Estimation of Hidden Markov Models: Risk-Sensitive Filter Banks and Qualitative Analysis of Their Sample Paths

Vahid Reza Ramezani and Steven I. Marcus, *Fellow, IEEE*

***Abstract*—A sequential filtering scheme for the risk-sensitive state estimation of partially observed Markov chains is presented. The previously introduced risk-sensitive filters are unified in the context of risk-sensitive maximum *a posterior* probability estimation. Structural results for the filter banks are given. The influence of the availability of information and the transition probabilities on the decision regions and the behavior of risk-sensitive estimators are studied.**

***Index Terms*—Hidden Markov models (HMMs), risk-sensitive estimation.**

I. Introduction

THE exponential (risk-sensitive) criterion of a quadratic function of the state and control for full state control was first proposed by Jacobson [2]. Whittle [13] produced the controller for the linear/quadratic partial observation case which required a state estimator *separated* from the control policy in a fashion analogous to the policy separation for the partial observation in the risk-neutral case. Speyer [6] treated the estimation problem for the linear/quadratic partial observation case and showed that a linear estimator is optimal among all nonlinear and linear estimators. The nonlinear discrete-time stochastic problem for the partially observed control problem was solved by a change of measure technique [4]. This so-called *reference probability* approach was later used by Dey and Moore [7] to solve an estimation problem for a *partially observed Markov chain* or, as it is commonly referred to in signal processing literature, a *hidden Markov model* (HMM). (Also see [14]). The robustness of the risk sensitive estimation under certain types of model uncertainties was shown by Boel *et al.* [9].

It is often said that risk-sensitive filters take into account the "higher order" moments of the estimation error. Roughly speaking, this follows from the analytic property of the exponential $e^x = \sum_{k=0}^{\infty} x^k/k!$ so that if Ψ stands for the sum of the error functions over some interval of time then

$$E[\exp(\gamma\Psi)] = E\left[1 + \gamma\Psi + (\gamma)^2(\Psi)^2/2 + \cdots\right].$$

Manuscript received July 10, 2001; revised May 3, 2002 and July 8, 2002. Recommended by Associate Editor Q. Zhang. This work was supported in part by the National Science Foundation under Grant DMI-9988867, by the Air Force Office of Scientific Research under Grant F496200110161, and by the Office of Naval Research (ONR) under Contract 01-5-28834 under the MURI Center for Auditory and Acoustics Research.

V. R. Ramezani is with the Institute for Systems Research, University of Maryland, College Park, MD 20742 USA (e-mail: rvahid@isr.umd.edu).

S. I. Marcus is with the Institute for Systems Research and the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: marcus@isr.umd.edu).

Digital Object Identifier 10.1109/TAC.2002.805686

Thus, at the expense of the mean error cost, the higher order moments are included in the minimization of the expected cost, reducing the "risk" of large deviations and increasing our "confidence" in the estimator. The parameter $\gamma > 0$ controls the extent to which the higher order moments are included. In particular, the first order approximation, $\gamma \to 0$, $E[\exp(\gamma\Psi] \cong 1 + \gamma E\Psi$, indicates that the original minimization of the sum criterion or the risk-neutral problem is recovered as the small risk limit of the exponential criterion.

Our work can be traced back to Speyer's paper [6] and is related to the Dey–Moore filter [7]. We are also interested in the estimation problem with the exponential criterion for HMMs. However, we have a different perspective; we view the estimator as a dynamical system whose dynamics are inherited from the Markov chain through the partial observation and an optimization criterion. We are not only interested in the computation of the optimal estimator for an HMM and its properties for exponential criterion, but also in the behavior of its sample paths.

In Section II, we introduce a filtering scheme for the risk sensitive maximum *a posterior* probability (MAP) estimation for HMMs. It can be considered as the discrete time finite dimensional generalization of the risk-sensitive filter for linear/quadratic partial observation case [6] and the "single-step" quadratic risk-sensitive filter introduced in [7].

In Section III, "structural results" are given which relate risk-sensitivity to the dynamics of the underlying process and expose relations among the transition probabilities, risk-sensitivity and the *decision regions*.

In Section IV, in the context of risk-sensitive MAP for HMMs, we clarify and quantify the influence of risk-sensitivity on the behavior of the sample paths of the estimator; we give risk-sensitivity a "flow" interpretation.

II. Risk Sensitive Filter Banks

A. The Estimation of Hidden Markov Models

Define an HMM as a five-tuple $\langle \mathbf{X}, \mathbf{Y}, \mathbf{X}, \mathcal{A}, Q\rangle$; here $\mathcal{A}$ is the transition matrix, $\mathbf{Y} = \{1, 2, \ldots, N_{\mathbf{Y}}\}$ is the set of observations and $\mathbf{X} = \{1, 2, \ldots, N_{\mathbf{X}}\}$ is the finite set of (internal) states as well as the set of estimates or decisions. In addition, we have that $Q := [q_{x,y}]$ is the $N_{\mathbf{X}} \times N_{\mathbf{Y}}$ state/observation matrix, i.e., $q_{x,y}$ is the probability of observing y when the state is x. We consider the following information pattern. At decision epoch t, the system is in the (unobservable) state $X_t = i$ and the corresponding observation Y_t is gathered, such that

$$P(Y_t = j | X_t = i) = c_{i,j}. \tag{1}$$

The estimators V_t are functions of observations $(Y_0, \ldots Y_t)$ and are chosen according to some specified criterion. Throughout this paper, we use upper case letters to denote estimators and script upper case letters to denote "estimation maps" from observations to the set $\mathbf{X}$. If Y_t is an observation and V_t an estimator: $V_t = \mathcal{V}_t \circ Y_t$. When it causes no confusion, we may use upper case letters for both.

B. MAP for HMMs

Consider a sequence of finite-dimensional random variables X_t and the corresponding observations Y_t defined on the common probability space $(\mathbf{\Omega}, F, \mathbf{P})$. The MAP estimator $\hat{X}_t$ is a Borel measurable function of the filtration generated by observations up to Y_t denoted by $\mathcal{Y}_t$ which satisfies for $\omega \in \Omega$

$$\hat{X}_t(\omega) = \operatorname*{argmin}_{\zeta \in \mathbf{X}} E[\rho(X_t, \zeta) \,|\, \mathcal{Y}_t = \mathcal{Y}_t(\omega)] \tag{2}$$

where

$$t = 0, 1, \ldots \qquad \rho(X_t, \zeta) = \begin{cases} 0 & \text{if } X_t = \zeta \\ 1 & \text{otherwise.} \end{cases}$$

The usual definition of MAP as the argument with the greatest probability given the observation follows from the above [10]. We will need the following simple lemma.

Lemma 1: MAP also results from the additive cost minimization

$$\left(\hat{X}_0, \ldots, \hat{X}_N\right)(\omega) = \operatorname*{argmin}_{\zeta_0 \cdots \zeta_N \in \mathbf{X}^N} E\left[\sum_{i=0}^{N} \rho(X_i, \zeta_i) \,\middle|\, \mathcal{Y}_i = \mathcal{Y}_i(\omega)\right] \tag{3}$$

where $\mathbf{X}^N$ is the product space and each $\hat{X}_i$ is $\mathcal{Y}_i$ measurable.

Proof: The proof follows easily from the linearity of the conditional expectation and term by term minimization of the resulting sum. ■

C. Change of Measure

To carry out the computations, we will use a change of measure technique introduced in [3] and [4]. Let $(\mathbf{\Omega}, F, \mathbf{P})$ be the canonical probability space on which all of our time series are defined. Let $\mathcal{Y}_t$ be the filtration generated by the available observations up to decision epoch t, and let $\mathcal{G}_t$ be the filtration generated by the sequence of states and observations up to that time. Then a new probability measure $\mathcal{P}^\dagger$ is defined by the restriction of the Radon–Nikodym derivative on $\mathcal{G}_t$ to

$$\left.\frac{d\mathcal{P}}{d\mathcal{P}^\dagger}\right|_{\mathcal{G}_t} = \lambda_t := N_{\mathbf{Y}}^t \cdot \prod_{k=1}^{t} q_{X_k, Y_k} \tag{4}$$

under which $\{Y_t\}$ is independently and identically distributed (i.i.d). Each distribution is uniform over the set $\mathbf{Y}$ and $\{Y_t\}$ is independent of $\{X_t\}$. That such a measure exists follows directly from the Kolmogorov extension theorem.

Before we can introduce our filter, we need an optimization result. Let V_t be measurable functions of observations up to t taking values in $\{X_t\}$ and $\rho(\cdot, \cdot)$ as above. Fix $V_0, \ldots, V_{k-1}$. We would like to find $\hat{V}_k, \ldots, \hat{V}_{H-1}$ such that the following criterion is minimized:

$$S^\gamma(V_0, \ldots, V_{H-1}) := E[\exp(\gamma \cdot \mathcal{C}_H)] \tag{5}$$

where

$$\mathcal{C}_H := \sum_{t=0}^{H-1} \rho(X_t, V_t) \tag{6}$$

and γ is a strictly positive parameter. The minimum value will be denoted by $\overline{S}^\gamma(\hat{V}_k, \ldots, \hat{V}_{H-1})$.

This optimization problem can be solved via dynamic programming. We need to define recursively an *information state*

$$\sigma_{t+1}^\gamma = N_{\mathbf{Y}} \cdot \overline{Q}(Y_{t+1}) \mathcal{D}^T(V_t) \cdot \sigma_t^\gamma \tag{7}$$

where $\overline{Q}(y) := diag(q_{i,y})$, $\mathcal{A}^T$ denotes the transpose of the matrix $\mathcal{A}$ and the matrix $\mathcal{D}$ is defined by

$$[\mathcal{D}(v)]_{i,j} := a_{i,j} \cdot \exp(\gamma \rho(i, v)). \tag{8}$$

σ_0^γ is set equal to $N_{\mathbf{Y}} \cdot \overline{Q}(Y_0) p_0$, where p_0 is the initial distribution of the state and is assumed to be known. In the context of the risk-sensitive estimation of Markov chains, the meaning of σ_t^γ will become clear.

We define the matrix

$$L(v, y) := N_{\mathbf{Y}} \cdot \overline{Q}(y) \mathcal{D}^T(v). \tag{9}$$

It can be shown that [1]

$$S^\gamma(V_0, \ldots, V_{H-1}) = E[\exp(\gamma \cdot \mathcal{C}_H)] = E^\dagger\left[\sum_{i=1}^{N_X} \sigma_H^\gamma(i)\right] \tag{10}$$

where $E^\dagger$ is the expectation with respect to the new measure. Define the value functions $J^\gamma(\cdot, H-j)\colon R_+^{N_{\mathbf{X}}} \to R$, $j = H, \ldots, H-k$, as

$$J^\gamma(\sigma, H-j) := \min_{V_{H-j} \cdots V_{H-1}} \left\{ E^\dagger \left\{ \sum_{i=1}^{N_X} \sigma_H^\gamma(i) \,\middle|\, \sigma_{H-j}^\gamma = \sigma \right\} \right\}. \tag{11}$$

Lemma 2: Let $V_0, \ldots, V_{k-1}$ be given. Then the value functions defined above are obtained from the following dynamic programming equations:

$$\begin{cases} \overline{J}^\gamma(\sigma, H) = \displaystyle\sum_{i=1}^{N_{\mathbf{X}}} \sigma(i); \\ \overline{J}^\gamma(\sigma, H-j) \\ \quad = \displaystyle\min_{v \in \mathbf{X}} \left\{ E^\dagger \left[\overline{J}^\gamma(L(v, Y_{H-j+1}) \cdot \sigma, H-j+1) \right] \right\} \end{cases}$$
$$j = 1, 2, \ldots, H-k. \tag{12}$$

The estimation maps $\hat{\mathcal{V}}_k, \ldots, \hat{\mathcal{V}}_{H-1}$ obtained from (12) are risk optimal, i.e., $\hat{\mathcal{V}}_k(\sigma_k^\gamma), \ldots, \hat{\mathcal{V}}_{H-1}(\sigma_{H-1}^\gamma)$ achieve the minimum in (5).

Proof: See [1]. ■

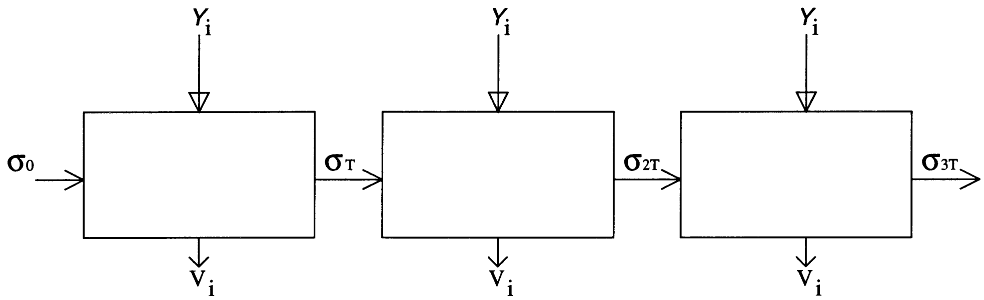

Fig. 1. The T-step risk sensitive filter banks.

D. The T-Step MAP (TMAP) Risk-Sensitive Estimator

The TMAP risk-sensitive estimator is defined by the following criterion:

$$\hat{V}_{NT}, \ldots, \hat{V}_{(N+1)T-1} = \underset{V_t \in X,\, t=NT, \ldots, (N+1)T-1}{\operatorname{argmin}} \cdot E\left\{\exp\left(\gamma \cdot \left[\sum_{t=0}^{NT-1} \rho\left(X_t, \hat{V}_t\right) + \sum_{t=NT}^{(N+1)T-1} \rho(X_t, V_t)\right]\right)\right\} \quad (13)$$

where V_t is $\mathcal{Y}_t$ measurable, T is the size of the filter and $N = 0, 1, \ldots$, is the index of filtering segments. This exponential criterion is a generalization of the risk-sensitive filtering idea introduced in [7] for the quadratic cost with the filtering performed in single steps, i.e., for $T = 1$; we will look at this special case for TMAP in Sections III and IV and show that it is essentially a "greedy algorithm."

Theorem 1: The TMAP can be computed recursively by the following procedure.

1) Set $\sigma_0 = N_{\mathbf{Y}} \cdot \overline{Q}(Y_0)p_0$.
2) Given σ_{NT}, use the minimizing sequence of the value functions obtained from the following dynamic programming equations:

$$\begin{cases} \overline{J}^\gamma(\sigma, T) = \sum_{i=1}^{N_{\mathbf{X}}} \sigma(i); \\ \overline{J}^\gamma(\sigma, T-j) \\ \quad = \min_{v \in \mathbf{X}} \left\{E^\dagger\left[\overline{J}^\gamma(L(v, Y_{NT+T-j+1}) \cdot \sigma, T-j+1)\right]\right\} \end{cases}$$
$$j = 1, 2, \ldots, T \quad (14)$$

to determine the value of the optimum estimates $\hat{V}_{NT}, \ldots, \hat{V}_{(N+1)T-1}$ as a function of the information state $\sigma_{NT}, \ldots, \sigma_{(N+1)T-1}$ obtained by (7).
3) Apply (7) once more to obtain $\sigma_{(N+1)T}$ and repeat steps 2) and 3) starting at $(N+1)T$.

Furthermore, for any given N as $\gamma \to 0$, TMAP (the previous algorithm) reduces to MAP.

Proof: The proof follows from repeated applications of Lemma 2. We will skip the details. The limiting result follows from the first order approximation of the exponential function and the observation that as $\gamma \to 0$, the matrix $\mathcal{D}(v) \to \mathcal{A}$ element wise. This implies that in the limit the input to each filtering step is the unnormalized conditional distribution and thus by Lemma 1 the filtering process reduces to the well-known MAP estimation of HMMs. ∎

Note that although the size of the sum $\sum_{i=0}^{NT-1} \rho(X_t, \hat{V}_t)$ increases with N, all we need to track is the information state, computed recursively. The optimal estimates $\hat{\mathcal{V}}_{NT}(\sigma_{NT}), \ldots, \hat{\mathcal{V}}_{(N+1)T-1}(\sigma_{(N+1)T-1})$ are measurable functions of the information state alone. Since our Markov chain is homogeneous and under the new measure the observations are i.i.d., (14) depends only on T and not on N. This justifies the cascade filter banks representation of Fig. 1.

We point out that Theorem 1 has no control counterpart. In this case, the estimators $\{V_t\}$ have no influence on the dynamics and thus estimation can be broken down into separate segments with the information state reinitialized. The same cannot be said for a controlled Markov chain due to the influence of the controller on the dynamics; the separate segments cannot be joined to represent the entire process in the above limiting sense as the "decoupling" Lemma 1 no longer holds. Note also that controlled Markov chains are not homogeneous.

III. Structural Results: The Filter Banks and the Information State

It is clear from the aforementioned information that to describe the behavior of TMAP we must understand the operation of each filtering segment and understand the meaning of the information state. The key in understanding the filter's operation is the analysis of the value functions which are obtained via dynamic programming.

Lemma 3: The value functions are continuous and concave functions of the information state $\sigma \in R_+^{N_{\mathbf{X}}}$.

Proof: Both statements are proved by induction. The continuity when $j = 0$ is obvious. Since $\{Y_i\}$ is i.i.d., uniformly distributed and finite dimensional, then (13) is taking a minimum over the average of composition of functions, each of which is continuous by the continuity of the linear functions and the induction hypothesis, and is, therefore, continuous. For the second statement, once again the case $j = 0$ is trivially verified. Assume concavity for $j - 1$. Let $0 \leq \lambda \leq 1$ and $\sigma_1, \sigma_2 \in R_+^{N_{\mathbf{X}}}$; define $\tilde{\sigma} := \lambda\sigma_1 + (1-\lambda)\sigma_2$. Then, by the induction hypothesis and by (13), we have (the first equality follows from the i.i.d. properties of the observations under the new measure)

$$\begin{aligned}\overline{J}^{\gamma}(\tilde{\sigma}, T-j) &= \min_{v\in\mathbf{X}}\left\{\frac{1}{N_{\mathbf{Y}}}\sum_{y=1}^{N_{\mathbf{Y}}}\overline{J}^{\gamma}(L(v, y)\cdot\tilde{\sigma}, T-j+1)\right\}\\ &\geq \min_{v\in\mathbf{X}}\left\{\frac{1}{N_{\mathbf{Y}}}\sum_{y=1}^{N_{\mathbf{Y}}}\left[\lambda\overline{J}^{\gamma}(L(v, y)\cdot\sigma_1, T-j+1)\right.\right.\\ &\qquad\left.\left. +(1-\lambda)\overline{J}^{\gamma}(L(v, y)\cdot\sigma_2, T-j+1)\right]\right\}\\ &\geq \lambda\overline{J}^{\gamma}(\sigma_1, T-j)+(1-\lambda)\overline{J}^{\gamma}(\sigma_2, T-j).\end{aligned}$$

■

Next, for P a finite set of vectors in $R_+^{N_\mathbf{X}}$, denote by $\mathcal{O}(P)$ the set

$$\mathcal{O}(P) := \left\{\frac{1}{N_{\mathbf{Y}}}\sum_{y=1}^{N_{\mathbf{Y}}}\alpha_y\cdot L(v, y)\,\middle|\,\alpha_y\in P, v\in\mathbf{X}\right\}. \quad (15)$$

Note that if P is finite, then so is $\mathcal{O}(P)$, since $|\mathcal{O}(P)| \leq |P|^{N_\mathbf{Y}}\cdot N_\mathbf{X}$.

Lemma 4: The value functions given by (14) are piecewise linear functions (hyper-planes through the origin) of $\sigma\in R_+^{N_\mathbf{X}}$, such that if P_{j-1} indicates the vectors in $R_+^{N_\mathbf{X}}$ which specify the set of hyper planes for $\overline{J}^{\gamma}(\sigma, T-j+1)$ then

$$\begin{aligned}\overline{J}^{\gamma}(\sigma, T-j+1) &= \min_{\alpha\in P_{j-1}}\{\alpha\cdot\sigma\}\\ \overline{J}^{\gamma}(\sigma, T-j) &= \min_{\alpha\in\mathcal{O}(P_{j-1})}\{\alpha\cdot\sigma\}\end{aligned} \quad (16)$$

where $P_0 = \overline{\mathbf{1}} := (\sum_{k=1}^{N_\mathbf{X}} e_k)^\mathbf{T}$ and $\{e_k\}$ are the unit vectors in $R^{\mathbf{N_x}}$.

Proof: The statement of the lemma is readily verified for $j=0$. Assume the lemma holds for $j-1$ then piecewise linearity implies that for each vector α_0 in P_{j-1} there is a point $\sigma^*\in R^{\mathbf{N_x}}$ and a disk $d(\sigma^*, r)$ such that on this disk $\overline{J}^{\gamma}(\sigma, T-j+1) = \alpha_0\cdot\sigma$. Consider a different point σ and $0<t\leq 1$ small enough so that $t(\sigma-\sigma^*)+\sigma^*\in d(\sigma^*, r)$. Then, by the concavity shown as before

$$\begin{aligned}&\overline{J}^{\gamma}(\sigma^*+t(\sigma-\sigma^*), T-j+1)\\ &\quad=\overline{J}^{\gamma}((1-t)\sigma^*+t\sigma, T-j+1)\\ &\quad\geq(1-t)\overline{J}^{\gamma}(\sigma^*, T-j+1)+t\overline{J}^{\gamma}(\sigma, T-j+1).\end{aligned}$$

Since $\overline{J}^{\gamma}(\sigma^*+t(\sigma-\sigma^*), T-j+1) = \alpha_0\cdot\sigma^*+t(\sigma-\sigma^*)$, after substitution and cancellations, we get

$$\overline{J}^{\gamma}(\sigma, T-j+1)\leq\alpha_0\cdot\sigma.$$

However, α_0 was arbitrary and so we have the first equality. To prove the second equality, note that

$$\begin{aligned}\overline{J}^{\gamma}(\sigma, T-j) &= \min_{v\in\mathbf{X}}\left\{\frac{1}{N_{\mathbf{Y}}}\sum_{y=1}^{N_{\mathbf{Y}}}\min_{\alpha\in P_{j-1}}\{\alpha\cdot L(v, y)\cdot\sigma\}\right\}\\ &= \min_{v\in\mathbf{X}}\left\{\left[\frac{1}{N_{\mathbf{Y}}}\sum_{y=1}^{N_{\mathbf{Y}}}\tilde{\alpha}(v, y, \sigma)\cdot L(v, y)\right]\cdot\sigma\right\}\\ &= \min_{\alpha\in\mathcal{O}(P_{j-1})}\{\alpha\cdot\sigma\}\end{aligned}$$

where $\tilde{\alpha}(v, y, \sigma)\in P_{j-1}$ minimizes $\alpha\cdot L(v, y)\cdot\sigma$ in the first equality, and the last equality follows since $\alpha\cdot L(u, y)\cdot\sigma > \tilde{\alpha}(v, y, \sigma)\cdot L(u, y)\cdot\sigma$, for all $\alpha\in P_{j-1}$, $v\in\mathbf{X}$, $y\in\mathbf{Y}$, $\sigma\in R_{N_\mathbf{X}}^+$. ■

Lemma 5: The optimal estimates $\{\hat{V}_t\}$ are constant along rays through the origin, i.e., let $\sigma\in R_+^{N_\mathbf{X}}$ then $\hat{\mathcal{V}}_t(\sigma') = \hat{\mathcal{V}}_t(\sigma)$, for all $\sigma' = \lambda\sigma$, $\lambda > 0$.

Proof: From Lemma 4, we see that $\overline{J}^{\gamma}(\sigma', T-j) = \lambda\overline{J}^{\gamma}(\sigma, T-j)$. The result follows from Theorem 1. ■

Definition 1: A cone in $R_+^{N_\mathbf{X}}$ is a set defined by $C_S := \{\sigma|\sigma=\lambda x, x\in S\subset R_+^{N_\mathbf{X}}, \lambda>0\}$.

Definition 2: For $j=1, 2, \ldots, T$ and $v\in X$, let

$$\begin{aligned}\overline{J}_v^{\gamma}(\sigma, T-j) &:= E^{\dagger}\left[\overline{J}^{\gamma}(L(v, Y_{T-j+1})\cdot\sigma, T-j+1)\right]\\ &= \frac{1}{N_{\mathbf{Y}}}\sum_{y=1}^{N_{\mathbf{Y}}}\left[\overline{J}^{\gamma}(L(v, y)\cdot\sigma, T-j+1)\right].\end{aligned} \quad (17)$$

Definition 3: The decision region $DR_v^j\subset R_+^{N_\mathbf{X}}$ for the estimate $v\in\mathbf{X}$, at the $T-j$ decision epoch, is defined as

$$DR_v^j := \left\{\sigma\,\middle|\,\sigma\in R_+^{N_\mathbf{X}}, \overline{J}^{\gamma}(\sigma, T-j) = \overline{J}_v^{\gamma}(\sigma, T-j)\right\}. \quad (18)$$

It follows from the definition of $\hat{\mathcal{V}}_{NT+T-j}(\sigma)$ that

$$DR_v^j := \left\{\sigma\,\middle|\,\sigma\in R_+^{N_\mathbf{X}}, \hat{\mathcal{V}}_{NT+T-j}(\sigma) = v\right\}. \quad (19)$$

We say a decision is made "strictly," if it is the only possible decision.

Theorem 2: For each $v=i\in\{1, 2, \ldots, \mathbf{N_X}\}$ and for every $j=1, 2, \ldots, T$, the decision region DR_i^j is always nonempty and includes a cone about the σ_i axis within which the decision made is (strictly) $\hat{\mathcal{V}}_{NT+T-j}(\sigma) = i$.

Proof: We state the proof for $N_X = 2$, from which the proof for the general case will become evident for the reader. On the σ_1 axis, we have by definition

$$\begin{aligned}&\overline{J}_1^{\gamma}(\sigma, T-j)\\ &\quad=\frac{1}{N_{\mathbf{Y}}}\sum_{y=1}^{N_{\mathbf{Y}}}\left[\overline{J}^{\gamma}\left(N_\mathbf{Y}\cdot\overline{Q}(y)\mathcal{A}^T\begin{bmatrix}\sigma_1\\0\end{bmatrix}, T-j+1\right)\right]\\ &\overline{J}_2^{\gamma}(\sigma, T-j)\\ &\quad=\frac{1}{N_{\mathbf{Y}}}\sum_{y=1}^{N_{\mathbf{Y}}}\left[\overline{J}^{\gamma}\left(N_\mathbf{Y}\cdot\overline{Q}(y)\mathcal{A}^T\begin{bmatrix}e^{\gamma}\sigma_1\\0\end{bmatrix}, T-j+1\right)\right].\end{aligned}$$

Applying Lemma 4 to each term of the summation on the right-hand side of $\overline{J}_2^{\gamma}(\sigma, T-j)$, we get

$$\overline{J}_2^{\gamma}(\sigma, T-j) = \frac{e^{\gamma}}{N_{\mathbf{Y}}} \sum_{y=1}^{N_{\mathbf{Y}}} \left[\overline{J}^{\gamma} \left(N_{\mathbf{Y}} \cdot \overline{Q}(y) \mathcal{A}^T \begin{bmatrix} \sigma_1 \\ 0 \end{bmatrix}, T-j+1 \right) \right].$$

Therefore, we can write

$$\overline{J}_2^{\gamma}(\sigma, T-j) - \overline{J}_1^{\gamma}(\sigma, T-j) = (e^{\gamma}-1) \left\{ \frac{1}{N_{\mathbf{Y}}} \sum_{y=1}^{N_{\mathbf{Y}}} \cdot \left[\overline{J}^{\gamma} \left(N_{\mathbf{Y}} \cdot \overline{Q}(y) \mathcal{A}^T \begin{bmatrix} \sigma_1 \\ 0 \end{bmatrix}, T-j+1 \right) \right] \right\}.$$

However, $e^{\gamma} > 1$ since $\gamma > 0$ and for every j, the value functions are strictly positive being integrals of the exponential functions. Thus, from the aforementioned on the σ_1 axis we have the strict inequality

$$\overline{J}_1^{\gamma}(\sigma, T-j) < \overline{J}_2^{\gamma}(\sigma, T-j)$$

which implies DR_1^j includes the σ_1 axis; fix σ on that axis, then by Lemma 3 and because sums and compositions of continuous functions are continuous, there exists a disk of positive radius in the R_+^2, i.e., $d(r, \sigma) \bigcap R_+^2$, $r > 0$ such that $\overline{J}_1^{\gamma}(\sigma, T-j) < \overline{J}_2^{\gamma}(\sigma, T-j)$ for every $x \in d(r, \sigma) \bigcap R_+^2$. Therefore, Lemma 5 implies that on the cone $\mathbf{C}_{d(r, \sigma) \cap R_+^2} \subset DR_1^j$ the decision is strictly $v = 1$.

The same proof works in higher dimensions by fixing an axis σ_l and making pairwise comparisons between $\overline{J}_l^{\gamma}(\sigma, T-j)$ and $\overline{J}_k^{\gamma}(\sigma, T-j)$, $k \neq l$ along the σ_l axis. The "strict" cone around the σ_l axis will be the intersection of all cones obtained from pairwise comparisons. ■

In general, the boundaries among the decision regions are not of "threshold type" (unlike MAP). We give a two-dimensional example. We consider the TMAP with $N_{\mathbf{X}} = 2$.

Remark: The transition cone $DR_1^j \bigcap DR_2^j$ is not, in general, of threshold type, i.e., the cone: $DR_1^j \bigcap DR_2^j$ does not degenerate to a line; we give a simple counterexample.

Example 1: Let $a_{11} = a_{22} = 1$ and $q_{xy} = 1/2$; then it can be shown (by straightforward induction) that the transition cones are not degenerate.

Let us look at the counterexample more closely (Fig. 2). The cone where the decision is strictly $v = 1$, i.e., $R_+^2 \cap (DR_2^j)^c$ is given by $\sigma_1 > \sigma_2 \cdot \exp(\gamma(j-1))$ and by symmetry $R_+^2 \cap (DR_1^j)^c$ is given by $\sigma_2 > \sigma_1 \cdot \exp(\gamma(j-1))$. The transition cone (where either decision is acceptable) is given by the complement of the union of these two regions (the colored area). The value functions are given by $j+1$ hyper-planes: $\sigma_1 + \exp(\gamma(j))\sigma_2$, $\sigma_1 \exp(\gamma(1)) + \exp(\gamma(j-1))\sigma_2$, $\sigma_1 \exp(\gamma(2)) + \exp(\gamma(j-2))\sigma_2 \ldots$, $\sigma_2 + \exp(\gamma(j))\sigma_1$ on the $j+1$ cones beginning with $\sigma_1 > \sigma_2 \cdot \exp(\gamma(j-1))$ and ending with $\sigma_2 > \sigma_1 \cdot \exp(\gamma(j-1))$. The transition cone between them is the union of $j-1$ cones whose boundaries are lines $\exp(-(j-1)\gamma)\sigma_1 = \sigma_2$, $\exp(-(j-2)\gamma)\sigma_1 = \sigma_2, \ldots, \sigma_1 = \sigma_2, \ldots, \exp(-(j-2)\gamma)\sigma_2 = \sigma_1, \ldots,$

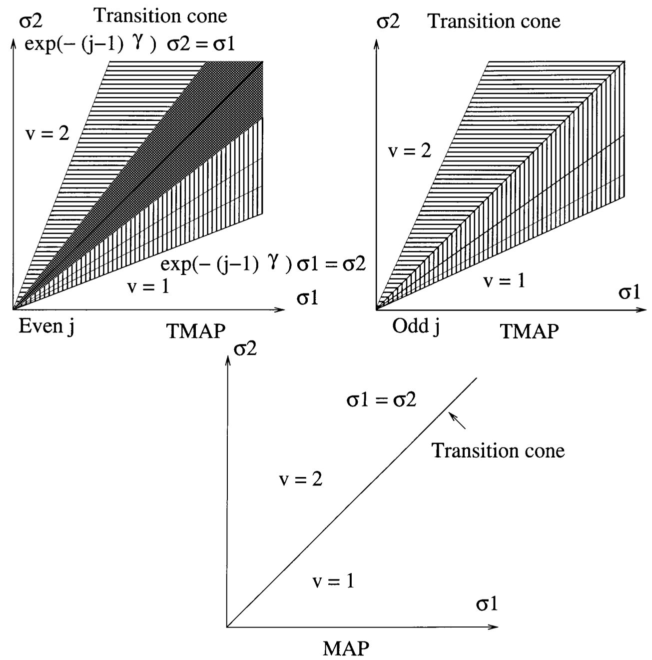

Fig. 2. Decision cones for $a_{11} = a_{22} = 1$, $q_{xy} = 1/2$.

$\exp(-(j-1)\gamma)\sigma_2 = \sigma_1$. When j is odd, the line $\sigma_1 = \sigma_2$ is a boundary (boundary in the sense of slope change in the cost and not decision); when j is even, it is not. (The solid cone which includes $\sigma_1 = \sigma_2$ for even values of j is meant to emphasize this.) On the transition cone either decision is allowed. We can interpret this region as the zone of uncertainty. For MAP and TMAP with $T = 1$ (we will show this later), this region is the threshold $\sigma_1 = \sigma_2$, but as the above example shows, for TMAP with $T > 1$, it may be a nondegenerate cone. We could interpret this as reflecting the "conservative" nature of the risk-sensitive estimation. We are expanding the zone of "uncertainty" at the expense of the region of "certainty." We will show in the subsequent sections that this is not always the manner in which risk-sensitivity manifests itself in the structure of decision regions. It is possible that the transition cone remains degenerate and either of the two other cones expands at the expense of the other or the decision regions are not affected by risk sensitivity at all, i.e., they remain identical to that of MAP. In two dimensions for example, $DR_1^j = \{\sigma | \sigma_1 > \sigma_2\}$ and $DR_2^j = \{\sigma | \sigma_2 > \sigma_1\}$.

The aforementioned theorem only guarantees the existence of nondegenerate cones around the σ_l axis but says nothing about their size. In fact, observe that in the above example the size of these cones becomes arbitrarily small as $\gamma \to \infty$ since the slopes of the lines $\exp(-(j-1)\gamma)\sigma_1 = \sigma_2$ and $\exp(-(j-1)\gamma)\sigma_2 = \sigma_1$, for every $j > 1$, converge to zero and infinity respectively.

Two special cases ($N = 0, T = M$) and ($T = 1, N = 0, \ldots, M-1$) are of interest. In both cases, the index t ranges from $t = 0$ to $t = M-1$. In the first case, TMAP reduces to the exponential/sum criterion for HMMs which is the discrete and finite dimensional version of the risk-sensitive L^2 filter introduced by Speyer and others. The second case would be the MAP version of the quadratic cost risk-sensitive filtering introduced first to our best knowledge by Dey and Moore in [7]. Obviously, the structural results obtained so far apply to these special cases.

Let $\mathcal{EX}(v)$ be the diagonal matrix $diag[\exp(\gamma\rho(i, v))]$, $i = 1, \ldots, N_{\mathbf{X}}$.

Theorem 3: The one step TMAP decision regions are given by

$$\hat{\mathcal{V}}_t(\sigma) = i \qquad \text{if } \sigma_i \geq \sigma_j,\ \forall j \neq i.$$

Proof: From the definition, we have $\mathcal{D}^T(v) = \mathcal{A}^T\mathcal{EX}(v)$ and that $\overline{J}^\gamma(\sigma, T) = \langle \sigma, \overline{\mathbf{1}}\rangle$ and, thus

$$\begin{aligned}\overline{J}^\gamma_v(\sigma, T-1) &= \frac{1}{N_{\mathbf{Y}}}\sum_{y=1}^{N_{\mathbf{Y}}}\left[\overline{J}^\gamma(L(v, y)\cdot\sigma, T)\right]\\ &= \frac{1}{N_{\mathbf{Y}}}\sum_{y=1}^{N_{\mathbf{Y}}}\left[\overline{J}^\gamma(N_{\mathbf{Y}}\cdot\overline{Q}(y)\mathcal{A}^T\mathcal{EX}(v)\cdot\sigma, T)\right]\\ &= \sum_{y=1}^{N_{\mathbf{Y}}}\left\langle\overline{Q}(y)\left(\mathcal{A}^T\mathcal{EX}(v)\cdot\sigma\right), \overline{\mathbf{1}}\right\rangle\\ &= \left\langle\left(\sum_{y=1}^{N_{\mathbf{Y}}}\overline{Q}(y)\right)\left(\mathcal{A}^T\mathcal{EX}(v)\cdot\sigma\right), \overline{\mathbf{1}}\right\rangle\\ &= \left\langle I\mathcal{A}^T\mathcal{EX}(v)\cdot\sigma, \overline{\mathbf{1}}\right\rangle\\ &= \left\langle\mathcal{EX}(v)\cdot\sigma, \overline{\mathbf{1}}\right\rangle. \end{aligned} \tag{20}$$

A little calculation shows that given σ, this is minimized, if we set v equal to the index of the largest component of σ, i.e., if σ_l is the largest component of σ then $v = l$. This is precisely how the decision regions for MAP are defined. ■

Note that TMAP for $T = 1$ is not reduced to MAP; although the decision regions are the same, the information states are different. In the case of MAP, the information state is the conditional distribution, while in the case of TMAP the information state is given by (7). The conditional distribution has no memory of decisions made in the past while the TMAPs information state does depend on these decisions. On the other hand, when γ is very small, (7) becomes the unnormalized conditional distribution. We can think of TMAPs information state as the conditional distribution modified by the sequence of the decisions made in the past. This modified information state is then put through the next decision region which itself is calculated to minimize a certain cost structure based on the averaged behavior of the future sample paths: $(E^\dagger[\overline{J}^\gamma(L(v, Y_{NT+T-j+1})\cdot\sigma, T - j + 1)])$. How far we look into the "averaged future" is determined by T. The information from the past is encapsulated in the information state. These ideas will become quantified later.

Theorem 4: The value function $\overline{J}^\gamma(\sigma, T-j)$ when $\forall(x, y)$, $q_{yx} = 1/N_{\mathbf{Y}}$ is given by

$$\begin{aligned}\overline{J}^\gamma(\sigma, T-j) = \min_{v_{T-j}, \ldots, v_{T-1}} &\langle\sigma, \mathcal{EX}(v_{T-j})\\ &\cdot\mathcal{AEX}(v_{T-j+1})\cdots\mathcal{AEX}(v_{T-1})\overline{\mathbf{1}}\rangle.\end{aligned}$$

The proof is based on the same technique used in the previous theorem and will be omitted.

In the aforementioned counterexample when $\mathcal{A} = I_{2\times 2}$ and $q_{yx} = 1/2$, by the previous theorem, we have

$$\begin{aligned}&\overline{J}^\gamma(\sigma, T-j)\\ &\quad= \min_{v_{T-j}, \ldots, v_{T-1}}\left\langle\sigma, \mathcal{EX}(v_{T-j})\cdot\cdots\mathcal{EX}(v_{T-j})\cdot\overline{\mathbf{1}}\right\rangle.\end{aligned}$$

If we let the number of times we choose $x = 2$ be n_2, and likewise for $x = 1$, $n_1 = T - n_2$, a little algebra will show that the total cost $\overline{J}^\gamma(\sigma, 0)$ is given by

$$\sigma_1\exp\{\gamma n_2\} + \sigma_2\exp\{\gamma(T - n_2)\}.$$

By differentiation with respect to n_2, a few rearrangements and taking logarithms, the minimum cost is obtained when (modulo the integer parts)

$$T/2 - \frac{1}{2\gamma}\log(\sigma_1/\sigma_2) = n_2; \qquad 0 \leq n_2 \leq T. \tag{21}$$

This suggests that for large values of γ regardless of σ, we choose the two states an approximately equal number of times. To see why this limiting behavior occurs first let $\sigma_1 = \sigma_2 = 1/2$, then according to the decision regions, we could choose either 1 or 2. Suppose we choose 1, then the information state evolves to $(\sigma_1, \sigma_2 e^\gamma) = (1/2, e^\gamma/2)$. We could continue to choose $v = 1$ and "move up" on the information state (successive) 2-dimensional planes. Note that the upper boundary of the transition cone is given by $\sigma_2\exp(\gamma) + \exp(\gamma(j-1))\sigma_1$. Thus, going forward in time, the information state moves up while the upper boundary, $\sigma_2\exp(\gamma)+\exp(\gamma(j-1))\sigma_1$, "moves down" since going forward, $j - 1$, shown before, changes to $(j-1) - 1 = j - 2$. Therefore, at about $j/2$ the information state moves out of the transition cone and enters the upper cone where we will have to choose 2 as our estimate. A similar argument shows that from that point on, the information state "moves forward" on the information state (successive) two-dimensional planes and as the upper boundary $\sigma_2\exp(\gamma)+\exp(\gamma(j-1))\sigma_1$ continues to decrease in slope, the information state σ remains on the decision region of the state 2. Hence, the decision sequence $\{1, 1, 1, \ldots 2, 2, 2\}$ with about the same number of 1s and 2s (exactly if j is even) is a possible solution. When γ is large, every point not on the axes falls into the transition cone and a similar argument shows that it follows the same decision pattern.

Now, consider the case $T = 1$. In this case the transition cone is reduced for all values of j, to the line $\sigma_1 = \sigma_2$ and a similar argument shows that for large values of γ the decision is $\{1, 2, 1, 2, 1, 2\ldots\}$. This simple example provides insight as to how the two special cases differ. The decision regions for $T = M > 1$ are more complex but it appears that this allows for a "smoother" solution. Note that if we let $\gamma \to 0$, assuming $\sigma_1 > \sigma_2$, $n_2 \to 0$ as expected.

Definition 4: A completely symmetric HMM is defined as an HMM whose transition probability matrix is symmetric in the sense that $a_{ii} = 1 - \epsilon$, $a_{ji} = \epsilon/(N_{\mathbf{X}} - 1)$ for all $j \neq i$, $N_{\mathbf{X}} = N_{\mathbf{Y}}$ (so that Q is a square matrix) and $q_{yx} = q$ for $x = y$ and otherwise $q_{yx} = (1-q)/(N_{\mathbf{Y}} - 1)$.

Note that the discrete metric which induces MAP (and in turn TMAP) is symmetric in the sense that $d(x_1, y_1) = d(x_2, y_2)$

$\forall x_i \neq y_i$. Therefore, for a completely symmetric HMM, all criteria upon which the determination of the value functions at a particular σ depend are symmetrically defined and thus the value functions are unchanged under the permutation of components of σ.

In two dimensions, a completely symmetric HMM is given by

$$\mathcal{A} = \begin{bmatrix} 1-\epsilon & \epsilon \\ \epsilon & 1-\epsilon \end{bmatrix} \qquad Q = \begin{bmatrix} q & 1-q \\ 1-q & q \end{bmatrix}.$$

In the previous example, we have set $\epsilon = 0$ and $q = 1/2$.

Theorem 5: For a completely symmetric HMM, the value functions $\overline{J}^\gamma(\sigma, T-j)$ restricted to the simplex (the set $\{\sigma | \sigma_1 + \sigma_2 \cdots + \sigma_{N_\mathbf{X}} = 1\}$) have their global maximum at the center of the simplex, i.e., at $\sigma_1 = \sigma_2 = \cdots = \sigma_{N_\mathbf{X}}$.

Proof: The restriction of a concave function to a convex region is certainly concave. The simplex is a compact set, and so the value function restricted to the simplex is concave and has a maximum point. Let us begin with the 1-simplex. By the complete symmetry of the system, a value function will be symmetric with respect to the center of the simplex, namely around the point $\sigma_1 = \sigma_2 = 1/2$. To see this let $\sigma = (\sigma_1, \sigma_2)$, $\sigma_1 + \sigma_2 = 1$ be an arbitrary point on the simplex. Consider this point and its mirror image $\sigma' = (\sigma_2, \sigma_1)$ obtained from σ by a permutation of its components ($\sigma_1 \to \sigma_2, \sigma_2 \to \sigma_1$) which leaves the value function unchanged. Because the value functions are concave, we can write

$$\begin{aligned}\overline{J}^\gamma(1/2, T-j) &= J((\sigma + \sigma')/2, T-j) \\ &\geq 1/2 \left\{ \overline{J}^\gamma(\sigma, T-j) + \overline{J}^\gamma(\sigma', T-j) \right\} \\ &= \overline{J}^\gamma(\sigma, T-j).\end{aligned}$$

Therefore, the center of the simplex is the global maximum.

In general, for an $(N_\mathbf{X} - 1)$-simplex, consider $N_\mathbf{X}$ permutations of σ

$$\sigma_1 \to \sigma_2, \sigma_2 \to \sigma_3 \ldots, \sigma_i \to \sigma_{i+1}, \quad \sigma_{N_\mathbf{X}} \to \sigma_1$$

repeated $N_\mathbf{X}$ times and generating $N_\mathbf{X}$ points σ^k $k = 1, \ldots, N_\mathbf{X}$. Then, by symmetry and by Jensen's inequality

$$\begin{aligned}&\overline{J}^\gamma \left(\left[\sigma^1 + \cdots + \sigma^{N_\mathbf{X}} \right] / N_\mathbf{X}, T-j \right) \\ &\geq 1/N_\mathbf{X} \cdot \overline{J}^\gamma(\sigma^1, T-j) + \cdots + 1/N_\mathbf{X} \cdot \overline{J}^\gamma \left(\sigma^{N_\mathbf{X}}, T-j \right) \\ &= \overline{J}^\gamma(\sigma, T-j).\end{aligned}$$

Let $\sigma^* = [\sigma^1 + \cdots + \sigma^{N_\mathbf{X}}]/N_\mathbf{X}$. Then, $\sigma_i^* = 1/N_\mathbf{X} \sum_{l=1}^{l=N_\mathbf{X}} \sigma_l = 1/N_\mathbf{X}$ since σ is a point of the simplex. Therefore, σ^* is the center of the simplex and by the above, the maximal point. ■

In fact, a stronger version of the theorem holds. We will only outline the proof. Consider the 2-simplex which is bounded by three 1-simplexes. Complete symmetry of the system implies that along these simplices and all the line segments parallel to them, the value function has its maximum at the middle of such lines. Thus, the maximum on the whole simplex can be found by considering only the values on the set of the midpoints of all these line segments joined at the center of the simplex (see Fig. 3).

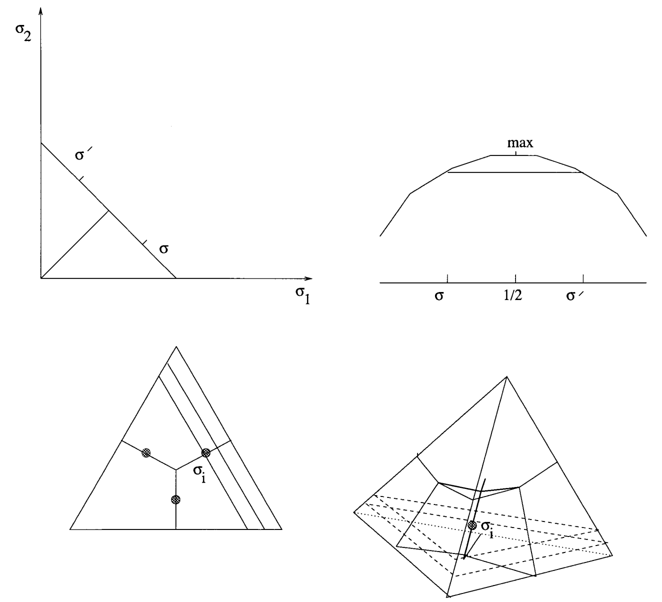

Fig. 3. The global maximum under complete symmetry.

Similarly, a $(N_\mathbf{X} - 1)$-simplex is bounded by $N_\mathbf{X}$, $(N_\mathbf{X} - 2)$-simplexes. Along these subsimplices and hyper-planes parallel to them and restricted to the simplex, the value functions have their maximum at the central points, reaching their global maximum at the center of the $(N_\mathbf{X} - 1)$-simplex. Note that the maximum point need not be unique. In fact, in the above example for even values of j the maximum point is not unique.

If the assumption of complete symmetry is removed, Theorem 5 no longer holds. We will show this shortly.

IV. Qualitative Analysis of the Sample Paths and the Information State

A. Sample Path Error Smoothing

It is often said that risk sensitive estimators take into account "the higher order moments" of the cost. How do higher order moments manifest themselves in the behavior of the estimator? To understand what risk sensitivity takes into account, we first explain what MAP does not.

Given a single observation and a single random variable, the MAP estimation is a reasonable thing to do: minimize the measure of the set where the estimate and the random variable disagree. But a problem arises when we consider a stochastic process (in our case a time series). The obvious thing to do then is to minimize the expected summation of the error functions for the whole time series. As shown before, this reduces back to finding the MAP estimate at each instant of time. Thus, at each instant of time our decision is not affected by our past or future decisions. This makes the MAP estimation insensitive to the accumulation of errors along sample paths. To see this evaluate X_k and the estimate $\hat{X}_k$ at some fixed value of $\omega \in \Omega$ to produce a sample path or a realization of the respective time series and its estimate. The sequence of decisions $\hat{X}_k$ $k = 0, 1, \ldots N$ partitions the sample space

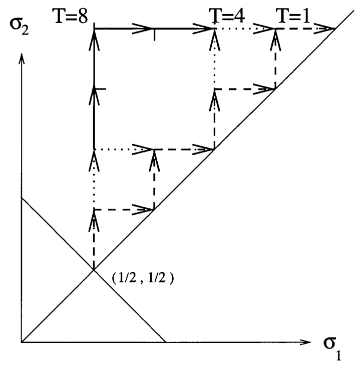

Fig. 4. Sample paths for $a_{11} = a_{22} = 1$, $q_{xy} = 1/2$ from (1/2, 1/2).

into 2^N subsets according to a binary tree with the branching criterion $X_k = \hat{X}_k$ or $X_k \neq \hat{X}_k$. Each ω belongs to exactly to one of these subsets. Some sample paths may end up on branches along which estimation errors accumulate for long stretches of time. Now, consider TMAP with $T = 1$ (which is the Dey–Moore filter equivalent for MAP). The exponential function turns the sum of the error functions into a product, the value of this product up to the last decision made for each fixed ω is $e^{\gamma \cdot m}$, where m counts the number of times an error has been made in estimation of that particular sample path. This product is then multiplied by either 1 or $e^{\gamma} > 1$ depending on the next decision made based on minimizing the measure of the error set at the present time *and* taking note of the errors made in the past. Thus the errors made in the past become more uniformly distributed over all sample paths. The information state condenses this error history and makes the computation recursive. For this reason perhaps *sample path error smoothing estimators* could be an alternative name for the risk-sensitive estimators. Dey–Moore filter is (in a sense) a greedy algorithm which does not consider (on the average) the accumulation of errors in the future but has the important benefit of computational simplicity. The exponential/sum filter does so for the entire path and in general TMAP looks into the future for T steps. In Fig. 4, observe that a particular sample path of our familiar example becomes more oscillatory as T is made smaller. Our simulations show that this is a general behavior no matter how complicated the chain: the smaller the filter size T, the bigger is the burden of error smoothing on the T next estimates. However, making T large comes at the cost of increased computational complexity.

The underlying mechanism of all these estimators is the coupling of the estimation errors in the product form. In an upcoming paper, we will show that for HMMs this coupling is possible in a broader context.

B. Risk-Sensitivity, Information, and Mixing

Through the sample path perspective, we can explain the behavior of the risk-sensitive estimators. In HMMs, all sample paths pass through a finite number of states. We can think of the transition probabilities as determining a flow in the system. Thus far, the only example we have considered was a nonmixing dynamical system. The transition probabilities were set to be zero and there was no flow between the states. This is important from the error smoothing point of view, for as the flow passes through the states, so does the history of the errors. If sample paths that have accumulated estimation error remain in a particular state, the estimator will be "attracted" to the state in order to relieve the error accumulated there. This explains the oscillatory behavior of our two state example in which no flow between the states was allowed. On the other hand, if the transition probabilities are nonzero this "attraction" is somewhat relieved; we have verified through simulations that mixing indeed inhibits the oscillatory behavior. This will also have an effect on the decision regions as we will see shortly. However, if we go through a state "too quickly," we cannot use that state to smoothen the error accumulated in the path effectively. Both these cases lead to certain type of singularities in the decision regions.

The second issue is the role of information. If we expect to receive good information about the system in the future, which will in turn reduce the error accumulation, we are likely to be less conservative at the present about our decisions. This means that we expect TMAPs decision regions to become less conservative and look more like MAPs under increased availability of information. This too will be shown in the following example.

We will study the decision regions for $T = 2$ TMAP for an HMM with $N_{\mathbf{X}} = N_{\mathbf{Y}} = 2$ and

$$\mathcal{A} = \begin{bmatrix} 1/2 & 1/2 \\ \delta & 1-\delta \end{bmatrix}; \qquad Q = \begin{bmatrix} 1/2+I & 1/2-I \\ 1/2-I & 1/2+I \end{bmatrix}. \quad (22)$$

The parameter I controls the availability of information. When $I = 0$, no information is available (the case of pure prediction) and as $I \to 1/2$, the HMM becomes perfectly observable. The parameter δ determines the transition probabilities of the second state and, in particular, $\delta = 0$ will make the Markov chain nonmixing.

As shown before for $T = 1$, the decision regions are identical to those of MAP. First, let $I = 0$, then it can be shown that for $T = 2$, the decision regions of $j = 2$ (the first stage of the two step filter) are of the threshold type determined by a transition line L with the slope $m(\delta)$ followed by the equipartition decision regions (identical to the decisions regions of MAP). The decision regions of the first stage are given by $\sigma_2 < m(\delta)\sigma_1$ choose 1 and if $\sigma_2 > m(\delta)\sigma_1$ choose 2. The slope $m(\delta)$ is given by

$$m(\delta) = \begin{cases} \dfrac{e^{\gamma}+1}{2} \cdot \dfrac{1}{\delta e^{\gamma} + 1 - \delta} & \delta < 1/2; \\ \dfrac{e^{\gamma}+1}{2} \cdot \dfrac{1}{(1-\delta) e^{\gamma} + \delta} & \delta > 1/2. \end{cases}$$

Simple calculations show that the slope is always greater than or equal to 1 (only when $\delta = 1/2$), so that the decision region of the first state is enlarged at the expense of the second. As expected, when $\gamma \to 0$, the decision regions equalize. When $\gamma \to \infty$, the slope is given by

$$m(\delta) = \begin{cases} \dfrac{1}{2\delta} & \delta < 1/2; \\ \dfrac{1}{2(1-\delta)} & \delta > 1/2. \end{cases}$$

When either $\delta = 0$ or $\delta = 1$, the slope becomes infinite. These are the singularities that we mentioned earlier. The equalization of the two regions at $\delta = 1/2$ is a general property which holds true even when no constraint is put on the available information as the following theorem demonstrates.

Theorem 6: Consider the HMM described by (1). The risk-sensitive decision regions are equalized if the rows of A are identical. Furthermore, TMAP reduces to MAP for every choice of T and γ.

Proof: Fix T and γ. Consider the calculations of the decision regions according to (13). For $j = 1$, we already saw that the decision regions are equalized. By assumption the rows of $\mathcal{A}$ are all the same and are give by a row vector whose transpose we denote by $\overline{\mathbf{r}}$. [Note that under this assumption, $\mathcal{A}^T X = (\sum_i X_i)\overline{\mathbf{r}}$, for every vector X]. For $j > 1$, we can write

$$\begin{aligned}
&\overline{J}^{\gamma}(\sigma, T-j) \\
&= \min_{v\in\mathbf{X}} \left\{ E\left[\overline{J}^{\gamma}(L(v, Y_{T-j+1})\sigma, T-j+1)\right]\right\} \\
&= \min_{v\in\mathbf{X}} \left\{ \frac{1}{N_{\mathbf{Y}}} \sum_{y=1}^{N_{\mathbf{Y}}} \left[\overline{J}^{\gamma}\left(N_{\mathbf{Y}}\overline{Q}(y)\mathcal{A}^T\mathcal{E}\mathcal{X}(v)\sigma, T-j+1\right)\right]\right\} \\
&= \min_{v\in\mathbf{X}} \left\{ \frac{1}{N_{\mathbf{Y}}} \sum_{y=1}^{N_{\mathbf{Y}}} \left[\overline{J}^{\gamma}\left(N_{\mathbf{Y}}\overline{Q}(y)\{\langle\mathcal{E}\mathcal{X}(v)\sigma, \overline{\mathbf{1}}\rangle\}\overline{\mathbf{r}},\right.\right.\right. \\
&\qquad\qquad \left.\left.\left. T-j+1\right)\right]\right\} \\
&= \min_{v\in\mathbf{X}} \left\{ \langle\mathcal{E}\mathcal{X}(v)\sigma, \overline{\mathbf{1}}\rangle \sum_{y=1}^{N_{\mathbf{Y}}} \left[\overline{J}^{\gamma}\left(\overline{Q}(y)\overline{\mathbf{r}}, T-j+1\right)\right]\right\} \\
&= \min_{v\in\mathbf{X}} \left\{\langle\mathcal{E}\mathcal{X}(v)\sigma, \overline{\mathbf{1}}\rangle\right\} \\
&\quad \cdot \left\{ \sum_{y=1}^{N_{\mathbf{Y}}} \left[\overline{J}^{\gamma}\left(\overline{Q}(y)\overline{\mathbf{r}}, T-j+1\right)\right]\right\}.
\end{aligned}$$

The second term does not depend on v, and so the minimization reduces to

$$\min_{v\in\mathbf{X}} \left\{\langle\mathcal{E}\mathcal{X}(v)\sigma, \overline{\mathbf{1}}\rangle\right\}.$$

This shows that the decision regions are equalized. Similar techniques applied to the evolution of the information state show that TMAPs and MAPs information states, under uniform flow, are multiples of each other. This fact, together with Lemma 5 and the previous result regarding the decision regions completes the proof. ■

In the previous result, the observation matrix $\overline{Q}(y)$ plays no role under the assumption of uniform flow in the determination of the decision regions. However, this is the exception. In general, the availability of information appears to have an elegant relation to the structure of the decision regions as the following shows.

Theorem 7: In (22), let $\delta = 0$ and $I \geq 1/(2(1+e^{-\gamma}))$; then the decision regions for TMAP, $T = 2$ are equalized.

Proof: The proof follows from the solution of a system of simultaneous inequalities defined by (13) with the constraints $\delta = 0$ and $I \geq 1/(2(1+e^{-\gamma}))$. We omit the tedious algebra. ■

As we mentioned earlier, this does not imply that TMAP for $T = 2$ reduces to MAP because the recursions governing the evolution of the information states for the two cases are different. But if for some T the decision regions are equalized then TMAP with filter size T does reduce to the TMAP with filter size $T = 1$. This would be significant both conceptually and computationally if we could determine conditions under which the decision regions are equalized. Note in the above for computational reasons, we had constrained the observation matrix to be symmetric. This produces only a sufficient condition as stated in Theorem 7. The right measure of the minimum quantity of information needed must be free of such constraints (for example, Shannon's mutual information among the states and the observations). Observe that in the previous example, the amount of needed information grows with increasing γ. Clearly $1/(2(1+e^{-\gamma})) \to 1/2$ as $\gamma \to \infty$ which implies under infinite risk, we need perfect observation to equalize the decision regions.

We proved earlier that the maximum expected error "cost to go" for a completely symmetric HMM occurs at the center of the simplex. This is not true in general. It can be shown that the maximum cost in the first case of the prior example ($I = 0$) occurs along the transition line L, with slope $m(\delta)$, which does not cross the center of the simplex $\sigma = (1/2, 1/2)$ unless $\delta = 1/2$, a special case of complete symmetry with $q = \epsilon = 1/2$. When $\delta = 0$ as $\gamma \to \infty$, this maximum point is moved arbitrarily closer to $\sigma = (0, 1)$. This is another example of what we termed a singularity under infinite risk.

C. Normalized Information State and Its Asymptotic Behavior

In this section, we give TMAP a "flow" interpretation. It is well known that the probability distribution of a Markov chain evolves according to

$$p_{n+1} = \mathcal{A}^T p_n \tag{23}$$

where p_0 is the initial probability distribution.

When no information is available other than the initial probability distribution, the probability distribution of the chain is the information state for the MAP estimator, the so called "prediction" case. We consider the TMAP estimator with $T = 1$, risk-sensitivity parameter γ and define the *risk-sensitive probability distribution* of a Markov chain U_t, as the normalized information state for the prediction case. Note that by Theorem 1, the prediction case is obtained by setting $\overline{Q}(y) = (1/N_{\mathbf{Y}})I$ (where I is the identity matrix) in (7).

Theorem 8: The risk-sensitive probability distribution evolves according to

$$U_{t+1} = F(U_t) := \mathcal{A}^T S(U_t) \tag{24}$$

where $\mathcal{A}$ is the transition probability matrix and given V on the $N_{\mathbf{X}}$-simplex, S is defined by

$$\begin{aligned}
&S(V) = W \\
&W_k = \begin{cases} V_k - \beta, & \text{if } k = \operatorname*{argmax}_i V_i \\ V_k + \beta_k, & \text{otherwise} \end{cases}
\end{aligned}$$

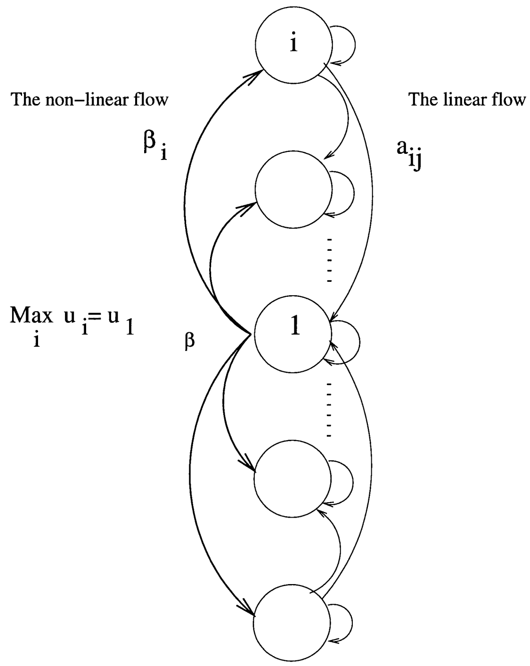

Fig. 5. The back-flow.

where

$$\beta = \frac{(e^\gamma - 1)V_m \sum_{i\neq 1} V_i}{e^\gamma \sum_{i\neq 1} V_i + V_m}, \qquad m = \underset{i}{\operatorname{argmax}}\, V_i$$

$$\beta_k = \beta \cdot \frac{V_k}{\sum_{i\neq m} V_i}.$$

Proof: The proof follows from Theorem 1, Lemma 5, Theorem 3, and straightforward algebra; for details, see [1]. ■

Theorem 8 shows that the evolution of the risk-sensitive probability distribution (rs-probability) is governed by two simplex preserving actions: one linear ($\mathcal{A}^T$) and the other nonlinear (S). (In Fig. 5, we assume $m = 1$.) The influence of this nonlinear action or "back-flow" is considered next. Clearly, $\beta = \sum_{i\neq m} \beta_i$; therefore, an amount β is subtracted from the state with the highest rs-probability and distributed among the other states according to the ratio $\beta_k = \beta \cdot (V_k / \sum_{i\neq m} V_i)$. In particular, if a state is "empty" ($V_k = 0$), it receives nothing ($\beta_k = 0$). In two dimensions, an amount β is simply moved from the state with the larger rs-probability (u_1) to the one with the smaller rs-probability (u_2)

$$\frac{(e^\gamma - 1)u_1 u_2}{e^\gamma u_2 + u_1} = \beta. \tag{25}$$

Note that $\gamma \to 0 \Rightarrow \beta \to 0$ and $\gamma \to \infty \Rightarrow \beta \to u_1$. In the context of risk-sensitive estimation, the operator $\mathcal{A}^T S$ plays exactly the same role which $\mathcal{A}^T$ plays in the risk-neutral case.

It is well known that if $\mathcal{A}$ is primitive [15], $p_{n+1} = \mathcal{A}^T p_n$ is *asymptotically stationary* and that the stationary distribution satisfies $p^* = \mathcal{A}^T p^*$. Our simulations show that under primitivity of $\mathcal{A}$, for small values of γ, (24) is also asymptotically stationary. For sufficiently large values of γ, however, (24) is *not* stationary but appears to exhibit *asymptotic periodicity*. First, we need a definition:

Definition 5: A cycle of rs-probabilities (CRP) is a finite set of probability vectors $\{v^1, \ldots, v^m\}$ such that $F(v^i) = v^{i+1}$ with $F(v^m) = v^1$.

Our simulations show that under primitivity (24) approaches a CRP.

Example 2: Let $\mathcal{A}$ be given by

$$\mathcal{A} = \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix} \qquad e^\gamma = 1.078.$$

Our simulations show that (24), for every initial condition, approaches

$$(0.655\,22,\ 0.344\,77)$$

we could say that the CRP has only one element. On the other hand, the fixed point of (23) is given by the linear equations

$$\mathcal{A}^T p = p$$
$$p_1 + p_2 = 1$$
$$p_1 = 0.666\,66 \qquad p_2 = 0.333\,33.$$

Solving the fixed point problem for (24), assuming $u_1 > u_2$, gives

$$u_1 = \frac{0.8u_1 + 0.4312u_2}{u_1 + 1.078u_2} \qquad u_1 + u_2 = 1$$
$$0.078{u_1}^2 - 0.7092u_1 + 0.4312 = 0$$
$$u_1 = 0.655\,22 \qquad u_2 = 0.344\,77.$$

So, the limiting behavior is determined by a fixed point problem where the nonlinearity S manifests itself in the quadratic form.

For large values of γ; however, the asymptotic behavior turns out to be periodic and not stationary. To understand why, consider

$$\mathcal{A} = \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix} \qquad e^\gamma \to \infty.$$

The CRP, for every initial condition, turns out to be

$$CRP\colon (v^1, v^2) \qquad F(v^1) = v^2 \qquad F(v^2) = v^1$$
$$v^1 = \begin{bmatrix} 0.8 \\ 0.2 \end{bmatrix} \qquad v^2 = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}.$$

Consider the "back-flow" β when $e^\gamma \to \infty$. Begin with some distribution (u_1, u_2) and suppose $u_1 > u_2$ then by (25), $\beta = u_1$ and, thus, $u_2 + \beta = 1$, and then the action $\mathcal{A}^T$ in (24) gives us (0.4, 0.6). However, we now have $u_2 > u_1$ and $u_1 + \beta = 1$ and so after the action of $\mathcal{A}^T$ we get (0.8, 0.2) and the cycle begins all over again.

The asymptotic behavior of (24) is the subject of our ongoing research; more examples and analysis can be found in [1].

V. Conclusion

We set out to study the risk-sensitive estimation of HMMs from a dynamical systems point of view. We unified the ideas of Speyer and Dey–Moore risk-sensitive filtering in the context of risk-sensitive filter banks and showed how and why the two may coincide. Structural results for the decision regions and an *error smoothing* perspective enabled us to understand the behavior of the risk-sensitive sample paths.

A more conservative or risk-sensitive estimator tends to move around the state space more rapidly from sample path to sample path and not for too long "trust" the correctness of the sample path it may be following. In this sense, the estimates are made more "uniform." This is evident, for example, in the asymptotic behavior of the rs-probability which appears to be periodic rather than stationary. It is likely that interesting problems remain.

An obvious set of problems are those of "classification:" the characterization of the decision regions in terms of the coupling of information, risk-sensitivity and the transition probabilities is incomplete. Solving the classification problems may in turn deepen our understanding as to how risk-sensitivity affects the behavior of the estimated sample paths.

Finally, we point out that many of the results of this paper depend on the properties of the discrete metric (used to define MAP) which is not the natural metric for R^n. Therefore, our structural results do not directly illuminate the linear-Gaussian risk-sensitive estimation case. However, the intuition gained in the discrete finite dimensional setting about the behavior of the sample paths may lead to a better understanding of the linear-Gaussian risk-sensitive estimator as well.

ACKNOWLEDGMENT

The first author would like to acknowledge the suggestions of J. P. Chawla.

REFERENCES

[1] V. R. Ramezani. (2001). Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. Maryland, College Park, Inst. Syst. Res. Tech. Rep. Ph.D. 2001-7. [Online]. Available: http://www.isr.umd.edu/TechReports/ISR/2001/PhD_2001-7.

[2] D. H. Jacobson, "Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic games," *IEEE Trans. Automat. Contr.*, vol. AC-18, no. 2, pp. 890–902, Apr. 1973.

[3] R. J. Elliot, L. Aggoun, and J. B. Moore, *Hidden Markov Models: Estimation and Control*. New York: Springer-Verlag, 1994.

[4] M. R. James, J. S. Baras, and R. J. Elliot, "Risk sensitive control and dynamic games for partially observed discrete-time nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 780–792, Apr. 1994.

[5] J. S. Baras and M. R. James, "Robust and risk-sensitive output feedback control for finite state machines and hidden Markov models," Inst. Syst. Res., Univ. Maryland, College Park, Tech. Res. Rep. T. R. 94-63.

[6] J. L. Spyer, C. Fan, and R. N. Banava, "Optimal stochastic estimation with exponential cost criteria," in *Proc. 31st Conf. Decision Control*, 1992, pp. 2293–2298.

[7] S. Dey and J. Moore, "Risk sensitive filtering and smoothing for hidden Markov models," *Syst. Control Lett.*, vol. 25, pp. 361–366, 1995.

[8] ——, "Risk sensitive filtering and smoothing via reference probability methods," *IEEE Trans. Automat. Contr.*, vol. 42, pp. 1587–1591, Nov. 1997.

[9] R. K. Boel, M. R. James, and I. R. Petersen, "Robustness and risk-sensitive filtering," in *Proc. 36th IEEE Conf. Decision Control*, 1997, Cat. no. 97CH36124.

[10] H. Vincent Poor, *Hidden Markov Models: Estimation and Control*. New York: Springer-Verlag, 1994.

[11] P. Whittle, *Risk-Sensitive Optimal Control*. New York: Wiley, 1990.

[12] E. Fernandez-Gaucherand and S. I. Marcus, "Risk-sensitive optimal control of hidden Markov models: A case study," *IEEE Trans. Automat. Contr.*, vol. 42, pp. 1418–1422, Oct. 1997.

[13] P. Whittle, "Risk-sensitive linear/quadratic/Gaussian control," *Adv. Appl. Probab. Trust*, vol. 13, pp. 746–777, 1981.

[14] W. H. Fleming and D. Hernandez-Hernandez, "Risk-sensitive control of finite state machines on an infinite horizon II," *SIAM J. Control Optim.*, vol. 37, no. 4, pp. 1048–1069.

[15] A. N. Shiryayev, *Probability*. New York: Springer-Verlag, 1984.

Vahid Reza Ramezani received the B.S., M.S., and Ph.D. degrees in electrical engineering. all from the University of Maryland, College Park, in 1981, 1994, and 2001, respectively.

During 1995–1997, he was a Design Engineer for LCC, Inc., McLean, VA, primarily concerned with the design, implementation, testing, and optimization of cellular telephony systems. His research interests include optical communication, risk-sensitive estimation and control, wireless communications, and intrusion detection. He is currently with the Institute for Systems Research, University of Maryland, College Park.

Steven I. Marcus (S'70–M'75–SM'83–F'86) received the B.A. degree in electrical engineering and mathematics from Rice University, Houston, TX, and the S.M. and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1971, 1972, and 1975, respectively.

From 1975 to 1991, he was with the Department of Electrical and Computer Engineering at the University of Texas at Austin, where he was the L.B. (Preach) Meaders Professor in Engineering. He was Associate Chairman of the Department from 1984 to 1989. In 1991, he joined the University of Maryland, College Park, where he was Director of the Institute for Systems Research until 1996. He is currently a Professor in the Electrical and Computer Engineering Department and the Institute for Systems Research, and Chair of the Electrical and Computer Engineering Department at the same university. Currently, his research is focused on stochastic control and estimation, with applications in semiconductor manufacturing, telecommunication networks, and preventive maintenance. He is Editor-in-Chief of the *SIAM Journal on Control and Optimization*.