Quantized control via locational optimization

Francesco Bullo, *Senior Member, IEEE,* Daniel Liberzon, *Senior Member, IEEE,*

***Abstract*— This paper studies state quantization schemes for feedback stabilization of control systems with limited information. The focus is on designing the least destabilizing quantizer subject to a given information constraint. We explore several ways of measuring the destabilizing effect of a quantizer on the closed-loop system, including (but not limited to) the worst-case quantization error. In each case, we show how quantizer design can be naturally reduced to a version of the so-called multicenter problem from locational optimization. Algorithms for obtaining solutions to such problems, all in terms of suitable Voronoi quantizers, are discussed. In particular, an iterative solver is developed for a novel weighted multicenter problem which most accurately represents the least destabilizing quantizer design. A simulation study is also presented.**

***Index Terms*— Feedback stabilization, locational optimization, quantized control.**

I. Introduction

In this paper we study control systems whose state variables are quantized. We think of a quantizer as a device that converts a real-valued signal into a piecewise constant one taking a finite set of values. The recent papers [3], [9], [14] discuss various situations where this type of quantization arises and provide references to the literature. Mathematically, a quantizer can be described by a piecewise constant function $q : \mathcal{D} \subset \mathbb{R}^n \to \mathcal{Q}$, where $\mathcal{Q}$ is a finite subset of $\mathbb{R}^n$ with a fixed number of elements N. Here n is the state dimension of a given system and $\mathcal{D}$ is a closed region of interest in the state space. We denote the elements of $\mathcal{Q}$ by $q_1, \ldots, q_N$ and refer to them as *quantization points*. The sets $W_i := \mathrm{cl}\{x \in \mathcal{D} : q(x) = q_i\}$, $i \in \{1, \ldots, N\}$ associated with fixed values of the quantizer form a partition[1] of the region $\mathcal{D}$ and are called *quantization regions* (cl denotes closure). We will sometimes identify a quantizer q with the corresponding pair $(\mathcal{Q}, \mathcal{W})$, where $\mathcal{W} := \{W_1, \ldots, W_N\}$. Quantized values of the state represent a limited information flow from the system to a feedback controller: the state is not completely known to the controller, but it is only known which one of a fixed number of quantization regions contains the current state at each instant of time. (Assuming that the quantization points are known to the controller, one can think of the information flow as a string of integers from 1 to N, transmitted at the times when the state trajectory crosses the boundaries between the quantization regions.)

In the literature it is usually assumed that quantization regions are fixed in advance and have specific shapes, most often rectilinear. Here we are interested in the situation where the number N of quantizer values is a given constraint in the control problem, but the control designer has flexibility in choosing a specific configuration of quantization regions (whose shapes can in principle be arbitrary) and quantization points. While there has been some research on systems with quantization regions of arbitrary shapes [19], [16] and on the relationship between the choice of quantization regions and the behavior of the closed-loop system [9], [14], the general problem of determining the "best" quantizer for a particular control task such as feedback stabilization remains largely open.

A feedback law which globally asymptotically stabilizes a given system in the absence of quantization will in general fail to provide global asymptotic stability of the closed-loop system that arises in the presence of state quantization. Two phenomena accounting for changes in the system's behavior caused by quantization will play the role in what follows. The first one is saturation: if the quantized signal is outside the range of the quantizer, then the quantization error is large, and the control law designed for the ideal case of no quantization may lead to instability. The second one is deterioration of performance near the equilibrium: as the difference between the current and the desired values of the state becomes small, higher precision is required, and so in the presence of quantization errors asymptotic convergence is typically lost. These phenomena manifest themselves in the existence of two nested invariant regions, $\mathcal{R}_1$ and $\mathcal{R}_2$, such that all trajectories of the quantized system starting in the bigger region $\mathcal{R}_1$ approach the smaller one $\mathcal{R}_2$ while no further convergence guarantees can be given. Chattering on the boundaries between quantization regions is possible, and solutions are to be interpreted in the sense of Filippov if necessary [11]; however, this issue will not play a significant role in the subsequent stability analysis, because we will work with a single $\mathcal{C}^1$ Lyapunov function on $\mathcal{R}_1 \setminus \mathcal{R}_2$. (One way to prevent chattering, and thus ensure a finite data rate, would be to introduce a dwell time; cf. [14].)

In Section II we explain how the destabilizing effect of a given quantizer can be measured. We introduce the concept of a *destabilization measure* which, in conjunction with an arbitrary stabilizing feedback law and a corresponding Lyapunov function, can be used to determine an ultimate bound on solutions. One example of such a destabilization measure is the *worst-case quantization error* $\max_{x \in \mathcal{D}} |q(x) - x|$. However, it turns out that there exist other destabilization measures which are actually more suitable in the present context. Although the parameters of the control system are used in the stability

F. Bullo is with Department of Mechanical and Environmental Engineering, University of California at Santa Barbara, Santa Barbara, CA 93106, USA. Email: bullo@engineering.ucsb.edu. Supported by DARPA/AFOSR Award F49620-02-1-0325.

D. Liberzon is with Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. Email: liberzon@uiuc.edu. Supported by NSF ECS-0134115 CAR and DARPA/AFOSR MURI F49620-02-1-0325 Awards.

[1] A collection $\{W_1, \ldots, W_N\}$ of subsets of $\mathcal{D}$ is a *partition* of $\mathcal{D}$ if the intersection between the relative interior of any two W_i is empty and the union of all W_i equals $\mathcal{D}$.

analysis, the destabilization measure itself is a function of the quantization regions and quantization points only. The quantizer design problem then naturally reduces to an optimization problem which consists in minimizing such a measure over all quantizers satisfying the information constraint. We describe this procedure for three different types of quantizers arising from uniform, radial and spherical, and radially weighted quantization.

After casting quantizer design as an optimization problem, we proceed to explain how techniques from *optimal facility location* (or *locational optimization*) yield new insights into this problem as well as efficient algorithms for solving it. Facility location problems concern the location of a fixed number of facilities that provide service demanded by users; the objective is to minimize the average or maximal distance from sets of demand points to facilities. We focus here on settings continuous in the location of both the facilities and the demand points (i.e., both facilities and demand points take values in a continuum of points, such as a polytope or an ellipsoid). Facility location problems are surveyed in [7]. Computational geometric aspects in continuous facility location are discussed in [24], [23] and indirectly in textbooks on computational geometry [6]. Relevant background on computational geometric methods in locational optimization is provided in Section III.

For example, a classical problem of interest in locational optimization is the so-called *multimedian problem*. It consists in choosing a collection of N points $q_1, q_2, \ldots, q_N$ in a bounded region $\mathcal{D} \subset \mathbb{R}^n$ so as to minimize the quantity $E(\min_{i\in\{1,\ldots,N\}} |q_i - x|^2)$, where the expected value is computed with respect to some probability density function on $\mathcal{D}$ and $|\cdot|$ denotes the Euclidean norm. Solutions of this problem are given by *centroidal Voronoi tessellations*; see [28], [8]. Within the context of quantization and information theory, the multimedian problem is known as the fixed-rate minimum-distorsion quantizer design [7], [13]. One of the early references on this problem is the classic work by Lloyd [18], who obtains optimality conditions and introduces a famous insightful algorithm. The multimedian problem is related to the problem of state moment stabilization of linear systems with limited data rate [21].

Since we are working in the deterministic setting, we will find that the problem relevant for our purposes is the *multicenter problem*, discussed in [28], [27]. This is a somewhat less frequently encountered variant of the multimedian problem, which is obtained by replacing the expected value by the worst-case value; it can also be stated as the problem of covering a given region with overlapping balls of minimal radius. The connection between the quantized control problem and the multicenter problem, although very natural, apparently has not been pursued before. In Section III we present a general formulation of the multicenter problem with weighting factors. We then discuss solutions of specific versions of this problem corresponding to the three types of quantization considered in Section II, all in terms of suitable Voronoi quantizers. We show how existing algorithms can handle the first two approaches, and then develop a new algorithm for the last one which gives less conservative results. We note that while weighted multimedian problems are commonly encountered, our formulation appears to be novel in that it introduces weighting factors in the context of the continuous multicenter problem.

In Section IV we provide a comparative simulation study of the three quantizer designs considered here and a standard rectangular quantizer for a two-dimensional linear system. These simulation results—as well as existing studies of the related multimedian problem, such as [12]—indicate that by solving the quantized feedback stabilization problem with the help of locational optimization techniques, one may obtain quite interesting quantization patterns. For the multicenter problem in the plane, for example, a typical Voronoi region is a hexagon. Consequently, hexagonal quantization regions for planar systems have some advantages over more traditional rectangular ones.

II. Quantization and stability

We assume that the stabilization problem in the absence of quantization has been solved, in the sense that a state feedback control law is known such that the origin is a globally asymptotically stable equilibrium point of the ideal closed-loop system. In the presence of quantization, we adopt the "certainty equivalence" control paradigm; namely, we let the same control law act on the quantized state $q(x)$, where q is a quantizer on $\mathbb{R}^n$ (or on a smaller region of interest). The problem under consideration is to characterize the destabilizing effect of the quantizer q, with the goal of obtaining an ultimate bound on solutions of the closed-loop system starting in a given bounded region. We first discuss this problem for general nonlinear systems and then develop more specific results for linear systems, moving from simpler but conservative to more complicated but sharper formulations.

A. Nonlinear systems

We start with the general situation where the process to be controlled is modeled by the system

$$\dot{x} = f(x, u), \qquad x \in \mathbb{R}^n,\ u \in \mathbb{R}^m. \tag{1}$$

All vector fields and control laws are understood to be sufficiently regular (e.g., $\mathcal{C}^1$) so that existence and uniqueness of solutions are ensured. Suppose that some nominal static feedback law $u = k(x)$ is given (with minor changes, dynamic feedback laws can also be used). In the presence of state quantization, we consider the feedback law $u = k(q(x))$ and the corresponding closed-loop system

$$\dot{x} = f(x, k(q(x))) = f(x, k(x + e)) \tag{2}$$

where $e := q(x) - x$ represents the *quantization error*.

Besides stabilizing the nominal system (1), the feedback law k clearly must possess some robustness property with respect to the measurement error e. To this end we impose the following assumption: there exists a $\mathcal{C}^1$ function $V : \mathbb{R}^n \to \mathbb{R}$ such that for some class $\mathcal{K}_\infty$ functions $\alpha_1, \alpha_2, \alpha_3, \rho$ and for

all $x, e \in \mathbb{R}^n$ we have[2] $\alpha_1(|x|) \leq V(x) \leq \alpha_2(|x|)$ and

$$|x| \geq \rho(|e|) \Rightarrow \frac{\partial V}{\partial x} f(x, k(x+e)) \leq -\alpha_3(|x|).$$

(Here and later, $|\cdot|$ denotes the standard Euclidean norm.) This amounts to the property that the control law $u = k(x)$ *input-to-state stabilizes* the closed-loop system with respect to the measurement error e [25], [26]. This assumption is rather restrictive and can be relaxed at the expense of obtaining weaker results (for linear systems, however, it is an automatic consequence of closed-loop asymptotic stability for $e \equiv 0$). There is also considerable research on designing control laws satisfying this assumption. These issues are further discussed elsewhere [15], [16], [17].

Pick a positive number M and consider the ball $\mathcal{B}_M := \{x \in \mathbb{R}^n : |x| \leq M\}$. Consider the worst-case quantization error

$$\Delta := \max_{x \in \mathcal{B}_M} |e| \tag{3}$$

(this quantity is sometimes also referred to as the *sensitivity* of the quantizer). The following result is fairly straightforward to obtain (see [16, Lemma 2] and [17, Lemma 5.2]).

Lemma 1 *Assume that*

$$\alpha_1(M) > \alpha_2 \circ \rho(\Delta). \tag{4}$$

Then the sets

$$\mathcal{R}_1 := \{x \in \mathbb{R}^n : V(x) \leq \alpha_1(M)\} \tag{5}$$

and

$$\mathcal{R}_2 := \{x \in \mathbb{R}^n : V(x) \leq \alpha_2 \circ \rho(\Delta)\}$$

are invariant regions for the system (2)*. Moreover, all solutions of* (2) *that start in the set $\mathcal{R}_1$ enter the smaller set $\mathcal{R}_2$ in finite time. An upper bound on this time is*

$$T = \frac{\alpha_1(M) - \alpha_2 \circ \rho(\Delta)}{\alpha_3 \circ \rho(\Delta)}. \tag{6}$$

This lemma implies, in particular, that all solutions starting in $\mathcal{R}_1$ at time $t = t_0$ satisfy the ultimate bound

$$|x(t)| \leq \alpha_1^{-1} \circ \alpha_2 \circ \rho(\Delta), \qquad \forall t \geq t_0 + T \tag{7}$$

with T given by the formula (6). We regard the quantity Δ defined by (3) as a *destabilization measure* of the quantizer q. For given feedback law k and Lyapunov function V, an ultimate bound on solutions can be described by a class $\mathcal{K}_\infty$ function of this measure as shown by (7). It is not hard to see that if the number N of quantization regions is sufficiently large, then Δ can be made small enough for the inequality (4) to hold. Minimizing Δ—and consequently the size of the attracting invariant region $\mathcal{R}_2$—over all possible choices of the quantizer q corresponds to the following optimization problem:

$$\min_{\mathcal{Q}, \mathcal{W}} \quad \max_{i \in \{1,\ldots,N\}} \quad \max_{x \in W_i} \quad |q_i - x| \tag{8}$$

[2]Recall that a function $\alpha : [0, \infty) \to [0, \infty)$ is said to be of *class* $\mathcal{K}$ if it is continuous, strictly increasing, and $\alpha(0) = 0$. If $\alpha \in \mathcal{K}$ is unbounded, then it is said to be of *class* $\mathcal{K}_\infty$. The existence of functions α_1 and α_2 with the indicated property simply means that V is positive definite and radially unbounded; it will be convenient to have these functions explicitly in the following analysis.

where $\mathcal{Q} = \{q_1, \ldots, q_N\}$ is a set of quantization points and $\mathcal{W} = \{W_1, \ldots, W_N\}$ is a partition of $\mathcal{B}_M$ into quantization regions. (We could work with partitions of $\mathcal{R}_1$ rather than $\mathcal{B}_M$, but this requires the knowledge of V and also may be less computationally feasible for non-quadratic V.) The optimization problem (8) is known as the *multicenter problem* in computational geometry; we defer its detailed discussion until Section III-A.

B. Worst-case quantization error for linear systems

We now specialize to the case when the process is described by the linear system

$$\dot{x} = Ax + Bu, \qquad x \in \mathbb{R}^n,\ u \in \mathbb{R}^m. \tag{9}$$

The linear system structure can be utilized to define a destabilization measure in several different ways. Suppose that the system (9) is stabilizable, so that for some matrix K the eigenvalues of $A + BK$ have negative real parts. Then there exists a unique positive definite symmetric matrix P such that

$$(A + BK)^T P + P(A + BK) = -I. \tag{10}$$

We let $\lambda_{\min}(P)$ and $\lambda_{\max}(P)$ denote the smallest and the largest eigenvalue of P, respectively. We assume that $B \neq 0$ and $K \neq 0$; this is no loss of generality because the case of interest is when A is not a stable matrix.

The quantized state feedback control law $u = Kq(x)$ yields the closed-loop system

$$\dot{x} = Ax + BKq(x). \tag{11}$$

Take a positive number M and consider the ellipsoids

$$\mathcal{R}_1 := \{x \in \mathbb{R}^n : x^T P x \leq \lambda_{\min}(P) M^2\} \tag{12}$$

and

$$\mathcal{R}_2 := \{x \in \mathbb{R}^n : x^T P x \leq \lambda_{\max}(P) 4(1+\varepsilon)^2 \|PBK\|^2 \Delta^2\} \tag{13}$$

where $\varepsilon > 0$ is arbitrary and Δ is the worst-case quantization error defined by (3). Then we have the following linear counterpart of Lemma 1. Although this result is known (see [16, Lemma 1] and [17, Lemma 5.1]), we sketch a proof because it will be needed in the sequel; this proof also differs slightly from the one given in the references.

Lemma 2 *Assume that*

$$\lambda_{\min}(P) M^2 > \lambda_{\max}(P) 4(1+\varepsilon)^2 \|PBK\|^2 \Delta^2. \tag{14}$$

Then the ellipsoids $\mathcal{R}_1$ and $\mathcal{R}_2$ defined by (12) *and* (13) *are invariant regions for the system* (11)*. Moreover, all solutions of* (11) *that start in the ellipsoid $\mathcal{R}_1$ enter the smaller ellipsoid $\mathcal{R}_2$ in finite time. An upper bound on this time is*

$$T = \frac{\lambda_{\min}(P) M^2 - \lambda_{\max}(P) 4(1+\varepsilon)^2 \|PBK\|^2 \Delta^2}{4\|PBK\|^2 \Delta^2 (1+\varepsilon)\varepsilon}. \tag{15}$$

Proof: Rewrite (11) as

$$\dot{x} = (A + BK)x + BKe$$

where $e := q(x) - x$ is the quantization error as before. In view of (10), the derivative of the function

$$V(x) := x^T P x$$

along solutions of this system satisfies

$$\dot{V} = -x^T x + 2x^T PBKe \leq -|x|^2 + 2|x||PBKe|. \quad (16)$$

We can rewrite this as

$$\begin{aligned}\dot{V} &\leq -\frac{\varepsilon}{1+\varepsilon}|x|^2 - \frac{1}{2(1+\varepsilon)}|x|^2 + 2(1+\varepsilon)|PBKe|^2 \\ &- \left(\frac{1}{\sqrt{2(1+\varepsilon)}}|x| - \sqrt{2(1+\varepsilon)}|PBKe|\right)^2 \\ &\leq -\frac{\varepsilon}{1+\varepsilon}|x|^2 - \frac{1}{2(1+\varepsilon)}\left(|x|^2 - \Big(2(1+\varepsilon)|PBKe|\Big)^2\right).\end{aligned}$$

Therefore, we have

$$|x| \geq 2(1+\varepsilon)|PBKe| \;\Rightarrow\; \dot{V} \leq -\frac{\varepsilon}{1+\varepsilon}|x|^2. \quad (17)$$

Using the inequality $|PBKe| \leq \|PBK\||e|$, where $\|\cdot\|$ stands for the matrix norm induced by the Euclidean norm, we obtain

$$|x| \geq 2(1+\varepsilon)\|PBK\||e| \;\Rightarrow\; \dot{V} \leq -\frac{\varepsilon}{1+\varepsilon}|x|^2. \quad (18)$$

It is now straightforward to conclude the result. □

As a consequence of this lemma, an ultimate bound on solutions starting in $\mathcal{R}_1$ at time $t = t_0$ is

$$|x(t)| \leq \sqrt{\frac{\lambda_{\max}(P)}{\lambda_{\min}(P)}}2(1+\varepsilon)\|PBK\|\Delta, \qquad \forall t \geq t_0 + T$$

with T given by the formula (15). Decreasing ε to 0, we see that solutions (asymptotically) approach the ellipsoid

$$\{x \in \mathbb{R}^n : x^T P x \leq \lambda_{\max}(P)4\|PBK\|^2\Delta^2\}.$$

Thus we still consider Δ as a destabilization measure. As in the nonlinear setting, this leads to the optimization problem (8). If N is large enough, then Δ can be made small enough so that the inequality (14) holds (for a given feedback gain K).

Remark 1 We note that another approach is to work with (17) directly, avoiding the use of the induced norm $\|PBK\|$. Define

$$\Delta_{PBK} := \max_{x \in \mathcal{R}_1} |PBKe|.$$

The result of Lemma 2 still holds if $\|PBK\|^2\Delta^2$ is replaced by Δ^2_{PBK} everywhere in the statement of that lemma. This yields a less conservative ultimate bound and motivates the following optimization problem:

$$\min_{\mathcal{Q},\,\mathcal{W}} \quad \max_{i \in \{1,\ldots,N\}} \quad \max_{x \in W_i} \quad |PBK(q_i - x)| \quad (19)$$

where $\mathcal{Q}$ is a set of quantization points as before and $\mathcal{W}$ is a partition of $\mathcal{B}_M$ (or $\mathcal{R}_1$) into N regions. This problem is in general lower-dimensional compared to (8) because the subspace $\ker(PBK)$ can be ignored. (Note that PBK is a singular matrix whenever $m < n$.) Therefore, for the same N the optimal value for this problem will be significantly lower than that for (8). However, Δ_{PBK} is not really a destabilization measure in the sense used in this paper, because it depends on the feedback gain matrix K. While it gives better results for a fixed feedback law, quantizer design based on this destabilization measure needs to be redone if the feedback law is changed, and is not suitable for switching between several feedback laws. For these reasons we henceforth focus on destabilization measures that are independent of a particular feedback law used. □

Remark 2 It is clear that in Lemma 2, the system's behavior is important only for $x \in \mathcal{R}_1 \setminus \mathcal{R}_2$. In other words, redefining the quantizer arbitrarily outside $\mathcal{R}_1 \setminus \mathcal{R}_2$ does not affect the result. This means that we can design the quantizer more efficiently and preserve or decrease the ultimate bound on solutions. For example, if the quantizer that solves the optimization problem (19) involves several quantization points inside $\mathcal{R}_2$, we can move some of them to $\mathcal{R}_1 \setminus \mathcal{R}_2$ to achieve better coverage there. We can also restate the problem (19) in terms of partitions of $\mathcal{R}_1 \setminus \mathcal{R}_2$ or of a spherical annulus containing this set. □

Remark 3 Lemma 2 suggests that among stabilizing state feedback gains K, the ones that provide smaller ultimate bounds for the solutions of the quantized system (for a fixed quantizer) are those with smaller values of the induced matrix norm $\|PBK\|$, where P is given by (10). In this regard, it is interesting to observe the following: if the open-loop system $\dot{x} = Ax$ is not asymptotically stable, then for every stabilizing feedback gain K and the corresponding positive definite symmetric matrix P satisfying (10) we have

$$\|PBK\| \geq 1/2 \quad (20)$$

and the inequality is strict if $\dot{x} = Ax$ is unstable. To see this, use (16) and the definition of e to write

$$\dot{V} \leq -|x|^2\left(1 - 2\|PBK\|\frac{|q(x) - x|}{|x|}\right). \quad (21)$$

This formula will be used again several times in the sequel. Now, note that if q is chosen to take the value 0 in a neighborhood of the origin, then the right-hand side of (21) equals $-|x|^2(1-2\|PBK\|)$ there, and so $1-2\|PBK\|$ cannot be positive since A is not stable. It is also straightforward to show (20) directly: just multiply (10) on the left by v^T and on the right by v, where v is a normalized eigenvector of $A^T P + PA$ with a nonnegative eigenvalue. □

C. Radial and spherical quantization for linear systems

In the previous developments, the required bounds on the quantization error do not depend on the size of the state. This leads to uniform quantization, in the sense that quantization points are distributed uniformly over the region of interest. However, it is well known that more efficient quantization schemes are those which provide lower precision far away from the origin and higher precision close to the origin. Quantizers with a logarithmic scale are particularly useful; see [9]. Loosely speaking, with logarithmic quantization one has the same number of quantization points in the vicinity of

every sphere centered at the origin in the state space, whereas with uniform quantization this number grows with the radius. This observation suggests introducing a "direct product" of one quantizer on a unit sphere and another along the radial direction, which is what we do next.

Let us write

$$x = |x|\text{vers}(x)$$

where

$$\text{vers}(x) := \frac{x}{|x|}.$$

We represent the quantizer accordingly as

$$q(x) = q^r(|x|)q^s(\text{vers}(x)) \tag{22}$$

where q^r takes N_1 positive real values, q^s takes N_2 values on or inside the unit sphere, and N_1 and N_2 are some positive integers such that $N_1N_2 \leq N$. This means that we introduce two separate quantizers, one for $|x|$ and the other for $\text{vers}(x)$. The set of quantization points for the resulting overall quantizer q is formed by the N_1N_2 pairwise products of the values of q^r and q^s.

Let us introduce the worst-case quantization error on the unit sphere corresponding to q^s:

$$\Delta_s := \max_{|x|=1} |q^s(x) - x|. \tag{23}$$

As before, pick a positive number M. We will take q^r to be a logarithmic quantizer, defined as follows: given a pair of numbers a, b satisfying

$$0 < a < 1 < b, \tag{24}$$

let

$$q^r(z) := (a^i/b^{i-1})M, \quad \text{for } z \in \left((a/b)^i M, (a/b)^{i-1}M\right), \; i \in \{1, \ldots, N_1\} \tag{25}$$

and define q^r at the endpoints of these intervals to make it continuous from the right or from the left. Consider the ellipsoid

$$\mathcal{R}_2 := \{x \in \mathbb{R}^n : x^T P x \leq \lambda_{\max}(P)(a/b)^{2N_1}M^2\}. \tag{26}$$

Lemma 3 *Assume that*

$$\lambda_{\min}(P)M^2 > \lambda_{\max}(P)(a/b)^{2N_1}M^2 \tag{27}$$

and

$$\Delta_s < \frac{1-\varepsilon}{2\|PBK\|} \tag{28}$$

for some $\varepsilon > 0$. Let

$$a := 1 - \frac{1-\varepsilon}{2\|PBK\|} + \Delta_s\,, \qquad b := 1 + \frac{1-\varepsilon}{2\|PBK\|} - \Delta_s. \tag{29}$$

Then (24) *holds. With q^r given by* (25) *and q given by* (22)*, the ellipsoids $\mathcal{R}_1$ and $\mathcal{R}_2$ defined by* (12) *and* (26) *are invariant regions for the system* (11)*. Moreover, all solutions of* (11) *that start in the ellipsoid $\mathcal{R}_1$ enter the smaller ellipsoid $\mathcal{R}_2$ in finite time. An upper bound on this time is*

$$T = \frac{\lambda_{\min}(P)M^2 - \lambda_{\max}(P)(a/b)^{2N_1}M^2}{(a/b)^{2N_1}M^2\varepsilon}. \tag{30}$$

Proof: The fact that the numbers a and b defined by (29) satisfy the inequalities (24) follows directly from (20) and (28). It is a simple matter to check from (25) that we have

$$(a/b)^{N_1}M \leq z \leq M \;\Rightarrow\; \left|\frac{q^r(z)}{z} - 1\right| \leq \max\left\{b-1, 1-a\right\}.$$

For the values of a and b given by (29), this becomes

$$(a/b)^{N_1}M \leq z \leq M \;\Rightarrow\; \left|\frac{q^r(z)}{z} - 1\right| \leq \frac{1-\varepsilon}{2\|PBK\|} - \Delta_s. \tag{31}$$

Now, from the triangle inequality and the fact that $|q^s(\text{vers}(x))| \leq 1$ for all x by construction, we obtain

$$\begin{aligned}
|q(x) - x| &\leq \left|q^r(|x|)q^s(\text{vers}(x)) - |x|q^s(\text{vers}(x))\right| \\
&\quad + \left||x|q^s(\text{vers}(x)) - |x|\text{vers}(x)\right| \\
&\leq \left|q^r(|x|) - |x|\right| + |x|\left|q^s(\text{vers}(x)) - \text{vers}(x)\right| \\
&\leq |x|\left(\left|\frac{q^r(|x|)}{|x|} - 1\right| + \Delta_s\right)
\end{aligned}$$

where the last inequality follows from (23). Substituting this into (21), we conclude that the derivative of V along solutions of the closed-loop system satisfies

$$\dot{V} \leq -|x|^2\left[1 - 2\|PBK\|\left(\left|\frac{q^r(|x|)}{|x|} - 1\right| + \Delta_s\right)\right]. \tag{32}$$

Combining this with the formula (31), we conclude that $\dot{V} \leq -\varepsilon|x|^2$ for all $x \in \mathcal{R}_1 \setminus \mathcal{R}_2$, from which it is straightforward to conclude the result. □

For fixed N_1 and N_2, the quantity Δ_s defined by (23) provides a destabilization measure (for q^s). When K is given and Δ_s satisfies the inequality (28) for some $\varepsilon > 0$, we can construct q^r via (25) and compute an ultimate bound on solutions using Lemma 3. Minimizing Δ_s corresponds to the following optimization problem:

$$\min_{\mathcal{Q}^s, \mathcal{W}^s} \quad \max_{i \in \{1,\ldots,N_2\}} \quad \max_{x \in W_i^s} \quad |q_i^s - x| \tag{33}$$

where $\mathcal{Q}^s = \{q_1^s, \ldots, q_{N_2}^s\}$ is a set of points on or inside the unit sphere and $\mathcal{W}^s = \{W_1^s, \ldots, W_{N_2}^s\}$ is a partition of the unit sphere. This problem can be solved by the same algorithm as the problem (8), as will be described in Section III-A. The quantity (33) will not exceed the right-hand side of (28) if N_2 is sufficiently large. For a given N, the values of N_1 and N_2 satisfying $N_1N_2 \leq N$ which yield the smallest ultimate bound are not easy to compute analytically and in general seem to depend on the stabilizing feedback gain K; however, there is a finite number of choices for these integers and we can find the optimal values by trying all of them. We remark that in the context of the multimedian problem, the idea of spherical coordinates quantization has been exploited before, and in particular the trade-off between the numbers of values for the radial and the spherical directions has been studied; see [29] and the references therein.

Remark 4 It is straightforward to derive similar results using the norm defined by the Lyapunov function, i.e., $\|x\| := \sqrt{x^T P x}$, instead of the Euclidean norm. This gives rise to an optimization problem on an ellipsoid rather than on a sphere. □

D. Radially weighted quantization for linear systems

The need for logarithmic quantization patterns is evidenced by the fact that it is the ratio $|e|/|x|$, and not the absolute value of the quantization error $|e|$ itself, that needs to be small. This is clear from the formulas (18) and (21). The approach of Section II-C leads to an "aligned" logarithmic quantization pattern, in the sense that quantization points on spheres of different radii are aligned along the same radial directions. However, it is not hard to see that non-aligned quantization patterns may achieve better coverage. This suggests proceeding from (21) in a more direct fashion.

To this end, pick two numbers $M > m > 0$ and consider the ellipsoids $\mathcal{R}_1$ given by (12) and

$$\mathcal{R}_2 := \{x \in \mathbb{R}^n : x^T P x \leq \lambda_{\max}(P) m^2\}. \tag{34}$$

Define

$$\Delta_{rw} := \max_{x \in \mathcal{R}_1 \setminus \mathcal{R}_2} \frac{|q(x) - x|}{|x|}. \tag{35}$$

The following result easily follows by virtue of (21), and the proof is omitted.

Lemma 4 *Assume that*

$$\lambda_{\min}(P) M^2 > \lambda_{\max}(P) m^2 \tag{36}$$

and

$$\Delta_{rw} \leq \frac{1 - \varepsilon}{2\|PBK\|} \tag{37}$$

for some $\varepsilon > 0$. Then the ellipsoids $\mathcal{R}_1$ and $\mathcal{R}_2$ defined by (12) *and* (34) *are invariant regions for the system* (11)*. Moreover, all solutions of* (11) *that start in the ellipsoid $\mathcal{R}_1$ enter the smaller ellipsoid $\mathcal{R}_2$ in finite time. An upper bound on this time is*

$$T = \frac{\lambda_{\min}(P) M^2 - \lambda_{\max}(P) m^2}{m^2 \varepsilon}. \tag{38}$$

The quantity Δ_{rw} defined by (35) provides another destabilization measure for q, in relation to a pair of numbers $M > m > 0$. Given a stabilizing feedback gain K, we can check the inequalities (36) and (37) and, if they are satisfied, obtain an ultimate bound on solutions from Lemma 4. (It is also clear from (21) that Δ_{rw} provides a lower bound on the rate of decay of solutions in $\mathcal{R}_1 \setminus \mathcal{R}_2$.) This leads us to the following optimization problem:

$$\min_{\mathcal{Q}, \mathcal{W}} \quad \max_{i \in \{1, \ldots, N\}} \quad \max_{x \in W_i} \quad \frac{|q_i - x|}{|x|} \tag{39}$$

where $\mathcal{Q} = \{q_1, \ldots, q_N\}$ is a set of quantization points and $\mathcal{W} = \{W_1, \ldots, W_N\}$ is a partition of the annulus $\{x \in \mathbb{R}^n : m \leq |x| \leq M\}$ into quantization regions. The inequality (37) will hold for a given K if N is sufficiently large.

The optimization problem (39) is different in structure from the ones we encountered earlier, and apparently has not been studied in the locational optimization literature. We henceforth call it the *radially weighted multicenter problem*. It turns out that while this problem is more challenging than the others, it is still computationally tractable. We will develop an algorithm for solving it in Section III-B.

Remark 5 In this paper, we are assuming that quantizer design can only be performed once and cannot be changed on-line. In situations where one can recompute the locations of quantization regions and quantization points on-line, it is possible to achieve global asymptotic stability of the quantized closed-loop system by using the dynamic quantization strategy developed in [3], [16]. In fact, a simple rescaling of the quantizer every T units of time would suffice. For this to work, we need to ensure that $\mathcal{R}_2$ is a strict subset of $\mathcal{R}_1$ in each of the previous schemes. This means that the optimization problems formulated earlier remain relevant, except that by passing from static to dynamic quantization we basically pass from an optimal to a suboptimal quantizer design objective. An interesting direction for future research is to incorporate a cost for recomputing the quantization parameters into the quantizer design problem. □

III. Continuous multicenter problems in facility location

In this section we present a class of optimization problems related to the field of facility location; see the discussion in Section I and the survey [7]. The facility location problem we consider will have as special cases the optimization problems studied in Section II, and in particular the problems (8) from Sections II-A and II-B, (33) from Section II-C, and (39) from Section II-D.

Let us review some preliminary concepts. Given a compact region $\mathcal{D} \subset \mathbb{R}^n$ and a set of N points $\mathcal{Q} = \{q_1, \ldots, q_N\}$ in $\mathbb{R}^n$, the *Voronoi partition* $\mathcal{V} = \{V_1, \ldots, V_N\}$ of $\mathcal{D}$ generated by $\mathcal{Q}$ is defined according to

$$V_i := \{x \in \mathcal{D} : \ |x - q_i| \leq |x - q_j|, \quad \forall j \neq i\}. \tag{40}$$

When it is useful to emphasize the dependency on $\mathcal{Q}$, we shall write $\mathcal{V}(\mathcal{Q})$ or $V_i(\mathcal{Q})$. When $\mathcal{D}$ is a polytope in $\mathbb{R}^n$, each Voronoi region V_i is a polytope, otherwise V_i is the intersection between a polytope and $\mathcal{D}$. The faces of the polytope which defines V_i are given by hyperplanes of points in $\mathbb{R}^n$ that are equidistant from q_i and q_j, $j \neq i$; among the latter, only "neighboring" points play a role. Note that this (standard) construction remains valid when $\mathcal{D}$ is a lower-dimensional subset of $\mathbb{R}^n$, such as a sphere. We refer to [6], [23] for comprehensive treatments of Voronoi partitions.

Let $\mathcal{Q} = \{q_1, \ldots, q_N\}$ be a collection of points in $\mathbb{R}^n$ and let $\mathcal{W} = \{W_1, \ldots, W_N\}$ be a partition of $\mathcal{D}$. In what follows, we shall concern ourselves with the function

$$\mathcal{H}(\mathcal{Q}, \mathcal{W}) := \max_{i \in \{1, \ldots, N\}} \quad \max_{x \in W_i} \quad \phi(x) f(|x - q_i|) \tag{41}$$

where $\phi : \mathcal{D} \to [0, \infty)$ is continuous non-negative and $f : [0, \infty) \to [0, \infty)$ is continuous, non-decreasing and unbounded. We also assume that ϕ does not identically vanish on $\mathcal{D}$. We investigate the optimization problem

$$\min_{\mathcal{Q}, \mathcal{W}} \quad \mathcal{H}(\mathcal{Q}, \mathcal{W}) \tag{42}$$

and refer to it as the *weighted multicenter problem*. In general, $\mathcal{H}$ is a nonlinear non-convex function of the locations $\mathcal{Q}$ and of the partition $\mathcal{W}$. Accordingly, its global minima can

be obtained only numerically via nonlinear programming algorithms. However, this and related facility location problems [28], [27], [8] have some peculiar structure that helps us characterize optimal solutions and design useful iterative algorithms. Let us start by considering the *weighted 1-center problem* over $\mathcal{D}$, i.e., take $N = 1$.

Lemma 5 *The function $\mathcal{H}_1 : \mathbb{R}^n \to [0, \infty)$ defined by*

$$\mathcal{H}_1(q) := \mathcal{H}(\{q\}, \{\mathcal{D}\}) = \max_{x \in \mathcal{D}} \quad \phi(x) f(|x - q|)$$

is continuous, radially unbounded, and quasiconvex.[3] *If f is convex and ϕ is constant, then $\mathcal{H}_1$ is convex.*

Proof: The function $\mathcal{H}_1$ is continuous because it is the maximum of a compact family of continuous functions. Further, $\mathcal{H}_1$ is radially unbounded because for every $x^* \in \mathcal{D}$ such that $\phi(x^*) > 0$ we have $\mathcal{H}_1(q) \geq \phi(x^*) f(|x^* - q|)$ and f is unbounded. To show the other statements, we invoke certain properties of convex and quasiconvex functions; see Sections 3.2 and 3.4 in [2]. At a fixed x, the function $q \mapsto f(|x - q|)$ is quasiconvex because it is the composition of a convex function with a non-decreasing function. Furthermore, if f is non-decreasing and convex, then $q \mapsto f(|x - q|)$ is convex because, at a fixed x, it is the composition of a convex function with a convex non-decreasing function. If f is convex, then $q \mapsto \max_{x \in \mathcal{D}} f(|x - q|)$ is convex because it is the pointwise supremum over a set of convex functions. For general ϕ and f, the function $\mathcal{H}_1$ is quasiconvex because it is the weighted pointwise supremum of quasiconvex functions. □

Next, we let $\mathrm{co}(\mathcal{D})$ denote the convex hull of $\mathcal{D}$ and study the global minima of $\mathcal{H}_1$.

Lemma 6 *The set of global minimum points for $\mathcal{H}_1$ is compact, convex and has a non-empty intersection with $\mathrm{co}(\mathcal{D})$. If f is strictly increasing, then all global minimum points belong to $\mathrm{co}(\mathcal{D})$.*

Proof: The fact that the set of global minimum points is compact and convex is an immediate consequence of continuity, radial unboundedness, and quasiconvexity. Let us prove the non-empty intersection with $\mathrm{co}(\mathcal{D})$. Suppose that $q^* \notin \mathrm{co}(\mathcal{D})$ is a global minimum point for $\mathcal{H}_1$. Let $p^* \in \mathrm{co}(\mathcal{D})$ be the closest point to q^*, i.e., $p^* := \mathrm{argmin}_{x \in \mathrm{co}(\mathcal{D})} |q^* - x|$. Then $|x - p^*| < |x - q^*|$ for all $x \in \mathcal{D}$, so that, for all $x \in \mathcal{D}$, we have $\phi(x) f(|x - p^*|) \leq \phi(x) f(|x - q^*|) \leq \max_{x \in \mathcal{D}} \phi(x) f(|x - q^*|) = \mathcal{H}_1(q^*)$. Therefore, $\mathcal{H}_1(p^*) = \max_{x \in \mathcal{D}} \phi(x) f(|x - p^*|) \leq \mathcal{H}_1(q^*)$ and p^* also belongs to the set of global minimum points. When f is strictly increasing, the previous argument leads to $\mathcal{H}_1(p^*) < \mathcal{H}_1(q^*)$, which contradicts the assumption that q^* is a global minimum. □

Lemmas 5 and 6 show that the weighted 1-center problem over $\mathcal{D}$ is a quasiconvex optimization problem, i.e., it consists in minimizing the quasiconvex function $\mathcal{H}_1$ over the convex set $\mathrm{co}(\mathcal{D})$. It is known that every quasiconvex optimization problem can be solved by iterative techniques (via a bisection algorithm solving a convex feasibility problem at each step; see Section 4.2.5 in [2]). We call $q^*(\mathcal{D})$ a *weighted center* of the region $\mathcal{D}$ if it is a (possibly non-unique) global minimum point:

$$q^*(\mathcal{D}) := \operatorname*{argmin}_{q \in \mathrm{co}(\mathcal{D})} \quad \max_{x \in \mathcal{D}} \quad \phi(x) f(|x - q|).$$

Now, it is useful to return to the general weighted multicenter problem (41), (42) and define $\mathcal{W} \mapsto \mathcal{Q}^*(\mathcal{W})$ as the map that associates to $\mathcal{W}$ a collection of N (possibly non-unique) global minimum points for the corresponding weighted 1-center problems; in other words, $\mathcal{Q}^*(\{W_1, \ldots, W_N\}) := \{q^*(W_1), \ldots, q^*(W_N)\}$. Note that these weighted centers are well defined since each W_i is compact. Finally, define the *Lloyd map* (or the Lloyd algorithm) $\mathcal{L} : (\mathcal{Q}, \mathcal{W}) \mapsto (\mathcal{Q}', \mathcal{W}')$ where $\mathcal{W}' := \mathcal{V}(\mathcal{Q})$ and $\mathcal{Q}' := \mathcal{Q}^*(\mathcal{W}')$. The following result is a relatively straightforward consequence of LaSalle Invariance Principle for discrete-time dynamical systems; further convergence properties are under current investigation in [5].

Lemma 7 *At a fixed $\mathcal{Q}$, a global minimum of $\mathcal{W} \mapsto \mathcal{H}(\mathcal{Q}, \mathcal{W})$ is achieved at $\mathcal{W} = \mathcal{V}(\mathcal{Q})$. At a fixed $\mathcal{W}$, a global minimum of $\mathcal{Q} \mapsto \mathcal{H}(\mathcal{Q}, \mathcal{W})$ is achieved at $\mathcal{Q} = \mathcal{Q}^*(\mathcal{W})$. The Lloyd map is a descent algorithm for the cost function $\mathcal{H}$, i.e., an application of the map is guaranteed not to increase $\mathcal{H}$. The cost is guaranteed to decrease in one iteration if no active*[4] *point $q_j \in \mathcal{Q}$ is a weighted center of its region W_j. Given an initial pair $(\mathcal{Q}_0, \mathcal{W}_0)$, the sequence $\{\mathcal{L}^k(\mathcal{Q}_0, \mathcal{W}_0),\ k \geq 0\}$ approaches the largest set invariant under $\mathcal{L}$ on which $\mathcal{H}(\mathcal{L}(\mathcal{Q}, \mathcal{W})) = \mathcal{H}(\mathcal{Q}, \mathcal{W})$.*

Remark 6

(i) Fixed points of the Lloyd map are *weighted central Voronoi quantizers*, i.e., pairs $(\mathcal{Q}, \mathcal{W})$ such that $\mathcal{W}$ is the Voronoi partition generated by $\mathcal{Q}$ and at the same time the points in $\mathcal{Q}$ are weighted centers for $\mathcal{W}$. It is an open conjecture that the iteration described in the lemma converges to local minima of $\mathcal{H}$. Nevertheless, the algorithm is of interest to us because it is guaranteed to improve a given quantizer design and provides a good indication as to whether or not N is large enough to achieve the control objective.

(ii) The classic Lloyd algorithm is tailored to the continuous multimedian problem as it appears, for example, in the problem of fixed-rate minimum-distorsion quantizer design; see [7], [13]. The classic Lloyd algorithm differs from the one considered here only in the fact that the points in $\mathcal{Q}$ are moved to the centroids—as opposed to the weighted centers—of the respective Voronoi regions. (Centroids are solutions of the 1-median problems.)

(iii) The results in Lemmas 5, 6, and 7 provide an algorithm for the solution of the relevant multicenter problems via 1) quasiconvex programming and 2) Voronoi partition computations. It is important to observe that intense research activity is ongoing on both problems and that numerical iterative algorithms are available for solving them; e.g., see [1], [2] and references therein. □

[3] Recall that a *quasiconvex* function is a function defined on a convex domain and with convex sublevel sets.

[4] We call q_j *active* if $\mathcal{H}(\mathcal{Q}, \mathcal{W}) = \max_{x \in W_j} \phi(x) f(|x - q_j|)$, i.e., the maximum over i is achieved at the index j.

Next, we consider the specific settings that arise in the quantizer design problems discussed in the previous section. We characterize additional properties of the multicenter problem (8), the spherical multicenter problem (33), and the radially weighted multicenter problem (39). To implement the Lloyd algorithm, two tasks must be carried out repeatedly. One consists in computing the Voronoi partition for a given set of points $\mathcal{Q}$, which is accomplished by the standard procedure described earlier. The other amounts to computing a weighted center for each set W_i in a given partition. Thus for each of the specific multicenter problems, we must now discuss how to solve the corresponding 1-center problem. Some additional remarks on the properties of these particular multicenter problems will also be provided.

A. Multicenter problem

Let us first consider the problem (8) arising in Sections II-A and II-B. The domain is a ball centered at the origin or, more generally, an ellipsoid, i.e., $\mathcal{D} = \{x \in \mathbb{R}^n : x^T P x \leq 1\}$ for some positive definite symmetric matrix P. Note also that the problem (19) arising in Section II-B reduces, via a linear change of coordinates, to the multicenter problem considered here in a lower dimension.

In the problem (8), the weighting function ϕ is identically equal to 1 and the performance function f is the identity map. Under these conditions, we refer to the optimization problem (42) simply as the multicenter problem; see [28], [27]. The multicenter problem can be equivalently restated as the problem of covering the region $\mathcal{D}$ with a given number of (possibly overlapping) balls of smallest radius. If $\mathcal{B}_1 \subset \mathbb{R}^n$ is the unit ball centered at the origin, and if $R\mathcal{B}_1 + q$ denotes the ball of radius R centered at a point q, the problem reads:

$$\min\ R, \qquad \text{subject to} \quad \bigcup_{i\in\{1,\ldots,N\}} (R\mathcal{B}_1 + q_i) \supseteq \mathcal{D}.$$

Let us analyze the 1-center problem. From Lemma 5 we know that this is a convex optimization problem. For each region V_i, the optimal solution $q^*(V_i)$ is the center of the minimal-radius enclosing sphere for V_i. This center is unique because the minimal-radius enclosing sphere is the intersection of all enclosing spheres. When $V_i \subset \mathbb{R}^2$ is a polygon, this sphere is referred to as the smallest enclosing circle and algorithms are available to compute it; see [6, Chapter 4]. When $V_i \subset \mathbb{R}^n$ is a polytope, the smallest enclosing ellipsoid (in particular, sphere) can be computed via iterative convex optimization algorithms; see [2, Section 8.4]. For a Voronoi region V_i near the boundary of $\mathcal{D}$, which is not a polytope, we can under-approximate it by a polytope generated by the vertices of V_i and suitable additional points on the intersection of V_i with the boundary of $\mathcal{D}$, and then compute the center of this polytope. For a sufficiently close under-approximation, this center will also be the center of V_i.

When $\mathcal{D}$ is a unit cube in $\mathbb{R}^n$, the optimal value of the problem (8) satisfies the bounds

$$\frac{1}{2\lfloor \sqrt[n]{N} \rfloor} \leq \Delta \leq \frac{\sqrt{n}}{2\lfloor \sqrt[n]{N} \rfloor}.$$

The upper bound is easily obtained by constructing a uniform cubical quantization pattern, while the lower bound is known as Sukharev's lower bound on dispersion [22], [20]. In the present case when $\mathcal{D}$ is a ball, it is straightforward to obtain similar bounds by considering inscribed and superscribed cubes for $\mathcal{D}$. The upper bound can be used to evaluate the convergence of the Lloyd algorithm. When the lower bound on Δ is not small enough for the inequality (4) or (14) to hold, it indicates that a different destabilization measure and/or a different stabilizing feedback law must be used, or that N must be increased.

It is also useful to recall some known facts about the multimedian problem. It is conjectured in [12] that for N sufficiently large, the optimal quantizer with respect to the uniform probability density is given by a tessellation (i.e., translation and rotation) of a fixed polytope, except near the boundary of the region of interest. In two dimensions, polygons that can give rise to such tessellations are equilateral triangles, rectangles, and regular hexagons. Among these, the hexagon is optimal, because it has the smallest mean-square quantization error with respect to its centroid per unit volume. This result remains true if we consider the worst-case rather than mean-square quantization error, which is the quantity being minimized in the multicenter problem. The hexagon achieves the smallest error with respect to its center. (For the unit volume regular hexagon this error is approximately 0.62, compared with 0.707 for the square and 0.936 for the equilateral triangle; the unit-volume disk gives the error of 0.564 but disks cannot be used to obtain tessellations.) In Section IV we will indeed see hexagonal patterns arising as solutions of the multicenter problem.

The spherical multicenter problem (33) from Section II-C corresponds to the setting where $\mathcal{D} = \{x \in \mathbb{R}^n : |x| = 1\}$ is the unit sphere in $\mathbb{R}^n$. Since the spherical multicenter problem is formulated in terms of the Euclidean distance in $\mathbb{R}^n$, Voronoi partitions of the sphere can be constructed as explained earlier for the general case. Voronoi regions will be intersections of polytopes with the unit sphere. The center of each Voronoi region V_i is the center of the minimal-radius enclosing sphere for V_i. We can consider a polytope in $\mathbb{R}^n$ generated by the vertices of V_i and perhaps some other points in V_i. If enough points are taken, then the center of this polytope will also be the center of V_i. As we explained earlier, computing the center of a polytope is a computationally tractable task.

B. Radially weighted multicenter problem

Here, we study the problem (39) formulated in Section II-D, where the domain is the spherical annulus $\mathcal{D} = \{x \in \mathbb{R}^n : m \leq |x| \leq M\}$. We consider the corresponding radially weighted 1-center problem over a set $V \subset \mathcal{D}$:

$$\min_{q\in \mathrm{co}(V)} \quad \max_{x\in V} \quad \frac{|q-x|}{|x|}. \tag{43}$$

The problem is well-posed because V is a subset of $\mathcal{D}$ and therefore does not contain the origin. In what follows, we take V to be a polytope; if it is not, we approximate it by a polytope as before. We begin by making the following observation.

Lemma 8 *The optimal cost in the problem* (43) *is smaller than* 1 *if and only if the set* V *is separated from the origin by a hyperplane.*

Proof: Suppose first that V is separated from the origin by a hyperplane, so that $0 \notin \mathrm{co}(V)$. Let $\tilde{q}$ be the projection of the origin onto $\mathrm{co}(V)$, i.e., $\tilde{q} := \mathrm{argmin}_{x\in\mathrm{co}(V)} |x|$. By construction, $|x-\tilde{q}| < |x|$ for all $x \in V$, hence $\max_{x\in V} |\tilde{q}-x|/|x| < 1$. This implies that the optimal cost in the problem (43) is less than 1. To prove the converse, suppose on the contrary that $0 \in \mathrm{co}(V)$. This means that the origin lies on the line segment between two points $x_1, x_2 \in V$. For the optimal cost to be less than 1, the optimal point q^* must belong to the open ball $\{q \in \mathbb{R}^n : |q - x_1| < |x_1|\}$ as well as to the open ball $\{q \in \mathbb{R}^n : |q-x_2| < |x_2|\}$. But the intersection between these two sets is empty, which is a contradiction. □

We shall henceforth assume that the set V is separated from the origin by a hyperplane. For N sufficiently large, the initial quantization points can be chosen in such a way that each of the resulting Voronoi regions indeed has this property. Since by Lemma 7 the Lloyd algorithm does not increase the cost, Lemma 8 implies that all Voronoi regions will then have this property at every step of the iteration. From Lemmas 5 and 6 we know that the problem (43) is quasiconvex and can thus be handled by iterative convex optimization algorithms; see the discussion in Remark 6(iii).

In what follows we investigate the structure of the problem (43) in order to obtain a solution more constructively. Let us first present an equivalent formulation of this optimization problem.

Lemma 9 *Let* V *be a polytope separated from the origin by a hyperplane. Consider the problem of finding the sphere with center* c *and radius* r *which encloses* V *and minimizes* $r/|c|$*. Let* (c^*, r^*) *be the parameters of the optimal sphere. Then the optimal value for the problem* (43) *is* $\gamma^* := r^*/|c^*|$ *and the optimal point is* $q^* := \big(1 - (\gamma^*)^2\big)c^*$.

Proof: Let $f_q(x) := |q - x|/|x|$. In the problem (43), we search for q that minimizes the value of the function f_q on its smallest level set enclosing V (we will see shortly that this is well defined). For each $\gamma > 0$, the γ-level set of f_q is described by

$$|q - x|^2 - \gamma^2|x|^2 = 0.$$

Because V is separated from the origin by a hyperplane, we know from Lemma 8 that the optimal value of γ is smaller than 1. Thus from here on we will only be interested in $\gamma < 1$. A square completion argument leads to

$$\frac{1}{1-\gamma^2}\left(|q - x|^2 - \gamma^2|x|^2\right) = \Big|x - \frac{q}{1-\gamma^2}\Big|^2 - \frac{\gamma^2}{(1-\gamma^2)^2}|q|^2,$$

so that the γ-level set of f_q is the sphere $|x - c|^2 = r^2$, with center $c := q/(1-\gamma^2)$ and radius $r := \gamma|c|$. In the new variables (c, r), we must minimize $\gamma = r/|c|$ among all spheres enclosing V. □

Note that the point q^* belongs to $\mathrm{co}(V)$ by Lemma 6, while c^* might not. Lemma 9 leads us to considering the problem

$$\begin{aligned} \min_{c\in\mathbb{R}^n, r\in\mathbb{R}} \quad & \gamma^2(c, r) := \frac{r^2}{|c|^2} \\ \text{subject to} \quad & |c - v_i|^2 \le r^2, \quad i \in \{1, \dots, p\}, \end{aligned} \tag{44}$$

where $v_1, \dots, v_p$ are the vertices of the polytope V. This is an optimization problem subject to inequality constraints, which can be solved with a finite number of computations. The idea is to enumerate active constraints, according to the procedure described in the following algorithm:

1: **for** all subsets S of the set of vertices of V **do**
2: compute the (c_S, r_S)-sphere minimizing γ^2 among all (c, r)-spheres touching all points in S
3: **end for**
4: discard all (c_S, r_S)-spheres not containing all vertices of V
5: find global minima for (44) by comparing the values of $r_S^2/|c_S|^2$ among all remaining candidate spheres

Steps 4 and 5 are straightforward comparison checks. Regarding step 1, it turns out we can restrict our search to sets S containing at least two vertices of V, by virtue of the following result.

Lemma 10 *The optimal sphere for the problem* (44) *touches at least two vertices of* V*, i.e., at least two constraints are active at the minimum.*

Proof: The proof is by contradiction. Suppose that the optimal sphere touches only one vertex. We denote this vertex by v and assume, performing an affine coordinate change, that it has coordinates $(1, 0, \dots, 0)^T$. Let $c = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)^T$. Then we are led to minimizing

$$\gamma^2(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n) = \frac{(\bar{x}_1 - 1)^2 + \sum_{i=2}^n \bar{x}_i^2}{\sum_{i=1}^n \bar{x}_i^2} = 1 + \frac{1 - 2\bar{x}_1}{\sum_{i=1}^n \bar{x}_i^2}. \tag{45}$$

Let us show that this function has no critical points besides the pole at the origin and the zero at v. We have

$$\frac{\partial\gamma^2}{\partial\bar{x}_1} = \frac{2\left(\bar{x}_1^2 - \sum_{i=2}^n \bar{x}_i^2 - \bar{x}_1\right)}{\left(\sum_{i=1}^n \bar{x}_i^2\right)^2} \tag{46}$$

and

$$\frac{\partial\gamma^2}{\partial\bar{x}_i} = \frac{-2\bar{x}_i(1 - 2\bar{x}_1)}{\left(\sum_{i=1}^n \bar{x}_i^2\right)^2}, \qquad i \neq 1. \tag{47}$$

In view of the formula (47), every critical point satisfies either $\bar{x}_1 = 1/2$ or $\bar{x}_i = 0$ for all $i \neq 1$. In the first case, the formula (46) implies that we must have $\sum_{i=2}^n \bar{x}_i^2 = -1/2$, and this equation has no solution. In the second case, (46) gives two solutions: $\bar{x}_1 = 0$ (pole at 0) and $\bar{x}_1 = 1$ (zero at v). The pole at the origin is not a minimum. The zero at v corresponds to the sphere of radius 0 centered at v, which is not a feasible solution because it does not enclose V. In summary, we have shown that the optimal sphere cannot touch only a single vertex of V. □

Regarding step 2, we need to minimize γ^2 over spheres passing through two or more vertices of V. Spheres passing through l generic points in $\mathbb{R}^n$ are parameterized by $n +$

$1-l$ variables. A convenient parameterization is obtained by intersecting hyperplanes of points equidistant from pairs of points from a given set. Coordinates of the points on the intersection are given by affine functions of $n+1-l$ free parameters. Note that the radius r of the sphere is uniquely determined by its center c and the vertices of V which lie on the sphere. It is straightforward to verify that the function γ^2 in (44) is a rational function whose numerator and denominator are quadratic inhomogeneous polynomials in these free parameters, and that critical points of γ^2 are solutions of $n+1-l$ quadratic equations in the same number of unknowns. According to Bezout's theorem, this generically gives 2^{n+1-l} candidate optimal spheres (see [4]). Step 2 is completed by choosing the one with the smallest radius. We emphasize that while this constructive solution has exponential complexity, a more efficient solver can be developed based on quasiconvex programming; see Remark 6(iii).

As an example of step 2, let us work out the planar case. When $n=2$, the problem reduces to finding critical points of γ for circles passing through l vertices of V, where $l>1$ by Lemma 8. Since for $l>2$ there is at most one circle passing through the corresponding vertices, we only need to explain how to solve this problem for $l=2$. For convenience, let us consider an affine change of coordinates which places the two vertices at $(1,0)^T$ and $(-1,0)^T$ and the origin at some point $(x_0,y_0)^T$. Without loss of generality, assume that $y_0\geq 0$. The center of the circle is denoted by $c=(\bar{x},\bar{y})^T$. We know that c must be equidistant from the two vertices, hence $\bar{x}=0$. Then we have

$$\gamma^2=\frac{1+\bar{y}^2}{x_0^2+(\bar{y}-y_0)^2}$$

and so

$$\begin{aligned}\frac{\partial\gamma^2}{\partial\bar{y}}&=\frac{2\bar{y}\big(x_0^2+(\bar{y}-y_0)^2\big)-2(\bar{y}-y_0)(1+\bar{y}^2)}{\big(x_0^2+(\bar{y}-y_0)^2\big)^2}\\&=\frac{-2\bar{y}^2y_0+2\bar{y}(x_0^2+y_0^2-1)+2y_0}{\big(x_0^2+(\bar{y}-y_0)^2\big)^2}.\end{aligned}$$

Equating the numerator to 0, we arrive at the equation

$$-\bar{y}^2y_0+\bar{y}(x_0^2+y_0^2-1)+y_0=0.$$

In the special case when $y_0=0$, this reduces to $\bar{y}(x_0^2-1)=0$. Since $x_0=\pm1$ corresponds to one of the vertices being at the origin, which cannot happen by our earlier assumption, the solution is $\bar{y}=0$ (as is also clear from symmetry). When $y_0\neq 0$, the minimum is achieved at

$$\bar{y}=\frac{x_0^2+y_0^2-1-\sqrt{(x_0^2+y_0^2-1)^2+4y_0^2}}{2y_0}<0.$$

(Note that this goes to 0 as y_0 approaches 0 or ∞.)

IV. Simulation results

For our simulation studies, we take the system (9) and work with the following data:

$$A=\begin{pmatrix}0 & 1\\ \frac{1}{2} & \frac{1}{2}\end{pmatrix},\qquad B=\begin{pmatrix}1\\1\end{pmatrix},\qquad K=\begin{pmatrix}-\frac{1}{2} & -1\end{pmatrix}.$$

This leads to $P=I$ from (10), which is convenient because the regions $\mathcal{R}_1$ and $\mathcal{R}_2$ will be balls around the origin. We have $\|PBK\|=\sqrt{5/2}$. In all simulations, we take $M=5$, so the outer ball shown in the following figures is $\mathcal{R}_1=\mathcal{B}_5$. We also fix the number N of quantization points to be 25.

We first consider a simple quantizer which divides the square $[-5,5]\times[-5,5]$ into 25 equal squares as shown in Figure 1(a). The worst-case quantization error is $\Delta=\sqrt{2}$ here, and the inequality (14) is satisfied for sufficiently small ε (in fact, $N=25$ is the smallest perfect square for which this is the case). Solutions of the quantized closed-loop system (11) are shown in Figure 1(b). Lemma 2 predicts that asymptotically, these solutions must approach the ball of radius $2\sqrt{5}\approx 4.472$ around the origin (this is obtained by setting $\varepsilon=0$ in the formula (13) for $\mathcal{R}_2$). The circle of this radius is drawn in the figure, and we see that solutions in fact enter it in finite time and then continue to approach the origin (we took $T=2$ in all simulations); this illustrates the conservativeness of the theory presented in Section II. We also clearly see chattering behavior on the boundaries between quantization regions.

Next, we design the quantizer by solving the multicenter problem on $\mathcal{B}_5$ using the Lloyd algorithm. Figure 2(a) illustrates the evolution of the quantization points starting from random initial conditions close to the origin, and Figure 2(b) depicts the quantization points obtained after 100 iterations and the corresponding Voronoi regions, predominantly of hexagonal shapes. The worst-case quantization error is $\Delta\approx 1.248$ in this case, which is smaller than the one for the rectangular quantizer. From Lemma 2, the ultimate bound on the norm of closed-loop solutions is approximately 3.948. We see from Figure 2(c) that solutions indeed enter the ball of this radius (again in finite time).

We now consider radial and spherical quantization. Let $N_1=1$ and $N_2=25$. This means that we must uniformly distribute 25 points on the unit circle, and the cordal quantization error Δ_s defined by (23) is computed from the formula $\Delta_s=\sqrt{2-2\cos(\pi/25)}\approx 0.126$. The inequality (28) is satisfied for sufficiently small ε because $1/(2\|PBK\|)\approx 0.316$. Since we are interested in an ultimate bound on solutions, we take $\varepsilon=0$. This gives $a\approx 0.809$, $b\approx 1.191$, and the required distance from the quantization points to the origin is determined from the formula (25) and approximately equals 4.045; see Figure 3(a). In view of Lemma 3, closed-loop trajectories must asymptotically approach a ball whose radius is now decreased to about 3.399. This indeed happens, as illustrated in Figure 3(b). Interestingly, $N_1=2$ gives a poorer guaranteed convergence radius (even if we take $N_2=13$), and it is easy to check that larger values of N_1 are not feasible because of (28).

Finally, we consider a quantizer that results from solving the radially weighted multicenter problem using the Lloyd algorithm. In this case we have to pick a desired ultimate bound on $|x(t)|$, which we take to be $m=3$. Figure 4(a) illustrates the evolution of the quantization points starting from random initial conditions in the annulus $\mathcal{B}_5\setminus\mathcal{B}_3$, and Figure 4(b) depicts the quantization points obtained after 100 iterations and the associated Voronoi regions (computed for the entire ball $\mathcal{B}_5$). The quantization points remain in the convex

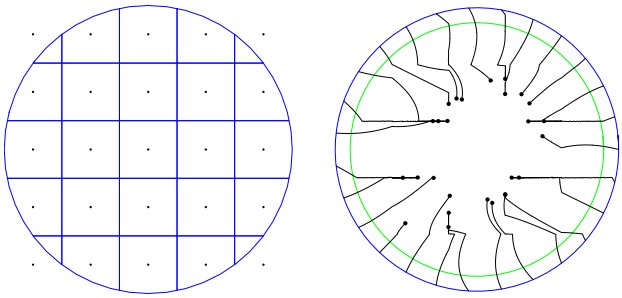

Fig. 1. Rectangular quantizer: (a) Quantization regions and points, (b) Closed-loop trajectories.

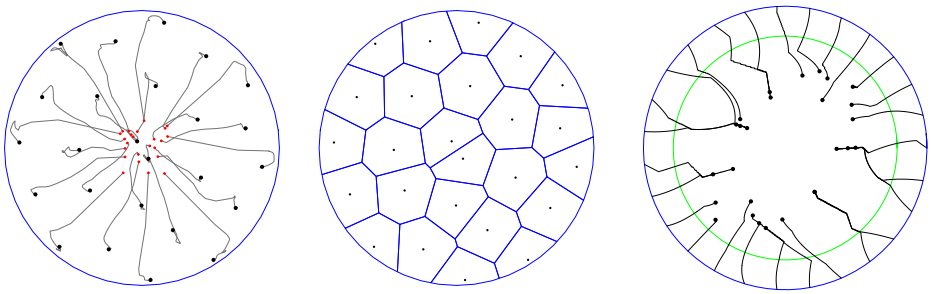

Fig. 2. Uniform quantizer: (a) Lloyd iterations, (b) Quantization regions and points, (c) Closed-loop trajectories.

hull of the intersections of the corresponding Voronoi regions with the annulus (Lemma 6). We have $\Delta_{rw} \approx 0.273$, hence the inequality (37) is satisfied for sufficiently small ε. Thus Lemma 4 guarantees that $|x(t)| \leq 3$ as $t \to \infty$, which is confirmed by Figure 4(c).

V. Conclusions

We discussed the problem of designing the most suitable quantizer for feedback stabilization subject to a given information constraint. We showed how a perturbation analysis based on a Lyapunov function can be used to define a destabilization measure of a quantizer and arrive at an ultimate bound on closed-loop solutions in several different ways. In each case, we demonstrated how the problem of minimizing this ultimate bound can be naturally cast as a weighted multicenter problem and can be tackled via quasiconvex programming and Voronoi diagram computations. We investigated uniform, radial and spherical, and radially weighted quantizer designs, developing a novel iterative solver for the latter. We compared these quantizers among themselves and to a standard rectangular quantizer in the context of a simple simulation example.

The general problem of designing an optimal quantizer for stabilization depends on many parameters, and the analysis presented here is quite conservative and not comprehensive. An important direction for future work is to address the issue of optimizing the design parameters and to move beyond stability by addressing specific performance criteria. (See [10] for some related recent work on discrete-time systems.) Some measure of "representation complexity" of a quantizer needs to be developed and taken into account in quantizer design. Topics for future work also include studying convergence properties of the Lloyd algorithm in nonsmooth settings and exploring the least destabilizing quantizer design for classes of nonlinear systems.

References

[1] F. Aurenhammer. Voronoi diagrams: A survey of a fundamental geometric data structure. *ACM Computing Surveys*, 23(3):345–405, 1991.

[2] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, UK, 2004.

[3] R. W. Brockett and D. Liberzon. Quantized feedback stabilization of linear systems. *IEEE Transactions on Automatic Control*, 45(7):1279–1289, 2000.

[4] J. L. Coolidge. *A Treatise on Algebraic Plane Curves*. Dover Publications, New York, 1959.

[5] J. Cortés and F. Bullo. Coordination and geometric optimization via distributed dynamical systems. *SIAM Journal on Control and Optimization*, June 2004. To appear.

[6] M. de Berg, M. van Kreveld, M. Overmars, and O. Schwarzkopf. *Computational Geometry: Algorithms and Applications*. Springer Verlag, New York, 2 edition, 2000.

[7] Z. Drezner, editor. *Facility Location: A Survey of Applications and Methods*. Springer Series in Operations Research. Springer Verlag, New York, 1995.

[8] Q. Du, V. Faber, and M. Gunzburger. Centroidal Voronoi tessellations: Applications and algorithms. *SIAM Review*, 41(4):637–676, 1999.

[9] N. Elia and S. K. Mitter. Stabilization of linear systems with limited information. *IEEE Transactions on Automatic Control*, 46(9):1384–1400, 2001.

[10] F. Fagnani and S. Zampieri. Performance evaluations of quantized stabilizers. In *IEEE Conf. on Decision and Control*, pages 1897–1901, Maui, HI, December 2003.

[11] A. F. Filippov. *Differential Equations with Discontinuous Righthand Sides*, volume 18 of *Mathematics and Its Applications*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1988.

[12] A. Gersho. Asymptotically optimal block quantization. *IEEE Transactions on Information Theory*, 25(7):373–380, 1979.

[13] R. M. Gray and D. L. Neuhoff. Quantization. *IEEE Transactions on Information Theory*, 44(6):2325–2383, 1998. Commemorative Issue 1948-1998.

[14] H. Ishii and B. A. Francis. Stabilizing a linear system by switching control with dwell time. *IEEE Transactions on Automatic Control*, 47(12):1962–1973, 2002.

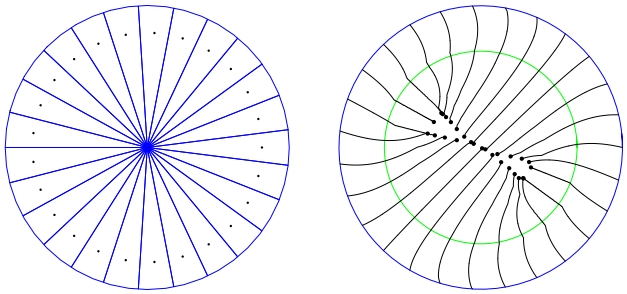

Fig. 3. Radial/spherical quantizer: (a) Quantization regions and points, (b) Closed-loop trajectories.

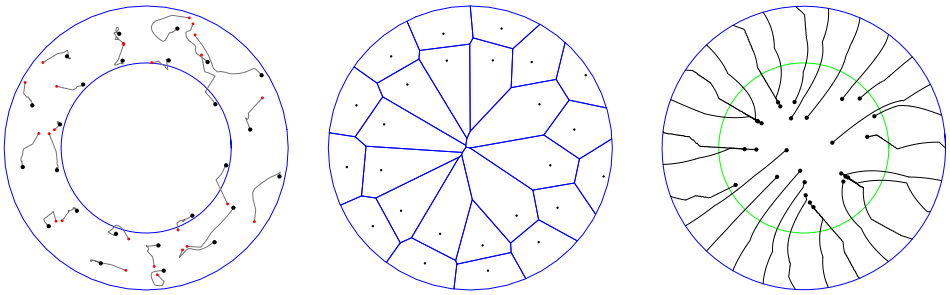

Fig. 4. Radially weighted quantizer: (a) Lloyd iterations, (b) Quantization regions and points, (c) Closed-loop trajectories.

[15] D. Liberzon. Nonlinear stabilization by hybrid quantized feedback. In N. A. Lynch and B. H. Krogh, editors, *Proc. Third International Workshop on Hybrid Systems: Computation and Control*, volume 1790 of *Lecture Notes in Computer Science*, pages 243–257. Springer, 2000.

[16] D. Liberzon. Hybrid feedback stabilization of systems with quantized signals. *Automatica*, 39(9):1543–1554, 2003.

[17] D. Liberzon. *Switching in Systems and Control*. Birkhäuser, Boston, MA, 2003.

[18] S. P. Lloyd. Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2):129–137, 1982. Presented as Bell Laboratory Technical Memorandum at a 1957 Institute for Mathematical Statistics meeting.

[19] J. Lunze, B. Nixdorf, and J. Schröder. Deterministic discrete-event representations of linear continuous-variable systems. *Automatica*, 35(3):395–406, 1999.

[20] J. Matousek. *Geometric Discrepancy: An Illustrated Guide*. Number 18 in Algorithms and Combinatorics. Springer Verlag, New York, 1999.

[21] G. N. Nair and R. J. Evans. Exponential stabilisability of finite-dimensional linear systems with limited data rates. *Automatica*, 39(4):585–593, 2003.

[22] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*. Number 63 in CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial & Applied Mathematics, 1992.

[23] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. Wiley Series in Probability and Statistics. John Wiley, New York, 2 edition, 2000.

[24] A. Okabe and A. Suzuki. Locational optimization problems solved through Voronoi diagrams. *European Journal of Operational Research*, 98(3):445–56, 1997.

[25] E. D. Sontag. Smooth stabilization implies coprime factorization. *IEEE Transactions on Automatic Control*, 34(4):435–443, 1989.

[26] E. D. Sontag and Y. Wang. On characterizations of the input-to-state stability property. *Systems & Control Letters*, 24(5):351–359, 1995.

[27] A. Suzuki and Z. Drezner. The p-center location problem in an area. *Location Science*, 4(1/2):69–82, 1996.

[28] A. Suzuki and A. Okabe. Using Voronoi diagrams. In Drezner [7], pages 103–118.

[29] P. F. Swaszek and J. B. Thomas. Multidimensional spherical coordinates quantization. *IEEE Transactions on Information Theory*, 29(4):570–6, 1983.