# A Framework for Mixed Estimation of Hidden Markov Models [1]

Subhrakanti Dey
Institute for Systems Research
University of Maryland
College Park, MD 20742

Steven I. Marcus
Institute for Systems Research
Department of Electrical Engineering
University of Maryland
College Park, MD 20742

## Abstract

In this paper, we present a framework for a mixed estimation scheme for hidden Markov models (HMM). A robust estimation scheme is first presented using the minimax method that minimizes a worst case cost for HMMs with bounded uncertainties. Then we present a mixed estimation scheme that minimizes a risk-neutral cost with a constraint on the worst-case cost. Some simulation results are also presented to compare these different estimation schemes in cases of uncertainties in the noise model.

## 1 Introduction

A hidden Markov model (HMM) is a stochastic process that usually consists of a state process that is a finite-state Markov chain and an observation or measurement process that is a function of the state process corrupted by noise. The observation process can be discrete-range or continuous-range. In this paper, we will be interested in discrete-time homogeneous Markov chains taking values in a finite-dimensional state space and observation processes that are continuous-range, i.e., they are observed in continuous-range noise. Precise details about our signal model will be given in the next section. In short, we are interested in developing a robust estimation algorithm for a class of HMMs with unknown but bounded uncertainties and a "mixed" estimation problem that minimizes a quadratic cost with a constraint on a worst-case cost for a class of HMMs with random disturbances. The background and mo-tivation behind formulating and solving such problems are given below.

HMMs are known to be good models of many random nonlinear physical processes and there are many applications of HMM signal processing in diverse areas like speech recognition, communication systems, biological signal processing, frequency tracking, fault detection etc. to name a few. In all these applications, the basic algorithm involves estimation of the state and the parameters of the Markov chain that describes the state process. State estimation of HMM is usually done by calculating the "forward variable" [1] which is essentially a conditional probability mass function of the state given the observations, which can be calculated recursively given the initial state distribution, the transition probability matrix of the Markov chain, the knowledge of the statistics of the measurement noise and the observations. One can then define a suitable state estimate (e.g. the "MAP" estimate or the conditional mean estimate) based on this conditional distribution of the state. This estimate is essentially a minimum-variance or a "risk-neutral" state estimate in the sense that it is not sensitive to uncertainties in the model. As opposed to this, a class of robust estimation algorithms, known as "risk-sensitive" estimation schemes, following the ideas of risk-sensitive control [2] [3] [4] [5] were developed in [6] (for linear Gaussian signal models) [7] (for a class of nonlinear signal models) and [8] (for hidden Markov models). Risk-sensitive estimation essentially minimizes an exponential of a quadratic (or more general convex) cost and thus penalizes the higher order moments of the estimation error energy to provide robustness against model uncertainties. Recently, a more meaningful insight into

3473

the robustness offered by risk-sensitive estimation has been given in [9].

However, the setting of risk-sensitive estimation schemes is stochastic in nature and in general, small noise limit results show that risk-sensitive estimation algorithms can be connected to a deterministic worst-case noise estimation problem given from a differential dynamic game [10] [5] [11] ($H_\infty$ estimation for linear Gaussian systems). Risk-sensitive output feedback control problems for HMMs have been treated in [12] [13] [14] and relations have been drawn to robust control for finite-state machines. In particular, in [12], a deterministic model for uncertainties is introduced leading to a dynamic game formulation of the robust control problem. A random perturbation of the deterministic system is treated as an HMM and the stochastic control problem for this HMM is shown to be related to the dynamic game problem for the deterministic model using small noise limits. However, in the general framework of [12], no specific choices of the cost functions associated with the disturbances are given. In addition, depending on the nature of the disturbances (often a mixture of random and unknown but bounded disturbances [15]), it is often necessary to introduce a trade-off between the risk-neutral and risk-sensitive or robust estimation objectives. One way to do this is to introduce a "mixed" criterion where a risk-neutral cost is minimized subject to a constraint on the worst-case cost. A mixed risk-neutral and minimax control problem is solved for HMMs in [16].

In our paper, we formulate a robust estimation problem for a hidden Markov model with unknown but bounded uncertainties. Following ideas similar to [12], we set up a dynamic game problem for the robust estimation scheme with appropriate choices for the cost functions associated with the disturbances in the state (reflected by bounded variations of the transition probability matrix) and the observation process (reflected by the additive continuous-range independent bounded noise) and the initial distribution of the state. The objective of the robust estimation problem is to obtain state estimates which minimize a worst-case cost over a finite horizon when the estimates are constrained to the vector space of unit vectors. Next, we extend the ideas of [16] to set up a "mixed" estimation problem that minimizes a risk-neutral or quadratic cost subject to a constraint satisfied by the worst-case cost described before.

Simulation results show that in the event of bounded disturbances being present, minimax estimation outperforms risk-neutral estimation and mixed estimation guarantees the worst case cost to be constrained where as risk-neutral estimation does not. We also compare the performance of minimax estimation with that of risk-sensitive estimation.

## 2 Signal Model

Consider a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ where $X_k$ is a discrete-time homogeneous Markov chain belonging to a finite-discrete set. Define $\mathcal{E} \overset{\triangle}{=} \{e_1, e_2, \ldots, e_N\}$ where $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)' \in \mathbb{R}^N$ with 1 in the $i$-th position. Without loss of generality, we can assume that $X_k \in \mathcal{E}$. Denote the transition probability matrix as $A = (a_{ij})$ where $a_{ij} = P(X_{k+1} = e_i \mid X_k = e_j)$. We assume that there exists an $\epsilon > 0$ such that $a_{ij} \geq \epsilon$. Also, $\sum_{i=1}^{N} a_{ij} = 1$, $\forall j$.

We observe a process $y_k \in \mathbb{R}^p$ such that

$$y_k = H(X_k) + v_k \tag{1}$$

where $v_k \in \mathbb{R}^p$, $k \in \mathbb{N}$ is the disturbance in the measurement process that may be random with known statistical information or unknown but bounded in $L_2$ with probability one, depending on the nature of our estimation problem. Define $\{\mathcal{Y}_k\} \overset{\triangle}{=} (y_0, y_1, \ldots, y_k)$. In case the disturbances are purely random, one can define $\{\mathcal{Y}_k\}$ to be the complete filtration generated by $\sigma\{y_0, y_1, \ldots, y_k\}$. In the following sections, we will be using the notation $\{\mathcal{Y}_k\}$ with their appropriate definitions relative to the context, without reiterating the definitions separately for each context.

Also, define $\pi_0 \in \mathbb{R}^N$ to be the initial probability distribution of the Markov chain, such that $P(X_0 = e_i) = \pi_0(i)$. We assume that there exists a $\delta > 0$ such that $\pi_0(i) \geq \delta$, $\forall i$. Obviously, $\sum_{i=1}^{N} \pi_0(i) = 1$.

## 3 Minimax state estimation for bounded uncertainties

In this section, we assume that $v_k$, defined in the previous section is unknown but bounded in $L_2$ with probability 1. Also, uncertainties in the $A$ matrix and the initial probability distribution vector $\pi_o$ are assumed

3474

to be such that the assumptions made earlier on the elements of $A$ and $\pi_0$ still hold.

Consider a specific state sequence $X_0 = e_{i_0}, X_1 = e_{i_1}, \ldots, X_k = e_{i_k}$ and an observation sequence $\{y_l\}$, $l = 0, 1, \ldots, k$. Define $Z_k = LX_k$, $L \in \mathbb{R}^{N \times N}$ and our objective is to obtain an estimate $\hat{Z}_k = L\hat{X}_k$ ($\hat{X}_k \in \mathcal{E}$) of $Z_k$ as a Borel measurable function of $\{\mathcal{Y}_k\}$, $k \geq 0$ such that the following worst case cost is minimized:

$$\max_{i_0, \ldots, i_k} \left[ \sum_{l=0}^{k} \phi(e_{i_l}, \hat{X}_l) - \beta(e_{i_0}) \right.$$
$$\left. - \frac{1}{\mu} \left\{ \sum_{l=0}^{k-1} U(e_{i_l}, e_{i_{l+1}}) + \sum_{l=0}^{k} V(y_l, e_{i_l}) \right\} \right] \quad (2)$$

where $\phi : \mathcal{E} \times \mathcal{E} \to \mathbb{R}$, $U : \mathcal{E} \times \mathcal{E} \to \mathbb{R}$, $\beta : \mathcal{E} \to \mathbb{R}$ and $V : \mathbb{R}^p \times \mathcal{E} \to \mathbb{R}$ with the following properties, $\phi(e_{i_l}, \hat{X}_l) \geq 0$, $\forall i_l$, $\forall l$, $\infty > \beta(e_{i_0}) > 0$, $\forall i_0 \in \{1, 2, \ldots, N\}$, $\infty > U(e_{i_l}, e_{i_{l+1}}) > 0$, $\forall i_l, i_{l+1}$, $\forall l$ and $\infty > \gamma \geq V(y_l, e_{i_l}) \geq 0$, $\forall i_l$, $\forall l$. Also, $\mu > 0$. We also make the assumptions that the above mentioned cost functions are infinite valued if any of their arguments do not belong to their respective domain spaces.

In other words, we find $\{\mathcal{Y}_k\}$-adapted $\hat{X}_k$, $k \geq 0$ such that

$$\hat{X}_k = \operatorname*{argmin}_{\xi \in \mathcal{E}} \max_{i_0, \ldots, i_k} \left[ \sum_{l=0}^{k-1} \phi(e_{i_l}, \hat{X}_l) + \phi(e_{i_k}, \xi) - \beta(e_{i_0}) \right.$$
$$\left. - \frac{1}{\mu} \left\{ \sum_{l=0}^{k-1} U(e_{i_l}, e_{i_{l+1}}) + \sum_{l=0}^{k} V(y_l, e_{i_l}) \right\} \right] \quad (3)$$

**Remark 3.1** Note above that at each time $k$, we only obtain $\hat{X}_k$, and do not obtain new values for $X_l$, $l < k$. In other words, this is a strict filtering problem.

Now, we make specific choices of the cost functions $\phi(.,.)$, $\beta(.)$, $U(.,.)$ and $V(.,.)$. Denoting $Q = L'L$, we make the following choices:

$$\phi(x, \hat{x}) = \frac{1}{2}(x - \hat{x})'Q(x - \hat{x})$$
$$\beta(e_j) = -ln(\pi_0(j))$$
$$U(e_i, e_j) = -ln(a_{ji})$$
$$V(y_k, e_j) = \|y_k - H(e_j)\|^2 \quad (4)$$

Here $\|.\|$ denotes the Euclidian distance between two vectors. Also, note that the above cost functions do satisfy the assumptions we made earlier.

With these specific choices, now we can define the following information state:

**Definition 3.1** *Define the information state* $s_k(j)$, $k \geq 1$, $j \in \{1, 2, \ldots, N\}$ *as*

$$s_k(j) = \max_i \left[ s_{k-1}(i) + \frac{1}{2}(e_i - \hat{X}_{k-1})'Q(e_i - \hat{X}_{k-1}) \right.$$
$$\left. - \frac{1}{\mu}(-ln(a_{ji}) + \|y_{k-1} - H(e_i)\|^2) \right]$$
$$s_0(i) = ln(\pi_0(i)), \; i \in \{1, 2, \ldots, N\} \quad (5)$$

The minimax state estimate is given by the following theorem:

**Theorem 3.1** *Consider the HMM signal model defined in Section 2 and the minimax dynamic game problem defined by (3) and (4). Then the minimax state estimate is given by*

$$\hat{X}_0 = e_{i^*}$$
$$i^* = \operatorname*{argmin}_l \max_i \left[ s_0(i) \right.$$
$$\left. + \frac{1}{2}(e_i - e_l)'Q(e_i - e_l) - \frac{1}{\mu}\|y_0 - H(e_i)\|^2 \right]$$
$$\hat{X}_k = e_{j^*}$$
$$j^* = \operatorname*{argmin}_l \max_i \left[ s_k(i) \right.$$
$$\left. + \frac{1}{2}(e_i - e_l)'Q(e_i - e_l) - \frac{1}{\mu}\|y_k - H(e_i)\|^2 \right] \quad (6)$$

**Proof:** The proof is straightforward once we use the method of forward dynamic programming and the definition of the information state 3.1. Substituting this in (3), we can obtain (6). ∎

**Remark 3.2** Note that, in the case of minimax control as in [16], one needs to use a backward dynamic programming and the concept of a value function, but the necessity of using such tools does not arise in the strict filtering problem mentioned above.

## 4 Mixed estimation

In this section, we formulate a mixed estimation problem for the HMM defined in Section 2. We briefly recapitulate the risk-neutral and the risk-sensitive state

3475

estimation algorithms. Then we present a solution to the mixed estimation problem.

For the purpose of this section, we return to the usual stochastic framework of the HMM observation model (1). We assume that $\{v_k\}$ is a sequence of *i.i.d.* random variables. In this paper, we assume that $v_k \sim N(0, \Sigma)$, $\forall k$. Recall from [1] that the risk-neutral estimate (which essentially is the conditional mean estimate) for the state of the Markov chain is given by $E[X_k \mid \mathcal{Y}_k]$. This is obtained from the unnormalized measure (denoted by $\alpha_k \in \mathbb{R}^N$) which can be computed recursively as follows:

$$\alpha_k(j) = E[< X_k, e_j > \mid \mathcal{Y}_k] = b_j(y_k) \sum_{i=1}^{N} a_{ji}\alpha_{k-1}(i)$$

$$\alpha_0(i) = b_i(y_0)\pi_0(i) \tag{7}$$

Or, in matrix notation,

$$\alpha_k = B(y_k)A\alpha_{k-1}, \quad \alpha_0 = B(y_0)\pi_0 \tag{8}$$

Here, $B(y_k) = diag(b_1(y_k), \ldots, b_N(y_k))$, $b_i(y_k) = \frac{1}{\sqrt{(2\pi \det \Sigma)^p}} \exp\left\{-\frac{1}{2}(y_k - H(e_i))'\Sigma^{-1}(y_k - H(e_i))\right\}$.

The above unnormalized estimate can be normalized to yield

$$P(X_k = e_j \mid \mathcal{Y}_k) = E[< X_k, e_j > \mid \mathcal{Y}_k] = \frac{\alpha_k(j)}{\sum_j \alpha_k(j)} \tag{9}$$

## Risk-sensitive estimation for HMMs

The risk-sensitive state estimate $\hat{X}_k^{rs}$ of a hidden Markov model is discussed in detail in [8]. We quote the main results here. The risk-sensitive cost for an HMM described in Section 2 is given by

$$\hat{X}_k^{rs} = \underset{\zeta \in \mathcal{E}}{\operatorname{argmin}} E[\exp(\frac{\theta}{2}\left\{\sum_{l=0}^{k-1}(X_l - \hat{X}_l^{rs})'Q(X_l - \hat{X}_l^{rs})\right. $$
$$\left. + (X_k - \zeta)'Q(X_k - \zeta)\right\}) \mid \mathcal{Y}_k], \quad k \geq 0 \tag{10}$$

Here, $\theta > 0$ plays a similar role as $\mu$ in (2).

One then defines a new measure $\bar{P}$, under which $\{y_k\}$ is a sequence of *i.i.d.* random variables with density $N(0, \Sigma)$. The corresponding Radon-Nikodym derivative is given by

$$\frac{dP}{d\bar{P}} = \Lambda_k = \prod_{0}^{k} \frac{g(y_k - H(X_k))}{g(y_k)}$$

where $g \triangleq N(0, \Sigma)$. Denoting the expectation under $\bar{P}$ as $\bar{E}$, and

$$\Psi_{0,k} = \frac{1}{2}\sum_{l=0}^{k-1}(X_l - \hat{X}_l^{rs})'Q(X_l - \hat{X}_l^{rs})$$

we can define the following unnormalized information state:

$$q_k(j) = \bar{E}[\Lambda_{k-1}\exp(\theta\Psi_{0,k-1}) < X_k, e_j > \mid \mathcal{Y}_{k-1}] \tag{11}$$

It can be shown that the information state obeys the following recursion

$$q_{k+1} = AD_kB(y_k)q_k \tag{12}$$

where

$$D_k = diag\left\{\exp\left(\frac{\theta}{2}(e_1 - \hat{X}_k^{rs})'Q(e_1 - \hat{X}_k^{rs})\right), \ldots, \right.$$
$$\left. \exp\left(\frac{\theta}{2}(e_N - \hat{X}_k^{rs})'Q(e_N - \hat{X}_k^{rs})\right)\right\} \tag{13}$$

and the optimal risk-sensitive state estimate is given by

$$\hat{X}_k^{rs} = e_{m^*}$$
$$m^* = \underset{m \in \{1, \ldots, N\}}{\operatorname{argmin}} \sum_{j=1}^{N} \frac{g(y_k - H(e_j))}{g(y_k)}$$
$$\exp\left(\frac{\theta}{2}(e_j - e_m)'Q(e_j - e_m)\right)q_k(j) \tag{14}$$

With this brief recapitulation of risk-neutral and risk-sensitive estimation for HMMs, we now define the mixed estimation problem.

The objective of the mixed estimation problem for the HMM described in Section 2. with $v_k$ in (1) random as described in the beginning of this section, is to find

$$\hat{X}_k^{mix} = \underset{\eta \in \mathcal{E}}{\operatorname{argmin}} E[(X_k - \eta)'Q(X_k - \eta) \mid \mathcal{Y}_k], \quad k \geq 0 \tag{15}$$

such that the following constraint is satisfied by the worst case cost for $k \geq 0$:

$$\max_{i_0, \ldots, i_k}\left[\sum_{l=0}^{k}\phi(e_{i_l}, \hat{X}_l^{mix}) - [\beta(e_{i_0}) \right.$$
$$\left. + \frac{1}{\mu}\left\{\sum_{l=0}^{k-1}U(e_{i_l}, e_{i_{l+1}}) + \sum_{l=0}^{k}V(y_l, e_{i_l})\right\}\right] \leq 0 \tag{16}$$

In the next subsection, we present the solution to the mixed estimation problem.

3476

## Solution to the mixed estimation problem for HMMs

Define $\mathcal{E}_k \subseteq \mathcal{E}$, $k \geq 0$ to be $\{e_m : \max_j \left[ s_k(j) + \frac{1}{2}(e_j - e_m)'Q(e_j - e_m) - \frac{1}{\mu}V(y_k, e_j) \right] \leq 0\}$. where $s_k(j)$ is as defined in (5) and $V(.,.)$ is as defined in (4).

Then, the state estimate $\hat{X}_k^{mix}$ for the mixed estimation problem is given by the following theorem:

**Theorem 4.1** *Consider the HMM signal model defined in Section 2 and the mixed estimation objective defined by (15), (16). Then the state estimate $\hat{X}_k^{mix}$ for the mixed estimation problem is given by*

$$\hat{X}_k^{mix} = \operatorname*{argmin}_{\eta \in \mathcal{E}_k} \sum_{i=1}^{N} (e_i - \eta)'Q(e_i - \eta)\alpha_k(i) \qquad (17)$$

**Proof:** The proof is rather straightforward once we note that $\mathcal{E}_k$ just denotes the admissible set for the state estimates such that the constraint on the worst case cost (16) is satisfied. One then applies (9) to obtain (17). ∎

## 5 Simulation results

In this section, we present some simulation results to demonstrate the differences among these different estimation methods, i.e., minimax, risk-sensitive and mixed estimation methods for a given HMM. The HMM under investigation has 10 states with the following $A$ where $a_{ii} = 0.19$, $a_{ij} = 0.09$, $\forall i \neq j$, $i, j \in \{1, 2, \ldots, 10\}$. The observation model is scalar where $v_k$ is Gaussian distributed. Note that when the measurement disturbance is bounded, it makes sense to use the robust estimation method in the minimax sense as presented in Section 3. Risk-sensitive or risk-neutral methods become suboptimal in that case. However, we ran some simulations with $v_k$ being a truncated Gaussian noise such that $|v_k| \leq 5\sigma$, where $\sigma$ is the standard deviation of the Gaussian distribution. The observed signal i.e., $H(x_k)$ is $H'X_k$ where $H = (1 2 3 4 5 6 7 8 9 10)'$. The performance criterion is the average squared error $\frac{1}{T}\sum_{k=1}^{T}(X_k - \hat{X}_k)'Q(X_k - \hat{X}_k)$ where $\hat{X}_k$ represents the risk-neutral, risk-sensitive or the minimax state estimate, as the case may be (with some abuse of notation) and $Q = HH'$.
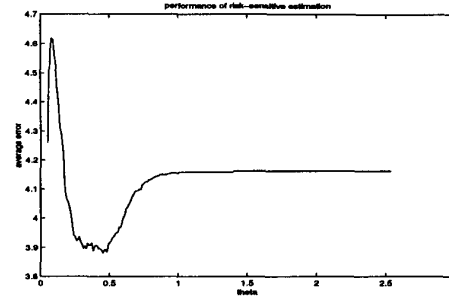


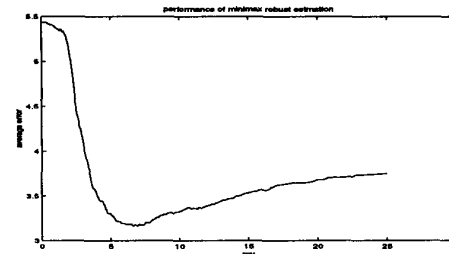**Figure 1:** Performance of the risk-sensitive filter for bounded measurement noise



**Figure 2:** Performance of the minimax robust filter for bounded measurement noise

Figure 1 shows how the suboptimal risk-sensitive filter performs with different values of $\theta$. Figure 2 shows the performance of the minimax robust estimate against various values of $\mu$. The risk-neutral average error was found to be 3.3555 over a run of $T = 1000$ time points.

We also simulated the mixed estimation algorithm. Note that when $v_k$ is purely random and the statistical information about $v_k$ is accurately known, the error performance achieved by the mixed algorithm is lower bounded by that of the risk-neutral estimation, since the mixed estimation optimizes over a constrained state space whereas the risk-neutral algorithm optimizes over the complete state space of which the constrained state space is only a subset. However, when $v_k$ contains a mixture of random and unknown but bounded noise or the statistical information about the noise is not accurately known, we observed that the constraint on the worst case cost (16) is repeatedly violated by the risk-neutral estimation scheme whereas the mixed estimation scheme always satisfies the constraint. We also observed similar results when the noise was generated according to a uniform distribution but was assumed to

3477

be Gaussian instead. Note that both the risk-neutral and the mixed estimation schemes become suboptimal for such uncertainties. However, the mixed estimation scheme guarantees an upper bound on the worst case cost whereas the risk-neutral estimation scheme fails to do so. We do not present any numerical results here for obvious reasons.

## 6 Conclusions

We addressed the problem of robust state estimation for hidden Markov models in this paper. We introduced a minimax robust estimation problem for HMMs with bounded uncertainties and presented a solution to this problem using the techniques of information states and forward dynamic programming methods. We also solve a mixed estimation problem that optimizes a quadratic cost with a constraint on the worst case cost. Some simulation results are presented that compare the performances of these different methods in case of uncertainties in the noise model.

## References

[1] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, pp. 257–285, 1989.

[2] P. Whittle, "Risk-sensitive linear-quadratic-Gaussian control," *Advances in Applied Probability*, vol. 13, pp. 764–777, 1981.

[3] D. Jacobson, "Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games," *IEEE Transactions on Automatic Control*, vol. 18, no. 2, pp. 124–131, 1973.

[4] A. Bensoussan and J. H. van Schuppen, "Optimal control of partially observable stochastic systems with an exponential-of-integral performance index," *SIAM J. Control Opt.*, vol. 23, pp. 599–613, 1985.

[5] M. James, J. Baras, and R. Elliott, "Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems," *IEEE Transaction on Automatic Control*, vol. 39, no. 4, pp. 780–792, 1994.

[6] J. Speyer, C. Fan, and R. Banavar, "Optimal stochastic estimation with exponential cost criteria," in *31st Conference on Decision and Control*, vol. 2, pp. 2293–2298, December 1992.

[7] S. Dey and J. Moore, "Risk-sensitive filtering and smoothing via reference probability methods," *IEEE Transaction on Automatic Control*, vol. 42, no. 11, pp. 1587–1591, 1997.

[8] S. Dey and J. Moore, "Risk-sensitive filtering and smoothing for hidden markov models," *Systems and Control Letters*, vol. 25, no. 5, pp. 361–366, 1995.

[9] R. K. Boel, M. R. James, and I. R. Petersen, "Robust filtering," in *Proceedings of the Conference on Decision and Control*, (San Diego), December 1997.

[10] K. Glover and J. C. Doyle, "State-space formulae for all stabilizing controllers that satisfy an $H_\infty$ norm bound and relations to risk-sensitivity," *Systems and Control Letters*, vol. 11, pp. 167–172, 1988.

[11] J. B. Moore, R. J. Elliott, and S. Dey, "Risk-sensitive generalizations of minimum variance estimation and control," in *IFAC Symposium on Nonlinear Control Systems Design*, (Tahoe City, CA), pp. 465–470, June 1995. also see summary in Journal of Mathematical Systems, Estimation and Control, volume 7, no: 1, pages 123-126, 1997.

[12] J. S. Baras and M. R. James, "Robust and risk-sensitive output feedback control for finite state machines and hidden markov models," *Journal of Mathematical Systems, Estimation and Control*, vol. 7, no. 3, pp. 371–374, 1997.

[13] W. Fleming and D. Hernandez-Hernandez, "Risk-sensitive control of finite state machines on an infinite horizon II." preprint, 1996.

[14] E. Fernandez-Gaucherand and S. I. Marcus, "Risk-sensitive optimal control of hidden markov models: structural results," *IEEE Transactions on Automatic Control*, vol. 42, no. 10, pp. 1418–1422, 1997.

[15] B. Hassibi and T. Kailath, "On nonlinear filters for mixed $H_2/H_\infty$ estimation," in *Proceedings of the American Control Conference*, (Albuquerque, New Mexico), June 1997.

[16] S. Coraluppi and S. I. Marcus, "Mixed risk-neutral/minimax control of Markov decision processes," in *Proceedings 31st Conference on Information Sciences and Systems*, (Baltmore, MD), March 1997.