Deconvolution of quantized-input linear systems: Analysis via Markov Processes of a low-complexity algorithm.

Sophie M. Fosson and Paolo Tilli

Abstract— **This paper is concerned with the problem of the deconvolution, which consists in recovering the unknown input of a linear system from a noisy version of the output. The case of a system with quantized input is considered and a low-complexity algorithm, derived from decoding techniques, is introduced to tackle it. The performance of such algorithm is analytically evaluated through the Theory of Markov Processes. In this framework, results are shown which prove the uniqueness of an invariant probability measure of a Markov Process, even in case of non-standard state space. Finally, the theoretic issues are compared with simulations' outcomes.**

I. INTRODUCTION

Consider the following time-invariant linear system

$$\begin{cases} x'(t) = ax(t) + bu(t) \qquad t \in [0,T] \\ y(t) = cx(t) \\ x(0) = 0 \end{cases} \tag{1}$$

where $u(t)$, $x(t)$ and $y(t)$ respectively are the input, the state function and the output; a, b and c are real constants, $[0,T]$ is a possibly infinite time horizon. $u(t)$ is supposed to be unknown, while $y(t)$ is accessible, but corrupted by an observational, additive noise $n(t)$: the available data is then

$$r(t) = y(t) + n(t). \tag{2}$$

The aim of this work is to reconstruct $u(t)$ from $r(t)$, that is, to reverse the input-output convolution integral:

$$r(t) = cx(t) + n(t) = cb\int_0^t e^{a(t-s)}u(s)\mathrm{d}s + n(t). \tag{3}$$

Such an inversion, properly known as *deconvolution*, has been receiving considerable attention since the last 1960s, [14], [15]. Its study, motivated by the ubiquitous engineering and technological applications (for example, seismology and geophysics, astronomy, image processing, industrial and medical systems, see, e.g., [2], [3], [4], [12], [13]), has been attracting interest among mathematicians for its ill-posed and ill-conditioned nature. Indeed, the integral (3) cannot be directly inverted, since this may produce non-unique and faulty solutions; the purpose is then to approximate the correct input $u(t)$ by the means of some estimation algorithm.

In order to achieve this, it is desirable to have some prior information on the input, which will drive the development of a suitable estimation procedure. In this work, we assume $u(t)$ to be quantized over a finite number of levels, this choice being motivated by the recent digitalization of many technological processes and devices. In the practice, this condition can model systems whose input is a switch, an actuator or a signal monitoring the (discrete) state of an industrial process. A goal of such study is, for instance, the design of fault detection devices.

The applications motivate also our interest for the *on-line* estimation algorithms, that is, for procedures that deconvolve the system during the course of transmission (see, e.g., [5], [6]); for instance, in a fault detection problem an eventual defect should be revealed as soon as possible and not when the process has been completed.

Furthermore, we will show that the quantized nature of the input makes our system analogous to a digital transmission paradigm. This inspires an Information-Coding description of the problem and suggests to exploit decoding techniques instead of classical estimation algorithms.

Finally, low-complexity is also desirable for the sake of the implementation.

All these elements are combined in the so-called One State Algorithm, that has been introduced in [8] for a differentiation system, i.e., in a simplified deconvolution problem. The aim of this work is to extend its use to general linear systems such as (1). In particular, in the next a theoretic analysis is developed, which provides a rigorous description of the algorithm. Such analysis is conducted in a probabilistic setting: the algorithm's pattern, in fact, can be modeled as a Markov Process and considering the case of long-time transmission (that is, $[0,T]$ is infinite), the Ergodic Theory of Markov Processes can be exploited to infer the behavior of the proposed deconvolution technique. Furthermore, a contribution is provided in this framework for what concerns the study of the invariant probability measures, that will be studied through the theory of metric spaces and by the application of the Banach Fixed Point Theorem.

The paper is organized follows. In the next section, we make some further assumptions and we describe the evolution of the system. In Section III, we introduce the Information-Coding description of the problem, which motivates the introduction of the One State Algorithm. The algorithm is presented in Section IV and theoretically analysed in Section V. Finally, Sections VI and VII are devoted to simulations, comparisons and conclusive considerations.

A. Notation, List of Symbols and Abbreviations

- Given a subset A of a set Ω, $\mathbb{1}_A : \Omega \to \{0,1\}$ is the indicator function, defined by $\mathbb{1}_A(x) = 1$ if x belongs

S.M. Fosson and P. Tilli are with the Department of Mathematics, Politecnico di Torino, Corso Duca Degli Abruzzi 24, 10129 Torino, Italy. sophie.fosson@polito.it, paolo.tilli@polito.it

S.M. Fosson's research is supported by Fondo Sociale Europeo, Regione Autonoma Valle d'Aosta.

ISBN 978-963-311-370-7

to A and $\mathbb{1}_A(x) = 0$ otherwise.
- erfc is the complementary error function, defined by $\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{+\infty} e^{-s} ds$ for any $x \in \mathbb{R}$.
- Discrete probabilities are indicated with P, while $P(\cdot,\cdot)$ denotes the transition probability kernel of a Markov Process. $\mathbb{E}$ is the stochastic mean.
- $C_b(\Omega)$ is the space of the bounded continuous functions on Ω.
- $\mathcal{B}(\Omega)$ is the Borel σ-algebra of Ω.
- Given a bounded measurable function v on $(\Omega, \mathcal{B}(\Omega))$, we denote by Pv the bounded measurable function on $(\Omega, \mathcal{B}(\Omega))$ defined by $(Pv)(x) = \int_\Omega v(y) P(x, dy)$.
- Given a measure μ on $(\Omega, \mathcal{B}(\Omega))$, we define the measure $(\mu P)(B) = \int_\Omega P(x, B)\mu(dx)$, $B \in \mathcal{B}(\Omega)$ (see [8, Paragraph 4.2.1]).
- $f \in m\text{-Lip}(\Omega)$ indicates that the function $f : \Omega \to \mathbb{R}$ is Lipschitz of constant m.

II. Problem Statement

The following assumptions are made on system (1):

Assumption 1: the input is quantized on two levels, say 0 and 1, and the switches from one level to the other one can occur only at time instants $k\tau$, $\tau > 0$, $k \in \mathbb{N}$. For simplicity, the time step τ is such that $T/\tau = K \in \mathbb{N}$. Therefore, we can write:

$$u(t) = \sum_{k=0}^{K-1} u_k \mathbb{1}_{[k\tau,(k+1)\tau[}(t), \quad u_k \in \mathcal{U} = \{0, 1\} \tag{4}$$

Now, $u(t)$ is completely determined by the bit sequence $(u_0, \dots, u_{K-1})$.

Assumption 2: the lecture of $r(t)$ is performed only at the same time instants $k\tau$, $k = 1, \dots, K$, hence the available information is the sequence $(r_1, \dots, r_K)$ where $r_k = r(k\tau) = y(k\tau) + n(k\tau)$.

Assumption 3: the system is stable, that is $a < 0$. The case $a = 0$ has already been studied in [8].

Assumption 1 induces a discrete nature also for $x(t)$: given the expression (4), we have

$$\begin{aligned} x(k\tau) &= b e^{ak\tau} \int_0^{k\tau} e^{-as} \sum_{h=0}^{K-1} u_h \mathbb{1}_{[h\tau,(h+1)\tau[}(s) ds \\ &= b e^{ak\tau} \sum_{h=0}^{k-1} u_h \int_{h\tau}^{(h+1)\tau} e^{-as} ds \\ &= \frac{b}{a}(e^{a\tau} - 1) e^{a(k-1)\tau} \sum_{h=0}^{k-1} u_h e^{-ah\tau}. \end{aligned} \tag{5}$$

Letting $x_k = x(\tau k)$, $\mathrm{q} = e^{a\tau}$, $\mathrm{w} = -\frac{b}{a}(1 - \mathrm{q})$, we can write the following recursive formula:

$$x_k = \mathrm{q} x_{k-1} + \mathrm{w} u_{k-1}. \tag{6}$$

By trivial computations, we obtain also

$$x_k = \mathrm{w} \sum_{h=0}^{k-1} u_{k-h-1} \mathrm{q}^h$$

which shows that each x_k assumes values in the set

$$\mathcal{X}_k = \mathrm{w} \left\{ \sum_{h=0}^{k-1} \mu_h \mathrm{q}^h, \ \mu_h \in \{0, 1\} \right\}$$

of the binary polynomials of degree at most $k-1$ multiplied by w. The cardinality of $\mathcal{X}_k$ is then 2^k and $\mathcal{X}_k \subset \mathcal{X}_{k+1}$ for any $k \in \mathbb{N}$. If we let K tend to infinity, we can define the more general state set

$$\mathcal{X} = \mathrm{w} \left\{ \sum_{h=0}^{\infty} \mu_h \mathrm{q}^h, \ \mu_h \in \{0, 1\} \right\}$$

that includes all the x_k's, $k \in \mathbb{N}$. The structure of $\mathcal{X}$ will play a fundamental role in the deconvolution's performance. For computational simplicity, from now onwards let

$$\tau = 1 \quad \text{and} \quad b > 0.$$

Moreover, for notational simplicity in this section we suppose $\mathrm{w} = 1$, so that $\mathcal{X} \subseteq [0, \frac{1}{1-\mathrm{q}}]$, with $\mathrm{q} \in (0, 1)$ by Assumption 3. Now, two possible cases have to be distinguished.

A. Case $\mathrm{q} \in \left[\frac{1}{2}, 1\right)$.

If $q \in \left[\frac{1}{2}, 1\right)$, then $\mathcal{X} \equiv \left[0, \frac{1}{1-q}\right]$. This can be proved as follows. Given any $x \in \left[0, \frac{1}{1-q}\right]$, we construct a series $\sum_{h=0}^{\infty} \mu_h q^h = x$, $\mu_h \in \{0, 1\}$ defining on by one the coefficients μ_h. The series being positive termed, the procedure is:

For $h = 0$, fix

$$\begin{cases} \mu_0 = 1 & \text{if } x \geq 1 \\ \mu_0 = 0 & \text{otherwise} \end{cases}$$

For $h = 1, 2, 3, \dots$, fix

$$\begin{cases} \mu_h = 1 & \text{if } x \geq \sum_{i=0}^{h-1} \mu_i + \mathrm{q}^h \\ \mu_h = 0 & \text{otherwise} \end{cases}$$

For any $h \in \mathbb{N}$, we then obtain a polynomial $\sum_{i=0}^{h} \mu_i q^i \leq x$. In case of equality, the property is proved; otherwise, we have to show that

$$x \leq \sum_{i=0}^{h} \mu_i \mathrm{q}^i + \sum_{i=h+1}^{\infty} \mathrm{q}^i \tag{7}$$

This is obvious if $\mu_i = 1$ for any $i = 0, \dots, h$. Otherwise, if there exists at least one null coefficient between 0 and h, let us consider the null coefficient with greater index, that is, pick j such that $\mu_j = 0$ and $\mu_i = 1$ for any $i = j+1, \dots, h$. Then, $x < \sum_{i=0}^{j-1} \mu_i \mathrm{q}^i + \mathrm{q}^j$, otherwise it should have been $\mu_j = 1$. Now, in order to prove the bound (7) it is sufficient to show that

$$\sum_{i=0}^{j-1} \mu_i \mathrm{q}^i + \mathrm{q}^j \leq \sum_{i=0}^{h} \mu_i \mathrm{q}^i + \sum_{i=h+1}^{\infty} \mathrm{q}^i \tag{8}$$

This is obtained by easy computations:

$$\begin{aligned} \mathrm{q}^j &\leq \sum_{i=j}^{h} \mu_i \mathrm{q}^i + \sum_{i=h+1}^{\infty} \mathrm{q}^i \\ &= \sum_{i=j+1}^{\infty} \mathrm{q}^i = \frac{\mathrm{q}^{j+1}}{1-\mathrm{q}}. \end{aligned} \tag{9}$$

Finally,

$$\mathrm{q}^j \leq \frac{\mathrm{q}^{j+1}}{1-\mathrm{q}} \quad \Leftrightarrow \quad \mathrm{q} \geq \frac{1}{2}$$

and this proves (7). Now, we know that

$$\sum_{i=0}^{h} \mu_i \mathrm{q}^i \leq x \leq \sum_{i=0}^{h} \mu_i \mathrm{q}^i + \frac{\mathrm{q}^{h+1}}{1-\mathrm{q}} \tag{10}$$

and in the limit case $h \to \infty$, this becomes $x = \sum_{i=0}^{\infty} \mu_i \mathrm{q}^i$. In conclusion, $[0, \frac{1}{1-\mathrm{q}}] \subseteq \mathcal{X}$ and given that the opposite inclusion holds by definition, we have the equivalence

$$\mathcal{X} \equiv \left[0, \frac{1}{1-\mathrm{q}}\right]. \tag{11}$$

B. Case $\mathrm{q} \in (0, \frac{1}{2})$.

If $q < \frac{1}{2}$, $\mathcal{X}$ is a Cantor set. It can be constructed from the interval $\left[0, \frac{1}{1-q}\right]$ by deleting the elements that cannot be represented by the series, that is, the subintervals $\left(\frac{q}{1-q}, 1\right)$, $\left(\frac{q^2}{1-q}, q\right) \cup \left(1 + \frac{q^2}{1-q}, 1+q\right)$, etc. More precisely,

$$\mathcal{X} = \left[0, \frac{1}{1-\mathrm{q}}\right] - \sum_{m=0}^{\infty} \bigcup_{n=1}^{2^m} \left(\mathrm{p}_{m,n} + \frac{\mathrm{q}^{m+1}}{1-q},\ \mathrm{p}_{m,n} + \mathrm{q}^m\right) \tag{12}$$

where $\mathrm{p}_{m,1}, \ldots, \mathrm{p}_{m,2^m}$ are the binary polynomials in q of degree at most $m-1$ ($\mathrm{p}_{-1,1} = 0$ by convention).

Notice that $\sum_{i=0}^{\infty} \mu_i \mathrm{q}^i$ is a bijective map from $\{0,1\}^{\mathbb{N}}$ to $\mathcal{X}$ if $\mathrm{q} < \frac{1}{2}$. The surjectivity is obvious, while as far as the injectivity is concerned, suppose that $\sum \mu_n \mathrm{q}^n = \sum \nu_n \mathrm{q}^n$, $\mu_n, \nu_n \in \{0,1\}$ but with some different coefficients; for instance, let $\mu_n = \nu_n$ for $n = 0, \ldots, m-1$, $\mu_m = 0$ and $\nu_m = 1$ for some m. Since $q < \frac{1}{2}$, $\frac{q^{m+1}}{1-q} < q^m$, hence $\sum \mu_n \mathrm{q}^n < \sum \nu_n \mathrm{q}^n$: this proves that there cannot be two series with the same sum, but different coefficients.

The geometrical characterization of $\mathcal{X}$ strongly affects the performance of our deconvolution algorithm, that will be shortly introduced. Before that, we need to change our perspective on the problem, describing it in Information theoretic, probabilistic terms.

III. Information-Coding Theoretic Approach

Given the discrete nature of u, our deconvolution problem can be interpreted as a digital transmission paradigm. By definition (4), recovering $u(t)$ in $[0, T]$ corresponds to recovering the binary message $(u_0, \ldots, u_{K-1})$ on the basis of the received real sequence $(r_1, \ldots, r_K)$, which is the task performed by a decoder in a digital transmission [11]. Moreover, the sequence $(x_1, \ldots, x_K)$ can be thought as an *encoded* version of the input, with encoding rule imposed by the system itself and defined by (6); $\mathcal{X}$ is then the *code's alphabet*. Furthermore, $y(t) = cx(t)$ can be interpreted as a signal amplification and $r(t) = y(t) + n(t)$ as the passage of $y(t)$ through an additive-noise channel.

This Information-Coding setting suggests to exploit some decoding technique to perform deconvolution, instead of a classical input estimation method. In [8] this approach has been proved to be successful in the case $a = 0$, that is, when the convolution reduces to an integration; in this work, we extend those results to the general convolution framework, using a suitable low-complexity decoding algorithm that we will introduce in the next section.

As mentioned above, the encoding rule is determined by the system itself. This constitutes a difference between our problem and the classic coding paradigm, for which the encoding is expressly conceived to improve the reliability of the transmission. In our case, instead, we undergo the encoding, whose influence on the quality of the transmission actually depends on the parameters a, b and c (the theoretic analysis in Section V will assess this issue).

A. Probabilistic Setting

In order to complete the description of our transmission paradigm, we suppose that the input and the additive noise $n(t)$ are stochastically generated according to known probabilistic distribution laws, in particular:

Assumption 4: The additive noise $n(t)$ is white gaussian, that is, $n(k\tau)$ are realizations of independent gaussian random variables N_k, $k = 1, \ldots, K$, with null mean and variance σ^2.

Assumption 5: $u_0, \ldots, u_{K-1}$ are realizations of independent Bernoulli random variables $U_0, \ldots, U_{K-1}$ with parameter $\frac{1}{2}$.

Assumption 4 is very common in digital transmissions: the so-called AWGN (Additive White Gaussian Noise) channels are based on this model [11]. Also the assumption on the input source statistics is often required, but the distribution law can vary: in this case, we have chosen the easiest one, however more common situations (for instance, distributions based on Markov Chains) will be studied in further works.

The randomness introduced by the noise and the input involves all the system: also x, y and r actually are random variables. It is then more appropriate to rewrite the whole problem in probabilistic terms (capital letters will be used to indicate random variables):

$$\begin{aligned} &U_k \sim \text{ Bernoulli } \left(\frac{1}{2}\right), \quad k = 0, \ldots, K-1; \\ &X_0 = 0; \\ &X_k = \mathrm{q} X_{k-1} + \mathrm{w} U_{k-1}, \quad k = 1, \ldots, K; \\ &Y_k = c X_k, \quad k = 1, \ldots, K; \\ &R_k = Y_k + N_k, \quad N_k = \mathcal{N}(0, \sigma^2), \quad k = 1, \ldots, K. \end{aligned}$$

Finally, we will denote by $\widehat{U}_{k-1}$ the estimate of the bit U_{k-1}.

B. Performance analysis

In order to check the performance of any deconvolution algorithm applied to our problem, we have to quantify how *far* the estimated sequence $\widehat{\mathbf{U}} = (\widehat{U}_0, \dots, \widehat{U}_{K-1})$ is from the correct one $\mathbf{U} = (U_0, \dots, U_{K-1})$, that is, to define a suitable distance between $\widehat{\mathbf{U}}$ and $\mathbf{U}$ and calculate it. In our context, we choose the mean square cost $\bar{d}(\mathbf{U}, \widehat{\mathbf{U}})$, defined as:

$$\bar{d}(\mathbf{U}, \widehat{\mathbf{U}}) = \frac{1}{K} \sum_{k=0}^{K-1} \mathbb{E}(U_k - \hat{U}_k)^2 \tag{13}$$

which is equivalent to

$$\bar{d}(\mathbf{U}, \widehat{\mathbf{U}}) = \frac{1}{K} \sum_{k=0}^{K-1} \mathrm{P}(U_k \neq \hat{U}_k).$$

The computation of the distance will be analytically achieved in Section V.

IV. One State Algorithm

In this section, we present our decoding algorithm, that is, the One State Algorithm introduced in [8]. It is a low-complexity, recursive scheme derived from the well-known BCJR technique [1]. The BCJR computes the probabilities of all the possible transmitted codewords from the observation of the noisy output and then decodes the received sequence with the most probable transmitted one. This task is performed through a double recursive procedure that implements a maximum a posteriori estimation (MAP for short, see [11]),

The BCJR has been proved to be optimal. Nevertheless, its original version is off-line and requires a finite number of states, hence it cannot be applied to problems such as the one we aim to solve. The One State Algorithm has been developed for this scope: it is a simplified, single recursive version of the BCJR that stores only one state at each step and takes account only of the past and present information. We refer the reader to [8] for the complete discussion on its genesis, here we just briefly remind its pattern:

1) Initialize state estimate: $\widehat{x}_0 = 0$;
2) For $k = 1, \dots, K$, given the received symbol $r_k \in \mathbb{R}$, estimate the current bit and the current state:

$$\begin{aligned} &\widehat{u}_{k-1} = \begin{cases} 0 \text{ if } |r_k - c\mathrm{q}\widehat{x}_{k-1}| \leq |r_k - (c\mathrm{q}\widehat{x}_{k-1} + c\mathrm{w})| \\ 1 \text{ otherwise.} \end{cases} \\ &\widehat{x}_k = \mathrm{q}\widehat{x}_{k-1} + \mathrm{w}\widehat{u}_{k-1}. \end{aligned} \tag{14}$$

Given the estimation of the state, the decoder estimates the possible transmitted signals according to the dynamics of the system. As the input is binary, at each step we have just two possible signals and we decide between them evaluating the distance between them and the acquired output sample r_k.

Naturally, the more the possible signals are pairwise distant, the more the decoding is reliable, the distance being $|c\mathrm{w}|$.

V. Theoretic Analysis

This section is devoted to the analytic description of the One State Algorithm and evaluation of its performance, which will be accomplished applying results of the Theory of Markov Processes. In particular, we will apply methods already used in [8] and also introduce new techniques to study Markov Processes.

In order to compute $\bar{d}(\mathbf{U}, \widehat{\mathbf{U}})$, let us define the stochastic process

$$D_k = \widehat{X}_k - X_k = \mathrm{q}D_{k-1} + \mathrm{w}(\widehat{U}_{k-1} - U_{k-1}) \tag{15}$$

where $k = 0, \dots, K$ and $D_0 = 0$. If $D_{k-1} = d$, then $D_k \in \{\mathrm{q}d - \mathrm{w}, \mathrm{q}d, \mathrm{q}d + \mathrm{w}\}$. Moreover,

$$D_k \in \mathcal{D}_k = \mathrm{w} \left\{ \sum_{i=0}^{k-1} \alpha_i \mathrm{q}^i, \alpha_i \in \{-1, 0, 1\} \right\}. \tag{16}$$

The key point now is that if K is finite, $(D_k)_{k=0,\dots,K}$ is a Markov Chain (with finite state space), while if we let K tend to infinity, $(D_k)_{k\in\mathbb{N}}$ is a Markov Process on $(\mathcal{D}, \mathcal{B}(\mathcal{D}))$ where

$$\mathcal{D} = \mathrm{w} \left\{ \sum_{i=0}^{\infty} \alpha_i e^{ai}, \alpha_i \in \{-1, 0, 1\} \right\} \tag{17}$$

(the definitions of Markov Chains and Markov Processes we are referring to are those introduced in [8]). As we are interested in the long-time behavior of the system, in the next we will focus on the Markov Process $(D_k)_{k\in\mathbb{N}}$. In particular, by the study of $(D_k)_{k\in\mathbb{N}}$ we can evaluate $\bar{d}(\mathbf{U}, \widehat{\mathbf{U}})$ for large K. In fact,

$$\begin{aligned} &\frac{1}{K} \sum_{k=0}^{K-1} \mathrm{P}(\widehat{U}_k \neq U_k) = \\ &= \frac{1}{K} \sum_{k=0}^{K-1} \int_{\xi\in\mathcal{D}} \mathrm{P}(\widehat{U}_k \neq U_k | D_k = \xi) P^k(0, d\xi) \end{aligned} \tag{18}$$

where $P^k(a, A)$ is the k-step transition probability from a to the set A. In the next, we will indicate

$$g(\xi) = \mathrm{P}(\widehat{U}_k \neq U_k | D_k = \xi)$$

since given at any step k, the probability of incorrect detection only depends on the value of D_K. Then,

Theorem 1 (Ergodic Theorem): If $(D_k)_{k\in\mathbb{N}}$ admits a unique invariant probability measure (say a probability measure μ such that $\mu(A) = \mu P(A)$ for any $A \in \mathcal{B}(\mathcal{D})$; i.p.m. for short), then

$$\lim_{K\to\infty} \bar{d}(\mathbf{U}, \widehat{\mathbf{U}}) = \int_{\mathcal{D}} g(\xi) \mu(d\xi). \tag{19}$$

This result is a consequence of the Ergodic Theorem of Markov Processes (see, e.g., [7] and [9, Theorem 2.3.4 - Proposition 2.4.2]), and states that it is sufficient to verify the existence and the uniqueness of an i.p.m. for $(D_k)_{k\in\mathbb{N}}$ to assess the mean square cost for large K (g can be analytically computed, while μ and the integral in (19) can be numerically evaluated whenever μ exists and is unique).

Our next goal is then to prove the existence and uniqueness of an i.p.m. for $(D_k)_{k\in\mathbb{N}}$. This requires first the computation of the transition probability kernel of $(D_k)_{k\in\mathbb{N}}$, which turns out to be discrete, in the sense that $D_k = \xi$ implies $D_{k+1} \in \{\mathrm{q}\xi - \mathrm{w}, \mathrm{q}\xi, \mathrm{q}\xi + \mathrm{w}\}$ for any $k \in \mathbb{N}$.

For notational brevity, throughout this section let us suppose

$$c = 1.$$

By the definition of $(D_k)_{k\in\mathbb{N}}$, by the One State procedure and given the assumptions 1-5 we obtain the following transition probabilities (the details of the computation are omitted for brevity) for any $k \in \mathbb{N}$:

$$\begin{aligned} P(\xi, \mathrm{q}\xi + \mathrm{w}) &= \mathrm{P}(\widehat{U}_k = 1, U_k = 0 | D_k = \xi) \\ &= \frac{1}{4}\mathrm{erfc}\left(\frac{\mathrm{q}\xi + \mathrm{w}/2}{\sigma\sqrt{2}}\right) \\ P(\xi, \mathrm{q}\xi - \mathrm{w}) &= \mathrm{P}(\widehat{U}_k = 0, U_k = 1 | D_k = \xi) \\ &= \frac{1}{4}\mathrm{erfc}\left(\frac{-\mathrm{q}\xi + \mathrm{w}/2}{\sigma\sqrt{2}}\right) \\ P(\xi, \mathrm{q}\xi) &= 1 - P(\xi, \mathrm{q}\xi + \mathrm{w}) - P(\xi, \mathrm{q}\xi - \mathrm{w}). \end{aligned} \tag{20}$$

Moreover,

$$g(\xi) = P(\xi, \mathrm{q}\xi + \mathrm{w}) + P(\xi, \mathrm{q}\xi - \mathrm{w}). \tag{21}$$

Now, we can proceed to the proof of the main theorems of this work.

A. Main Results

Theorem 2 (Existence of an i.p.m.): The Markov Process $(D_k)_{k\in\mathbb{N}}$ admits at least one invariant probability measure.

Theorem 3 (Uniqueness of the i.p.m.): The conditions $\mathrm{q} < \frac{1}{3+\sqrt{\frac{2}{e\pi}}}$ and $\mathrm{q} > \frac{1}{3-\sqrt{\frac{2}{e\pi}}}$ are sufficient so that $(D_k)_{k\in\mathbb{N}}$ admits a unique i.p.m..

B. Proof of Theorem 2

First, we notice that the structure of $\mathcal{D}$ is analogous to the structure of $\mathcal{X}$ described in Section II: if $\mathrm{q} \geq \frac{1}{3}$, $\mathcal{D} = \left[-\frac{\mathrm{w}}{1-\mathrm{q}}, \frac{\mathrm{w}}{1-\mathrm{q}}\right]$ while if $\mathrm{q} < \frac{1}{3}$, the state space $\mathcal{D}$ is a Cantor space homeomorphic to $\{-1, 0, 1\}^{\mathbb{N}}$ with the natural product topology (the construction is the same of $\mathcal{X}$). In both cases $\mathcal{D}$ is compact, which along with the *weak-Feller* property guarantees the existence of an i.p.m. (see for example [9, Theorem 7.2.3]). A Markov Process is said to be weak-Feller if $Pf \in C_b(\mathcal{D})$ whenever $f \in C_b(\mathcal{D})$ (see [9, Paragraph 7.2.1]). In our case, if $f \in C_b(\mathcal{D})$, then $Pf(\xi) = \sum_{i\in\{-\mathrm{w},0,\mathrm{w}\}} P(\xi, \mathrm{q}\xi + i) f(\mathrm{q}\xi + i) \in C_b(\mathcal{D})$ since $P(\xi, \mathrm{q}\xi + i)$, $i \in \{-\mathrm{w}, 0, \mathrm{w}\}$, are continuous and bounded as functions of ξ. Hence, $(D_k)_{k\in\mathbb{N}}$ is weak-Feller and admits an i.p.m..

C. Proof of Theorem 3

Let μ and ν be two probability measures on $\mathcal{D}$ and consider the Wasserstein distance between measures defined by:

$$d(\mu, \nu) = \sup_{f\in 1\text{-Lip}(\mathcal{D})} \left(\int f d\mu - \int f d\nu\right). \tag{22}$$

Moreover, let us name T the operator

$$T : \mu \to T\mu = \mu P$$

on the space of the probability measures on $\mathcal{D}$. Now, can prove the following lemmas.

Lemma 1: The conditions of Theorem 3 are sufficient so that for any $f \in 1$ -Lip$(\mathcal{D})$, then $Pf \in h$-Lip$(\mathcal{D})$ with $h < 1$.

Proof Given any $f \in 1$-Lip$(\mathcal{D})$ and $\xi, \zeta \in \mathcal{D}$,

$$\begin{aligned} &Pf(\xi) - Pf(\zeta) \\ &= \sum_{i\in\{-\mathrm{w},0,\mathrm{w}\}} f(\mathrm{q}\xi + i)P(\xi, \mathrm{q}\xi + i) - f(\mathrm{q}\zeta + i)P(\zeta, \mathrm{q}\zeta + i) \\ &= f(\mathrm{q}\xi) + \sum_{i\in\{-\mathrm{w},\mathrm{w}\}} [f(\mathrm{q}\xi + i) - f(\mathrm{q}\xi)]P(\xi, \mathrm{q}\xi + i) + \\ &- f(\mathrm{q}\zeta) - \sum_{i\in\{-\mathrm{w},\mathrm{w}\}} [f(\mathrm{q}\zeta + i) - f(\mathrm{q}\zeta)]P(\zeta, \mathrm{q}\zeta + i). \end{aligned}$$

Adding and removing the quantity $[f(\mathrm{q}\xi + \mathrm{w}) - f(\mathrm{q}\xi)]P(\zeta, \mathrm{q}\zeta + \mathrm{w}) + [f(\mathrm{q}\xi - \mathrm{w}) - f(\mathrm{q}\xi)]P(\zeta, \mathrm{q}\zeta - \mathrm{w})$, using the Lipschtiz property of f and recalling that $P(\zeta, \mathrm{q}\zeta + \mathrm{w}) + P(\zeta, \mathrm{q}\zeta - \mathrm{w}) \leq \frac{1}{2}$ for any $\zeta \in \mathcal{D}$, we obtain

$$\begin{aligned} &Pf(\xi) - Pf(\zeta) = \\ &[f(\mathrm{q}\xi + \mathrm{w}) - f(\mathrm{q}\xi)][P(\xi, \mathrm{q}\xi + \mathrm{w}) - P(\zeta, \mathrm{q}\zeta + \mathrm{w})] \\ &\ + [f(\mathrm{q}\xi - \mathrm{w}) - f(\mathrm{q}\xi)][P(\xi, \mathrm{q}\xi - \mathrm{w}) - P(\zeta, \mathrm{q}\zeta - \mathrm{w})] \\ &\ + [f(\mathrm{q}\xi + \mathrm{w}) - f(\mathrm{q}\xi) - f(\mathrm{q}\zeta + \mathrm{w}) + f(\mathrm{q}\zeta)]P(\zeta, \mathrm{q}\zeta + \mathrm{w}) \\ &\ + [f(\mathrm{q}\xi - \mathrm{w}) - f(\mathrm{q}\xi) - f(\mathrm{q}\zeta - \mathrm{w}) + f(\mathrm{q}\zeta)]P(\zeta, \mathrm{q}\zeta - \mathrm{w}) \\ &\ + f(\mathrm{q}\xi) - f(\mathrm{q}\zeta) \leq \\ &\leq \mathrm{w}\big|P(\xi, \mathrm{q}\xi + \mathrm{w}) - P(\zeta, \mathrm{q}\zeta + \mathrm{w})\big| \\ &\ + \mathrm{w}\big|P(\xi, \mathrm{q}\xi - \mathrm{w}) - P(\zeta, \mathrm{q}\zeta - \mathrm{w})\big| + \mathrm{q}|\xi - \zeta| \end{aligned}$$

Hence,

$$\begin{aligned} |Pf(\xi) - Pf(\zeta)| &\leq \mathrm{w} \max_{z\in\mathcal{D}} \left|\frac{dP}{dz}(z, \mathrm{q}z + \mathrm{w})\right| |\xi - \zeta| + \\ &+ \left(\mathrm{w} \max_{z\in\mathcal{D}} \left|\frac{dP}{dz}(z, \mathrm{q}z - \mathrm{w})\right| + \mathrm{q}\right) |\xi - \zeta| \end{aligned} \tag{23}$$

provided that P is differentiable by definition. The derivatives are

$$\begin{aligned} \frac{d}{dz}P(z, \mathrm{q}z + \mathrm{w}) &= -\frac{\mathrm{q}}{2\sigma\sqrt{2\pi}} \exp\left(\frac{(\mathrm{q}z + \mathrm{w}/2)^2}{2\sigma^2}\right) \\ \frac{d}{dz}P(z, \mathrm{q}z - \mathrm{w}) &= \frac{\mathrm{q}}{2\sigma\sqrt{2\pi}} \exp\left(\frac{(\mathrm{q}z - \mathrm{w}/2)^2}{2\sigma^2}\right) \end{aligned}$$

The absolute values of the derivatives achieve their maximum in the endpoints of $\mathcal{D}$, respectively in $z = -\frac{\mathrm{w}}{1-\mathrm{q}} = \frac{b}{a}$ and $z = \frac{\mathrm{w}}{1-\mathrm{q}} = -\frac{b}{a}$, that is:

$$\begin{aligned} &\max_{z\in\mathcal{D}} \left|\frac{d}{dz}P(z, \mathrm{q}z + \mathrm{w})\right| = \max_{z\in\mathcal{D}} \left|\frac{d}{dz}P(z, \mathrm{q}z - \mathrm{w})\right| = \\ &= \frac{\mathrm{q}}{2\sigma\sqrt{2\pi}} \exp\left(\frac{(b/a)^2(1 - 3\mathrm{q})^2}{8\sigma^2}\right). \end{aligned}$$

By the inequality (23), a sufficient condition to obtain the thesis is

$$\frac{\mathrm{wq}}{\sigma\sqrt{2\pi}}\exp\left(\frac{(b/a)^2(1-3\mathrm{q})^2}{8\sigma^2}\right)+q<1. \quad (24)$$

Recalling that $\mathrm{w}=-\frac{b}{a}(1-\mathrm{q})$ and $\mathrm{q}\in(0,1)$, the inequality (24) is equivalent to

$$\mathrm{q}\frac{-b/a}{\sigma\sqrt{2\pi}}\exp\left(\frac{(b/a)^2(1-3\mathrm{q})^2}{8\sigma^2}\right)<1. \quad (25)$$

Let us consider the case $q\neq 1/3$. As the function te^{-t^2}, $t\in[0,+\infty)$ has global maximum at $t=\frac{1}{\sqrt{2}}$ where it assumes the value $\frac{1}{\sqrt{2e}}$, we have

$$\mathrm{q}\frac{-b/a}{\sigma\sqrt{2\pi}}e^{-\frac{(b/a)^2(1-3\mathrm{q})^2}{8\sigma^2}}\leq\frac{2\mathrm{q}}{\sqrt{\pi}|1-3\mathrm{q}|}\frac{1}{\sqrt{2e}}$$

Then, a sufficient condition to fulfill (25) is

$$\sqrt{\frac{2}{e\pi}}\frac{\mathrm{q}}{|1-3\mathrm{q}|}<1 \quad (26)$$

and finally, this corresponds to $\mathrm{q}<\frac{1}{3+\sqrt{\frac{2}{e\pi}}}$ if $q<\frac{1}{3}$ and to $\mathrm{q}>\frac{1}{3-\sqrt{\frac{2}{e\pi}}}$ if $q>\frac{1}{3}$. ■

Lemma 2: In the hypotheses of Lemma 1, T is a contraction, that is,

$$d(\mu P,\nu P)\leq h\ d(\mu,\nu),\quad h<1. \quad (27)$$

Proof Given a function f on $\mathcal{D}$,

$$\int f\ d(\mu P)=\int_{\xi\in\mathcal{D}}\int_{\zeta\in\mathcal{D}}f(\xi)P(\zeta,d\xi)\mu(d\zeta) = \int Pf\ d\mu$$

Hence

$$d(\mu P,\nu P)=\sup_{f\in 1\text{-Lip}(\mathcal{D})}\left(\int f\ d(\mu P)-\int f\ d(\nu P)\right) = \sup_{f\in 1\text{-Lip}(\mathcal{D})}\left(\int Pf\ d\mu-\int Pf\ d\nu\right).$$

By Lemma 1, if $f\in 1\text{-Lip}(\mathcal{D})$, then $Pf\in h\text{-Lip}(\mathcal{D})$, $h<1$. Then, for any probability measures μ,ν on $(\mathcal{D},\mathcal{B}(\mathcal{D}))$ and $f\in 1\text{-Lip}(\mathcal{D})$:

$$\int Pf\ d\mu-\int Pf\ d\nu\leq \leq\sup_{g\in h\text{-Lip}(\mathcal{D})}\left(\int g d\mu-\int g d\nu\right)=h\ d(\mu,\nu)$$

In particular,

$$d(\mu P,\nu P)= =\sup_{f\in 1\text{-Lip}(\mathcal{D})}\left(\int Pf\ d\mu-\int Pf\ d\nu\right)\leq h\ d(\mu,\nu)$$

with $h<1$. ■

At this point, we can conclude the proof of Theorem 3. In fact, under the required conditions, the operator T is a contraction in the space of the probability measures on $(\mathcal{D},\mathcal{B}(\mathcal{D}))$, then it admits a unique fixed point, i.e., there is a unique probability measure μ such that $T\mu=\mu$, or equivalently $\mu P=\mu$. In conclusion, μ is the unique i.p.m. for $(D_k)_{k\in\mathbb{N}}$.

D. Further conditions for the uniqueness of the i.p.m.

Theorem 3 does not assure the uniqueness when $\mathrm{q}\in\left[\frac{1}{3+\sqrt{\frac{2}{e\pi}}},\frac{1}{3-\sqrt{\frac{2}{e\pi}}}\right]$. However, for these values of q the inequality (25) holds whenever

$$\frac{-b/a}{\sigma}<6.43 \quad (28)$$

(this result has been obtained by numerical resolution of (25)). In other terms, the condition (28) is sufficient to achieve the uniqueness also for $\mathrm{q}\in\left[\frac{1}{3+\sqrt{\frac{2}{e\pi}}},\frac{1}{3-\sqrt{\frac{2}{e\pi}}}\right]$. The inequality (28) states that the diameter of $\mathcal{D}$ cannot be too larger than the noise variance. This has sense, since uniqueness is connected to the idea that the process can spread the state space, which is more likely when the space has small dimensions and the noise is considerable (for our process, a very small noise concentrates all the mass about 0, the center of $\mathcal{D}$).

Notice also that condition (28) is only sufficient and one can reasonably expect that uniqueness holds for any value of $\mathrm{q}\in(0,1)$, with no further conditions.

Given existence and uniqueness of the i.p.m. we can apply the Ergodic Theorem 1 to compute the mean square cost for large values of K and then to evaluate the performance of the One State algorithm. The integral in (19) can be numerically computed; in particular, one obtains an i.p.m. which is symmetric and with a maximum in the center of $\mathcal{D}$. Moreover, the value of the maximum increases as far as the noise decreases.

In the next section, we present a few significant simulations and the comparison between simulations' and theoretic results.

VI. A FEW SIMULATIONS

In the next, we report the outcomes of some simulations of our transmission system. Recalling the pattern of the One State Algorithm, notice that $|c\mathrm{w}|=|c\frac{b}{a}(1-\mathrm{q})|$, which represents the distance between the two possible transmitted signals at each step, plays a fundemental role. A larger value of $|c\mathrm{w}|$ is then desirable, since, as already mentioned, a larger distance improves the reliability of our estimation technique. On the other hand, $c^2\mathrm{w}^2$ can be interpreted as the energy per channel use of our transmission system, then for the applications its value cannot be increased too much.

In the next, we will represent the mean square cost $\bar{d}(\mathbf{U},\widehat{\mathbf{U}})$ in function of the so-called the Signal-To-Noise Ratio (SNR for short) of our transmission, that is, $c^2\mathrm{w}^2/\sigma^2$. This quantity represents the proportion between signal and noise energies and is usually assumed as reference parameter to study the quality of a transmission. A more detailed discussion about this point can be retrieved in [8], while in [11]

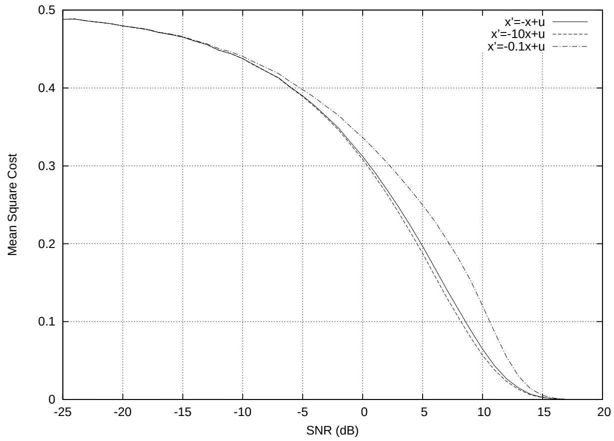

Fig. 1. Simulation with $b = c = 1$, $a = -1, -0.1, -10$.

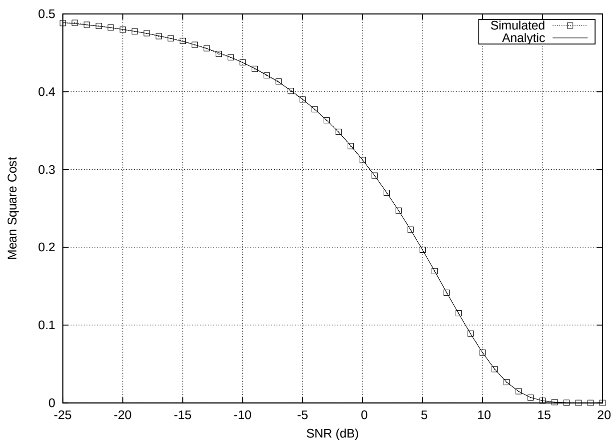

Fig. 2. Analytic vs Simulated Mean Square Cost (system with $b = c = 1$, $a = -1$): the results are consistent.

the reader can find the basic concepts of this information-theoretic issue.

Given that $c\mathrm{w}$ is the leading parameter, in the simulations we fix and $b = c = 1$, while a can vary. In Figure 1 the cases $a = -1$, $a = -0.1$ and $a = -10$ are represented: the graphs show the mean square cost $\bar{d}(\mathbf{U}, \widehat{\mathbf{U}})$ in function of $SNR = c^2\mathrm{w}^2/\sigma^2$ expressed in dB. We can notice that the performance improves (that is, the mean square cost decreases) as a decreases.

In Figure 2 we show a comparison between the mean square cost obtained by analytic computation and by the simulations in the case $b = c = 1$ and $a = -1$: the graphs are coincident. This is only an example, but a perfect consistency between simulated and analytic results has been observed to hold in every case.

VII. Conclusions

In this paper, the One State Algorithm has been introduced to solve the deconvolution problem for linear systems with quantized input. This algorithm, derived from the optimal BCJR decoding method, is very low-complexity. Moreover, its performance can be analytically evaluated in terms of a mean square cost through the theory of Markov Processes: indeed, the distance between the estimated and the correct state can be interpreted as a Markov Process, because of the recursive pattern of the One State Algorithm. In particular, for long-time transmissions an Ergodic Theorem can be applied to assess the mean square cost.

The theoretic analysis of the One State technique requires the proof of the existence and uniqueness of an invariant probability measure for a Markov Process. It is well known that uniqueness is difficult to prove when the process has not a positive probability of spreading the state space in one step; in this paper, we have undertaken the problem interpreting the i.p.m. as the fixed point of a contraction operator on the space of the probability measures. This approach allows to state sufficient conditions to have the uniqueness and can be applied even when the state space is not standard, for instance, when it is a Cantor set. This would be more difficult to achieve through the classic theorems of [7], [10], [9], which require a precise topological knowledge of the state space.

In the last part of this work, a few simulations have been presented and $\bar{d}(\mathbf{U}, \widehat{\mathbf{U}})$ has been computed for some instances of the problem. As far as the performance is concerned, notice that in [8] a more reliable method was individuated, the so-called Two States algorithm; nevertheless, this algorithm is more reliable only when the elements of $\mathcal{X}$ are isolated and equidistanced.

We finally remark that the One State Algorithm can be exploited in any stochastic framework, that is, with other input and noise distributions. Furthermore, it can be implemented also in multi-dimensional linear systems, which are of greater interest for the applications. In the latter case, the theoretic analysis via Markov Process can be analogously undertaken, but the dynamics of the process and the structure of the state space become more difficult to describe as far as the dimensions increase. This instance will be dealt with in some future work, in which also applications to the Fault Tolerant Control will be discussed.

References

[1] L. Bahl, J. Cocke, F. Jelinek and J. Raviv. Optimal decoding of linear codes for minimizing symbol error rate. IEEE Trans. on Information Theory, volume IT-20(2), pages 284–287, 1974.

[2] M.R. Banham and A.K. Katsaggelos. Digital image restoration. IEEE Signal Proc. Magazine, volume 14(2), pages 24–41, 1997.

[3] M. Bertero and P. Boccacci. Introduction to Inverse Problems in Imaging. Institute of Physics, Bristol and Philadelphia, 1998.

[4] C.L. Byrne and R.M. Fitzgerald. Reconstruction from partial information with applications to tomography. SIAM J. Appl. Math., volume 42(4), pages 933–940, 1982.

[5] F. Fagnani and L. Pandolfi. A singular perturbation approach to a recursive deconvolution problem. SIAM J. Control Optim., volume 40, pages 1384–1405, 2002.

[6] F. Fagnani and L. Pandolfi. A recursive algorithm for the approximate solution of Volterra integral equations of the first kind of convolution type. Inverse problems, volume 19, pages 23–47, 2003.

[7] S.R. Foguel. The ergodic theory of Markov Processes. Van Nostrand, Princeton, 1969.

[8] F. Fagnani, S. Fosson. Deconvolution of linear systems with quantized input: a coding theoretic viewpoint. Submitted to *Mathematics of Control, Signals, and Systems.* (MCSS), March 2009. Pre-Print: http://arxiv.org/abs/1001.3550.

[9] O. Hernández-Lerma and J.B. Lasserre. Markov Chains and Invariant Probabilities. Birkhauser-Verlag, Basel, 2003.

[10] S.P. Meyn and R.L. Tweedie. Markov Chains and Stochastic Stability. Springer-Verlag, London, 1993.

[11] T. Richardson and R. Urbanke. Modern Coding Theory. Cambridge University Press, 2008.

[12] G. Sparacino and C. Colbelli. A stochastic deconvolution method to reconstruct insulin secretion rate after a glucose stimulus. IEEE Trans. on biomedical engineering, volume 43(5), pages 512–529, 1996.

[13] J.L. Starck, E. Pantin and F. Murtagh. Deconvolution in Astronomy: A Review. Publications of the Astronomical Society of the Pacific, pages 1051-1069, 2002.

[14] A.N. Tikhonov. On the solution of ill-posed problems and the method of regularization. Soviet Math. Dokl., volume 4 , pages 1035–1038, 1963.

[15] A.N. Tikhonov and V.Y. Arsenin. Solution of ill-posed problems. Winston and Wiley, Washington, 1977.