Optimal Stochastic Control of Discrete-Time Systems Subject to Total Variation Distance Uncertainty

Charalambos D. Charalambous, Farzad Rezaei and Ioannis Tzortzis

***Abstract*— This paper presents another application of the results in [1], [2], where existence of the maximizing measure over the total variation distance constraint is established, while the maximizing pay-off is shown to be equivalent to an optimization of a pay-off which is a linear combination of L_1 and L_∞ norms. Here emphasis is geared towards to uncertain discrete-time controlled stochastic dynamical system, in which the control seeks to minimize the pay-off while the measure seeks to maximize it over a class of measures described by a ball with respect to the total variation distance centered at a nominal measure. Two types of uncertain classes are considered; an uncertainty on the joint distribution, an uncertainty on the conditional distribution. The solution of the minimax problem is investigated via dynamic programming.**

I. INTRODUCTION

The objective of this paper is to employ the definition of uncertainty via the total variational distance introduced in [1], [2] and the results therein to

- investigate stochastic optimal control of systems governed by discrete-time dynamical systems subject to total variational distance uncertainty.

The theory developed in [1], [2] is applied to formulate and solve minimax stochastic uncertain controlled systems described by discrete-time nonlinear stochastic controlled systems, in which the control seeks to minimize the pay-off while the measure seeks to maximize it over the total variational distance constraint. The main objective is to characterize the solution of the minimax game via dynamic programming. It turns out that once the maximizing measure is found and substituted into the pay-off the equivalent optimization problem to be solved is a stochastic optimal control problem. There is however, a fundamental difference from the classical pay-off of stochastic control problems treated in the literature in that the pay-off is a non-linear functional of measure induced by the stochastic system, contrary to the classical which is a linear functional.

The rest of the paper is organized as follows. In Section II various models of uncertainty are introduced and their relation to total variation distance is described. In Sections III the abstract formulation is introduced while in Section III-A the solution of the maximization problem over the total variation norm constraint set is presented. In section IV, the abstract setup is applied to stochastic discrete-time uncertain controlled systems. A dynamic programming equation is derived to characterize the optimality of minimax strategies.

II. MOTIVATION

Below, the total variation distance uncertainty class is described, while its relation to other uncertainty classes is explained.

Let (Σ, d_Σ) denote a complete separable metric space (a Polish space), and $(\Sigma, \mathcal{B}(\Sigma))$ the corresponding measurable space, in which $\mathcal{B}(\Sigma)$ is the σ-algebra generated by open sets in Σ. Let $\mathcal{M}_1(\Sigma)$ denote space of countably additive probability measures on $(\Sigma, \mathcal{B}(\Sigma))$.

Total Variational Distance Uncertainty. Given a known or nominal probability measure $\mathbb{P} \in \mathcal{M}_1(\Sigma)$ the uncertainty set based on total variation distance is defined by

$$B_R(\mathbb{P}) \triangleq \Big\{\mathbb{Q} \in \mathcal{M}_1(\Sigma) : ||\mathbb{Q} - \mathbb{P}||_{var} \leq R\Big\}$$

where $R \in [0, \infty)$. The total variational distance[1] on $\mathcal{M}_1(\Sigma) \times \mathcal{M}_1(\Sigma)$ is defined by

$$||\mathbb{Q} - \mathbb{P}||_{var} \triangleq \sup_{P \in \Pi(\Sigma)} \sum_{F_i \in P} |\mathbb{Q}(F_i) - \mathbb{P}(F_i)|, \mathbb{Q}, \mathbb{P} \in \mathcal{M}_1(\Sigma)$$

where $\Pi(\Sigma)$ denotes the collection of all finite partitions of Σ. Note that the distance metric induced by the total variation norm does not require absolute continuity of measures when defining the uncertainty ball (i.e., singular measures are admissible), that is, the measures need not be defined on the same space. It can very well be the case that $\tilde{\mathbb{P}} \in \mathcal{M}_1(\tilde{\Sigma}), \tilde{\Sigma} \subseteq \Sigma$ and $\mathbb{P} \in \mathcal{M}_1(\Sigma)$ is the extension of $\tilde{\mathbb{P}}$ on Σ. Additionally, since $\mathcal{M}_1(\Sigma)$ are probability measures then it follows that the radius of uncertainty belongs to the restricted set $R \in [0, 2]$.

Recall that the relative entropy of $\mathbb{Q} \in \mathcal{M}_1(\Sigma)$ with respect to $\mathbb{P} \in \mathcal{M}_1(\Sigma)$ is a mapping $\mathbb{D}(\cdot||\cdot) : \mathcal{M}_1(\Sigma) \times \mathcal{M}_1(\Sigma) \to [0, \infty]$ defined by

$$\mathbb{D}(\mathbb{Q}||\mathbb{P}) \triangleq \begin{cases} \int_\Sigma \log(\frac{d\mathbb{Q}}{d\mathbb{P}}) d\mathbb{Q}, & \text{if } \mathbb{Q} << \mathbb{P} \\ +\infty, & \text{otherwise} \end{cases}$$

Here $\mathbb{Q} << \mathbb{P}$ is the notation often used to denote that measure $\mathbb{Q} \in \mathcal{M}_1(\Sigma)$ is absolutely continuous with respect to measure $\mathbb{P} \in \mathcal{M}_1(\Sigma)$.[2]

This work was supported by the Cyprus Research Promotion Foundation project ARTEMIS-0506/20 and a University of Cyprus Research grant.

F. Rezaei is with School of Information Technology and Engineering, University of Ottawa, 800 King Edward Ave., Ottawa, Canada. E-mail:frezaei@site.uottawa.ca

C.D. Charalambous is with the Department of Electrical Engineering, University of Cyprus, Nicosia, Cyprus. E-mail:chadcha@ucy.ac.cy

I. Tzortzis is with the Department of Electrical Engineering, University of Cyprus, Nicosia, Cyprus. E-mail:tzortzis.ioannis@ucy.ac.cy

[1]The definition of total variational distance applies to signed measures as well.

[2]If $\mathbb{P}(A) = 0$ for some $A \in \mathcal{B}(\Sigma)$ then $\mathbb{Q}(A) = 0$.

Relative Entropy Uncertainty. Given a known or nominal probability measure $\mathbb{P} \in \mathcal{M}_1(\Sigma)$ the uncertainty set based on relative entropy is defined by

$$A_{R_1}(\mathbb{P}) \triangleq \Big\{\mathbb{Q} \in \mathcal{M}_1(\Sigma) : \mathbb{D}(\mathbb{Q}|\mathbb{P}) \leq R_1\Big\}$$

where $R_1 \in [0, \infty)$ (when $R_1 = 0$ then $\mathbb{Q} = \mathbb{P}, \mathbb{Q} - a.s$).
Relative entropy uncertainty model has connections to risk sensitive pay-off, minimax games, and large deviations [4], [5], [6], [7], [8], [9], [10].
Unfortunately, relative entropy uncertainty modeling has two disadvantages. 1) it does not define a true metric on the space of measures; 2) relative entropy between two measures takes the value infinity when the measures are not absolutely continuous. The latter rules out the possibility of measures $\mathbb{Q} \in \mathcal{M}_1(\Sigma)$ and $\mathbb{P} \in \mathcal{M}_1(\Sigma)$ to be initially defined on different spaces (e.g., one being defined on a higher dimension space than the other measure). Clearly, the total variation distance uncertainty set is larger than the relative entropy uncertainty set. This can be concluded from Pinsker's inequality [3] as well.

$$||\mathbb{Q} - \mathbb{P}||_{var}^2 \leq 2\mathbb{D}(\mathbb{Q}||\mathbb{P}), \quad \mathbb{Q}, \mathbb{P} \in \mathcal{M}_1(\Sigma), \quad \text{if} \quad \mathbb{Q} << \mathbb{P}$$

Hence, even for those measures $\mathbb{Q} << \mathbb{P}$ the uncertainty set described by relative entropy is a subset of the much larger total variation distance uncertainty set, that is, $A_{\frac{R^2}{2}}(\mathbb{P}) \subset B_R(\mathbb{P})$.
L_1 *Distance Uncertainty.* Suppose measures $\mathbb{Q}, \mathbb{P} \in \mathcal{M}_1(\Sigma)$ are absolutely continuous with respect to a fixed measure $\mathbb{P}_o \in \mathcal{M}_1(\Sigma)$ (e.g., $\mathbb{P} << \mathbb{P}_o, \mathbb{Q} << \mathbb{P}_o$)[3]. Under these conditions it can be shown that total variation distance reduces to $L_1(\mathbb{P}_o)$ distance as follows.

$$\begin{aligned} C_R(\mathbb{P}) &\triangleq \Big\{\mathbb{Q} \in \mathcal{M}_1(\Sigma) : ||\mathbb{Q} - \mathbb{P}||_{var} \leq R\Big\} \\ &= \Big\{\varphi \in L_1(\mathbb{P}_o) : \int_\Sigma |\varphi(x) - \psi(x)| \mathbb{P}_o(dx) \leq R\Big\} \\ &\equiv C_R(\varphi) \end{aligned}$$

where existence of $\varphi = \frac{d\mathbb{Q}}{d\mathbb{P}_o} \in L_1(\mathbb{P}_o), \psi = \frac{d\mathbb{P}}{d\mathbb{P}_o} \in L_1(\mathbb{P}_o)$ follows from the absolute continuity $\mathbb{P} << \mathbb{P}_o, \mathbb{Q} << \mathbb{P}_o$. Hence, the $L_1(\mathbb{P}_o)$ distance set $C_R(\varphi)$ is a smaller set that the total variation distance set $B_R(\mathbb{P})$. Robustness via L_1 distance uncertainty on the space of spectral densities is investigated in the context of Wiener -Kolmogorov theory in an estimation and decision framework in [18], [19].

III. Abstract Formulation

The material of this section are derived in [2]. Let (Σ, d_Σ) denote a complete separable metric space (a Polish space), and $(\Sigma, \mathcal{B}(\Sigma))$ the corresponding measurable space, in which $\mathcal{B}(\Sigma)$ is the σ-algebra generated by open sets in Σ. Let $\mathcal{X} \triangleq BC(\Sigma)$ denote the Banach space of bounded continuous functions on Σ, equipped with the sup-norm. It is known that the dual space [17] $\mathcal{X}^*$ is isometrically isomorphic to $\mathcal{M}_{rba}(\Sigma)$, the Banach space of finitely additive finite signed regular measures on $(\Sigma, \mathcal{B}(\Sigma))$.

[3]They can be generalized to spectral measures as well.

At the abstract level, systems are represented by measures in $\mathcal{M}_1(\Sigma)$ induced by the underlying random processes, which are defined on an appropriate Polish space. Similarly, controls denoted by u, are defined on a subset $\mathcal{U}_0$ of an appropriate Polish space $(\mathcal{U}, d_\mathcal{U})$, while we choose a suitable subset $\mathcal{U}_{ad} \subset \mathcal{U}_0$ for the class of admissible controls. The pay-off is represented by a linear functional on the space of probability measures $\mathcal{M}_1(\Sigma)$.

Nominal System. The nominal system is defined as follows. By choosing a control policy $u \in \mathcal{U}_{ad}$ for the nominal system (which is perfectly known), then the nominal system induces a nominal probability measure $\mathbb{P}^u \in \mathcal{M}_1(\Sigma)$.

Uncertain System. For a given $u \in \mathcal{U}_{ad}$, let $M(u) \subset \mathcal{M}_1(\Sigma)$ denote the set of probability measures induced by the perturbed system while control $u \in \mathcal{U}_{ad}$ is applied. The perturbed system or uncertain system $\mathbb{Q}^u \in M(u)$ is further restricted to the following constraint described by the variational norm.

$$B_R(\mathbb{P}^u) = \{\mathbb{Q}^u \in M(u) : ||\mathbb{Q}^u - \mathbb{P}^u||_{var} \leq R\}$$

Mini-Max Optimization. Let $\ell^u : \Sigma \to \Re$ be a real-valued bounded non-negative measurable function. The uncertainty tries to maximize the average pay-off functional denoted by $\int_\Sigma \ell^u(x) d\mathbb{Q}^u(x)$ over the set $B_R(\mathbb{P}^u)$ for a given $u \in \mathcal{U}_{ad}$. The effect of uncertainty leads to the following maximization problem:

$$\sup_{\mathbb{Q}^u \in B_R(\mathbb{P}^u)} \int_\Sigma \ell^u(x) d\mathbb{Q}^u(x)$$

for every control $u \in \mathcal{U}_{ad}$.
The designer on the other hand, tries to choose a control policy to minimize the worst case average pay-off. This gives rise to the min-max problem

$$\inf_{u \in \mathcal{U}_{ad}} \sup_{\mathbb{Q}^u \in B_R(\mathbb{P}^u)} \int_\Sigma \ell^u(x) d\mathbb{Q}^u(x) \tag{III.1}$$

As a first step, we present the existence of a $\mathbb{Q}^{u,*} \in B_R(\mathbb{P}^u)$ at which the supremum in (III.1) is attained. Subsequently, we present an explicit characterization of this measure.

A. Characterization of the Maximizing Measure

In this section we drop the dependence on the control u of the various measures and functions. Suppose ℓ is a non negative element in $BC(\Sigma)$. Clearly, $\mathcal{M}_1(\Sigma) \subset \mathcal{M}_{rba}(\Sigma)$. Let $\mathbb{P} \in \mathcal{M}_1(\Sigma)$ be a given probability measure referred to as the nominal measure. Define the uncertainty set by

$$B_R(\mathbb{P}) \triangleq \Big\{\mathbb{Q} \in \mathcal{M}_1(\Sigma) : ||\mathbb{Q} - \mathbb{P}||_{var} \leq R\Big\}$$

The objective is to find the worst case (supremum) of average pay-off over the uncertain set $B_R(\mathbb{P})$. The average pay-off is defined as a linear functional acting on $\ell \in BC(\Sigma)$, i.e., $\int_\Sigma \ell(x) \mathbb{Q}(dx)$, where $\mathbb{Q} \in B_R(\mathbb{P})$. Hence the problem is the following

$$J_\ell(\mathbb{Q}^*) = \sup_{\mathbb{Q} \in B_R(\mathbb{P})} \int_\Sigma \ell(x) \mathbb{Q}(dx), \qquad \mathbb{P} \in \mathcal{M}_1(\Sigma) \tag{III.2}$$

The set $B_R(\mathbb{P})$ is weak*-compact, while the pay-off is weak* continuous. Hence, there exists a maximizing measure in $B_R(\mathbb{P})$. The optimization in $(III.2)$ is solved by appealing to the Hahn-Banach theorem [14]. Since $\ell \in BC(\Sigma)$ is fixed, then there exists $\mathbb{V} \in (BC(\Sigma))^* \simeq \mathcal{M}_{rba}(\Sigma)$ such that

$$\mathbb{V}(\ell) \triangleq \int_\Sigma \ell d\mathbb{V} = ||\ell||_\infty, \quad \text{with} \quad ||\mathbb{V}||_{var} = 1 \tag{III.3}$$

Define $\mathbb{V} \triangleq \mathbb{Q} - \mathbb{P} \in \mathcal{M}_{rba}(\Sigma)$. Then from (III.2) and the above result we have the following.

$$\begin{aligned} &\sup_{\mathbb{Q}\in B_R(\mathbb{P})} \int_\Sigma \ell(x)\mathbb{Q}(dx) \\ &= \sup_{\mathbb{V}\in B_R(\mathcal{M}_{rba}(\Sigma))} \int_\Sigma \ell(x)\mathbb{V}(dx) + E_\mathbb{P}(\ell) \\ &\leq R||\ell||_\infty + E_\mathbb{P}(\ell) \end{aligned} \tag{III.4}$$

where $B_R(\mathcal{M}_{rba}(\Sigma)) = \{\mathbb{V} \in \mathcal{M}_{rba}(\Sigma) : ||\mathbb{V}||_{var} \leq R\}$. The supremum on the right hand side of (III.4) is attained by a signed measure $\mathbb{V}^* \in \mathcal{M}_{rba}(\Sigma)$, having the property $||\mathbb{V}^*||_{var} = R$. Clearly, if $\mathbb{Q}^* = \mathbb{V}^* + \mathbb{P}$ is a probability measure, then the upperbound in (III.4) is attained by $\mathbb{Q}^* \in \mathcal{M}_1(\Sigma)$. Therefore, it remains to establish that $\mathbb{Q}^*$ is a probability measure and $\mathbb{Q}^* \in B_R(\mathbb{P})$.
It can be shown that $\mathbb{Q}^*$ is non-negative.

Lemma 3.1: [2] Suppose $\ell \in BC(\Sigma)$ is non-negative. The maximizing measure $\mathbb{Q}^* = \mathbb{V}^* + \mathbb{P}$, where $\mathbb{V}^* \in \mathcal{M}_{rba}(\Sigma), \mathbb{P} \in \mathcal{M}_{rba}(\Sigma)$, $||\mathbb{V}^*||_{var} = R$ in (III.4) is a non-negative measure.

The maximizing measure $\mathbb{Q}^*$ is not unique. The next lemma is crucial in the characterization of the class of maximizing measures.

Lemma 3.2: [2]. Suppose $\ell : \Sigma \to \Re$ is a bounded non-negative measurable function, and $\mathbb{E}$ is a finitely additive non-negative finite measure defined on $(\Sigma, \mathcal{B}(\Sigma))$.
Assume
i) The total weight of the measure $\mathbb{E}$ is not concentrated on any bounded measurable subset of Σ;
ii) The support of $\mathbb{E}$ contains the point at which ℓ attains its maximum.
Then

$$\sup_{s>0} \frac{\int_\Sigma \ell(x)e^{s\ell(x)}\mathbb{E}(dx)}{\int_\Sigma e^{s\ell(x)}\mathbb{E}(dx)} = ||\ell||_\infty$$

Next, we state the main theorem which characterizes the maximizing measure as a convex combination of two probability measures.

Theorem 3.3: [2] Under the assumptions of Lemma 3.2, there exists a family of probability measures which attain the supremum in (III.2) given by

$$\mathbb{Q}^*(E) = \frac{\beta}{\beta+1} \frac{\int_E e^{s_0\ell(x)}\mathbb{E}(dx)}{\int_\Sigma e^{s_0\ell(x)}\mathbb{E}(dx)} + \frac{1}{1+\beta}\mathbb{P}(E)$$

where $E \in \mathcal{B}(\Sigma), \beta \in (2, \infty)$ is arbitrary, and $\mathbb{E}$ is an arbitrary finite non-negative finitely additive measure defined on $(\Sigma, \mathcal{B}(\Sigma))$. Moreover, β and $\mathbb{E}$ are chosen such that $||\mathbb{Q}^* - \mathbb{P}|| = R$.
Clearly, $\mathbb{Q}^*(dx)$ is a convex combination of the tilted measure $\frac{e^{s_0\ell(x)}\mathbb{E}(dx)}{\int_\Sigma e^{s_0\ell(x)}\mathbb{E}(dx)}$ and $\mathbb{P}(dx)$. Moreover, the initial optimization problem is also a convex combination of L_1 and L_∞ optimization problems in view of

$$\begin{aligned} \sup_{\mathbb{Q}\in B_R(\mathbb{P})} \int_\Sigma \ell(x)\mathbb{Q}(dx) &= R.||\ell||_\infty + E_\mathbb{P}(\ell) \\ &= (1+\beta)\int_\Sigma \ell(x)\mathbb{Q}^*(dx) \end{aligned}$$

where $\frac{\int_\Sigma \ell(x)e^{s_0\ell(x)}\mathbb{E}(dx)}{\int_\Sigma e^{s_0\ell(x)}\mathbb{E}(dx)} = \frac{R}{\beta}||\ell||_\infty$. The rest of the paper deals with the application of the above results to discrete-time controlled stochastic systems.

IV. Fully Observed Uncertain Control Systems

Define $\mathbb{N}_+ \triangleq \{0,1,2,3,\ldots\}$, $\mathbb{N}_+^n \triangleq \{0,1,2,\ldots,n\}, n \in \mathbb{N}_+$. All processes are defined on the probability space $(\Omega, \mathbb{F}, \mathbb{Q})$ with filtration $\{\mathbb{F}_{0,j}\}_{j=0}^n, n \in \mathbb{N}_+$. Let $\mathcal{F}_{0,j} \subset \mathbb{F}_{0,j}, j = 0,1,\ldots,n$ be a sub-sigma field. The state space, the control space, and the noise spaces are sequences of Polish spaces $\{\mathcal{X}_j : j = 0,1,\ldots,n\}$, $\{\mathcal{U}_j : j = 0,1,\ldots,n-1\}$, $\{\mathcal{W}_j : j = 1,2,\ldots,n-1\}$, respectively. These spaces are associated with their corresponding measurable spaces $(\mathcal{X}_j, \mathcal{B}(\mathcal{X}_j)), j \in \mathbb{N}_+^n$, $(\mathcal{U}_j, \mathcal{B}(\mathcal{U}_j))$, $(\mathcal{W}_j, \mathcal{B}(\mathcal{W}_j))$, $j \in \mathbb{N}_+^{n-1}$. Thus, sequences of state spaces, control spaces, noise spaces are identified with their product spaces $(\mathcal{X}_{0,n}, \mathcal{B}(\mathcal{X}_{0,n})) \triangleq \times_{j=0}^n (\mathcal{X}_j, \mathcal{B}(\mathcal{X}_j))$, $(\mathcal{U}_{0,n-1}, \mathcal{B}(\mathcal{U}_{0,n-1})) \triangleq \times_{j=0}^{n-1}(\mathcal{U}_j, \mathcal{B}(\mathcal{U}_j))$, $(\mathcal{W}_{0,n-1}, \mathcal{B}(\mathcal{W}_{0,n-1})) \triangleq \times_{j=0}^{n-1}(\mathcal{W}_j, \mathcal{B}(\mathcal{W}_j))$, respectively, $n \in \mathbb{N}^n$. The state process is denoted by $x \triangleq \{x_j : j = 0,1,\ldots,n\}$, $x : \mathbb{N}_+^n \times \Omega \longmapsto \mathcal{X}_j$, the control process is denoted by $u \triangleq \{u_j : j = 0,1,\ldots,n-1\}$, $u : \mathbb{N}_+^{n-1} \times \Omega \longmapsto \mathcal{U}_j$, and noise process is denoted by $w \triangleq \{w_j : 0 = 1,2,\ldots,n-1\}$, $w : \mathbb{N}_+^{n-1} \times \Omega \longmapsto \mathcal{W}_j$.
Denote by $\tilde{\mathcal{U}}_{ad}[0,n-1]$ the set of the $\mathcal{U}_{0,n-1}-$valued control processes u such that u_j is $\mathcal{F}_{0,j}-$measurable, $j \in \mathbb{N}_+^{n-1}$. Note that state constrained controls may be included in the formulation by further assuming that u_j take values in a nonempty subset $\mathcal{U}_j(x_j) \subset \mathcal{U}_j, \forall x_j \in \mathcal{X}_j, j = 0,1,\ldots,n-1$.

Define two additional classes of admissible controls as follows. $\mathcal{U}_{ad}[0,n-1] \subseteq \tilde{\mathcal{U}}_{ad}[0,n-1]$ denoting those controls u_j which are $\mathcal{G}_{0,j} \triangleq \sigma\{x_0^u,\ldots,x_j^u,u_0,\ldots,u_{j-1}\}-$measurable. These are called feedback control strategies. $\mathcal{U}_{ad}^w[0,n-1] \subseteq \tilde{\mathcal{U}}_{ad}[0,n-1]$ denoting those controls u_j which are $\mathcal{F}_{0,j}^w \triangleq \sigma\{w_0,\ldots,\ldots,w_{j-1}\}-$measurable.
Conditional distributions are represented by stochastic kernels defined below.

Definition 4.1: Given the measurable spaces $(\mathcal{X}, \mathcal{B}(\mathcal{X})), (\mathcal{Y}, \mathcal{B}(\mathcal{Y}))$, a stochastic Kernels on $(\mathcal{Y}, \mathcal{B}(\mathcal{Y}))$ conditioned on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ is a mapping $P : \mathcal{B}(\mathcal{Y}) \times \mathcal{X} \to [0,1]$ satisfying the following two

properties:
1) For every $x \in \mathcal{X}$, the set function $P(\cdot;x)$ is a probability measure (possibly finitely additive) on $\mathcal{B}(\mathcal{Y})$;
2) for every $A \in \mathcal{B}(\mathcal{Y})$, the function $P(A;\cdot)$ is $\mathcal{B}(\mathcal{X})$-measurable.
The set of all stochastic Kernels $(\mathcal{Y}, \mathcal{B}(\mathcal{Y}))$ conditioned on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ are denoted by $\mathcal{Q}(\mathcal{Y};\mathcal{X})$.

A. Problem Formulation

Below, the stochastic dynamics, pay-off, assumptions and uncertain system definitions are introduced.

Stochastic Dynamical System.

For each $u \in \tilde{\mathcal{U}}_{ad}[0,n-1]$ the nominal state process giving rise to a nominal measure is described by the following discrete-time difference equation.

Definition 4.2: (Nominal System). A nominal system family of state processes $\{x^u = x_0^u, x_1^u, \ldots, x_n^u : u \in \tilde{\mathcal{U}}_{ad}[0,n-1]\}$ corresponds to a sequence of stochastic kernels $\{P_{w_j}(dw;x,u) : j = 0,1,\ldots,n-1\}$, and functions $\{b_j : \mathcal{X}_j \times \mathcal{U}_j \times \mathcal{W}_j \longmapsto \mathcal{X}_{j+1} : j = 0,1,\ldots,n-1\}$ if for all $u \in \tilde{\mathcal{U}}_{ad}[0,n-1]$, there exists noise processes $\{w_j : j = 0,2,\ldots,n-1\}$ such that the following hold.

1) For each $j \in \mathbb{N}_+^{n-1}$, w_j is $\mathcal{F}_{0,j}-$ measurable and $\{x_0^u, x_1^u, \ldots, x_n^u\}$ are generated by by the recursion

$$x_{j+1}^u = b_j(x_j^u, u_j, w_j),\ x_0^u = x_0 \tag{IV.5}$$

which implies that if x_0 is $\mathcal{F}_{0,0}-$measurable then x_j^u is $\mathcal{F}_{0,j-1}-$measurable.

2) For every $A \in \mathcal{B}(\mathcal{W}_j), j \in \mathbb{N}_+^{n-1}$

$$Prob(w_j \in A|\mathcal{G}_{0,j}) = P_{w_j}(A; x_j^u, u_j), a.s.$$

3) $Prob(x_0^u = x_0) = 1, \forall u \in \tilde{\mathcal{U}}_{ad}[0,n-1]$.

Uncertain Models.

Two types of uncertainty models are described and analyzed. Type 1: uncertainty on the joint distribution of the noise sequence $\mathbb{Q}_w(dw_1,...,dw_{n-1}) \in \mathcal{M}_1(\mathcal{W}_{0,n-1})$. Type 2: uncertainty on the conditional distribution $\mathbb{Q}_{w_j|\mathcal{G}_{0,j}}(dw_j|\mathcal{G}_{0,j}) \in \mathcal{M}_1(\mathcal{W}_j)$, $0 \leq j \leq n-1$.

Definition 4.3: Given a nominal system of Definition 4.2, a fixed nominal joint measure $\mathbb{P}_w \in \mathcal{M}_1(\mathcal{W}_{0,n-1})$ and $R \in [0,2]$, the class of measures is defined by

$$\mathcal{A}_R(\mathbb{P}_w) \triangleq \Big\{\mathbb{Q}_w \in \mathcal{M}_1(\mathcal{W}_{0,n-1}) : ||\mathbb{Q}_w - \mathbb{P}_w||_{var} \leq R\Big\}$$

Definition 4.4: Given a nominal system of Definition 4.2 a fixed nominal stochastic kernel $P_{w_i}(dw_i; x_i^u, u_i) \in \mathcal{Q}(\mathcal{W}_i; \mathcal{X}_i \times \mathcal{U}_i)$, and $R_i \in [0,2]$, the class of measures is defined by

$$\mathcal{B}_{R_i}(P_{w_i})(\mathcal{G}_{0,i}) \triangleq \Big\{\mathbb{Q}_{w_i}(\cdot|\mathcal{G}_{0,i}) \in \mathcal{M}_1(\mathcal{W}_{0,i-1}) : ||\mathbb{Q}_{w_i}(\cdot;\mathcal{G}_{0,i}) - P_{w_i}(\cdot; x_i^u, u_i)||_{var} \leq R_i\Big\}$$

for $i = 0,1,\ldots,n-1$.

Pay-Off Functional.

The sample pay-off is functional of x^u, u, w, and for each $u \in \tilde{\mathcal{U}}_{ad}[0,n-1]$ the average pay-off is defined by

$$J_{0,n}(u,\mathbb{Q}) \triangleq E_{\mathbb{Q}}\Big\{\sum_{j=0}^{n-1} f_j(x_j^u, u_j, w_j) + h_n(x_n^u)\Big\} \tag{IV.6}$$

where $E_{\mathbb{Q}}\Big(\cdot\Big)$ denotes expectation with respect to the true joint measure $\mathbb{Q} \in \mathcal{M}_1(\mathcal{X}_{0,n} \times \mathcal{U}_{0,n-1} \times \mathcal{W}_{0,n-1})$
The following assumptions are introduced.

Assumptions 4.5: The nominal system family satisfies the following assumptions:
1) $(\mathcal{U}_{0,n-1}, d)$ is Polish space. The control $\{u_j : j \in \mathbb{N}_+^{n-1}\}$ is non anticipative.
2) The maps $\{b_j : \mathcal{X}_j \times \mathcal{U}_j \times \mathcal{W}_j \longmapsto \mathcal{X}_{j+1} : j = 0,1,\ldots,n-1\}$ are bounded continuous, and the maps $\{f_j : \mathcal{X}_j \times \mathcal{U}_j \times \mathcal{W}_j \longmapsto \mathbb{R} : j = 0,1,\ldots,n-1\}$, $f_n : \mathcal{X}_n \longmapsto \mathbb{R}$ are bounded continuous and non-negative.
Notice that for $u \in \tilde{\mathcal{U}}_{ad}[0,n-1]$ the nominal system state x_j^u is a measurable function of $\{w_k : k = 0,1,\ldots,j-1\}$ and $\{u_k : k = 0,1,\ldots,j-1\}$, and hence x_j^u is $\mathcal{F}_{0,j-1}-$measurable for $j \in \mathbb{N}_+^n$.
For $u \in \mathcal{U}_{ad}[0,n-1]$ the nominal system state x_j^u is a measurable function of $\{w_k : k = 0,1,\ldots,j-1\}$ and $\{u_k : k = 0,1,\ldots,j-1\}$.
For $\mathcal{U}_{ad}^w[0,n-1]$ the nominal system state and control (x_j^u, u_j) are measurable functions of the noise sample path $(w_0, w_1, \ldots, w_j - 1)$.

B. Maximization Stochastic Control: Maximization over Class of Measures and Dynamic Programming

In Section III it is described at the abstract level, how to construct the maximizing measure of a linear functional over a total variation distance constraint. Similar arguments can be carried out for the case of discrete time stochastic controlled systems, to deal with uncertainty of the measure $\mathbb{Q}_w \in \mathcal{M}_1(\mathcal{W}_{0,n-1})$ which under certain assumption on the nature of the measures induced by the true system, defines the average pay-off IV.6. Two types of uncertainty models are described and analyzed, Type 1: uncertainty on the joint distribution of the noise sequence $\mathbb{Q}_w(dw_1,...,dw_{n-1}) \in \mathcal{M}_1(\mathcal{W}_{0,n-1})$. Type 2: uncertainty on the conditional distribution $\mathbb{Q}_{w_j|\mathcal{G}_{0,j}}(dw_j|\mathcal{G}_{0,j}) \in \mathcal{M}_1(\mathcal{W}_j)$, $0 \leq j \leq n-1$.

Uncertainty on Joint Distribution.

Given the above formulation and Type 1 uncertainty a minimax stochastic controlled problem can be formulated over a total variation distance uncertainty ball, centered at the nominal joint measure $\mathbb{P}_w \in \mathcal{M}_1(\mathcal{W}_{0,n-1})$ having radius $R \in [0,2]$ with respect to the total variation distance metric. The precise problem statement should thus, be as follows.

Problem 4.6: (Type 1 Uncertainty) For a given $u \in \mathcal{U}_{ad}^{w}[0,n-1]$ assume that the measures $M(u)$ induced by the true system while control $u \in \mathcal{U}_{ad}^{w}[0,n-1]$ is applied are $M(u) \subset \mathcal{M}_1(\mathcal{W}_{0,n-1})$. Given a nominal system of Definition 4.2 an admissible control set $\mathcal{U}_{ad}^{w}[0,n-1]$ and Type 1 uncertainty class $\mathcal{A}_R(\mathbb{P}_w)$ find a $u^* \in \mathcal{U}_{ad}^{w}[0,n-1]$ and a probability measure $\mathbb{Q}_w^* \in \mathcal{A}_R(\mathbb{P}_w)$ which solve the minimax optimization problem

$$J_{0,n}(u^*,\mathbb{Q}_w^*) = \inf_{u\in\mathcal{U}_{ad}^{w}[0,n-1]} \sup_{\mathbb{Q}_w\in\mathcal{A}_R(\mathbb{P}_w)} \Big\{ E_{\mathbb{Q}_w}\Big(\sum_{k=0}^{n-1} f_k(x_k^u,u_k,w_k) + h_n(x_n^u)\Big)\Big\} \quad \text{(IV.7)}$$

Since $M(u) \subset \mathcal{M}_1(\mathcal{W}_{0,n-1})$ then the expectation in (IV.5) is with respect to $\mathbb{Q}_w \in \mathcal{M}_1(\mathcal{W}_{0,n-1})$. Using Lemma 3.1, Lemma 3.2 and Theorem 3.3 the maximization over the above problem can be re-written as follows.

$$J_{0,n}(u,\mathbb{Q}_w^*) = R.\parallel \sum_{k=0}^{n-1} f_k(x_k^u,u_k,w_k) + h_n(x_n^u) \parallel_\infty$$
$$+E_{\mathbb{P}_w}\Big\{\sum_{k=0}^{n-1} f_k(x_k^u,u_k,w_k) + h_n(x_n^u)\Big\} \quad \text{(IV.8)}$$
$$= (1+\beta)E_{\mathbb{Q}_w^*}\Big\{\sum_{k=0}^{n-1} f_k(x_k^u,u_k,w_k) + h_n(x_n^u)\Big\} \quad \text{(IV.9)}$$

where above $\parallel\cdot\parallel_\infty$ is defined over the space $\mathcal{W}_{0,n-1}$, and $\mathbb{Q}_w^*$ is given by the following measure.

$$\mathbb{Q}_w^*(A) = \gamma.\frac{e^{\alpha\sum_{j=0}^{n-1} f_j(x_j^u,u_j)+h_n(x_n^u)}}{E_\eta\Big(e^{\alpha\sum_{j=0}^{n-1} f_j(x_j^u,u_j)+h_n(x_n^u)}\Big)}\mathbb{E}_w(A) + (1-\gamma)\mathbb{P}_w(A) \quad \text{(IV.10)}$$

where $A \in \mathcal{B}(\mathcal{W}_{0,n-1})$, $\gamma = \frac{\beta}{1+\beta}$, and α depends on the choice of control u. Notice that $\mathbb{E}_w \in \mathcal{M}_1(\mathcal{W}_{0,n-1})$ is an arbitrary measure which is not unique. Clearly, the minimization problem (IV.9) is not standard. Further investigation is needed to derive a principle of optimality and a dynamic programming equation.

Uncertainty on Conditional Distribution.

Given the above formulation and Type 2 uncertainty a minimax stochastic controlled problem can be formulated over a total variation distance uncertainty ball, centered at the nominal conditional distribution $P_{w_i}(dw_i;x_i^u,u_i) \in \mathcal{Q}(\mathcal{W}_i;\mathcal{X}_i\times\mathcal{U}_i)$ having radius $R_i \in [0,2]$, for $i = 0,1,\dots,n-1$ with respect to the total variation distance metric. The precise problem statement should thus, be as follows.

Problem 4.7: (Type 2 Uncertainty) For a given $u \in \mathcal{U}_{ad}[0,n-1]$ assume that the measures $M(u)$ induced by the true system while control $u \in \mathcal{U}_{ad}[0,n-1]$ is applied are $M(u) \subset \mathcal{M}_1(\mathcal{W}_{0,n-1})$. Given a nominal system of Definition 4.2 an admissible control set $\mathcal{U}_{ad}[0,n-1]$ and Type 2 uncertainty class $\mathcal{B}_{R_k}(\mathbb{P}_{w_k})(\mathcal{G}_{0,k}), k=0,1,...,n-1$ find a $u^* \in \mathcal{U}_{ad}[0,n-1]$ and a sequence of stochastic kernels $\mathbb{Q}_{w_k}^*(dw_k;\mathcal{G}_{0,k}) \in \mathcal{B}_{R_k}(P_{w_k})(\mathcal{G}_{0,k})$, $k=0,1,...,n-1$ which solve the minimax optimization problem.

$$J_{0,n}(u^*,\{\mathbb{Q}_{w_k}^*\}_{k=0}^{n-1}) = \inf_{u\in\mathcal{U}_{ad}[0,n-1]} \sup_{\substack{\mathbb{Q}_{w_k}(dw_k;\mathcal{G}_{0,k})\in\mathcal{B}_{R_k}(P_{w_k})(\mathcal{G}_{0,k})\\ k=0,1,\dots,n-1}} \Big[E_{\mathbb{Q}}\Big\{\sum_{k=0}^{n-1} f_k(x_k^u,u_k,w_k) + h_n(x_n^u)\Big\}\Big] \quad \text{(IV.11)}$$

First, it is essential to determine whether there exist maximizing measures $\mathbb{Q}_{w_i}^*(dw_i;\mathcal{G}_{0,i}) \in \mathcal{B}_{R_i}(P_{w_i})(\mathcal{G}_{0,i})$, a.s., $i=0,1,...,n-1$ and whether dynamic programming can be used. Consider the average pay-off

$$\sup_{\substack{\mathbb{Q}_{w_k}(dw_k;\mathcal{G}_{0,k})\in\mathcal{B}_{R_k}(P_k)(\mathcal{G}_{0,k})\\ k=0,1,\dots,n-1}} J_{0,n}(u,\{\mathbb{Q}_{w_i}\}_{i=0}^{n-1})$$
$$= \sup_{\mathbb{Q}_{w_0}(dw_0;\mathcal{G}_{0,0})\in\mathcal{B}_{R_0}(P_{w_0})(\mathcal{G}_{0,0})} E_{\mathbb{Q}_{w_0}}\Big\{f_0(x_0^u,u_0,w_0)$$
$$+ \sup_{\mathbb{Q}_{w_1}(dw_1;\mathcal{G}_{0,1})\in\mathcal{B}_{R_1}(P_{w_1})(\mathcal{G}_{0,1})} E_{\mathbb{Q}_{w_1}}\Big\{f_1(x_1^u,u_1,w_1) + ...$$
$$+ \sup_{\mathbb{Q}_{w_{n-1}}(dw_{n-1};\mathcal{G}_{0,n-1})\in\mathcal{B}_{R_{n-1}}(P_{w_{n-1}})(\mathcal{G}_{0,n-1})} E_{\mathbb{Q}_{w_{n-1}}}\Big\{$$
$$f_{n-1}(x_{n-1}^u,u_{n-1},w_{n-1}) + h_n(b_{n-1}(x_{n-1}^u,u_{n-1},w_{n-1}))$$
$$|\mathcal{G}_{0,n-1}\Big\}...|\mathcal{G}_{0,1}\Big\}|\mathcal{G}_{0,0}\Big\} \quad \text{(IV.12)}$$

By invoking Lemma 3.2 the maximization of inner conditional expectations can be done by assuming that these expectations result at each stage to a function which is bounded continuous. For example, the maximization over the ball $\mathcal{B}_{R_{n-1}}(P_{w_{n-1}})(\mathcal{G}_{0,n-1})$ yields

$$R_{n-1}\max_{w_{n-1}\in\mathcal{W}_{n-1}}\Big\{f_{n-1}(x_{n-1}^u,u_{n-1},w_{n-1})$$
$$+h_n(b_{n-1}(x_{n-1}^u,u_{n-1},w_{n-1}))\Big\}$$
$$+E_{\mathbb{Q}_{w_{n-1}}}\Big\{f_{n-1}(x_{n-1}^u,u_{n-1},w_{n-1})$$
$$+h_n(b_{n-1}(x_{n-1}^u,u_{n-1},w_{n-1}))|\mathcal{G}_{0,n-1}\Big\}$$

By Theorem 3.3, the optimal conditional distribution is characterized by

$$\mathbb{Q}_{w_{n-1}}^*(dw_{n-1};\mathcal{G}_{0,n-1})$$
$$= \frac{\gamma_{n-1}e^{a_{n-1}\Big((f_{n-1}+h_n(b_{n-1}))(x_{n-1}^u,u_{n-1},w_{n-1})\Big)}}{\mathbb{E}_{w_{n-1}|\mathcal{G}_{0,n-1}}e^{a_{n-1}\Big((f_{n-1}+h_n(b_{n-1}))(x_{n-1}^u,u_{n-1},w_{n-1})\Big)}}$$
$$\times\mathbb{E}_{w_{n-1}}(dw_{n-1};\mathcal{G}_{0,n-1})$$
$$+(1-\gamma_{n-1})P_{w_{n-1}}(dw_{n-1};x_{n-1}^u,u_{n-1})$$

Let $\mathbb{E}_{w_{n-1}}(dw_{n-1};\mathcal{G}_{0,n-1}) = P_{w_{n-1}}(dw_{n-1};x^u_{n-1},u_{n-1})$, a.s. then $\mathbb{Q}^*_{w_{n-1}}(dw_{n-1};\mathcal{G}_{0,n-1}) = \mathbb{Q}^*_{w_i}(dw_i;x^u_i,u_i), a.s.$. Let $V_j(x)$ represent the minimax cost on the future time horizon $\{j,j+1,...,n\}$ at time $j\in\mathbb{N}^n_+$ defined by

$$V_j(x) \stackrel{\triangle}{=} \inf_{u\in\mathcal{U}_{ad}[j,n-1]} \sup_{\substack{Q_{w_k}(dw_k;\mathcal{G}_{0,k})\in\mathcal{B}_{R_k}(P_{w_k})(\mathcal{G}_{0,k})\\ k=j,j+1,\dots,n-1}} \Big[E_{\mathbb{Q}}\Big\{ \sum_{k=j}^{n-1} f_k(x^u_k,u_k,w_k) + h_n(x^u_n)|\mathcal{G}_{0,k}\Big\}\Big] \tag{IV.13}$$

Then by reconditioning one obtains

$$V_j(x) \stackrel{\triangle}{=} \inf_{u\in\mathcal{U}_{ad}[j,n-1]} \sup_{\substack{Q_{w_k}(dw_k;\mathcal{G}_{0,k})\in\mathcal{B}_{R_k}(P_{w_k})(\mathcal{G}_{0,k})\\ k=j,j+1,\dots,n-1}} \Big[E_{\mathbb{Q}}\Big\{ f_k(x^u_k,u_k,w_k) + E_{\mathbb{Q}}\Big\{ \sum_{k=j+1}^{n-1} f_k(x^u_k,u_k,w_k) + h_n(x^u_n)|\mathcal{G}_{0,j+1}\Big\}|\mathcal{G}_{0,k}\Big\}\Big] \tag{IV.14}$$

Hence, the following dynamic programming recursion

$$V_j(x) \stackrel{\triangle}{=} \inf_{u\in\mathcal{U}_{ad}[j,j]} \sup_{Q_{w_j}(dw_j;\mathcal{G}_{0,j})\in\mathcal{B}_{R_j}(P_{w_j})(\mathcal{G}_{0,j})} \Big[E_{Q_{w_j}(dw_j;\mathcal{G}_{0,j})}\Big\{ f_j(x^u_j,u_j,w_j) + V_{j+1}(b_j(x_j,u_j,w_j))\Big\}\Big] \tag{IV.15}$$

$$V_n(x) = h_n(x_n) \tag{IV.16}$$

Assuming that the right side term $f_j(x^u_j,u_j,\cdot) + V_{j+1}(b_j(x_j,u_j,\cdot)) : \mathcal{W}_j \rightarrow \Re_+$ of (IV.15) is bounded continuous then the sup in the right side of (IV.15) is given by

$$\begin{aligned} &\sup_{\mathbb{Q}_{w_j}(dw_j;\mathcal{G}_{0,j})\in\mathcal{B}_{R_j}(P_{w_j})(\mathcal{G}_{0,j})} E_{\mathbb{Q}_{w_j}}(dw_j;\mathcal{G}_{0,j})\Big\{ f_j(x^u_j,u_j,w_j) + V_{j+1}(b_j(x_j,u_j,w_j))\Big\} \\ &= R_j \max_{w_j\in\mathcal{W}_j} \Big\{ f_j(x^u_j,u_j,w_j) + V_{j+1}(b_j(x_j,u_j,w_j))\Big\} \\ &+ E_{P_{w_j}}(dw_j;x^u_j,w_j)\Big\{ f_j(x^u_j,u_j,w_j) + V_{j+1}(b_j(x_j,u_j,w_j))\Big\} \end{aligned} \tag{IV.17}$$

The point to be made here is that the dynamic programming equation involves in its right hand side the supremum of the cost-to-go in addition to the standard term. This is a new equation and to the best of our knowledge it has not appeared in the literature.

V. Future Work

Some examples should be worked out to understand the the implications of the maximization over uncertainty on distributions with respect to the total variation distance.

References

[1] F. Rezaei, C.D. Charalambous, N.U. Ahmed, Optimization of Stochastic Uncertain Systems with Variational Norm Constraints, *in 2007 Proceedings of the 46th IEEE Conference on Decision and Control*, New Orleans, U.S.A., December 12-14, 2007.

[2] F. Rezaei, C.D. Charalambous, N.U. Ahmed, Optimal Control of Uncertain Stochastic Systems Subject to Toatal Variation Distance Uncertainty, *SIAM Journal on Control and Optimization*, Submitted, 2010, pp.35.

[3] M.S. Pinsker, *Information and Information Stability of Random Variables and Processes*, Holden-Day, San Francisco, 1964.

[4] P. Dai Pra, L. Meneghini, W. J. Runggaldier, Connections between Stochastic control and Dynamic Games, *Matehmatics of Control, Signals and Systems*, vol.9, No.4, 1996, pp.303-326.

[5] V. A. Ugrinovskii, I.R. Petersen, Finite Horizon Minimax Optimal Control of Stochastic Partially Observed Time Varying Uncertain Systems, *Mathematics of Control, Signal and Systems*, vol.12, 1999, pp.1-23.

[6] I.R. Petersen, M.R. James, P. Dupuis, Minimax Optimal Control of Stochastic Uncertain Systems with Relative Entropy Constraints, *IEEE Transactions on Automatic Control*, Vol. 45, No. 3, March 2000, pp.398-412.

[7] L.P. Hansen, T. J. Sargent, Robust Control and Model Uncertainty, *The American Economic Review*, Vol.91, No.2, pp.60-66, 2001.

[8] F. Rezaei, C.D. Charalambous, A. Kyprianou, Optimization of fully observable nonlinear stochastic uncertain controlled diffusion, *Proceeding of the 43rd IEEE Conference on Decision and Control*, pp.2555-2560, Dec. 2004.

[9] N.U. Ahmed, C.D. Charalambous, Relative Entropy Applied to Optimal Control of Stochastic Uncertain Systems on Hilbert Space, *Proceeding of the 44th IEEE Conference on Decision and Control*, pp.1776-1781, Dec. 2005.

[10] C.D. Charalambous, F. Rezaei, Stochastic Uncertain Systems Subject to Relative Entropy Constraints: Induced Norms and Monotonicity Properties of Minimax Games, *IEEE Transactions on Automatic Control*, Vol.52, No.4, pp.647-663, April 2007.

[11] P.R. Kumar, J.H., 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.

[12] W.M. McEneany, Connections between Risk-Sensitive Stochastic Control, Differential Games, and H^∞ Control: The Nonlinear Case, PhD Dissertation, Brown University, 1993.

[13] C.D. Charalambous, Partially Observable Nonlinear Risk-Sensitive Control Problems: Dynamic Programming and Verification Theorems, *IEEE Transactions on Automatic Control*, Vol. 42, No. 8, August 1997, pp.1130-1138.

[14] W. Rudin, *Functional Analysis*, McGraw-Hill, New York, 1991.

[15] E. Yang, Z. Zhang, On the Redundancy of Lossy Source Coding with Abstract Alphabets, *IEEE Transactions on Information Theory*, Vol. IT-45, pp. 1092-1110, 1999.

[16] J. Yong and X. Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, New York: Springer-Verlag, 1999.

[17] N. Dunford and J. T. Schwartz, *Linear Operators: Part 1: General Theory*, Interscience Publishers, Inc., New York, 1957.

[18] H.V. Poor, On Robust Wiener Filtering, *IEEE Transactions on Automatic Control*, Vol. 25, No. 3, June 1980, pp.531-536.

[19] K.S. Vastola and H.V. Poor, On Robust Wiener-Kolmogorov Theory, *IEEE Transactions on Information Theory*, Vol. 30, No. 2, March 1984, pp.315-327.