

Correspondence

A Constrained MDP Approach to Dynamic Quantizer Design for HMM State Estimation

Nader Ghasemi and Subhrakanti Dey

Abstract—This paper is concerned with dynamic quantizer design for state estimation of hidden Markov models (HMM) using multiple sensors under a sum power constraint at the sensor transmitters. The sensor nodes communicate with a fusion center over temporally correlated flat fading channels modelled by finite state Markov chains. Motivated by energy limitations in sensor nodes, we develop optimal quantizers by minimizing the long term average of the mean square estimation error with a constraint on the long term average of total transmission power across the sensors. Instead of introducing a cost function as a weighted sum of our two objectives, we propose a constrained Markov decision formulation as an average cost problem and employ a linear programming technique to obtain the optimal policy for the constrained problem. Our experimental results assert that the constrained approach is quite efficient in terms of computational complexity and memory requirements for our average cost problem and leads to the same optimal deterministic policies and optimal cost as the unconstrained approach under an irreducibility assumption on the underlying Markov chain and some mild regularity assumptions on the sensor measurement noise processes. We illustrate via numerical studies the performance results for the dynamic quantization scheme. We also study the effect of varying degrees of channel and measurement noise on the performance of the proposed scheme.

Index Terms—Hidden Markov models, linear programming, Markov decision process, power control, state estimation.

I. INTRODUCTION

Due to recent advances in very large scale integration (VLSI) and wireless technology, miniature sensor nodes, deployed over a wide area, have been employed in a wide variety of applications ranging from security applications, military and civil surveillance, target tracking, environmental and health monitoring and control tasks, to name a few. Motivated by limited battery power and communication capability of sensor nodes, estimation at the fusion center may only rely on the set of quantized (e.g., binary) observations. One can achieve additional energy/bandwidth savings by employing a censoring scheme [1] whereby only the sensors with informative observations are allowed to transmit to the fusion center while others with uninformative observations stay inactive. In this paper, we do not consider any censoring schemes, we focus our attention to a dynamic binary quantization approach and we present an algorithm for optimal quantizer design for state estimation of hidden Markov

Manuscript received October 23, 2007; revised October 03, 2008. First published December 02, 2008; current version published February 13, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Pramod K. Varshney. This work was partially supported by the Australian Research Council.

The authors are with the ARC Special Research Centre for Ultra-Broadband Information Networks (CUBIN), an affiliated program of National ICT Australia (NICTA), Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Melbourne, VIC 3010, Australia (e-mail: {n.ghasemi@ee.unimelb.edu.au; s.dey@ee.unimelb.edu.au}).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSP.2008.2010381

models (HMMs) under an average total transmission power constraint across the sensors.

To date, extensive research have been reported in context of distributed estimation of state/parameter of *dynamical* systems. In [2] (single sensor) and [3] (multiple sensors), a heuristic criterion based on mutual information between the state and the sensor measurement sequences was numerically maximized to obtain near-optimal quantizer thresholds for state estimation of a hidden binary Markov chain. Moreover, recent works [4], [5] have employed a stochastic control approach in order to design optimal quantizers for general HMMs. While these results do not take into account other constraints such as channel noise and fading in wireless sensor systems, a more recent study [6] extends these works by considering a more realistic case of additive channel noise and wireless channel fading. This work has been further extended in [7] which considers a minimization of a tradeoff between the long term average of mean square estimation error and expected total power consumption using an *unconstrained* Markov decision process (MDP) formulation.¹ In [7] the constrained problem is transformed into a Lagrangian² problem which is then solved using a dynamic programming approach. Typical of most dynamic programming based algorithms, this approach is computationally demanding particularly for Markov chains with large state spaces and networks with a large numbers of sensors.

As an alternative approach to the problem considered in [7], in this paper, we address the issue of optimizing³ quantizer design at each sensor node, including the quantization threshold and the transmission power, subject to a fixed total transmission power (across all sensors) constraint in order to minimize the state estimation error at the fusion center using a *constrained* approach. In particular, we propose a constrained Markov decision process formulation⁴ for distributed state estimation of an underlying finite state discrete time Markov chain measured via multiple sensors in noise. We show how one can use linear-programming (LP) based techniques to find optimal policies for the quantization thresholds and transmit powers, for the constrained MDP (CMDP) problem. Comparing to the unconstrained scheme in [7], the benefit of our constrained approach is that it is more efficient both in terms of number of computations and memory requirements. In general, linear programming can be applied in problems with large number of variables. Moreover, in certain cases, the LP approach could also result in deterministic policies which have the advantage of lower implementation complexity than randomized policies. It is shown based on our experimental results that this is achieved under some mild assumptions on the measurement noise at the sensors and a standard irreducibility assumption on the underlying Markov chain. We assume that the number of required sensors is fixed, determined by the intended quality for the estimation error. One should note however that in cases

¹For a formulation using unconstrained MDP approach see also [8] for a dynamic power control problem in cellular CDMA systems.

²The Lagrangian is the combination of the original cost (i.e., mean square estimation error) to be minimized and the total power constraint weighted by some constant called Lagrange multiplier or tradeoff parameter.

³It is worth mentioning that, for numerical tractability, we later consider an optimization problem which is a variant of the original quantization problem in which the range space for the quantization threshold is restricted to a finite subset \mathcal{R} of the set of real values \mathbb{R} .

⁴For constrained MDP formulations, see also [9] and [10] for dynamic resource allocation problems.

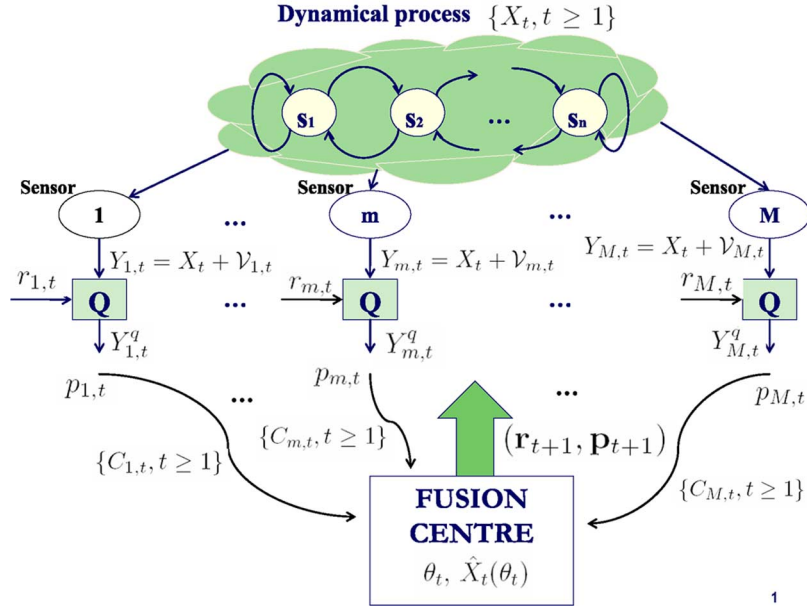


Fig. 1. Distributed setup for state estimation of an underlying dynamical process represented by a discrete time finite state Markov chain.

where a large scale sensor system is needed, our algorithms are still useful in situations where the fusion center plays the role of a powerful mobile agent which traverses the sensor network and at each dwell it interacts only with a few sensors.⁵

It is assumed that the optimal quantizer thresholds and power levels are determined at the fusion center (equipped with perfect channel information for all sensors) and communicated back to the sensors via noiseless feedback channels (of negligible delay) to be used for transmission of quantized measurements in the next time slot. In practice, under a channel reciprocity assumption (e.g., with a time-division duplex protocol), channel estimation can be carried out at the fusion center by periodically sending out pilot tones to the sensors with high power, so that the sensors can compute the channel gains and send them back to the fusion center as an overhead during data transmission. We do not consider power consumption incurred due to channel estimation in our analysis. This is justified by the fact that the fusion center usually has access to a replenishable power supply and the average power consumed by the sensors in communication of overhead information to the fusion center is constant and does not affect the quality of the estimation error at the fusion centre. Thus, given an average total power budget, one can easily subtract away the overhead average power cost at the sensors and then consider the leftover power budget as the new average total power budget across the sensors for transmission of quantized measurements for state estimation.

The paper is organized as follows. In Section II, we present the problem formulation of state estimation of a hidden Markov model using quantized measurements from multiple sensors with an average sum power constraint across the sensors. In Section III, we formulate the constrained optimization problem using a constrained approach for the presented model. We characterize the solutions to the constrained problem using a linear programming technique in order to find the optimal cost and also the optimal policies which determine the optimal quantization thresholds and power levels. We also evaluate the optimal error/power tradeoff curve as a way to designate the optimal value of one of the two criteria while the other one is determined to be less than or equal to some value of interest. Some comments on computational complexity of the constrained MDP approach as compared to

the unconstrained approach of [7] are also provided. The simulation results are given in Section IV followed by some concluding remarks in Section V.

II. PROBLEM STATEMENT

A. System Model

In this section, we present the problem formulation for state estimation of an underlying dynamical system modeled by a discrete time Markov chain. The state measurements are obtained from a sensor system consisting of M sensors. We consider a typical structure for a sensor system depicted in Fig. 1 in which we assume that the channels from the sensors to the fusion center are mutually independent error-prone bandlimited flat fading channels. For the dynamical process represented by a stationary ergodic Markov chain $\{X_t, t \geq 1\}$ with state space $\mathcal{X} = \{s_1, \dots, s_n\}$, the transition probability matrix is given by $\bar{\mathbf{P}}_{\mathcal{X}} = [\bar{p}_{ij}^{\mathcal{X}}]$, where $\bar{p}_{ij}^{\mathcal{X}} = Pr(X_{t+1} = s_j | X_t = s_i)$, $1 \leq i, j \leq n$. The state variable X_t is related to the measurements by $\mathbf{Y}_t = \mathbf{1}X_t + \mathbf{V}_t$, where $\mathbf{Y}_t = [Y_{1,t}, \dots, Y_{M,t}]^T$, $\mathbf{1} = [1, \dots, 1]^T$, and $\mathbf{V}_t = [V_{1,t}, \dots, V_{M,t}]^T$ is an i.i.d. additive noise sequence with joint probability density function $f_{\mathbf{V}}$.

We consider a threshold-based binary quantization⁶ as our quantization scheme in which the sequence $\{\mathbf{r}_t, t \geq 1\} = \{(r_{1,t}, \dots, r_{M,t}), t \geq 1\}$ represents the sequence of quantization parameters. Let the quantized data at time t be denoted by $\mathbf{Y}_t^q = [Y_{1,t}^q, \dots, Y_{M,t}^q]^T$. Each sensor transmits its quantized output to the fusion center over a discrete time flat fading channel. Let $\mathbf{C}_t = (C_{1,t}, \dots, C_{M,t})^T$ be the sensors' channel state vector at time t . We assume that the sequence of channel states $\{C_{m,t}, t \geq 1\}$ is a stationary ergodic Markov chain with state space $\mathcal{C} = \{c_1, \dots, c_k\}$ and transition probability matrix⁷ $\bar{\mathbf{P}}_{\mathcal{C}} = [\bar{p}_{ij}^{\mathcal{C}}]$,

⁶We refer to regular quantizers (convex quantization region) in which one may take any two distinct binary symbols a_1 and a_2 such that the events $\{Y_{m,t} < r_{m,t}\}$ and $\{Y_{m,t} \geq r_{m,t}\}$ are equivalent to $\{Y_{m,t}^q = a_1\}$ and $\{Y_{m,t}^q = a_2\}$ respectively. There are other possible forms of binary quantization, however, our objective here is to find the optimal quantization thresholds and transmit power levels for the given quantization scheme rather than finding the optimal quantization scheme.

⁷To simplify our subsequent analysis, we assume that $\bar{p}_{ij}^{\mathcal{C}} > 0$ for all $1 \leq i, j \leq k$.

⁵This architecture for large low power sensor networks called SENSOR NETWORKS WITH MOBILE AGENTS (SENMA) has been proposed in [11].

where $\bar{p}_{ij}^m = Pr(C_{m,t+1} = c_j | C_{m,t} = c_i)$, $1 \leq i, j \leq k$, $1 \leq m \leq M$. Each channel state c_i , $1 \leq i \leq k$ may represent a value of the channel gain. In the estimation problem, the channel states of different sensors are assumed to be known at the fusion center, but not at the sensor nodes—which is consistent with the notion of channel state information at the receiver (CSIR), a requirement that is standard in the wireless communications community. Denote the m th sensor's transmission power at time t by $p_{m,t}$. Let $\mathbf{p}_t = (p_{1,t}, \dots, p_{M,t})$ and define the sequence $\{\mathbf{p}_t, t \geq 1\} = \{(p_{1,t}, \dots, p_{M,t}), t \geq 1\}$. Henceforth \mathbf{p}_t will be called the power level parameter. Transmission powers for all sensors are chosen from the same set of available power levels $\mathcal{P} = \{p_1, \dots, p_\ell\}$, which is chosen to be a set of finitely many discrete power levels. This is motivated by the fact that in real applications, the feedback channel from the fusion center to the sensors have only limited bandwidth and hence a finite bit rate, thus being capable of feeding back the index of a finite size power codebook only. The received symbol at the fusion center (from the m th sensor) after decoding is denoted by $Y_{m,t}^f$. Let $\mathbf{Y}_t^f = [Y_{1,t}^f, \dots, Y_{M,t}^f]^T$. $Y_{m,t}^f$ is described by the following channel transition probability $Pr(Y_{m,t}^f = a_j | Y_{m,t}^q = a_i, C_{m,t} = c, p_{m,t} = p) := q_{ij}^m(c, p)$, where $i, j \in \{1, 2\}$, $c \in \mathcal{C}$, $p \in \mathcal{P}$, and a_1, a_2 denote the symbols for a binary-quantized measurement.

We assume that the sensors use binary phase shift keying (BPSK) to transmit their binary quantized measurements over orthogonal additive white Gaussian noise channels. The resulting input-output transition matrix is denoted by $\mathbf{Q}^m(c, p) = [q_{ij}^m(c, p)]$ conditioned on the channel state $c \in \mathcal{C}$ and power level $p \in \mathcal{P}$. The off-diagonal entries in $\mathbf{Q}^m(c, p)$ are called crossover probabilities. Let κ denote the path loss exponent of the wireless channel, and d_m be the distance between the m th sensor and the fusion center. Also, define γ_t^m as the signal-to-noise ratio (SNR) of the received signal at the fusion center from the m th sensor at time t . Then, $\gamma_t^m = g_{m,t} p_{m,t} / \sigma_\nu^2$, where σ_ν^2 is the variance of the additive white Gaussian channel noise $\nu_{m,t}$, and $g_{m,t}, p_{m,t}$ are the channel gain and transmission power respectively. We can then express the crossover probability after detection at the fusion center as $\varepsilon^m(\gamma_t^m) = Q(\sqrt{\alpha \gamma_t^m d_m^\kappa}) = \int_0^\infty \sqrt{\alpha \gamma_t^m d_m^\kappa} (1/\sqrt{2\pi}) e^{-(t^2/2)} dt$ with α being a constant. For more details on how to compute these crossover probabilities, see [7].

For each pair of sequences $\mathbf{r} = \{\mathbf{r}_t, t \geq 1\}$ and $\mathbf{p} = \{\mathbf{p}_t, t \geq 1\}$, the long term average of the mean square error in state estimation is given by

$$J(\mathbf{r}, \mathbf{p}) := \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E} |X_t - \hat{X}_t|^2 \quad (1)$$

where the state estimate \hat{X}_t is a Borel measurable function of the sequence $\{\mathbf{Y}_l^f, \mathbf{C}_l, l \leq t\}$. Note that the expectation in (1) is with respect to all sources of randomness, namely, the state X_t , the additive observation noise, the channel state, and channel noise. The objective of this paper is to solve the dynamic quantization optimization problem at the fusion center by obtaining the optimal sequences \mathbf{r} and \mathbf{p} such that the mean square estimation error $J(\mathbf{r}, \mathbf{p})$ is minimized constrained on the long term average of total power consumption across all the sensors. These optimal values of $\mathbf{r}_t, \mathbf{p}_t$ are then fed back to the sensors to be used for transmission at the next time step. This optimization problem can be formulated as

$$\min_{\mathbf{r}, \mathbf{p}} J(\mathbf{r}, \mathbf{p}), \text{ subject to: } \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E} Q(\mathbf{p}_t) \leq P_{\text{avg}} \quad (2)$$

where P_{avg} is the maximum allowable per-stage sum power, and $Q(\mathbf{p}_t)$ is a power consumption cost function defined as $Q(\mathbf{p}_t = \mathbf{v}) :=$

$\sum_{m=1}^M v_m$ and $\mathbf{v} = [v_1, \dots, v_M]^T$, $v_m \in \mathcal{P}$ is the value for the transmission power vector to be used at the next time step. Notice that the fusion center can not find the optimal quantization scheme using the optimization problem (2) by directly minimizing the cost function $J(\mathbf{r}, \mathbf{p})$. This is because it does not have exact knowledge on X_t due to the problem being a partially observed MDP problem. Therefore, in the following, we define the information state θ_t , which can be computed at the fusion center using a recursive formula, and then in Section III we restate the definition of the cost function $J(\mathbf{r}, \mathbf{p})$ based on the information state, which essentially converts the partially observed MDP to a fully observed MDP in terms of the information state space. Define the information state vector $\theta_t = [\theta_{1,t}, \dots, \theta_{n,t}]^T$, where $\theta_{i,t} = Pr(X_t = s_i | \mathcal{D}_t)$, $i \in \{1, \dots, n\}$, $t \geq 1$ and $\mathcal{D}_t := \mathcal{D}(\mathbf{Y}_l^f, \mathbf{C}_l, l \leq t)$ is the σ -algebra generated by the fusion center observations up to time t . The recursion for the information state θ_t is given by

$$\begin{aligned} \theta_t(\mathbf{r}_t, \mathbf{p}_t) &= \frac{p(\mathbf{c}_t | \mathbf{c}_{t-1})}{v_t'} \Psi(s_1, \dots, s_n, \mathbf{r}_t, \mathbf{p}_t, \mathbf{y}_t^f, \mathbf{c}_t) \bar{\mathbf{P}}_{\mathcal{X}}^T \theta_{t-1} \\ &:= \frac{1}{v_t} \Lambda(\mathbf{r}_t, \mathbf{p}_t, \mathbf{y}_t^f, \mathbf{c}_t) \theta_{t-1} \end{aligned} \quad (3)$$

where $\bar{\mathbf{P}}_{\mathcal{X}}$ is the transition matrix of $\{X_t\}$, v_t' and v_t are normalizing factors such that $\|\theta_t\|_1 = 1$, and $p(\mathbf{c}_t | \mathbf{c}_{t-1})$ is the product of M channel transition probabilities defined as $p(\mathbf{C}_t = \mathbf{c}_t | \mathbf{C}_{t-1} = \mathbf{c}_{t-1}) = \prod_{m=1}^M \bar{p}_{i_m j_m}^m$ in which $c_{m,t-1} = c_{i_m}$, $c_{m,t} = c_{j_m}$ for $i_m, j_m \in \{1, \dots, k\}$. Ψ is a diagonal matrix with i th diagonal entry called state-to-observation probability being defined as $\psi_i(s_i, \mathbf{r}_t, \mathbf{p}_t, \mathbf{y}_t^f, \mathbf{c}_t) = Pr(\mathbf{Y}_t^f = \mathbf{y}_t^f | X_t = s_i, \mathbf{r}_t, \mathbf{C}_t = \mathbf{c}_t, \mathbf{p}_t)$.⁸ It is obvious that the state space for the information state θ can be defined as the simplex $\mathcal{T}_\theta = \{\Omega \in \mathbb{R}_+^n | \|\Omega\|_1 = 1\} \subset \mathbb{R}^n$. See [7] for further details on deriving the recursion (3) for the information state.

For numerical tractability, we discretize the continuum information state θ using the discretization procedure presented in [4]. We choose a step size $1/N$ and approximate the continuum information state $\theta \in \mathcal{T}_\theta$ by discretized values $\vartheta \in \mathcal{T}$, where $\mathcal{T} \subset \mathcal{T}_\theta$ is the state space of the discretized information state. Note that after discretization, we still have that $\|\vartheta\|_1 = 1$, $\forall \vartheta \in \mathcal{T}$. Also, in further analysis for numerical tractability, we restrict⁹ the range space for the variable \mathbf{r} to a finite set \mathcal{R}^M of discrete values in \mathbb{R}^M . For notational and computational simplicity, the same finite subset $\mathcal{R} \subset \mathbb{R}$ is employed for optimizing each entry $r_{m,t}$ in \mathbf{r}_t . Note that the *optimal* choice of the set \mathcal{R} is a key factor in further minimizing the optimal value of the state estimation error. More on how to optimize the set \mathcal{R} can be found later at the end of the next section.

III. OPTIMAL QUANTIZATION

A. Constrained MDP Formulation

As discussed earlier, in [7], we solved an unconstrained MDP formulation of the constrained optimization problem (2) using a weighted combination of the two cost functions involving the average estimation error and the average sum power across the sensors. This unconstrained version was solved using a relative value iteration algorithm—see [7] for further details. Motivated by the need to reduce computational complexity, in this section, we formulate a variation (due to the restriction of the quantizer threshold space to a finite discrete set) of the constrained optimization problem (2) as a *constrained* expected average cost MDP problem in which the state estimation error is minimized subject to a constraint on the average total power consumption per time step

⁸See [7] for details on deriving the state-to-observation probabilities.

⁹The restriction to a finite set is by no means less appealing, in fact, without such an assumption on the set of thresholds (and also on the set of power levels), the wireless feedback channels from the fusion center to the sensors must have infinite capacity.

across the sensors. In order to proceed, assume that the quantization and power level parameters $(\mathbf{r}_t, \mathbf{p}_t)$ at each time instance t are specified by a stationary randomized Markov policy $\mu^\infty = \{\mu, \mu, \dots\}$ which specifies a probability distribution $q_{\mu^\infty(\vartheta, \mathbf{C})}(\mathbf{r}, \mathbf{p})$ on the action space $\mathcal{A} = \mathcal{R}^M \times \mathcal{P}^M$ and determines the probability of applying action (\mathbf{r}, \mathbf{p}) when the system occupies the state (ϑ, \mathbf{C}) . In fact, a deterministic policy μ_d which specifies the action $(\mathbf{r}_t, \mathbf{p}_t) = \mu_d(\vartheta_{t-1}, \mathbf{C}_{t-1})$ as a function of the discretized information state ϑ_{t-1} and the channel state \mathbf{C}_{t-1} may be considered as a special case of a randomized policy in which the probability distribution on the set of actions is degenerate [12]. We restrict our attention to Markov policies while seeking optimal policies, because for each distribution of the initial state $(\vartheta^\circ, \mathbf{c}^\circ)$ and any history-dependent policy, there exists a Markov policy with the same expected average cost (cf. [12, Theorem 8.1.2]). For brevity, henceforth, μ^∞ may be referred to as μ .

We now present the following optimization problem which can be solved to obtain the optimal stationary policy μ^* for minimizing the average estimation error subject to the average sum power constraint. The constrained optimization problem can be expressed as

$$\begin{aligned} \min_{\mu} J(\mu) &:= \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}^\mu \left[\sum_{t=1}^T \phi(\vartheta_t(\mu(\vartheta_{t-1}, \mathbf{C}_{t-1}))) \right] \\ &\quad \vartheta_1 = \vartheta^\circ, \mathbf{C}_0 = \mathbf{c}^\circ \\ \text{subject to :} \\ J^s(\mu) &:= \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}^\mu \left[\sum_{t=1}^T \mathcal{Q}(\mathbf{p}_t = \mu(\vartheta_{t-1}, \mathbf{C}_{t-1})) \right] \\ &\quad \vartheta_1 = \vartheta^\circ, \mathbf{C}_0 = \mathbf{c}^\circ] \leq P_{\text{avg}} \end{aligned} \quad (4)$$

where $\phi(\theta_t)$ is the conditional estimation error cost defined as $\mathbb{E}[|X_t - \hat{X}_t|^2 | \mathcal{D}_t] = \sum_{i=1}^n [s_i - \sum_{j=1}^n s_j \theta_{j,t}]^2 \theta_{i,t}$. Also, $(\vartheta^\circ, \mathbf{c}^\circ)$ is the initial condition, ϑ_1 is assumed to be independent of the action (\mathbf{r}, \mathbf{p}) and the initial action $(\mathbf{r}_1, \mathbf{p}_1)$ is determined independent of the policy μ . The optimization problem (4) is, indeed, an average cost constrained MDP problem with finite state space $\mathcal{S} = \mathcal{T} \times \mathcal{C}^M$ and finite action space $\mathcal{A} = \mathcal{R}^M \times \mathcal{P}^M$ with (ϑ, \mathbf{C}) , and (\mathbf{r}, \mathbf{p}) being the state and action respectively.

It has been established¹⁰ in [12] that for finite state and finite action unichain models, given any policy, there exists a stationary randomized policy with the same limiting average state-action frequency. The limiting average state-action frequency corresponding to a policy is defined as the steady state probability that the system occupies state (ϑ, \mathbf{C}) and action (\mathbf{r}, \mathbf{p}) is applied under that policy. In fact, it is a probability measure over the set of state-action pairs and it has the property that both of the long-run expected average costs $J(\mu)$ and $J^s(\mu)$ corresponding to the policy μ can be expressed as the expectation of the immediate costs $\phi(\cdot)$ and $\mathcal{Q}(\cdot)$ with respect to this measure.

We now present an LP formulation which is equivalent to the constrained problem (4) with a finite set of decision variables ξ and a finite set of constraints. The constrained optimization problem (4) can be related to the following equivalent dual linear programming (LP) problem

$$\begin{aligned} \min_{\xi} \mathcal{L}(\xi) &:= \sum_{\mathbf{y}^f, \mathbf{c}} \sum_{\substack{\mathbf{r} \in \mathcal{R}^M \\ \mathbf{p} \in \mathcal{P}^M}} \phi(\vartheta(\mathbf{y}^f, \mathbf{c}, \mathbf{r}, \mathbf{p})) \xi(\vartheta, \mathbf{c}, \mathbf{r}, \mathbf{p}) \\ \text{subject to :} \\ \sum_{\mathbf{y}^f, \mathbf{c}} \sum_{\substack{\mathbf{r} \in \mathcal{R}^M \\ \mathbf{p} \in \mathcal{P}^M}} \mathcal{Q}(\mathbf{c}, \mathbf{p}) \xi(\vartheta, \mathbf{c}, \mathbf{r}, \mathbf{p}) &\leq P_{\text{avg}} \end{aligned}$$

¹⁰cf. [12, Theorem 8.9.4].

$$\begin{aligned} \sum_{\substack{\mathbf{r} \in \mathcal{R}^M \\ \mathbf{p} \in \mathcal{P}^M}} \xi(\vartheta^f, \mathbf{c}^f, \mathbf{r}, \mathbf{p}) &= \sum_{\mathbf{y}^f, \mathbf{c}} \sum_{\substack{\mathbf{r} \in \mathcal{R}^M \\ \mathbf{p} \in \mathcal{P}^M}} p(\mathbf{c}^f | \mathbf{c}) \\ &\quad \times \left\| \mathbf{A}(\mathbf{r}, \mathbf{p}, \mathbf{y}^f, \mathbf{c}^f) \vartheta \right\|_1 \xi(\vartheta, \mathbf{c}, \mathbf{r}, \mathbf{p}), \forall (\vartheta^f, \mathbf{c}^f) \in \mathcal{S} \\ \sum_{\mathbf{y}^f, \mathbf{c}} \sum_{\substack{\mathbf{r} \in \mathcal{R}^M \\ \mathbf{p} \in \mathcal{P}^M}} \xi(\vartheta, \mathbf{c}, \mathbf{r}, \mathbf{p}) &= 1 \\ \xi(\vartheta, \mathbf{c}, \mathbf{r}, \mathbf{p}) &\geq 0, \quad \forall (\vartheta, \mathbf{c}) \in \mathcal{S}, \forall (\mathbf{r}, \mathbf{p}) \in \mathcal{A}. \end{aligned} \quad (5)$$

It has been established in [12] that every feasible solution to the dual LP problem (5) corresponds to a randomized stationary policy in the constrained MDP problem (4). Furthermore, it has been shown that the dual variable $\xi(\vartheta, \mathbf{c}, \mathbf{r}, \mathbf{p})$ equals the limiting average state-action frequency. However, the dual LP problem is feasible if and only if there is a stationary optimal policy as the solution to the constrained problem. The optimal randomized policy $\mu^{*\infty}$ can be determined using the optimal basic feasible solution ξ^* to the LP problem (5). The following theorem establishes the existence of solutions to the constrained unichain average cost MDP problem (4).

Theorem 3.1: Suppose there exists a solution ξ^* to the dual LP problem. Then there exists an optimal stationary randomized policy μ^* for the constrained MDP problem, where μ^* satisfies

$$q_{\mu^*(\vartheta, \mathbf{c})}(\mathbf{r}, \mathbf{p}) = \frac{\xi^*(\vartheta, \mathbf{c}, \mathbf{r}, \mathbf{p})}{\sum_{\substack{\mathbf{r}' \in \mathcal{R}^M \\ \mathbf{p}' \in \mathcal{P}^M}} \xi^*(\vartheta, \mathbf{c}, \mathbf{r}', \mathbf{p}')}, \quad \text{if } \sum_{\substack{\mathbf{r}' \in \mathcal{R}^M \\ \mathbf{p}' \in \mathcal{P}^M}} \xi^*(\vartheta, \mathbf{c}, \mathbf{r}', \mathbf{p}') > 0 \quad (6)$$

where $q_{\mu^*(\vartheta, \mathbf{c})}(\mathbf{r}, \mathbf{p})$ denotes the probability that randomized policy μ^* takes action (\mathbf{r}, \mathbf{p}) when the system is in the state (ϑ, \mathbf{c}) . Otherwise, $q_{\mu^*(\vartheta, \mathbf{c})}(\mathbf{r}, \mathbf{p})$ takes an arbitrary value if $\sum_{\substack{\mathbf{r}' \in \mathcal{R}^M \\ \mathbf{p}' \in \mathcal{P}^M}} \xi^*(\vartheta, \mathbf{c}, \mathbf{r}', \mathbf{p}') = 0$. This means that in this case, an action which transfers the MDP to the recurrent class of states, defined as $\{(\vartheta, \mathbf{c}) \in \mathcal{S} : \sum_{\substack{\mathbf{r} \in \mathcal{R}^M \\ \mathbf{p} \in \mathcal{P}^M}} \xi^*(\vartheta, \mathbf{c}, \mathbf{r}, \mathbf{p}) > 0\}$, is chosen. \square

Remark: This theorem is essentially an adaptation of the standard existence theorem for optimal policies for the constrained unichain average cost MDPs. (cf. [12, Theorem 8.9.6]). The gain λ of the optimal policy μ^* , which is the value of the optimal average estimation error subject to a constraint on the average sum power, is computed as the value of LP objective function \mathcal{L} evaluated at the optimal feasible solution ξ^* , that is, $\lambda = \mathcal{L}(\xi^*)$.

We introduce the following assumption:

A1) For any $\mathbf{r} \in \mathcal{R}^M$, $\mathbf{p} \in \mathcal{P}^M$, $\mathbf{y}^f \in \{a_1, a_2\}^M$, and $\mathbf{c} \in \mathcal{C}^M$, the matrix $\mathbf{A}(\mathbf{r}, \mathbf{p}, \mathbf{y}^f, \mathbf{c})$ is non-singular and strictly positive.¹¹ \square

Under assumption **A1)**, all stationary policies have a single recurrent class and no transient states. This means that the transition probability matrix of every stationary policy is irreducible. In fact, under assumption **A1)**, basic feasible solutions to the dual LP problem (5) correspond to *deterministic* stationary policies.¹² This is because, under assumption **A1)**, the discretized constrained optimization problem (4) forms a recurrent (or ergodic) MDP for stationary policies. Therefore, based on Theorem 3.1 and under assumption **A1)**, one can conclude that for

¹¹Notice that **A1)** holds under very mild conditions for the noise. For the example of two sensors ($M = 2$), **A1)** holds for non-singular and positive $\bar{\mathbf{P}}_{\mathcal{X}}$ and any i.i.d. noise sequence $\{\mathcal{V}_t : t \geq 1\}$ such that for any $\Omega > 0$, each of the four events $\{\mathbf{V}_t \in \mathbb{X}_{-\Omega} \times \mathbb{X}_{-\Omega}\}$, $\{\mathbf{V}_t \in \mathbb{X}_{-\Omega} \times \mathbb{X}_{\Omega}\}$, $\{\mathbf{V}_t \in \mathbb{X}_{\Omega} \times \mathbb{X}_{-\Omega}\}$, and $\{\mathbf{V}_t \in \mathbb{X}_{\Omega} \times \mathbb{X}_{\Omega}\}$, for $t \geq 1$ has strictly positive probability, where $\mathbb{X}_{-\Omega} = (-\infty, -\Omega]$ and $\mathbb{X}_{\Omega} = [\Omega, \infty)$. This criterion is obviously satisfied for any two dimensional non-degenerate Gaussian noise distribution.

¹²cf. [12, Corollary 8.8.3].

a given value of $P_{\text{avg}} = P^\circ$, there exists a bounded optimal basic feasible solution to the LP problem (5) which corresponds to a *deterministic* stationary policy μ^* that is average optimal.¹³ Define the expected long-term average estimation error with the optimal policy μ^* as

$$\bar{E}^{\mu^*} = \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E}^{\mu^*} [\phi(\vartheta_t(\mathbf{r}_t, \mathbf{p}_t))]$$

and similarly the expected long-term average sum power as

$$\bar{P}^{\mu^*} = \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E}^{\mu^*} [Q(\mathbf{p}_t)]$$

where again the initial action $(\mathbf{r}_1, \mathbf{p}_1)$ is independent of the policy μ^* . It is clear that \bar{E}^{μ^*} is the minimum average estimation error such that the average power \bar{P}^{μ^*} is less than P° . Define $E^*(P^\circ) = \bar{E}^{\mu^*}$, where $E^*(P)$ is defined as the minimum average estimation error such that the average power is less than P . $E^*(P)$ can be referred to as the optimal error/power curve. This curve is optimal in that no quantization scheme can result in a error/power point below this curve. By varying P_{avg} , we can find the optimal policy μ^* and the corresponding average error \bar{E}^{μ^*} for each value which is a point on the optimal error/power curve.

Now we briefly provide some details on the computational complexity of the unconstrained approach of [7] and the CMDP approach presented in this paper. Let $|\cdot|$ denote the cardinality of a set. The number of operations required per iteration in order to obtain a solution to the problem using the LP method is of order $O(mn)$, where $m = |\mathcal{S}| + 2 \approx |\mathcal{S}|$ (when $|\mathcal{S}|$ is large) is the number of constraints in the LP problem (5), and $n = |\mathcal{S}| \times |\mathcal{A}| + 1 \approx |\mathcal{S}| \times |\mathcal{A}|$ is the number of decision and slack variables.¹⁴ In the unconstrained approach, the number of operations per iteration needed to find the optimal policy and optimal cost using relative value iteration algorithm is $O(|\mathcal{S}|^2 \times |\mathcal{A}|^2 + 2|\mathcal{S}|)$. Note that the relative value iteration needs at least 100 iterations in order to converge to the optimal cost with a reasonable accuracy whereas the LP method converges within less than 20 iterations. This certifies the fact that LP approach is computationally superior to the unconstrained approach particularly for large state and action spaces.

Optimal design of the set \mathcal{R} : The optimal solution μ^* to the constrained problem (4) determines the optimal value of the quantization thresholds \mathbf{r} from the finite set \mathcal{R}^M . Therefore, choosing the set $\mathcal{R} = \{\rho_1, \dots, \rho_d\}$, *optimally* rather than arbitrarily, plays a major role in improving the performance of the dynamic quantization system. In order to obtain the optimal set \mathcal{R}^* for a given value $P_{\text{avg}} = P^\circ$ of the sum power upper bound, we may consider the dynamic quantization system as a scalar-valued map function $\mathcal{F} : \mathbb{R}^d \mapsto \mathbb{R}_+$ in which we have $\lambda = \mathcal{F}(\eta; P^\circ)$, where λ is the gain of ϵ -optimal¹⁵ policy found from the solution to the LP problem (5) for a given set of quantization thresholds η . Here, η may be viewed as a vector in \mathbb{R}^d , where d is the cardinality of the set \mathcal{R} . We have that the measurements of the objective function \mathcal{F} are available at any value of the parameter η . However, no direct measurements (with or without noise) of the gradient of $\mathcal{F}(\cdot)$ are

¹³cf. [12, Corollary 8.8.4].

¹⁴Note that only one slack variable is required which is the one for the power constraint inequality.

¹⁵that is, for an $\epsilon > 0$, a policy μ_ϵ^* with the property that $\lambda^{\mu_\epsilon^*}(\vartheta^\circ, \mathbf{c}^\circ) + \epsilon > \lambda^\mu(\vartheta^\circ, \mathbf{c}^\circ)$, for all $(\vartheta^\circ, \mathbf{c}^\circ) \in \mathcal{S}$ and all randomized history-dependent policies μ . Here, for a policy μ , λ^μ is referred to as the gain of policy μ and it denotes the lim sup average estimation error defined by $\lambda^\mu(\vartheta^\circ, \mathbf{c}^\circ) = \limsup_{T \rightarrow \infty} (1/T) \sum_{t=1}^T \mathbb{E}^\mu[\phi(\vartheta_t(\mathbf{C}_{t-1}, \mathbf{p}_t)) | \vartheta_1 = \vartheta^\circ, \mathbf{C}_0 = \mathbf{c}^\circ]$ constrained on the long term average sum power defined by $\limsup_{T \rightarrow \infty} (1/T) \sum_{t=1}^T \mathbb{E}^\mu[Q(\mathbf{p}_t) | \vartheta_1 = \vartheta^\circ, \mathbf{C}_0 = \mathbf{c}^\circ] \leq P_{\text{avg}}$.

assumed available. Therefore, the method we apply is a gradient-free stochastic optimization algorithm called adaptive stochastic approximation (ASP) studied in [13]. Further specific implementation details can be found in [13]. Moreover, the strong asymptotic convergence results for this algorithm has been established in [13]. See also [7] for details on the implementation of this stochastic optimization algorithm for the unconstrained Lagrangian based version of the problem (4).

IV. SIMULATION STUDIES

In this section, we present some numerical examples illustrating the tradeoff between the average estimation error and the average total power consumption achieved by the proposed dynamic quantization algorithm. The simulations are performed under various fading statistics, wireless channel noise, and measurement noise of the sensors. Throughout the following experiments, some of the parameters are assumed fixed unless otherwise mentioned: the step size $1/N$ in discretization of the information state θ is chosen $1/N = 0.02$; κ , the path loss exponent of the wireless channel, is considered $\kappa = 2$ for the deployment of the sensors in an open rural area; and the constant coefficient α for computing crossover probabilities is taken to be $\alpha = 2$.

The experiments are performed for both constrained and unconstrained approaches. The results for the unconstrained scheme proposed in [7] are obtained using a relative value iteration algorithm which determines the optimal deterministic policies and the corresponding optimal costs. On the other hand, linear programming (LP) is employed to find optimal policies for the constrained model. The optimal solution ξ^* to the LP problem (5) is obtained using interior-point methods. Then, using Theorem 3.1, we relate basic feasible solutions of the dual LP problem to stationary randomized policies for the constrained MDP problem. It was evident from the experiments that, under assumption **A1**, the set of states in the recurrent class where optimal randomized policy is applied is the null set. Consequently both the constrained and the unconstrained approaches result in the same optimal deterministic policy and the same optimal cost. Our simulations demonstrated that the highest relative difference between the optimal cost obtained from LP and the one from relative value iteration is 10^{-8} . In general, the benefit of using the unconstrained approach lies in the fact that it results in deterministic policies which involve less complexity in implementation than randomized policies. Whereas, in the context of the dynamic quantization problem, this benefit has been also achieved by our constrained MDP approach under assumption **A1**). Moreover, our experimental results assert that the LP approach is more efficient with regard to computational complexity and it is faster than the relative value iteration approach. Furthermore, the LP approach is also more suitable to problems where there are several constraints, since each additional constraint increments the rank of the basis matrix in the dual LP problem only by one. Therefore, in the context of the dynamic quantization for optimal estimation problem, one may prefer the constrained MDP approach to the unconstrained scheme in other cases, for example, when there are individual average power constraints for each sensor rather than only one average sum power constraint for all the sensors.

A. State Estimation of a Two-State Markov Chain

In this section, we examine state estimation of a two-state Markov chain $\{X_t, t \geq 1\}$ using two sensors. The sensors are located at different distances from the fusion center with distance vector $d = [90.0, 165.0]^T$, where the distance figures are given in meters. For the Markov chain $\{X_t\}$, the state space is $\mathcal{X} = \{s_1 = 1.1, s_2 = 3.2\}$ and the transition probability matrix is given by $\bar{\mathbf{P}}_{\mathcal{X}} = \begin{bmatrix} 0.78 & 0.22 \\ 0.44 & 0.56 \end{bmatrix}$. The measurement noise of the sensors are assumed to be independent and normally distributed with zero mean and variances $\sigma^2 = [0.45, 0.3]^T$.

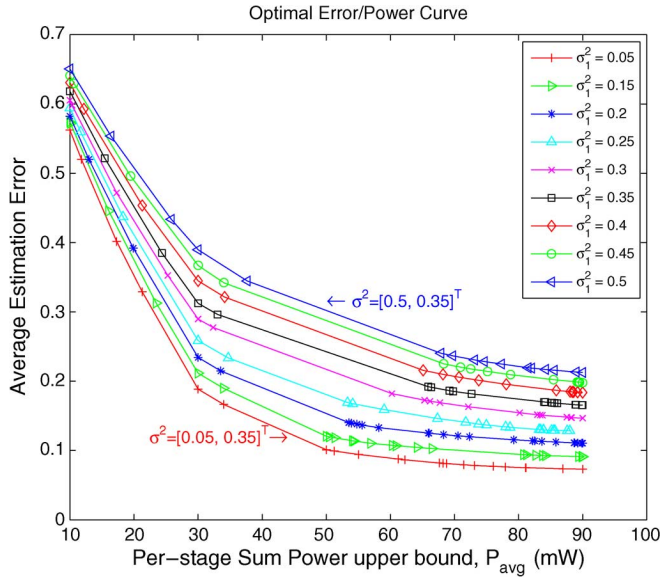


Fig. 2. Optimal error/power curves for different values of the measurement noise variance.

The wireless flat fading channels from the sensors to the fusion center are assumed to be independent and each channel is modeled by a two state Markov chain with state space $\mathcal{C} = \{c_1, c_2\}$. We consider asymmetric channels for the two sensors, that is, the channels have different fading statistics given by the following transition probability matrices $\mathbf{P}_C^1 = \begin{bmatrix} 0.62 & 0.38 \\ 0.52 & 0.48 \end{bmatrix}$, $\mathbf{P}_C^2 = \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix}$. The channel states c_1 and c_2 represent the corresponding channel gains $g_1^2 = 2 \times 10^{-10}$ and $g_2^2 = 4 \times 10^{-11}$ respectively. The noise power of the wireless channel for every sensor is $\sigma_v^2 = 2 \times 10^{-16} W$. The action space of the power level parameter for every sensor is considered as $\mathcal{P} = \{45, 25, 5\}$, where the power levels are given in mW. It is worth mentioning that the chosen power levels are typical values, and the devised algorithm can be applied for any technology-specific values of power levels.

In order to study the effect of the measurement noise variance σ^2 on the optimal error/power curve $E^*(P)$, we found the long term average of both state estimation error and sum power (i.e., \bar{E}^{μ^*} and \bar{P}^{μ^*} respectively) by generating a sample path¹⁶ using Monte Carlo simulations for the ϵ -optimal policy μ^* found from the optimal solution to the LP problem. Fig. 2 illustrates the optimal error/power curves $E^*(P)$ for various values of measurement noise variance in the range of $0.05 \leq \sigma_1^2 \leq 0.5$. For different values of long run per-stage sum power upper bound P_{avg} along every error/power curve $E^*(P)$, the optimal set of quantization thresholds \mathcal{R}^* with cardinality $d = 4$ has been found using the gradient-free stochastic optimization method ASP.

In Fig. 2, for the same average power, a trend of consistent decrease of the estimation error with the decrease of the measurement noise variance can be observed. Moreover, in order to achieve the same estimation error, larger transmission powers are needed at the sensor nodes for higher values of the measurement noise. This is due to the fact that higher power level results in higher received SNR at the fusion center which then leads to a lower crossover probability that compensates for the highly noisy sensors' observations.

Similarly, Fig. 3 demonstrates the influence of noise power of the wireless channel on the optimal error/power curve. It can be seen in Fig. 3 that the increase in the channel noise power increases the estimation error for the same value of power level. The reason is because higher channel noise gives lower SNR at the fusion center and thus

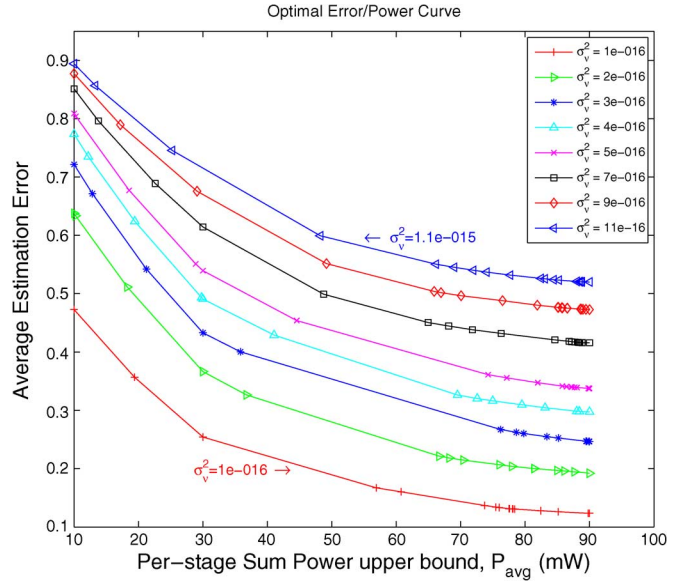


Fig. 3. Optimal error/power curves for different channel noise variances.

higher crossover probability which leads to more uncertainty in the state estimate $\hat{X}_t(\theta_t)$. In order to achieve the same estimation error, the effect of a high channel noise can be directly compensated also by using a higher transmit power level than the one required in a low channel noise condition.

V. CONCLUSION AND DIRECTIONS FOR FURTHER RESEARCH

In this paper, we have analyzed optimal quantization and power allocation in an HMM state estimation problem to optimize average estimation error constrained on average sum power consumption. We formulated the problem as a constrained Markov decision process problem. In the constrained approach, a linear programming technique has been employed to find optimal policies of the constrained unichain model. In this work, we considered the wireless channel from each sensor to the fusion center as a time-varying flat fading channel which is modeled by independent ergodic finite state Markov chains. Under some mild assumptions on the measurement noise of the sensors, the discretized version of the optimization problem forms a recurrent MDP for stationary policies. The solution to the discretized problem provides optimal quantization thresholds and power levels. The effect of the channel noise and the measurement noise on the optimal error/power curve is shown by simulation results. In order to improve the performance of the optimal quantization system, we employed a gradient-free stochastic optimization technique to determine the optimal set of quantization thresholds.

Finally, we indicate some future directions as extensions of this study. However, it is worth mentioning that these open issues may not be necessarily addressed using the same tools and methods presented in this manuscript. One can consider the problem of finding the optimal number of sensors required to achieve some specified estimation error and power consumption budget. Moreover, in this context, one can study the asymptotic behavior of the estimation error as the number of sensors grows. Furthermore, in this paper, we assumed that the fusion center has a perfect feedback channel to the sensors. However, models which relax this assumption also need to be considered. One can also address similar issues in scenarios that extend beyond the model studied in this paper, in particular, to distributed tracking problems where the underlying model is a jump Markov linear system rather than just a finite state Markov chain.

¹⁶For further description about generating sample paths see [7].

REFERENCES

- [1] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: A low-communication-rate scheme for distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 554–568, Apr. 1996.
- [2] L. Shue, S. Dey, B. Anderson, and F. De Bruyne, "On state-estimation of a two-state hidden Markov model with quantization," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 202–208, Jan. 2001.
- [3] F. A. Galati and S. Dey, "Information theoretic quantiser design for decentralised estimation of hidden Markov models," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Hong Kong, Apr. 2003, vol. 6, pp. 749–752.
- [4] M. Huang and S. Dey, "Dynamic quantizer design for hidden Markov state estimation via multiple sensors with fusion center feedback," *IEEE Trans. Signal Process.*, vol. 54, no. 8, pp. 2887–2896, Aug. 2006.
- [5] M. Huang and S. Dey, "Distributed state estimation for hidden Markov models by sensor networks with dynamic quantization," in *Proc. Intelligent Sensors, Sensor Networks, Information Processing Conf.*, Melbourne, Australia, Dec. 2004, pp. 355–360.
- [6] M. Huang and S. Dey, "Dynamic quantization for multisensor estimation over bandlimited fading channels," *IEEE Trans. Signal Process.*, vol. 55, no. 9, pp. 4696–4702, Sep. 2007.
- [7] N. Ghasemi and S. Dey, "Power-efficient dynamic quantization for multisensor hmm state estimation over fading channels," in *Proc. IEEE Int. Symp. Communications, Control, Signal Processing (ISCCSP)*, Mar. 12–14, 2008, pp. 1553–1558.
- [8] J. F. Chamberland and V. V. Veeravalli, "Decentralized dynamic power control for cellular CDMA systems," *IEEE Trans. Wireless Commun.*, vol. 2, no. 3, pp. 549–559, May 2003.
- [9] D. V. Djonin and V. Krishnamurthy, "MIMO transmission control in fading channels—a constrained Markov decision process formulation with monotone randomized policies," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 5069–5083, Oct. 2007.
- [10] A. K. Karmokar, D. V. Djonin, and V. K. Bhargava, "Optimal and suboptimal packet scheduling over correlated time varying flat fading channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 2, pp. 446–456, Feb. 2006.
- [11] L. Tong, Q. Zhao, and S. Adireddy, "Sensor networks with mobile agents," in *Proc. IEEE Military Communications Conf. (MILCOM)*, Boston, MA, Oct. 2003, vol. 1, pp. 688–693.
- [12] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York: Wiley, 1994.
- [13] J. C. Spall, "Adaptive stochastic approximation by the simultaneous perturbation method," *IEEE Trans. Autom. Control*, vol. 45, no. 10, pp. 1839–1853, Oct. 2000.

A Fast Robust Recursive Least-Squares Algorithm

Leonardo Rey Vega, Hernán Rey, Jacob Benesty, and Sara Tressens

Abstract—We present a fast robust recursive least-squares (FRRLS) algorithm based on a recently introduced new framework for designing robust adaptive filters. The algorithm is the result of minimizing a cost function subject to a time-dependent constraint on the norm of the filter update. Although the characteristics of the exact solution to this problem are known, there is no closed-form solution in general. However, the approximate solution we propose is very close to the optimal one. We also present some theoretical results regarding the asymptotic behavior of the algorithm. The FRRLS is then tested in different environments for system identification and acoustic echo cancellation applications.

Index Terms—Acoustic echo cancellation, impulsive noise, recursive least-squares algorithm, robust filtering, system identification.

I. INTRODUCTION

The recursive least-squares algorithm has the ability to solve the least-squares estimation problem recursively. Through its link with Kalman estimation [1], it can lead to the optimal estimate in the mean-square error sense. However, this is based on the assumption that the error signal e_i , which is by definition the difference between the system and model filter outputs, is Gaussian. In real-world environments, this assumption can be false. Perturbations such as background and impulsive noise can deteriorate the performance of many adaptive filters under a system identification setup. In echo cancellation, double-talk situations can also be viewed as impulsive noise sources. The performance of the RLS can be significantly deteriorated in these cases.

Several algorithms have been proposed attempting to overcome this issue [2]–[5]. In this work, we use a recently introduced new framework for the construction of robust adaptive filters [6] in order to design a robust RLS algorithm. Throughout this correspondence, the term robust will be used as "slightly sensitive to large perturbations (outliers)".

Particularly, we use a universal cost function introduced in [7] that preserves the system estimate from the effect of impulsive noise (or double talk) through the memory factor in the classical RLS cost function. Then, we propose to optimize this function subject to a constraint on the norm of the adaptive filter update. However, the exact solution to this problem has no closed form. Therefore, we propose an approximate solution to the problem. This solution is actually very close to the optimal one.

Manuscript received February 25, 2008; revised November 07, 2008. First published December 09, 2008; current version published February 13, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Roberto Lopez-Valcarce. This work was supported in part by the Universidad de Buenos Aires under project UBACYT I005.

L. Rey Vega is with the Department of Electronics and CONICET, University of Buenos Aires, Buenos Aires 1063, Argentina (e-mail: lrey@fi.uba.ar).

H. Rey is with the Instituto de Ingeniería Biomédica (FIUBA) and CONICET, Buenos Aires 1063, Argentina (e-mail: hrey@fi.uba.ar).

J. Benesty is with the INRS-EMT, Université du Québec, Montréal, QC H5A 1K6, Canada (e-mail: benesty@emt.inrs.ca).

S. Tressens is with the Department of Electronics, University de Buenos Aires, Buenos Aires 1063, Argentina (e-mail: stres@fi.uba.ar).

Color versions of one or more of the figures in this correspondence are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSP.2008.2010643