Continuity and Robustness to Incorrect Priors in Estimation and Control

Graeme Baker and Serdar Yüksel

***Abstract*—This paper studies continuity properties of single and multi stage estimation and stochastic control problems with respect to initial probability distributions and applications of these results to the study of robustness of control policies applied to systems with incomplete probabilistic models. We establish that continuity and robustness cannot be guaranteed under weak and setwise convergences, but the optimal cost is continuous under the more stringent topology of total variation for stage-wise cost functions that are nonnegative, measurable, and bounded. Under further conditions on either the measurement channels or the source processes, however, weak convergence is sufficient. We also discuss similar properties under the Wasserstein distance. These results are shown to have direct implications, positive or negative, for robust control: If an optimal control policy is applied to a prior model $\tilde{P}$, and if $\tilde{P}$ is close to the true model P, then the application of the incorrect optimal policy to the true model leads to a loss that is continuous in the distance between $\tilde{P}$ and P under total variation, and under some setups, weak convergence distance measures.**

I. Introduction

We start with the probabilistic setup of the problem. Let $\mathbb{X} \subset \mathbb{R}^n$, be a Borel set in which elements of a controlled Markov process $\{X_t,\, t \in \mathbb{Z}_+\}$ live. Here and throughout the paper $\mathbb{Z}_+$ denotes the set of nonnegative integers and $\mathbb{N}$ denotes the set of positive integers. Let $\mathbb{Y} \subset \mathbb{R}^m$ be a Borel set, and let an observation channel Q be defined as a stochastic kernel (regular conditional probability) from $\mathbb{X}$ to $\mathbb{Y}$, such that $Q(\,\cdot\,|x)$ is a probability measure on the (Borel) σ-algebra $\mathcal{B}(\mathbb{Y})$ on $\mathbb{Y}$ for every $x \in \mathbb{X}$, and $Q(A|\,\cdot\,) : \mathbb{X} \to [0,1]$ is a Borel measurable function for every $A \in \mathcal{B}(\mathbb{Y})$. Throughout the paper we let $\mathcal{Q}$ denote the set of all stochastic kernels from $\mathbb{X}$ to $\mathbb{Y}$. Let a decision maker (DM) be located at the output an observation channel $Q \in \mathcal{Q}$, with inputs X_t and outputs Y_t. Let $\mathbb{U}$, the action space, be a Borel subset of some Euclidean space. An *admissible policy* γ is a sequence of control functions $\{\gamma_t,\, t \in \mathbb{Z}_+\}$ such that γ_t is measurable with respect to the σ-algebra generated by the information variables

$$I_t = \{Y_{[0,t]}, U_{[0,t-1]}\}, \quad t \in \mathbb{N}, \qquad I_0 = \{Y_0\}.$$

where $U_t = \gamma_t(I_t)$ are the $\mathbb{U}$-valued control actions and we use the notation

$$Y_{[0,t]} = \{Y_s,\, 0 \le s \le t\}, \quad U_{[0,t-1]} = \{U_s,\, 0 \le s \le t-1\}.$$

The authors are with the Department of Mathematics and Statistics, Queen's University, Kingston, Ontario, Canada, K7L 3N6. Email: 11gsb6@queensu.ca,yuksel@mast.queensu.ca. This research was partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

We define Γ to be the set of all such admissible policies.

Given a control policy, the joint distribution of the state, control, and observation processes is defined by:

$$\Pr\big((X_0,Y_0) \in B\big) = \int_B P(dx_0)Q(dy_0|x_0), \quad B \in \mathcal{B}(\mathbb{X}\times\mathbb{Y}),$$

where P is the (prior) distribution of the initial state X_0, and

$$\Pr\bigg((X_t,Y_t) \in B \,\bigg|\, x_{[0,t-1]}, y_{[0,t-1]}, u_{[0,t-1]}\bigg) = \int_B \mathcal{T}(dx_t|x_{t-1},u_{t-1})Q(dy_t|x_t), \quad B \in \mathcal{B}(\mathbb{X}\times\mathbb{Y}),$$

where $\mathcal{T}(\cdot|x,u)$ is a stochastic kernel from $\mathbb{X}\times\mathbb{U}$ to $\mathbb{X}$ and $Q(\cdot|x)$ is a stochastic kernel from $\mathbb{X}$ to $\mathbb{Y}$.

Using stochastic realization results (see Lemma 1.2 in [6], or Lemma 3.1 of [3]), the above dynamics can be represented in a functional form as follows: we can consider a dynamical system described by the discrete-time equations

$$\begin{aligned} X_{t+1} &= f(X_t,U_t,W_t), \\ Y_t &= g(X_t,V_t) \end{aligned} \tag{1}$$

for some measurable functions f, g, with $\{W_t\}$ being an independent and identically distributed (i.i.d) system noise process and $\{V_t\}$ an i.i.d. disturbance process, which are independent of X_0 and each other. Here, the second equation represents the communication channel Q, as it describes the relation between the state and observation variables. We let the objective be the minimization of

$$J(P,Q,\gamma) = E_P^{Q,\gamma}\left[\sum_{t=0}^{T-1} c(X_t,U_t)\right]$$

over the set of admissible policies $\gamma \in \Gamma$, where $c : \mathbb{X}\times\mathbb{U} \to \mathbb{R}$ is a Borel-measurable stagewise cost function and $E_P^{Q,\gamma}$ denotes the expectation with initial state probability measure P and channel Q under policy γ. Note that $P \in \mathcal{P}(\mathbb{X})$, where we let $\mathcal{P}(\mathbb{X})$ denote the space of all probability measures on $\mathbb{X}$.

Even though the primary focus in the paper is on the single and finite-stage cost setups, we will also briefly discuss the discounted cost infinite horizon setting and the average cost settings. For these, the criteria to be minimized are given respectively as

$$J_\beta(P,Q,\gamma) = E_P^{Q,\gamma}\left[\sum_{t=0}^{\infty} \beta^t c(X_t,U_t)\right]$$

for $0 < \beta < 1$, and

$$J_\infty(P, Q, \gamma) = \limsup_{T\to\infty} \frac{1}{T} E_P^{Q,\gamma} \left[\sum_{t=0}^{T-1} c(X_t, U_t)\right].$$

We define the optimal cost for the finite-stage, discounted cost, and average cost cases by:

$$J^*(P, Q) = \inf_{\gamma\in\Gamma} J(P, Q, \gamma),$$
$$J^*_\beta(P, Q) = \inf_{\gamma\in\Gamma} J_\beta(P, Q, \gamma),$$
$$\text{and } J^*_\infty(P, Q) = \inf_{\gamma\in\Gamma} J_\infty(P, Q, \gamma),$$

respectively. We are interested in the following problems:

Problem P1: Continuity of $J^*(P, Q)$ on the space of prior distributions. Suppose $\{P_n, n \in \mathbb{N}\}$ is a sequence of priors converging in some sense to P. When does $P_n \to P$ imply

$$J^*(P_n, Q) \to J^*(P, Q)?$$

Problem P2: Robustness due to mismatch and robust Bayesian control. Another problem of importance is robustness of an optimal controller to modeling errors. Suppose that an optimal policy is constructed according to a model which is incorrect: Suppose that γ_n is an optimal policy designed for P_n, an incorrect model for a true model P. Is it the case that if $P_n \to P$ then $J(P, Q, \gamma_n) \to J^*(P, Q)$? A related problem is to obtain explicit bounds on the mismatch error

$$|J(P, Q, \gamma_n) - J^*(P, Q)|$$

as a function of some distance measure between P_n and P.

II. Single Stage: Continuity of Optimal Cost with Respect to Initial Distribution

Given a Borel-measurable cost function, $c : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$, we denote the optimal cost with channel $Q \in \mathcal{Q}$ and input distribution $P \in \mathcal{P}(\mathbb{X})$ by

$$J^*(P, Q) = \inf_{\gamma\in\Gamma} \int_{\mathbb{X}\times\mathbb{Y}} c(x, \gamma(y))Q(\mathrm{d}y|x)P(\mathrm{d}x).$$

The results in this section will utilize one or more of the following assumptions on the cost function c and the (Borel) set $\mathbb{U} \subset \mathbb{R}^k$:

Assumption II.1.

1. *The function $c : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ is nonnegative, bounded, and continuous on $\mathbb{X} \times \mathbb{U}$.*
2. *The function $c : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ is nonnegative, measurable, and bounded.*
3. *The function $c : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ is nonnegative, bounded, and continuous on $\mathbb{U}$ for every $x \in \mathbb{X}$.*
4. *$\mathbb{U}$ is a compact set.*
5. *$\mathbb{U}$ is a convex set.*

Before proceeding further, we state the following conditions under which an optimal control policy exists, i.e., when the infimum in $\inf_{\gamma\in\Gamma} E_P^{Q,\gamma}[c(x_0, u_0)]$ is attainable.

Theorem II.1 ([16]). *Suppose Assumptions II.1.3 and II.1.4 hold. Then, there exists an optimal control policy for any channel Q.*

We now review three notions of convergence for sequences of probability measures: weak, setwise, and convergence in total variation. Firstly, for some $N \in \mathbb{N}$ a sequence $\{\mu_n, n \in \mathbb{N}\}$ in $\mathcal{P}(\mathbb{R}^N)$ is said to converge to $\mu \in \mathcal{P}(\mathbb{R}^N)$ *weakly* if

$$\int_{\mathbb{R}^N} c(x)\mu_n(\mathrm{d}x) \to \int_{\mathbb{R}^N} c(x)\mu(\mathrm{d}x) \qquad (*)$$

for every continuous and bounded $c : \mathbb{R}^N \to \mathbb{R}$. Secondly, $\{\mu_n\}$ is said to converge *setwise* to $\mu \in \mathcal{P}(\mathbb{R}^N)$ if $(*)$ holds for all measurable and bounded $c : \mathbb{R}^N \to \mathbb{R}$. For two probability measures $\mu, \nu \in \mathcal{P}(\mathbb{R}^N)$, the *total variation* metric is $\|\mu - \nu\|_{TV} = \sup_{B\in\mathcal{B}(\mathbb{R}^N)} |\mu(B) - \nu(B)|$. A sequence $\{\mu_n\}$ is said to converge in total variation to $\mu \in \mathcal{P}(\mathbb{R}^N)$ if $\|\mu_n - \mu\|_{TV} \to 0$. One could also consider other measures such as relative entropy; however since relative entropy is stronger than total variation by Pinsker's inequality, the positive continuity results that we will obtain will be applicable to convergence under relative entropy as well. We will also discuss the Wasserstein metric later in the paper and its connection with weak convergence.

For $P \in \mathcal{P}(\mathbb{X})$ and $Q \in \mathcal{Q}$ we let PQ denote the joint distribution induced on $(\mathbb{X} \times \mathbb{Y}, \mathcal{B}(\mathbb{X} \times \mathbb{Y}))$ by channel Q with input distribution P:

$$PQ(A) = \int_A Q(\mathrm{d}y|x)P(\mathrm{d}x), \; A \in \mathcal{B}(\mathbb{X} \times \mathbb{Y}).$$

Lemma II.1. *Let $P \in \mathcal{P}(\mathbb{X})$ be an initial distribution, $Q \in \mathcal{Q}$ a measurement channel, and $\{P_n, n \in \mathbb{N}\}$ a sequence in $\mathcal{P}(\mathbb{X})$.*

(i) Assume that $Q(dy|x)$ is weakly continuous in x in the sense that $\int Q(dy|x)c(y)$ is continuous in x for every continuous and bounded c. If $P_n \to P$ weakly then $P_nQ \to PQ$ weakly.

(ii) If $P_n \to P$ setwise then $P_nQ \to PQ$ setwise.

(iii) If $P_n \to P$ in total variation then $P_nQ \to PQ$ in total variation. In particular,

$$\|P_nQ - PQ\|_{TV} = \|P_n - P\|_{TV}.$$

A. Connection with the convergence of measurement channels under an absolute continuity condition

We present a characterization relating sequences with changing input distributions to sequences with changing measurement channels. This characterization connects continuity on input distributions to the results from our earlier work [16] on the continuity of $J^*(P, Q)$ on the space of measurement channels. Recall that for $\mu, \nu \in \mathcal{P}(\mathbb{R}^N)$, ν is absolutely continuous with respect to μ, written $\nu \ll \mu$, if for all $A \in \mathcal{B}(\mathbb{R}^N)$, $\mu(A) = 0$ implies $\nu(A) = 0$. The next result follows from the Radon-Nikodym theorem [9].

Lemma II.2. *If $P_nQ \to PQ$ (weakly/setwise/in total variation) and $P_n \ll P$ for all n then there exists a sequence of* *measurement channels $\{Q_n\}$ such that $P_nQ(A) = PQ_n(A)$ for all $A \in \mathcal{B}(\mathbb{X}\times\mathbb{Y})$ and $PQ_n \to PQ$.*

B. Weak convergence

Theorem II.2. *Let a channel $Q \in \mathcal{Q}$ be given. $J^*(P,Q)$ is not necessarily continuous in P under weak convergence. This holds even when $\mathbb{X}$, $\mathbb{Y}$, and $\mathbb{U}$ are compact and when c is bounded and continuous in both x and u.*

Proof. This result is implied by Theorem II.7. ☐

The next lemma shows that the optimal cost is unchanged when γ is restricted to the class of continuous policies.

Lemma II.3 ([16])**.** *Let μ be an arbitrary probability measure on $(\mathbb{X}\times\mathbb{Y}, \mathcal{B}(\mathbb{X}\times\mathbb{Y}))$ and let $\mathcal{C}$ be the set of continuous functions from $\mathbb{Y}$ to $\mathbb{U}$. If Assumptions II.1.2 and II.1.5 hold, then*

$$\inf_{\gamma\in\Gamma}\int c(x,\gamma(y))\mu(\mathrm{d}x,\mathrm{d}y) = \inf_{\gamma\in\mathcal{C}}\int c(x,\gamma(y))\mu(\mathrm{d}x,\mathrm{d}y)$$

The following theorem is related to [16] and Theorem 3 of Wu and Verdú [14].

Theorem II.3. *Let a channel $Q \in \mathcal{Q}$ be given. If $Q(B|x)$ is lower semi-continuous in x for each fixed $B \in \mathcal{B}(\mathbb{Y})$ and if Assumptions II.1.1 and II.1.5 hold, then $J^*(P,Q)$ is upper semi-continuous on $\mathcal{P}(\mathbb{X})$ under weak convergence.*

Theorem II.4. *Let $c : \mathbb{X}\times\mathbb{U} \to \mathbb{R}$ be given and Assumptions II.1.1 and II.1.4 hold. If $Q \in \mathcal{Q}$ is so that*

$$\lim_{k\to\infty}\sup_{\gamma}\left|\int Q(\mathrm{d}y|x_k)c(x_k,\gamma(y)) - \int Q(\mathrm{d}y|x)c(x,\gamma(y))\right| = 0 \tag{2}$$

when $x_k \to x$ and if $P_n \to P$ weakly then

$$J^*(P_n,Q) \to J^*(P,Q).$$

Example II.1. Consider the additive noisy channel:

$$y = x + w,$$

where $w \sim \mu$ with μ admitting a density, η, which is continuous. An example is the Gaussian density. Suppose that Assumptions II.1.1 and II.1.4 hold. For $x_k \to x$ we have that

$$\lim_{k\to\infty}\sup_{\gamma}\left|\int \eta(y-x_k)c(x_k,\gamma(y))\mathrm{d}y - \int \eta(y-x)c(x,\gamma(y))\mathrm{d}y\right|$$
$$\leq \lim_{k\to\infty}\|\eta(\cdot - x_k) - \eta(\cdot - x)\|_{TV}\cdot\|c\|_\infty \tag{3}$$
$$+ \lim_{k\to\infty}\sup_{\gamma}\left|\int \eta(y-x)\left(c(x_k,\gamma(y)) - c(x,\gamma(y))\right)\mathrm{d}y\right| = 0$$

where $\|c\|_\infty$ denotes the supremum norm of c. We note that the term in (3) tends to zero since $\eta(\cdot - x_k)$ converges to $\eta(\cdot - x)$ pointwise and therefore by Scheffé's theorem it converges in L_1 and thus in total variation. Additionally, the last equality holds since c is uniformly continuous by Assumptions II.1.1 and II.1.4. Therefore (2) is satisfied and Theorem II.4 holds. ◇

We note here that a related result due to Wu and Verdú [14] establishes continuity of the MMSE error under weak convergence when the channel is additive. For further properties in this context, we refer the reader to [7]. In general, however, the following example shows that the boundedness condition cannot be relaxed.

Example II.2. Let $\mathbb{X} = \mathbb{U} = \mathbb{R}$, $\mathbb{Y} = [0,1]$, and $c(x,u) = (x-u)^2$. With the given cost function this is a mean-square error problem; therefore, the optimal policy is $\gamma^*(y) = E[x|y]$. We let the channel be distributed uniformly on $[0,1]$, that is, $Q \sim U([0,1])$. Note that this channel is non-informative. Let P_n be the following atomic distribution,

$$P_n = \left(\frac{1}{2} - \frac{1}{n}\right)\cdot\left(\delta_{\frac{1}{n}} + \delta_{-\frac{1}{n}}\right) + \frac{1}{2n}\cdot\left(\delta_{a_n} + \delta_{-a_n}\right)$$

where δ_s is the delta measure at point s, that is, $\delta_s(A) = 1_{\{s\in A\}}$, for any Borel set A, and a_n is the sequence of numbers in $\mathbb{N}$ defined by $a_n = \sqrt{n - \left(\frac{1}{n} + \frac{2}{n^2}\right)}$. It can be shown that $P_n \to \delta_0$ weakly. By symmetry and the non-informative nature of Q, the optimal policy is $\gamma^*(y) = E[X|Y] = 0$ for all P_n and for $P = \delta_0$. With initial distribution P, we have

$$J^*(P,Q) = E_P^{Q,\gamma^*}[(X-U)^2] = E_P^{Q,\gamma^*}[(X)^2] = 0$$

For all $n \in \mathbb{N}$, $J^*(P_n,Q) = E_{P_n}^{Q,\gamma^*}[(X)^2] = 1$, so $J^*(P_n,Q) \not\to J^*(P,Q)$ as $n \to \infty$. ◇

1) Weak convergence through the Wasserstein metric and a simulation argument: The Wasserstein metric is useful for the following *simulation* based analysis.

Definition II.1 (Wasserstein metric)**.** The *Wasserstein metric* of order 1 for two distributions $\mu,\nu \in \mathcal{P}(\mathbb{X})$ is defined as

$$W_1(\mu,\nu) = \inf_{\eta\in\mathcal{H}(\mu,\nu)}\int_{\mathbb{X}\times\mathbb{X}}\eta(\mathrm{d}x,\mathrm{d}y)|x-y|,$$

where $\mathcal{H}(\mu,\nu)$ denotes the set of probability measures on $\mathbb{X}\times\mathbb{X}$ with first marginal μ and second marginal ν.

To facilitate the analysis, we will enlarge the class of admissible policies to include those that are possibly randomized. By Blackwell's Irrelevant Information Theorem ([2]; or see [15, p. 457]), randomization cannot improve the system performance when compared with the set of deterministic policies. However, this randomization allows us to *simulate* an alternative model (through arguments based on Borkar [4]) and generate a policy corresponding to the simulated system, as we detail next.

Assumption II.2. $Q(\cdot|x) = 1_{\{f(x)\in\cdot\}}$ *for some f.*

Assumption II.3. *The cost function is such that*

$$|c(x,u) - c(x',u)| \leq \alpha|x - x'|$$

for some $\alpha \in \mathbb{R}_+$ for all $u \in \mathbb{U}$.

Theorem II.5. *Let Assumptions II.2 and II.3 hold. As $W_1(P_n, P) \to 0$, $J^*(P_n, Q) \to J^*(P, Q)$.*

Proof. This follows from an argument in Borkar [4] and [13] (see Section IV.C in [13]) under Assumption II.2. □

For compact $\mathbb{X}$, the Wasserstein distance of order 1 metrizes the weak topology on the set of probability measures on $\mathbb{X}$ (see [11, Theorem 6.9]). In particular, on the set of probability measures μ with $\int |x|\mu(dx) < \infty$, weak convergence is equivalent to the convergence under the Wasserstein metric.

Assumption II.4. *There exists a measurable function f so that for some probability measure ν, the following holds:*

$$Q(Y \in A|x) = \int_A f(x, y)\nu(dy)$$

Theorem II.6. *Under Assumptions II.3 and II.4,*

$$|J^*(P_n, Q) - J^*(P, Q)| \to 0,$$

as $W_1(P_n, P) \to 0$, $J^*(P_n, Q) \to J^*(P, Q)$.

Proof. Here, we essentially build on [1]. □

C. Setwise convergence

In this section we investigate the continuity properties of $J^*(P, Q)$ on the space of initial distributions under setwise convergence.

Theorem II.7. *Let a channel $Q \in \mathcal{Q}$ be given. $J^*(P, Q)$ is not necessarily continuous in P under setwise convergence. This holds even when $\mathbb{X}$, $\mathbb{Y}$, and $\mathbb{U}$ are compact and when c is bounded and continuous in both x and u.*

Proof. We present the following counterexample: Let $\mathbb{X} = \mathbb{Y} = \mathbb{U} = [0, 1]$ and let $c(x, u) = (x - u)^2$. For $n \in \mathbb{N}$ and $k = 1, \ldots, n$, we define

$$L_{n,k} = \left[\frac{2k-2}{2n}, \frac{2k-1}{2n}\right), \; R_{n,k} = \left[\frac{2k-1}{2n}, \frac{k}{n}\right).$$

For ease of notation, we shall let $L = \{y \in \cup_{k=1}^n L_{n,k}\}$ and $R = \{y \in \cup_{k=1}^n R_{n,k}\}$. Next, we define the square-wave function by $h_n(t) = 1_{\{t \in L\}} - 1_{\{t \in R\}}$. As $\int_0^1 h_n(t)\mathrm{d}t = 0$ and $|h_n(t)| \leq 1$, the function

$$f_n(t) = (1 + h_n(t))1_{\{t \in [0,1]\}}$$

is a probability density function. Building on [16], by the proof of the Riemann-Lebesgue lemma, we have therefore

$$\lim_{n\to\infty} \int_0^1 f_n(t)g(t)\mathrm{d}t = \int_0^1 g(t) \text{ for all } g \in L_1([0, 1], \mathbb{R}).$$

So if we let $P_n \sim f_n$ for $n \in \mathbb{N}$, we have that $P_n \to P \sim U([0, 1])$ setwise. Next we let the channel be

$$Q(\cdot|x) \sim \frac{1}{2} \cdot \delta_x + \frac{1}{2} \cdot U([0, 1]).$$

For initial distribution P, the optimal policy is

$$\gamma_P^*(y) = E[X|Y] = \frac{1}{2}\left(\frac{1}{2} + y\right).$$

This gives $J^*(P, Q) = \frac{1}{16}$. By routine calculations, the optimal policy for initial distribution P_n is

$$\gamma_{P_n}^*(y) = \begin{cases} \frac{1}{2} - \frac{1}{4n} & \text{if } y \in \cup_{k=1}^n R_{n,k} \\ \frac{1}{3} \cdot \left(\frac{1}{2} - \frac{1}{4n}\right) + \frac{2}{3}y & \text{if } y \in \cup_{k=1}^n L_{n,k} \end{cases}.$$

This gives $J^*(P_n, Q) = \frac{1}{18} - \frac{1}{24n^2}$. So we have $J^*(P_n, Q) \to \frac{1}{18} \neq \frac{1}{16}$ as $n \to \infty$, and we see that the optimal cost is clearly not continuous on the space of initial distributions under setwise convergence. □

Theorem II.8. *Let a channel Q be given. If Assumption II.1.2 holds, then $J^*(P, Q)$ is upper semi-continuous on $\mathcal{P}(\mathbb{X})$ under setwise convergence.*

D. Continuity under total variation

We present the following theorem. This essentially builds on [16].

Theorem II.9. *Under Assumption II.1.2, the optimal cost $J^*(P, Q)$ is continuous on the set of input distributions, $\mathcal{P}(\mathbb{X})$, under the topology of total variation.*

E. Multi-Stage Case: Continuity of Optimal Cost with Respect to Initial Distribution

We now consider equation (1) for multi-stage continuity problems. Clearly, the lack of continuity for single-stage problems implies the lack of continuity of multi-stage problems. In the following, the emphasis will be on developing a setup where continuity can be established.

For stochastic control problems, *strategic measures* are defined (see [10]) as the set of probability measures induced on the product spaces of the state and action pairs by measurable control policies: Given an initial distribution on the state, and a policy, one can uniquely define a probability measure on the infinite product space consistent with finite dimensional distributions. Now, define a *strategic measure* under a policy $\gamma^n = \{\gamma_0^n, \gamma_1^n, \cdots, \gamma_k^n, \cdots\}$ as a probability measure defined on $\mathcal{B}(\mathbb{X} \times \mathbb{Y} \times \mathbb{U})^{\mathbb{Z}_+}$ by:

$$\begin{aligned} &P_{P_n}^{\gamma^n}(\mathrm{d}(x_0, y_0, u_0), \mathrm{d}(x_1, y_1, u_1), \cdots) \\ &\quad = P_n(\mathrm{d}x_0)Q(\mathrm{d}y_0|x_0)1_{\{\gamma^n(y_0) \in \mathrm{d}y_0\}}P_n(\mathrm{d}x_1|x_0, u_0) \\ &\quad \times Q(\mathrm{d}y_1|x_1)1_{\{\gamma^n(y_0,y_1) \in \mathrm{d}y_1\}} \cdots \end{aligned} \tag{4}$$

Under a strategic measure P^{γ^n} we define,

$$J_\beta^*(P_n, Q) = E_{P_n}^{\gamma^n}\left[\sum_k \beta^k c(x_k, \gamma_k^n(y_{[0,k]}))\right]$$

Theorem II.10.

$$|J_\beta^*(P_n, Q) - J_\beta^*(P, Q)| \leq \|P_n(\mathrm{d}x_0) - P(\mathrm{d}x_0)\|_{TV} \frac{\|c\|_\infty}{1 - \beta}$$

The analysis so far had assumed no structural properties on the controlled model or the observation model. If one knows further properties regarding the process to be controlled, a more refined bound can be attained, and in fact continuity even under weak convergence can be established. Similar to the analysis in Section II-B1, one could obtain stronger results when the source model satisfies certain coupling properties.

III. Robustness to Inconsistent Priors and Mismatch Bounds

A. Robustness to inconsistent priors under total variation

Throughout this section we are interested in a single stage stochastic control problem with cost function c, initial distribution P, and channel Q. We shall denote this problem by $\Xi = (c, P, Q)$. Consider the following problem: let $\widetilde{P}$ be another initial distribution. Decision maker DM^1 finds an optimal (or ε-optimal) policy, $\widetilde{\gamma}^*$, for the problem $\widetilde{\Xi} = (c, \widetilde{P}, Q)$ and applies it to Ξ. Can we find a bound on $J(P, Q, \widetilde{\gamma}^*) - J^*(P, Q)$?

Proposition III.1. *Assume that II.1.2 holds. Let $\widetilde{\gamma}^*$ be an optimal (or ε-optimal) control policy for the single stage stochastic control problem $\widetilde{\Xi} = (c, \widetilde{P}, Q)$, where c is a cost function, $\widetilde{P}$ is an initial distribution, and Q is a measurement channel. Let P be another probability distribution on $\mathbb{X}$. Then $|J(P, Q, \widetilde{\gamma}) - J^*(P, Q)| \leq 2\|c\|_\infty \|P - \widetilde{P}\|_{TV}$ $(+2\varepsilon)$.*

B. Lack of robustness to inconsistent priors under weak or setwise convergence

Proposition III.2. *Let a channel $Q \in \mathcal{Q}$, an initial distribution $P \in \mathcal{P}(\mathbb{X})$, and a cost function $c : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ be given. Assume that $P_n \to P$ setwise and let $\gamma^*_{P_n}$ be an optimal (or ε-optimal) policy for the control problem $\Xi_n = (c, P_n, Q)$. It does not follow that $J(P, Q, \gamma^*_{P_n}) \to J^*(P, Q)$ as $n \to \infty$. This result holds even if c is bounded and continuous in both x and u.*

C. Robustness under weak convergence

Theorem III.1. *Suppose that Assumptions II.1.1, II.1.4 and* (2) *hold. Then, as $P_n \to P$ weakly $\lim_{n\to\infty} |J(P, Q, \gamma^*_{P_n}) - J^*(P, Q)| \to 0$, that is the system is robust to errors in the priors under weak convergence.*

We recall from Example II.1 that the class of channels that satisfy (2) include the additive Gaussian channel.

D. Finite-horizon or discounted-horizon setups

Theorem III.2. *Let Assumptions II.1.1, II.1.4 and* (2) *hold. Then, $|J_\beta(P, Q, \gamma^*_{P_n}) - J^*(P, Q)| \leq \frac{1}{1-\beta}\|c\|_\infty$. Also,*

$$|J^*_\beta(P_n, Q) - J^*_\beta(P, Q)| \leq \|P_n(\mathrm{d}x_0) - P(\mathrm{d}x_0)\|_{TV} \frac{2\|c\|_\infty}{1-\beta}$$

Proof. We use that

$$J_\beta(P, Q, \gamma^*_{P_n}) - J^*(P, Q) = J_\beta(P, Q, \gamma^*_{P_n}) \\ - J_\beta(P_n, Q, \gamma^*_{P_n}) + J_\beta(P_n, Q, \gamma^*_{P_n}) - J^*(P, Q).$$

The proof is complete by Theorem II.10. □

E. Mismatch for multi-stage estimation and relations to stability of non-linear filters with incorrect initializations

In this section we consider a variant on the multistage stochastic control problem developed in Section I. For a fixed $T \in \mathbb{N}$ we consider the control-free system:

$$\begin{cases} x_{t+1} &= f(x_t, w_t) \\ y_t &= g(x_t, v_t),\ t \in \{1, \ldots, T\} \end{cases}$$

for some measurable functions f and g with $\{w_t\}$ and $\{v_t\}$ being independent and identically distributed (i.i.d.) noise processes, which are independent of x_0 and each other.

We consider the estimation problem $\Xi = (c, P^1, Q)$, where c is a cost function, P^1 is an initial distribution, and Q is a measurement channel. Suppose decision maker DM^1 has an incorrect prior, $P^2 \in \mathcal{P}(\mathbb{X})$, so he selects his optimality policy, γ^2.We are here interested in what happens when DM^1 applies γ^2 to estimation problem Ξ. We let γ^1 be the optimal policy under P^1. A detailed analysis leads to

$$E_{P^1}\left[\sum_{t=0}^{T-1} c(x_t, \gamma^2_t(I_t)) - c(x_t, \gamma^1_t(I_t))\right] \\ \leq 2\|c\|_\infty E_{P^1}\left[\sum_{t=0}^{T-1} \left\|P^1(\mathrm{d}x_t|I_t) - P^2(\mathrm{d}x_t|I_t)\right\|_{TV}\right] \quad (5)$$

The boundedness of (5) as $T \to \infty$ has been established under restrictive conditions such as strong mixing or observability/invertibility conditions, see [5].

References

[1] A. Arapostathis, V. S. Borkar, and M. K. Ghosh. *Ergodic control of diffusion processes*, volume 143. Cambridge University Press, 2012.

[2] D. Blackwell. Memoryless strategies in finite-stage dynamic programming. *Annals of Mathematical Statistics*, 35:863–865, 1964.

[3] V. S. Borkar. White-noise representations in stochastic realization theory. *SIAM J. on Control and Optimization*, 31:1093–1102, 1993.

[4] V. S. Borkar. Average cost dynamic programming equations for controlled Markov chains with partial observations. *SIAM Journal on Control and Optimization*, 39(3):673–681, 2000.

[5] P. Chigansky, R. Liptser, and R. van Handel. Intrinsic methods in filter stability. *Handbook of Nonlinear Filtering*, 2009.

[6] I. I. Gihman and A. V. Skorohod. *Controlled stochastic processes*. Springer Science & Business Media, 2012.

[7] D. Guo, Y. Wu, S. Shamai, and S. Verdú. Estimation in gaussian noise: Properties of the minimum mean-square error. *IEEE Trans. on Information Theory*, 57(4):2371–2385, 2011.

[8] O. Hernandez-Lerma and J. Lasserre. *Discrete-time Markov control processes*. Springer, 1996.

[9] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, New York, 3rd edition, 1987.

[10] M. Schäl. Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal. *Z. Wahrscheinlichkeitsth*, 32:179–296, 1975.

[11] C. Villani. *Optimal transport: old and new*. Springer, 2008.

[12] Richard L. Wheeden and Antoni Zygmund. *Measure and Integral*. Marcel Dekker, New York, 1977.

[13] R. G. Wood, T. Linder, and S. Yüksel. Optimality of Walrand-Varaiya type policies and approximation results for zero delay coding of markov sources. In *IEEE International Symposium on Information Theory*, Hong Kong, 2015.

[14] Y. Wu and S. Verdú. Functional properties of minimum mean-square error and mutual information. *IEEE Transactions on Information Theory*, 58(3):1289–1301, 2012.

[15] S. Yüksel and T. Başar. *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. Springer-Birkhäuser, New York, NY, 2013.

[16] S. Yüksel and T. Linder. Optimization and convergence of observation channels in stochastic control. *SIAM Journal on Control and Optimization*, 50(2):864–887, 2012.