KALMAN-LIKE STATE TRACKING AND CONTROL IN POMDPS WITH APPLICATIONS TO BODY SENSING NETWORKS

Daphney-Stavroula Zois *Marco Levorato* *Urbashi Mitra*

Ming Hsieh Department of Electrical Enginering
University of Southern California
Los Angeles, CA 90089–2565, USA
Emails: {zois, levorato, ubli}@usc.edu

ABSTRACT

In this paper, the problem of state tracking with controlled observations is considered for a system modeled by a discrete-time, finite-state Markov chain. The system state is 'hidden' and observed via conditionally Gaussian measurements that are shaped by the underlying state and an exogenous control input. Following an innovations approach, a Kalman-like filter is derived to estimate the Markov chain system state. To optimize the control strategy, the associated mean-squared error is used as an optimization criterion for a partially observable Markov Decision Process (POMDP). The optimal solution is determined via stochastic dynamic programming. Numerical results are presented for the application of physical activity detection in heterogeneous, wireless body area networks.

***Index Terms*—** Markov models, approximate MMSE, POMDP, stochastic dynamic programming, innovations approach

1. INTRODUCTION

In a plethora of applications, a key task is to accurately infer and / or track an underlying phenomenon of interest by adaptively exploiting different sensing capabilities, *e.g.* sensor type, location or number of samples, of the underlying sensing system. To maximize the information content at each time step, the following resource allocation problem must be solved: *which sensing mode should be employed at each step to provide the next observation?* Relevant applications include sensor scheduling for object tracking and classification [1, 2], adaptive estimation of sparse signals [3], health care [4] and localization in robotics [5].

In this work, the problem of system state inference with observation control is considered for a system modeled by a discrete-time, finite-state Markov chain. The 'hidden' system state is observed through a conditionally Gaussian measurement vector that depends on the underlying system state and an exogenous control input, which shapes the observations' quality. To accurately track the time-evolving system state, we address the joint problem of determining recursive formulae for a Minimum Mean-Squared Error (MMSE) state estimate and designing a control strategy.

Our motivation stems from the problem of physical activity detection in Wireless Body Area Networks (WBANs) [4]. WBANs constitute a novel class of sensor networks comprising of heterogeneous, biometric sensors, *e.g.* accelerometer (ACC), electrocardiograph (ECG), and a fusion center, usually a personal device. An individual wearing the WBAN switches between a set of physical activities, *e.g.* Stand, Walk, and as a result, the sensors generate and communicate biometric signals to the fusion center. Sensor heterogeneity implies that certain sensors can better discriminate between specific activities [4]; and to accurately infer the underlying time-evolving activity of the individual, we need to jointly design a <u>state estimator</u> and a <u>sensor selection strategy</u> deciding to which sensors to communicate and how many samples to request.

The current work extends our prior work [4], which assumed discrete observations, performed Maximum Likelihood detection of the system state and as an optimization metric, employed a worst-case error probability bound. In fact, the framework proposed herein is much more general.

The classical Kalman filter (KF) [6] is suitable for estimating Gauss-Markov processes in discrete-time, linear systems, while its extensions, *i.e.* the Extended KF (EKF) [6] and the Unscented KF (UKF) [6] are suitable for general, nonlinear, (non)-Gaussian systems, where they assume a Gaussian approximation for the state distribution and their performance depends significantly on either some kind of linearization or the careful selection of a set of sample points. The proposed Sign-of-Innovation KF (SOI-KF) [7] and its extensions (see [5] and references therein) are based on quantized versions of the measurement innovation and / or real measurements for linear Gaussian process and measurement models as well as nonlinear systems. The work in [8] proposes an approximate MMSE estimator starting from a maximum $\grave{a}$ posteriori (MAP) detector in the case of discrete memoryless channels. The Kalman-like Markov chain filters designed in [9, 10] assume discrete-time, finite-state observations but exercise no control. In contrast to these prior works and building on the innovations approach followed by [9, 10], we propose an approximate MMSE state estimator for the case of discrete-time, finite-state Markov chains observed via <u>controlled</u> conditionally Gaussian measurements.

Our work allows fusion of multiple samples from heterogeneous sensors and thus, generalizes prior frameworks that assume one observation from a single sensor [11–13] or θ samples from θ sensors [1, 2, 14]. Furthermore, our filter's MSE performance is intertwined with the control policy design since the trace of the conditional filtering error covariance matrix constitutes the cost functional of a partially observable Markov Decision Process (POMDP) [15], allowing us to focus on the estimation error explicitly. In contrast, prior work assumes a general convex distance measure for the tracking cost [1, 2, 11, 12], some performance bounds [4, 14, 16] or some information-theoretic measures [13]. The resulting POMDP proves

This research has been funded in part by the following grants and organizations: ONR N00014-09-1-0700, NSF CNS-0832186, NSF CCF-0917343, AFOSR FA9550-12-1-0215, the National Center on Minority Health and Health Disparities (NCMHD) (supplement to P60 MD002254), Nokia and Qualcomm.

to be non-standard due to non-linear dependence on the predicted belief state in contrast to [1,2]. Still, the optimal policy can be derived via stochastic dynamic programming versus [14], where a suboptimal scheme is proposed. Our framework also encompasses a large class of problems that fall under the category of sequential multiple hypothesis testing, *e.g.* [3,13,16], and in fact, it is more general since we allow the underlying hypothesis to change with time. Last but not least, the performance of the proposed framework is validated via real data collected from an implemented WBAN [4].

2. PROBLEM FORMULATION

2.1. System Model

We consider a special class of dynamical systems known as Partially Observable Markov Decision Processes (POMDPs) [15], where time is divided in discrete time slots denoted by $k \in \{0, 1, \dots\}$. We define as system state $\mathbf{x}_k$ the activity of the individual at time k. We model the system state as a finite state, first-order Markov chain with n states *i.e.* $\mathcal{X} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$, where $\mathbf{e}_i$ is the unit vector with 1 in the i-th position and zero everywhere else. The Markov chain is characterized by a state transition probability matrix $\mathbf{P}$ with probabilities $P_{j|i} = P(\mathbf{x}_{k+1} = \mathbf{e}_j | \mathbf{x}_k = \mathbf{e}_i), \mathbf{e}_i, \mathbf{e}_j \in \mathcal{X}$ that do not change with time, *i.e.* the chain is stationary.

At each time step, the WBAN sensors generate a set of biometric signals; feature extraction and selection techniques [17, 18] are employed to produce a set of samples that correspond to extracted features' values, *e.g.* ACC mean, ECG period. The fusion center determines the number of samples to receive from each sensor by selecting the appropriate control input $\mathbf{u}_k = [N_1^{\mathbf{u}_k}, \dots, N_K^{\mathbf{u}_k}]^T$, where $N_l^{\mathbf{u}_k}$ denotes the number of samples to receive from sensor l and K the total number of sensors in the WBAN. We assume that each control input $\mathbf{u}_k$ satisfies the constraint $\sum_{l=1}^{K} N_l^{\mathbf{u}_k} \leqslant N$, where N is fixed, and thus, $\mathbf{u}_k \in \mathcal{U} = \{\mathbf{u}^1 \dots, \mathbf{u}^{\sum_{i=1}^{N} \binom{i+K-1}{i}}\}$. At time step k, the fusion center receives a measurement vector $\mathbf{y}_k$ comprising of the selected samples indicated by control input $\mathbf{u}_{k-1}$. Each such measurement vector conditioned on the state and control input follows an AR(1)–correlated multivariate Gaussian model of the form

$$\mathbf{y}_k \big| \mathbf{e}_i, \mathbf{u}_{k-1} \sim f(\mathbf{y}_k | \mathbf{e}_i, \mathbf{u}_{k-1}) = \mathcal{N}(\mathbf{m}_i^{\mathbf{u}_{k-1}}, \mathbf{Q}_i^{\mathbf{u}_{k-1}}), \forall \mathbf{e}_i \in \mathcal{X}, \tag{1}$$

where $\mathbf{m}_i^{\mathbf{u}_{k-1}} = [\mu_{i,S_1}^{\mathbf{u}_{k-1}}, \dots, \mu_{i,S_K}^{\mathbf{u}_{k-1}}]^T$ denotes the mean vector and $\mathbf{Q}_i^{\mathbf{u}_{k-1}} = \mathrm{diag}(\mathbf{Q}_i^{\mathbf{u}_{k-1}}(S_1), \dots, \mathbf{Q}_i^{\mathbf{u}_{k-1}}(S_K))$ the covariance matrix of the Gaussian model for state $\mathbf{e}_i$ and control input $\mathbf{u}_{k-1}$. For a particular sensor S_l, the mean vector $\mu_{i,S_l}^{\mathbf{u}_{k-1}}$ is of size $N_l^{\mathbf{u}_{k-1}} \times 1$ and the covariance matrix is defined as $\mathbf{Q}_i^{\mathbf{u}_{k-1}}(S_l) = \frac{\sigma_{S_l,i}^2}{1-\phi^2}\mathbf{T} + \sigma_z^2 \mathbf{I}$, where $\mathbf{T}$ is a Toeplitz matrix with first row/column $[1, \phi, \phi^2, \dots, \phi^{N_l^{\mathbf{u}_{k-1}}-1}]$, $\mathbf{I}$ is the $N_l^{\mathbf{u}_{k-1}} \times N_l^{\mathbf{u}_{k-1}}$ identity matrix, ϕ is the parameter of the AR(1) model and σ_z^2 accounts for sensing and communication noise. This model has been validated with real data in [18].

2.2. Innovations Representation

We introduce the source sequence of true states $X^k = \{\mathbf{x}_0, \dots, \mathbf{x}_k\}$, the control sequence $U^k = \{\mathbf{u}_0, \dots, \mathbf{u}_k\}$ and the observations sequence $Y^k = \{\mathbf{y}_0, \dots, \mathbf{y}_k\}$. We also define the *global history* $\mathcal{B}_k = \sigma\{X^k, Y^k, U^k\}$, the *history* $\mathcal{B}_k^- = \sigma\{X^k, Y^{k-1}, U^{k-1}\}$ and the *observation-control history* $\mathcal{F}_k = \sigma\{Y^k, U^{k-1}\}$, where $\sigma\{z\}$ denotes the σ-algebra generated by z. The total information available to the fusion center at time k is $\mathcal{F}_k$ and the control input at time k is a function of $\mathcal{F}_k$ *i.e.* $\mathbf{u}_k = \eta_k(\mathcal{F}_k)$.

The *innovations sequence* $\{\mathbf{w}_k\}$ related to $\{\mathbf{x}_k\}$ with respect to $\mathcal{B}_k$ [19] is defined as

$$\mathbf{w}_{k+1} \doteq \mathbf{x}_{k+1} - \mathbb{E}\{\mathbf{x}_{k+1} | \mathcal{B}_k\} = \mathbf{x}_{k+1} - \mathbf{P}\mathbf{x}_k, \tag{2}$$

where the last equality is due to the Markov property. Similarly, the *innovations sequence* $\{\mathbf{v}_k\}$ for the process $\{\mathbf{y}_k\}$ with respect to $\mathcal{B}_k^-$ [19] is defined as

$$\mathbf{v}_k \doteq \mathbf{y}_k - \mathbb{E}\{\mathbf{y}_k | \mathcal{B}_k^-\} = \mathbf{y}_k - \mathcal{M}(\mathbf{u}_{k-1})\mathbf{x}_k, \tag{3}$$

where $\mathcal{M}(\mathbf{u}_{k-1}) = [\mathbf{m}_1^{\mathbf{u}_{k-1}}, \dots, \mathbf{m}_n^{\mathbf{u}_{k-1}}]$ and we have exploited the signal model in (1). Thus, the *Doob–Meyer decompositions* [20] of $\{\mathbf{x}_k\}$ and $\{\mathbf{y}_k\}$ with respect to $\mathcal{B}_k$ and $\mathcal{B}_k^-$, respectively, are

$$\mathbf{x}_{k+1} = \mathbf{P}\mathbf{x}_k + \mathbf{w}_{k+1},\ k \geqslant 0, \tag{4}$$

$$\mathbf{y}_k = \mathcal{M}(\mathbf{u}_{k-1})\mathbf{x}_k + \mathbf{v}_k,\ k \geqslant 1, \tag{5}$$

where $\{\mathbf{w}_k\}$ is a $\{\mathcal{B}\}$–Martingale Difference (MD) sequence and $\{\mathbf{v}_k\}$ is a $\{\mathcal{B}^-\}$–MD sequence [19].

3. SYSTEM STATE ESTIMATOR

In this section, we develop a Kalman-like filter for estimating the discrete-time, finite-state Markov chain system state from past observations and control inputs based on the theory introduced in [9, 19]. We begin by defining the probability distribution of $\mathbf{x}_k$ conditioned on $\mathcal{F}_k$, known as the *belief state* [15], as

$$\mathbf{p}_{k|k} \doteq \left[p_{k|k}^1, \dots, p_{k|k}^n\right]^T, \tag{6}$$

where $p_{k|k}^i = P(\mathbf{x}_k = \mathbf{e}_i | \mathcal{F}_k), \forall \mathbf{e}_i \in \mathcal{X}$. The expected value of $\mathbf{x}_k$ conditioned on $\mathcal{F}_k$ coincides with the belief state since

$$\mathbf{x}_{k|k} \doteq \mathbb{E}\{\mathbf{x}_k | \mathcal{F}_k\} = \sum_{i=1}^{n} \mathbf{e}_i P(\mathbf{x}_k = \mathbf{e}_i | \mathcal{F}_k) = \mathbf{p}_{k|k}, \tag{7}$$

Both notations will be used interchangeably in the sequel.

At this point, we define the *estimate innovations sequence* $\{\mu_k\}$ and the *observation innovations sequence* $\{\lambda_k\}$ as follows [9, 10]

$$\mu_k \doteq \mathbf{x}_{k|k} - \mathbf{x}_{k|k-1} = \mathbb{E}\{\mathbf{x}_k | \mathcal{F}_k\} - \mathbb{E}\{\mathbf{x}_k | \mathcal{F}_{k-1}\}, \tag{8}$$

$$\lambda_k \doteq \mathbf{y}_k - \mathbf{y}_{k|k-1} = \mathbf{y}_k - \mathbb{E}\{\mathbf{y}_k | \mathcal{F}_{k-1}\}, \tag{9}$$

where $\mathbf{x}_{k|k-1}, \mathbf{y}_{k|k-1}$ are the predicted state and measurement estimates, respectively. We can easily prove that both sequences are $\{\mathcal{F}\}$–MD sequences [19]. The MD representation theorem [9, 19] constitutes a powerful tool for developing recursive nonlinear MMSE Kalman-like estimators. This theorem states that the sequences $\{\mu_k\}$ and $\{\lambda_k\}$ can be related via $\mu_k = \mathbf{G}_k \lambda_k$ assuming we can determine an $\{\mathcal{F}\}$–predictable sequence $\{\mathbf{G}_k\}$. In such a case, $\mathbf{G}_k$ is given by

$$\mathbf{G}_k = \mathbb{E}\{\mu_k \lambda_k^T | \mathcal{F}_{k-1}\} \left[\mathbb{E}\{\lambda_k \lambda_k^T | \mathcal{F}_{k-1}\}\right]^{-1}. \tag{10}$$

Inspired by [21] and since a recursive solution is desired in our case, we impose recursivity as a design constraint and use (10) as an approximation. This approximation along with the Doob–Meyer decompositions (4)–(5) and the definitions in (8)–(9) allow us to determine a suboptimal Kalman-like nonlinear MMSE filtered estimator for the Markov chain system state. Note that for the set of recursive estimators with a Kalman-like structure, the proposed estimator is an optimal MMSE estimator.

Theorem 1 *The Markov chain system estimate at time step k is recursively defined as*

$$\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + \mathbf{G}_k[\mathbf{y}_k - \mathbf{y}_{k|k-1}],\ k \geqslant 0 \tag{11}$$

with

$$\mathbf{x}_{k|k-1} = \mathbf{P}\mathbf{x}_{k-1|k-1}, \tag{12}$$

$$\mathbf{y}_{k|k-1} = \mathcal{M}(\mathbf{u}_{k-1})\mathbf{x}_{k|k-1}, \tag{13}$$

$$\mathbf{G}_k = \boldsymbol{\Sigma}_{k|k-1}\mathcal{M}^T(\mathbf{u}_{k-1})\times (\mathcal{M}(\mathbf{u}_{k-1})\boldsymbol{\Sigma}_{k|k-1}\mathcal{M}^T(\mathbf{u}_{k-1}) + \mathbf{J}_k)^{-1}, \tag{14}$$

where $\mathbf{x}_{0|-1} = \pi$, and π is the initial distribution over the system states, $\boldsymbol{\Sigma}_{k|k-1}$ is the conditional covariance matrix of the prediction error and $\mathbf{J}_k = \sum_{i=1}^{n} x^i_{k|k-1}\mathbf{Q}_i^{\mathbf{u}_{k-1}}$.

At this point, we underscore that even though the proposed filter is *formally* similar to the classical KF, it is not a standard KF and the gain $\mathbf{G}_k$ is not a standard Kalman gain; in fact, *it depends on the observations*. The *conditional filtering* and *prediction error covariance matrices* have the following form

$$\boldsymbol{\Sigma}_{k|k} \doteq \mathbb{E}\{(\mathbf{x}_k - \mathbf{x}_{k|k})(\mathbf{x}_k - \mathbf{x}_{k|k})^T|\mathcal{F}_k\}, \tag{15}$$

$$\boldsymbol{\Sigma}_{k|k-1} \doteq \mathbb{E}\{(\mathbf{x}_k - \mathbf{x}_{k|k-1})(\mathbf{x}_k - \mathbf{x}_{k|k-1})^T|\mathcal{F}_{k-1}\}. \tag{16}$$

Since no constraint is imposed on the individual components of $\mathbf{x}_{k|k}$, there is no guarantee that they lie on the $[0, 1]$ interval. To overcome this issue without incorporating additional constraints that may challenge the determination of a solution to our problem, we adopt the approach of [10], *i.e.* apply a suitable memoryless (linear or nonlinear) transformation of $\mathbf{x}_{k|k}$ to ensure feasible solutions are determined.

4. OPTIMAL CONTROL POLICY DESIGN

Our goal is to determine an admissible control policy $\gamma = \{\eta_0, \eta_1, \ldots, \eta_{L-1}\}$ [15] that minimizes the MSE

$$J_\gamma = \mathop{\mathbb{E}}_{\mathbf{y}_0,\mathbf{y}_1,\ldots,\mathbf{y}_L}\left\{\sum_{k=1}^{L}\mathrm{tr}\left(\boldsymbol{\Sigma}_{k|k}(\mathbf{y}_k, \mathbf{u}_{k-1})\right)\right\}, \tag{17}$$

where L represents the horizon length, $\mathrm{tr}(\cdot)$ denotes the trace operator and the dependence of $\boldsymbol{\Sigma}_{k|k}$ on $\mathbf{y}_k$ and $\mathbf{u}_{k-1}$ has been stated explicitly. Thus, we have the following *finite horizon, partially observable stochastic control problem*: $\min_{\mathbf{u}_0,\mathbf{u}_1,\ldots,\mathbf{u}_{L-1}} J_\gamma$; and the optimal solution may be obtained via dynamic programming (DP) [15]. However, in contrast to the standard problems of this type, *e.g.* see [15, 22], our cost function is defined with respect to the observations sequence and not the state sequence. To determine the optimal policy, we adapt methods from [15, 22] to our case.

The information $\mathcal{F}_k$ for decision making at each time step k is of expanding dimension [15], hence we can show that a sufficient statistic for control purposes, which is bounded in memory, is the conditional probability distribution $\mathbf{p}_{k+1|k}$ of the next state $\mathbf{x}_{k+1}$ given the observation-control history $\mathcal{F}_k$ and designate this as the *predicted belief state*. The predicted belief state can be updated recursively based on the new measurement vector and the current control input as shown in Lemma 2.

Lemma 2 *Let $\mathbf{p}_{k|k-1}$ denote the predicted belief state at time $k-1$. Assume that the control input $\mathbf{u}_{k-1}$ is selected and at time step k, the measurement vector $\mathbf{y}_k$ is generated. Then, the predicted belief state $\mathbf{p}_{k+1|k}$ is given by the following recursion*

$$\mathbf{p}_{k+1|k} = \frac{\mathbf{Pr}(\mathbf{y}_k, \mathbf{u}_{k-1})\mathbf{p}_{k|k-1}}{\mathbf{1}_n^T\mathbf{r}(\mathbf{y}_k, \mathbf{u}_{k-1})\mathbf{p}_{k|k-1}}, \tag{18}$$

where $\mathbf{1}_n$ is a column vector consisting of n ones and $\mathbf{r}(\mathbf{y}_k, \mathbf{u}_{k-1}) =$ $\mathrm{diag}\left(f(\mathbf{y}_k|\mathbf{e}_1, \mathbf{u}_{k-1}), \ldots, f(\mathbf{y}_k|\mathbf{e}_n, \mathbf{u}_{k-1})\right)$ denotes the $n \times n$ diagonal matrix of measurement vector probability density functions.

Theorem 3 gives the finite-horizon DP equations in terms of the predicted belief state.

Theorem 3 *For $k = L-1, \ldots, 1$ the cost-to-go function $\overline{J}_k(\mathbf{p}_{k|k-1})$ is related to $\overline{J}_{k+1}(\mathbf{p}_{k+1|k})$ through the recursion*

$$\overline{J}_k(\mathbf{p}_{k|k-1}) = \min_{\mathbf{u}_{k-1}\in\mathcal{U}}\Bigg[\mathbf{p}_{k|k-1}^T\mathbf{h}(\mathbf{p}_{k|k-1}, \mathbf{u}_{k-1}) + \int \mathbf{1}_n^T\mathbf{r}(\mathbf{y}, \mathbf{u}_{k-1})\mathbf{p}_{k|k-1}\overline{J}_{k+1}\left(\frac{\mathbf{Pr}(\mathbf{y}, \mathbf{u}_{k-1})\mathbf{p}_{k|k-1}}{\mathbf{1}_n^T\mathbf{r}(\mathbf{y}, \mathbf{u}_{k-1})\mathbf{p}_{k|k-1}}\right)d\mathbf{y}\Bigg], \tag{19}$$

where $\mathbf{h}(\mathbf{p}_{k|k-1}, \mathbf{u}_{k-1})$ is a column vector with components $h(\mathbf{e}_1, \mathbf{p}_{k|k-1}, \mathbf{u}_{k-1}), \ldots, h(\mathbf{e}_n, \mathbf{p}_{k|k-1}, \mathbf{u}_{k-1})$ with $h(\mathbf{e}_i, \mathbf{p}_{k|k-1}, \mathbf{u}_{k-1}) = 1 - \mathrm{tr}\left(\mathbf{G}_k^T\mathbf{G}_k\mathbf{Q}_i^{\mathbf{u}_{k-1}}\right) - \|\mathbf{p}_{k|k-1} + \mathbf{G}_k(\mathbf{m}_i^{\mathbf{u}_{k-1}} - \mathbf{y}_{k|k-1})\|^2$. The cost-to-go function for $k = L$ is given by

$$\overline{J}_L(\mathbf{p}_{L|L-1}) = \min_{\mathbf{u}_{L-1}\in\mathcal{U}}\left[\mathbf{p}_{L|L-1}^T\mathbf{h}(\mathbf{p}_{L|L-1}, \mathbf{u}_{L-1})\right]. \tag{20}$$

The DP equations stated in Theorem 3 result in high computational complexity to determine the optimal solution. Specifically, as with traditional POMDPs, the predicted belief state $\mathbf{p}_{k|k-1}$ is uncountably infinite [15]. Furthermore, the control input definition suggests that the control space size can be exponentially large, while determining the expected future cost is challenging since it requires, in the worst-case, an N–dimensional integration. Last but not least, in contrast to standard POMDP problems [15], the term $\mathbf{p}_{k|k-1}^T\mathbf{h}(\mathbf{p}_{k|k-1}, \mathbf{u}_{k-1})$ is not a linear function of the predicted belief state $\mathbf{p}_{k|k-1}$ and thus, existing techniques such as [23] and [24] cannot be directly employed. Still, for small problem sizes, an approximate solution via numerical computation is feasible.

5. NUMERICAL RESULTS

In this section, we provide numerical results to illustrate the performance of the proposed framework. We use real data collected through an implemented WBAN, the KNOWME network [18]. For clarity of exposition, we focus on the three-sensor case, two ACCs and one ECG, with four physical activity states (Sit, Stand, Run, Walk), and underscore that our methods are directly applicable to multiple sensors and physical states. The features used are: 1) the ACC mean from the first ACC, 2) the ACC variance from the second ACC and 3) the ECG period from the ECG. The state distributions for the three sensors for a single individual is shown in Fig. 1. The Markov chain transition probabilities are described by the matrix

$$\mathbf{P} = \begin{bmatrix} 0.6 & 0 & 0.2 & 0.4 \\ 0.1 & 0.4 & 0.1 & 0 \\ 0 & 0.1 & 0.3 & 0.3 \\ 0.3 & 0.3 & 0.6 & 0.3 \end{bmatrix}.$$

The numerical results presented herein were performed for $N = 2$ samples. Even though the number of samples are few, patterns still emerge.

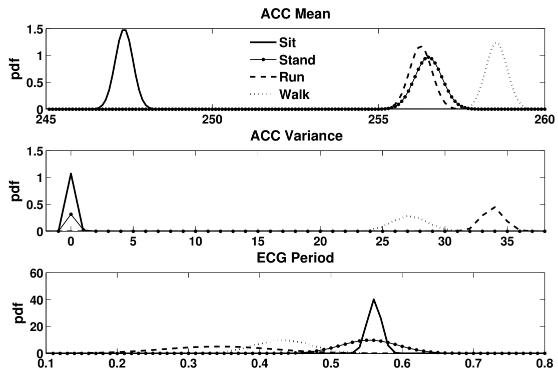

Fig. 1. Gaussian distributions associated with each of the four activity states for the ACC mean, ACC variance and ECG period features, respectively. The plots indicate that a combination of samples from the ACC mean and the ACC variance can help us discriminate between the physical activities of interest. On the other hand, the ECG Period is not very informative.

In Fig. 2, we present the tracking performance of the proposed system by showing the true and estimated state sequences. The output of our system is an estimate of the belief state and we detect the activity state via a MAP rule. We observe that the proposed framework tracks significantly well the underlying, time-evolving activity state even though the total number of samples used are few. Furthermore, we note that the Stand state is usually not detected since according to the stationary distribution of the Markov chain, it corresponds to an ephemeral state. Modifications of the tracking cost similar to the ones presented in [4] can be employed to detect ephemeral states.

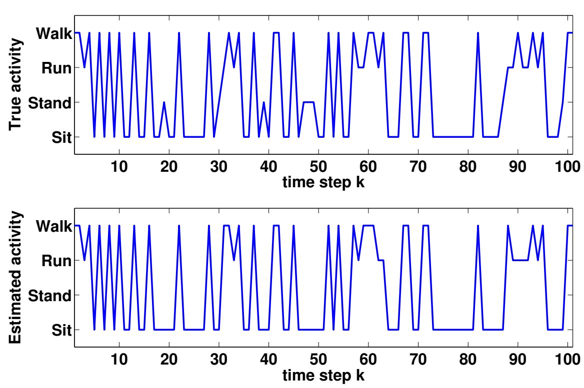

Fig. 2. Tracking performance of the proposed framework: the upper plot shows the individual's true activity while the lower plot the estimated activity.

At this point, we wish to comment on the form of the optimal control policy. The optimal control policy consists of three types of control inputs: 1) 2 ACC mean samples, 2) 1 ACC mean sample and 1 ACC variance sample, and 3) 2 ACC mean samples. The first type of control input is selected for most of the predicted belief states and this is due to the fact that it can discriminate between the more likely states, *i.e.* Sit, Run, Walk. The second and third types of control input are primarily selected for detecting the least likely state, Stand. Specifically, when the Sit state has low probability ($\leqslant 0.5$), the second control input is selected since one sample from each of the informative sensors can help us discriminate Stand from the rest of the states. However, when the Run and Walk states have zero probability, samples from the ACC mean are enough to detect Stand, as verified by Fig. 1.

Table 1. Detection accuracy for different control policies (A: 1 ACC mean sample, B: 1 ACC variance sample, Γ: 1 ECG Period sample, Optimal: determined by DP.)

Control policy	A	B	Γ	Optimal
Detection accuracy	74%	77%	40%	87%

Table 1 summarizes the detection accuracy achieved by employing different control policies. Control strategies A, B and Γ always request 1 sample from the associated sensor irrespective of the predicted belief state. We find that this approach leads to inferior detection performance compared to the optimal control policy, which considers the predicted belief state. Furthermore, adaptively fusing samples from sensors of different capabilities, as done by the optimal control policy, can boost detection performance significantly. The observed performance improvement is a direct outcome of the optimized control input selection process since the current control input determines which measurements are selected that in turn (via the predicted belief state) determine the next control. Finally, we expect that larger values of the total number N of available samples will give rise to higher detection accuracy.

6. CONCLUSIONS AND FUTURE WORK

In this work, we addressed the problem of joint MMSE state estimation and control policy design for a discrete-time, finite-state Markov chain observed via controlled Gaussian measurements. Specifically, we derived a suboptimal Kalman-like nonlinear MMSE estimator for the Markov chain system state exploiting an innovations approach. We also derived a stochastic dynamic programming algorithm to determine the optimal control policy with the cost functional being the filters' MSE performance. We validated the performance of the proposed framework via numerical simulations on a WBAN application using real data.

At this early stage, approximate optimal control policies were determined by numerically solving the DP equation. Future work will focus on designing efficient algorithms for calculating control policies as well as incorporating sensing usage costs into the proposed framework and applications to other problems admitting the proposed framework.

7. REFERENCES

[1] G. K. Atia, V. V. Veeravalli, and J. A. Fuemmeler, "Sensor Scheduling for Energy-Efficient Target Tracking in Sensor Networks," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4923–4937, October 2011.

[2] J. A. Fuemmeler, G. K. Atia, and V. V. Veeravalli, "Sleep Control for Tracking in Sensor Networks," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4354–4366, September 2011.

[3] D. Wei and A. O. Hero, "Multistage adaptive estimation of sparse signals," in *IEEE Statistical Signal Processing Workshop (SSP)*, Aug. 2012, pp. 153–156.

[4] D.-S. Zois, M. Levorato, and U. Mitra, "Energy–Efficient, Heterogeneous Sensor Selection for Physical Activity Detection in Wireless Body Area Networks," *IEEE Tansactions on Signal Processing*, vol. 61, no. 7, pp. 1581–1594, April 2013.

[5] E. D. Nerurkar, K. X. Zhou, and S. I. Roumeliotis, "Hybrid Estimation Framework for Multi-robot Cooperative Localization using Quantized Measurements," Tech. Rep., University of Minnesota, 2011.

[6] Z. Chen, "Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond," Tech. Rep., McMaster University, 2003.

[7] A. Ribeiro, G. B. Giannakis, and S. I. Roumeliotis, "SOI-KF: Distributed Kalman Filtering With Low-Cost Communications Using the Sign of Innovations," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4782–4795, Dec. 2006.

[8] N. Phamdo and N. Farvardin, "Optimal Detection of Discrete Markov Sources Over Discrete Memoryless Channels–Applications to Combined Source-Channel Coding," *IEEE Transactions on Information Theory*, vol. 40, no. 1, pp. 186–193, Jan. 1994.

[9] A. Segall, "Recursive estimation from discrete-time point processes," *IEEE Transactions on Information Theory*, vol. 22, no. 4, pp. 422–431, Sep. 1976.

[10] E. Baccarelli and R. Cusani, "Recursive Kalman-type optimal estimation and detection of hidden Markov chains," *Signal Process.*, vol. 51, no. 1, pp. 55–64, May 1996.

[11] V. Krishnamurthy, "Algorithms for optimal scheduling and management of hidden Markov model sensors," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1382–1397, June 2002.

[12] V. Krishnamurthy and D. Djonin, "Structured Threshold Policies for Dynamic Sensor Scheduling - a Partially Observed Markov Decision Process Approach," *IEEE Transactions on Signal Processing*, vol. 55, no. 10, pp. 4938–4957, October 2007.

[13] M. Naghshvar and T. Javidi, "Extrinsic Jensen-Shannon divergence with application in active hypothesis testing," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, july 2012, pp. 2191–2195.

[14] E. Masazade, R. Niu, and P. K. Varshney, "An approximate dynamic programming based non-myopic sensor selection method for target tracking," in *46th Annual Conference on Information Sciences and Systems (CISS)*, March 2012, pp. 1–6.

[15] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 1, Athena Scientific, 2005.

[16] S. Nitinawarat, G .K. Atia, and V. V. Veeravalli, "Controlled sensing for hypothesis testing," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2012, pp. 5277–5280.

[17] M. Li, V. Rozgic, G. Thatte, A. Sangwon, L.and Emken, M. Annavaram, U. Mitra, D. Spruijt-Metz, and S. Narayanan, "Multimodal Physical Activity Recognition by Fusing Temporal and Cepstral Information," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, pp. 369 –380, August 2010.

[18] G. Thatte, M. Li, S. Lee, B. A. Emken, M. Annavaram, S. Narayanan, D. Spruijt-Metz, and U. Mitra, "Optimal Time-Resource Allocation for Energy-Efficient Physical Activity Detection," *IEEE Tansactions on Signal Processing*, vol. 59, no. 4, pp. 1843 –1857, April 2011.

[19] A. Segall, "Stochastic Processes in estimation theory," *IEEE Transactions on Information Theory*, vol. 22, no. 3, pp. 275–286, May 1976.

[20] G. Kallianpur, *Stochastic Filtering Theory*, Springer, 1980.

[21] S. Galli, "A new family of soft-output adaptive receivers exploiting nonlinear MMSE estimates for TDMA-based wireless links," *IEEE Transactions on Communications*, vol. 50, no. 12, pp. 1935–1945, Dec. 2002.

[22] J. L. Speyer and W. H. Chung, *Stochastic Processes, Estimation and Control*, SIAM, 2008.

[23] R. D. Smallwood and E. J. Sondik, "Optimal control of partially observable Markov processes over a finite horizon," *Operations Research*, vol. 21, pp. 1071–1088, 1973.

[24] W. S. Lovejoy, "A survey of algorithmic methods for partially observed Markov decision processes," *Annals of Operations Research*, vol. 28, pp. 47–66, April 1991.