Proceedings of the 41st IEEE
Conference on Decision and Control
Las Vegas, Nevada USA, December 2002

TuP08-5

Action Functional Stochastic H^∞ Estimation for Nonlinear Discrete Time Systems

[†]Charalambos .D. Charalambous[1],[†]Alireza Farhadi, and [‡]Seddik M. Djouadi

[†] School of Information Technology and Engineering, University of Ottawa
Ottawa, Ontario, K1N 6N5, Canada
E-mail: {chadcha , afarhadi }@site.uottawa.ca

[‡] System Engineering Department, University of Arkansas at Little rock
Little Rock, AR 72204
E-mail: msdjoudi@ualr.edu

Abstract

This paper presents an action functional, sample path optimization technique, for formulating and solving nonlinear discrete-time stochastic H^∞ estimation problems. These H^∞ problems are formulated as minimax dynamic games in which the maximizing players are stochastic square summable disturbances, while the minimizing players are the state estimates. Certain action functionals are defined which play the role of information state and its adjoint in converting the minimax game into a fully observable game. Subsequently, a verification theorem is derived.

Key Words: Stochastic, Nonlinear, Minimax Games, Discrete, Information State, Estimation.

1 Introduction

Since the publication of Zames[1] seminal paper on H^∞ optimization, several approaches have been proposed to extend the techniques of robust design, with respect to unknown disturbances and unmodeled dynamics, to nonlinear stochastic as well as deterministic systems. This generalization leads to a minimax formulation in which the exogenous inputs or disturbances are the maximizing players and the controllers or estimators are the minimizing players. Previous work is formed in [2, 3, 4, 5, 6, 7, 8]. Unfortunately, little work is done in formulating and analyzing stochastic minimax partially observable problems.

The difficulty is encountered in identifying the infor-

mation state. However, under the so-called matching condition which implies that square summable disturbances and color noises are entering the dynamics and observation through the same channel, then an application of certain results from Large Deviation yields an equivalent problem with an exponential pay-off, known as risk-sensitive problem [9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20]. Thus, risk-sensitive measures and large deviations theory offers an indirect method for solving stochastic minimax games under the matching condition. However, since the matching condition, is a rather severe requirement for any typical system to satisfy, one would like to remove the matching condition by providing a direct and more general approach in analyzing minimax stochastic dynamic games.

This paper introduces an information state approach in formulating and analyzing stochastic minimax estimation problems, without imposing the matching condition. The adjoint of the information state is also introduced and recursions are derived using dynamic programming.

2 The Minimax Estimation Problem

2.1 Dynamics

Let $\left(\Omega, \mathcal{F}, \mathbf{P}^\nu\right)$ be a basis probability space on which the state process $x_{[0,N]} \stackrel{\triangle}{=} \{x_n\}_{n=0}^N$ and the observation process $y_{[0,N]} \stackrel{\triangle}{=} \{y_n\}_{n=0}^N$ are defined as follows.

$$x_{k+1} = f_{\gamma,w}(k, x_k, \gamma_k, w_k), \quad x_0 \in \Re^n, \qquad (1)$$

$$y_k = h_{\delta,v}(k, x_k, \delta_k, v_k), \quad y_0 \in \Re^d, \qquad (2)$$

in which $x_0 : \Omega \to \Re^n$, $w : [0, N-1] \times \Omega \to \Re^{n_1}$, $v : [0, N] \times \Omega \to \Re^{d_1}$ $\gamma : [0, N-1] \times \Omega \to \mathcal{D}^\gamma$, $\delta : [0, N] \times \Omega \to \mathcal{D}^\delta$. Here $w_{[0,N-1]} \stackrel{\triangle}{=} \{w_n\}_{n=0}^{N-1}$, $v_{[0,N]} \stackrel{\triangle}{=} \{v_n\}_{n=0}^N$ are

finite-dimensional independent sequences of random variables, and $\gamma_{[0,N-1]} \triangleq \{\gamma_n\}_{n=0}^{N-1}$, $\delta_{[0,N]} \triangleq \{\delta_n\}_{n=0}^{N}$ are square summable disturbances. Assumptions on the vectors $f_{\gamma,\omega}$, $h_{\delta,v}$ ensuring unique weak solutions are given under Assumptions 2.1

Notation 2.1 *Let $\mathcal{G}_{0,n}^0 = \sigma\{x_k, y_k; k = 0, 1, \ldots, n\}$ and $\mathcal{Y}_{0,n}^0 = \sigma\{y_k; k = 0, 1, \ldots, n\}$ denote the sigma-algebras generated by the complete and incomplete data, respectively, and denote by $\{\mathcal{G}_{0,k}\}, \{\mathcal{Y}_{0,k}\}, k \in [0, N]$ their complete filtration with respect to $(\Omega, \mathcal{F}, \mathbf{P}^\nu)$.*
Let H be a Hilbert space with norm $\|\cdot\|_H$ and let $\{\mathcal{Z}_{0,k}\}, k \in [0, N]$ denote a complete filtration with respect to $(\Omega, \mathcal{F}, \mathbf{P}^\nu)$; define the Banach space of stochastic processes as follows.

$$\ell_{\mathcal{Z}}^p\Big([0,m];H\Big) \triangleq \Big\{ \phi_{[0,m]} = \Big\{ \phi_n; 0 \le n \le m \Big\};$$
$$\phi : [0,m] \times \Omega \to H \;\; such \; that \; \phi_n \; is \; an \; \mathcal{Z}_{0,n}$$
$$measurable \; random \; variable \; on \;\; [0,m] \;\; with$$
$$E^\nu\Big\{ \sum_{k=0}^m \|\phi_k\|_H^p < \infty, \;\; 1 \le p < \infty \Big\}.$$

For $\phi \in \ell_{\mathcal{Z}}^p([k,m];H)$ the norm is defined by $\|\phi\|_{\mathcal{Z},p} \triangleq$
$\Big\{ \sum_{n=k}^m E^\nu \|\phi_k\|_H^p \Big\}^{\frac{1}{p}}$.

Assumption 2.1 *(i) X is a compact subset of $\Re^n$ and $D^\gamma = \Re^n, D^\delta = \Re^d$. (ii) $y_0 = 0 \in \Re^d$, x_0 is unknown deterministic (or random with distribution $d\Pi(x)$ such that $x_0 \in L^2(\Omega)$). (iii) $p_0 : \Re^n \to (-\infty, 0], p_0 \in C(\Re^n)$.(and if x_0 is random then $\int |P_0| d\Pi(x) < \infty)$. (iv) $w_{[0,N-1]} \triangleq \{w_k\}_{n=0}^{N-1}$ is an $\Re^{n_1}$-valued independent sequence of random variables with density $\{\phi_{w_n}(w)\}_{n=0}^{N-1}$. (v) $v_{[0,N]} \triangleq \{v_k\}_{n=0}^{N}$ is an $\Re^{d_1}$-valued independent sequence of random variables with strictly positive density $\{\psi_{v_n}(v)\}_{n=0}^{N}$. and mutually independent x_0, if x_0 is assumed to be a random variable (vi) $f_{\gamma,w} : [0, N-1] \times \Re^n \times D^\gamma \times \Re^{n_1} \to \Re^n$ is a Borel measurable function such that $f_{\gamma,w}(\cdot, x, \gamma, w) = f_w(\cdot, x, w) + \sigma(\cdot, x)\gamma$, and three exists $c_1 > 0$ such that $\sigma(k,x)\sigma'(k,x) \ge I_n \times c_1, \forall k \in [0, N-1], x \in \Re^n$. where I_k denotes an $k \times k$ identity map. (viii) $h_{\delta,v} : [0, N] \times \Re^n \times D^\gamma \times \Re^{d_1} \to \Re^d$ is a Borel measurable function such that $h_{\delta,v}(\cdot, x, \delta, v) = h_v(\cdot, x, v) + \eta(\cdot, x)\delta$, and threre exists $c_2 > 0$ such that $\eta(k,x)\eta'(k,x) \ge I_d \times c_2, \forall k \in [0, N], x \in \Re^n$. (ix) There are inverse maps $\tilde{f}_{\gamma,w} : [0, N-1] \times \Re^n \times \Re^n \times D^\gamma$ and $\tilde{h}_{\delta,v} : [0, N] \times \Re^d \times \Re^n \times D^\delta$ such that $w_k = \tilde{f}_{\gamma,x}(k, x_{k+1}, x_k, \gamma_k), \;\; \forall k \in [0, N-1]$, which is differentiable in the second variable for each k, and $v_k = \tilde{h}_{\delta,y}(k, y_k, x_k, \delta_k), \;\; \forall k \in [0, N]$. (x) The matrix-valued functions $H_{\delta,v}(k, x, \delta_k, v_k) \triangleq \frac{\partial}{\partial v} h_{\delta,v}(k, x, \delta_k, v)|_{v=v_k}, \;\; \forall k \in [0, N], \tilde{H}_{\delta,v}(k, x, \delta_k, v_k) \triangleq \frac{\partial}{\partial x} \tilde{h}_{\delta,y}(k, y, x_k, \delta)|_{y=y_k}, \forall k \in$*

$[0, N]$. *are non-singular. (xi) $\lambda : [0, N-1] \times \Re^n \times X \to [0, \infty), \kappa : N \times \Re^n \to [0, \infty)$ are Borel measurable, $\lambda \in \ell_{\mathcal{G}}^2([0, N-1]; \Re)$ and $\kappa \in \ell_{\mathcal{G}}^2([N, N]; \Re)$.*

Definition 2.1 *The set of admissible estimations and disturbances are defined as follows.*

$\hat{\mathcal{X}}_{[0,N-1]} \triangleq \Big\{ \hat{x} \; : \; [0, N] \times \Omega \to X \subseteq \Re^k; \; \{\hat{x}_k\}, \; is$ *progressively measurable with respect to $\{\mathcal{Y}_{0,k}\}, k \in [0, N-1]$ and $E^\nu\Big\{ \sum_{k=0}^{N-1} \|\hat{x}_k\|_{\Re^n}^2 \Big\} < \infty \Big\}$. $\mathcal{D}_{[0,N-1]}^\gamma \triangleq \Big\{ \gamma : [0, N-1] \times \Omega \to D^\gamma = \Re^n; \; \{\gamma_k\} \; is \; progressively$ *measurable with respect to $\{\mathcal{G}_{0,k}\}, k \in [0, N-1]$ and $E^\nu\Big\{ \sum_{k=0}^{N-1} \|\gamma_k\|_{\Re^n}^2 \Big\} < \infty \Big\}$.*

Under Assumptions 2.1 (i)-(xi) it can be shown that the recursions (1), (2) have unique weak solutions.

2.2 Pay-Off Functional
For each sample path $y(\cdot, \omega) \in \ell^2([0, N]; \Re^d)$ and for each sample path $x(\cdot, \omega) \in \ell^2([0, N]; \Re^n)$ corresponding to admissible $(\gamma, \delta) \in \mathcal{D}_{[0,N-1]}^\gamma \times \mathcal{D}_{[0,N-1]}^\delta$ and $x_0 \in \Re^n$, the real-valued pay-off functional $J_{0,N}^\theta(\hat{x}, \gamma, \delta)$ is defined by

$$J_{0,N}^\theta(\hat{x}_{[0,N-1]}, \gamma_{[0,N-1]}, \delta_{[0,N-1]}) \triangleq E^\nu\Big\{ p(x_0) +$$
$$\sum_{k=0}^{N-1} \Big[\lambda(k, x_k, \hat{x}_k) - \frac{1}{\theta} \|\gamma_k\|_{\Re^n}^2 - \frac{1}{\theta} \|\delta_k\|_{\Re^d}^2 \Big] \Big\}. \quad (3)$$

where $E^\nu = E_{w,v}$ denotes expectation with respect to the distribution of the processes $(w_{[0,N-1]}, v_{[0,N-1]})$.
For $(\hat{x}, \gamma^*, \delta^*) \in \hat{\mathcal{X}}_{[0,N-1]} \times \ell_{\mathcal{G}}^2([0, N-1]; \Re^n) \times \ell_{\mathcal{G}}^2([0, N-1]; \Re^d)$ define

$$J_{0,N}^\theta(\hat{x}) \triangleq J_{0,N}^\theta(\hat{x}, \gamma^*, \delta^*) =$$
$$\sup_{(\gamma,\delta) \in \mathcal{D}_{[0,N-1]}^\delta \times \mathcal{D}_{[0,N-1]}^\gamma} E^\nu\Big\{ p(x_0) + \sum_{k=0}^{N-1} \Big[\lambda(k, x_k, \hat{x}_k)$$
$$-\frac{1}{\theta} \|\gamma_k\|_{\Re^n}^2 - \frac{1}{\theta} \|\delta_k\|_{\Re^d}^2 \Big] \Big\}. \quad (4)$$

The estimation problem is to determine $\hat{x}^* \in \hat{\mathcal{X}}_{[0,N-1]}$ which impacts

$$J_{0,N}^\theta(\hat{x}^*) = \inf_{\hat{x} \in \hat{\mathcal{X}}_{[0,N-1]}} J_{0,N}^\theta(\hat{x}, \gamma^*; \delta^*), \quad (5)$$

subject to the constraints (1), (2).

This problem will be formulated using an action functional as follows.

Definition 2.2 *For each sample path of the complete data $(y(\cdot, \omega), x(\cdot, \omega)) \in \ell^2([0, N]; \Re^d) \times \ell^2([1, N]; \Re^n)$,*

1029

generated by some admissible $(\gamma, \delta) \in \mathcal{D}_{[0,N-1]}^{\gamma} \times \mathcal{D}_{[0,N]}^{\delta}$ and $x_0 \in \Re^n$, the sample path pay-off functional with respect to $x_{[1,N]}, y_{[0,N-1]}$ is defined by

$$I_{0,N}^{\theta}(\hat{x}, x, y) \triangleq E_{w,v} \Big\{ p(x_0) + \sum_{k=0}^{N-1} \Big[\lambda(k, x_k, \hat{x}_k)$$

$$-\frac{1}{\theta} || \sigma^{-1}(k, x_k) \Big(x_{k+1} - f_w(k, x_k, w_k) \Big) ||_{\Re^n}^2$$

$$-\frac{1}{\theta} || \eta^{-1}(k, x_k) \Big(y_k - h_v(k, x_k, v_k) \Big) ||_{\Re^d}^2 \Big] \Big\}, \quad (6)$$

For a given $\hat{x} \in \hat{\mathcal{X}}_{[0,N-1]}$, let $y^*(\cdot, \omega), x^*(\cdot, \omega) \in (\ell^2([0, N-1]); \Re^d) \times (\ell^2([1, N]); \Re^n)$ denote the most likely sample path associated with the pay-off functional $I_{0,N}^{\theta}(\hat{x}, x, y)$ defined by

$$\Big(y_{[0,N-1]}^*, x_{[0,N]}^* \Big) \in \arg$$

$$\sup_{y \in \ell^2([0,N];\Re^n)} \sup_{x \in \ell^2([0,N-1];\Re^n)} I_{0,N-1}^{\theta}(\hat{x}, x, y) \quad (7)$$

Then

$$J_{0,N}^{\theta}(\hat{x}) \triangleq \sup_{y \in \ell^2([0,N-1];\Re^n)} \sup_{x \in \ell^2([0,N-1];\Re^n)} I_{0,N}^{\theta}(\hat{x}, x, y)$$

$$= E_{w,v} \sup_{y \in \ell^2([0,N-1];\Re^d)} \sup_{x \in \ell^2([0,N];\Re^n)} \Big\{ p_0(x_0)$$

$$+ \sum_{k=0}^{N-1} \Big[\lambda(k, x_k, \hat{x}_k) - \frac{1}{\theta} \int_{\Re^n} || \sigma^{-1}(k, x_k) \Big(x_{k+1} -$$

$$f_w(k, x_k, w) \Big) ||_{\Re^n}^2 \phi_{w_k}(w) dw - \frac{1}{\theta} \int_{\Re^d} || \eta^{-1}(k, x_k)$$

$$\Big(y_k - h_v(k, x_k, v_k) \Big) ||_{\Re^d}^2 \psi_v(v) dv \Big] \Big\}. \quad (8)$$

We shall call $I_{0,N}^{\theta}(\hat{x}, x, y)$ the action functional associated with (1), (2) and $(y^*(\cdot, \omega), x^*(\cdot, \omega)) \in \ell^2([0, N-1]; \Re^d) \times \ell^2([1, N]; \Re^n)$ the most likely sample paths associated with the observed and unobserved process, respectively, obtained by maximizing the action functional.

2.3 Sample Path Information State and Adjoint
In this section we introduce the sample path information state and its adjoint which we shall use to re-cast the partially observed minimax dynamic game (1), (2), $J_{0,N}^{\theta}(\hat{x}, \gamma, \delta)$ into a fully observed dynamic game.

Definition 2.3 For $\mathcal{A} \triangleq \Big\{ \alpha \in \Re^2; \alpha_1 > 0, \alpha_2 \geq 0 \Big\}$ define the spaces

$$\mathcal{B}^{\alpha} \triangleq \Big\{ p \in C(\Re^n); p(x) \leq -\alpha_1 ||x||_{\Re^n}^2 + \alpha_2, \ \alpha \in \mathcal{A} \Big\},$$

$$\mathcal{B} \triangleq \Big\{ p \in C(\Re^n); p(x) \leq -\alpha_1 ||x||_{\Re^n}^2 + \alpha_2, for \ \alpha \in \mathcal{A} \Big\},$$

$$C_b(\Re^n) \triangleq \Big\{ p \in C(\Re^n); |p(x)| \leq k, \ for \ some \ k \geq 0 \Big\}.$$

Define the sup pairing on the product space $\mathcal{B} \times C_b(\Re^n)$ by

$$< \pi, \zeta >_{sup} \triangleq \sup_{x \in \Re^n} \Big\{ \pi(x) + \zeta(x) \Big\}, \quad \pi \in \mathcal{B}, \zeta \in C_b(\Re^n).$$

For each $n \in [0, N]$ define the operators $T^{*,\theta} : \mathcal{B} \to \mathcal{B}, T^{\theta} : C_b(\Re^n) \to C_b(\Re^n)$ by

$$T^{*,\theta}(\hat{x}, y) \pi^{\theta}(z) \triangleq \sup_{x \in \Re^n} \Big\{ \lambda(n, x, \hat{x}) -$$

$$\frac{1}{\theta} \int_{\Re^{d_2}} || \sigma^{-1}(n, x) \Big(z - f_w(n, x, w) \Big) ||_{\Re^n}^2 \phi_{w_n}(w) dw$$

$$-\frac{1}{\theta} \int_{\Re^d} || \eta^{-1}(n, x) \Big(y - h_v(n, x) \Big) ||_{\Re^d}^2 \psi_{v_n}(v) dv$$

$$+ \pi^{\theta}(x) \Big\}. \quad (9)$$

The adjoint is defined by

$$T(\hat{x}, y) \zeta^{\theta}(x) \triangleq \sup_z \Big\{ \lambda(n, x, \hat{x}) -$$

$$\frac{1}{\theta} \int_{\Re^{d_2}} || \sigma^{-1}(n, x) \Big(z - f_w(n, x, w) \Big) ||_{\Re^n}^2 \phi_{w_n}(w) dw$$

$$-\frac{1}{\theta} \int_{\Re^d} || \eta^{-1}(n, x) \Big(y - h_v(n, x) \Big) ||_{\Re^d}^2 \psi_{v_n}(v) dv$$

$$+ \zeta^{\theta}(z) \Big\}. \quad (10)$$

Finally, with respect to the sup pairing $< \cdot, \cdot >_{sup}$ the above operators satisfy

$$< T^{\theta,*} \pi, \zeta >_{sup} = < \pi, T^{\theta} \zeta >_{sup}, \quad \forall \pi \in \mathcal{B}, \zeta \in C_b(\Re^n).$$

For some $0 < \theta \leq \theta^*$, and for each fixed observation path $y(\cdot, \omega) \in \ell^2([0, N]; \Re^d)$ and given a fixed state z, let $(x_0 \in \Re^n, \gamma \in \mathcal{D}_{[0,m]}^{\gamma}, \delta \in \mathcal{D}_{[0,m]}^{\delta})$ be the restrictions of $(x_0 \in \Re^n, \ \gamma \in \mathcal{D}_{[0,N-1]}^{\gamma}, \ \delta \in \mathcal{D}_{[0,N-1]}^{\delta})$ which generate a trajectory which at time m is $x_m = z$. Define the information state $\pi_m^{\theta}(\cdot) \in \mathcal{B}$ by

$$\pi_m^{\theta}(z) \triangleq \sup_{x \in \ell^2([0,m-1];\Re^n)} E_{w,v} \Big\{ p(x_0) + \sum_{k=0}^{m-1} \Big[$$

$$\lambda(k, x_k, \hat{x}_k) - \frac{1}{\theta} || \sigma^{-1}(k, \hat{x}_k) \Big(x_{k+1} - f_w(k, x_k, w_k) \Big) ||_{\Re^n}^2$$

$$-\frac{1}{\theta} || \eta^{-1}(k, x_k) \Big(y_k - h_v(k, x_k, v_k) \Big) ||_{\Re^d}^2 \Big] | x_m = z \Big\}. \quad (11)$$

This is the cost-to-go from stage $k = m$ to stage $k = 0$, in which $x_m = z$ is optimal.

Theorem 2.1 Consider a fixed sample path $y(\cdot, \omega) \in \ell^2([0, N]; \Re^d)$, $0 < \theta \leq \theta^*$, and $\hat{x} \in \hat{\mathcal{X}}_{[0,N-1]}$. The information state $\pi_m^{\theta}(\cdot) \in \mathcal{B}$ satisfies the following recursion.

$$\pi_m^{\theta}(z) = T^{\theta,*}(\hat{x}_{m-1}, y_{m-1}) \pi_{m-1}^{\theta}(z),$$

$$\pi_0^{\theta}(z) = p_0(z), m \in [1, N]. \quad (12)$$

1030

Proof. Follows from Dynamic programming.

Next, we consider the adjoint information state. For some $0 < \theta \le \theta^*$, and for each fixed observation path $y(\cdot, \omega) \in \ell^2([0,N]; \Re^d)$ and given a fixed state z, let $(x_0 \in \Re^n, \gamma \in \mathcal{D}_{[m,N-1]}^{\gamma}, \delta \in \mathcal{D}_{[m,N-1]}^{\delta})$ be the restrictions of $(x_0 \in \Re^n, \gamma \in \mathcal{D}_{[0,N-1]}, \delta \in \mathcal{D}_{[0,N-1]}^{\delta})$ which generate a trajectory which at time m is $x_m = z$. Define the state $\zeta_m^{\theta}(\cdot) \in C_b(\Re^n)$ by

$$\zeta_m^{\theta}(z) \triangleq \sup_{x \in \ell^2([m+1,N];\Re^n)} E_{w,v} \Big\{ \sum_{k=m}^{N-1} \Big[\lambda(k, x_k, \hat{x}_k) $$
$$- \frac{1}{\theta} \| \sigma^{-1}(k, x_k) \big(x_{k+1} - f_w(k, x_k, w_k) \big) \|_{\Re^n}^2 $$
$$- \frac{1}{\theta} \| \eta^{-1}(k, x_k) \big(y_k - h_v(k, x_k, v_k) \big) \|_{\Re^d}^2 \Big] $$
$$\Big| x_m = z \Big\}. \tag{13}$$

This is the cost-to-go from stage $k = m$ to stage $k = N$, in which $x_m = z$ is optimal.

Theorem 2.2 *Consider a fixed sample path $y(\cdot, \omega) \in \ell^2([0,N]; \Re^d)$, $0 < \theta \le \theta^*$, and $\hat{x} \in \hat{\mathcal{X}}_{[0,N-1]}$ The state $\zeta_m^{\theta}(\cdot) \in C_b(\Re^n)$ satisfies the following recursion.*

$$\zeta_m^{\theta}(z) = T^{\theta}(\hat{x}_m, y_m) \zeta_{m+1}^{\theta}(z), \quad \zeta_N^{\theta}(z) = 0, $$
$$m \in [0, N-1]. \tag{14}$$

Proof. Follows from Dynamic programming.

3 An Information State Stochastic Minimax Game

Next, we dsicuss the invariant property of the pay-off as a functional of the information state and its adjoint, with respect to the suppairing.

Corollary 3.1 *Consider a fixed sample path $y(\cdot, \omega) \in \ell^2([0,N]; \Re^d)$, $0 < \theta \le \theta^*$, and $\hat{x} \in \hat{\mathcal{X}}_{[0,N-1]}$.*
Then the following time-invariant property holds.

$$< \pi_N >_{sup} = < \pi_m, \zeta_m >_{sup} = < \pi_0, \zeta_0 >_{sup}, $$
$$\forall \pi_m \in \mathcal{B}, \zeta_m \in C_b(\Re^n), m \in [0, N]. \tag{15}$$

Proof. Follows from (12),(13).

3.1 Representation of the Pay-Off Functional and Deterministic Optimization

For a fixed observation sample path $y = y(\cdot, \omega) \in \ell^2([0, N-1]; \Re^d)$ and state sample path $x(\cdot, \omega) \in \ell^2([0,N]; \Re^n)$ generated by an admissible $(\gamma, \delta) \in$

$\mathcal{D}_{[0,N-1]}^{\gamma} \times \mathcal{D}_{[0,N-1]}^{\delta}, x_0 \in \Re^n$ and $\hat{x} \in \hat{\mathcal{X}}_{[0,N-1]}$, the extremal of the action functional with respect to $x_{[1,N]} \in \ell^2([1,N]; \Re^n)$ is given by

$$I_{0,N}^{\theta}(\hat{x}, y, x^*) = \sup_{x \in \ell^2([0,m-1];\Re^n)} \sup_{x \in \ell^2([m,m];\Re^n)} $$
$$\sup_{x \in \ell^2([m+1,N];\Re^n)} E_{w,v} \Big\{ p(x_0) + \sum_{k=0}^{N-1} \Big[\lambda(k, x_k, \hat{x}_k) $$
$$- \frac{1}{\theta} \| \sigma^{-1}(k, x_k) \big(x_{k+1} - f_w(k, x_k, w_k) \big) \|_{\Re^n}^2 $$
$$- \frac{1}{\theta} \| \eta^{-1}(k, x_k) \big(y_k - h_v(k, x_k, v_k) \big) \|_{\Re^d}^2 \Big] \Big\}. \tag{16}$$

Using the definitions of $\pi_m(\cdot), \zeta_m(\cdot)$ we deduce

$$I_{0,N}^{\theta}(\hat{x}, y, x^*) = \sup_{x \in \ell^2([m,m];\Re^n)} \Big\{ \pi_m(x_m) + \zeta_m(x_m) \Big\}, $$
$$y \in \ell^2([0,N-1]; \Re^d), \hat{x} \in \hat{\mathcal{X}}_{[0,N-1]}. \tag{17}$$

Consequently, we have the following result.

Theorem 3.1 *For $\hat{x} \in \hat{\mathcal{X}}_{[0,N-1]}$, $\pi_k^{\theta}(\cdot) \in \mathcal{B}$ and $\zeta_k^{\theta}(\cdot) \in C_b(\Re^n)$, $k \in [0, N]$ the pay-off functional has the following representation*

$$J_{0,N}^{\theta}(\hat{x}) = I_{0,N}^{\theta}(\hat{x}, y^*, x^*) \triangleq $$
$$\sup_{y \in \ell^2([0,N-1];\Re^d)} \Big\{ < \pi_N >_{sup} \Big\} = $$
$$\sup_{y \in \ell^2([0,N-1];\Re^d)} \Big\{ < \pi_m, \zeta_m >_{sup} \Big\} = $$
$$\sup_{y \in \ell^2([0,N-1];\Re^d)} \Big\{ < \pi_0, \zeta_0 >_{sup} \Big\}, m \in [0, N] \tag{18}$$

Moreover, the new optimization problem with respect to the admissible states $\hat{x} \in \hat{\mathcal{X}}_{[0,N-1]}$ is defined as follows.

$$J_{0,N}^{\theta}(\hat{x}^*) \triangleq \inf_{\hat{x} \in \hat{\mathcal{X}}} \sup_{y \in \ell^2} \Big\{ < \pi_N >_{sup} \Big\} $$
$$subject \ to \quad \pi_m^{\theta}(z) = T^{\theta}(\hat{x}_{m-1}, y_{m-1}) \pi_{m-1}^{\theta}(z), $$
$$\pi_0^{\theta}(z) = p_0, \ m \in [1, N]. \tag{19}$$

Proof. Follow from definition(2.2) and Corollary follows from theorem (2.1).

Remark 3.1 *Notice that the stochastic optimization problem is reduced to a completely observable optimization problem, because the pay-off functional is expressed in term of the information states π, ζ.*

Definition 3.1 *Let $\hat{\mathcal{X}}_{[k,m]}$ denote the set of state estimators defined on the interval $[k, m]$ which are adapted to the $\sigma-$ algebra $\sigma \Big\{ \pi_j^{\theta}; k \le j \le m \Big\}$.*

1031

For $\hat{x} \in \hat{\mathcal{X}}_{[0,N-1]}$ *(e.g., $\hat{x}_k = \mu(k,\pi_k^\theta), \forall k \in [0, N-1]$)* *the new optimization problem is defined as follows.*

$$J_{0,N}^\theta(\hat{x}^*) \triangleq \inf_{\hat{x} \in \hat{\mathcal{X}}_{[0,N-1]}} \sup_{y \in \ell^2([0,N-1];\Re^d)} \left\{ <\pi_N>_{sup} \right\}$$
$$subject \; to \quad \pi_m^\theta(z) = T^\theta(\hat{x}_{m-1}, y_{m-1})\pi_{m-1}^\theta(z),$$
$$\pi_0^\theta(z) = p_0, \quad m \in [1, N]. \tag{20}$$

Let $\pi_m^\theta = \pi$ be the state obtained from the recursion (12), using the optimal strategy $(\hat{x}^*, y^*)$ during the interval $0 \le k \le m-1$, and for $\pi_N^\theta(\cdot) \in \mathcal{B}, m \in [0, N]$, define

$$W^{\theta,\hat{x}}(\pi,m) \triangleq \sup_{y \in \ell^2([m,N-1];\Re^d)} \left\{ <\pi_N^\theta>_{sup}; \pi_m^\theta = \pi \right\}, \tag{21}$$

The value function associated with (20) is

$$W^\theta(\pi,m) \triangleq \inf_{\hat{x} \in \hat{\mathcal{X}}_{[m,N-1]}} W^{\theta,\hat{x}}(\pi,m) =$$
$$\inf_{\hat{x} \in \hat{\mathcal{X}}_{[m,N-1]}} \sup_{y \in \ell^2([m,N-1];\Re^d)} \left\{ <\pi_N^\theta>_{sup}; \right.$$
$$\left. \pi_m^\theta = \pi \right\}. \tag{22}$$

By Theorem 3.1 we deduce the following theorem.

Theorem 3.2 *(Dynamic programming equation) For each $\pi_m^\theta(\cdot) \in \mathcal{B}, \zeta_m^\theta(\cdot) \in C_b(\Re^n), m \in [0, N]$, the value function $W^\theta(\cdot)$ satisfies the following dynamic programming equation (recursion).*

$$W^\theta(\pi,k) = \inf_{\hat{x} \in \hat{\mathcal{X}}_{[m,m]}} \sup_{y \in \ell^2([m,m];\Re^d)}$$
$$\left\{ W^\theta(T^{\star,\theta}(\hat{x},y)\pi, k+1) \right\}, \quad k \in [0, N-1],$$
$$W^\theta(\pi, N) = <\pi>_{sup}. \tag{23}$$

Proof. Similar to [20].

4 Explicit Solution for Linear Systems

In this section, we consider linear systems and a quadratic pay-off functional to derive solutions to the information state equation, and then to solve the resulting minimax game to deduce in the optimal estimator.
The system dynamics are

$$x_{k+1} = A_k x_k + \sigma_k \gamma_k + D_k w_k \tag{24}$$

$$y_k = C_k x_k + \eta_k \delta_k + N_k v_k \tag{25}$$

and the pay-off functional is:

$$J_{0,N}^\theta = E^\nu \{ \frac{1}{2}(x_0 - \bar{x}_0)\bar{P}_0^{-1}(x_0 - \bar{x}_0)'$$
$$+ \sum_{k=0}^{N-1} [\frac{1}{2}(x_k - \hat{x}_k)'Q_k(x_k - \hat{x}_k)$$
$$- \frac{1}{2\theta}||\gamma_k||^2 - \frac{1}{2\theta}||\delta_k||^2] \}, \tag{26}$$

Where $(\bar{P}_0^{-1})' = (\bar{P}_0^{-1}) \ge 0, Q_k = Q_k' > 0, \forall k \in [0, N-1], w \in N(0,\Sigma_w), v \in N(0,\Sigma_v), \Sigma_w \ge 0, \Sigma_v > 0$
The information state solution is:

$$\pi_k^\theta = -\frac{1}{2}x_k'P_k x_k + x_k'\Gamma_k + \frac{1}{2}\delta_k \tag{27}$$

Where P, Γ, δ are given by some recursion equations. The Pay-off is now given by $J_{0,N}^\theta = \sup_{\hat{x}_N} \pi_N^\theta(x_N)$ Thus

$$x_N^* = arg \sup_{\hat{x}_N \in \Re^n} J_{0,N}^\theta(\hat{x}) = P_N^{-1}\Gamma_N. \tag{28}$$

Then the completely observable problem is linear quadratic. Thus, we can use Dynamic programming to obtain the explicit expression for the optimal estimator.

References

[1] G. Zames, "Feedback and Optimal Sensitivity: Model Reference Transformations, Multiplicative Semi norms, and Approximate Inverses", *IEEE Transactions on Automatic Control* Vol. 26, No. 2, pp.301-320, 1981.

[2] E.N. Barron and R. Jensen, " Total Risk-Aversion, Stochastic Optimal Control, and Differential Games", *Applied Mathematics and Optimization*, Vol. 19, pp. 313-327, 1989.

[3] M.R. James, "Asymptotic Analysis of Nonlinear Stochastic Risk-Sensitive Control and Differential Games", *Mathematics of Control, Signals, and Systems*, Vol. 5, No. 3, pp. 401-417, 1992.

[4] J. Willems, "Dissipative dynamical systems part I: General theory," *Arch Rational Mech. Anal.*, Vol. 45, pp. 321-351, 1972.

[5] A. Isidori and A. Astofi, "Disturbance attenuation and H^∞-control via measurement feedback in nonlinear systems," *IEEE Transaction on Automatic Control, Vol. 37. No. 9, pp. 1283-1293 , 1992.*

[6] P. Whittle, "A risk-sensitive maximum principle," Systems and Control Letters, Vol. 15, pp. 183-192, 1990.

[7] W. H. Fleming and W. M. McEneaney, "Risk-sensitive control and differential games, pp. 185-197," in Stochastic Theory and Adaptive Control, eds., T.E. Duncan and B. Pasik-duncan,(New York), Springer-Verag 1992.

[8] T. Runolfsson, "The equivalence between infinite-horizon optimal control of stochastic system with exponential of integral performance index and stochastic differential games," *IEEE Transaction on Automatic Control, Vol. 39, No. 4, pp. 1551-1563, 1994.*

[9] D. Jacobson, *"Optimal stochastic linear system with exponential performance criteria and their relation to deterministic differential games," IEEE Transaction on Automatic Control,* Vol. 18, No. 2, pp. 124-131, 1973.

[10] D. Jacobson, *Extensions of Linear-Quadratic Control, optimization and Matrix Theory.* London, NY, San Francisco: academic Press Inc., Mathematics in Science and Engineering, Vol. 133, 1997.

[11] J. Speyer, "An adaptive terminal guidance scheme based on an exponential cost criterion with application to homing missile guidance," *IEEE Transactions on Automatic Control, Vol. 21, pp. 371-375, 1976.*

[12] P. Whittle, *"Risk-sensitive Linear/Quadratic/Gaussian control", Advances in Applied Probability,* Vol. 13, pp. 764-777, 1981.

[13] P. Kumar and J. van Schuppen, "On the optimal control of stochastic systems with an exponential-of-integral performance index, " *Journal of mathematical analysis and applications,* Vol. 80, pp. 312-332, 1981.

[14] A Bensoussan and J. H. van Schuppen, "Optimal control of partially observable stochastic systems with an exponential-of-integral performance index," *SIAM Journal on Control and Optimization, Vol. 23, No. 4, pp. 599-613, 1985.*

[15] *K. Glover and J. Doyle, "State-space formulae for all stabilizing controllers that satisfy an H^∞-norm bound and relations to risk-sensitive," System and Control Letters,* Vol. 11, pp. 167-172, 1988.

[16] C. Charalambous and J. Hibey, "Minimum Principle for partially observable nonlinear risk-sensitive control problems using measure-valued decompositions," *Stochastic and Stochastic Reports,* Vol. 57, 1996.

[17] C. Charalambous, "The role of information state and adjoint in relating nonlinear output feedback risk-sensitive control and dynamic games," *IEEE Transaction on Automatic Control, Vol. 42, No. 8, pp. 1163-1170, 1997.*

[18] *P. D. Pra, L. Meneghini, and W. Runggaldier, "Some connections between stochastic control and dynamic games." Preprint, 1997.*

[19] *P. Whittle, "A Risk-Sensitive Maximum Principle: The case of Imperfect State Observations", IEEE Transactions on Automatic Control,* Vol. 36, No. 7, pp. 793-801, 1991.

[20] M.R. James and J.S. Baras and R.J. Elliott, "Risk-Sensitive Control and Dynamic Games for Partially Observed Discrete-Time Nonlinear Systems", *IEEE Transaction on Automatic Control,* Vol. 39, No. 4, pp. 780-792, 1994.