

Maximum Entropy Models, Dynamic Games, and Robust Output Feedback Control for Automata

John S. Baras and Maben Rabi

Abstract—In this paper, we develop a framework for designing controllers for automata which are robust with respect to uncertainties. A deterministic model for uncertainties is introduced, leading to a dynamic game formulation of the robust control problem. This problem is solved using an appropriate information state. We derive a Hidden Markov Model as the maximum entropy stochastic model for the automaton. A risk-sensitive stochastic control problem is formulated and solved for this Hidden Markov Model. The two problems are related using small noise limits.

I. INTRODUCTION

An *automaton* or *finite state machine* (FSM) [8] is a discrete-time system defined by the model

$$\left. \begin{aligned} x_{k+1} &= f(x_k, u_k), \\ y_{k+1} &= g(x_k), \quad k = 0, 1, \dots, M, \end{aligned} \right\} \quad (1)$$

where the “state” x_k evolves dynamically in a finite set \mathbf{X} , and the control u_k and output y_k take values in finite sets \mathbf{U} and \mathbf{Y} , respectively. These sets have n , m , and p elements, respectively. The behavior of the FSM is described by a *state transition map* $f: \mathbf{X} \times \mathbf{U} \rightarrow \mathbf{X}$ and an *output map* $g: \mathbf{X} \rightarrow \mathbf{Y}$.

FSM models, and their extensions EFSM (extended FSM), together with accompanying optimization, optimal control and decision problems, have been used widely in applications. However, it is typically the case that conventional treatments of such application problems do not specifically deal with *model uncertainties* or *disturbances*; e.g., as arising from modelling errors, sensor noise, parametric variations, etc.. Yet in almost all applications *performance and operational robustness* is of paramount importance. By that we mean the design of system structures (in this case FSM structures) and input strategies that can sustain desired performance and operation despite model uncertainties and/or signal disturbances. Such problems are widely known as *robust control-communication-signal processing* problems. In this paper we refer collectively to such problems as robust decision and control problems, or simply *robust control* problems. We propose and solve a general robust control problem for FSMs, and in the process we develop a new and deeper understanding of the fundamental principles that support the framework that has been developed for linear

systems (e.g. [4], [10]), as well as nonlinear systems [2], [24], [25], [26], [27].

The starting point of our approach is motivated by the method developed in [3], [11], [6], [2], [10], [1], [4], [7], [24], [25], [26], [27]. We then develop a general framework for robust output feedback control of FSMs, by carefully studying two basic methodologies for representing model uncertainties and signal disturbances: a deterministic one and a stochastic one. We investigate the relationship between the two resulting design methodologies for robust output feedback control, and establish a certain “duality” between them. Key to this linkage is the formulation of the robust control problem as a dynamic game between two players: *nature* (who selects model uncertainties and signal disturbances) and the *control designer* (who selects control strategies). When we use deterministic models for the system uncertainties the corresponding game is a deterministic game, while when we use stochastic models for the system uncertainties the corresponding game is a stochastic one. The relationship between the two design methods is a consequence of the relationship between the deterministic and the stochastic games.

When we model system uncertainties stochastically, the FSM model of (1) is transformed to a Hidden Markov Model (HMM). Over the last fifteen years the output robust control problem has been investigated for various classes of systems, within the context of the so-called H^∞ control. This substantial body of research results established the equivalence of the output robust control problem (or “four-block problem”), with the problem of solving a non-cooperative two player deterministic dynamic game, as well as with the problem of solving a single player *risk-sensitive* and partially observed stochastic optimal control problem. The equivalence of these three problems has been established for a great variety of system and performance metric models, but it has principally been interpreted and understood as a means for obtaining the solution of any of these problems in terms of the solutions of the other. A key conclusion from these earlier research results is that risk-sensitive controllers are very closely related to robust controllers, see [4], [7], [24], [25], [26], [27].

Yet, despite these developments, from a deeper systems perspective, a key question that remained unanswered was the following. It is clear that the risk-sensitive stochastic control problem involved in these equivalences, represents a particular *randomization* of the robust output feedback control problem. As is true in many other problems this randomization reduces the computational complexity of

This research was supported by the Army Research Office under the ODDR&E MURI01 Program Grant No. DAAD19-01-1-0465 to the Center for Networked Communicating Control Systems (through Boston University), and under ARO Grant No. DAAD190210319.

J. S. Baras and M. Rabi are with the Institute for Systems Research and the Department of Electrical & Computer Engineering, University of Maryland, College Park, MD 20742, USA.

computing robust output feedback controls. In the present paper we investigate exactly this deeper question. Namely, what is the deeper fundamental principle leading to the particular randomization used in the risk-sensitive stochastic control formulation of the robust control problem? The answer, established here, is that this randomization is in fact equivalent to the construction of a *maximum entropy model* [19], [20], [21], [22], which is a carefully constructed HMM. In this paper we establish this result for automata and HMMs. Establishing the result for more general systems including nonlinear systems and hybrid systems, will be done elsewhere.

We establish the result by first reviewing our earlier work [26]. We formulate the robust output feedback control problem for a FSM and summarize the description of its equivalence to a deterministic partially observed dynamical game. We then formulate the robust output feedback control problem for the maximum entropy randomization of the (1), following ideas from [16], [28]. We then solve the *risk-sensitive* stochastic optimal control problem for the resulting HMM. Our solution, which is interesting in itself, leads us to the solution of the robust output feedback control problem for FSMs. Finally, we link the two problems in yet another way by employing large deviation limits as in [7].

The robust output feedback control problem for FSMs is formulated in Section II; this entails defining a deterministic uncertainty (disturbance) mechanism with associated cost functions. In Section III, a stochastic uncertainty (disturbance) model is derived via the principle of maximum entropy modelling [19], [20], [21], [22]. In the same Section we also derive the duality of this randomization with a risk-sensitive control problem. The risk-sensitive problem is solved, and the large deviation principle is invoked (i.e. a small noise limit is evaluated) and used in Section IV to solve the robust output feedback control problem of Section II.

II. OUTPUT ROBUST CONTROL PROBLEM WITH DETERMINISTIC UNCERTAINTY MODELS

A. Deterministic Perturbation

We model the influence of disturbances by a FSM model obtained by augmenting (1) with two additional (disturbance) inputs w and v :

$$\left. \begin{aligned} x_{k+1} &= b(x_k, u_k, w_k), \\ y_{k+1} &= h(x_k, v_k), \quad k = 0, 1, \dots, M, \end{aligned} \right\} \quad (2)$$

where, w_k and v_k take values in finite sets \mathbf{W} and \mathbf{V} respectively. The range spaces of the maps b and h are \mathbf{X} and \mathbf{Y} respectively. The functions $b : \mathbf{X} \times \mathbf{U} \times \mathbf{W} \rightarrow \mathbf{X}$ and $h : \mathbf{X} \times \mathbf{V} \rightarrow \mathbf{Y}$ are required to satisfy the following consistency conditions:

$$\left\{ \begin{array}{l} \text{there exists } w_\emptyset \in \mathbf{W} \text{ such that} \\ b(x, u, w_\emptyset) = f(x, u) \text{ for all } x \in \mathbf{X}, u \in \mathbf{U}, \end{array} \right.$$

$$\left\{ \begin{array}{l} \text{there exists } v_\emptyset \in \mathbf{V} \text{ such that} \\ h(x, v_\emptyset) = g(x) \text{ for all } x \in \mathbf{X}. \end{array} \right.$$

The symbols w_\emptyset and v_\emptyset [8] play the role of “zero inputs”, so that when no disturbances are present (i.e. $w_k \equiv w_\emptyset$, and $v_k \equiv v_\emptyset$), the behavior of (2) is the same as (1).

We will assume that there exists a null control $u_\emptyset \in \mathbf{U}$ and an equilibrium or rest state $x_\emptyset \in \mathbf{X}$ such that

$$x_\emptyset = f(x_\emptyset, u_\emptyset).$$

The set of possible initial states is denoted $N_0 \subset \mathbf{X}$, and assumed to contain x_\emptyset , while the set of possible future states for the disturbance model (2) is

$$N_{\mathbf{X}}(x, u) = \{b(x, u, w) : w \in \mathbf{W}\} \subset \mathbf{X},$$

and the corresponding set of possible future outputs is

$$N_{\mathbf{Y}}(x) = \{h(x, v) : v \in \mathbf{V}\} \subset \mathbf{Y}.$$

These sets can be thought of as “neighborhoods” of the nominal future values $f(x, u)$, $g(x)$, and are determined by the maps b and h . These can be designed as appropriate for the application at hand.

B. Cost Functions

To quantify the effect of the disturbances, a measure of their “sizes” is required. To this end, one specifies functions

$$\phi_w : \mathbf{W} \times \mathbf{X} \times \mathbf{U} \rightarrow \mathbf{R}, \quad \phi_v : \mathbf{V} \times \mathbf{X} \rightarrow \mathbf{R}, \quad \beta : \mathbf{X} \rightarrow \mathbf{R},$$

with the following properties:

$$\left\{ \begin{array}{l} \phi_w(w_\emptyset; x, u) = 0 \quad \text{for all } x \in \mathbf{X}, u \in \mathbf{U}, \\ \phi_w(w; x, u) \geq 0 \quad \text{for all } w \neq w_\emptyset \in \mathbf{W}, x \in \mathbf{X}, u \in \mathbf{U}, \\ \phi_v(v_\emptyset; x) = 0 \quad \text{for all } x \in \mathbf{X}, \\ \phi_v(v; x) \geq 0 \quad \text{for all } v \neq v_\emptyset \in \mathbf{V}, x \in \mathbf{X}, u \in \mathbf{U}, \end{array} \right.$$

and,

$$\left\{ \begin{array}{l} \beta(x_\emptyset) = 0, \\ +\infty > \beta(x_0) \geq 0 \quad \text{for all } x_0 \neq x_\emptyset \in N_0, \\ \beta(x_0) = +\infty \quad \text{for all } x_0 \notin N_0, \end{array} \right.$$

We think of $\phi_w(w; x, u)$ as the magnitude of the disturbance w as it affects the system when it is in state x with control u applied, and $\phi_v(v; x)$ as the magnitude of the disturbance v when in state x . The cost function β specifies the amount of uncertainty regarding the initial state.

Associated with these cost functions are quantities which define the optimal cost of transferring from x to x'' and the optimal cost of producing the output y'' . These quantities will be used in the solution of the robust control problem below. They are defined by

$$\begin{aligned} U(x, x'') &\triangleq \min_{w \in \mathbf{W}} \{\phi_w(w; x, u) : x'' = b(x, u, w)\}, \\ V(x, y'') &\triangleq \min_{v \in \mathbf{V}} \{\phi_v(v; x) : y'' = h(x, v)\}. \end{aligned} \quad (3)$$

We adopt the convention that the minimum over an empty set equals $+\infty$. Thus U and V are extended real valued functions. Note that

$$U(x, f(x, u); u) = 0 \quad \text{for all } x \in \mathbf{X}, u \in \mathbf{U}, \quad (4)$$

$$U(x, b(x, u, w); u) \geq 0 \quad \text{for all } w \neq w_\emptyset \in \mathbf{W}, x, u, \quad (5)$$

$$U(x, x''; u) = +\infty \quad \text{if } x'' \notin N_{\mathbf{X}}(x, u). \quad (6)$$

and

$$V(x, g(x)) = 0 \quad \text{for all } x \in \mathbf{X}, \quad (7)$$

$$V(x, h(x, v)) \geq 0 \quad \text{for all } v \neq v_\emptyset \in \mathbf{V}, x \in \mathbf{X}, \quad (8)$$

$$V(x, y'') = +\infty \quad \text{if } y'' \notin N_{\mathbf{Y}}(x). \quad (9)$$

C. Robust Control

As part of the problem specification, one defines an additional output quantity

$$z_{k+1} = \ell(x_k, u_k), \quad (10)$$

where z_k takes values in a finite set \mathbf{Z} , and $\ell : \mathbf{X} \times \mathbf{U} \rightarrow \mathbf{Z}$. We assume there exists a specific null element $z_\emptyset \in \mathbf{Z}$ such that

$$\ell(x_\emptyset, u_\emptyset) = z_\emptyset.$$

A cost function for this output is also specified, with the properties

$$\begin{cases} \phi_z(z_\emptyset) = 0, \\ \phi_z(z) \geq 0 \quad \text{for all } z \in \mathbf{Z}. \end{cases}$$

The output quantity z and its associated cost function ϕ_z encode the performance objective of the problem at hand. To summarize, the complete system is described by the equations

$$\left. \begin{aligned} x_{k+1} &= b(x_k, u_k, w_k), \\ z_{k+1} &= \ell(x_k, u_k), \\ y_{k+1} &= h(x_k, v_k), \quad k = 0, 1, \dots, M. \end{aligned} \right\} \quad (11)$$

The state variable x_k is not measured directly, and so the controller must make use of information available in the output signal $y_{0,k}$; i.e., the controller must be an *output feedback* controller. We denote by $\mathcal{O}_{k,l}$ the set of non-anticipating control policies defined on the interval $[k, l]$; i.e., those controls for which there exist functions $\bar{u}_j : \mathbf{Y}^{j-k+1} \rightarrow \mathbf{U}$ such that $u_j = \bar{u}_j(y_{k+1,j})$ for each $j \in [k, l]$.

The *output feedback robust control problem* is the following [26]: given $\gamma > 0$ and a finite time interval $[0, M]$ find an output feedback controller $u \in \mathcal{O}_{0,M-1}$ such that

$$\sum_{k=0}^{M-1} \phi_z(z_{k+1}) \leq \beta(x_0) + \gamma \sum_{k=0}^{M-1} (\phi_w(w_k; x_k, u_k) + \phi_v(v_k; x_k)) \quad (12)$$

for all $(w, v) \in \mathbf{W}^M \times \mathbf{V}^M, x_0 \in \mathbf{X}$.

D. Dynamic Game

The robust control problem formulated above can be recast as a dynamic game problem, see, e.g., [24], [25], [26]. The payoff function for a controller $u \in \mathcal{O}_{0,M-1}$ (player 1) and disturbances $(w, v, x_0) \in \mathbf{W}^M \times \mathbf{V}^M \times \mathbf{X}$ (player 2) is given by

$$J^\gamma(u, w, v, x_0) \triangleq -\beta(x_0) + \sum_{k=0}^{M-1} \phi_z(z_{k+1}) - \gamma (\phi_w(w_k; x_k, u_k) + \phi_v(v_k; x_k)). \quad (13)$$

We consider the upper payoff for this game given the dynamics (2). Define

$$J^\gamma(u) \triangleq \max_{(w,v) \in \mathbf{W}^M \times \mathbf{V}^M} \max_{x_0 \in \mathbf{X}} \{J^\gamma(u, w, v, x_0)\}.$$

The bound

$$0 \leq J^\gamma(u) \leq M \max_{z \in \mathbf{Z}} \phi_z(z) \quad (14)$$

is readily verified. The dynamic game problem is to find an output feedback controller $u^* \in \mathcal{O}_{0,M-1}$ such that

$$J^\gamma(u^*) = \min_{u \in \mathcal{O}_{0,M-1}} J^\gamma(u). \quad (15)$$

Then if

$$J^\gamma(u^*) = 0, \quad (16)$$

the robust control objective (12) is achieved.

We will solve this dynamic game problem in Section IV.

III. OUTPUT ROBUST CONTROL PROBLEM WITH STOCHASTIC UNCERTAINTY MODELS

A. Random Perturbation

The random perturbation defined below is a stochastic analog of the deterministic perturbation introduced in Section II. Let $S = \{(x_0, y_0), (x_1, y_1), \dots, (x_M, y_M)\}$, and let,

$$P_1^M(s) = \text{Prob} \left[\{(x_0, y_0), (x_1, y_1), \dots, (x_M, y_M)\} = s \right].$$

We choose the statistics of the stochastic system in the least biased way [19], [20] by picking the joint process probabilities $\{P_1^M(s)\}$ that generate the *maximum entropy* over the set S while at the same time describing the observed levels of disturbances.

We want to maximize

$$-\mathbf{E} \log (P_1^M(s))$$

over all distributions on S that generate specified levels of disturbances which are measured through the average costs of the disturbances. So the distribution we pick must satisfy:

$$\mathbf{E}U(x_i, x_{i+1}; u) = \alpha_i \quad \text{for } 1 \leq i \leq M-1,$$

$$\mathbf{E}V(x_i, y_i) = \beta_i \quad \text{for } 1 \leq i \leq M,$$

$$\mathbf{E}\beta(x_0) = \gamma.$$

In addition, the probabilities $P_1^M(s)$ must add up to one. We thus have to maximize a strictly concave function on the

simplex subject to linear constraints. The function achieves the global maximum at the only critical point it has. The Lagrangian is:

$$J = \sum_{i=1}^{M-1} \lambda_i [\alpha_i - \mathbf{E}U(x_i, x_{i+1}; u)] + \sum_{i=1}^M \mu_i [\beta_i - \mathbf{E}V(x_i, y_i)] + \nu [\gamma - \mathbf{E}\beta(x_0)] + \kappa \left[1 - \sum_{s \in S} P_1^M(s) \right] - \mathbf{E} \log (P_1^M(s)).$$

The critical point is found by solving the following equations:

$$- (1 + \log (P_1^M(s))) - \sum_{i=1}^{M-1} \lambda_i U(x_i, x_{i+1}; u) - \sum_{i=1}^{M-1} \mu_i V(x_i, y_i) - \nu \beta(x_0) - \kappa = 0,$$

for $s \in S$, and,

$$\begin{aligned} \alpha_i - \sum_{s \in S} P_1^M(s) U(x_i, x_{i+1}; u) &= 0, \quad \text{for } 1 \leq i < M, \\ \beta_i - \sum_{s \in S} P_1^M(s) V(x_i, y_i) &= 0, \quad \text{for } 1 \leq i \leq M, \\ \gamma - \sum_{s \in S} P_1^M(s) \beta(x_0) &= 0, \\ 1 - \sum_{s \in S} P_1^M(s) &= 0, \end{aligned}$$

which gives us the Gibbs distribution:

$$P_1^M(s) = \exp \left\{ - \sum_{i=1}^{M-1} \lambda_i U(x_i, x_{i+1}; u) - \sum_{i=1}^{M-1} \mu_i V(x_i, y_i) - \nu \beta(x_0) - \kappa - 1 \right\} \quad (17)$$

for $s \in S$, and,

$$\alpha_i \sum_{s \in S} P_1^M(s) - \sum_{s \in S} P_1^M(s) U(x_i, x_{i+1}; u) = 0, \quad (18)$$

for $1 \leq i < M$, and,

$$\beta_i \sum_{s \in S} P_1^M(s) - \sum_{s \in S} P_1^M(s) V(x_i, y_i) = 0, \quad (19)$$

for $1 \leq i \leq M$, and,

$$\gamma \sum_{s \in S} P_1^M(s) - \sum_{s \in S} P_1^M(s) \beta(x_0) = 0, \quad (20)$$

and,

$$\exp \{ \kappa + 1 \} = \sum_{s \in S} \exp \left\{ - \sum_{i=1}^{M-1} \lambda_i U(x_i, x_{i+1}; u) - \sum_{i=1}^{M-1} \mu_i V(x_i, y_i) - \nu \beta(x_0) \right\}. \quad (21)$$

Now we introduce some simplifying notation.

For $x_i \in \mathbf{X}, x_{i+1} \in \mathbf{X}$:

$$\begin{aligned} \sum_{x_i \in \mathbf{X}, x_{i+1} \in \mathbf{X}} \exp \{ -\lambda_i U(x_i, x_{i+1}; u) \} &= \sum_j E_j, \\ \sum_{x_i \in \mathbf{X}, x_{i+1} \in \mathbf{X}} U(x_i, x_{i+1}; u) &= \sum_j F_j. \end{aligned}$$

The multiplier (λ_i) can be known by solving the following equation:

$$\alpha_i \sum_j E_j - \sum_j E_j F_j = 0, \quad \text{for } 1 \leq i \leq M-1. \quad (22)$$

Now, we want to show that the multiplier (λ_i) is a monotonic function of the corresponding disturbance level α_i .

Consider the function:

$$f(\lambda) = \alpha_i \sum_j E_j - \sum_j E_j F_j = \sum_j E_j (\alpha_i - F_j).$$

If λ_i is a zero then, $0 = \sum_j E_j (\alpha_i - F_j)$, and, $0 = \sum_j E_j \left(1 - \alpha_i \frac{d\lambda_i}{d\alpha_i} F_j + \frac{d\lambda_i}{d\alpha_i} F_j^2 \right)$. This gives us:

$$\frac{d\lambda_i}{d\alpha_i} = \frac{\sum_j E_j}{\sum_j E_j (\alpha_i F_j - F_j^2)}.$$

The above expression is negative if its denominator D is.

$$\begin{aligned} D &= \alpha_i \sum_j E_j F_j - \sum_j E_j F_j^2 \\ &= \frac{\sum_j E_j F_j}{\sum_j E_j} \sum_j E_j F_j - \sum_j E_j F_j^2 \\ &= \frac{(\sum_j E_j F_j)^2}{\sum_j E_j} - (\sum_j E_j F_j^2) \leq 0. \end{aligned}$$

In a manner entirely parallel to this, we can solve for the multipliers $\{\mu_i\}, \nu$ as well as prove that they decrease monotonically as the disturbance levels are raised.

Also, when $\alpha_i = \min_{x_i, x_j} \{U(x_i, x_j; u)\}$, $\lambda_i = \infty$. Thus when all the disturbance levels go down to zero, the Gibbs distribution we get as the maximum entropy solution, reduces to the unperturbed FSM we started with.

When we parametrize the expected disturbance levels so that the Lagrange multipliers ($\{\lambda_i\}, \{\mu_i\}, \nu$) are all equal to ε , the maximum entropy model we derived (17) is a *controlled Hidden Markov Model*. It consists of an \mathbf{X} valued controlled Markov chain x_k^ε together with a \mathbf{Y} valued output process y_k^ε . We have:

$$\begin{aligned} \mathbf{P}^u (x_{k+1}^\varepsilon = x'' \mid x_k = x, x_{0,k-1}, u_k = u, u_{0,k-1}) &= A^\varepsilon(u)_{x, x''}, \\ \mathbf{P}^u (y_{k+1}^\varepsilon = y'' \mid x_k = x) &= B^\varepsilon(x)_{y''}, \end{aligned}$$

where,

$$\left. \begin{aligned} A^\varepsilon(u)_{x, x''} &\triangleq \frac{1}{Z_{x, u}^\varepsilon} \exp \left(-\frac{1}{\varepsilon} U(x, x''; u) \right), \\ B^\varepsilon(x)_{y''} &\triangleq \frac{1}{Z_x^\varepsilon} \exp \left(-\frac{1}{\varepsilon} V(x, y'') \right), \end{aligned} \right\} \quad (23)$$

where the functions U and V are defined by (3), and the normalizing constants $Z_{x,u}^\varepsilon$ and Z_x^ε are chosen so that

$$\sum_{x'' \in \mathbf{X}} A^\varepsilon(u)_{x,x''} = 1, \quad \sum_{y'' \in \mathbf{Y}} B^\varepsilon(x)_{y''} = 1.$$

Similarly, the initial distribution is

$$\rho^\varepsilon(x_0) = \frac{1}{Z_{x_0}^\varepsilon} \exp\left(-\frac{1}{\varepsilon}\beta(x_0)\right).$$

Thus where \mathbf{P}^u is the probability distribution on $\mathbf{X}^{M+1} \times \mathbf{Y}^M$ defined by a control policy $u \in \mathcal{O}_{0,M-1}$:

$$\mathbf{P}^u(x_{0,M}, y_{1,M}) = \prod_{k=0}^{M-1} A^\varepsilon(u_k)_{x_k, x_{k+1}} B^\varepsilon(x_k)_{y_{k+1}} \rho^\varepsilon(x_0)$$

The probability distribution \mathbf{P}^u is equivalent to a distribution \mathbf{P}^\dagger under which $\{y_k^\varepsilon\}$ is iid uniformly distributed on \mathbf{Y} , independent of $\{x_k^\varepsilon\}$, and $\{x_k^\varepsilon\}$ is a controlled Markov chain as above:

$$\mathbf{P}^\dagger(x_{0,M}, y_{1,M}) = \prod_{k=0}^{M-1} \left(A^\varepsilon(u_k)_{x_k, x_{k+1}} \frac{1}{p} \right) \rho^\varepsilon(x_0),$$

where p here denotes the number of outputs. Note that

$$\frac{d\mathbf{P}^u}{d\mathbf{P}^\dagger} |_{\mathcal{G}_k} = \lambda_k^\varepsilon \triangleq \prod_{l=1}^k \Psi^\varepsilon(x_{l-1}^\varepsilon, y_l^\varepsilon),$$

where

$$\Psi^\varepsilon(x, y'') \triangleq p B^\varepsilon(x)_{y''},$$

and \mathcal{G}_k is the filtration generated by $(x_{0,k}^\varepsilon, y_{1,k}^\varepsilon)$.

B. Cost

The cost function is defined for admissible $u \in \mathcal{O}_{0,M-1}$ by

$$J^{\gamma,\varepsilon}(u) = \mathbf{E}^u \left[\exp \frac{1}{\gamma\varepsilon} \sum_{l=0}^{M-1} \phi_z(\ell(x_l^\varepsilon, u_l)) \right] \quad (24)$$

and the *output feedback risk-sensitive stochastic control problem* for the HMM (23) is to find $u^* \in \mathcal{O}_{0,M-1}$ so that

$$J^{\gamma,\varepsilon}(u^*) = \min_{u \in \mathcal{O}_{0,M-1}} J^{\gamma,\varepsilon}(u).$$

In terms of the reference measure, the cost can be expressed as

$$J^{\gamma,\varepsilon}(u) = \mathbf{E}^\dagger \left[\lambda_M^\varepsilon \exp \frac{1}{\gamma\varepsilon} \sum_{l=0}^{M-1} \phi_z(\ell(x_l^\varepsilon, u_l)) \right]. \quad (25)$$

C. Information State

Following [7], [24], [25], [26], we define an information state process $\sigma_k^{\gamma,\varepsilon} \in \mathbf{R}^n$ by the relation

$$\sigma_k^{\gamma,\varepsilon}(x) = \mathbf{E}^\dagger \left[I_{\{x_k^\varepsilon = x\}} \exp \frac{1}{\gamma\varepsilon} \sum_{l=0}^{k-1} \phi_z(\ell(x_l^\varepsilon, u_l)) \lambda_k^\varepsilon | \mathcal{Y}_k \right], \quad (26)$$

where \mathcal{Y}_k is the filtration generated by the observation process $y_{1,k}^\varepsilon$, and $\sigma_0^{\gamma,\varepsilon}(x) = I_{\{x = x_0\}}$.

The evolution of this process is determined by a matrix $\Sigma^{\gamma,\varepsilon}(u, y'')$ whose entries are defined by

$$\Sigma^{\gamma,\varepsilon}(u, y'')_{x,x''} \triangleq A^\varepsilon(u)_{x,x''} \Psi^\varepsilon(x, y'') \exp \frac{1}{\gamma\varepsilon} \phi_z(\ell(x, u)). \quad (27)$$

Indeed, the information state is the solution of the recursion (c.f. [7], [26])

$$\left. \begin{aligned} \sigma_k^{\gamma,\varepsilon} &= \Sigma^{\gamma,\varepsilon*}(u_{k-1}, y_k^\varepsilon) \sigma_{k-1}^{\gamma,\varepsilon} \\ \sigma_0^{\gamma,\varepsilon} &= \rho^\varepsilon, \end{aligned} \right\} \quad (28)$$

where the $*$ denotes matrix transpose. We can also define an adjoint process $\nu_k^{\gamma,\varepsilon} \in \mathbf{R}^n$ by the recursion

$$\left. \begin{aligned} \nu_{k-1}^{\gamma,\varepsilon} &= \Sigma^{\gamma,\varepsilon}(u_{k-1}, y_k^\varepsilon) \nu_k^{\gamma,\varepsilon} \\ \nu_M^{\gamma,\varepsilon} &= 1. \end{aligned} \right\} \quad (29)$$

Then relative to the inner product

$$\langle \sigma, \nu \rangle = \sum_{x \in \mathbf{X}} \sigma(x) \nu(x)$$

for $\sigma, \nu \in \mathbf{R}^n$, it is straightforward to establish the adjoint relationships

$$\begin{aligned} \langle \Sigma^{\gamma,\varepsilon*} \sigma, \nu \rangle &= \langle \sigma, \Sigma^{\gamma,\varepsilon} \nu \rangle, \\ \langle \sigma_k^{\gamma,\varepsilon}, \nu_k^{\gamma,\varepsilon} \rangle &= \langle \sigma_{k-1}^{\gamma,\varepsilon}, \nu_{k-1}^{\gamma,\varepsilon} \rangle \end{aligned}$$

for all $\sigma \in \mathbf{R}^n$, $\nu \in \mathbf{R}^n$, and all k .

Remark 3.1: The reason for introducing the information state $\sigma_k^{\gamma,\varepsilon}$ is to replace the original output feedback risk-sensitive stochastic control problem with an equivalent stochastic control problem with a state variable $\sigma_k^{\gamma,\varepsilon}$ which is completely observed, and to solve this new problem using dynamic programming. This will yield a state feedback controller for the new problem, or equivalently, an output feedback controller for the original problem which is *separated* through the information state [9], [3], [7], [24], [25], [26].

As in [7], [26], the cost function can be expressed purely in terms of the information state:

$$J^{\gamma,\varepsilon}(u) = \mathbf{E}^\dagger [\langle \sigma_M^{\gamma,\varepsilon}, 1 \rangle]. \quad (30)$$

D. Dynamic Programming

Consider the state $\sigma^{\gamma,\varepsilon}$ on the interval k, \dots, M with initial condition $\sigma_k^{\gamma,\varepsilon} = \sigma \in \mathbf{R}^n$:

$$\left. \begin{aligned} \sigma_l^{\gamma,\varepsilon} &= \Sigma^{\gamma,\varepsilon*}(u_{l-1}, y_l^\varepsilon) \sigma_{l-1}^{\gamma,\varepsilon}, \quad k+1 \leq l \leq M, \\ \sigma_k^{\gamma,\varepsilon} &= \sigma. \end{aligned} \right\} \quad (31)$$

The corresponding value function for this control problem is defined for $\sigma \in \mathbf{R}^n$ by

$$S^{\gamma,\varepsilon}(\sigma, k) = \min_{u \in \mathcal{O}_{k,M-1}} \mathbf{E}^\dagger [\langle \sigma_M^{\gamma,\varepsilon}, 1 \rangle | \sigma_k^{\gamma,\varepsilon} = \sigma]. \quad (32)$$

The dynamic programming equation for this problem is as follows [7], [26]:

$$\left. \begin{aligned} S^{\gamma,\varepsilon}(\sigma, k) &= \min_{u \in U} \mathbf{E}^\dagger [S^{\gamma,\varepsilon}(\Sigma^{\gamma,\varepsilon*}(u, y_{k+1}^\varepsilon) \sigma, k+1)] \\ S^{\gamma,\varepsilon}(\sigma, M) &= \langle \sigma, 1 \rangle. \end{aligned} \right\} \quad (33)$$

The next theorem is a statement of the dynamic programming solution to the output feedback risk-sensitive stochastic control problem.

Theorem 3.2: [26] The value function $S^{\gamma,\varepsilon}$ defined by (32) is the unique solution to the dynamic programming equation (33). Conversely, assume that $S^{\gamma,\varepsilon}$ is the solution of the dynamic programming equation (33). Suppose that $u^* \in \mathcal{O}_{0,M-1}$ is a policy such that, for each $k = 0, \dots, M-1$, $u_k^* = \bar{u}_k^*(\sigma_k^{\gamma,\varepsilon})$, where $\bar{u}_k^*(\sigma)$ achieves the minimum in (33). Then u^* is an optimal output feedback controller for the risk-sensitive stochastic control problem (§III-B).

Remark 3.3: Note that the controller u_k^* is defined as a function of the information state $\sigma_k^{\gamma,\varepsilon}$, and since $\sigma_k^{\gamma,\varepsilon}$ is a non-anticipating function of $y_{0,k}^\varepsilon$, u_k^* is an output feedback controller for the risk-sensitive stochastic control problem; indeed, u^* is an information state feedback controller.

E. Small Noise Limit

In [7], [24], [25], [26] it was shown that a deterministic dynamic game problem is obtained as a small noise limit of a risk-sensitive stochastic control problem. In this subsection, we carry out this limit procedure for the risk-sensitive stochastic control problem defined above. We first obtain a limit for the information state, and use this to evaluate the appropriate limit for the value function. This yields an information state and value function for the dynamic game problem of Section II-D. These results will be used in Section IV in the solution of the output feedback robust control problem of Section II.

Define the matrix $\Lambda^\gamma(u, y'')$ by its entries

$$\Lambda^\gamma(u, y'')_{x,x''} \triangleq \phi_z(\ell(x, u)) - \gamma(U(x, x''; u) + V(x, y'')). \quad (34)$$

Then we have

$$\lim_{\varepsilon \rightarrow 0} \gamma \varepsilon \log \Sigma^{\gamma,\varepsilon}(u, y'')_{x,x''} = \Lambda^\gamma(u, y'')_{x,x''}. \quad (35)$$

The action of the matrix $\Sigma^{\gamma,\varepsilon}$ and its adjoint (transpose) on vectors σ, ν in \mathbf{R}^n is given by the usual matrix multiplication, i.e., sums of products of entries. The action of the matrix $\Lambda^\gamma(u, y'')$ and its adjoint on vectors p, q in \mathbf{R}^n is instead defined in terms of maximization operations as follows:

$$\begin{aligned} \Lambda^{\gamma*}(u, y'')p(x'') &\triangleq \max_{x \in \mathbf{X}} \{\Lambda^\gamma(u, y'')_{x,x''} + p(x)\}, \\ \Lambda^\gamma(u, y'')q(x) &\triangleq \max_{x'' \in \mathbf{X}} \{\Lambda^\gamma(u, y'')_{x,x''} + q(x'')\}. \end{aligned}$$

The inner product $\langle \cdot, \cdot \rangle$ is replaced by the ‘‘sup-pairing’’

$$(p, q) \triangleq \max_{x \in \mathbf{X}} \{p(x) + q(x)\}, \quad (36)$$

and in fact we have

$$\lim_{\varepsilon \rightarrow 0} \gamma \varepsilon \log \langle e^{\frac{1}{\gamma \varepsilon} p}, e^{\frac{1}{\gamma \varepsilon} q} \rangle = (p, q). \quad (37)$$

The actions corresponding to the matrix $\Lambda^\gamma(u, y'')$ are ‘‘adjoint’’ in the sense that

$$(\Lambda^{\gamma*}p, q) = (p, \Lambda^\gamma q). \quad (38)$$

The limit result for the information state is the following:

Theorem 3.4: [26] We have

$$\left. \begin{aligned} \lim_{\varepsilon \rightarrow 0} \gamma \varepsilon \log \Sigma^{\gamma,\varepsilon*}(u, y) e^{\frac{1}{\gamma \varepsilon} p} &= \Lambda^{\gamma*}(u, y)p, \\ \lim_{\varepsilon \rightarrow 0} \gamma \varepsilon \log \Sigma^{\gamma,\varepsilon}(u, y) e^{\frac{1}{\gamma \varepsilon} q} &= \Lambda^\gamma(u, y)q \end{aligned} \right\} \quad (39)$$

in \mathbf{R}^n uniformly on $\mathbf{U} \times \mathbf{Y} \times \mathbf{R}^n$.

In view of this theorem, we define a limit information state and its adjoint by the recursions

$$\left. \begin{aligned} p_k^\gamma &= \Lambda^{\gamma*}(u_{k-1}, y_k)p_{k-1}^\gamma \\ p_0^\gamma &= -\beta, \end{aligned} \right\} \quad (40)$$

and

$$\left. \begin{aligned} q_{k-1}^\gamma &= \Lambda^\gamma(u_{k-1}, y_k)q_k^\gamma \\ q_M^\gamma &= 0. \end{aligned} \right\} \quad (41)$$

Note that

$$(p_k^\gamma, q_k^\gamma) = (p_{k-1}^\gamma, q_{k-1}^\gamma)$$

for all k .

Turning now to the value function, we have:

Theorem 3.5: [26] The function $W^\gamma(p, k)$ defined for $p \in \mathbf{R}^n$ by

$$W^\gamma(p, k) \triangleq \lim_{\varepsilon \rightarrow 0} \gamma \varepsilon \log S^{\gamma,\varepsilon}(e^{\frac{1}{\gamma \varepsilon} p}, k) \quad (42)$$

exists (i.e. the sequence converges uniformly on \mathbf{R}^n), is continuous, and satisfies the recursion

$$\left. \begin{aligned} W^\gamma(p, k) &= \min_{u \in \mathbf{U}} \max_{y \in \mathbf{Y}} \{W^\gamma(\Lambda^{\gamma*}(u, y)p, k+1)\} \\ W^\gamma(p, M) &= (p, 0). \end{aligned} \right\} \quad (43)$$

IV. SOLUTION TO THE ROBUST CONTROL PROBLEM

A. Equivalent Game Problem

We now replace the deterministic output feedback game problem (Section II) with an equivalent deterministic game problem with p_k^γ , defined in Section III-E, as a completely observed state variable. The solution of this new problem will result in an information state feedback controller, and thus an output feedback controller for the original game problem which is *separated* through the information state.

The next theorem shows that the cost function can be expressed in terms of the information state [7], [24], [25], [26].

Theorem 4.1: We have for all $u \in \mathcal{O}_{0,M-1}$

$$J^\gamma(u) = \max_{y \in \mathbf{Y}^M} \{(p_M^\gamma, 0)\}. \quad (44)$$

Proof: We have, for all $u \in \mathcal{O}_{0,M-1}$,

$$\begin{aligned} \max_{y \in \mathbf{Y}^M} \{(p_M^\gamma, 0)\} &= \max_{y \in \mathbf{Y}^M} \max_{x'' \in \mathbf{X}} \{p_M^\gamma(x'')\} \\ &= \max_{w \in \mathbf{W}^M} \max_{v \in \mathbf{V}^M} \max_{x_0 \in \mathbf{X}} \{J^\gamma(u, w, v, x_0)\} = J^\gamma(u), \end{aligned}$$

where we have made use of the definitions for the cost functions U and V (Section II-B).

B. Dynamic Programming

Consider now the state p^γ on the interval k, \dots, M with initial condition $p_k^\gamma = p \in \mathbf{R}^n$:

$$\left. \begin{aligned} p_l^\gamma &= \Lambda^{\gamma*}(u_{l-1}, y_l)p_{l-1}^\gamma, \quad k+1 \leq l \leq M, \\ p_k^\gamma &= p. \end{aligned} \right\} \quad (45)$$

The value function is defined for $p \in \mathbf{R}^n$ by

$$W^\gamma(p, k) = \min_{u \in \mathcal{O}_{k, M-1}} \max_{y \in \mathbf{Y}^{M-k}} \{(p_M^\gamma, 0) : p_k^\gamma = p\}. \quad (46)$$

The solution of the game problem is expressed as follows.

Theorem 4.2: The value function $W^\gamma(p, k)$ defined by (46) is the unique solution to the dynamic programming equation (43). Further, if $W^\gamma(p, k)$ is the solution of (43), and if $u^* \in \mathcal{O}_{0, M-1}$ is a policy such that, for each $k = 0, \dots, M-1$, $u_k^* = \bar{u}_k^*(p_k^\gamma)$, where $\bar{u}_k^*(p)$ achieves the minimum in (43), then u^* is an optimal policy for the output feedback dynamic game problem (Section II-D).

Proof: Standard dynamic programming arguments.

C. Robust Control

The solution to the state feedback robust control problem was expressed in terms of the solution $\bar{f}_k^\gamma(x)$ of a dynamic programming equation, and a state feedback controller $\bar{u}_k^*(x)$ was obtained. The framework we have developed in this paper allows us to characterize the solution of the output feedback robust control problem in terms of the solution $W^\gamma(p, k)$ of a dynamic programming equation, and obtain an output feedback controller $\bar{u}_k^*(p_k^\gamma(\cdot; y_{1,k}))$. Note that the information state p_k^γ is also the solution of a dynamic programming equation (40).

Theorem 4.3: (Necessity) Assume that there exists a controller $u^o \in \mathcal{O}_{0, M-1}$ solving the output feedback robust control problem. Then there exists a solution $W^\gamma(p, k)$ of the dynamic programming equation (43) such that $W^\gamma(-\beta, 0) = 0$. (Sufficiency) Assume that there exists a solution $W^\gamma(p, k)$ of the dynamic programming equation (43) such that $W^\gamma(-\beta, 0) = 0$, and let $\bar{u}_k^*(p)$ be a control value achieving the minimum in (43). Then $\bar{u}_k^*(p_k^\gamma(\cdot; y_{1,k}))$ is an output feedback controller which solves the output feedback robust control problem.

V. CONCLUSIONS AND DISCUSSION

In this paper we showed that the robust output feedback control problem for automata is solvable using deterministic and stochastic models for the uncertainties. The stochastic model was shown to be a HMM derived via the maximum entropy principle.

For future work, it would be interesting to investigate connections with model complexity and extensions to more general dynamical systems, including hybrid systems.

REFERENCES

- [1] J.A. Ball and J.W. Helton, "Nonlinear H^∞ Control for Stable Plants", *Math. Cont., Signals, and Syst.*, Vol. 5, pp. 233–261, 1992.
- [2] T. Basar and P. Bernhard, *H^∞ -Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*, Birkhauser, 1991.
- [3] A. Bensoussan and J.H. van Schuppen, "Optimal Control of Partially Observable Stochastic Systems with an Exponential-of-Integral Performance Index", *SIAM J. Cont. Optim.*, Vol. 23, pp. 599–613, 1985.
- [4] J.C. Doyle, K. Glover, P.P. Khargonekar and B.A. Francis, "State-Space Solutions to Standard H^2 and H^∞ Control Problems", *IEEE Trans. Aut. Control*, Vol. AC-34 (8), pp. 831–847, 1989.
- [5] M.I. Friedlin and A.D. Wentzell, *Random Perturbations of Dynamical Systems*, Springer-Verlag, New York, 1984.
- [6] M.R. James, "Finite Time Observer Design by Probabilistic – Variational Methods", *SIAM J. Cont. Optim.*, Vol. 29 (4), pp. 954–967, 1991.
- [7] M.R. James, J.S. Baras and R.J. Elliott, "Risk-Sensitive Control and Dynamic Games for Partially Observed Discrete-Time Nonlinear Systems", *IEEE T. Aut. Cont.*, Vol. AC-39 (4), pp. 780–792, 1994.
- [8] R.E. Kalman, P.L. Falb and M.A. Arbib, *Topics in Mathematical System Theory*, McGraw-Hill, New York, 1969.
- [9] P.R. Kumar and P. Varaiya, *Stochastic Systems: Estimation, Identification, and Adaptive Control*, Prentice-Hall, Englewood Cliffs, 1986.
- [10] D.J.N. Limebeer, B.D.O. Anderson, P.P. Khargonekar and M. Green, "A Game Theoretic Approach to H^∞ Control for Time Varying Systems", *SIAM J. Control Optim.*, Vol. 30, pp. 262–283, 1992.
- [11] P. Whittle, *Risk-Sensitive Optimal Control*, Wiley, New York, 1990.
- [12] C. Martin-Vide and V. Mitrana (Edts), *Grammars and Automata for String Processing*, Taylor and Francis: 2003.
- [13] R. Alur and G.J. Pappas (Edts), *Hybrid Systems: Computation and Control*, LNCS 2993, Springer: 2004.
- [14] J.W. Helton and M. James, *Extending H^∞ Control to Nonlinear Systems*, SIAM, Philadelphia PA, 1999.
- [15] D. Haussler, "A General Minimax Result for Relative Entropy." *IEEE Trans. Inform. Theory*, vol. 43, pp. 1276-1280.
- [16] J.O. Berger, *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. Springer: 1985, New York.
- [17] P. Harremoës and F. Topsøe, "Maximum Entropy Fundamentals", *Entropy*, vol. 3, pp. 191-226, 2001.
- [18] P. Harremoës and F. Topsøe, "Unified Approach to Optimization Techniques in Shannon Theory", in *Proc. 2002 IEEE International Symposium on Information Theory* p. 238, IEEE, New York, 2002.
- [19] E.T. Jaynes, "Information Theory and Statistical Mechanics, I", *Phys. Rev.*, vol. 106, pp. 620-630, 1957.
- [20] E.T. Jaynes, "Information Theory and Statistical Mechanics, II", *Phys. Rev.*, vol. 106, pp. 171-190, 1957.
- [21] I. Csiszar, "Why Least Squares and Maximum Entropy? An Axiomatic Approach to Inference for Linear Inverse Problems", *Ann. Statist.*, vol. 19, pp. 2032-2066, 1991.
- [22] J.N. Kapur, *Maximum Entropy Models in Science and Engineering*, Wiley: 1993, New York.
- [23] J.S. Baras, R. Elliott and M. Kohlmann, "The Partially Observed Stochastic Minimum Principle", *SIAM Journal on Control and Optimization*, Vol. 27, No. 6, pp. 1279-1292, November 1989.
- [24] M.R. James and J.S. Baras, "Robust H^∞ Output Feedback Control for Nonlinear Systems", *IEEE Transactions on Automatic Control*, pp. 1007-1017, Vol. 40, No. 6, June 1995.
- [25] M.R. James and J.S. Baras, "Partially Observed Differential Games, Infinite Dimensional HJI Equations, and Nonlinear H^∞ Control", *SIAM Journal on Cont. and Optim.*, Vol. 34, No. 4, pp. 1342-1364, July 1996.
- [26] J.S. Baras and M.R. James, "Robust and Risk-Sensitive Output Feedback Control for Finite State Machines and Hidden Markov Models", *Journal of Math. Systems, Estim., and Contr.*, Vol. 7, No. 3, pp. 371-374, 1997.
- [27] J.S. Baras and N.S. Patel, "Robust Control of Set-Valued Discrete Time Dynamical Systems", *IEEE T. on Aut. Cont.*, Vol. 43, No. 1, pp. 61-75, January 1998.
- [28] P.D. Grünwald and A.P. Dawid, "Game Theory, Maximum Entropy, Minimum Discrepancy and Robust Decision Theory", *the Annals of Statistics*, Vol. 32, No. 4, pp. 1367-1433, 2004.