Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Filtering and Parameter Estimation for a Class of Hidden Markov Models with Application to Bubble-Counting in Microfluidics ⋆

Francesco Carravetta * **Costanzo Manes** ** **Pasquale Palumbo** *

* *Istituto di Analisi dei Sistemi e Informatica A. Ruberti, CNR, Viale Manzoni 30, 00185 Roma, Italy (e-mail: francesco.carravetta@iasi.cnr.it, pasquale.palumbo@iasi.cnr.it).*
** *Dipartimento di Ingegneria e Scienze dell'Informazione e Matematica, Università degli Studi dell'Aquila, 67100 Coppito (AQ)-Italy (e-mail: costanzo.manes@univaq.it).*

Abstract: Motivated by a practical problem, in this work we investigate the problem of simultaneous estimation of state and parameters of an Hidden Markov Model with a particular structure. The motivating application is the problem of automatic counting of bubbles or droplets flowing into a microfluidic channel, where the noisy output of a photodiode has to be processed in order to detect the transit of bubbles. The goal is achieved through the recursive computation of a pseudo-max-likelihood estimate.

Keywords: Markov Models, Parameter Identification, System State Estimation

1. INTRODUCTION

Hidden Markov Models (HMM) are effectively used in many applications for modeling systems with discrete state space and observation space. Speech processing and recognition is one of the first applications where HMM have been successfully applied (see eg. Rabiner (1989)). More recent audio applications are listed in Pikrakis et al. (2006). In general, HMM have revealed useful in many pattern recognition problems, such as radar image classification, text processing, active learning, image segmentation, and others (see the recent paper Zhang et al. (2013) and references therein). Estimation problems oh HMM are investigated in Elliot et al. (1995), White and Carravetta (2011).

In this paper we consider the state estimation problem on a class of HMMs that we have developed for dealing with a specific microfluidic application. The problem is the automatic counting of bubbles or droplets flowing into a microfluidic channel. The microchannel is made of transparent material, and the transit of bubbles causes a change in the transparency of the fluid in the channel. A photodiode can be mounted on the external wall of the microchannel in order to detect the transit of bubbles by capturing the variations of luminosity. We have been inspired by the experimental setup described in Sapuppo et al. (2011) and Schembri et al. (2012)

Theoretically, three levels of light are detected: high level when there is no bubble, low level when the border of the bubble is in correspondence with the diode position (entering or outcoming bubble), intermediate level when the bubble is in transit. The problem is to elaborate an algorithm that processing the low-resolution and noisy output of the phothodiode correctly detects the passing bubbles or droplets.

In this paper the problem is formulated as a state and parameter estimation problem in a HMM, and the bubble detection goal is achieved through the recursive computation of a pseudo-max-likelihood estimate.

2. THE CLASS OF HMM AT ISSUE

We consider a physical system characterized by a finite number of states that evolve according to a finite state Continuous-Time Markov Chain (CTMC) model. We assume that sensor captures noisy and quantized measurements on the system at discrete times kT, $k \in \mathbb{N}$. The behavior of the CTMC at the discrete times kT can be modeled as a (discrete-time) Markov Chain (MC). Le $n \in \mathbb{N}$ denote the number of states of the MC, and let us encode the n states using the n vectors of the canonical basis of $\mathbb{R}^n$, i.e. $\{e_1, \dots, e_n\}$. Let $X_k \in \{e_1, \dots, e_n\}$ denote the state of the MC a time $k \in \mathbb{N}$. It is known that the MC behavior can be stochastically described by means of the vector $p_0 = [p_{0,1}, \dots, p_{0,n}]^T$ of the initial probabilities of the states, i.e. $p_{0,i} = \mathbf{P}\{X_0 = e_i\}$, and by the matrix of the conditional probabilities $\mathbf{P}\{X_{k+1} = e_i | X_k = e_j\}$, $i, j = 1, \dots, n$. As well known, Elliot et al. (1995), the

⋆ This work is partially supported by MIUR-PRIN 2009 grant N. 2009J7FWLX_002.

chain obeys a difference stochastic equation of the following kind:

$$X_{k+1} = A_k X_k + V_k, \tag{1}$$

where the $n \times n$ matrix $A_k = \{a_{ij}(k)\}_{i,j=1,\ldots,n}$ is such that $a_{ij}(k) = \mathbf{P}\{X_{k+1} = e_i | X_k = e_j\}$, and V_k is a white noise ($V_k = X_{k+1} - \mathrm{E}\{X_{k+1}|X_k\}$). We assume that at each state e_i, $i = 1,\ldots,n$ correspond an intensity L_i of a physical quantity to be measured. In the ideal case of measurements taken at discrete-times kT by a noisefree analog sensor, we would have the following output signal

$$S_k = LX_k, \quad \text{where } L = [L_1 \ \cdots \ L_n]. \tag{2}$$

Thus, the row vector L collects the discrete admissible output levels. The output levels L_i are not necessarily distinct (it can happen that $L_i = L_j$ for some pairs $(i,j) \in \mathbb{N}^2$). We assume that the signal $S(t)$, corrupted by an additive white noise sequence W_k, is measured by a coarse sensor, that provides a discretization of the noisy signal (quantizer). Let $m \in \mathbb{N}$ be the number of possible output values of the quantizer. We encode such finite output values of the quantizer by means of the canonical basis of R^m, denoted $\{f_1, \cdots, f_m\}$. Thus, the considered model of generation of the output sequence Y_k, $k \in \mathbb{N}$, is

$$Y_k = \mathcal{Q}(Z_k), \tag{3}$$

$$Z_k = S_k + W_k. \tag{4}$$

$\mathcal{Q}$ in (3) is the quantizer, i.e. $\mathcal{Q} : \mathbb{R} \to \{f_1,\ldots,f_m\}$. For a given partition $\{I_1,\ldots,I_m\}$ of $\mathbb{R}$ (i.e. such that $\mathbb{R} = \cup_{i=1}^m I_i$ and $I_i \cap I_j = \emptyset$, $\forall i \neq j \in \mathbb{N}$) the quantizer is such that:

$$\mathcal{Q}(z) = f_i, \quad \text{if} \quad z \in I_i \subset \mathbb{R}. \tag{5}$$

We assume that the set $\{I_i\}_{i=1}^m$ of subsets $I_i \in \mathbb{R}$, is a partition of $\mathbb{R}$, i.e. it is a set of m disjoint intervals as follows:

$$\begin{aligned} I_1 &= (-\infty, z_1], \quad I_m = (z_{m-1}, +\infty) \\ I_{i+1} &= (z_i, z_{i+1}], \qquad i = 1,\ldots,m-2. \end{aligned} \tag{6}$$

In (4), W_k is a zero mean, signal-independent, Gaussian noise, such that $\mathrm{E}\{W_k^2\} = \sigma^2$, for all $k \in \mathbb{N}$, and $\mathrm{E}\{W_k W_h\} = 0$, for $k \neq h$. Thus, the distribution of the noise W_k is $p_{W_k}(w) = (2\pi\sigma)^{-1/2} e^{-\frac{w^2}{2\sigma^2}}$.

We assume that the transition probability matrix A_k is constant (let's say A) but unknown, and that the m levels L_i are unknown too. The additive noise variance σ^2 is assumed known. Even the initial state distribution p_0 is unknown.

Fig. 1 represents schematically the states of the MC, the output levels L_i associated to the states, the additive noise and the quantization process.

The model given by eqs. (1) and (2)–(4) can be transformed into a standard HMM as follows. Consider the random variable Z_k (analog sensor noisy output), and let us denote with $p_Z(z; k|\mathcal{E})$ the probability density of Z_k, conditional to some event $\mathcal{E}$. As for $p_Z(z; k|X_k = e_i)$ one has, using the Gaussian assumption on the noise W_k:

$$\begin{aligned} p_{Z_k}(z|X_k = e_i) &= p_{Z_k}(z|S_k = L_i) = p_{W_k}(z - L_i) \\ &= \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(z-L_i)^2}{2\sigma^2}}. \end{aligned} \tag{7}$$

From this, we can compute all the conditional output probabilities:

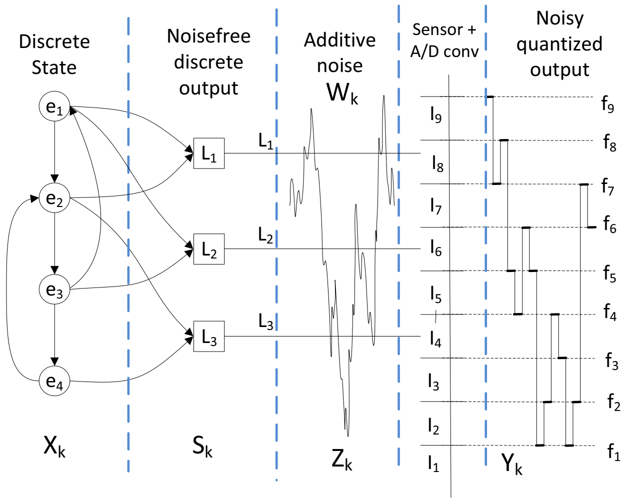

Fig. 1. Schematic of the HMM under investigation

$$\begin{aligned} c_{i,j}(L) &= \mathbf{P}\{Y_{k+1} = f_i | X_k = e_j\} \\ &= \int_{I_i} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(z-L_j)^2}{2\sigma^2}} dz. \end{aligned} \tag{8}$$

Also, we need to define the following quantity:

$$d_{i,j}(L) = \int_{I_i} \frac{z}{\sqrt{2\pi\sigma}} e^{-\frac{(z-L_j)^2}{2\sigma^2}} dz. \tag{9}$$

Note that from the definitions we have

$$\sum_{i=1}^m c_{i,j} = 1; \quad \sum_{i=1}^m d_{i,j} = L_j, \tag{10}$$

as the first summation results in the integral of a probability density over $\mathbb{R}$, and the second summation in the expectation of the probability density given in (7). Under the above setting the posterior observation satisfies the conditional independence property:

$$\mathbf{P}\{Y_{k+1},\ldots,Y_1 | X_k,\ldots,X_0\} = \prod_{i=0}^k \mathbf{P}\{Y_{i+1}|X_i\}. \tag{11}$$

and the observation satisfies:

$$Y_{k+1} = C(L)X_k + N_{k+1}, \tag{12}$$

where $C(L)$ is the $m \times n$ matrix collecting the coefficients $c_{i,j}$ defined in (8), and N_k is a zero mean, white sequence ($N_{k+1} = Y_{k+1} - \mathrm{E}\{Y_{k+1}|X_k\}$).

3. THE FILTERING AND IDENTIFICATION PROBLEMS

The problem we aim to solve is the optimal (in the least square sense) filtering of the MC $\{X_k\}$ under the discrete observations Y_k given in (3). Also some variables such as the *occupation times* of all states, or the *number of jumps* from i to j (for any pair $(i,j) \in \mathbb{N}^2$) are to be estimated. We want to solve this estimation problem without the knowledge of the transition matrix A in the model (1) or of the conditional output matrix C in (8)

The estimation task will be accomplished by simultaneously identificating the parameters $a_{i,j}$ (i.e. the matrix A) and the components of the row vector L (i.e, the levels L_i). Note that the choice of estimating the row vector

$L \in \mathrm{R}^{1\times m}$, and then computing $C(L)$ is more efficient that the choice of directly estimating the matrix $C \in \mathrm{R}^{m\times n}$.

Let $\Theta = \{p_0, A, L\}$, denote the set of unknown parameters that characterize the system (1) and (12)

The max-log-likelihood estimation of Θ, namely $\widehat{\Theta} = [\hat{p}_0,\, \widehat{A},\, \widehat{L}]^T$, is defined by:

$$\hat{p}_0, \widehat{A}, \widehat{L} = \arg\max_{p_0,A,L} \mathcal{L}(p_0, A, L) \tag{13}$$

$$\mathcal{L}(p_0, A, L) = \mathbf{E}_\Theta\{\log \mathbf{P}_\Theta\{Y_{k+1}, \ldots, Y_1, X_k, \ldots, X_0\}|\mathcal{Y}_k\}\,, \tag{14}$$

where $\mathbf{E}_\Theta, \mathbf{P}_\Theta$ are the expectation and probability corresponding to a given parameter's set of values and $\mathcal{Y}_k$ is the σ-algebra generated by $Y_k, \ldots, Y_0$. Notice that the estimate defined in (14) actually depends on k, and improves as k increases, with new observations.

We can obtain a significant computational simplification with a slight modification of the likelihood functional given in (14) as follows. let us denote $\Theta^{(k+1)}$ the parameters' set estimated at step k, i.e. with respect to $\mathcal{Y}_k$, and obtained by maximization of the functional $\mathcal{L}_k(p_0, A, L)$ defined as:

$$\mathcal{L}_k(p_0, A, L) = \mathbf{E}_{\Theta^{(k)}}\{\log \mathbf{P}_\Theta\{Y_{k+1}, \ldots, Y_1, X_k, \ldots, X_0\}|\mathcal{Y}_k\}\,, \tag{15}$$

with [1]

$$\mathcal{L}_0(p_0, A, L) = \mathbf{E}_{\Theta^{(0)}}\left(\mathbf{P}_{\Theta^{(0)}}(Y_1)\right). \tag{16}$$

We call the estimate $\Theta^{(k)}$ a *pseudo* max-likelihood (PML) estimate. In the following Theorem we show that the sequence of PML estimates $\{\Theta^{(k)}\}$ can be calculated recursively.

Before doing so, let us recall some known facts about optimal filtering of Markov chains that can be found in the literature. There are many ways indeed for build up an optimal filter of a Markov chain X under observations Y satisfying the conditional independence property (11). Bayesan methods allows the recursive computation of the joint probability $\mathbf{P}\{Y_{k+1}, \ldots, Y_1, X_k\}$ for given MC's transitions $\mathbf{P}\{X_{k+1}|X_k\}$ and posterior observations $\mathbf{P}\{Y_{k+1}|X_k\}$, from which both the LS as well as the MAP (maximum a posteriori) estimates can be obtained Elliot et al. (1995). Also, within the Bayesan framework, a smoothing algorithm, for the more general class of *reciprocal* chains (*including* the Markov case) can be found in White and Carravetta (2011). As for methods relying on a state-space approach, we here report the recursive algorithm for the computation of the *conditional distribution* (CD) of a MC X generated by a model as in (1), under the observation model (12), namely $p_k(e_i) = \mathbf{P}\{X_k = e_i|\mathcal{Y}_k\}$ (details and proofs can be found in Elliot et al. (1995)). The algorithm consists in finding first an *unnormalized conditional distribution* (UCD) namely $q_k(\cdot)$, which satisfies the recursive equations:

$$q_{k+1}(e_r) = m\sum_{j=1}^{n} a_{rj}\, q_k(e_j)(Y_{k+1}^T C_j), \tag{17}$$

[1] note that $\mathcal{L}_0$ is constant with respect to A

where C_j denotes the j-th column of the output matrix C, and then p_k by normalization: $p_k(e_i) = \frac{q_k(e_i)}{\sum_{i=1}^n q_k(e_i)}$. Equation (17) is the *optimal filter* of the HMM (1)-(12) provided the parameters are given. It is the finite-range, discrete-time- version of the well known Zakai equation for stochastic differential equations. In addition to the filter, we need to introduce three functionals of the MC $\{X_k\}$ that will play a role in the identification procedure: *Number of jumps in k steps* (of the MC X) from e_r to e_s, namely $\mathcal{I}_k^{r,s}$:

$$\mathcal{I}_k^{r,s} = \sum_{l=1}^{k} \langle X_{l-1}, e_r\rangle\langle X_l, e_s\rangle. \tag{18}$$

the *Occupation time* (of a state e_r, for the MC X), namely $\mathcal{O}_k^r$:

$$\mathcal{O}_{k+1}^r = \sum_{l=1}^{k+1} \langle X_{l-1}, e_r\rangle. \tag{19}$$

and the *state-to-observations transitions number*, namely $\mathcal{T}_k^{\beta,t}$:

$$\mathcal{T}_k^{\beta,r} = \sum_{l=1}^{k} \langle Y_l, f_\beta\rangle\langle X_{l-1}, e_r\rangle, \tag{20}$$

which is the number of times, up to time k, that Y is in the state β given the underlying MC X, at the preceding step, is in e_r. Notice that

$$\sum_{\beta=1}^{m} \mathcal{T}_k^{\beta,r} = \sum_{s=1}^{n} \mathcal{I}_k^{rs} = \mathcal{O}_k^r. \tag{21}$$

Let H_k be one of the above functionals $\mathcal{I}_k, \mathcal{O}_k, \mathcal{T}_k$. For known values of parameters, it is possible to compute recursively the least-square estimate $\widehat{H}_k = \mathbf{E}\{H_k|\mathcal{Y}_k\}$ by Elliot et al. (1995):

$$\widehat{H}_k = \mathbf{E}(H_k|\mathcal{Y}_k) = \langle \gamma_k(H_k), \mathbf{1}\rangle / \sum_{i=1}^{n} q_k(i), \tag{22}$$

$$\gamma_k(\mathcal{I}_k^{r,s}) = m\sum_{j=1}^{n} \left(Y_k^T C_j\right) \langle \gamma_{k-1}(\mathcal{I}_{k-1}^{r,s}), e_j\rangle A_j + m\left(Y_k^T C_r\right) \langle q_{k-1}, e_r\rangle a_{sr} e_s, \tag{23}$$

$$\gamma_k(\mathcal{O}_k^r) = m\sum_{j=1}^{n} \left(Y_k^T C_j\right) \langle \gamma_{k-1}(\mathcal{O}_{k-1}^r), e_j\rangle A_j + m\left(Y_k^T C_r\right) \langle q_{k-1}, e_r\rangle A_r. \tag{24}$$

$$\gamma_k(\mathcal{T}_k^{\beta,r}) = m\sum_{j=1}^{n} (Y_k^T C_j)\langle \gamma_{k-1}(\mathcal{T}_{k-1}^{\beta,r}), e_j\rangle A_j + m\langle q_{k-1}, e_r\rangle\langle Y_k, f_\beta\rangle c_{\beta,k} A_r. \tag{25}$$

where A_j denote (similarly to C_j) the j-th column of A.

Theorem 1. *The PML estimates $\hat{p}^0(k), \widehat{A}^{(k)}, \widehat{L}^{(k)}$ are given by the following recursive algorithm:*
step 0) $\widehat{\Theta}^{(0)} = \{\hat{p}_0^{(0)}, \hat{\theta}^{(0)}, \widehat{L}^{(0)}\}$ *is obtained by solving the discrete optimization problem:*

$$\widehat{\Theta}^{(0)} = \arg\max_{p_0,L} = \sum_{i=1}^{m}\sum_{j=1}^{n} C_{i,j}^2(L_j)\,(p_{0,j})^2\,, \tag{26}$$

$\widehat{A}^{(0)}$ *is set to any value.*

step k) *Let* $C = C(\widehat{L}^{(k)})$, *and* $a_{i,j} = \widehat{a}_{i,j}^{(k)}$, *in the right hand sides of (23)–(25), then we have:*

$$\widehat{a}_{sr}^{(k+1)} = \frac{\widehat{\mathcal{I}}_k^{rs}}{\widehat{\mathcal{O}}_k^r}, \quad \hat{p}_{0,j}^{(k+1)} = \mathrm{E}_k\{\langle X_0, e_j\rangle|\mathcal{Y}_k\}, \qquad (27)$$

whereas the estimate $\widehat{L}^{(k+1)}$ *is obtained by finding the zero of the function:*

$$\Phi(L) = L_r \widehat{\mathcal{O}}_k^r - \sum_{\beta=1}^m \frac{d_{\beta,r}(L)}{c_{\beta,r}(L)} \widehat{\mathcal{T}}_k^{\beta,r}, . \qquad (28)$$

Proof. (Step 0)- Since $\mathbf{P}\{Y_1\} = \mathbf{P}\{Y_1|X_0\}p_0^{(0)}$, by (16) it is

$$\mathcal{L}_0(p_0, L) = \sum_{i=1}^m \sum_{j=1}^n \mathbf{P}^2_{\Theta^{(0)}}\{Y_1 = f_i|X_0 = e_j\}\left(p_{0,j}\right)^2$$

from which (26) follows. Moreover, since $\mathcal{L}_0$ does not depend of A, the latter can be chosen at any initial value. (Step k). By using (11), and the Markov Property of X:

$$\mathbf{P}_\Theta\{Y_{k+1}, \dots, Y_1, X_k, \dots, X_0\}$$
$$= \mathbf{P}_\Theta\{X_0\}\prod_{l=1}^k \mathbf{P}_\Theta\{Y_l|X_{l-1}\}\mathbf{P}_\Theta\{X_l|X_{l-1}\}$$
$$= \prod_{r,s,j=1}^n \prod_{\beta=1}^m p_\alpha^{\langle X_0, e_\alpha\rangle} \prod_{l=1}^k c_{\beta,r}^{\langle Y_l, f_\beta\rangle\langle X_{l-1}, e_r\rangle} a_{sr}^{\langle X_l, e_s\rangle\langle X_{l-1}, e_r\rangle}$$

and taking the log and the conditional expectation $E_k(\cdot|\mathcal{Y}_k) = E_{\Theta^{(k)}}(\cdot|\mathcal{Y}_k)$:

$$\mathcal{L}(p_0, A, L) = \sum_{\beta=1}^m \sum_{r=1}^n \widehat{\mathcal{T}}_k^{\beta,r} \log c_{\beta,r}(t, L)$$
$$+ \sum_{r,s=1}^n \widehat{\mathcal{I}}_k^{r,s} \log a_{sr} + \sum_{\alpha=1}^n \mathbf{E}_k\{\langle X_0, e_\alpha\rangle|\mathcal{Y}_k\} \log p_{0,\alpha} \qquad (29)$$

The above functional has to be maximized with respect to the variables a_{sr}, $p_{0,\alpha}$, L_r, and θ_i, with $s, r, \alpha = 1, \dots, n$, $i = 1, \dots, q$, with the following $2n+1$ equality constraints (the latter two for $r = 1, \dots, n$):

$$\sum_{\alpha=1}^m p_{0,\alpha} = 1; \quad \sum_{s=1}^n a_{sr} = 1; \quad \sum_{\beta=0}^m c_{\beta,r}(L) = 1 \qquad (30)$$

Thus, let us introduce Lagrange multipliers: λ_r, μ, γ_r: relative to the constraints (30) and build up the Lagrangian:

$$\mathbf{L}(p_0, A, L, \lambda, \mu, \gamma) = \mathcal{L}(A, \theta, L) + \sum_{r=1}^n \lambda_r \left(\sum_{s=1}^n a_{sr} - 1,\right)$$
$$+ \mu\left(\sum_{\alpha=1}^n p_\alpha - 1\right) + \sum_{r=1}^n \gamma_r \left(\sum_{\beta=1}^m c_{\beta,r}(t, \theta, L) - 1\right), \qquad (31)$$

Differentiating in a_{sr} and setting to 0, we have

$$\frac{1}{a_{sr}} \widehat{\mathcal{I}}_k^{rs} + \lambda_r = 0, \qquad (32)$$

and since $\sum_{s=1}^n \widehat{\mathcal{I}}_k^{rs} = \widehat{\mathcal{O}}_k^r$, by using the second of (30) we can solve with respect to λ_r, and then derive the estimation at time k of a_{sr}, namely $a_{sr}^{(k+1)}$, which is given by the first of (27). Similarly, by differentiating with respect to p_α^0 we have $\mu + \mathbf{E}_k\{\langle X_0, e_\alpha\rangle|\mathcal{Y}_k\}/p_\alpha^0 = 0$, which, on account of the first of (30), and $\sum_\alpha \mathbf{E}_k\{\langle X_0, e_\alpha\rangle|\mathcal{Y}_k\} = 1$, gives $\mu = -1$, and then the second of (27). As for the remaining parameters, by differentiating (31) with respect to L_r and setting to zero, we obtain:

$$\sum_{\beta=1}^m \frac{1}{c_{\beta r}} \frac{\partial c_{\beta,r}}{\partial L_r} \widehat{\mathcal{T}}_k^{\beta r} + \gamma_r \sum_{\beta=1}^m \frac{\partial c_{\beta,r}}{\partial L_r} = 0, \qquad (33)$$

From (8) an easy calculation leads to the following identities:

$$\frac{\partial c_{\beta,r}}{\partial L_r} = \frac{1}{\sigma^2} d_{\beta,r} - \frac{1}{\sigma^2}(L_r + F(t, \theta)) c_{\beta,r}, \qquad (34)$$

By (34), using (10), one has

$$\sum_{\beta=1}^m \frac{\partial c_{\beta,r}}{\partial L_r} = 0. \qquad (35)$$

Substituting (35) and (34) in (33) results in

$$(L_r) \sum_{\beta=1}^m \widehat{\mathcal{T}}_k^{\beta,r} = \sum_{\beta=1}^m \frac{d_{\beta,r}}{c_{\beta,r}} \widehat{\mathcal{T}}_k^{\beta,r}.$$

On account of (21), the above condition amounts to finding a zero of the function defined in (28). •

4. A MARKOV CHAIN BASED MODEL OF BUBBLE-MOTION IN A MICRO CHANNEL

The problem is the automatic counting of bubbles or droplets flowing into a microfluidic channel. The microchannel is made of transparent material, and the transit of bubbles causes a change in the transparency of the fluid in the channel. A photodiode can be mounted on the external wall of the microchannel in order to detect the transit of bubbles by capturing the variations of luminosity. More details on this experimental setup can be found in Sapuppo et al. (2011) and Schembri et al. (2012).

4.1 Simulation results

The following numerical simulations have been carried out in order to show *in silico* the effectiveness of the proposed theoretical approach. A 4-state Markov chain is here proposed

- *absence of drop* (AD, in short), corresponding to a high value coming from the sensor
- *presence of drop* (PD, in short), corresponding to a low value coming from the sensor
- *border of entering drop* (BED, in short), corresponding to the lowest value coming from the sensor
- *border of outcoming drop* (BOD, in short), corresponding to the lowest value coming from the sensor

It is assumed that only the following change of states are allowed (bubbles never come back in our experiments)

$$AD \longrightarrow BED \longrightarrow PD \longrightarrow BOD \longrightarrow AD$$

with the following Markov probabilities (unknown to the filter)

$$\begin{array}{ll} P(BED|AD) = 0.05, & P(PD|BED) = 0.01, \\ P(BOD|PD) = 0.05, & P(AD|BOD) = 0.01 \end{array} \qquad (36)$$

By ordering the states as (AD, BED, BOD, PD), the Markov transition matrix is:

$$A = \begin{bmatrix} 0.99 & 0 & 0.05 & 0 \\ 0.01 & 0.95 & 0 & 0 \\ 0 & 0 & 0.95 & 0.01 \\ 0 & 0.05 & 0 & 0.99 \end{bmatrix} \quad (37)$$

As far as the measurements, we assume the four Markov states provide the sensor the following signals:

$$AD \mapsto 60, \qquad PD \mapsto 45, \qquad BED, BOD \mapsto 45$$

and the sensor provides a 32 levels quantized signal within the range $[0.75, 12]$. The values coming from sensors are affected by an additive noise, modeled by a zero-mean Gaussian distribution with standard deviation equal to the quantization bit 2.5.

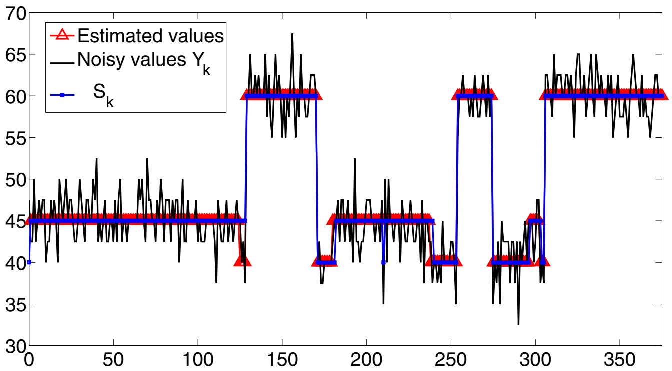

Fig. 2. Original Markov chain values and the output coming from the noisy digital sensor.

In order to test the performance of the identification algorithm, we assume the following initial estimate for matrix A:

$$\widehat{A}_0 = \begin{bmatrix} 0.5 & 0 & 0.5 & 0 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0.5 & 0 & 0.5 \end{bmatrix} \quad (38)$$

and an initial estimate for matrix L given by:

$$\widehat{L}_0 = \begin{bmatrix} 70 & 30 & 30 & 50 \end{bmatrix} \quad (39)$$

The proposed algorithm provides very nice estimates of the model parameters:

$$\widehat{A} = \begin{bmatrix} 0.9883 & 0 & 0.2 & 0 \\ 0.0117 & 0.9582 & 0 & 0 \\ 0 & 0 & 0.8 & 0.046 \\ 0 & 0.0418 & 0 & 0.9954 \end{bmatrix} \quad (40)$$

$$\widehat{L}_0 = \begin{bmatrix} 58.8635 & 39.1832 & 34.6028 & 43.8772 \end{bmatrix} \quad (41)$$

As far as the state estimate, we have about the 90% of success, as it can also appreciated by Fig.1.

In case of *a priori* knowledge of matrices A and L, the filtering algorithm improves the percentage of fitting to about 98%.

5. CONCLUSIONS

In this work we investigate the problem of simultaneous estimation of state and parameters of a class of Hidden Markov Models. The structure of the model is such to solve a specific problem in the area of microfluidic: the problem of detecting and counting the bubbles flowing in microchannels. A photodiode detects changes of intensity of light caused by the presence or not of a bubble. The problem is that, unfortunately, such changes are not very significant and do not alow to discriminate the presence of a bubble. Moreover there is a good amount of sensor noise, and the analog to digital conversion is made with a coarse quantization. Thus, it is not easy to detect the transit of bubbles with simple processing of the signal from the photodioe. For this reason, we investigate an estimator scheme, that perform on line identification of an HMM in order to improve the quality of the estimation. The proposed signal processing scheme is based on the recursive maximization of a pseudo-Log-likelihood function. Computer simulation provide satisfactory results. Future work could be the test of the algorithm on real data.

REFERENCES

R.J. Elliott, L. Aggoun, and J.B. Moore (1995). Hidden Markov Models - Estimation and Control. Springer, 1995.

Pikrakis, A., Theodoridis, S., and Kamarotos, D. (2006). "Classification of musical patterns using variable duration hidden Markov models," IEEE Trans. on Audio, Speech, and Language Processing, 14 (5), pp. 1795–1807.

L. Rabiner (1989). "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, 77 (2) pp. 257–286

F. Sapuppo, F. Schembri, L. Fortuna, A. Llobera and M. Bucolo (2011). "A polymeric micro-optical system for the spatial monitoring in two-phase microfluidics," *Microfluidics and NanoFluidics*, 12 (1-4) pp 165–174.

F. Schembri, F. Sapuppo, and M. Bucolo (2012). "Experimental classification of nonlinear dynamics in microfluidic bubbles flow," *Nonlinear Dynamics*, 67 (4) pp 2807–2819.

L.B. White, and F. Carravetta (2011). "Optimal Smoothing for Hidden Reciprocal Processes," *IEEE Trans. Automat. Contr.*, 56 (9), pp. 2156–2161.

H. Zhang, Q.M.J. Wu, and T.M. Nguyen (2013). "Modified student's t-hidden Markov model for pattern recognition and classification," *IET Signal Processing,* 7 (3), pp 219–227