densities. The multi-dimensional case is mathematically more delicate but can be treated in a similar way using the results in [3] and [17] which allows to have a similar spectral theory working on the quasi-Holder densities. Details will be discussed elsewhere.

Let us go back to the nesting. Notice that $C = \delta^{-nr}$. From this it follows that, if N_r denotes the number of quantization subsets and T_r the expected entrance time, then

$$N_r \sim ra/\delta = arC^{\frac{1}{rn}} \qquad T_r \sim \bar{T}r$$

where $\bar{T}$ is defined in Theorem 1. Notice that $\bar{T} \le n$. The number r of nestings can be considered as a supplementary parameter which can be varied with C, as C goes to ∞. In fact, if we keep r fixed, we obtain that the number of quantization subsets grows slower with respect to the strategy presented in Proposition 1, while the expected entrance time is still independent of C, but it is bigger. Another possibility is to vary r with C keeping the parameter δ fixed. In this way r grows logarithmically in C as

$$r = \frac{1}{n} \frac{\ln C}{\ln(\delta^{-1})}$$

which implies that

$$N_r \sim \frac{a}{\delta n} \frac{\ln C}{\ln(\delta^{-1})} \qquad T_r \sim \frac{\bar{T}}{n} \frac{\ln C}{\ln(\delta^{-1})} \tag{11}$$

The growth of the number of quantization subsets needed here is sharply smaller with respect to the previous cases. This is paid in terms of performance. Indeed, while in the previous case we obtained an expected entrance time which was independent of C, here we have that the expected entrance time grows logarithmically with C. This is an alternative way to obtain a logarithmic quantized feedback with respect to what is proposed in [5].

Remark: The nested quantizer, by the way it has been defined, will have the quantization subsets obtained by cutting each single subset $Q_{\delta^i\Delta} \setminus Q_{\delta^{i+1}\Delta}$ with a bunch of parallel hyperplanes. In this way, in general, the final quantization of Q_Δ will not be obtained by cutting Q_Δ with parallel hyperplanes. However there is an easy way to obtain such a quantization with the same complexity and performance parameters. First observe that

$$k_i(x) = \Delta\delta^{i+1} \; q\left(\frac{kx}{\Delta\delta^{i+1}}\right) \;,$$

where $k \in \mathbb{R}^{1\times n}$ is the linear dead-beat controller. This implies that the nested quantized feedback defined as

$$\bar{k}(x) := k_i(x) \quad \text{if} \quad kx \in I_{\delta^i\Delta} \setminus I_{\delta^{i+1}\Delta} \;,$$

where $I_{\delta^i\Delta} := \{kx \in \mathbb{R} : x \in Q_{\delta^i\Delta}\}$, has the same properties of the nested quantized feedback defined in (9) and moreover its quantization subsets is obtained by cutting Q_Δ with hyperplanes which are orthogonal to k.

3 Quantized controllers with memory

In the literature various time-varying quantized feedbacks have been proposed. They are all based on the same intuitive idea which both the logarithmic quantized feedback strategy and the strategy based on nesting is based on. Namely, the most convenient quantized feedback should be accurate only when the state is closed to the target, while it may be quite imprecise when the state is far from the objective. All these time-varying strategies can be seen in the unified framework proposed in [2] called zooming. The idea is quite simple. Use the same feedback but scaled in such a way that it is rough when the state is big and it become fine when the state approaches the origin. Here we recall this method putting it in our framework. Then we will analyze the properties and performance of this method. We will present in detail two techniques which are two particular cases of the general scheme presented at the end of the section.

3.1 One step zooming

In this subsection we will present the simplest possible version of the zooming method. Assume that the controller state space $S = \mathbb{Z}$ and that we are given a family of open subsets $L_s \subseteq \mathbb{R}^n$, $s \in \mathbb{Z}$, such that

(i) $L_s \subseteq L_{s-1}$ for all $s \in \mathbb{Z}$;

(ii) $\bigcup_{s \in \mathbb{Z}} L_s = \mathbb{R}^n$;

(iii) $\bigcap_{s \in \mathbb{Z}} L_s = \{0\}$.

Assume moreover that we have found a map $k : S \times \mathbb{R}^n \to \mathbb{R}$ such that the closed loop map $\Gamma_s(x) = \Gamma(s, x) = Ax + Bk(s, x)$ is such that

$$\Gamma_s(L_s) \subseteq L_{s+1}.$$

The controller is defined as

$$s_{t+1} = \begin{cases} s_t + 1 & \text{if } \Gamma(s_t, x_t) \in L_{s_t+1} \\ s_t - \alpha & \text{otherwise} \end{cases}$$

$$u_t = k(s_t, x_t)$$

(12)

This strategy ensures the convergence to zero for any initial state $x_0 \in \mathbb{R}^n$ and $s_0 \in S$, if the integer α is chosen big enough compared to the degree of instability of the system (1). Notice that, in our setting, we can more precisely say that the controller (S, f, k, s_0) is (L_{s_0}, L_{s_0+r})-stabilizing for every $r \in \mathbb{N}$. Or also that, for every $s_1 \in \mathbb{Z}$, there exists $r_1 \in \mathbb{N}$ such that (S, f, k, s_0) is $(L_{s_1}, L_{s_1-r_1}, L_{s_1+r})$-stabilizing for every $r \in \mathbb{N}$.

This is the general description of the zooming method. Now we give more concrete example of this technique by presenting a specific choice of the family of subsets $L_s \subseteq \mathbb{R}^n$, $s \in \mathbb{Z}$ and of the map $k : S \times \mathbb{R}^n \to \mathbb{R}$ satisfying the previous conditions. Assume again that, with no loss of generality, (A, B) is in the controller canonical form. In the previous section we have shown how to construct a quantized feedback map $\bar{k} : \mathbb{R}^n \to \mathbb{R}$ such that the closed loop map $\bar{\Gamma}(x) := Ax + B\bar{k}(x)$ is such that

$$\bar{\Gamma}^n(Q_\Delta) \subseteq Q_{\delta\Delta}$$

where $\Delta > 0$, $\delta < 1$. Notice that this quantized feedback required $N = \lceil a/\delta \rceil$ quantization subsets inside Q_Δ, where a is defined in (8). We modify $\bar{k}$ by deleting all the quantization

hyperplanes of $\bar{k}$ not intersecting Q_Δ. There is only one quantized map which can be obtained in this way and which is equal to the old $\bar{k}$ in Q_Δ. Notice that the quantized map obtained in this way has still $N = \lceil a/\delta \rceil$ quantization subsets.

Exactly as we did in the previous section, we can obtain a scaled version $k_i(x) := \delta^i \bar{k}(\delta^{-i} x)$ of the quantized feedback $\bar{k}(x)$. The corresponding closed loop map $\Gamma_i(x) := \delta^i \bar{\Gamma}(\delta^{-i} x)$ contracts the states in $Q_{\delta^i \Delta}$ into $Q_{\delta^{i+1} \Delta}$ in n steps. Define now the subsets $L_s \subseteq \mathbb{R}^n$ the quantized maps $k(s, \cdot)$, $s \in \mathbb{Z}$ as follows

$$\begin{aligned} L_{in+j} &:= \Gamma_i^j(Q_{\delta^i \Delta}) \\ k(in+j, x) &:= k_i(x) \end{aligned} \qquad \forall i \in \mathbb{Z}, \ j = 0, 1, \ldots, n-1$$

It is easy to verify that this is a zooming strategy satisfying the above conditions and that it yields (L_0, L_{nr})-stability for every $r \in \mathbb{N}$.

We want to evaluate now the entrance time of this strategy. Fix $r \geq 0$ and assume that $s_0 = 0$ and that $x_0 \in L_0$. We recall that

$$T_{(L_0, L_{nr})}(x_0) = T_{(L_0, \bar{L}_{nr})}(x_0) = \min\{t_0 \in \mathbb{N} \mid (s_t, x_t) \in \bar{L}_{nr}, \ \forall t \geq t_0\} \ ,$$

where $\bar{L}_s$ denotes $[s, \infty) \times L_s \subseteq S \times \mathbb{R}^n$. It is clear that in this case we have simply $T_{(L_0, \bar{L}_{nr})}(x_0) = nr$ and so the evaluation of the expected time is trivial. In conclusion this (L_0, L_{nr})-stabilization strategy exhibits the following parameters

1. The contraction is $C = \delta^{-nr}$

2. Number of states is $L = nr$.

3. Number of quantization subsets of $f(s, x)$ is $M = 2$.

4. Number of quantization subsets of $k(s, x)$ is $N = \lceil a/\delta \rceil$.

5. Expected entrance time $T = nr$.

In fact, the combined map (f, k) has only $N + 1$ quantization subsets.

3.2 Two steps zooming

In this subsection we will present another version of the zooming method which yields different performance. Assume in this case that we have found a map $k : S \times \mathbb{R}^n \to \mathbb{R}$ such that the closed loop map $\Gamma_s(x) = \Gamma(s, x) = Ax + Bk(s, x)$ is such that

a) $\Gamma_s(L_s) \subseteq L_s$ and $\Gamma_s(L_{s+1}) \subseteq L_{s+1}$.

b) For any $x \in L_s$ there exists $t \in \mathbb{N}$ such that $\Gamma_s^t(x) \in L_{s+1}$.

These two conditions are equivalent to the requirement that $k(s, x)$ is (L_s, L_{s+1})-stabilizing. The controller is defined as

$$s_{t+1} = \begin{cases} s_t + 1 & \text{if } \Gamma(s_t, x_t) \in L_{s_t + 1} \\ s_t & \text{if } \Gamma(s_t, x_t) \in L_{s_t} \setminus L_{s_t + 1} \\ s_t - \alpha & \text{otherwise} \end{cases} \tag{13}$$

$$u_t = k(s_t, x_t)$$

Also this strategy ensures the convergence to zero for any initial state $x_0 \in \mathbb{R}^n$ and $s_0 \in S$, if the integer α is chosen big enough. Interpretation in terms of our stability concepts can be done exactly as in the case of one-step zooming.

Also in this case we analyze this method for a specific choice of $L_s \subseteq \mathbb{R}^n$ and $k : S \times \mathbb{R}^n \to \mathbb{R}$. Take again the same quantized feedback map $\bar{k} : \mathbb{R}^n \to \mathbb{R}$ considered above and its scaled versions $k_i(x)$. Define $L_s \subseteq \mathbb{R}^n$ and $k(s, \cdot)$, $s \in \mathbb{Z}$ as follows

$$
\begin{aligned}
L_s &:= Q_{\delta^s \Delta} \\
k(s, x) &:= k_s(x)
\end{aligned} \qquad \forall s \in \mathbb{Z}
$$

The evaluation of the expected entrance time for this strategy is less obvious. Fix $r \geq 0$ and assume that $s_0 = 0$ and that $x_0 \in L_0$. We want to evaluate $\mathbb{E}[T_{(L_0, \bar{L}_r)}]$ assuming that x_0 is distributed in $L_0 = Q_\Delta$ according to the uniform probability density g.

Consider the map

$$
\Phi : Q_\Delta \to Q_\Delta : x \mapsto \delta^{-1} \bar{\Gamma}^{n(x)}(x),
$$

where $n(x) := T_{(L_0, \bar{L}_1)}(x)$ is the one step entrance time function. Notice that in this case we have that $1 \leq T_{(L_0, \bar{L}_1)}(x) \leq n$. In this way we obtain that

$$
T_{(L_0, \bar{L}_r)}(x) = \sum_{i=0}^{r-1} T_{(L_0, \bar{L}_1)}(\Phi^i(x)) \ .
$$

The map Φ is a piecewise affine map. Let

$$
\mathcal{Q} : L^1(Q_\Delta) \to L^1(Q_\Delta)
$$

be the associated Perron-Frobenius operator (see (10)). We can write

$$
T_r := \mathbb{E}[T_{(L_0, \bar{L}_r)}] = \sum_{k=0}^{r-1} \mathbb{E}_{g_k}[T_{(L_0, \bar{L}_1)}]
$$

where $g_k = \mathcal{Q}^k g$ and g is the uniform probability density in Q_Δ. We have the following result which can be proved as Theorem 1.

Theorem 2 *Assume that $s_0 = 0$ and x_0 is uniformly distributed on Q_Δ. Then there exists a $\mathcal{Q}$-invariant probability density $\widetilde{g}$ such that*

$$
T_r = r\widetilde{T} + a_r \qquad \forall r \in \mathbb{N}
$$

where $\{a_r\}$ is a bounded sequence and where

$$
\widetilde{T} := \mathbb{E}_{\widetilde{g}}[T_{(L_0, \bar{L}_1)}]
$$

is the expected entrance time from L_0 to $\bar{L}_1$ with respect to the probability density $\widetilde{g}$.

Notice that this strategy appears to be more efficient with respect to the 1-step zooming also in terms of the number of the quantized controller states. Indeed, in the 2-steps zooming strategy each controller state transition corresponds to a contraction of δ^n while to obtain the same in the 1-step zooming case it is necessary to have n state transitions. In other words, the set L_r in the 2-steps zooming coincides with the set L_{nr} in the 1-step zooming and so to reach we need only r state transitions in the 2-steps zooming, and nr state transitions in the 1-step zooming.

In conclusion in this case we have the parameters

1. The contraction is $C = \delta^{-nr}$

2. Number of states is $L = r$.

3. Number of quantization subsets of $f(s, x)$ is $M = 3$.

4. Number of quantization subsets of $k(s, x)$ is $N = \lceil a/\delta \rceil$.

5. Expected entrance time $T \sim \widetilde{T}r$.

In fact, the combined map (f, k) has only $2N + 1$ quantization subsets.

In the following table the performance and the complexity of the nesting technique, the one the step zooming technique and the two steps zooming technique are compared.

		nesting	zooming1	zooming2
contraction	C	δ^{-nr}	δ^{-nr}	δ^{-nr}
expected time	T	$\bar{T}r$	nr	$\widetilde{T}r$
states	L	1	nr	r
quantization of k	N	$r\lceil a/\delta \rceil$	$\lceil a/\delta \rceil$	$\lceil a/\delta \rceil$
quantization of f	M	1	2	3

$$(14)$$

Observe that, if we consider the index L not critical in the quantized controller design, then the zooming strategies clearly overperform the nesting strategies. This is true if we assume noiseless communication. As we pointed out above, in case of noisy communication, the problem of maintaining the synchronization between the encoder and the decoder state makes the design of the quantized controller with memory much more complex.

Remark The zooming control strategy presented above can be generalized by modifying the state updating map given in (12) and in (13). Indeed, we can take more in general an updating map like

$$
s_{t+1} = \begin{cases}
s_t + \beta & \text{if } \Gamma(s_t, x_t) \in L_{s_t+\beta} \\
\quad \vdots \\
s_t + 2 & \text{if } \Gamma(s_t, x_t) \in L_{s_t+2} \setminus L_{s_t+3} \\
s_t + 1 & \text{if } \Gamma(s_t, x_t) \in L_{s_t+1} \setminus L_{s_t+2} \\
s_t & \text{if } \Gamma(s_t, x_t) \in L_{s_t} \setminus L_{s_t+1} \\
s_t - 1 & \text{if } \Gamma(s_t, x_t) \in L_{s_t-1} \setminus L_{s_t} \\
s_t - 2 & \text{if } \Gamma(s_t, x_t) \in L_{s_t-2} \setminus L_{s_t-1} \\
\quad \vdots \\
s_t - \gamma & \text{if } \Gamma(s_t, x_t) \in L_{s_t-\gamma} \setminus L_{s_t-\gamma+1} \\
s_t - \alpha & \text{otherwise}
\end{cases}
$$

where α, β, γ are integers such that $\beta > 0$ and $\alpha > \gamma \geq 0$. The zooming and the nesting can be seen as extremes of this general quantized control methodology.

4 General performance bounds

In this section we study some general performance limitations of quantized control schemes, showing in particular that the controllers proposed in the previous section can not be improved much. To obtain these results we have to make some further assumptions on the quantized controllers we consider. These assumptions are verified by all examples treated so far.

Assume that the map (Λ, Γ) is (Z, W, V)-stable with Z bounded. In the examples treated in Section 3 the set S was infinite. However, in this section S denotes only the subset of states really used by the controller in shrinking W into V and this is assumed to be finite. The number of its element is denoted by L. We also assume that N_s and M_s are finite numbers for any $s \in S$. Moreover we assume that the map $k(s, x)$ has the form

$$k(s, x) = q_s(k_s x),$$

where $q_s : \mathbb{R} \to \mathbb{R}$ is a scalar quantized map and $k_s \in \mathbb{R}^{1 \times n}$. This hypotheses is equivalent to the fact that the quantization subsets K_s^j are obtained by cutting $\mathbb{R}^n$ with parallel hyperplanes orthogonal to a given vector. For what concerns the partitions $\mathcal{F}_s$, we introduce the integer ν_s which is the maximum number of intervals in which the straight line $\rho \mathbb{R}$ is divided by the partition $\mathcal{F}_s$ as ρ varies in $\mathbb{R}^n$ and we define $\nu = \max_s \nu_s$. In the results we are going to establish we assume that ν is a finite fixed parameter. Notice that in the example on 1 step zooming we gave in the previous section, ν_s is periodically varying in s, while in the example on 2 steps zooming ν_s is constant.

We then make a geometric assumption on the sets V and W which roughly amounts to require that both V and W are not 'too skinny' in some dimension. Formally, for a fixed $\gamma > 0$, we assume that

$$\frac{\lambda(W)}{d(W)^n} \geq \gamma, \ \frac{\lambda(V)}{d(V)^n} \geq \gamma,$$

where $\lambda(\cdot)$ is the Lebesgue measure and $d(\cdot)$ is the diameter of a set in $\mathbb{R}^n$.

We also assume A to have a real non-zero eigenvalue μ. This may be a serious limitation, but it is mainly made for the sake of simplicity. The general case leads to problems similar to considering quantization in more than one direction and will be treated elsewhere.

To resume, we have introduced two new parameters: ν which measures the complexity of the family of partitions $\mathcal{F}_s$ and γ which is connected with the shape of V and W. Moreover we have also considered a real eigenvalue μ of the matrix A. These three parameters will be kept fixed in the sequel, as well M which, on the other hand, is strictly linked to ν.

In the rest of this section we are going to establish some trade-off inequality constraints involving the performance parameters C and $T = \mathbb{E}[T_{(W, \bar{V})}]$, where $\bar{V}$ was defined (5), and the two complexity parameters L and N. Actually, it is possible and useful to condense the role of L and N into just one complexity parameter $\mathbf{N}$ defined as the number of quantization subsets $F_s^i \cap K_s^j$ effectively visited by the controller in transient evolution from W into V. More precisely, $\mathbf{N}$ is the number of quantization subsets $F_s^i \cap K_s^j$ for which

$$\{s\} \times (F_s^i \cap K_s^j) \not\subseteq \bar{V}$$

and for which there exists $x \in W$ and $t \in \mathbb{N}$ such that

$$(\Lambda, \Gamma)^t(0, x) \in \{s\} \times (F_s^i \cap K_s^j).$$

The inequalities we will prove will be in terms of T, C and $\mathbf{N}$, while the constants appearing will only depend on μ, ν and γ. Notice that

$$\mathbf{N} \leq \sum_{s \in S} N_s M_s \leq NML. \tag{15}$$

This implies that, if we keep also M fixed, all the inequalities in terms of T, C, $\mathbf{N}$ we will obtain can be translated into inequalities in terms of T, C, N and L.

Our strategy in obtaining the trade-off inequalities will be the following. We will first use a one-dimensional reduction technique and some ideas in [8] to lower bound the expected

entrance time T in terms of the number of the words of a language emerging from the symbolic description of the closed loop dynamics. We will then use a general result in [8] to lower bound this number of words in terms of $\mathbf{N}$. Finally we will combine the various pieces together and obtain the inequalities.

4.1 A one-dimensional reduction

With no loss of generality we can assume that the eigenvector corresponding to the eigenvalue μ is e_1, the first element of the canonical basis of $\mathbb{R}^n$. Given any $W \subseteq \mathbb{R}^n$, we define W_2 to be the projection of W on the last $n-1$ components. For any fixed $x_2 \in \mathbb{R}^{n-1}$, define

$$W_{x_2} := \left\{ x \in \mathbb{R} : \begin{bmatrix} x \\ x_2 \end{bmatrix} \in W \right\} \tag{16}$$

We will denote by $l(W_{x_2})$ the one-dimensional Lebesgue measure of W_{x_2}. Define the length of W as

$$l(W) = \sup_{x_2 \in \mathbb{R}^{n-1}} l(W_{x_2}) \,.$$

If $\bar{V} \subseteq S \times \mathbb{R}^n$ we define

$$\bar{V}_s = \{ x \in \mathbb{R}^n \mid (s, x) \in \bar{V} \}$$

and

$$l(\bar{V}) = \sup_{s \in S} l(\bar{V}_s) \,.$$

The linear contraction of the pair $(W, \bar{V})$ is defined as

$$C_l(W, \bar{V}) = \frac{l(W)}{l(\bar{V})} \,.$$

We can decompose

$$\mathbb{E}\left[T_{(W,\bar{V})} \right] \;=\; \int_{W_2} \left(\int_{W_{x_2}} T_{(W,\bar{V})}(x_1, x_2)\, dx_1 \right) \frac{1}{\lambda(W)}\, dx_2 = \int_{W_2} \mathbb{E}\left[T_{(W_{x_2},\bar{V})} \right] \frac{l(W_{x_2})}{\lambda(W)}\, dx_2 \tag{17}$$

where

$$\mathbb{E}\left[T_{(W_{x_2},\bar{V})} \right] = \int_{W_{x_2}} T_{(W,\bar{V})}(x_1, x_2) \frac{dx_1}{l(W_{x_2})}$$

depends on x_2 and is calculated with respect to the 1-dimensional uniform probability density on W_{x_2}. This shows that, if we obtain bounds on $\mathbb{E}[T_{(W_{x_2},\bar{V})}]$, then we can transfer them on $\mathbb{E}[T_{(W,\bar{V})}]$ using (17). To estimate $\mathbb{E}[T_{(W_{x_2},\bar{V})}]$ we will use a symbolic dynamics technique as in [8]. As a first step we need to introduce a symbolic description of the dynamics.

4.2 A symbolic description of the dynamics

Let

$$\mathcal{A} = \{ (s, i, j) \in S \times \mathbb{N}^2 \mid s \in S \,, \; i = 1, \ldots, M_s \,, \; j = 1, \ldots, N_s \} \,.$$

For any $(s, i, j) \in \mathcal{A}$, the subset $F_s^i \cap K_s^j$ can be decomposed as the union of maximal disjoint affine intervals in the direction of e_1. The family of all such intervals is denoted by $\mathcal{I}_s^{i,j}$. Let

$$\mathcal{I} := \bigcup_{(s,i,j) \in \mathcal{A}} \mathcal{I}_s^{i,j}$$

17

We can associate to (Λ, Γ), a sublanguage of $(\mathcal{A} \times \mathcal{I})^*$ as follows. From any initial state $x_0 \in W$ and $s_0 = 0$, the closed loop system generates an evolution (s_t, x_t), $t \in \mathbb{N}$. With this evolution we can associate the word $(s_0, i_0, j_0, I_0)(s_1, i_1, j_1, I_2) \cdots (s_t, i_t, j_t, I_t) \in (\mathcal{A} \times \mathcal{I})^*$ such that for every $k = 0, 1, \dots, t$ we have that $x_k \in F_{s_k}^{i_k}$, $x_k \in K_{s_k}^{j_k}$ and $x_k \in I_k \in \mathcal{I}_s^{i_k, j_k}$. The sublanguage of $(\mathcal{A} \times \mathcal{I})^*$ constituted by all these words as x_0 varies in W and $t \in \mathbb{N}$ will be denoted by Σ^*.

If $\bar{V} \subseteq S \times \mathbb{R}^n$, we define

$$\Sigma^*(\bar{V}) = \{(s_0, i_0, j_0, I_0)(s_1, i_1, j_1, I_1) \cdots (s_t, i_t, j_t, I_t) \in \Sigma^* \mid \{s_k\} \times I_k \not\subseteq \bar{V}, \forall k = 0, 1, \dots, t\}.$$

For any $I \in \mathcal{I}_0^{i,j}$ with $I \subseteq W$ the symbol $\gamma_t(I, \bar{V})$ denotes the number of distinct words of length t in $\Sigma^*(\bar{V})$ starting from the symbol $(0, i, j, I)$. Moreover, if $I \subseteq W$ is any interval in the direction of e_1. We define

$$\gamma_t(I, \bar{V}) = \sum_{i=1}^{M_0} \sum_{j=1}^{N_0} \gamma_t(I \cap F_0^i \cap K_0^j, \bar{V}). \tag{18}$$

The following result shows why the symbolic representation of the closed loop system is helpful for estimating the probability distribution of the entrance time.

Lemma 1 *Let $I \subseteq W$ be any interval in the direction of the first coordinate. For any $t \in \mathbb{N}$ we have that*

$$\mathbb{P}[T_{(I, \bar{V})} = t] \leq C_l(I, \bar{V})^{-1} \frac{\gamma_t(I, \bar{V})}{|\mu|^t} \tag{19}$$

$$\mathbb{P}[T_{(I, \bar{V})} \geq t] \geq 1 - C_l(I, \bar{V})^{-1} - C_l(I, \bar{V})^{-1} \sum_{k=1}^{t-1} \frac{\gamma_k(I, \bar{V})}{|\mu|^k}. \tag{20}$$

where $\mathbb{P}$ here denotes the uniform probability on I.

Proof Fix $i_0 \in \{1, \dots, M_0\}$ and $j_0 \in \{1, \dots, N_0\}$ and consider the word

$$(0, i_0, j_0, I_0)(s_1, i_1, j_1, I_1) \cdots (s_{t-1}, i_{t-1}, j_{t-1}, I_{t-1}) \in \Sigma^*(\bar{V}),$$

where $I_0 = I \cap F_0^{i_0} \cap K_0^{j_0}$. By definition of $\Sigma^*(\bar{V})$, the interval

$$(I_0) \cap (\Gamma_0^{-1} I_1) \cap \cdots \cap (\Gamma_0^{-1} \cdots \Gamma_{s_{t-2}}^{-1} I_{t-1}) \tag{21}$$

is non-empty and the map $\Gamma_{s_{t-1}} \Gamma_{s_{t-2}} \cdots \Gamma_0$ is affine on it. Notice now that the value of the state s_t is completely determined from s_{t-1} from the fact that we know that $x_{t-1} \in F_{s_{t-1}}^{i_{t-1}}$. Hence,

$$l(I_0 \cap (\Gamma_0^{-1} I_1) \cap \cdots \cap (\Gamma_0^{-1} \cdots \Gamma_{s_{t-2}}^{-1} I_{t-1}) \cap (\Gamma_0^{-1} \cdots \Gamma_{s_{t-1}}^{-1} \bar{V}_{s_t})) \leq \frac{l(\bar{V}_{s_t})}{|\mu|^t} \leq \frac{l(\bar{V})}{|\mu|^t}.$$

Since intervals (21) relative to different words are pairwise disjoint, we obtain

$$\mathbb{P}[T_{(I, \bar{V})} = t] = \frac{l(\{x \in I \mid T_{(I, \bar{V})}(x) = t\})}{l(I)} \leq \frac{l(\bar{V})}{|\mu|^t} \frac{\gamma_t(I, \bar{V})}{l(I)} = C_l(I, \bar{V})^{-1} \frac{\gamma_t(I, \bar{V})}{|\mu|^t}.$$

This proves (19). Estimation (20) immediately follows from (19). ∎

From this we obtain that, for any choice of $\bar{t} \in \mathbb{N}$, we have

$$\mathbb{E}\left[T_{(I, \bar{V})}\right] = \sum_{t=1}^{+\infty} \mathbb{P}[T_{(I, \bar{V})} \geq t] \geq \sum_{t=1}^{\bar{t}} \mathbb{P}[T_{(I, \bar{V})} \geq t] \geq \bar{t}(1 - C_l(I, \bar{V})^{-1}) - C_l(I, \bar{V})^{-1} \sum_{t=1}^{\bar{t}} \sum_{k=1}^{t-1} \frac{\gamma_k(I, \bar{V})}{|\mu|^k}. \tag{22}$$

Our fundamental goal is now to determine upper bounds on $\gamma_k(I, \bar{V})$, when I is any interval in the direction of e_1. In order to do this, we will first exhibit a graph representation of the language $\Sigma^*(\bar{V})$ and we will then use a result established in [8].

18

4.3 The Markovian representation

The language $\Sigma^*(\bar{V})$ can be represented by a graph as follows. Consider the directed graph with set of vertices $\Sigma^*(\bar{V})$ and set of edges $\mathcal{E}$ given by

$$(\omega_0 \omega_1 \cdots \omega_{t-1} \to \omega_0 \omega_1 \cdots \omega_{t-1}\omega_t) \in \mathcal{E} \quad \Longleftrightarrow \quad \omega_0 \omega_1 \cdots \omega_{t-1}\omega_t \in \Sigma^*(\bar{V}) \ . \qquad (23)$$

Notice that the words in $\Sigma^*(\bar{V})$ starting from $\omega_0 \in \Sigma^*(\bar{V})$ are in one-to-one correspondence with the finite paths on the graph starting from the vertex ω_0. This representation of $\Sigma^*(\bar{V})$ will be called a *Markov representation*. This can be simplified by considering an equivalence relation on the vertices. With each word $\omega_0 \omega_1 \cdots \omega_t \in \Sigma^*(\bar{V})$, we associate its *symbolic future*

$$\mathrm{fut}_\Sigma(\omega_0 \omega_1 \cdots \omega_t) = \left\{ \overline{\omega}_0 \overline{\omega}_1 \cdots \overline{\omega}_k \mid \overline{\omega}_0 = \omega_t \text{ and } \omega_0 \omega_1 \cdots \omega_t \overline{\omega}_1 \cdots \overline{\omega}_k \in \Sigma^*(\bar{V}) \right\} \ ,$$

which is a subset of $\Sigma^*(\bar{V})$. More roughly, the symbolic future of a word $\omega_0 \omega_1 \cdots \omega_t$ is the set of words whose concatenation with $\omega_0 \omega_1 \cdots \omega_t$ is in $\Sigma^*(\bar{V})$.

The concept of *geometric future* of a word $\omega_0 \omega_1 \cdots \omega_t \in \Sigma^*(\bar{V})$ can be introduced as follows. Assume that $\omega_k = (s_k, i_k, j_k, I_k)$ for $k = 0, 1, \ldots, t$. Then

$$\mathrm{fut}(\omega_0 \omega_1 \cdots \omega_t) := (\mathrm{fut}_d(\omega_0 \omega_1 \cdots \omega_t)), \mathrm{fut}_c(\omega_0 \omega_1 \cdots \omega_t))$$

where

$$\mathrm{fut}_d(\omega_0 \omega_1 \cdots \omega_t)) := s_t$$

is called the *discrete geometric future* and

$$\mathrm{fut}_c(\omega_0 \omega_1 \cdots \omega_t) = \Gamma_{s_{t-1}} \cdots \Gamma_{s_0} \left(I_0 \cap (\Gamma_{s_0}^{-1} I_1) \cap \ldots \cap (\Gamma_{s_0}^{-1} \cdots \Gamma_{s_{t-1}}^{-1} I_t) \right)$$

the *continuous geometric future*.

Straightforward considerations (see also [3, 8]) show that two words have the same symbolic future if and only if they have the same geometric future.

Using these equivalences it is possible to simplify the Markov representation of $\Sigma^*(\bar{V})$ introduced above (see [11] for a general treatment of this reduction method). Define $\mathcal{X}$ to be the quotient of the set $\Sigma^*(\bar{V})$ by the equivalence relation

$$\omega'_0 \cdots \omega'_{t'} \equiv \omega''_0 \cdots \omega''_{t''} \ \Leftrightarrow \ \mathrm{fut}_\Sigma(\omega'_0 \cdots \omega'_{t'}) = \mathrm{fut}_\Sigma(\omega''_0 \cdots \omega''_{t''}) \ . \qquad (24)$$

The elements of $\mathcal{X}$ will be called *states* and they will be denoted by the symbol $\mathbf{x}$. The symbol $\langle \omega_0 \omega_1 \cdots \omega_t \rangle$ represents the state consisting of the equivalent class which contains the word $\omega_0 \omega_1 \cdots \omega_t$. The equivalence relation defining $\mathcal{X}$ ensures that any state $\mathbf{x} \in \mathcal{X}$ has a well defined geometric future $\mathrm{fut}(\mathbf{x})$. In fact, the geometric future $\mathrm{fut}(\mathbf{x})$ uniquely determines the state $\mathbf{x}$. Edges can be naturally redefined on $\mathcal{X}$ to obtain a new labelled graph denoted $\mathcal{G}$ which is still a Markov representation of $\Sigma^*(\bar{V})$. The words in $\Sigma^*(\bar{V})$ starting from $\omega_0 \in \Sigma^*(\bar{V})$ are in one-to-one correspondence with the finite paths on $\mathcal{G}$ starting from the state $\langle \omega_0 \rangle$. This implies that if $I \in \mathcal{I}_0^{i,j}$, then $\gamma_t(I, \bar{V})$ is exactly the number of distinct paths of length t in $\mathcal{G}$ starting from the vertex $\langle (0, i, j, I) \rangle$.

4.4 Estimations of paths on the graph

In this subsection we will recall a result established in [8] which estimates the number of paths on a graph with some structure. Consider a direct graph $\mathcal{G}$ on a vertex set $\mathcal{X}$ (which is not necessarily finite or denumerable). We set from some notation. If $\mathcal{X}_1, \ldots, \mathcal{X}_k \subset \mathcal{X}$, we define

$\gamma_k[\mathbf{x}_1 \in \mathcal{X}_1, \dots, \mathbf{x}_k \in \mathcal{X}_k]$ to be the number of paths $\mathbf{x}_1 \cdots \mathbf{x}_k \in \mathcal{X}^*$ on the graph $\mathcal{G}$ such that $\mathbf{x}_1 \in \mathcal{X}_1, \dots, \mathbf{x}_k \in \mathcal{X}_k$.

The conditions that $\mathcal{G}$ must satisfy are the following. We assume there exists a finite partition

$$\mathcal{X} = \mathcal{X}_1 \cup \mathcal{X}_2 \cup \cdots \cup \mathcal{X}_{\mathbf{N}},$$

a subset $\mathcal{X}^P \subseteq \mathcal{X}$ and a function $q : \mathcal{X} \to]0, 1[$ with the following properties:

(A) There exist numbers $q_1, \dots, q_{\mathbf{N}} \in]0, 1[$ such that

$$q(\mathbf{x}) \le q_i, \quad \forall \mathbf{x} \in \mathcal{X}_i$$
$$q(\mathbf{x}) = q_i, \quad \forall \mathbf{x} \in \mathcal{X}_i^P := \mathcal{X}^P \cap \mathcal{X}_i.$$

(B) There exist positive numbers D_1 and α_1 such that, for every $\mathbf{x}' \in \mathcal{X}$ and $q \in]0, 1[$,

$$\gamma_t[\mathbf{x}_1 = \mathbf{x}', \mathbf{x}_2, \dots, \mathbf{x}_t \in \mathcal{X}, q(\mathbf{x}_t) \ge q] \le D_1 \frac{q(\mathbf{x}')}{q} \alpha_1^t.$$

(C) There exist positive numbers D_2 and α_2 such that, for every $\mathbf{x}' \in \mathcal{X}$ and $i \in \{1, \dots, \mathbf{N}\}$,

$$\gamma_t[\mathbf{x}_1 = \mathbf{x}', \mathbf{x}_2, \dots, \mathbf{x}_{t-1} \in \mathcal{X} \setminus \mathcal{X}^P, \mathbf{x}_t \in \mathcal{X}_i] \le D_2 \alpha_2^t.$$

Then, if we define

$$\gamma_t = \sum_{h=1}^{\mathbf{N}} \sup_{\mathbf{x}' \in \mathcal{X}_h} \gamma_t[\mathbf{x}_1 = \mathbf{x}', \mathbf{x}_2, \dots, \mathbf{x}_t \in \mathcal{X}], \qquad h = 1, \dots, \mathbf{N}$$

we have the following result.

Theorem 3 *[8] There exists $r \in \{1, \dots, \mathbf{N}\}$ and a constant H only depending on α_1, α_2, D_1, and D_2 such that:*

(1) If $\alpha_1 > \alpha_2$, then

$$\gamma_t \le \left[\sum_{s=1}^{r \wedge t} \binom{t-1}{s-1} \binom{r}{s} \left(\frac{s}{r}\right)^s \right] \left(\frac{\mathbf{N}H}{t \wedge \frac{\mathbf{N}H}{e}} \right)^{t \wedge \frac{\mathbf{N}H}{e}} \alpha_1^t. \tag{25}$$

(2) If $\alpha_1 \le \alpha_2$, then

$$\gamma_t \le \left[\sum_{s=1}^{r \wedge t} \binom{t+s-1}{2s-1} \binom{r}{s} \left(\frac{s}{r}\right)^s \right] \left(\frac{\mathbf{N}H}{t \wedge \frac{\mathbf{N}H}{e}} \right)^{t \wedge \frac{\mathbf{N}H}{e}} \alpha_2^t. \tag{26}$$

where e denotes the Neper constant and $a \wedge b$ denotes the minimum of the two numbers a and b.

4.5 Application to the expected entrance time estimate

We now show that the graph $\mathcal{G}$ constructed in Subsection 4.3 indeed satisfies the assumptions of previous theorem. We start by defining the partition. For any $(s, i, j) \in \mathcal{A}$ such that $\{s\} \times (F_s^i \cap K_s^j) \not\subseteq \bar{V}$ define

$$\mathcal{X}_{s,i,j} = \{\mathbf{x} \in \mathcal{X} \mid \text{fut}(\mathbf{x}) \subseteq \{s\} \times F_s^i \cap K_s^j\} .$$

The number of nonempty $\mathcal{X}_{s,i,j}$ as defined above is bounded from above by the number $\mathbf{N}$ defined at the begining of Section 4.

We now define $\mathcal{X}^P$ as the subclass of $\mathcal{X}$ consisting of the states of the type $\langle (s, i, j, I) \rangle$, where I is any interval which coincides with the intersection of K_s^j and any straight line in the direction of e_1. Finally, for any $\mathbf{x} \in \mathcal{X}$, we define

$$q(\mathbf{x}) = \frac{l(\text{fut}_c(\mathbf{x}))}{l(Z)} .$$

Let

$$q_{s,i,j} = \sup\{q(\mathbf{x}) \mid \mathbf{x} \in \mathcal{X}_{s,i,j}\} .$$

Elementary geometric considerations show that $q(\mathbf{x})$ is constant for all $\mathbf{x} \in \mathcal{X}_{s,i,j}^P := \mathcal{X}_{s,i,j} \cap \mathcal{X}^P$. Let $q_{s,i,j}$ be its value. This shows that assumption (A) holds. The assumptions (B) and (C) follow from the following lemmas.

Lemma 2 *Let $\mathbf{x}' \in \mathcal{X}$. Then*

$$\gamma_t[\mathbf{x}_1 = \mathbf{x}', \mathbf{x}_2, \ldots, \mathbf{x}_t \in \mathcal{X}, q(\mathbf{x}_t) \geq q] \leq \frac{q(\mathbf{x}')}{q} |\mu|^{t-1} .$$

Proof Let $s_j = \text{fut}_d(\mathbf{x}_j)$. Notice that the intervals of the form

$$\text{fut}_c(\mathbf{x}') \cap \Gamma_0^{-1} \text{fut}_c(\mathbf{x}_2) \cap \cdots \cap \Gamma_0^{-1} \cdots \Gamma_{s_{t-1}}^{-1} \text{fut}_c(\mathbf{x}_t) \qquad \mathbf{x}_2, \ldots, \mathbf{x}_t \in \mathcal{X}$$

constitute a family of disjoint subintervals of $\text{fut}_c(\mathbf{x}')$. This shows that

$$l[\text{fut}_c(\mathbf{x}')] \geq \sum_{\mathbf{x}_2, \ldots, \mathbf{x}_t \in \mathcal{X}} l\left[\text{fut}_c(\mathbf{x}') \cap \Gamma_0^{-1} \text{fut}_c(\mathbf{x}_2) \cap \cdots \cap \Gamma_0^{-1} \cdots \Gamma_{s_{t-1}}^{-1} \text{fut}_c(\mathbf{x}_t)\right] .$$

Notice moreover that $\Gamma_{s_{t-1}} \cdots \Gamma_0$ is affine on each of these intervals and so we have that

$$l\left[\text{fut}_c(\mathbf{x}') \cap \Gamma_0^{-1} \text{fut}_c(\mathbf{x}_2) \cap \cdots \cap \Gamma_0^{-1} \cdots \Gamma_{s_{t-1}}^{-1} \text{fut}_c(\mathbf{x}_t)\right] \leq \frac{l[\text{fut}_c(\mathbf{x}_t)]}{|\mu|^{t-1}}$$

if $\mathbf{x}' \mathbf{x}_2 \cdots \mathbf{x}_{t-1} \mathbf{x}_t$ is a possible path on the graph and it is 0 otherwise. This yields the result. ∎

Lemma 3 *For every $x' \in \mathcal{X}$, and $(s, i, j) \in \mathcal{A}$ we have that*

$$\gamma_t[\mathbf{x}_1 = x', \mathbf{x}_2, \ldots, \mathbf{x}_{t-1} \in \mathcal{X} \setminus \mathcal{X}^P, \mathbf{x}_t \in \mathcal{X}_{s,i,j}] \leq (2\nu)^{t-2} .$$

Proof We notice that there is an edge connecting a state $\mathbf{x}' =< (s, i, j, I) >$ to another state $\mathbf{x}''$ if and only if $\Gamma_s(\text{fut}_c(\mathbf{x}')) \cap \text{fut}_c(\mathbf{x}'') \neq \emptyset$. Since the map Γ_s is affine on the interval $\text{fut}_c(\mathbf{x}')$, then $\Gamma_s(\text{fut}_c(\mathbf{x}'))$ is also an interval which will be split by at most ν subintervals by the partition $\mathcal{F}_{s'}$ where $s' = \Lambda(s, F_s^i)$. Therefore at most 2ν followers of the state $\mathbf{x}'$ can be in $\mathcal{X} \setminus \mathcal{X}^P$. The result follows by applying this argument. ∎

Using Theorem 3, estimation (22) and decomposition (17) we obtain a lower bound estimation on $\mathbb{E}[T_{(W,\bar{V})}]$.

Theorem 4 *Assume that $|\mu| > 2\nu$. There exists $r \in \{1, \ldots, \mathbf{N}\}$ and a constant H, only depending on μ, ν and γ, such that, for all $\bar{t} \in \mathbb{N}$, we have the following:*

$$T \geq \bar{t}\left[1 - (\gamma C^{1/n})^{-1}\right] - (\gamma C^{1/n})^{-1}\left[\sum_{s=1}^{r \wedge \bar{t}-1} \binom{\bar{t}}{s+1}\binom{r}{s}\left(\frac{s}{r}\right)^s\right]\left(\frac{\mathbf{N}H}{\bar{t}-1 \wedge \frac{\mathbf{N}H}{e}}\right)^{\bar{t}-1 \wedge \frac{\mathbf{N}H}{e}}. \quad (27)$$

Proof It follows from definition (18), Theorem 3 and standard combinatorial techniques that

$$
\begin{aligned}
\sum_{k=1}^{t-1} \frac{\gamma_k(I,\bar{V})}{|\mu|^k} &\leq \sum_{k=1}^{t-1}\left[\sum_{s=1}^{r \wedge k}\binom{k-1}{s-1}\binom{r}{s}\left(\frac{s}{r}\right)^s\right]\left(\frac{\mathbf{N}H}{k \wedge \frac{\mathbf{N}H}{e}}\right)^{k \wedge \frac{\mathbf{N}H}{e}} \\
&\leq \left[\sum_{k=1}^{t-1}\sum_{s=1}^{r \wedge k}\binom{k-1}{s-1}\binom{r}{s}\left(\frac{s}{r}\right)^s\right]\left(\frac{\mathbf{N}H}{t-1 \wedge \frac{\mathbf{N}H}{e}}\right)^{t-1 \wedge \frac{\mathbf{N}H}{e}} \\
&\leq \left[\sum_{s=1}^{r \wedge t-1}\binom{r}{s}\left(\frac{s}{r}\right)^s\sum_{k=s}^{t-1}\binom{k-1}{s-1}\right]\left(\frac{\mathbf{N}H}{t-1 \wedge \frac{\mathbf{N}H}{e}}\right)^{t-1 \wedge \frac{\mathbf{N}H}{e}} \\
&\leq \left[\sum_{s=1}^{r \wedge t-1}\binom{t-1}{s}\binom{r}{s}\left(\frac{s}{r}\right)^s\right]\left(\frac{\mathbf{N}H}{t-1 \wedge \frac{\mathbf{N}H}{e}}\right)^{t-1 \wedge \frac{\mathbf{N}H}{e}}.
\end{aligned}
$$

With similar techniques it follows that

$$\sum_{n=1}^{\bar{t}}\sum_{k=1}^{n-1} \frac{\gamma_k(I,\bar{V})}{|\mu|^k} \leq \left[\sum_{s=1}^{r \wedge \bar{t}-1}\binom{\bar{t}}{s+1}\binom{r}{s}\left(\frac{s}{r}\right)^s\right]\left(\frac{\mathbf{N}H}{\bar{t}-1 \wedge \frac{\mathbf{N}H}{e}}\right)^{\bar{t}-1 \wedge \frac{\mathbf{N}H}{e}}.$$

A simple substitution in (22) now yields

$$\mathbb{E}\left[T_{(I,\bar{V})}\right] \geq$$

$$\bar{t}\left[1 - C_l(I,\bar{V})^{-1}\right] - C_l(I,\bar{V})^{-1}\left[\sum_{s=1}^{r \wedge \bar{t}-1}\binom{\bar{t}}{s+1}\binom{r}{s}\left(\frac{s}{r}\right)^s\right]\left(\frac{\mathbf{N}H}{\bar{t}-1 \wedge \frac{\mathbf{N}H}{e}}\right)^{\bar{t}-1 \wedge \frac{\mathbf{N}H}{e}}. \quad (28)$$

Using (17) and (28) we obtain

$$
\begin{aligned}
\mathbb{E}\left[T_{(W,\bar{V})}\right] &= \int_{v \in W_2} \mathbb{E}\left[T_{(W_{x_2},\bar{V})}\right]\frac{l(W_{x_2})}{\lambda(W)}\,dx_2 \geq \int_{v \in W_2}\left[\bar{t}\left[1 - C_l(W_{x_2},\bar{V})^{-1}\right] - C_l(W_{x_2},\bar{V})^{-1}\cdot\right. \\
&\qquad \left.\cdot \left[\sum_{s=1}^{r \wedge \bar{t}-1}\binom{\bar{t}}{s+1}\binom{r}{s}\left(\frac{s}{r}\right)^s\right]\left(\frac{\mathbf{N}H}{\bar{t}-1 \wedge \frac{\mathbf{N}H}{e}}\right)^{\bar{t}-1 \wedge \frac{\mathbf{N}H}{e}}\right]\frac{l(W_{x_2})}{\lambda(W)}\,dx_2 \\
&= \bar{t}(1 - \tilde{C}^{-1}) - \tilde{C}^{-1}\left[\sum_{s=1}^{r \wedge \bar{t}-1}\binom{\bar{t}}{s+1}\binom{r}{s}\left(\frac{s}{r}\right)^s\right]\left(\frac{\mathbf{N}H}{\bar{t}-1 \wedge \frac{\mathbf{N}H}{e}}\right)^{\bar{t}-1 \wedge \frac{\mathbf{N}H}{e}}
\end{aligned}
$$

$$(29)$$

where

$$\tilde{C}^{-1} = \int_{v \in W_2} C_l(W_{x_2},\bar{V})^{-1}\frac{l(W_{x_2})}{\lambda(W)}\,dx_2 = \frac{l(\bar{V})\lambda(W_2)}{\lambda(W)}$$

and where $\lambda(W_2)$ denotes the $(n-1)$-dimensional Lebesgue measure of W_2. Since, $l(\bar{V}) \leq d(V)$ and $\lambda(W_2) \leq d(W)^{n-1}$, we obtain

$$\tilde{C}^{-1} \leq \frac{d(V)d(W)^{n-1}}{\lambda(W)} \leq \gamma^{-1} \left(\frac{\lambda(V)}{\lambda(W)} \right)^{\frac{1}{n}} .$$

Inserting it in (29) we obtain the result. ∎

Remark: In the case $|\mu| \leq 2\nu$ an estimation very similar to (27) can be obtained using, in an analogous way, the second part of Theorem 3.

4.6 The main results

Using the technical results of previous subsection, we can now obtain inequality constraints involving T, C, and $\mathbf{N}$.

Theorem 5 *There exist $H_1 > 0$, $\beta_1 > 0$ and $C_1 > 1$, only depending on μ, ν and γ, such that*

$$C \geq C_1 \quad \text{and} \quad \frac{\lceil T \rceil}{\ln C} \leq \beta_1 \implies \mathbf{N} \geq H_1 \lceil T \rceil C^{\frac{1}{\lceil nT \rceil}} , \tag{30}$$

where $\lceil T \rceil$ is the smallest integer bigger than or equal to T.

Proof We prove it in the case when $|\mu| > 2\nu$, the other case being completely analogous. Using the inequality

$$\binom{r}{s} \leq \left(1 + \frac{r-s}{s} \right)^s e^s$$

which can be deduced from the Stirling approximation, we obtain

$$\sum_{s=1}^{r \wedge \bar{t}-1} \binom{\bar{t}}{s+1} \binom{r}{s} \left(\frac{s}{r} \right)^s \leq \sum_{s=1}^{\bar{t}-1} \binom{\bar{t}}{s+1} e^s \leq A_1 (1+e)^{\bar{t}} ,$$

for a suitable constant $A_1 > 0$. Inserting it in (27) and assuming $\bar{t} \geq 2$, we obtain

$$T \geq \bar{t}(1 - (\gamma C^{1/n})^{-1}) - (\gamma C^{1/n})^{-1} A_2^{\bar{t}-1} \left(\frac{\mathbf{N}H}{\bar{t}-1 \wedge \frac{\mathbf{N}H}{e}} \right)^{\bar{t}-1 \wedge \frac{\mathbf{N}H}{e}} , \tag{31}$$

for some constant A_2 which, with no loss of generality, we can assume to be greater than 1.

We now show that we can find $C' > 1$ and $\beta_1 > 0$ such that

$$C \geq C' \quad \text{and} \quad \frac{\lceil T \rceil}{\ln C} \leq \beta_1 \implies \lceil T \rceil \leq \frac{\mathbf{N}H}{e} . \tag{32}$$

Indeed, if $\lceil T \rceil > \mathbf{N}H/e$, then, choosing $\bar{t} := \lceil T \rceil + 1$, it follows from (31) that

$$\begin{aligned}
\lceil T \rceil &\geq (\lceil T \rceil + 1)(1 - (\gamma C^{1/n})^{-1}) - \gamma^{-1}(eA_2)^{\lceil T \rceil}(\gamma C^{1/n})^{-1} \\
&\geq (\lceil T \rceil + 1)(1 - (\gamma C^{1/n})^{-1}) - \gamma^{-1}(eA_2)^{\beta_1 \ln C}(\gamma C^{1/n})^{-1}
\end{aligned} \tag{33}$$

By taking the limit for $C \to +\infty$ in (33) we obtain that necessarily $\beta_1 \ln(eA_2) - 1/n \geq 0$. The claim is thus proven by choosing $\beta_1 < [n \ln(eA_2)]^{-1}$ and C' sufficiently large. Assuming now that (32) holds true and choosing again $\bar{t} := \lceil T \rceil + 1$ in (31), we obtain

$$\lceil T \rceil \geq (\lceil T \rceil + 1)(1 - (\gamma C^{1/n})^{-1}) - (\gamma C^{1/n})^{-1} \left(\frac{\mathbf{N}HA_2}{\lceil T \rceil} \right)^{\lceil T \rceil} . \tag{34}$$

Solving with respect to $\mathbf{N}$, we obtain

$$\mathbf{N} \geq \frac{\lceil T \rceil}{A_2 H}[\gamma C^{1/n} - \lceil T \rceil - 1]^{1/\lceil T \rceil} \geq \frac{\lceil T \rceil}{A_2 H}[\gamma C^{1/n} - \beta_1 \ln C - 1]^{1/\lceil T \rceil} . \tag{35}$$

Let $C'' > 0$ be such that

$$C > C'' \;\Rightarrow\; \gamma C^{1/n} - \beta_1 \ln C - 1 \geq \frac{\gamma}{2} C^{1/n} .$$

If we choose $C_1 := C' \vee C''$ and $H_1 = (A_2 H)^{-1}(1 \vee (\gamma/2))$ (where $\vee$ denotes the maximum) we obtain the result. $\blacksquare$

Remark The immediate application of the previous theorem shows that, for the class of memoryless deadbeat quantized feedback such as the ones considered in the previous section we have that T is constant and so we obtain the bound on $N = \mathbf{N} \geq$ const $C^{1/nT}$ which resembles the performance obtained there which was $N \sim$ const $C^{1/T}$. The difference in the exponents of C is due to the fact that the bound is consequence of an essentially one dimensional analysis and that we did not use the hypothesis that the system was controllable. In other words it is easy see that for the n-dimensional system in which

$$A = \begin{bmatrix} \mu & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

it is possible to obtain the same performance as the scalar system $x_{t+1} = \mu x_t + u_t$.

Theorem 5 yields the following important corollary.

Corollary 1 *There exist $\beta_1 > 0$ and $C_1 > 1$ and a positive continuous decreasing map $\delta : \mathbb{R}_+ \to \mathbb{R}_+$ only depending on μ, ν and γ, such that*

$$C \geq C_1 \quad \text{and} \quad \frac{\lceil T \rceil}{\ln C} \leq \beta \leq \beta_1 \;\Longrightarrow\; \frac{\mathbf{N}}{\ln C} \geq \delta(\beta) . \tag{36}$$

Proof Consider H_1, C_1 and β_1 as defined in Theorem 5 and assume, without loss of generality that $\beta_1 \leq 1/n$. Result then immediately follows from Theorem 5 using the fact that the function $x \mapsto xC^{\frac{1}{nx}}$ is decreasing on $(0, \ln C/n]$. $\blacksquare$

From this last result we can also easily obtain the following corollary.

Corollary 2 *There exists $C_1 > 1$ and a positive continuous decreasing map $\beta : \mathbb{R}_+ \to \mathbb{R}_+$, only depending on μ, ν and γ, such that*

$$C \geq C_1 \quad \text{and} \quad \frac{\mathbf{N}}{\ln C} \leq \delta \;\Longrightarrow\; \frac{\lceil T \rceil}{\ln C} \geq \beta(\delta) . \tag{37}$$

In the the remaining of the section we will show that when $|\mu| > 2\nu$ other results can be obtained.

Theorem 6 *Assume that $|\mu| > 2\nu$. Then there exist $H_2 > 0$, $\beta_2 > 0$ and $C_2 > 1$, only depending on μ, ν and γ, such that*

$$C \geq C_2 \quad \text{and} \quad \frac{\mathbf{N}}{\ln C} \leq \beta_2 \;\Longrightarrow\; T \geq H_2 \mathbf{N} C^{\frac{1}{n\mathbf{N}}} . \tag{38}$$

Proof Using again Stirling approximation and standard combinatorics we have that

$$\sum_{s=1}^{r \wedge \bar{t}-1} \binom{\bar{t}}{s+1}\binom{r}{s}\left(\frac{s}{r}\right)^s \leq \sum_{s=1}^{\mathbf{N}} \binom{\bar{t}}{s+1}\binom{N}{s} = \binom{\bar{n}+\mathbf{N}}{\mathbf{N}+1}$$

$$= \frac{\bar{n}}{\mathbf{N}+1}\binom{\bar{n}+\mathbf{N}}{\mathbf{N}} \leq \bar{t}\left(1+\frac{\bar{t}}{N}\right)^{\mathbf{N}} e^{\mathbf{N}},$$

We obtain in this way

$$T \geq \bar{t}\left[1 - (\gamma C^{1/n})^{-1} - (\gamma C^{1/n})^{-1}\left(1+\frac{\bar{t}}{\mathbf{N}}\right)^{\mathbf{N}} A^{\mathbf{N}}\right], \tag{39}$$

for some $A > 0$. If in (39) we choose $\bar{t} = \left\lceil D\mathbf{N}C^{1/n\mathbf{N}}\right\rceil$ for some constant $D > 0$ which will be fixed later, we have that

$$\frac{T}{\mathbf{N}C^{1/n\mathbf{N}}} \geq D\left[1 - (\gamma C^{1/n})^{-1} - \left(2 + DC^{1/n\mathbf{N}}\right)^{\mathbf{N}} A^{\mathbf{N}}(\gamma C^{1/n})^{-1}\right]$$

$$= D\left[1 - (\gamma C^{1/n})^{-1} - \left(2C^{-1/n\mathbf{N}} + D\right)^{\mathbf{N}} A^{\mathbf{N}}\gamma^{-1}\right]. \tag{40}$$

Assume now that $\mathbf{N} \leq \beta \ln C$ for some $\beta > 0$ which will be chosen later. This implies that

$$\frac{T}{\mathbf{N}C^{1/n\mathbf{N}}} \geq D\left[1 - (\gamma C^{1/n})^{-1} - \left((2e^{-1/n\beta} + D)A\right)^{\mathbf{N}}\gamma^{-1}\right] \tag{41}$$

If we choose D and β small enough, we obtain

$$\left((2e^{-1/n\beta} + D)A\right)^{\mathbf{N}}\gamma^{-1} \leq 1/3 \quad \forall \mathbf{N}$$

If we now choose C_1 in such a way that $(\gamma C^{1/n})^{-1} < 1/3$ for $C > C_1$, we obtain that (38) holds with $D_1 = D/3$ and $\beta_1 = \beta$. ∎

From previous result we easily obtain the following generalization of Corollary 1.

Corollary 3 *Assume that $|\mu| > 2\nu$. Then there exists $C_2 > 1$ and a positive continuous decreasing map $\delta : \mathbb{R}_+ \to \mathbb{R}_+$, only depending on μ, ν and γ, such that*

$$C \geq C_2 \quad \text{and} \quad \frac{T}{\ln C} \leq \beta \implies \frac{\mathbf{N}}{\ln C} \geq \delta(\beta). \tag{42}$$

4.7   Some interpretative comments

Theorem 5 and Corollary 1 provide inequality constraints between the three parameters $\mathbf{N}$, T and C assuming that $T/\ln C$ is sufficiently small. In particular Corollary 1 shows that, if we want to obtain stabilization with sufficiently small logarithmic time rate, then the complexity parameter $\mathbf{N}$ has to grow at least logarithmically in C. Actually Corollary 1 also establishes an explicit quantitative link between $T/\ln C$ and $\mathbf{N}/\ln C$: indeed, it follows from the proof of Corollary 1, that the map $\delta(\beta)$ in (36), for β sufficiently small, has the form

$$\delta(\beta) = H_1\beta\omega^{1/\beta} \tag{43}$$

for some $\omega > 1$. Notice that this fact applies even if the matrix A of the system to be controlled is already asymptotically stable but not nilpotent. Of course, in that case we would obtain logarithmic time rate without any control. However, if we want to decrease the logarithmic time rate below a certain threshold, then we must use in any case a logarithmic number of quantization subsets. Corollary 2 instead says that if $\mathbf{N}/\ln C$ remains bounded, we will obtain expected entrance times T which will grow at least logarithmically in C. In other words, sublogarithmic time growth, namely faster convergence, can not be obtained if we do not allow superlogarithmic growth in the complexity parameter $\mathbf{N}$. In the case when $|\mu| > 2\nu$ we have more precise results. In fact Corollary 3 says that in order to obtain any type of logarithmic time rate behavior, then we must have $\mathbf{N}$ also growing logarithmically in C. We expect this to be true for any non asymptotically stable situation, namely even when $1 \le |\mu| \le 2\nu$, but we have not been able to prove this, yet. Notice also that since $\mathbf{N} \le NML$ and M is considered to be constant, Theorem 5, Corollaries 1 and 3 still holds true substituting NL at the place of $\mathbf{N}$. This yields the results we had anticipated in the introduction and shows that we can not improve very much what it was obtained in the examples shown in the previous section. Both in the nesting and in the zooming case we have indeed logarithmic growth with respect to C both of the mean entrance time T and of the product NL as it can be seen from (11) and from table 14. Moreover the relation between $T/\ln C$ and $NL/\ln C$ in these examples has the same form of the one in (43): if we call $T/\ln C = \beta$, we have $NL/\ln C = H\beta\tilde{\omega}^{1/\beta}$ for suitable constants H and $\tilde{\omega}$ which vary in the various cases.

Finally observe that, in principle, Theorem 6 leaves the possibility of having stabilization with $\mathbf{N}/\ln C \to 0$ for $C \to +\infty$. Actually, memoryless quantized stabilization strategies with a fixed N, not depending on C, do indeed exist and have been studied in [7, 8]. They only yield what has been called 'almost stability' in the sense that almost every point in the initial set W is driven into the target set V. However their expected time T is finite and it was proven in [6] to grow linearly with C in the one-dimensional case $n = 1$. This perfectly agrees with the result expressed in Theorem 6.

5   Conclusions

We have introduced a general setting which allowed us to analyze quantized feedback stabilizers of a linear discrete time system. We have introduced three indices N, M, and L describing the complexity of the quantized feedback, a performance index T which is the expected time used by the controller to drive the state of the system from an initial set W to a final target set V, and finally the contraction rate C which is the ratio between the volumes of W and V. We have compared various examples on the basis of these indices and we have proven some results expressing fundamental bounds among the above indices.

All our analysis has been carried on in the assumption that the system had only one input and that we had full state observation. First goal of our future research is to remove these assumptions and to consider general input output linear discrete time systems. This will also force us to consider quantizations in more than one dimension which was one of the basic assumptions for the bounds we have obtained.

Another interesting open problem is related to the possibility of error transmission in the control with communication constraints framework we proposed. As we pointed out, such errors have catastrophic consequences for the quantized controllers with memory proposed in the literature. Therefore different strategies need to be developed for solving this problem.

References

[1] J. Baillieul, "Feedback designs in information-based control," in *Proc. of the Workshop on Stochastic Theory and Control*, Kansas, 2001, pp. 35–57, Springer-Verlag.

[2] R.W. Brockett and D. Liberzon. Quantized feedback stabilization of linear systems. *IEEE Trans. Automatic Control*, AC-45:1279–1289, 2000.

[3] J. Buzzi. Intrinsic ergodicity of affine maps in $[0, 1]^d$. *Monat. fur Mathematik*, 124:97–118, 1997.

[4] D.F. Delchamps. Stabilizing a linear system with quantized state feedback. *IEEE Trans. Automat. Control*, AC-35:916–924, 1990.

[5] N. Elia and S.K Mitter. Stabilization of linear systems with limited information. *IEEE Trans. Automat. Control*, AC-46:1384–1400, 2001.

[6] F. Fagnani A performance analysis of chaotic quantized feedback stabilizers. submitted.

[7] F. Fagnani and S. Zampieri Stability analysis and synthesis for scalar linear systems with a quantized feedback. *IEEE Trans. Automat. Control*, AC-48:1569–1584, 2003.

[8] F. Fagnani and S. Zampieri A symbolic dynamics approach to performance analysis of quantized feedback systems: the scalar case. submitted.

[9] A. Lasota and M.C. Mackey. *Chaos, fractals, and noise.* Springer Verlag, 1994.

[10] D. Liberzon. On stabilization of linear systems with limited information. *IEEE Trans. Automat. Control*, AC-48:304–307, 2003.

[11] D. Lind and B. Marcus. *Symbolic Dynamics and Coding.* Cambridge Univ., 1995.

[12] C. Liverani. Rigorous numerical investigation of the statistical properties of piecewise expanding maps. A feasibility study. *Nonlinearity*, 3:463–490, 2001.

[13] G.N. Nair and R.J. Evans. Stabilization with data-rate-limited feedback: tightest attainable bounds. *Systems and Control Letters*, 41:49–56, 2000.

[14] G.N. Nair and R.J. Evans. Exponential stabilisability of finite-dimensional linear systems with limited data rates. *Automatica*, 39:585–593, 2002.

[15] I.R. Petersen and A.V. Savkin Multirate stabilization of multivariable discrete-time linear systems via a limited capacity communication channel, Proc. of CDC Conf., pp. 304–309, Las Vegas, 2002.

[16] B. Picasso, F. Gouaisbaut, A. Bicchi Construction of invariant and attractive sets for quantized-input linear systems, Proc. of CDC Conf., pp. 824–829, Las Vegas, 2002.

[17] B. Saussol. Absolutely continuous invariant measures for multidimensional expanding maps. *Israel J. Math*, 116:223–248, 2000.

[18] S. Tatikonda, *Control under communication constraints*, Ph.D. thesis, MIT, Cambridge, 2000.

[19] W.S. Wong and R.W. Brockett. Systems with finite communication bandwidth constraints II: Stabilization with limited information feedback. *IEEE Trans. Automatic Control*, AC-44:1049–1053, 1999.