

An Investigation on the Optimal Sampling Pattern for Linear Control Systems

Enrico Bini, Giuseppe Buttazzo

Abstract

In digital control systems the state is sampled at given sampling instants and the input is kept constant between two consecutive instants. By *optimal sampling problem* we mean the selection of sampling instants and control inputs, such that a given function of the state and input is minimized.

In this paper we formulate the optimal sampling problem and we propose a numerical method to solve it. Being the solving procedure time consuming, we also propose a new *quantization-based* sampling strategy that is computationally tractable and capable to achieve a near-optimal cost. Finally, we prove that the quantization-based sampling is optimal in scalar systems for large number of samples.

I. INTRODUCTION

We consider optimal LQ control problems of the form

$$\min_{\bar{u}} \left\{ \int_0^T (x'Qx + \bar{u}'R\bar{u}) dt + x(T)'Sx(T) : \dot{x} = Ax + B\bar{u}, x(0) = x_0 \right\}, \quad (1)$$

where x and \bar{u} are the state and input signals (moving over \mathbb{R}^n and \mathbb{R}^m , resp.), $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $Q \in \mathbb{R}^{n \times n}$, $R \in \mathbb{R}^{m \times m}$, $S \in \mathbb{R}^{n \times n}$ are matrices, with Q and S positive semi-definite, R positive definite (to denote the transpose of any matrix M we use the compact Matlab-like notation M'). The control input signal \bar{u} is constrained to be piecewise constant:

$$\bar{u}(t) = u_k \quad \forall t \in [t_{k-1}, t_k) \quad (2)$$

E. Bini is with Lund University, Sweden. G. Buttazzo is with University of Pisa, Italy. Email: `bini@control.lth.se`, `buttazzo@dm.unipi.it`

This work has been partially supported by the project Marie-Curie **ad exact citation** and by the project 2008K7Z249 “Trasporto ottimo di massa, disuguaglianze geometriche e funzionali e applicazioni” funded by the Italian Ministry of Research and University.

with $0 = t_0 < t_1 < \dots < t_N = T$. The sequence $\{t_0, t_1, \dots, t_{N-1}, t_N\}$ is called *sampling pattern*, while t_k are called *sampling instants*. Often, we represent a sampling pattern by the values that separates two consecutive instants that are called *interarrivals* τ_k . The sampling instants and the interarrivals are in one-to-one correspondence thanks to the relations

$$\begin{cases} t_0 = 0 \\ t_k = \sum_{i=0}^{k-1} \tau_i \quad k \geq 1, \end{cases} \quad \tau_k = t_{k+1} - t_k.$$

In periodic sampling we have $\tau_k = \tau$ for all k , with $\tau = \frac{T}{n}$ the period of the sampling.

In our formulation, we intentionally ignore disturbances to the system. While accounting for disturbances would certainly make the problem more adherent to the reality, it would also prevent us from deriving the analytical results that we propose in this paper. Moreover we believe that the ideas in this paper could certainly influence the development of the extension to the case with noise.

In continuous-time systems, the optimal u that minimizes the cost of (1) can be found by solving the Riccati differential equation

$$\begin{cases} \dot{K} = KBR^{-1}B'K - A'K - KA - Q \\ K(T) = S \end{cases} \quad (3)$$

and then setting the input u as

$$u(t) = -R^{-1}B'K(t)x(t). \quad (4)$$

In this case, the achieved cost is

$$J_\infty = x_0'K(0)x_0. \quad (5)$$

For *given* sampling instants, the *optimal* values u_k of the input that minimize the cost (1) can

be analytically determined through the classical discretization process described below. If we set

$$\Phi(\tau) = e^{A\tau}, \quad \bar{A}_k = \Phi(\tau_k), \quad (6)$$

$$\Gamma(\tau) = \int_0^\tau e^{A(\tau-t)} dt B, \quad \bar{B}_k = \Gamma(\tau_k), \quad (7)$$

$$\bar{Q}(\tau) = \int_0^\tau \Phi'(t) Q \Phi(t) dt, \quad \bar{Q}_k = \bar{Q}(\tau_k), \quad (8)$$

$$\bar{R}(\tau) = \tau R + \int_0^\tau \Gamma'(t) Q \Gamma(t) dt, \quad \bar{R}_k = \bar{R}(\tau_k), \quad (9)$$

$$\bar{P}(\tau) = \int_0^\tau \Phi(t)' Q \Gamma(t) dt, \quad \bar{P}_k = \bar{P}(\tau_k), \quad (10)$$

then the problem of minimizing the cost (1) can be written as a discrete time-variant problem

$$\begin{cases} x_{k+1} = \bar{A}_k x_k + \bar{B}_k u_k \\ \text{given } x_0 \end{cases}$$

with the cost

$$J = x'_N S x_N + \sum_{k=0}^{N-1} (x'_k \bar{Q}_k x_k + u'_k \bar{R}_k u_k + 2x'_k \bar{P}_k u_k).$$

This problem is then solved using dynamic programming [5], [11]. The solution requires the backward recursive definition of the sequence of matrices \bar{K}_k

$$\begin{cases} \bar{K}_N = S \\ \bar{K}_k = \hat{Q}_k - \hat{B}_k \hat{R}_k^{-1} \hat{B}'_k, \end{cases} \quad (11)$$

with \hat{Q}_k , \hat{R}_k , and \hat{B}_k , functions of \bar{K}_{k+1} as well, defined by

$$\hat{Q}_k = \bar{Q}_k + \bar{A}'_k \bar{K}_{k+1} \bar{A}_k, \quad \hat{R}_k = \bar{R}_k + \bar{B}'_k \bar{K}_{k+1} \bar{B}_k, \quad \hat{B}_k = \bar{P}_k + \bar{A}'_k \bar{K}_{k+1} \bar{B}_k. \quad (12)$$

Then, the optimal input sequence u_k is then determined by

$$u_k = -\hat{R}_k^{-1} \hat{B}'_k x_k, \quad (13)$$

with minimal cost equal to

$$J = x'_0 \bar{K}_0 x_0. \quad (14)$$

Equation (13) allows to compute the optimal input signal u_k for *given sampling instants* t_0, t_1, \dots, t_N . In fact, the optimal input sequence depends on \bar{A}_k , \bar{B}_k , \bar{Q}_k , \bar{R}_k , \bar{P}_k , \hat{Q}_k , \hat{R}_k , and

\hat{B}_k which are all function of the inter-sample separations $\tau_k = t_{k+1} - t_k$. However, to our best knowledge, the problem of determining the optimal sampling pattern is still open.

In Sections II-A and II-B we recall some commonly used sampling techniques: periodic and Lebesgue sampling. In Section III we formulate the problem of optimal sampling and we report some results. Being the optimal sampling time consuming, in Section IV we propose a new sampling method that we call *quantization-based sampling* being related to quantization theory. In Section V we demonstrate that quantization-based sampling is optimal in the scalar case. Finally, in Section VI we evaluate all presented sampling methods.

A. Related works

As computational resources are more and more constrained, much research efforts have been recently dedicated to the investigation of control techniques that are capable of sampling less frequently with an unaltered control cost. Triggering the activation of controllers by *events*, rather than by time, is an attempt to reduce the consumed computational resources. A first example of event-based controller was proposed by Årzén [2]. Xu and Cao [14] proposed a method to optimally design a control law that selects among a finite set of control inputs. The input is applied when the (scalar) state reaches a threshold. The number of thresholds is finite.

In self-triggered controller [12], the control task determines the next instant when it will be activated. Wang and Lemmon addressed self-triggered linear \mathcal{H}_∞ controllers [13]. Self-triggered controllers have also been analyzed and proved stable also for state-dependent homogeneous systems and polynomial systems [1].

A problem related to the one considered here was addressed by Kowalska and von Mohrenschildt [7] who proposed the variable time control (VTC). Similarly to our approach, they also perform the cost minimization over the sampling instants as well. However, the authors perform a linearization of the discrete-time system in a neighborhood of every sampling instant, losing then optimality.

The contributions of this paper are:

- the formulation of the optimal sampling problem and its numerical implementation through a gradient descent algorithm;
- the introduction of the *quantization-based sampling*, which is capable to provide a cost very close to the optimal one with a small computational effort;

- the proof that quantization-based sampling is optimal when the state is scalar and the number of samples is large.

II. SAMPLING METHODS

The temporal distribution of the sampling instants of any given sampling method is evaluated with the sampling density, whereas the capacity to reduce the cost is determined by the normalized cost. Both metrics are formally defined below.

Definition 1: Given a problem, specified by x_0 , A , B , Q , R , and S , an interval length T , and a number of samples N , we define the *sampling density* $\sigma_{m,N} : [0, T] \rightarrow \mathbb{R}^+$ of any sampling method m as

$$\sigma_{m,N}(t) = \frac{1}{N\tau_k} \quad \forall t \in [t_k, t_{k+1}).$$

Notice that the sampling density is normalized since:

$$\int_0^T \sigma_{m,N}(t) dt = \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} \sigma_{m,N}(t) dt = \sum_{k=0}^{N-1} \frac{1}{N\tau_k} \tau_k = 1.$$

To remove the dependency on N we also define the following density.

Definition 2: Given a problem, specified by x_0 , A , B , Q , R , and S , and an interval length T , we define the *asymptotic sampling density* $\sigma_m : [0, T] \rightarrow \mathbb{R}^+$ of any sampling method m as

$$\sigma_m(t) = \lim_{N \rightarrow \infty} \sigma_{m,N}(t). \quad (15)$$

The density σ_m provides significant information only when N is large compared to the size of the interval $[0, T]$, while in reality it is often more desirable to have a small N to reduce the amount of computational resource consumed by the controller. Nonetheless we think that the asymptotic density can still provide informative results that can guide the design of more efficient sampling techniques.

While the density σ_m provides an indication of how samples are distributed over time, the quantity defined below returns a measure of the cost associated to any sampling method.

Definition 3: Given a problem, specified by x_0 , A , B , Q , R , S , an interval length T , a number of samples N , we define the *normalized cost* of any sampling method m as

$$c_{m,N} = \frac{N^2}{T^2} \frac{J_m - J_\infty}{J_\infty} \quad (16)$$

where J_m is the minimal cost of the sampling method m with N samples, J_∞ is the minimal cost of the continuous-time systems.

The scaling factor $\frac{N^2}{T^2}$ is motivated by the observation (proved by Melzer and Kuo [9]) that the cost $J(\tau)$ of periodic sampling with period τ can be approximated by $J_\infty + k\tau^2 + o(\tau^2)$ for small values of the period τ (the Taylor expansion in a neighborhood of $\tau \rightarrow 0$ has not first order term).

To remove the dependency on N , we can also compute the limit of the normalized cost.

Definition 4: Given a problem, specified by $x_0, A, B, Q, R,$ and S , an interval length T , we define the *asymptotic normalized cost* of any sampling method m as

$$c_m = \lim_{N \rightarrow \infty} c_{m,N}. \quad (17)$$

The asymptotic normalized cost (17) is also very convenient from an ‘‘engineering’’ point of view. In fact, it can be readily used to estimate the number of samples to achieve a bounded cost increase with respect to the continuous-time case. If for a given sampling method m we can tolerate at most a (small) factor r of cost increase w.r.t. the continuous-time optimal controller, then

$$(1 + r)J_\infty \geq J_m \approx T^2 \frac{c_{m,N} J_\infty}{N^2} + J_\infty,$$

from which we deduce

$$N \geq T \sqrt{\frac{c_{m,N}}{r}} \approx T \sqrt{\frac{c_m}{r}}, \quad (18)$$

by assuming a weak dependency of $c_{m,N}$ on N . Equation (18) constitutes a good hint for assigning the number of samples in a given interval.

We recall here below the characteristics of the periodic sampling and of the Lebesgue sampling, well known in the literature. In Section III we will then discuss the optimal sampling, while in Section IV we propose a sampling method based on quantization theory.

A. Periodic sampling

The simplest (and almost universally used) sampling method is the one obtained by dividing the interval $[0, T]$ in N intervals of equal size; it corresponds to the choice $t_k = kT/N$ and it is called *periodic sampling*. We then have that all the inter-sampling periods are equal: $\tau_k = \tau = \frac{T}{N}$ for all k , and the sampling density is, obviously, constant, with

$$\sigma_{\text{per},N}(t) = 1/T, \quad \forall N.$$

For the periodic case, it is possible to determine analytically the asymptotic normalized cost c_{per} . In 1971, Melzer and Kuo [9] approximated to the second order the solution $\bar{K}(\tau)$ of the Discrete Algebraic Riccati Equation corresponding to a sampling period τ , in a neighborhood of $\tau = 0$. They showed that

$$\bar{K}(\tau) = K_{\infty} + X \frac{\tau^2}{2} + o(\tau^2), \quad (19)$$

with K_{∞} being the solution of the ARE of the continuous-time problem (3) and X the second order derivative of $\bar{K}(\tau)$ in 0, that is the solution of the following Lyapunov equation

$$(A' - K_{\infty} B R^{-1} B') X + X (A - B R^{-1} B' K_{\infty}) + \frac{1}{6} (A' - K_{\infty} B R^{-1} B') K_{\infty} B R^{-1} B' K_{\infty} (A - B R^{-1} B' K_{\infty}) = 0. \quad (20)$$

Melzer and Kuo [9] also proved that such a solution is positive semidefinite. Hence the normalized asymptotic cost in the periodic case is

$$\begin{aligned} c_{\text{per}} &= \frac{1}{T^2} \lim_{N \rightarrow \infty} N^2 \frac{x_0' (K_{\infty} + X \frac{T^2}{2N^2} + o(1/N^2)) x_0 - x_0' K_{\infty} x_0}{x_0' K_{\infty} x_0} \\ &= \frac{x_0' X x_0}{2x_0' K_{\infty} x_0}. \end{aligned} \quad (21)$$

In the case of a scalar system ($n = 1$), assuming without loss of generality $B = R = 1$, the Lyapunov equation (20) has solution $X = \frac{1}{12} (K_{\infty} - A) K_{\infty}^2$ and the ARE has solution $K_{\infty} = A + \sqrt{A^2 + Q}$. Hence the asymptotic cost becomes

$$c_{\text{per}} = \frac{1}{24} (A \sqrt{A^2 + Q} + A^2 + Q). \quad (22)$$

B. Lebesgue sampling

As shown by Åström and Bernhardsson [3], sampling only when the state varies beyond a certain threshold, can provide significant benefits in terms of cost reduction, compared to the periodic sampling. This method was called *Lebesgue sampling*. The most intuitive way to translate their method (that applies to scalar systems with disturbances) into our context, is to choose the sampling instants t_k such that:

$$\forall k = 0, \dots, N-1 \quad \int_{t_k}^{t_{k+1}} |\dot{u}(t)| dt = \frac{1}{N} \int_{t_0}^T |\dot{u}(t)| dt, \quad (23)$$

where u is the optimal continuous-time control input. Following this sampling rule, $|u(t_k) - u(t_{k-1})|$ is constant for all k (except when \dot{u} changes its sign in (t_k, t_{k+1})), and, by construction, the asymptotic density is

$$\sigma_{\text{leb}}(t) = \frac{|\dot{u}(t)|}{\int_0^T |\dot{u}(s)| ds}.$$

After the sampling instants t_1, t_2, \dots, t_{N-1} are determined according to Eq. (23), the values of the control input u_k are optimally assigned according to (13).

For the Lebesgue sampling we are unable to determine the normalized cost c_{leb} , in general. However, in Section V we analytically compute c_{leb} for scalar systems ($n = 1$).

III. OPTIMAL SAMPLING

We now investigate the optimal solution of the problem (1). Let us introduce a notation that is useful in the context of this section. For any vector $x \in \mathbb{R}^n$, let us denote by \mathbf{x} the following vector in \mathbb{R}^p , with $p = \frac{n(n+1)}{2}$,

$$\mathbf{x} = [x_1^2, 2x_1x_2, \dots, 2x_1x_n, x_2^2, 2x_2x_3, \dots, 2x_2x_n, \dots, x_{n-1}^2, 2x_{n-1}x_n, x_n^2]',$$

and for any matrix $M \in \mathbb{R}^{n \times n}$, let us denote by $\mathbf{M} \in \mathbb{R}^p$

$$\mathbf{M} = [M_{1,1}, M_{1,2}, \dots, M_{1,n}, M_{2,2}, M_{2,3}, \dots, M_{2,n}, \dots, M_{n-1,n-1}, M_{n-1,n}, M_{n,n}]'.$$

This notation allows writing the cost (14) as $J = \mathbf{x}'_0 \bar{\mathbf{K}}_0$, and the Riccati recursive equation (11), as

$$\begin{cases} \bar{\mathbf{K}}_N = \mathbf{S} \\ \bar{\mathbf{K}}_k = r(\tau_k, \bar{\mathbf{K}}_{k+1}), \end{cases} \quad (24)$$

with $r : \mathbb{R} \times \mathbb{R}^p \rightarrow \mathbb{R}^p$ properly defined from (11).

Since we search for stationary point of J , let us investigate the partial derivatives $\frac{\partial \bar{\mathbf{K}}_k}{\partial \tau_h}$. Firstly

$$h < k \quad \Rightarrow \quad \frac{\partial \bar{\mathbf{K}}_k}{\partial \tau_h} = 0,$$

because $\bar{\mathbf{K}}_k$ depends only on the current and the future sampling intervals $\{\tau_k, \tau_{k+1}, \dots, \tau_{N-1}\}$.

Then we have

$$\begin{cases} \frac{\partial \bar{\mathbf{K}}_k}{\partial \tau_k} = \frac{\partial r}{\partial \tau}(\tau_k, \bar{\mathbf{K}}_{k+1}) \\ \frac{\partial \bar{\mathbf{K}}_k}{\partial \tau_h} = \frac{\partial r}{\partial \bar{\mathbf{K}}}(\tau_k, \bar{\mathbf{K}}_{k+1}) \frac{\partial \bar{\mathbf{K}}_{k+1}}{\partial \tau_h} \quad h > k \end{cases}$$

from which it follows

$$\frac{\partial \bar{\mathbf{K}}_k}{\partial \tau_h} = \left[\prod_{i=k}^{h-1} \frac{\partial r}{\partial \bar{\mathbf{K}}}(\tau_i, \bar{\mathbf{K}}_{i+1}) \right] \frac{\partial r}{\partial \tau}(\tau_h, \bar{\mathbf{K}}_{h+1}) \quad h \geq k \quad (25)$$

Notice that $\frac{\partial r}{\partial \bar{\mathbf{K}}}(\tau_i, \bar{\mathbf{K}}_{i+1}) \in \mathbb{R}^{p \times p}$.

Since the problem (1) is constrained by $\sum_{k=0}^{N-1} \tau_k = T$, from the KKT conditions it follows that at the optimal point the gradient ∇J must be proportional to $[1, 1, \dots, 1]$, meaning that all components of ∇J have to be equal to each other. Hence, a necessary condition for the optimum is that

$$\forall h = 0, \dots, N-2, \quad \frac{\partial J}{\partial \tau_h} = \frac{\partial J}{\partial \tau_{h+1}} \Leftrightarrow \mathbf{x}'_0 \frac{\partial \bar{\mathbf{K}}_0}{\partial \tau_h} = \mathbf{x}'_0 \frac{\partial \bar{\mathbf{K}}_0}{\partial \tau_{h+1}},$$

which can be rewritten as

$$\forall h = 0, \dots, N-2,$$

$$\mathbf{x}'_0 \left[\prod_{i=0}^{h-1} \frac{\partial r}{\partial \bar{\mathbf{K}}}(\tau_i, \bar{\mathbf{K}}_{i+1}) \right] \left[\frac{\partial r}{\partial \tau}(\tau_h, \bar{\mathbf{K}}_{h+1}) - \frac{\partial r}{\partial \bar{\mathbf{K}}}(\tau_h, \bar{\mathbf{K}}_{h+1}) \frac{\partial r}{\partial \tau}(\tau_{h+1}, \bar{\mathbf{K}}_{h+2}) \right] = 0. \quad (26)$$

Finding the analytical solution of (26) is an overwhelming task. Hence below we propose some special cases that provide some insights on how the general solution should be. In Section III-C we describe a numerical algorithm to find the solution.

A. Two sampling instants: optimality of periodic sampling

If $N = 2$, then (26) must be verified only for $h = 0$. In this special case condition (26) becomes

$$\mathbf{x}'_0 \left[\frac{\partial r}{\partial \tau}(T - \tau_1, r(\tau_1, \mathbf{S})) - \frac{\partial r}{\partial \bar{\mathbf{K}}}(T - \tau_1, r(\tau_1, \mathbf{S})) \frac{\partial r}{\partial \tau}(\tau_1, \mathbf{S}) \right] = 0 \quad (27)$$

in which the only unknown is τ_1 .

Eq. (27) also allows checking whether periodic sampling can be optimal or not. Let $\bar{\mathbf{K}}_\tau$ be the solution of the DARE associated to the discretized system with period τ . Then a necessary condition for the optimality of the periodic sampling with period τ is

$$\mathbf{x}'_0 \left[I_p - \frac{\partial r}{\partial \bar{\mathbf{K}}}(\tau, \bar{\mathbf{K}}_\tau) \right] \frac{\partial r}{\partial \tau}(\tau, \bar{\mathbf{K}}_\tau) = 0 \quad (28)$$

where I_p denotes the identity matrix in $\mathbb{R}^{p \times p}$. If (28) is false, then we are certain that when the system state is at x_0 , periodic sampling with period τ is not best.

B. Scalar case

In the case of a scalar state ($n = 1$ and then $p = 1$) we can avoid using the **bold face** notation introduced at the beginning of Section III, since both M and \mathbf{M} are the same scalar value.

The necessary condition for optimality (26) requires to compute the Riccati recurrence function $r(\tau, \bar{K})$ and its partial derivatives. In the scalar case, the Riccati recurrence function r is

$$r(\tau, \bar{K}) = \frac{\bar{Q}_k \bar{R}_k - \bar{P}_k^2 + (\bar{A}_k^2 \bar{R}_k - 2\bar{A}_k \bar{B}_k \bar{P}_k + \bar{B}_k^2 \bar{Q}_k) \bar{K}}{\bar{R}_k + \bar{B}_k^2 \bar{K}} \quad (29)$$

with the following partial derivatives

$$\begin{aligned} \frac{\partial r}{\partial K}(\tau, \bar{K}) &= \frac{\bar{A}_k^2 \bar{R}_k^2 - 2\bar{A}_k \bar{B}_k \bar{R}_k \bar{P}_k + \bar{B}_k^2 \bar{P}_k^2}{(\bar{R}_k + \bar{B}_k^2 \bar{K})^2} = \frac{(\bar{A}_k \bar{R}_k - \bar{B}_k \bar{P}_k)^2}{(\bar{R}_k + \bar{B}_k^2 \bar{K})^2} \\ \frac{\partial r}{\partial \tau}(\tau, \bar{K}) &= \frac{1}{(\bar{R}_k + \bar{B}_k^2 \bar{K})^2} \left(R(\bar{P}_k + \bar{A}_k \bar{B}_k \bar{K})^2 \right. \\ &\quad \left. + (\bar{A}_k \bar{R}_k - \bar{B}_k \bar{P}_k)(Q(\bar{A}_k \bar{R}_k - \bar{B}_k \bar{P}_k) + 2\bar{A}_k(A\bar{R}_k - B\bar{P}_k)K + 2\bar{A}_k \bar{B}_k(A\bar{B}_k - B\bar{A}_k)\bar{K}^2) \right) \end{aligned}$$

Let us now investigate the condition on $\tau_0, \dots, \tau_{N-1}$ to satisfy (26). Since we assume $x_0 \neq 0$, we may have any of the two factors in (26) equal to zero.

First, we observe that if τ_k is such that $\bar{A}_k \bar{R}_k = \bar{B}_k \bar{P}_k$, then $\frac{\partial r}{\partial K}(\tau_k, \bar{K}) = 0$ for any possible \bar{K} . Let us set k^* as the minimum indices among the k such that $\bar{A}_k \bar{R}_k = \bar{B}_k \bar{P}_k$. From (25), it follows that $\frac{\partial \bar{K}_0}{\partial \tau_h} = 0$, for all $h \geq k^* + 1$. In fact, for such special τ_{k^*} , the value of \bar{K}_{k^*} is

$$\begin{aligned} \bar{K}_{k^*} &= \frac{\bar{Q}_{k^*} \bar{R}_{k^*} - \bar{P}_{k^*}^2 + \frac{1}{\bar{R}_{k^*}} (\bar{A}_{k^*}^2 \bar{R}_{k^*}^2 - 2\bar{A}_{k^*} \bar{B}_{k^*} \bar{P}_{k^*} \bar{R}_{k^*} + \bar{B}_{k^*}^2 \bar{P}_{k^*}^2 + \bar{B}_{k^*}^2 \bar{Q}_{k^*} \bar{R}_{k^*} - \bar{B}_{k^*}^2 \bar{P}_{k^*}^2) \bar{K}_{k^*+1}}{\bar{R}_{k^*} + \bar{B}_{k^*}^2 \bar{K}_{k^*+1}} \\ &= \frac{\bar{Q}_{k^*} \bar{R}_{k^*} - \bar{P}_{k^*}^2 + \frac{\bar{B}_{k^*}^2}{\bar{R}_{k^*}} (\bar{Q}_{k^*} \bar{R}_{k^*} - \bar{P}_{k^*}^2) \bar{K}_{k^*+1}}{\bar{R}_{k^*} + \bar{B}_{k^*}^2 \bar{K}_{k^*+1}} = \frac{\bar{Q}_{k^*} \bar{R}_{k^*} - \bar{P}_{k^*}^2}{\bar{R}_{k^*}} \end{aligned}$$

that is independent of \bar{K}_{k^*+1} and then independent of any $\tau_{k^*+1}, \dots, \tau_{N-1}$. These are all potential critical points that need to be explicitly tested.

If instead all intersample separations are such that $\bar{A}_k \bar{R}_k$ never equals $\bar{B}_k \bar{P}_k$ (this happens if the minimum τ_k such that $\bar{A}_k \bar{R}_k = \bar{B}_k \bar{P}_k$ is larger than T), then from (26) it follows that an optimal sampling pattern must satisfy the following condition

$$\frac{\partial r}{\partial \tau}(\tau_h, r(\tau_{h+1}, \bar{K}_{h+2})) - \frac{\partial r}{\partial K}(\tau_h, r(\tau_{h+1}, \bar{K}_{h+2})) \frac{\partial r}{\partial \tau}(\tau_{h+1}, \bar{K}_{h+2}) = 0. \quad (30)$$

This relationship allows finding all intersample separations $\tau_0, \dots, \tau_{N-2}$ starting from any τ_{N-1} using the backward recursive equations (30) and (24). The only issue here is that by choosing

τ_{N-1} we cannot guarantee that $\sum_{k=1}^{N-1} \tau_k = T$. Hence we implemented an iterative procedure that scales the value τ_{N-1} until the constraint $\sum_{k=1}^{N-1} \tau_k = T$ is verified.

C. Numerical solution

In general, finding the $\tau_0, \dots, \tau_{N-1}$ that solve Eq. (26) is very hard. Hence we did implement a gradient descent algorithm, which iteratively performs the following steps:

- 1) computes the gradient $\nabla J = \left(\frac{\partial \bar{K}_0}{\partial \tau_0}, \dots, \frac{\partial \bar{K}_0}{\partial \tau_{N-1}} \right)$ at the current solution;
- 2) project ∇J onto the equality constraint $\sum_{k=0}^{N-1} \tau_k = T$;
- 3) performs a step along the projected gradient and then update the solution if the cost has been reduced or reduce the length of the step if the cost is not reduced.

As it will be later shown in Section VI this numerical optimization procedure is capable to find solutions that are much better than both periodic and Lebesgue sampling. However, this considerable cost reduction has a price. The major drawback of the numerical algorithm is certainly its complexity.

Only the computation of the gradient of the matrix \bar{K}_0 with respect to all sampling instants has the complexity of $O(N^2 n^3)$. This step needs to be computed over and over until numerical stopping criteria of the gradient descent algorithm are reached. While still giving interesting insights on the optimal sampling pattern problem, this considerable computational cost prevents computing the asymptotic behavior for large N and practical application of this result. For this reason we propose below another solution which demonstrated astonishingly efficient characteristics.

IV. QUANTIZATION-BASED SAMPLING

In this section we describe a sampling method that is capable to provide a near-minimal cost (certainly lower than Lebesgue sampling) without requiring to execute a heavy optimization routines. The basic idea is to approximate with a piecewise constant function the optimal continuous-time control input u .

This approach is well studied under the name of *quantization*, a discretization procedure which aims to approximate a function, in the L^p sense, by means of piecewise constant functions. Given a function $u \in L^p(\Omega)$ the goal is to find a piecewise constant function \bar{u} taking only N values,

which realizes the best approximation of u , in the sense that the $L^p(\Omega)$ norm

$$\int_{\Omega} |u(x) - \bar{u}(x)|^p dx$$

is minimal. In our case, the generic quantization problem can be formulated as minimizing the quantization error E_{qua}

$$E_{\text{qua}} = \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} \|u(t) - u_k\|^2 dt. \quad (31)$$

with $t_0 = 0$ and $t_N = T$.

In this problem the unknowns are the constants $\{u_0, \dots, u_{N-1}\}$ to approximate the function u , as well as the intermediate instants $\{t_1, \dots, t_{N-1}\}$. If we differentiate the quantization error E_{qua} with respect to u_k we find that

$$u_k = \frac{1}{t_{k+1} - t_k} \int_{t_k}^{t_{k+1}} u dt. \quad (32)$$

which, not surprisingly, states that the constant u_k that better approximates u in the interval $[t_k, t_{k+1}]$ is its average value over the interval. Thanks to (32) the quantization cost can be rewritten as

$$E_{\text{qua}} = \int_0^T \|u\|^2 dt - \sum_{k=0}^{N-1} (t_{k+1} - t_k) \|u_k\|^2$$

Now we differentiate the error with respect to t_k , with $k = 1, \dots, N-1$. We find

$$\begin{aligned} \frac{\partial E_{\text{qua}}}{\partial t_k} &= -\frac{\partial}{\partial t_k} \left(\frac{1}{t_{k+1} - t_k} \left\| \int_{t_k}^{t_{k+1}} u dt \right\|^2 \right) - \frac{\partial}{\partial t_k} \left(\frac{1}{t_k - t_{k-1}} \left\| \int_{t_{k-1}}^{t_k} u dt \right\|^2 \right) \\ &= -\frac{1}{(t_{k+1} - t_k)^2} \left\| \int_{t_k}^{t_{k+1}} u dt \right\|^2 + 2u'(t_k)u_k + \frac{1}{(t_k - t_{k-1})^2} \left\| \int_{t_{k-1}}^{t_k} u dt \right\|^2 - 2u'(t_k)u_{k-1} \\ &= -\|u_k\|^2 + 2u'(t_k)u_k + \|u_{k-1}\|^2 - 2u'(t_k)u_{k-1} \\ &= \|u_{k-1} - u(t_k)\|^2 - \|u_k - u(t_k)\|^2 \end{aligned}$$

from which it follows that the sampling sequence that minimizes the quantization error must be such that

$$\|u_{k-1} - u(t_k)\|^2 = \|u_k - u(t_k)\|^2 \quad (33)$$

We can then define the quantization-based sampling method (abbreviated with qua) as follows:

- 1) the optimal continuous-time input u is computed;

- 2) the piecewise-constant function \bar{u} that minimizes E_{qua} of (31) is found by applying the gradient condition of (33);
- 3) for the sampling instants $t_0(= 0), t_1, \dots, t_{N-1}, t_N(= T)$ of this solution \bar{u} , we compute the optimal input sequence from (13), since the inputs of (32) are not optimal for the minimization of J .

Later in Section VI we will report the performance of this algorithm.

A. Asymptotic behavior

As shown in [6], [10] the quantization problem of a function $u \in L^p(\Omega)$, is equivalent to minimize the Wasserstein's distance $W_p(\mu, \nu)$ where μ is the image measure $u^\#(dx/|\Omega|)$ and ν is a sum of Dirac masses

$$\nu = \frac{1}{N} \sum_{k=1}^N \delta_{y_k}.$$

As $N \rightarrow \infty$, the asymptotic density of points y_k can be computed and is equal to

$$\frac{f(y)^{m/(m+p)}}{\int f(y)^{m/(m+p)} dy}$$

where m is the dimension of the space of values of u and f is the density of the absolutely continuous part of the measure μ .

In the scalar case, we have $m = 1$ and we find $f(y) = 1/|\dot{u}|(v^{-1}(y))$, provided u is regular enough. From the asymptotic density of values y_k , which is

$$\frac{|\dot{u}|^{-1/(1+p)}(u^{-1}(y))}{\int |\dot{u}|^{-1/(1+p)}(u^{-1}(y)) dy}$$

we can pass to the asymptotic sampling density, which is then

$$\sigma_{\text{qua}}(t) = \frac{|\dot{u}(t)|^{p/(1+p)}}{\int_0^T |\dot{u}(t)|^{p/(1+p)} dt}.$$

Taking $p = 2$, i.e. optimizing the L^2 norm $\int_0^T |u - \bar{u}|^2 dt$, we end up with the asymptotic sampling density

$$\sigma_{\text{qua}}(t) = \frac{|\dot{u}(t)|^{2/3}}{\int_0^T |\dot{u}(s)|^{2/3} ds} \quad (34)$$

Equation (34) provides a very interesting intuition, which can be used as follows to determine the sampling instants:

- 1) the optimal continuous-time input u is computed;

- 2) the sampling instants t_1, \dots, t_{N-1} are determined such that their asymptotic density is (34), by construction. That is, we set

$$\forall k = 0, \dots, N-1 \quad \int_{t_k}^{t_{k+1}} |\dot{u}(t)|^{\frac{2}{3}} dt = \frac{1}{N} \int_0^T |\dot{u}(t)|^{\frac{2}{3}} dt \quad (35)$$

with the usual hypothesis of $t_0 = 0, t_N = T$.

This method is abbreviated with q23 to remind the $\frac{2}{3}$ exponent in (34). While this method is a good approximation of the quantization-based sampling only when N is large, it has the clear advantage being not iterative. In fact, even the qua method cannot be solved analytically and then requires an iterative procedure to find the minimum of the quantization error.

V. THE SCALAR CASE: QUANTIZATION IS OPTIMAL

These quantization-based sampling techniques (methods qua and q23) follows the intuitive idea that the optimal discrete-time input should mimic the optimal continuous-time input. As it will be shown later in Section VI, their excellent capability to reduce the cost appears in all the performed experiments. Unfortunately, we were unable to formally prove a general result that relates the costs J_{qua} or J_{q23} to the minimal continuous-time cost J_{∞} or to the cost of optimal sampling J_{opt} . Nonetheless, if the system is scalar ($n = 1$), we did actually prove that the asymptotic sampling density σ_{qua} of the quantization (Equation (34)) is actually equal to the asymptotic density of the optimal sampling σ_{opt} . The only additional hypothesis we are using is that the weight S of the final state at the instant T is equal to the continuous ARE solution. Although we believe that the following results hold even without this hypothesis, this assumption simplified the mathematical development. Moreover, assuming the weight to the final state equal to the solution of the ARE is not very stringent, since the state at time $t = T$ is going to be small anyway, especially for large T .

Also, throughout this section we assume that if $Q = 0$ then $A > 0$, otherwise the optimal input is obviously $\bar{u}(t) = 0$, which is a constant function independently of the sampling instants.

For such a scalar system we are actually capable to compute analytically the asymptotic optimal sampling density.

Lemma 5: Let assume a scalar system ($n = 1$) with weight to the final state S equal to the solution of the continuous ARE. Then the optimal sampling pattern has asymptotic density

$$\sigma_{\text{opt}} \propto |\dot{u}|^{\frac{2}{3}} \quad (36)$$

with u being the optimal continuous-time input.

Proof: At the price of opportune normalizations of the cost expression and of the system dynamics, we can assume, without loss of generality, that $B = 1$, and $R = 1$.

From (3) it follows that the solution of the ARE, and then the weight to the final state, is

$$S = A + \sqrt{A^2 + Q} \quad (37)$$

This assumption enables us to have a simple expression for the optimal continuous-time input u , that is

$$u(t) = -x_0(A + \sqrt{A^2 + Q})e^{-\sqrt{A^2 + Q}t}. \quad (38)$$

From (6)–(10), the discretized system with an intersample separation of τ_k gives the following discrete-time model

$$\begin{aligned} \bar{A}_k &= 1 + A\tau_k + \frac{A^2}{2}\tau_k^2 + \frac{A^3}{6}\tau_k^3 + \frac{A^4}{24}\tau_k^4 + o(\tau_k^4), \\ \bar{B}_k &= \tau_k + \frac{A}{2}\tau_k^2 + \frac{A^2}{6}\tau_k^3 + \frac{A^3}{24}\tau_k^4 + o(\tau_k^4), \\ \bar{Q}_k &= Q(\tau_k + A\tau_k^2 + \frac{2A^2}{3}\tau_k^3 + \frac{A^3}{3}\tau_k^4) + o(\tau_k^4), \\ \bar{R}_k &= \tau_k + Q(\frac{1}{3}\tau_k^3 + \frac{A}{4}\tau_k^4 + \frac{7A^2}{60}\tau_k^5) + o(\tau_k^5), \\ \bar{P}_k &= Q(\frac{1}{2}\tau_k^2 + \frac{A}{2}\tau_k^3 + \frac{7A^2}{24}\tau_k^4) + o(\tau_k^4). \end{aligned}$$

Since we investigate the asymptotic optimal sampling density ($N \rightarrow \infty$ and then $\tau_k \rightarrow 0$), we realize that $\bar{A}_k \bar{R}_k = \tau_k + o(\tau_k)$ never equals $\bar{B}_k \bar{P}_k = \frac{Q}{2}\tau_k^3 + o(\tau_k^3)$, for small τ_k . This remark allows us to state that the optimal solution must satisfy Eq. (30). Such an equation establishes a relationship between τ_h and τ_{h+1} . Since they both tend to zero, we write τ_h as the following function of τ_{h+1}

$$\tau_h = \alpha\tau_{h+1} + \beta\tau_{h+1}^2 + o(\tau_{h+1}^2), \quad (39)$$

with α and β opportune constants to be found from Eq. (30).

By approximating the Riccati recurrence function r of (29) to the fourth order¹ w.r.t. τ , we

¹For computing this and next expressions, we made use of the symbolic manipulation tool “Maxima” (<http://maxima.sourceforge.net/>).

find that

$$\begin{aligned}
r(\tau, \bar{K}) = & \bar{K} - (\bar{K}^2 - 2A\bar{K} - Q)\tau + (\bar{K}^3 - 3A\bar{K}^2 + (2A^2 - Q)\bar{K} + AQ)\tau^2 \\
& - (\bar{K}^4 - 4A\bar{K}^3 + (\frac{55}{12}A^2 - \frac{4}{3}Q)\bar{K}^2 + (-\frac{4}{3}A^3 + \frac{5}{2}QA)\bar{K} - \frac{2}{3}QA^2 + \frac{1}{4}Q^2)\tau^3 \\
& + (\bar{K}^5 - 5A\bar{K}^4 + (\frac{49A^2}{6} - \frac{5}{3}Q)\bar{K}^3 + \frac{19}{4}(-A^3 + QA)\bar{K}^2 \\
& + (\frac{2A^4}{3} - \frac{13QA^2}{4} + \frac{7Q^2}{12})\bar{K} + \frac{QA^3}{3} - \frac{Q^2A}{2})\tau^4 + o(\tau^4)
\end{aligned} \tag{40}$$

with the following partial derivatives

$$\begin{aligned}
\frac{\partial r}{\partial \tau} = & -(\bar{K}^2 - 2A\bar{K} - Q) + 2(\bar{K}^3 - 3A\bar{K}^2 + (2A^2 - Q)\bar{K} + AQ)\tau \\
& - 3(\bar{K}^4 - 4A\bar{K}^3 + (\frac{55}{12}A^2 - \frac{4}{3}Q)\bar{K}^2 + (-\frac{4}{3}A^3 + \frac{5}{2}QA)\bar{K} - \frac{2}{3}QA^2 + \frac{1}{4}Q^2)\tau^2 \\
& + 4(\bar{K}^5 - 5A\bar{K}^4 + (\frac{49A^2}{6} - \frac{5}{3}Q)\bar{K}^3 + \frac{19}{4}(-A^3 + QA)\bar{K}^2 \\
& + (\frac{2A^4}{3} - \frac{13QA^2}{4} + \frac{7Q^2}{12})\bar{K} + \frac{QA^3}{3} - \frac{Q^2A}{2})\tau^3 + o(\tau^3)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial r}{\partial \bar{K}} = & 1 - 2(\bar{K} - A)\tau + (3\bar{K}^2 - 6A\bar{K} + 2A^2 - Q)\tau^2 \\
& - (4\bar{K}^3 - 12A\bar{K}^2 + 2(\frac{55}{12}A^2 - \frac{4}{3}Q)\bar{K} - \frac{4}{3}A^3 + \frac{5}{2}QA)\tau^3 + o(\tau^3)
\end{aligned}$$

It is now possible to replace in the necessary condition for optimality (30) these expressions of partial derivatives. If we write τ_h as a function of τ_{h+1} (Eq. (39)), we find the following equation

$$\begin{aligned}
& -\tau_{h+1}^2 \frac{Q + A\bar{K}_{h+2}}{12} ((6(\alpha - 1)(\alpha + 1)^2 A\bar{K}_{h+2}^2 \\
& + 2((\alpha + 1)((2\alpha^2 - 2\alpha - 1)Q + (-4\alpha^2 - 2\alpha + 5)A^2) - 3\alpha\beta A)\bar{K}_{h+2} \\
& - 6((\alpha + 1)^2(\alpha - 1)A + \alpha\beta)Q)\tau_{h+1} - 3(\alpha^2 - 1)(Q + A\bar{K}_{h+2})) + o(\tau_{h+1}^3) = 0
\end{aligned}$$

from which we have

$$\begin{aligned}
& (6(\alpha - 1)(\alpha + 1)^2 A\bar{K}_{h+2}^2 + 2((\alpha + 1)((2\alpha^2 - 2\alpha - 1)Q + (-4\alpha^2 - 2\alpha + 5)A^2) - 3\alpha\beta A)\bar{K}_{h+2} \\
& - 6((\alpha + 1)^2(\alpha - 1)A + \alpha\beta)Q)\tau_{h+1} - 3(\alpha^2 - 1)(Q + A\bar{K}_{h+2}) + o(\tau_{h+1}) = 0 \tag{41}
\end{aligned}$$

A necessary condition to have Equation (41) true, is that both coefficient of τ_{h+1} and the constant are zero. By setting the constant term of (41) equal to zero we find $\alpha^2 = 1$. However,

from (39), we observe that $\alpha = -1$ is not acceptable, since it will lead to negative intersample separations. Hence we have $\alpha = 1$. By replacing $\alpha = 1$ in the coefficient of τ_{h+1} in (41), we find

$$(2Q + 2A^2 + 3\beta A)\bar{K}_{h+2} + 3\beta Q = 0$$

from which we find

$$\beta = -\frac{2(Q + A^2)\bar{K}_{h+2}}{3(Q + A\bar{K}_{h+2})}.$$

Recalling the expression (39), we can now assert that a necessary condition for the optimality of a sampling pattern is that

$$\tau_h = \tau_{h+1} - \frac{2(Q + A^2)\bar{K}_{h+2}}{3(Q + A\bar{K}_{h+2})}\tau_{h+1}^2 + o(\tau_{h+1}^2). \quad (42)$$

We are now going to exploit (42) to find the asymptotic sampling density of the optimal pattern.

Let us compute the derivative of the asymptotic density of the optimal sampling at a generic instant t_{h+1} . By Definition 1 and 2, we have

$$\begin{aligned} \dot{\sigma}_{\text{opt}}(t_{h+1}) &= \lim_{N \rightarrow \infty} \frac{\sigma_{\text{opt}}(t_{h+2}) - \sigma_{\text{opt}}(t_{h+1})}{\tau_{h+1}} = \lim_{N \rightarrow \infty} \frac{\frac{1}{N\tau_{h+1}} - \frac{1}{N\tau_h}}{\tau_{h+1}} = \lim_{N \rightarrow \infty} \frac{1 - \frac{1}{1 - \frac{2(Q+A^2)\bar{K}_{h+2}}{3(Q+A\bar{K}_{h+2})}\tau_{h+1}}}{N\tau_{h+1}^2} \\ &= \lim_{N \rightarrow \infty} \frac{-\frac{2(Q+A^2)\bar{K}_{h+2}}{3(Q+A\bar{K}_{h+2})}\tau_{h+1}}{N\tau_{h+1}^2} = -\frac{2(Q+A^2)\bar{K}_{h+2}}{3(Q+A\bar{K}_{h+2})} \lim_{N \rightarrow \infty} \frac{1}{N\tau_{h+1}} = -\frac{2(Q+A^2)\bar{K}_{h+2}}{3(Q+A\bar{K}_{h+2})} \sigma_{\text{opt}}(t_{h+2}) \end{aligned}$$

from which it follows the following differential equation

$$\dot{\sigma}_{\text{opt}}(t) = -\frac{2(Q + A^2)K(t)}{3(Q + AK(t))} \sigma_{\text{opt}}(t) \quad (43)$$

with $K(t)$ being the solution of the Riccati differential equation (3). If $S = A + \sqrt{Q + A^2}$, then $K(t)$ is constantly equal to S . Then in such special case, Eq. (43) becomes

$$\dot{\sigma}_{\text{opt}}(t) = -\frac{2(Q + A^2)(A + \sqrt{Q + A^2})}{3(Q + A(A + \sqrt{Q + A^2}))} \sigma_{\text{opt}}(t) = -\frac{2}{3} \sqrt{Q + A^2} \sigma_{\text{opt}}(t) \quad (44)$$

which is solved by

$$\sigma_{\text{opt}}(t) = c e^{-\frac{2}{3} \sqrt{Q + A^2} t} \quad (45)$$

with c opportune constant such that $\int_0^T \sigma_{\text{opt}}(t) dt = 1$.

By observing the expression (38) of the optimal continuous-time input we realize that $\sigma_{\text{opt}} \propto |\dot{u}|^{\frac{2}{3}}$. The Lemma is then proved. ■

We believe that Lemma 5 is a quite interesting result. Basically it states that, by tolerating the weak assumption that the weight S to the final state $x(T)$ is set equal to the solution of the continuous ARE, the asymptotic density σ_{opt} of the optimal sampling is the same as the asymptotic density σ_{q23} of the quantization-based sampling. In addition to this result, the next Lemma also provides an exact computation of the asymptotic normalized cost (see Definition 4) of the optimal sampling. This result allows quantifying the benefit of optimal sampling.

The following Lemma provides a more general result from which the asymptotic normalized cost c_{opt} of the optimal sampling is derived later in Corollary 7.

Lemma 6: Let us assume a scalar system ($n = 1$). Let us also assume, at the price of opportune normalizations, that $B = 1$ and $R = 1$ and that the weight to the final state is equal to the solution of the continuous ARE, $S = A + \sqrt{A^2 + Q}$.

Then, the asymptotic normalized cost of the sampling method $m\alpha$ with asymptotic sampling density

$$\begin{aligned}\sigma_{m\alpha}(t) &\propto |\dot{u}(t)|^\alpha \\ \sigma_{m\alpha}(t) &= \frac{\alpha(S-A)}{1 - e^{-\alpha(S-A)T}} e^{-\alpha(S-A)t}\end{aligned}\quad (46)$$

is

$$c_{m\alpha} = \frac{S}{12(S-A)T^2} \frac{1 - e^{-2(1-\alpha)(S-A)T}}{2(1-\alpha)} \left(\frac{1 - e^{-\alpha(S-A)T}}{\alpha} \right)^2 \quad (47)$$

Proof: Within these hypothesis, the solution of the Riccati differential equation is $K(t) = S = A + \sqrt{Q + A^2}$, for all $t \in [0, T]$. Hence the optimal continuous-time cost is

$$J_\infty = x_0^2 S \quad (48)$$

Since we are investigating the normalized cost c_{opt} (see Definition 3), we study the following sequence:

$$\delta_k = \frac{N^2}{T^2} \left(\frac{\bar{K}_k}{S} - 1 \right)$$

so that $c_{\text{opt}} = \lim_{N \rightarrow \infty} \delta_0$. From the definition of δ_k it follows that

$$\bar{K}_k = S \left(\frac{T^2}{N^2} \delta_k + 1 \right)$$

From (40), by approximating \bar{K}_k to the third order of τ_k , we have

$$\begin{aligned}\bar{K}_k &= \bar{K}_{k+1} - (\bar{K}_{k+1} - S)(\bar{K}_{k+1} - 2A + S)\tau_k + (\bar{K}_{k+1} - A)(\bar{K}_{k+1} - S)(\bar{K}_{k+1} - 2A + S)\tau_k^2 \\ &\quad - (\bar{K}_{k+1}^4 - 4A\bar{K}_{k+1}^3 + (\frac{55}{12}A^2 - \frac{4}{3}Q)\bar{K}_{k+1}^2 + (-\frac{4}{3}A^3 + \frac{5}{2}QA)\bar{K}_{k+1} - \frac{2}{3}QA^2 + \frac{1}{4}Q^2)\tau_k^3 + o(\tau_k^3)\end{aligned}$$

which allows to find a recurrent relationship that defines δ_k

$$\begin{cases} \delta_N = 0 \\ \delta_k = \delta_{k+1} - \delta_{k+1}(S\frac{T^2}{N^2}\delta_{k+1} + 2(S-A))\tau_k \\ \quad + \delta_{k+1}(S\frac{T^2}{N^2}\delta_{k+1} + 2(S-A))(S\frac{T^2}{N^2}\delta_{k+1} + S-A)\tau_k^2 + \frac{S(S-A)^2}{12}\frac{N^2}{T^2}\tau_k^3 + o(\tau_k^3) \end{cases} \quad (49)$$

From (49), it follows that the discrete derivative of δ_k is

$$\frac{\delta_{k+1} - \delta_k}{\tau_k} = \delta_{k+1}(S\frac{T^2}{N^2}\delta_{k+1} + 2(S-A)) - \delta_{k+1}(S\frac{T^2}{N^2}\delta_{k+1} + 2(S-A))(S\frac{T^2}{N^2}\delta_{k+1} + S-A)\tau_k - \frac{S(S-A)^2}{12}\frac{N^2}{T^2}\tau_k^2 + o(\tau_k^2)$$

By definition of asymptotic sampling density (see Definitions 1 and 2), as $N \rightarrow \infty$, the intersample separation τ_k tends to zero with

$$\tau_k = \frac{1}{N\sigma_{m\alpha}(t_k)} + o\left(\frac{1}{N}\right).$$

From this observation, as $N \rightarrow \infty$, the discrete derivative of δ_k becomes the following differential equation

$$\dot{\delta}(t) = 2(S-A)\delta(t) - \frac{S(S-A)^2}{12T^2}\sigma_{m\alpha}^{-2}(t),$$

where $\delta(t)$ is the limit of δ_k .

With the sampling density $\sigma_{m\alpha}(t)$ of (46), the differential equation becomes

$$\dot{\delta}(t) = 2(S-A)\delta(t) - \frac{S(1 - e^{-\alpha(S-A)T})^2}{12\alpha^2T^2}e^{2\alpha(S-A)t}$$

which is a first-order linear non-homogeneous differential equation. Its solution then is the solution of the homogeneous one plus a particular integral

$$\delta(t) = ce^{2(S-A)t} + \delta^*(t)$$

with the constant c to be found from the boundary condition $\delta(T) = 0$ (as required by the condition $\delta_N = 0$ in (49)).

Let us search for a particular integral $\delta^*(t)$ of the form $\delta^*(t) = de^{2\alpha(S-A)t}$. Then the constant d must be such that

$$\begin{aligned} 2\alpha(S-A)de^{2\alpha(S-A)t} &= 2(S-A)de^{2\alpha(S-A)t} - \frac{S(1 - e^{-\alpha(S-A)T})^2}{12\alpha^2T^2}e^{2\alpha(S-A)t} \\ \frac{S(1 - e^{-\alpha(S-A)T})^2}{12\alpha^2T^2} &= 2(S-A)(1-\alpha)d \\ d &= \frac{S(1 - e^{-\alpha(S-A)T})^2}{24(S-A)(1-\alpha)\alpha^2T^2} \end{aligned}$$

Now we find the constant c from the boundary condition $\delta(T) = 0$

$$\begin{aligned}\delta(T) &= c e^{2(S-A)T} + \delta^*(T) = 0 \\ c &= -\frac{\delta^*(T)}{e^{2(S-A)T}} = -\frac{S(1 - e^{-\alpha(S-A)T})^2 e^{-2(1-\alpha)(S-A)T}}{24(S-A)(1-\alpha)\alpha^2 T^2}\end{aligned}$$

from which we can finally find the exact expression of the function δ

$$\delta(t) = \frac{S(1 - e^{-\alpha(S-A)T})^2}{24(S-A)(1-\alpha)\alpha^2 T^2} (e^{2\alpha(S-A)t} - e^{-2(1-\alpha)(S-A)T} e^{2(S-A)t})$$

and the asymptotic normalized cost $c_{m\alpha} = \delta(0)$, which is the same as (47). Hence the Lemma is proved. \blacksquare

Lemma 6 provides the asymptotic normalized cost for any sampling method whose asymptotic sampling density is the one of (47). By simply observing that

- the periodic sampling has constant sampling density, hence it corresponds to the case with $\alpha = 0$;
- the Lebesgue sampling corresponds, by construction, to the case with $\alpha = 1$;
- from Lemma 5, the optimal sampling corresponds to the case with $\alpha = \frac{2}{3}$;

then the following Corollary follows.

Corollary 7: Let us assume a scalar system ($n = 1$). Let us also assume, at the price of opportune normalizations, that $B = 1$ and $R = 1$ and that the weight to the final state equal to the solution of the continuous ARE, $S = A + \sqrt{A^2 + Q}$.

Then, the asymptotic normalized costs of the periodic, Lebesgue, and optimal sampling are, respectively

$$c_{\text{per}} = c_{m0} = \frac{(A\sqrt{A^2 + Q} + A^2 + Q)}{24} (1 - e^{-2\sqrt{A^2 + Q}T}), \quad (50)$$

$$c_{\text{leb}} = c_{m1} = \frac{A + \sqrt{A^2 + Q}}{12T} (1 - e^{-\sqrt{A^2 + Q}T})^2, \quad (51)$$

$$c_{\text{opt}} = c_{m\frac{2}{3}} = \frac{9}{32T^2} \left(\frac{A}{\sqrt{A^2 + Q}} + 1 \right) (1 - e^{-\frac{2}{3}\sqrt{A^2 + Q}T})^3. \quad (52)$$

Notice that as $T \rightarrow \infty$, the cost c_{per} coincides with the one derived earlier in (22).

VI. EVALUATION

EB: to be revised

In the case of the three-dimensional system of (??) already considered in the last sections, this gradient descent algorithm can improve also the solution found with the Lebesgue sampling. In Figure 1 we draw the optimal continuous-time input (in thick gray) and the optimal control input found with the gradient descent algorithm (in black). For this value of $N = 32$ the normalized

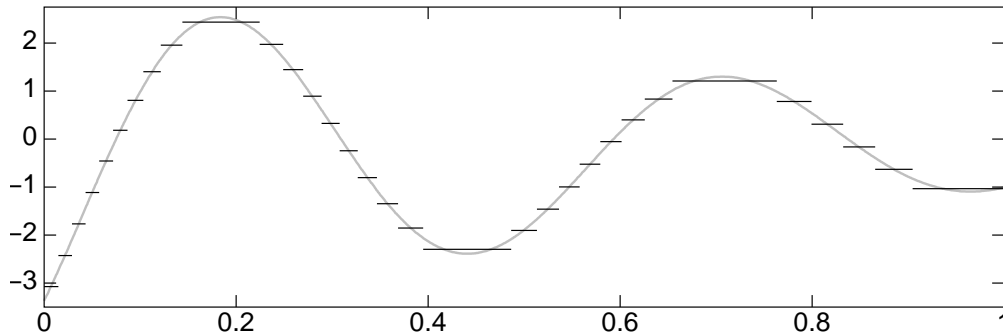


Fig. 1: Optimal sampling.

cost achieved by the optimal numerical algorithm is 6.1378 compared to 7.1979 of the Lebesgue sampling (and 13.070 of the periodic one). If the optimal sampling is chosen, then it is enough to use 17 samples to reach the threshold of the 2% of cost increase w.r.t. to the continuous-time optimal input (compared to the 21 samples with Lebesgue sampling and the 27 samples with periodic sampling). In fact, when $N = 17$ we have $c_{\text{opt},17} = 5.6997$ and by applying (18) we have $T\sqrt{\frac{c_{\text{opt},17}}{r}} = \sqrt{\frac{5.6997}{0.02}} = 16.881 \leq 17$.

It can be noticed that the sampling instant of the optimal input are still related to the derivative of the continuous-time optimal control. In Figure 2 we plot the sampling densities of both the Lebesgue sampling (in gray) and the optimal one (in black), when the number of samples is 32 in $[0, 1]$. The optimal sampling appears to be bound to \dot{u} in a looser way. In fact the optimal instants are a bit less dense than the Lebesgue sampling pattern when \dot{u} has a steep slope, while they are slightly more dense in proximity of local maxima/minima of u . Hence, the optimal sampling pattern tends to underweight the instants when the derivative of u is high compared to the Lebesgue sampling. Roughly speaking we can say that the optimal sampling pattern is “in between” periodic and Lebesgue sampling.

With the same system used in the last sections, in Table I we compare the normalized costs of the methods described so far:

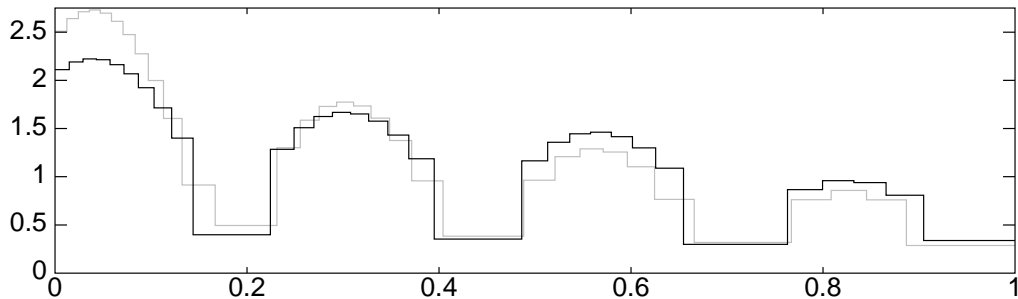


Fig. 2: Comparing Lebesgue and optimal sampling densities.

- periodic sampling (per),
- Lebesgue sampling (leb),
- quantization based on the theoretical asymptotic density of (35) (q23),
- quantization based on the exact condition of gradient equal to zero of Equation (33) (for large N this method tends to the one with asymptotic density) (qua),
- optimal numerical solution (opt).

For example computing the optimal pattern of Figure 1 (hence with $n = 3$ and $N = 32$) required 2 hours on a 2.4 GHz laptop.

We also investigate the behavior of the sampling methods with varying weight on the cost: $Q = 0$ (Table Ia), $Q = 10 I$ (Table Ib), and $Q = 100 I$ (Table Ic). The optimal continuous-time

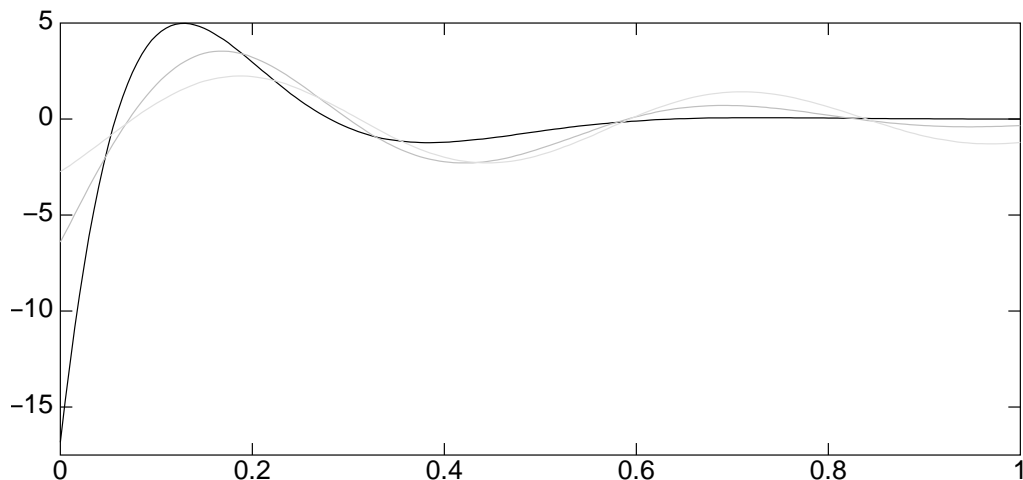


Fig. 3: Optimal continuous-time input with varying Q .

input is also drawn in Figure 3: the light gray plot depicts the optimal input when $Q = 0$, the dark gray one when $Q = 10 I$, and the black one corresponds to the case with $Q = 100 I$. It can be observed that the magnitude of the input u is larger as its relative cost (w.r.t. the state cost) is smaller. Below we provide some comments on the data reported in Table I:

N	10	20	40	N	10	20	40	N	10	20	40
per	14.488	13.382	13.136	per	16.615	14.640	14.221	per	30.959	25.509	24.352
leb	9.4078	8.0191	8.4132	leb	7.2274	6.3719	4.8349	leb	10.161	24.900	4.5791
q23	8.2823	8.1115	7.8431	q23	4.9684	3.7374	3.6936	q23	1.7244	1.1534	1.2179
qua	5.9003	6.6710	7.2797	qua	2.9200	3.2204	3.6597	qua	0.96569	1.1478	1.2095
opt	5.8911	6.5765	7.1273	opt	2.9139	3.1786	3.4165	opt	0.94476	0.99627	1.1631

(a) $Q = 0$ (b) $Q = 10 I$ (c) $Q = 100 I$

TABLE I: Comparison of normalized costs $c_{m,N}$.

- The experiments confirm the validity of the asymptotic density of the quantization-based sampling of (34), since the cost achieved by q23 tends to the cost of the numerical quantization qua as N grows;
- The cost achieved by the quantization-based sampling (qua and, q23 for larger N) is very close to the optimal one.
- The capacity of both quantization-based sampling and Lebesgue sampling to reduce the cost w.r.t. periodic sampling is much higher in all those circumstances with high variation of the optimal continuous-time input u (such as when Q is larger compared to R). The increase of c_{per} with Q was also analytically demonstrated in (22), for uni-dimensional systems.

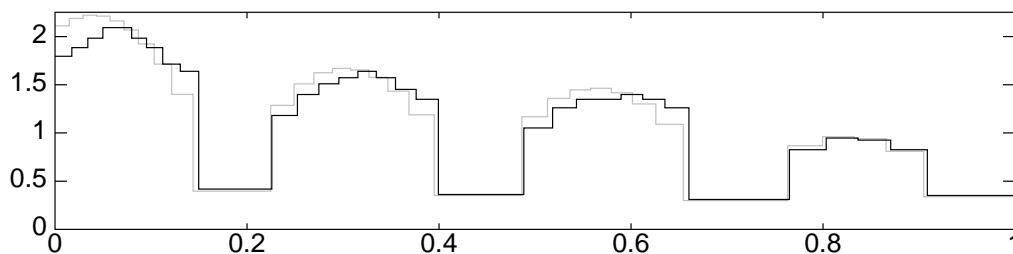


Fig. 4: Comparing the densities of optimal and quantization-based sampling.

In Figure 4 we compare the density of the optimal sampling (in gray) with the density of the

quantization-based sampling (in black). The two densities are very similar. From Eq. (34) we have the confirmation that the quantization-based sampling (and then the optimal one since they are similar) is “in between” periodic and Lebesgue sampling. In fact its asymptotic density is $|\dot{u}|^{2/3}$ rather than $|\dot{u}|$ (for the Lebesgue sampling) or the constant density $|\dot{u}|^0$ (for the periodic sampling).

While the two numerical sampling methods (qua and opt) have little chances to be implemented on a controller because of their heavy computational requirement, the sampling method based on the asymptotic density (q23) is always better than Lebesgue sampling and is considerably more efficient to be implemented.

VII. CONCLUSIONS AND FUTURE WORKS

In this paper we have investigated in detail, the relationship between the cost and the sampling instants. We formulate the problem for determining the optimal sampling sequence and we proposed a very efficient sampling method which reaches very close performance to the optimal one.

Being this research quite new, there are more open issues than questions with answers. Among the open problems we mention:

- a deeper investigation of the dependency of the optimal sampling pattern on the many system features;
- possible more efficient implementation of the gradient optimization procedure;
- the investigation of global minimization procedures which could lead to a higher cost reduction (gradient descent algorithms could indeed fall into local minima).

REFERENCES

- [1] Adolfo Anta and Paulo Tabuada. To sample or not to sample: Self-triggered control for nonlinear systems. *IEEE Transactions on Automatic Control*, 55(9):2030–2042, September 2010.
- [2] Karl-Erik Årzén. A simple event-based pid controller. In *14th World Congress of IFAC*, Beijing, China, January 1999.
- [3] Karl Johan Åström and Bo Bernhardsson. Comparison of Riemann and Lebesgue sampling for first order stochastic systems. In *Proceedings of the 41st IEEE Conference on Decision and Control*, December 2002.
- [4] Karl Johan Åström and Bo Bernhardsson. Systems with Lebesgue sampling. In *Directions in Mathematical Systems Theory and Optimization*, volume 286 of *Lecture Notes in Computer Science*. Springer-Verlag, 2003.
- [5] Richard Bellman. The theory of dynamic programming. *Bulletin of the American Mathematical Society*, 60:503–516, 1954.

- [6] Siegfried Graf and Harald Luschgy. *Foundations of Quantization for Probability Distributions*. Springer, Berlin, 1999.
- [7] Kaska Kowalska and Martin von Mohrenschildt. An approach to variable time receding horizon control. *Optimal Control Applications and Methods*, pages n/a–n/a, 2011.
- [8] Alan John Laub. A schur method for solving algebraic Riccati equations. *IEEE Transactions on Automatic Control*, 24(6):913–921, December 1979.
- [9] Stuart M. Melzer and Benjamin C. Kuo. Sampling period sensitivity of the optimal sampled data linear regulator. *Automatica*, 7:367–370, 1971.
- [10] Gilles Pagès. Quadratic optimal functional quantization of stochastic processes and numerical applications. In *Monte Carlo and quasi-Monte Carlo methods in scientific computing*, pages 101–142, Berlin, 2008. Springer-Verlag.
- [11] Lev Semenovich Pontryagin, Vladimir Grigorevich Boltyanskii, Revaz V. Gamkrelidze, and Evgenii Frolovich Mishchenko. The mathematical theory of optimal processes. *New York: Interscience Publisher, Inc.*, 1962. (translation from a Russian book).
- [12] Manel Velasco, Pau Martí, and Joseph M. Fuertes. The self triggered task model for real-time control systems. In *24th IEEE Real-Time Systems Symposium (Work-in-Progress session)*, December 2003.
- [13] Xiaofeng Wang and Michael D. Lemmon. Self-triggered feedback control systems with finite-gain \mathcal{L}_2 stability. *IEEE Transactions on Automatic Control*, 54(3):452–467, March 2009.
- [14] Yan-Kai Xu and Xi-Ren Cao. Lebesgue-sampling-based optimal control problems with time aggregation. *IEEE Transactions on Automatic Control*, 56(5):1097–1109, May 2011.