Filter Stability, Observability and Robustness for Partially Observed Stochastic Dynamical Systems

by

Curtis James McDonald

A thesis submitted to the
Department of Mathematics and Statistics
in conformity with the requirements for
the degree of Master of Applied Science

Queen's University
Kingston, Ontario, Canada
August 2019

Abstract

Filter stability refers to the correction of an incorrectly initialized filter for a partially observed stochastic dynamical system with increasing measurements. In this thesis, we study the filter stability problem, develop new methods and results for both controlled and control-free stochastic dynamical systems, and study the implications of filter stability on robustness of optimal solutions for partially observed stochastic control problems. We introduce a definition of non-linear stochastic observability and through this notion of observability, we provide sufficient conditions for when a falsely initialized filter merges with the correctly initialized filter over time. We study stability under different notions such as the weak topology, total variation, and relative entropy. Additionally, we investigate properties of the transition kernel and measurement kernel which result in stability with an exponential rate of merging. We generalize our results to the controlled case, which is an unexplored area in the literature, to our knowledge.

Stability results are then applied to stochastic control problems. Under filter stability, we bound the difference in the expected cost incurred for implementing an incorrectly designed control policy compared to an optimal policy and relate filter stability, robustness, and unique ergodicity of non-linear filters.

Acknowledgments

I am grateful to my parents for their unwavering support and encouragement of my academic pursuits. From my childhood to this day they have provided me with the environment to achieve my goals.

I would like to thank my fellow classmates and colleagues during my undergraduate and graduate studies. Through countless hours of work we learned together, and through hours of leisure built lasting friendships.

I am most grateful to my supervisor Serdar Yüksel. His example and guidance has taught me much about how to be a good academic, but more importantly a good person.

Finally, I would like to thank Fady Alajaji, Francesco Cellarosi, and Gregory Smith for taking the time and effort to participate in my thesis committee.

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

In this thesis, we consider stochastic processes that are observed through noisy measurements. We will study both controlled and control-free modes. In the control-free setup, an observer watches the development of a Markov chain through imperfect or noisy measurements, but has no influence on the underlying Markov chain. Such models are known as Partially Observed Markov Processes (POMPs) or Hidden Markov Models (HMMs). In a controlled model, the observer is now known as a controller or decision maker (DM) who can influence the development of the system through control actions. Such models are called Partially Observed Markov Decision Processes (POMDP).

The goal of the observer or the DM is either to track the hidden state in a POMP or design a control policy to minimize some expected cost criterion in a POMDP. While doing this, a standard result in the theory of partially observed processes is that the DM needs to keep computing the conditional probability measure on the state given the past information: this measure is called the filter. A further result in the field, to be stated explicitly in the paper, is that these computations can be carried out in a recursive manner which is known as the non-linear filtering equation. In particular, if $\{X_n\}_{n=0}^{\infty}$ is a Markov process and $\{Y_n\}_{n=0}^{\infty}$ is a noisy measurement of X_n, then the process $\{P(X_n \in \cdot|Y_0, \cdots, Y_n)\}_{n=0}^{\infty}$ itself is a probability measure valued Markov process.

However, this recursion is dependent on the initial condition of the filter process. This is the initial belief or distribution the observer has about the starting point of the state process before the observer has made any measurements. If the initial belief of the observer does not match the true initial distribution of the state process, we say that the filter has been incorrectly initialized.

Rarely in practice is the prior selected correctly. Filter stability is a property of a POMP where, over time, the measurements made by the observer will correct any incorrectly initialized filter and the observer's estimate of the state will be accurate. Filter stability is defined rigorously in Section 1.4.

A notable case in the theory is the linear Gaussian POMP and the associated Kalman filter. Kalman filtering is very powerful since it allows for the filter update equation to be finite dimensional, as a Gaussian measure is uniquely characterized by the mean and the covariance of the associated random variable. In this case, under suitable controllability and observability assumptions, a standard and very powerful result is that the Kalman Filter is robust to initialization errors [29].

On the other hand, in general, the filter update equation is rather non-trivial and the problem of filter stability requires much more sophisticated derivations when compared with the analysis for the stability of the Kalman Filter. In this thesis, we are primarily interested in what properties of a POMP and POMDP give rise to filter stability. As part of our analysis, we will define when measurements are “informative” about the underlying state process resulting in stability.

In a POMP the observer is a passive part of the process, the observer simply sees the measurements and records them. The observer’s presence does not affect the development of the process. In a POMDP, the DM or controller takes an active role and at each time stage takes an action that affects the development of the process. The way the DM maps measurements to control actions is called a control policy.

We also study the stability of the filter in POMDPs, an area which does not have much work in the current literature. One of the main differences between a POMDP and a POMP is that a POMP is a Markov chain, while a POMDP is generally not since the DM’s control policy may depend on past measurements. This complicates the dependency structure of the POMDP and therefore results from the POMP literature do not directly apply to the controlled setup. This is perhaps one reason why there is hardly any study on filter stability for controlled stochastic models, except for the standard machinery involving the Kalman Filter.

Furthermore, in a POMDP the DM has an objective: to minimize an expected cost incurred based on the control actions and the state realizations. Filter stability then plays a crucial role in the DM making optimal decisions. In many stochastic control problems, the filter is a sufficient statistic for an optimal control policy. If the DM has an incorrectly initialized filter, the DM may be selecting what it believes to be the optimal control action, but it has the wrong information and is actually selecting a very poor control action. It is then crucial that the filter—the DM’s vital source of information—becomes accurate over time so that the DM can apply good control actions and achieve a cost close to the optimal cost. This involves a problem of robustness in stochastic control, which we show to be closely tied to filter stability.

1.2 Partially Observed Markov Process

The components of a POMP are as follows. Let $\mathcal{X}, \mathcal{Y}$ be Polish spaces (that is, complete, separable, metric) equipped with their Borel sigma fields $\mathcal{B}(\mathcal{X})$ and $\mathcal{B}(\mathcal{Y})$. $\mathcal{X}$ will be called the state space, and $\mathcal{Y}$ the measurement space.

Given a measurable space $(X, \mathcal{B}(\mathcal{X}))$ we denote the space of probability measures on this space as $\mathcal{P}(\mathcal{X})$. We will denote random variables by capital letters and their realizations with lower case letters. Further, we will express contiguous sets of random variables such as $Y_0, Y_1, \cdots, Y_n$ with a subscript $Y_{[0,n]}$ indicating the starting and ending index of the collection. Infinite sequences $Y_0, Y_1, \cdots$ will be expressed as $Y_{[0,\infty)}$. We then define two probability kernels, the transition kernel T and the measurement kernel Q:

$$\begin{aligned} T : \mathcal{X} &\to \mathcal{P}(\mathcal{X}) & \qquad Q : \mathcal{X} &\to \mathcal{P}(\mathcal{Y}) \\ x &\mapsto T(dx'|x) & x &\mapsto Q(dy|x) \end{aligned}$$

where for a set $A \in \mathcal{B}(\mathcal{Y})$ we write $Q(A|x) = \int_A Q(dy|x)$. For these kernel operators, we can overload the notation to define them as mappings from a space of probability measures to another space of probability measures as follows

$$\begin{aligned} T : \mathcal{P}(\mathcal{X}) &\to \mathcal{P}(\mathcal{X}) & \qquad Q : \mathcal{P}(\mathcal{X}) &\to \mathcal{P}(\mathcal{Y}) \\ \pi(dx) &\mapsto \int_{\mathcal{X}} T(dx'|x)\pi(dx) & \pi(dx) &\mapsto \int_{\mathcal{X}} Q(dy|x)\pi(dx) \end{aligned}$$

In practice, the form of the kernel operator is clear via context if the input is a probability measure or an element of the state space. Note that T and Q are time invariant kernels in a POMP as we study.

A POMP is initialized with a state $x_0 \in \mathcal{X}$ drawn from a prior measure μ on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$. However, the state is not available at the observer, instead the observer sees the sequence $Y_n \sim Q(dy|X_n)$. That is, each Y_n is a noisy measurement of the hidden random variable X_n via the measurement channel Q. We then have for any set $A \in \mathcal{B}(\mathcal{X} \times \mathcal{Y})$,

$$P\bigg((X_0, Y_0) \in A\bigg) = \int_A Q(dy|x)\mu(dx) \tag{1.1}$$

and the POMP updates via the transition kernel $T : \mathcal{X} \to \mathcal{P}(\mathcal{X})$

$$P((X_n, Y_n) \in A|(X, Y)_{[0,n-1]} = (x, y)_{[0,n-1]}) = \int_A Q(dy|x_n)T(dx_n|x_{n-1}) \tag{1.2}$$

It follows that $\{(X_n, Y_n)\}_{n=0}^{\infty}$ itself is a Markov chain, and we will denote P^{μ} as the probability measure on $\Omega = \mathcal{X}^{\mathbb{Z}_+} \times \mathcal{Y}^{\mathbb{Z}_+}$, endowed with the product topology where $X_0 \sim \mu$ (this of course means $\omega \in \Omega$ is a sequence of states and measurements $\omega = \{(x_i, y_i)\}_{i=0}^{\infty}$). A diagram of the flow of the POMP is seen in Figure 1.1. The nodes represent random variables, and the arrows are labelled with the kernel that defines the conditional measure between two random variables. That is, the distribution of Y_1 is fully determined by the realization of X_1 and the measurement channel Q, and the distribution of X_2 is fully determined by the realization of X_1 and the transition kernel T.

We now introduce some additional notation that will be useful when dealing with

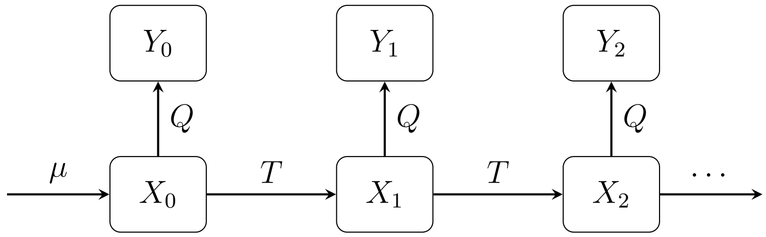

Figure 1.1: Chain of Implications in POMP

sigma fields rather than random variables directly in the context of conditional probabilities or expectations. Strictly speaking, we have a probability measure P^μ on $(\mathcal{X}^{\mathbb{Z}_+} \times \mathcal{Y}^{\mathbb{Z}_+}, \mathcal{B}(\mathcal{X}^{\mathbb{Z}_+} \times \mathcal{Y}^{\mathbb{Z}_+}))$ where the infinite product space is endowed with the product topology, which makes each of the product spaces Polish. We denote by $\mathcal{F}^{\mathcal{X}}_{a,b}$ the sigma field generated by $(X_a, \cdots, X_b)$ and similarly for $\mathcal{Y}$. We also write $\mathcal{F}^{\mathcal{X}}_n$ for the sigma field generated by X_n. We then have $\mathcal{F}^{\mathcal{X}}_{0,\infty} \vee \mathcal{F}^{\mathcal{Y}}_{0,\infty}$ as the sigma field generated by all state and measurement sequences. We note that we will condition on random variables or sigma fields interchangeably in the thesis when most convenient.

When we write $P^\mu(X_{[0,n]} \in \cdot)$ we will be discussing the measure P^μ restricted to the sigma field $\mathcal{F}^{\mathcal{X}}_{0,n}$ which we will denote $P^\mu|_{\mathcal{F}^{\mathcal{X}}_{0,n}}$. Similarly for some set $A \in \mathcal{F}^{\mathcal{X}}_{0,\infty} \vee \mathcal{F}^{\mathcal{Y}}_{0,\infty}$ we write $P^\mu((X_{[0,\infty)}, Y_{[0,\infty)}) \in A | Y_{[0,n]})$ as the conditional measure of P^μ with respect to the sigma field $\mathcal{F}^{\mathcal{Y}}_{0,n}$, which we denote $P^\mu|\mathcal{F}^{\mathcal{Y}}_{0,n}$. We can also consider restricting and conditioning simultaneously and we use $P^\mu(X_n \in \cdot | Y_{[0,n]})$ and $P^\mu|_{\mathcal{F}^{\mathcal{X}}_n}|\mathcal{F}^{\mathcal{Y}}_{0,n}$ to represent the same probability measure.

1.3 The Non-Linear Filter

Definition 1.1. *We define the one step predictor as the sequence of conditional probability measures*

$$\pi^{\mu}_{n-}(\cdot) = P^{\mu}(X_n \in \cdot | Y_{[0,n-1]}) \quad n \in \{1, 2, \cdots\} \tag{1.3}$$

Definition 1.2. *We define the filter as the sequence of conditional probability measures*

$$\pi^{\mu}_{n}(\cdot) = P^{\mu}(X_n \in \cdot | Y_{[0,n]}) \quad n \in \{0, 1, 2, \cdots\} \tag{1.4}$$

Calculating the filter or predictor realizations can be performed in a recursive manner. That is, given the previous filter realization $\pi^{\mu}_{n} \in \mathcal{P}(\mathcal{X})$ and a new observation $y_{n+1} \in \mathcal{Y}$ we can compute the next filter realization π^{μ}_{n+1} via the filter update function $\phi : \mathcal{P}(\mathcal{X}) \times \mathcal{Y} \to \mathcal{P}(\mathcal{X})$.

This update can be quite complicated for a general process, however under a simplifying assumption it is a direct application of Bayes' theorem. Often in the literature it is assumed that the measurement channel Q is non-degenerate. That is, there exists a dominating measure $\lambda \in \mathcal{P}(\mathcal{Y})$ and for every $x \in \mathcal{X}$, $Q(dy|x) \ll \lambda$. Note that "$\ll$" means absolute continuity, so that for any set $A \in \mathcal{B}(\mathcal{Y})$ we have $\lambda(A) = 0 \implies Q(A|x) = 0 \ \forall x \in \mathcal{A}$. Then there exists a conditional probability density function (pdf) with respect to λ called the likelihood function $\frac{dQ}{d\lambda}(x, y) = g(x, y)$. Then we can define the *Bayesian update operator*

$$\psi : \mathcal{P}(\mathcal{X}) \times \mathcal{Y} \to \mathcal{P}(\mathcal{X}) \cup \{0\}$$

$$(\pi(dx), y) \mapsto \begin{cases} \frac{g(x,y)\pi(dx)}{\int_{\mathcal{X}} g(x,y)\pi(dx)} & \text{if } \int_{\mathcal{X}} g(x,y)\pi(dx) > 0 \\ 0 & \text{else} \end{cases}$$

and we can explicitly write the filter update operator as the composition of the Bayesian update operator with the transition kernel

$$\begin{aligned} \pi^{\mu}_{n+1}(dx) &= \phi(\pi^{\mu}_n, y_{n+1})(dx) = \psi(T(\pi^{\mu}_n), y_{n+1})(dx) \\ &= \frac{g(x, y_{n+1}) \int_{\mathcal{X}} T(dx|x')\pi^{\mu}_n(dx')}{\int_{\mathcal{X}} g(x, y_{n+1}) \int_{\mathcal{X}} T(dx|x')\pi^{\mu}_n(dx')} \end{aligned} \tag{1.5}$$

where (1.5) is often referred to as the filter update equation in the literature.

1.4 Filter Stability

Since the filter update is a recursive process, it is sensitive to the initial distribution of X_0 which is the starting point of the recursion. Suppose that an observer computes the non-linear filter assuming that the initial prior is ν, when in reality the prior distribution is μ. The observer receives the measurements and computes the filter π^{ν}_n for each n, but the measurement process is generated according to the true measure μ. The question we are interested in is that of filter stability, namely, if we have two different initial probability measures μ and ν, when do we have that the filter processes π^{μ}_n and π^{ν}_n merge in some appropriate sense as $n \to \infty$.

In the literature, there are a number of merging notions when one considers stability which we enumerate here. Let $C_b(\mathcal{X})$ represent the set of continuous and bounded functions from $\mathcal{X} \to \mathbb{R}$.

Definition 1.3. *Two sequences of probability measures P_n, Q_n merge weakly if* $\forall\ f \in$

$C_b(\mathcal{X})$ *we have* $\lim_{n\to\infty} \left|\int f dP_n - \int f dQ_n\right| = 0$.

Definition 1.4. *For two probability measures* P *and* Q *we define the total variation norm as* $\|P - Q\|_{TV} = \sup_{\|f\|_\infty \le 1} \left|\int f dP - \int f dQ\right|$ *where* f *is assumed measurable and bounded with norm* 1*. We say two sequences of probability measures* P_n*,* Q_n *merge in total variation if* $\|P_n - Q_n\|_{TV} \to 0$ *as* $n \to \infty$*.*

Note that merging in total variation implies weak merging since $C_b(\mathcal{X})$ is a subset of the set of measurable and bounded functions. We will also utilize the information theoretic notion of relative entropy (Kullback-Leibler divergence). Relative entropy is often utilized as a notion of distance between two probability measure as it is non-negative, although it is not a metric since it is not symmetric.

Definition 1.5.

(i) For two probability measures P *and* Q *we define the relative entropy as* $D(P\|Q) = \int \log \frac{dP}{dQ} dP = \int \frac{dP}{dQ} \log \frac{dP}{dQ} dQ$ *where we assume* $P \ll Q$ *and* $\frac{dP}{dQ}$ *denotes the Radon-Nikodym derivative of* P *with respect to* Q*.*

(ii) Let X *and* Y *be two random variables, let* P *and* Q *be two different joint measures for* (X, Y) *with* $P \ll Q$*. Then we define the (conditional) relative entropy between* $P(X|Y)$ *and* $Q(X|Y)$ *as*

$$\begin{aligned} D(P(X|Y)\|Q(X|Y)) &= \int \log\left(\frac{dP_{X|Y}}{dQ_{X|Y}}(x,y)\right) dP(x,y) \\ &= \int \left(\int \log\left(\frac{dP_{X|Y}}{dQ_{X|Y}}(x,y)\right) dP(x|Y=y)\right) dP(y) \quad (1.6) \end{aligned}$$

Some notational discussion is in order. For some probability measures such as $P^\mu(Y_{[0,n]} \in \cdot)$ or $P^\mu(X_n \in \cdot)$, it will be convenient to denote the random variable inside the measure and take out the set argument. When we take the relative entropy of such measures, to make the notation shorter, we will drop the "$\in \cdot$" argument and write $D(P^\mu(Y_{[0,n]})\|P^\nu(Y_{[0,n]}))$.

Note that in a conditional relative entropy, we are integrating the logarithm of the Radon-Nikodym derivative of the conditional measures $P(X|Y)$ and $Q(X|Y)$ over the joint measure of P on (X,Y). The second equality (1.6) shows that this can be thought of as the expectation of the relative entropy $D(P(X|Y=y)\|Q(X|Y=y))$ at specific realizations of $Y = y$, where the expectation is over the marginal measure of P on Y. When we apply this to the filter, π_n^μ and π_n^ν are realizations of the filter for specific measurements, therefore when we discuss their relative entropy, we take the expectation over the marginal of P^μ on $Y_{[0,n]}$. We write this as $E^\mu[D(\pi_n^\mu\|\pi_n^\nu)]$ where $D(\pi_n^\mu\|\pi_n^\nu)$ plays the role of the inner integral in (1.6).

The key relationship between relative entropy and total variation is Pinsker's inequality (see e.g., [34, 14, 27]) which states that for two probability measures P and Q we have that

$$\|P-Q\|_{TV} \leq \sqrt{\frac{2}{\log_2(e)}D(P\|Q)} \tag{1.7}$$

1.4.1 Notions of Stability

Definition 1.6. *A filter process is said to be stable in the sense of weak merging in expectation if for any $f \in C_b(\mathcal{X})$ and any prior ν with $\mu \ll \nu$ we have* $\lim_{n\to\infty} E^\mu\left[\left|\int f d\pi_n^\mu - \int f d\pi_n^\nu\right|\right] = 0$.

Definition 1.7. *A filter process is said to be stable in the sense of weak merging P^μ almost surely (a.s.) if there exists a set of measurement sequences $A \subset \mathcal{Y}^{\mathbb{Z}_+}$ with P^μ*

probability 1 such that for any sequence in A, for any $f \in C_b(\mathcal{X})$ and any prior ν with $\mu \ll \nu$ we have $\lim_{n\to\infty} \left|\int f d\pi_n^\mu - \int f d\pi_n^\nu\right| = 0$.

Definition 1.8. *A filter process is said to be stable in the sense of total variation in expectation if for any measure ν with $\mu \ll \nu$ we have* $\lim_{n\to\infty} E^\mu[\|\pi_n^\mu - \pi_n^\nu\|_{TV}] = 0$.

Definition 1.9. *A filter process is said to be stable in the sense of total variation P^μ a.s. if for any measure ν with $\mu \ll \nu$ we have* $\lim_{n\to\infty} \|\pi_n^\mu - \pi_n^\nu\|_{TV} = 0 \quad P^\mu$ *a.s..*

Definition 1.10. *A filter process is said to be stable in relative entropy if for any measure ν with $\mu \ll \nu$ we have* $\lim_{n\to\infty} E^\mu[D(\pi_n^\mu\|\pi_n^\nu)] = 0$.

Definition 1.11. *Given $f : \mathcal{X} \to \mathbb{R}$ we define the Lipschitz norm*

$$\|f\|_L = \sup\left\{\frac{|f(x) - f(y)|}{d(x,y)}\,\middle|\, d(x,y) \neq 0\right\}$$

With $BLip := \{f : \|f\|_L \leq 1, \|f\|_\infty \leq 1\} \subset C_b(\mathcal{X})$ we define the bounded Lipschitz (BL) metric as $\|P - Q\|_{BL} = \sup_{f\in BLip}\left|\int f dP - \int f dQ\right|$. *A system is then said to be stable in the sense of BL-merging P^μ a.s. if we have* $\|\pi_n^\mu - \pi_n^\nu\|_{BL} \to 0 \quad P^\mu$ *a.s.*

Here we make a cautionary remark about the *merging* of probability measures compared to the *convergence* of a sequence of probability measures to a limit measure. In convergence, we have some sequence P_n and a static limit measure P and we wish to show $P_n \to P$ under some convergence notion. However, in merging we have two sequences P_n and Q_n which may not individually have limits, but come closer together for large n in one of the merging notions defined previously.

The distinction is important. Let us assume that $\mathcal{X}$ is a finite dimensional real space and let $C_0(\mathcal{X})$ denote the space of all continuous functions which decay to zero as $|x| \to \infty$ under the standard supremum norm. The topological dual space of such a space of functions is the set of finite signed measures endowed with total variation [22, Chapter 1] and when the space is compact, merging under the weak-* topology for two sequences of finite measures coincides with the merging notion given in Definition 1.3, that is considering all $C_b(\mathcal{X})$ functions. Likewise, in Definition 1.3, if Q_n were replaced with a single probability measure (i.e. considering converging instead of merging), due to Prokhorov's theorem [6] and resulting tightness, the convergence notions under $C_0(\mathcal{X})$ and $C_b(\mathcal{X})$ would still be equivalent. However, in general both π_n^μ and π_n^ν are time-varying and in this case, as elaborately noted in [15], weak-* merging (that is, considering only $C_0(\mathcal{X})$ functions) is strictly weaker than merging under all $C_b(\mathcal{X})$ functions (Definition 1.3) as the following example reveals:

Example 1.1. *[15, Example 1.1] Consider two sequences of point masses $P_n = \delta_n$ and $Q_n = \delta_{n+\frac{1}{n}}$. These measures merge in the weak-* sense since they both converge to the trivial (all zero) measure in the weak-* sense. However, there exists a continuous and bounded function f such that for large n we have $\int f dP_n = 1$ but $\int f dQ_n = 0$, so P_n and Q_n do not merge in the sense of Definition 1.3.*

From [15] we have that if $\mathcal{X}$ is compact (or if $Q_n = Q$ for a fixed probability measure Q), the merging notions are identical. We note that such subtleties involving merging notions were elaborately investigated by van Handel [39] who focused on merging in the bounded Lipschitz norm, Definition 1.11, which is strictly weaker than Definition 1.7 when the space considered is not compact. In our analysis, we will consider weak merging as opposed to that in the bounded Lipschitz sense.

1.5 Partially Observed Markov Decision Process

1.5.1 Definition of a POMDP

To define a POMDP, begin with the state and measurement spaces from the POMP, $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ and $(\mathcal{Y}, \mathcal{B}(\mathcal{Y}))$, but also add an action space $\mathcal{U}$. Redefine the transition kernel T as

$$\begin{aligned} T :& \mathcal{X} \times \mathcal{U} \to \mathcal{P}(\mathcal{X}) \\ & (x, u) \mapsto T(dx'|x, u) \end{aligned}$$

T defines a regular conditional probability measure on $\mathcal{B}(\mathcal{X})$ given X and U, meaning that for every fixed x and u, $T(\cdot|x, u)$ is a probability measure and for every fixed Borel set $A : T(A|\cdot, \cdot)$ is a measurable function on $\mathcal{X} \times \mathcal{U}$. Then we have for a set $A \in \mathcal{B}(\mathcal{X} \times \mathcal{Y})$

$$P\left((X_0, Y_0) \in A\right) = \int_A Q(dy|x)\mu(dx) \tag{1.8}$$

$$P((X_n, Y_n) \in A|(X, Y, U)_{[0,n-1]} = (x, y, u)_{[0,n-1]}) = \int_A Q(dy|x)T(dx|x_{n-1}, u_{n-1}) \tag{1.9}$$

The DM determines its control actions via an admissible control policy $\gamma = \{\gamma_n\}_{n\geq 0}$ which is a sequence of mappings where each $\gamma_n : \mathcal{Y}^{n+1} \times \mathcal{U}^n \to \mathcal{U}$ maps the past and present measurements $y_{[0,n]}$ and the past control actions $u_{[0,n-1]}$ to a control action u_n such that for every $n \in \mathbb{Z}_+, u_n = \gamma_n(y_{[0,n]}, u_{[0,n-1]})$.

Recursively we see u_0 is a function of y_0 and u_1 is a function of y_0, y_1, and u_0. Yet since u_0 is itself a function of y_0, we have that u_1 is really just a function of

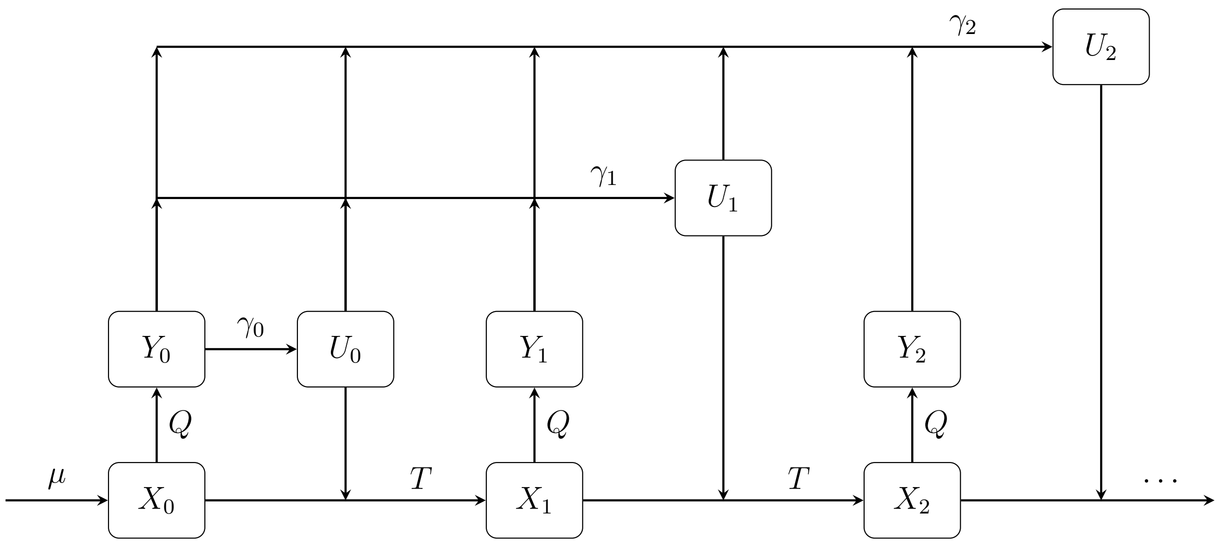

Figure 1.2: Chain of Implications in POMDP

$y_{[0,1]}$. In other words, we can restrict ourselves to considering control policies that are only a function of $y_{[0,n]}$ and not the past control actions $u_{[0,n-1]}$ without any loss of generality. We will denote the collection of such *admissible* control policies as $\gamma \in \Gamma$.

A diagram of the development of the POMDP and the chain of dependence of the random variables is shown in Figure 1.2.

Now, consider the measurable space $\Omega = \mathcal{X}^{\mathbb{Z}_+} \times \mathcal{Y}^{\mathbb{Z}_+}$, endowed with the product topology, (this of course means $\omega \in \Omega$ is a sequence of states and measurements $\omega = \{(x_i, y_i)\}_{i=0}^{\infty}$).

Definition 1.12. *For a fixed initial measure $\mu \in \mathcal{P}(\mathcal{X})$ and a policy $\gamma \in \Gamma$, we define the strategic measure $P^{\mu,\gamma}$ as the probability measure on $(\Omega, \mathcal{F}^X_{[0,\infty)} \vee \mathcal{F}^{\mathcal{Y}}_{[0,\infty)})$ such that*

i) For all $A \in \mathcal{B}(\mathcal{X} \times \mathcal{Y})$ we have

$$P^{\mu,\gamma}\left((X_0, Y_0) \in A\right) = \int_A Q(dy|x)\mu(dx) \tag{1.10}$$

ii) For every $n \geq 1$, for all $A \in \mathcal{B}(\mathcal{X} \times \mathcal{Y})$ let $u_{n-1} = \gamma_{n-1}(y_{[0,n-1]})$ then we have

$$P^{\mu,\gamma}((X_n, Y_n) \in A|(X,Y)_{[0,n-1]} = (x,y)_{[0,n-1]}) = \int_A Q(dy|x)T(dx|x_{n-1}, u_{n-1}) \tag{1.11}$$

Remark 1.1. *Note that $(X,Y)_{[0,\infty}$ is in general not a Markov chain under $P^{\mu,\lambda}$ as u_{n-1} depends on the past measurements in equation (1.11).*

Given a prior $\mu \in \mathcal{P}(\mathcal{X})$ and a policy $\gamma \in \Gamma$ we can then define the filter and predictor for a POMDP using the strategic measure $P^{\mu,\lambda}$.

Definition 1.13. *We define the one step predictor as the sequence of conditional probability measures*

$$\begin{aligned}\pi^{\mu,\gamma}_{n-}(\cdot) &= P^{\mu,\gamma}(X_n \in \cdot|Y_{[0,n-1]}, U_{[0,n-1]} = \gamma_n(Y_{[0,n-1]})) \\ &= P^{\mu,\gamma}(X_n \in \cdot|Y_{[0,n-1]}) \quad n \in \{1,2,\cdots\}\end{aligned}$$

Definition 1.14. *We define the filter as the sequence of conditional probability measures*

$$\begin{aligned}\pi^{\mu,\gamma}_{n}(\cdot) &= P^{\mu,\gamma}(X_n \in \cdot|Y_{[0,n]}, U_{[0,n-1]} = \gamma_n(Y_{[0,n-1]})) \\ &= P^{\mu,\gamma}(X_n \in \cdot|Y_{[0,n]}) \quad n \in \{0,1,2,\cdots\}\end{aligned}$$

Remark 1.2. *Recall that the $U_{[0,n-1]}$ are all functions of the $Y_{[0,n-1]}$, so conditioning on the control actions is not really necessary in the above definitions. Yet this conditional probability is policy dependent, if we condition on the past actions, this* *conditioning is independent of the policy. This distinction will be important when we study filter stability for controlled models.*

Like a POMP, the filter update is recursive, however now the filter update requires a control action as well to determine the transition kernel. However, the control action u_n is a function of $y_{[0,n]}$ which is already known to the filter at time n, therefore the filter also has knowledge of u_n since the control policy is known by the filter. If we assume the measurement channel is non-degenerate with likelihood function $g(x,y)$ then we have

$$\begin{aligned}\pi_{n+1}^{\mu,\gamma}(dx) &= \phi(\pi_n^{\mu,\gamma}, u_n, y_{n+1}) = \psi(T(\pi_n^{\mu}, u_n), y_{n+1})(dx) \\ &= \frac{g(x, y_{n+1}) \int_{\mathcal{X}} T(dx|x', u_n)\pi_n^{\mu}(dx')}{\int_{\mathcal{X}} g(x, y_{n+1}) \int_{\mathcal{X}} T(dx|x', u_n)\pi_n^{\mu}(dx')}\end{aligned} \tag{1.12}$$

Say a prior $\mu \in \mathcal{P}(\mathcal{X})$ and a policy $\gamma \in \Gamma$ are chosen, an observer sees measurements $Y_{[0,\infty)}$ generated via the strategic measure $P^{\mu,\gamma}$. The observer is aware that the policy applied is γ, but incorrectly thinks the prior is $\nu \neq \mu$. The filter stability problem for a POMDP is then concerned with the merging of $\pi_n^{\mu,\gamma}$ and $\pi_n^{\nu,\gamma}$ as n goes to infinity. We must then slightly modify our previous definitions of stability of a POMP to appropriate definitions for a POMDP.

We say a POMDP is stable in any of Definitions 1.6 to 1.10 *with respect to a policy* γ if the definition holds with the measure P^{μ} replaced with the strategic measure $P^{\mu,\gamma}$.

We say the filter process is *universally* stable in the above definitions if it is stable for any control policy $\gamma \in \Gamma$.

1.5.2 Types of Control Problems

A POMDP as described above is a well defined stochastic process, but the DM does not have an objective or purpose. In a stochastic control problem the objective of the DM is to minimize an expected cost. We include a cost function $c : \mathcal{X} \times \mathcal{U} \to \mathbb{R}_+$ which penalizes the DM (or rewards if one wishes to consider a maximization problem) for their actions at each stage of the problem in relation to the state realization. We can then consider three different types of control problems

i) The single stage control cost:

$$J(\mu, \gamma) = E^{\mu,\gamma}[c(X_0, U_0)]$$

ii) The infinite horizon discounted cost problem for some $\beta \in [0, 1)$:

$$J_\beta(\mu, \gamma) = E^{\mu,\gamma}\left[\sum_{n=0}^{\infty} \beta^n c(X_n, U_n)\right]$$

iii) The infinite horizon average cost problem:

$$J_\infty(\mu, \gamma) = \limsup_{N\to\infty} \frac{1}{N} E^{\mu,\gamma}\left[\sum_{n=0}^{N-1} c(X_n, U_n)\right]$$

For each type of problem, we can consider the optimal cost for a given prior μ:

$$J^*(\mu) = \inf_{\gamma\in\Gamma} J(\mu, \gamma)$$

with similar definitions for J^*_β, J^*_∞.

1.5.3 Robustness

Say γ^ν is the optimal control policy with respect to the prior ν, that is $J^*_\beta(\nu) = J_\beta(\nu, \gamma^\nu)$. However, suppose that the true prior is μ, but a controller believes the prior is ν. Then the controller will design the control policy γ^ν and apply it to the system and incur a cost $J_\beta(\mu, \gamma^\nu)$. The robustness problem we consider here studies the difference between $J^*_\beta(\mu)$ and $J_\beta(\mu, \gamma^\nu)$.

1.6 Literature Review

Filter stability is a classical problem and we refer the reader to [10] for a comprehensive review. As discussed in [10], filter stability arises via two separate mechanisms:

1. The transition kernel is in some sense *sufficiently* ergodic, forgetting the initial measure and therefore passing this insensitivity (to incorrect initializations) on to the filter process.

2. The measurement channel provides sufficient information about the underlying state, allowing the filter to track the true state process.

A number of works focus on the first of the two mechanisms, see [20]. By ergodicity here, we mean that the successive applications of the transition kernel T brings to any two different priors closer together with increasing time.. These results usually rely on some form of mixing, pseudo-mixing, or a similar condition on the transition kernel. These results are often paired with a study on the Hilbert metric to deduce filter stability. Results in [8], [13] study the signal to noise ratio to establish sufficient conditions for filter stability. Many results in the literature impose a non-degeneracy condition on the measurement channel, see [10, 39, 17] and utilize the likelihood function $g(x, y)$.

One benefit of studying filter stability through the ergodic properties of the transition kernel is these results often lead to an exponential rate of merging. We discuss such results in Chapter 4. An exponential rate of merging is useful since it allows for an explicit rate of merging to the correct filter; just asymptotic stability may not be strong enough for finite horizon and a class of infinite horizon discounted cost problems.

In Chapters 2 and 3 our primary focus is on the latter of the two mechanisms mentioned above. The question of interest is to find sufficient conditions for "observability": some property of the measurements that implies filter stability, along the same spirit as that of the Kalman Filter in the linear case. Thus, we seek to define a notion of observability for non-linear controlled and control-free stochastic dynamical systems. The method adopted in this paper sees its origins in Chigansky and Liptser [9] and a series of papers by Van Handel [36, 37, 39]. Chigansky and Liptser were not interested in proving full stability, arguing that such results usually rely on ergodicity conditions (i.e. mechanisms of type 1). Instead, they focused on *informative observations* for a specific continuous function f, rather than over all continuous functions in the criterion of weak merging. Nonetheless, [9, Equation 1.7] captures the essence of our definition of one step observability. The idea is to express a continuous function $f(x)$ by integrating a measurable function $g(y)$ over the conditional distribution for Y given $X = x$. That is, consider the functional $S(g)(x) \mapsto \int_{\mathcal{Y}} g(y)Q(dy|x)$. We wish to take a continuous function f and solve for a measurable function g such that $f \approx S(g)$.

A fundamental result which pairs with observability is that of Blackwell and Dubins [5], an implication of which Chigansky and Liptser independently arrived at. Blackwell and Dubins use martingale convergence theorem to show that if P and Q are two measures on a fully observed stochastic process $\{X_n\}_{n=0}^{\infty}$ with $P \ll Q$, then the conditional distributions on the future based on the past merge in total variation P a.s., that is

$$\|P(X_{[n+1,\infty)} \in \cdot | X_{[0,n]}) - Q(X_{[n+1,\infty)} \in \cdot | X_{[0,n]})\|_{TV} \to 0 \quad P \ a.s.$$

In [37], van Handel introduces a definition of observability for POMP. Namely, a system is observable if every prior results in a unique probability measure on the measurement sequences, $P^{\mu}|_{\mathcal{F}^{\mathcal{Y}}_{0,\infty}} = P^{\nu}|_{\mathcal{F}^{\mathcal{Y}}_{0,\infty}} \implies \mu = \nu$. In [39], van Handel extends these results to non-compact state spaces, where *uniform observability* is introduced. Given a uniformity class $\mathcal{G} \subset C_b(\mathcal{X})$, for two measures P, Q define $\|P - Q\|_{\mathcal{G}} = \sup_{g \in \mathcal{G}} \left| \int g dP - \int g dQ \right|$. A filtering model is $\mathcal{G}$-uniformly observable if:

$$\|P^{\mu}(Y_n | Y_{[0,n-1]}) - P^{\nu}(Y_n | Y_{[0,n-1]})\|_{TV} \to 0 \quad P^{\mu} \ a.s. \tag{1.13}$$

$$\implies$$

$$\|\pi_n^{\mu} - \pi_n^{\nu}\|_{\mathcal{G}} \to 0 \quad P^{\mu} a.s. \tag{1.14}$$

If $\mathcal{G}$ is the uniformly bounded Lipschitz functions, the POMP is simply called uniformly observable and $\|\cdot\|_{\mathcal{G}}$ is the bounded Lipschitz distance. This condition is quite difficult to prove in itself, and the question of "informative measurements" is finding sufficient conditions for uniform observability to hold.

The result of Blackwell and Dubins [5] pairs with uniform observability, in that

(1.13) directly follows from Blackwell and Dubins. Then uniform observability would imply filter stability in bounded Lipschitz distance [36]. van Handel proves this in [36], however the author only studied the measurement channel where $Y_n = f(X_n) + Z_n$ where f^{-1} is uniformly continuous and Z_n must have an everywhere non-zero characteristic function (e.g. a Gaussian distribution) and so the results cannot be easily applied to other system models.

For a compact state space, uniform observability and observability are equivalent notions [39]. We also note that for a finite state space with a non-degenerate measurement channel, stability can be fully characterised via observability and a detectability condition [37] , [40, Theorem V.2] or [11, Theorems 2.7 and 3.1].

In this thesis we study a number of different stability notions introduced in Definition 1.6-1.10. Note that observability only implies weak merging almost surely, and for the discrete time case as studied here, observability only implies weak merging of the predictor almost surely, not the filter directly. Methods are then needed to extend observability to imply more stringent notions of stability. A useful tool is the condition discovered by Kunita [28] and derived in full in [10] which states a necessary and sufficient condition for the merging of the filter in total variation in expectation based on comparing the sigma fields $\mathcal{F}^{\mathcal{Y}}_{0,\infty}$ and $\bigcap_{n\geq 0}\mathcal{F}^{\mathcal{X}}_{n,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}$. That is the filter merges in total variation in expectation if and only if:

$$E^{\nu}[\frac{d\mu}{d\nu}(X_0)|\mathcal{F}^{\mathcal{Y}}_{0,\infty}] = E^{\nu}[\frac{d\mu}{d\nu}(X_0)|\bigcap_{n\geq 0}\mathcal{F}^{\mathcal{X}}_{n,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}] \quad P^{\mu}\ a.s. \tag{1.15}$$

Unfortunately, as first observed in [4], Kunita later went on to incorrectly assume that the order of operators could be changed so that:

$$\bigcap_{n\geq 0} \mathcal{F}^{\mathcal{X}}_{n,\infty} \vee \mathcal{F}^{\mathcal{Y}}_{0,\infty} = \mathcal{F}^{\mathcal{Y}}_{0,\infty} \vee \left(\bigcap_{n\geq 0} \mathcal{F}^{\mathcal{X}}_{n,\infty} \right)$$

This however is not true, but this mistake does not affect the earlier insight of Kunita's work nor the results we will use in this thesis.

Relative entropy as a measure of discrepancy between the true filter and the incorrectly initialized filter is studied by Clark, Ocone, and Coumarbatch in [12]. Here they consider the filtering problem in continuous time with the associated non-degeneracy assumptions. The authors establish the relative entropy of the true filter and the incorrect filter as a supermartingale. The paper does not establish convergence to zero, however. A notable setup where actual convergence (of the relative entropy difference) to zero is established is the (rather specific) Beneš filter studied in [33].

Robustness as a general concept does not have a singular definition in the literature. Robustness problems study when the controller has the incorrect system model, or has uncertainty about the specifics of the system model [3]. In the later scenario, many works assume the incorrect model satisfies some relative entropy or total variation bound with the true model. The goal is then to design a controller which works well over all the models in this set of possible system models. The approach taken is to treat the system as a game theory problem where the controller is the minimizer and the uncertainty is the maximizer [41].

However, these problems are different form what we study in that the DM is "aware" they have the wrong system specifications and have some limited idea of what the possible mdoels may be. They design a control policy with this in mind and consciously plan their policy to perform decently well over all the uncertain possible system models. However, in the robustness problem as we study here the DM is blindly applying control as if it has the correct system model. This is similar to the results in [26] where the controller has the wrong prior and acts as if it were the true prior. We in fact use these results in Chapter 5, however these results show that if the priors merge, the robustness difference goes to zero. However, it is not studied how this prior merging is accomplished. In our problem, we start with two disparate priors μ and ν and filter stability brings about the required merging.

1.7 Organization of Thesis

In Chapter 2 we study filter stability in (control-free) POMPs. We provide a new observability definition characterized by the conditional measurement channels $Y_{[n,n+N]}|X_n$ which is a sufficient condition for the weak merging of the predictor. We then provide conditions on the measurement channel and transition kernel to extend to filter stability and stability in total variation and relative entropy.

In Chapter 3 we adapt our control-free results to POMDPs. Here, an important distinction with the control-free setup is that, while the one step observability applies almost identically, the N-step observability does not apply in a policy independent sense. The measurement channel $Y_n|X_n \sim Q(dy|x)$ is unaffected by control actions, hence one step observability still holds. However, due to the past dependence and time variance of control actions the channel $Y_{[n,n+N]}|X_n$ cannot be utilized to achieve stability. However, if stability can be achieved via the one step channel than we provide similar conditions leading to merging in the weak sense, in total variation and in relative entropy.

In Chapter 4 we achieve stability not by observability of the measurement channel, but by studying the filter update as a contraction. Under suitable conditions involving the Dobrushin coefficient of the transition kernel T and measurement kernel Q, we show the filter merges in total variation in expectation with an exponential rate of merging.

In Chapter 5 we apply our stability results to study the cost a DM incurs for utilizing an incorrectly designed control policy compared to an optimal control policy. For the single stage cost problem and the discounted cost problem, under exponential filter stability we can bound the robustness difference. For the average cost problem, under total variation merging in expectation we show the robustness difference is determined solely by the maximum difference of the optimal cost operator $J^*_\infty(\mu)$ under different priors.

1.8 Contributions

A number of studies exist in the literature which define observability, [37], [39], [17]. However, these definitions are either difficult to compute for typical systems or are so specific in their application to one system model that they cannot be easily applied in general. Our definition of observability given in Definition 2.1 is computable for a variety of systems as we show in Section 2.4. More importantly, the definition presented here leads to very general convergence results. Figure 2.1 efficiently summarizes our results in Chapter 2.

Furthermore, in the literature there are various notions of stability that are not always compared directly. One author's notion of filter stability may be different than another's and no easy comparison between the two may be present. By analysing weak, total variation, and relative entropy merging as well as the implications between them we provide a rather unified presentation of filter stability.

In controlled environments, while filter stability is understood for the Kalman filter and its simple recursive structure, on the general non-linear filtering recursions we are not aware of any studies. In particular, research results in the non-linear filtering literature for POMPs are not directly adapted to controlled environments. As we see in Section 3.3, the dependency structure of a POMDP can significantly complicate results that hold in a POMP and thus the adaptation of results is not always possible.

As we review in Chapter 4, most results on exponential filter stability in the literature rely on the Hilbert metric approach and the mixing condition. This is useful as the Hilbert metric is a projective distance and thus the non-linear normalizing term in the filter recursion can be ignored. However, mixing in the sense of the Hilbert metric is a very restrictive assumption to place on the filter update function, and severely limits the applicable system models for this approach. We are the first to establish, to our knowledge, the bound on the Bayesian update utilizing the Dobrushin coefficient, and arrive at conditions that are more widely applicable than mixing.

Implementing an incorrectly designed policy in a true system is studied in [26], however these results show that the robustness cost goes to zero as the priors merge, but not what happens when two disparate priors are implemented. Aside from this result, the general question of the difference in expected cost between an optimal policy and an incorrectly designed policy is not well known. We show that under filter stability, the merging filters act as new priors for a restarted control problem that can be analysed in light of [26]. For the discounted cost problem, we present a bound which depends on the exponential rate of filter stability α, the discount factor

β, and the maximum distance of the optimal cost operator $J^*_\beta(\mu)$ under different priors. For the average cost problem, we show the robustness difference is upper bound by the maximum difference in $J^*_\infty(\mu)$ under different priors.

Note that the filter itself and the measurements $\{(\pi^\mu_n, Y_n)\}_{n=0}^\infty$ can be thought of as a stochastic process taking values in the space of probability measures and the measurement space $(\mathcal{P}(\mathcal{X}) \times \mathcal{Y})$. Under mild conditions, optimal control policies may be taken to be functions of only the filter realization, and under optimal control, $\{(\pi^{\mu,\gamma^\mu}_n, Y_n)\}_{n=0}^\infty$ is a Markov chain [31]. We say the filter process is uniquely ergodic if the transition kernel of this process admits a unique invariant measure that is a probability measure on the space of probability measures $\mathcal{P}(\mathcal{X})$.

If the filter proces is uniquely ergodic, then under additional mild conditions on where the priors may start from $J^*_\infty(\mu) = J^*_\infty(\nu)$ for every such prior [24] and the robustness difference for an average cost problem is 0. Therefore, our robustness result establishes explicitly the relations between filter stability, robustness and unique ergodicity, and hopefully will generate further interest in addressing the problem of unique ergodicity for controlled non-linear filters.

Chapter 2

Filter Stability in Control-Free Environments

2.1 Introduction

In this chapter, we study filter stability in a POMP. As mentioned in the literature review, filter stability may arise, as is the case for the celebrated Kalman Filter [39], due to the informative nature of the measurements Y_n in relation to the hidden process X_n. We show that this informative nature is captured by our definition of observability in Definition 2.1. There exist other notions of observability in the literature, however these are either difficult to verify for most problems, or are so functionally abstract that they lack application expect in specific system structures.

A POMP has a relatively simple dependency structure, and as such conditional probability measures can be manipulated easily in equations due to the conditional Markovian independence of random variables. As we will later see, POMDPs have a more complicated dependency structure that may violate some of the key properties used in the proof of Theorem 2.1. We will revisit these properties later to see what results carry easily from the POMPs to POMDPs and which do not.

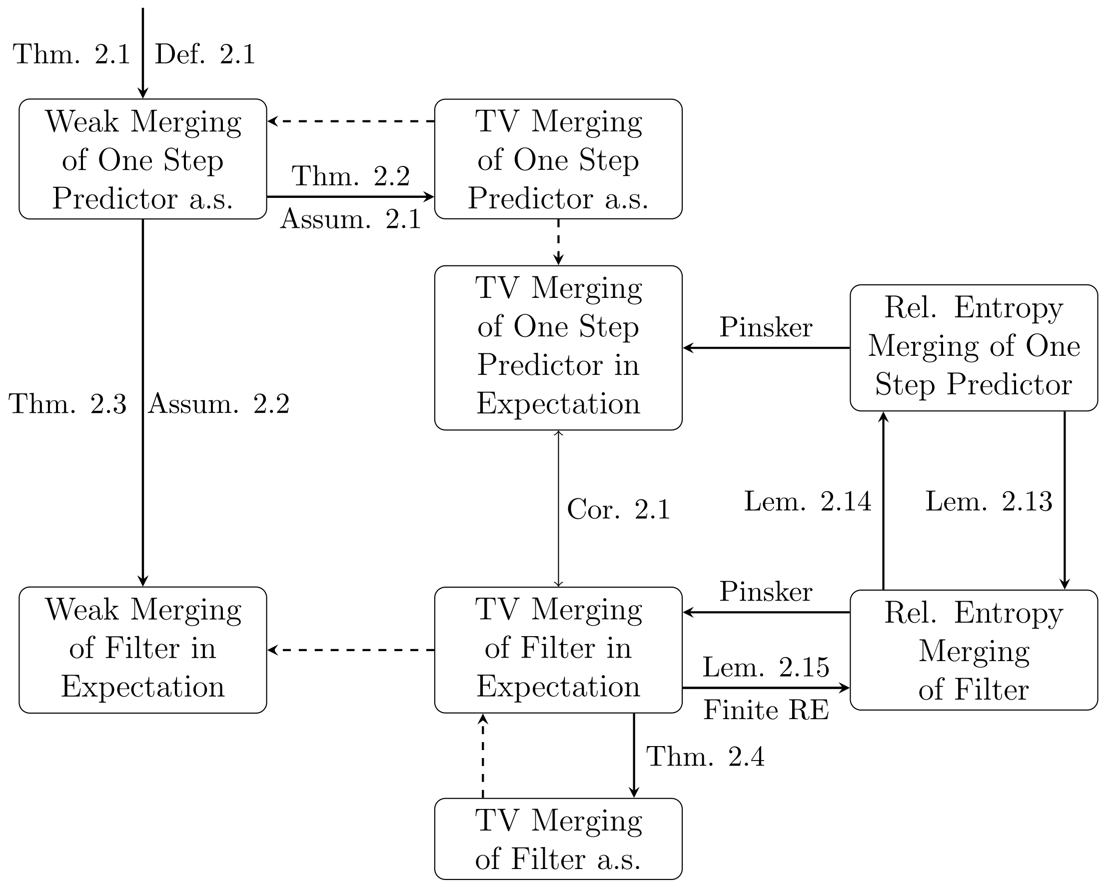

Figure 2.1: Proof Program and Flow of Ideas and Conditions for Filter Stability

Observability and the dependency structure of the POMP result in the weak merging of the predictor almost surely, and not stability of the filter directly. To extend to weak merging of the filter as well as more stringent notions of stability, conditions must be placed on the measurement channel or the transition kernel. A number of such implications are studied throughout the chapter to relate weak merging, total variation, and relative entropy. These are summarized in Figure 2.1. The dashed lines represent implications that are always true, and the solid lines are labelled with the theorems and assumptions that prove the implication.

2.2 Predictor Stability

We define observability as follows:

Definition 2.1.

(i) [One Step Observability] A POMP is said to be one step observable if for every $f \in C_b(\mathcal{X}), \epsilon > 0$, $\exists$ a measurable and bounded function $g : \mathcal{Y} \to \mathbb{R}$ such that

$$\left\| f(\cdot) - \int_{\mathcal{Y}} g(y_1) Q(dy_1 | X_1 = \cdot) \right\|_\infty < \epsilon$$

(ii) [N Step Observability] A POMP is said to be N step observable if for every $f \in C_b(\mathcal{X}), \epsilon > 0$, $\exists$ a measurable and bounded function $g : \mathcal{Y}^N \to \mathbb{R}$ such that

$$\left\| f(\cdot) - \int_{\mathcal{Y}^N} g(y_{[1,N]}) P(dy_{[1,N]} | X_1 = \cdot) \right\|_\infty < \epsilon$$

(iii) [Observability] A POMP is said to be observable if it is N step observable for some finite $N \in \mathbb{N}$.

Remark 2.1. *One step observability is the specific case of N step observability when $N = 1$. However, it is unique in that the distribution of $Y_1 | X_1$ is determined only by the measurement channel Q, whereas the distribution of $Y_{[1,n]} | X_1$ is determined by both the measurement channel Q as well as the transition kernel T. In a POMP this distinction is not very important since $X_{n+1} | X_n \sim T(\cdot | X_n = x_n)$ is the same for every time index n. However once we introduce control, this will no longer hold and in most control problems one step observability is the only workable definition to achieve filter stability.*

Theorem 2.1. *Let $\mu \ll \nu$ and let Definition 2.1 (iii) be satisfied. Then π^{μ}_{n-} and π^{ν}_{n-} merge weakly as $n \to \infty$, P^{μ} a.s.*

Proof. Fix any $f \in C_b(\mathcal{X})$. We wish to show $\forall \epsilon > 0, \exists N$ such that $\forall n > N$, $|\int f d\pi^{\mu}_{n-} - \int f d\pi^{\nu}_{n-}| < \epsilon$. By assumption of the model being N' step observable, we can find some measurable and bounded function g such that

$$\tilde{f}(x) = \int_{\mathcal{Y}^{N'}} g(y_{[1,N']}) P^{\mu}(dy_{[1,N']}|X_1 = x) \qquad \|f - \tilde{f}\|_{\infty} < \frac{\epsilon}{3}$$

Then we have for any $n \in \mathbb{N}$

$$\begin{aligned}\left|\int f d\pi^{\mu}_{n-} - \int f d\pi^{\nu}_{n-}\right| &\leq \left|\int \tilde{f} d\pi^{\mu}_{n-} - \int \tilde{f} d\pi^{\nu}_{n-}\right| + \left|\int (f - \tilde{f}) d\pi^{\mu}_{n-}\right| + \left|\int (f - \tilde{f}) d\pi^{\nu}_{n-}\right| \\ &\leq \left|\int \tilde{f} d\pi^{\mu}_{n-} - \int \tilde{f} d\pi^{\nu}_{n-}\right| + \frac{2}{3}\epsilon\end{aligned}$$

Let us now focus in on the expressions

$$\begin{aligned}\int_{\mathcal{X}} \tilde{f}(x) \pi^{\mu}_{n-}(dx) &= \int_{\mathcal{X}} \int_{\mathcal{Y}^{N'}} g(y_{[1,N']}) P^{\mu}(dy_{[1,N']}|X_1 = x_n) P^{\mu}(dx_n|Y_{[0,n-1]}) \\ \int_{\mathcal{X}} \tilde{f}(x) \pi^{\nu}_{n-}(dx) &= \int_{\mathcal{X}} \int_{\mathcal{Y}^{N'}} g(y_{[1,N']}) P^{\mu}(dy_{[1,N']}|X_1 = x_n) P^{\nu}(dx_n|Y_{[0,n-1]})\end{aligned}$$

Now, the conditional channel $Y_{[n,n+N'-1]}|X_n$ has a few important properties. Firstly,

$$P^{\mu}(Y_{[1,N']} \in \cdot|X_1 = x) = P^{\mu}(Y_{[n,n+N'-1]} \in \cdot|X_n = x)$$

for any $x \in \mathcal{X}$ and any $n \in \mathbb{N}$; this demonstrates that the conditional channel is time invariant. Secondly, the prior is irrelevant conditioned on $X_n = x$ therefore

$$P^\mu(Y_{[n,n+N'-1]} \in \cdot | X_n = x) = P^\nu(Y_{[n,n+N'-1]} \in \cdot | X_n = x)$$

for any $x \in \mathcal{X}$ and any $n \in \mathbb{N}$. Thirdly, $Y_{[n,n+N'-1]} | X_n$ is conditionally independent of $Y_{[0,n-1]}$ therefore

$$P^\mu(Y_{[n,n+N'-1]} \in \cdot | X_n = x) = P^\mu(Y_{[n,n+N'-1]} \in \cdot | X_n = x, Y_{[0,n-1]} = y_{[0,n-1]})$$

for any $x \in \mathcal{X}, y_{[0,n-1]} \in \mathcal{Y}^n$ and any $n \in \mathbb{N}$. Putting these together and applying the chain rule for conditional probability we have

$$\begin{aligned}
\int_{\mathcal{X}} \tilde{f}(x)\pi^\mu_{n-}(dx) &= \int_{\mathcal{X}} \int_{\mathcal{Y}^{N'}} g(y_{[n,n+N'-1]}) P^\mu(dy_{[n,n+N'-1]} | x_n, Y_{[0,n-1]}) P^\mu(dx_n | Y_{[0,n-1]}) \\
&= \int_{\mathcal{Y}^{N'}} g(y_{[n,n+N'-1]}) P^\mu(dy_{[n,n+N'-1]} | Y_{[0,n-1]}) \\
\int_{\mathcal{X}} \tilde{f}(x)\pi^\nu_{n-}(dx) &= \int_{\mathcal{Y}^{N'}} g(y_{[n,n+N'-1]}) P^\nu(dy_{[n,n+N'-1]} | Y_{[0,n-1]})
\end{aligned}$$

By assumption $\mu \ll \nu$ and therefore $P^\mu|_{\mathcal{F}^{\mathcal{Y}}_{0,\infty}} \ll P^\nu|_{\mathcal{F}^{\mathcal{Y}}_{0,\infty}}$. By the result of Blackwell and Dubins [5] we have that $P^\mu(Y_{[n,n+N'-1]} \in \cdot | Y_{[0,n-1]})$ and $P^\nu(Y_{[n,n+N'-1]} \in \cdot | Y_{[0,n-1]})$ merge in total variation P^μ a.s. as $n \to \infty$. Define $\tilde{g} = \frac{g}{\|g\|_\infty}$. Then we can find an N such that for all $n > N$ we have

$$\left| \int_{\mathcal{Y}^{N'}} \tilde{g}(y_{[n,n+N'-1]}) P^\mu(dy_{[n,n+N'-1]} | Y_{[0,n-1]}) - \int_{\mathcal{Y}^{N'}} \tilde{g}(y_{[n,n+N'-1]}) P^\nu(dy_{[n,n+N'-1]} | Y_{[0,n-1]}) \right|$$
$$\leq \frac{\epsilon}{3\|g\|_\infty}$$

therefore for $n > N$ we have

$$\begin{aligned}
&\left|\int_{\mathcal{Y}^{N'}} g(y_{[n,n+N'-1]})P^{\mu}(dy_{[n,n+N'-1]}|Y_{[0,n-1]}) - \int_{\mathcal{Y}^{N'}} g(y_{[n,n+N'-1]})P^{\nu}(dy_{[n,n+N'-1]}|Y_{[0,n-1]})\right| \\
=&\|g\|_\infty \left|\int_{\mathcal{Y}^{N'}} \tilde{g}(y_{[n,n+N'-1]})P^{\mu}(dy_{[n,n+N'-1]}|Y_{[0,n-1]}) - \int_{\mathcal{Y}^{N'}} \tilde{g}(y_{[n,n+N'-1]})P^{\nu}(dy_{[n,n+N'-1]}|Y_{[0,n-1]})\right| \\
\leq&\|g\|_\infty \frac{1}{3\|g\|_\infty}\epsilon = \frac{1}{3}\epsilon
\end{aligned}$$

and therefore for every $\epsilon > 0$, we have an N such that for all $n > N$ we have

$$\left|\int f d\pi^{\mu}_{n-} - \int f d\pi^{\nu}_{n-}\right| \leq \frac{1}{3}\epsilon + \frac{2}{3}\epsilon = \epsilon$$

□

Remark 2.2. *The question may arise of why work with the predictor first instead of considering the stability of the filter directly. One of the key steps is that* $P^{\mu}(Y_{[n,n+N'-1]} \in \cdot|Y_{[0,n-1]})$ *and* $P^{\nu}(Y_{[n,n+N'-1]} \in \cdot|Y_{[0,n-1]})$ *merge in total variation* P^{μ} *a.s. as* $n \to \infty$. *To achieve this in a POMP, we apply the theorem of Blackwell and Dubins to the measurement process* $\{Y_n\}_{n=0}^{\infty}$. *However, the theorem of Blackwell and Dubins is fundamentally about predictive measures of the future given the past, and hence only directly implies stability results for the predictor and not the filter immediately.*

Therefore, informative measurements alone imply the weak merging of the predictor almost surely. However, we would like to consider total variation merging of the predictor, and this will require an assumption on the transition kernel of the underlying Markov process.

Assumption 2.1. *Let our state space* $\mathcal{X}$, *which is a complete, separable, metric space by assumption, have metric* d. *Assume* $T(\cdot|x)$ *is absolutely continuous with* *respect to a dominating σ-finite measure λ measure for every $x \in \mathcal{X}$ and denote the resulting pdf as $t(\cdot|x)$. Further, assume the family $\{t(\cdot|x)\}_{x\in\mathcal{X}}$ is uniformly bounded and equicontinuous. That is for every $x' \in \mathcal{X}$ and every $\epsilon > 0$, we can find a $\delta > 0$ such that if $d(y, x') < \delta$ we have that $|t(y|x) - t(x'|x)| < \epsilon$ for every $x \in \mathcal{X}$.*

Lemma 2.1. *Let there exist some measure $\bar{\mu}$ such that $T(\cdot|x) \ll \bar{\mu}$ for every $x \in \mathcal{X}$. Then we have that $\pi^{\mu}_{n-}, \pi^{\nu}_{n-} \ll \bar{\mu}$ for every $n \in \mathbb{N}$.*

Proof. For all $n \geq 1$ we have

$$\begin{aligned}\pi^{\mu}_{n-}(A) &= \int_{\mathcal{X}} T(A, x)\pi^{\mu}_{n-1}(dx) = \int_{\mathcal{X}}\int_{A} \frac{dT(\cdot|x)}{d\bar{\mu}}(a)\bar{\mu}(da)\pi^{\mu}_{n-1}(dx)\\ &= \int_{A}\left(\int_{\mathcal{X}} \frac{dT(\cdot|x)}{d\bar{\mu}}(a)\pi^{\mu}_{n-1}(dx)\right)\bar{\mu}(da)\end{aligned}$$

where we have applied Fubini's theorem in the final equality. Therefore π^{μ}_{n-} is absolutely continuous with respect to $\bar{\mu}$ for every $n \geq 1$. □

Lemma 2.2. *Let Assumption 2.1 hold and let f^{μ}_{n-} denote the density function of π^{μ}_{n-}. Fix any sequence of measurements $y_{[0,\infty)}$ and denote the collection of probability density functions $\mathscr{F}^{\mu} = \{f^{\mu}_{n-}|n \in \mathbb{N}\}, \mathscr{F}^{\nu} = \{f^{\nu}_{n-}|n \in \mathbb{N}\}$. Then $\mathscr{F}^{\mu}, \mathscr{F}^{\nu}$ are uniformly bounded equicontinuous families.*

Proof. As we see from Lemma 2.1,

$$f^{\mu}_{n-}(x_n) = \frac{d\pi^{\mu}_{n-}}{d\lambda}(x_n) = \int_{\mathcal{X}} t(x_n|x_{n-1})\pi^{\mu}_{n-1}(dx_{n-1})$$

Where t is the pdf of the transition kernel with respect to our dominating measure λ. We require $\forall\epsilon > 0,\ x^* \in \mathcal{X}\ \exists\ \delta > 0$ such that $\forall\ d(x, x^*) < \delta$, $\forall n \in \mathbb{N}$ we have $|f^{\mu}_{n-}(x) - f^{\mu}_{n-}(x^*)| < \epsilon$. By Assumption 2.1, clearly f^{μ}_{n-} is uniformly bounded since t

is uniformly bounded. Then, for any $\epsilon > 0$, $\forall x^* \in \mathcal{X}$ we can find a $\delta > 0$ such that $|t(x_2|x_1) - t(x^*|x_1)| < \epsilon$ when $d(x_2, x^*) < \delta$. Now, assume $d(x_2, x^*) < \delta$, we have

$$\begin{aligned} |f^{\mu}_{n-}(x_2) - f^{\mu}_{n-}(x^*)| &= \left| \int_{\mathcal{X}} t(x_2|x_1) - t(x^*|x_1) \pi^{\mu}_{n-}(dx_1) \right| \\ &\leq \int_{\mathcal{X}} |t(x_2|x_1) - t(x^*|x_1)| \pi^{\mu}_{n-}(dx_1) \leq \epsilon \end{aligned}$$

which proves that $\mathscr{F}^{\mu}$ and $\mathscr{F}^{\nu}$ are uniformly bounded and equicontinuous families.

□

Theorem 2.2. *Let Assumption 2.1 hold. If π^{μ}_{n-} and π^{ν}_{n-} merge weakly P^{μ} a.s., then $\|\pi^{\mu}_{n-} - \pi^{\nu}_{n-}\|_{TV} \to 0$, P^{μ} a.s.*

Proof. By assumption we have a set of measurement sequences $B \subset \mathcal{Y}^{\mathbb{Z}_+}$ with $P^{\mu}(B) = 1$ such that for every measurement sequence in B we have the predictor is stable in the weak sense along this measurement sequence. Choose any $y_{[0,\infty)} \in B$ and fix this measurement sequence for the remainder of the proof. Via Lemma 2.1, and 2.2, $\mathscr{F}^{\mu}$ and $\mathscr{F}^{\nu}$ are uniformly bounded and equicontinuous families. Let $\mathscr{F}^{\mu-\nu} = \{f_n | f_n = f^{\mu}_{n-} - f^{\nu}_{n-}\}$, then the sequence $\{f_n\}_{n=1}^{\infty}$ is a uniformly bounded and equicontinuous class of integrable functions. As in the proof of [30, Lemma 2], now pick a sequence of compact sets $K_j \subset \mathcal{X}$ such that $K_j \subset K_{j+1}$. By the Arzela-Ascoli theorem [35], for any subsequence we can find further subsequences $f_{n^j_k}$ such that

$$\lim_{k\to\infty} \sup_{x\in K_j} |f_{n^j_k}(X) - f^j(x)| = 0$$

for some continuous function $f^j : K_j \to [0, \infty)$. Via the K_j being nested, we can have $\{f_{n^{j+1}_k}\}$ be a subsequence of $\{f_{n^j_k}\}$, and therefore $f^{j+1} = f^j$ over K_j. Then define the function $\tilde{f}$ on $\mathcal{X}$ by $\tilde{f}(x) = f^j(x), x \in K_j$. Using Cantor's diagonal method, we can find an increasing sequence of integers $\{m_i\}$ which is a subsequence of $\{n^j_k\}$ for every j. Therefore

$$\lim_{i\to\infty} f_{m_i}(x) = \tilde{f}(x) \;\; \forall x \in \mathcal{X}$$

and the convergence is uniform over each K_j and $\tilde{f}$ is continuous. Now, f_{m_i} converges weakly to the zero measure by assumption, and via uniform convergence for any Borel set $\mathcal{B}$ we have

$$\int_{\mathcal{B}} f_{m_i}(x)dx \to \int_{\mathcal{B}} \tilde{f}(x)dx,$$

i.e. set-wise convergence. Yet this implies weak convergence, so $\tilde{f} = 0$ almost everywhere, yet $\tilde{f}$ is continuous so it is 0 everywhere.

Now, via Prokhorov's theorem (Theorem 8.6.2 in [6]) we have that $\mathscr{F}^{\mu-\nu}$ is a tight family. Therefore, for every $\epsilon > 0$ we can find a compact set K_ϵ such that

$$|\pi^{\mu}_{n-} - \pi^{\nu}_{n-}|(\mathcal{X} \setminus K_\epsilon) < \epsilon \; \forall \; n \in \mathbb{N}.$$

then we have

$$\begin{aligned}
&\lim_{i\to\infty} \|\pi^{\mu}_{m_i-} - \pi^{\nu}_{m_i-}\|_{TV} \leq \lim_{i\to\infty} |\pi^{\mu}_{m_i-} - \pi^{\nu}_{m_i-}|(\mathcal{X} \setminus K_\epsilon) + |\pi^{\mu}_{m_i-} - \pi^{\nu}_{m_i-}|(K_\epsilon) \\
&\leq \lim_{i\to\infty} \sup_{\|g\|_\infty \leq 1} \left| \int_{K_\epsilon} g(x) f_{m_i}(x) dx \right| + \epsilon \\
&\leq \lim_{i\to\infty} \sup_{\|g\|_\infty \leq 1} \left| \int_{K_\epsilon} g(x)(\tilde{f} - f_{m_i})(x) dx \right| + \left| \int_{K_\epsilon} g(x)\tilde{f}(x) dx \right| + \epsilon \\
&\leq \lim_{i\to\infty} \|\tilde{f} - f_{m_i}\|_\infty \lambda(K_\epsilon) + \epsilon
\end{aligned}$$

since we have already argued $\tilde{f} = 0$. Now, over the compact set K_ϵ, f_{m_i} converges to $\tilde{f}$ uniformly, therefore $\exists N$ such that $\forall k > N$, $\|\tilde{f} - f_{m_k}\|_\infty < \frac{\epsilon}{\lambda(K_\epsilon)}$. We then conclude that

$$\lim_{i\to\infty} \|\pi^{\mu}_{m_i-} - \pi^{\nu}_{m_i-}\|_{TV} = 0$$

Thus, for every subsequence of $\{f_n\}_{n=1}^{\infty}$, we can find a subsequence that converges in total variation, which implies that the original sequence converges in total variation.

□

2.3 Filter Stability

As we saw in Theorem 2.1, observability alone only results in predictor stability. Some assumptions and further analysis is required to extend the predictor stability results to the filter itself.

2.3.1 Weak Merging

Here we will utilize results from [25]. This paper was concerned with a different topic than filter stability, namely the weak Feller property of the "filter update" kernel. That is, one can view the filter π^{μ}_{n} and the measurement Y_n as its own Markov chain $\{(\pi^{\mu}_{n}, Y_n)\}_{n=0}^{\infty}$ which takes values in $\mathcal{P}(\mathcal{X}) \times \mathcal{Y}$. The filter update kernel is the transition kernel of this Markov chain. We will not study this kernel, but some of the analysis in [25] is useful in providing concise arguments to connect the filter to the predictor.

Assumption 2.2. *The measurement channel Q is continuous in total variation. That is, for any sequence $a_n \to a \in \mathcal{X}$ we have $\|Q(\cdot|a_n) - Q(\cdot|a)\|_{TV} \to 0$ or in other words $\|P(Y_0 \in \cdot|X_0 = a_n) - P(Y_0 \in \cdot|X_0 = a)\|_{TV} \to 0$.*

Theorem 2.3. *Let Assumption 2.2 hold. If the predictor merges weakly P^μ a.s., then the filter merges weakly in expectation.*

Proof. Begin by assuming that the predictor merges weakly almost surely. As is argued in [25], one can view the filter π_n^μ as a function of π_{n-1}^μ (the previous filter) and the current observation $Y_n = y_n$, that is $\pi_n^\mu = \phi(\pi_{n-1}^\mu, y_n)$. Pick any continuous and bounded function f, we have

$$\begin{aligned} &E^\mu\left[\left|\int_{\mathcal{X}} f(x)\pi_n^\mu(dx) - \int_X f(x)\pi_n^\nu(dx)\right|\right] \\ =&E^\mu\left[E^\mu\left[\left|\int_{\mathcal{X}} f(x)\phi(\pi_{n-1}^\mu, y_n)(dx) - \int_X f(x)\phi(\pi_{n-1}^\nu, y_n)(dx)\right| Y_{[0,n-1]}\right]\right] \end{aligned} \tag{2.1}$$

Now, define the set $I^+(y_{[0,n-1]}) \subset \mathcal{Y}$ as:

$$I^+(y_{[0,n-1]}) = \left\{y_n \in \mathcal{Y} \,\middle|\, \int_{\mathcal{X}} f(x)\phi(\pi_{n-1}^\mu, y_n)(dx) > \int_X f(x)\phi(\pi_{n-1}^\nu, y_n)(dx)\right\}$$

where the argument $y_{[0,n-1]}$ is the sequence on which the previous filters π_{n-1}^μ and π_{n-1}^ν are realized. Define the complement of this set as $I^-(y_{[0,n-1]})$. Then for every fixed realization $y_{[0,n-1]}$ we can break the inner expectation (which is an integral) into two parts and with the appropriate sign, drop the absolute value showing the inner conditional expectation of (2.1) is equivalent to

$$E^\mu\left[1_{I^+(y_{[0,n-1]})}\left(\int_{\mathcal{X}} f(x)\phi(\pi_{n-1}^\mu, y_n)(dx) - \int_{\mathcal{X}} f(x)\phi(\pi_{n-1}^\nu, y_n)(dx)\right)\middle| Y_{[0,n-1]}\right] \tag{2.2}$$

$$-E^\mu\left[1_{I^-(y_{[0,n-1]})}\left(\int_{\mathcal{X}} f(x)\phi(\pi_{n-1}^\mu, y_n)(dx) - \int_{\mathcal{X}} f(x)\phi(\pi_{n-1}^\nu, y_n)(dx)\right)\middle| Y_{[0,n-1]}\right] \tag{2.3}$$

Let us focus for now in the term (2.2). We can add and subtract

$$E^{\nu}[1_{I^+(y_{[0,n-1]})}\int_{\mathcal{X}} f(x)\phi(\pi^{\nu}_{n-1},y_n)(dx)|Y_{[0,n-1]}] \tag{2.4}$$

to (2.2) and we have

$$\begin{aligned}
&E^{\mu}\left[1_{I^+(y_{[0,n-1]})}\int_{\mathcal{X}} f(x)\phi(\pi^{\mu}_{n-1},y_n)(dx)\middle| Y_{[0,n-1]}\right]\\
-&E^{\nu}\left[1_{I^+(y_{[0,n-1]})}\int_{\mathcal{X}} f(x)\phi(\pi^{\nu}_{n-1},y_n)(dx)\middle| Y_{[0,n-1]}\right]\\
+&E^{\nu}\left[1_{I^+(y_{[0,n-1]})}\int_{\mathcal{X}} f(x)\phi(\pi^{\nu}_{n-1},y_n)(dx)\middle| Y_{[0,n-1]}\right]\\
-&E^{\mu}\left[1_{I^+(y_{[0,n-1]})}\int_{\mathcal{X}} f(x)\phi(\pi^{\nu}_{n-1},y_n)(dx)\middle| Y_{[0,n-1]}\right]
\end{aligned}$$

terms 3 and 4 have the same inner argument, which is a bounded and measurable function of Y_n. Therefore the difference of these terms is upper bound by $\|f\|_\infty \|P^{\mu}(Y_n|Y_{[0,n-1]}) - P^{\nu}(Y_n|Y_{[0,n-1]})\|_{TV}$ which decays to zero by Blackwell Dubins [5].

For the first two terms, the prior measure in the conditional expectation (i.e. E^{μ} or E^{ν}) and the filter argument in ϕ (i.e π^{μ}_{n-1} or π^{ν}_{n-1}) agree within each term, hence we can apply [25, Equation 4]. We rewrite the difference of these as:

$$\int_{\mathcal{X}} f(x)Q(I^+(y_{[0,n-1]})|x)\pi^{\nu}_{n-}(dx) - \int_{\mathcal{X}} f(x)Q(I^+(y_{[0,n-1]})|x)\pi^{\mu}_{n-}(dx) \tag{2.5}$$

where $f(\cdot)Q(I^+(y_{[0,n-1]}|\cdot) : \mathcal{X} \to \mathbb{R}$. We can then consider the family of functions $\mathcal{F} = \{f(\cdot)Q(I^+(y_{[0,n-1]}|\cdot)\}$ indexed by an integer n and an infinite sequence $y_{[0,\infty)}$. The family is uniformly bounded by $\|f\|_\infty < \infty$, and by assumption 2.2 the family is equicontinuous. Then (2.5) is less than

$$\sup_{\tilde{f}\in\mathcal{F}}\left|\int_x \tilde{f}(x)\pi^{\mu}_{n-}(dx) - \int_X \tilde{f}(x)\pi^{\nu}_{n-}(dx)\right|$$

by [18, Corollary 11.3.4] we have that the above goes to zero as $n \to \infty$.

Therefore the limit of (2.2) is zero, and by a similar argument that same can be said for (2.3). Both are upper bound by $\|f\|_\infty \leq \infty$ hence dominated convergence theorem can be applied to state that the limit of (2.1) is 0. □

Therefore, with a small continuity assumption on the transition kernel, we can extend the weak merging of the predictor to the filter. For total variation merging, we will discover that the merging of the filter and predictor are actually equivalent. We will demonstrate this by studying the structure of the Radon Nikodym derivatives and connections with different limiting sigma fields.

2.3.2 Radon Nikodym Derivatives For the True and False Measures

Up to this point, we have established the total variation merging of the predictor a.s. and the weak merging of the filter in expectation. However, we would like to consider more stringent notions of stability for the filter, as well as stability in relative entropy for both the predictor and filter. Under different assumptions specific results can be developed. For example, [36, Lemma 4.2] establishes the total variation merging of the filter in expectation from that of the predictor using non-degeneracy. However, by examining the form of the Radon Nikodym derivative of P^μ and P^ν restricted and conditioned on different sigma fields, we can gain significant insight into how these different notions of stability relate to one another. These results are inspired as a generalization of Lemma 5.6 and Corollary 5.7 in [38], or a similar derivation in the introduction of [10], which establish the specific form of $\frac{d\pi_n^\mu}{d\pi_n^\nu}$.

Lemma 2.3. *Assume $\mu \ll \nu$. For any sigma field $\mathcal{G} \subseteq \mathcal{F}^{\mathcal{X}}_{0,\infty} \vee \mathcal{F}^{\mathcal{Y}}_{0,\infty}$ we have:*

$$\frac{dP^\mu|_{\mathcal{G}}}{dP^\nu|_{\mathcal{G}}} = E^\nu\left[\frac{d\mu}{d\nu}(X_0)\middle|\mathcal{G}\right] \quad P^\mu \ a.s.$$

Proof. Begin with the largest sigma field, $\mathcal{G} = \mathcal{F}^{\mathcal{X}}_{0,\infty} \vee \mathcal{F}^{\mathcal{Y}}_{0,\infty}$. Pick any $A \in \mathcal{F}^{\mathcal{X}}_{0,\infty} \vee \mathcal{F}^{\mathcal{Y}}_{0,\infty}$ we have

$$\begin{aligned} P^\mu((X_{[0,\infty)}, Y_{[0,\infty)}) \in A) = E^\mu\left[1_A\right] = E^\mu[E^\mu[1_A|\mathcal{F}^{\mathcal{X}}_0]] = \int E^\mu[1_A|X_0 = x_0]P^\mu(dx_0) \\ = \int E^\mu[1_A|X_0 = x_0]\mu(dx_0) \end{aligned}$$

Now, conditioned on $X_0 = x_0$ the prior is irrelevant, therefore $E^\mu[1_A|X_0 = x_0] = E^\nu[1_A|X_0 = x_0]$ and we have:

$$\begin{aligned} \int E^\nu[1_A|X_0 = x_0]\frac{d\mu}{d\nu}(x_0)\nu(dx_0) = \int E^\nu[1_A|X_0 = x_0]\frac{d\mu}{d\nu}(x_0)P^\nu(dx_0) \\ = E^\nu[E^\nu[1_A|\mathcal{F}^{\mathcal{X}}_0]\frac{d\mu}{d\nu}(X_0)] = E^\nu[E^\nu[\frac{d\mu}{d\nu}(X_0)1_A|\mathcal{F}^{\mathcal{X}}_0]] = E^\nu[1_A\frac{d\mu}{d\nu}(X_0)] \end{aligned}$$

where $\frac{d\mu}{d\nu}(X_0)$ is $\mathcal{F}^{\mathcal{X}}_0$ measurable so we can move it inside the conditional expectation. It follows that

$$\frac{dP^\mu|_{\mathcal{F}^{\mathcal{X}}_{0,\infty} \vee \mathcal{F}^{\mathcal{Y}}_{0,\infty}}}{dP^\nu|_{\mathcal{F}^{\mathcal{X}}_{0,\infty} \vee \mathcal{F}^{\mathcal{Y}}_{0,\infty}}} = \frac{d\mu}{d\nu}(X_0) = E^\nu\left[\frac{d\mu}{d\nu}(X_0)\middle|\mathcal{F}^{\mathcal{X}}_{0,\infty} \vee \mathcal{F}^{\mathcal{Y}}_{0,\infty}\right] \quad P^\mu \ a.s.$$

Now pick some other field $\mathcal{G} \subset \mathcal{F}^{\mathcal{X}}_{0,\infty} \vee \mathcal{F}^{\mathcal{Y}}_{0,\infty}$, pick $A \in \mathcal{G}$ we have:

$$P^\mu((X_0^\infty, Y_0^\infty) \in A) = E^\mu[1_A] = E^\nu[\frac{d\mu}{d\nu}(X_0)1_A] = E^\nu[E^\nu[\frac{d\mu}{d\nu}(X_0)1_A|\mathcal{G}]] = E^\nu[1_A E^\nu[\frac{d\mu}{d\nu}(X_0)|\mathcal{G}]]$$

Since 1_A is $\mathcal{G}$ measurable. It follows that

$$\frac{dP^\mu|_\mathcal{G}}{dP^\nu|_\mathcal{G}} = E^\nu\left[\frac{d\mu}{d\nu}(X_0)\middle|\mathcal{G}\right] \quad P^\mu \ a.s.$$

□

Lemma 2.4. *Assume $\mu \ll \nu$. For any two sigma fields $\mathcal{G}_1, \mathcal{G}_2 \subset \mathcal{F}^\mathcal{X}_{0,\infty} \vee \mathcal{F}^\mathcal{Y}_{0,\infty}$, let $P^\mu|_{\mathcal{G}_1}|\mathcal{G}_2$ represent the probability measure P^μ restricted to $\mathcal{G}_1$, conditioned on field $\mathcal{G}_2$. We then have*

$$\frac{dP^\mu|_{\mathcal{G}_1}|\mathcal{G}_2}{dP^\nu|_{\mathcal{G}_1}|\mathcal{G}_2} = \frac{E^\nu[\frac{d\mu}{d\nu}(X_0)|\mathcal{G}_1 \vee \mathcal{G}_2]}{E^\nu[\frac{d\mu}{d\nu}(X_0)|\mathcal{G}_2]} \quad P^\mu \ a.s.$$

Proof. For any set $A \in \mathcal{G}_1$ we have:

$$\begin{aligned} P^\mu((X_{[0,\infty)}, Y_{[0,\infty)}) \in A) = E^\mu[1_A] = E^\mu[E^\mu[1_A|\mathcal{G}_2]] &= E^\nu[E^\mu[1_A|\mathcal{G}_2]\frac{dP^\mu|_{\mathcal{G}_2}}{dP^\nu|_{\mathcal{G}_2}}] \\ &= E^\nu[E^\mu[1_A\frac{dP^\mu|_{\mathcal{G}_2}}{dP^\nu|_{\mathcal{G}_2}}|\mathcal{G}_2]] \\ &= E^\nu[E^\mu[1_A E^\nu[\frac{d\mu}{d\nu}(X_0)|\mathcal{G}_2]|\mathcal{G}_2]] \end{aligned} \tag{2.6}$$

where we can move $\frac{dP^\mu|_{\mathcal{G}_2}}{dP^\nu|_{\mathcal{G}_2}}$ in between expectations since it is $\mathcal{G}_2$ measurable and we have applied Lemma 2.3 in the final equality.

The Radon Nikodym derivative $\frac{dP^\mu|_{\mathcal{G}_1}|\mathcal{G}_2}{dP^\nu|_{\mathcal{G}_1}|\mathcal{G}_2}$ is then the unique function (up to dominating sets of measure 0) f such that

$$E^\nu[E^\mu[1_A E^\nu[\frac{d\mu}{d\nu}(X_0)|\mathcal{G}_2]|\mathcal{G}_2]] = E^\nu[E^\nu[1_A f E^\nu[\frac{d\mu}{d\nu}(X_0)|\mathcal{G}_2]|\mathcal{G}_2]] \tag{2.7}$$

we claim that this will be satisfied by $f = \frac{E^\nu[\frac{d\mu}{d\nu}(X_0)|\mathcal{G}_1 \vee \mathcal{G}_2]}{E^\nu[\frac{d\mu}{d\nu}(X_0)|\mathcal{G}_2]}$ (note the denominator is positive P^μ *a.s.*). This claim is proven here

$$E^\nu[E^\nu[1_A f E^\nu[\frac{d\mu}{d\nu}(X_0)|\mathcal{G}_2]|\mathcal{G}_2]] = E^\nu[E^\nu[1_A \left(\frac{E^\nu[\frac{d\mu}{d\nu}(X_0)|\mathcal{G}_1 \vee \mathcal{G}_2]}{E^\nu[\frac{d\mu}{d\nu}(X_0)|\mathcal{G}_2]}\right) E^\nu[\frac{d\mu}{d\nu}(X_0)|\mathcal{G}_2]|\mathcal{G}_2]]$$
$$= E^\nu[E^\nu[1_A E^\nu[\frac{d\mu}{d\nu}(X_0)|\mathcal{G}_1 \vee \mathcal{G}_2]|\mathcal{G}_2]]$$
$$= E^\nu[E^\nu[E^\nu[1_A \frac{d\mu}{d\nu}(X_0)|\mathcal{G}_1 \vee \mathcal{G}_2]|\mathcal{G}_2]]$$

where we can move 1_A inside the expectation since it is $\mathcal{G}_1$ and hence $\mathcal{G}_1 \vee \mathcal{G}_2$ measurable. Then $\mathcal{G}_2$ is a sub-field of $\mathcal{G}_1 \vee \mathcal{G}_2$ and we can apply the law of iterated expectations

$$E^\nu[E^\nu[E^\nu[1_A \frac{d\mu}{d\nu}(X_0)|\mathcal{G}_1 \vee \mathcal{G}_2]|\mathcal{G}_2]] = E^\nu[1_A \frac{d\mu}{d\nu}(X_0)] = E^\mu[1_A]$$

□

Lemma 2.5. *For any two sigma fields $\mathcal{G}_1, \mathcal{G}_2 \subset \mathcal{F}^{\mathcal{X}}_{0,\infty} \vee \mathcal{F}^{\mathcal{Y}}_{0,\infty}$ we have:*

$$\|P^\mu|_{\mathcal{G}_1}|\mathcal{G}_2 - P^\nu|_{\mathcal{G}_1}|\mathcal{G}_2\|_{TV} = \frac{E^\nu\left[\ \left|E^\nu\left[\frac{d\mu}{d\nu}(X_0)|\mathcal{G}_1 \vee \mathcal{G}_2\right] - E^\nu\left[\frac{d\mu}{d\nu}(X_0)|\mathcal{G}_2\right]\right|\ \middle|\mathcal{G}_2\right]}{E^\nu\left[\frac{d\mu}{d\nu}(X_0)|\mathcal{G}_2\right]} \qquad P^\mu \ a.s.$$

Proof. An equivalent way to express total variation as opposed to that presented in Definition 1.4 is

$$\|P^\mu|_{\mathcal{G}_1}|\mathcal{G}_2 - P^\nu|_{\mathcal{G}_1}|\mathcal{G}_2\|_{TV} = \int \left|\frac{dP^\mu|_{\mathcal{G}_1}|\mathcal{G}_2}{dP^\nu|_{\mathcal{G}_1}|\mathcal{G}_2} - 1\right| dP^\nu|_{\mathcal{G}_1}|\mathcal{G}_2$$
$$= \int \left|\frac{E^\nu[\frac{d\mu}{d\nu}(X_0)|\mathcal{G}_1 \vee \mathcal{G}_2]}{E^\nu[\frac{d\mu}{d\nu}(X_0)|\mathcal{G}_2]} - 1\right| dP^\nu|_{\mathcal{G}_1}|\mathcal{G}_2$$

via Lemma 2.4. We can then cross multiply which completes the proof. □

For the specific case of the non-linear filter, that is $\mathcal{G}_1 = \mathcal{F}^{\mathcal{X}}_n$ and $\mathcal{G}_2 = \mathcal{F}^{\mathcal{Y}}_{0,n}$, the results presented above imply the following known results in the literature.

Lemma 2.6. *[38, Lemma 5.6] Assume $\mu \ll \nu$. Then we have that $\pi^\mu_n \ll \pi^\nu_n$ a.s. and we have*

$$\frac{d\pi^\mu_n}{d\pi^\nu_n}(x) = \frac{E^\nu[\frac{d\mu}{d\nu}(X_0)|Y_{[0,n]}, X_n = x]}{E^\nu[\frac{d\mu}{d\nu}(X_0)|Y_{[0,n]}]} \qquad P^\mu \text{ a.s.} \tag{2.8}$$

Lemma 2.7. *[38, Corollary 5.7] Assume $\mu \ll \gamma$ for some measure γ, then we can express*

$$\|\pi^\mu_n - \pi^\gamma_n\|_{TV} = \frac{E^\gamma\left[\left|E^\gamma[\frac{d\mu}{d\gamma}(X_0)|Y_{[0,\infty)}, X_{[n,\infty)}] - E^\gamma[\frac{d\mu}{d\gamma}(X_0)|Y_{[0,n]}]\right|\middle| Y_{[0,n]}\right]}{E^\gamma\left[\frac{d\mu}{d\gamma}(X_0)\middle| Y_{[0,n]}\right]} \tag{2.9}$$

Lemma 2.8. *[10, Equation 1.10] The filter merges in total variation in expectation if an only if*

$$E^\nu\left[\frac{d\mu}{d\nu}(X_0)\middle|\bigcap_{n\geq 0}\mathcal{F}^{\mathcal{Y}}_{0,\infty}\vee\mathcal{F}^{\mathcal{X}}_{n,\infty}\right] = E^\nu\left[\frac{d\mu}{d\nu}(X_0)\middle| F^{\mathcal{Y}}_{0,\infty}\right] \;\; P^\nu \text{ a.s.} \tag{2.10}$$

In the following sections, we will build on Lemmas 2.3-2.5 to obtain stability results for the predictor as well as stability results under the relative entropy criterion.

2.3.3 Total Variation Merging

Since our results apply to any general sigma field, not just the fields used in the analysis of the filter, we can analyse the predictor to establish Lemmas 2.9, 2.10, and 2.11. We then conclude in Corollary 2.1 that the total variation merging of the predictor in expectation is equivalent to that of the filter.

Lemma 2.9. *Assume $\mu \ll \nu$. Then we have that $\pi^{\mu}_{n-} \ll \pi^{\mu}_{n-}$ P^{μ} a.s. and we have*

$$\frac{d\pi^{\mu}_{n-}}{d\pi^{\nu}_{n-}}(x) = \frac{E^{\nu}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,n-1]}, X_n = x]}{E^{\nu}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,n-1]}]} \qquad P^{\mu} \ a.s. \tag{2.11}$$

Proof. These results become clear from Lemma 2.4 when we state the predictor as P^{μ} restricted to $\mathcal{F}^{\mathcal{X}}_n$ conditioned on $\mathcal{F}^{\mathcal{Y}}_{0,n-1}$. □

Lemma 2.10. *Assume $\mu \ll \gamma$ for some measure γ, then we can express*

$$\|\pi^{\mu}_{n-} - \pi^{\gamma}_{n-}\|_{TV} = \frac{E^{\gamma}\left[\left|E^{\gamma}[\frac{d\mu}{d\gamma}(X_0)|Y_{[0,\infty)}, X_{[n,\infty)}] - E^{\gamma}[\frac{d\mu}{d\gamma}(X_0)|Y_{[0,n-1]}]\right|\,\middle|\, Y_{[0,n-1]}\right]}{E^{\gamma}\left[\frac{d\mu}{d\gamma}(X_0)\,\middle|\, Y_{[0,n-1]}\right]} \tag{2.12}$$

Proof. By Lemma 2.5 we can write

$$\|\pi^{\mu}_{n-} - \pi^{\gamma}_{n-}\|_{TV} = \frac{E^{\gamma}\left[|E^{\gamma}[\frac{d\mu}{d\gamma}(X_0)|Y_{[0,n-1]}, X_n] - E^{\gamma}[\frac{d\mu}{d\gamma}(X_0)|Y_{[0,n-1]}]|\,\middle|\, Y_{[0,n-1]}\right]}{E^{\gamma}\left[\frac{d\mu}{d\gamma}(X_0)\,\middle|\, Y_{[0,n-1]}\right]}$$

Since Y_n is a random function of X_n, we have that $\sigma(Y_{[0,n-1]}, X_n) = \sigma(Y_{[0,n]}, X_n)$. Further, by the Markov property we have that we have that $(X_{[0,n-1]}, Y_{[0,n-1]})$ are independent of $(X_{[n+1,\infty)}, Y_{[n+1,\infty)})$ conditioned on (X_n, Y_n) therefore we can state

$$E^{\gamma}[\frac{d\mu}{d\gamma}(X_0)|Y_{[0,n-1]}, X_n] = E^{\gamma}[\frac{d\mu}{d\gamma}(X_0)|Y_{[0,\infty)}, X_{[n,\infty)}]$$

□

Lemma 2.11. *The predictor merges in total variation in expectation if and only if*

$$E^{\nu}\left[\frac{d\mu}{d\nu}(X_0)\,\middle|\, \bigcap_{n\geq 1} \mathcal{F}^{\mathcal{Y}}_{0,\infty} \vee \mathcal{F}^{\mathcal{X}}_{n,\infty}\right] = E^{\nu}\left[\frac{d\mu}{d\nu}(X_0)\,\middle|\, F^{\mathcal{Y}}_{0,\infty}\right] \ P^{\nu} \ a.s. \tag{2.13}$$

Proof.

$$E^{\mu}\left[\|\pi^{\mu}_{n-}-\pi^{\nu}_{n-}\|_{TV}\right]=E^{\nu}\left[\frac{dP^{\mu}|_{\mathcal{F}^{\mathcal{Y}}_{0,n-1}}}{dP^{\nu}|_{\mathcal{F}^{\mathcal{Y}}_{0,n-1}}}\|\pi^{\mu}_{n-}-\pi^{\nu}_{n-}\|_{TV}\right]$$
$$=E^{\nu}\left[E^{\nu}\left[\frac{d\mu}{d\nu}(X_0)\middle|Y_{[0,n-1]}\right]\|\pi^{\mu}_{n-}-\pi^{\nu}_{n-}\|_{TV}\right]$$
$$=E^{\nu}\left[E^{\nu}\left[|E^{\nu}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,\infty)},X_{[n,\infty)}]-E^{\nu}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,n-1]}]|\middle|Y_{[0,n-1]}\right]\right]$$
$$=E^{\nu}\left[|E^{\nu}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,\infty)},X_{[n,\infty)}]-E^{\nu}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,n-1]}]|\right]$$

We then see that $A_n=E^{\nu}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,n-1]}]$ is a non-negative uniformly integrable martingale (with respect to the measure P^{ν}) adapted to the increasing filtration $\mathcal{F}^{\mathcal{Y}}_{0,n-1}$. Hence the limit as $n\to\infty$ in $L^1(P^{\nu})$ is $E^{\nu}[\frac{d\mu}{d\nu}(X_0)|\mathcal{F}^{\mathcal{Y}}_{0,\infty}]$. Similarly, we can view $B_n=E^{\nu}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,\infty)},X_{[n,\infty)}]$ as a backwards non-negative uniformly integrable martingale (with respect to the measure P^{ν}) adapted to the decreasing sequence of filtrations $\mathcal{F}^{\mathcal{Y}}_{0,\infty}\vee\mathcal{F}^{\mathcal{X}}_{n,\infty}$. Then by the backwards martingale convergence theorem, the limit as $n\to\infty$ in $L^1(P^{\nu})$ is $E^{\nu}[\frac{d\mu}{d\nu}(X_0)|\bigcap_{n=1}^{\infty}\mathcal{F}^{\mathcal{Y}}_{1,\infty}\vee\mathcal{F}^{\mathcal{X}}_{n,\infty}]$. It is then clear the the total variation in expectation is zero if and only if equation (2.13) holds. □

Corollary 2.1. *The filter merges in total variation in expectation if and only if the predictor merges in total variation in expectation.*

Proof. The sigma fields $\mathcal{F}^{\mathcal{X}}_{n,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}$ are a decreasing sequence, that is $\mathcal{F}^{\mathcal{X}}_{n+1,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}\subset\mathcal{F}^{\mathcal{X}}_{n,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}$. Therefore, when we take their intersection, removing the first or largest sigma field $\mathcal{F}^{\mathcal{X}}_{0,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}$ from the intersection of a deceasing set of sigma fields does not change the overall intersection. From Lemma 2.8 and 2.11, it is clear that the two conditions for merging in total variation in expectation are equivalent since the sigma fields on the LHS of Equation (2.10) and (2.13) are equal. □

Remark 2.3. *Corollary 2.1 is a new result in view of the existing literature. We note first that much of the literature focuses on continuous time, where the predictor is not used in the analysis. In discrete time, [36, Lemma 4.2] proves that the merging of the predictor in total variation in expectation implies that of the filter. However this result relies on a non-degeneracy assumption in the measurement channel and the specific structure of the filter recursion equation [10, Equation 1.4].*

We have now established that the filter merges in total variation in expectation, but we would like to extend this result to hold almost surely. By a simple application of Fatou's lemma, we can argue the liminf of the total variation of the filter is zero P^μ a.s. Hence if the limit exists, it must be zero, yet it is not immediate that the limit will exist. In [36, p. 572], a technique is established to prove the existence of this limit. We now recall the following, where a proof is included for completeness.

Theorem 2.4. *[36, p. 572] Assume the filter is stable in total variation in expectation. Then the filter is stable in total variation P^μ a.s.*

Proof. Let $\gamma = \frac{\mu+\nu}{2}$, then we have that $\mu \ll \gamma, \nu \ll \gamma$ and furthermore $\|\frac{d\mu}{d\gamma}\|_\infty < 2, \|\frac{d\nu}{d\gamma}\|_\infty < 2$. The boundedness of the Radon-Nikodym derivatives is key, as this makes the expressions in the numerator of equation (2.9) uniformly integrable martingales with respect tot he measure P^μ. This is different than being uniformly integrable with respect to the measure P^ν. The latter holds even if the Radon Nikodym derivative $\frac{d\mu}{d\nu}$ is unbounded since $\frac{d\nu}{d\mu}$ is an $L^1(P^\nu)$ function which closes the martingale $E^\nu\left[\frac{d\mu}{d\nu}(X_0|Y_{[0,n-1]}\right]$ in the right, thus making it uniformly integrable. However, this no longer holds we we consider the measure P^μ and thus we need a hard upper bound on the derivative $\frac{d\mu}{d\gamma}$ to achieve uniform integrability.

By the martingale convergence theorem (see [5, Theorem 2]) the expressions in the numerator of (2.9) converge as $n \to \infty$. Furthermore, the denominator converges to a non-zero quantity. Therefore $\|\pi_n^\mu - \pi_n^\gamma\|_{TV}$ and $\|\pi_n^\nu - \pi_n^\gamma\|_{TV}$ admit limits P^μ a.s. We have by assumption,

$$\lim_{n\to\infty} E[\|\pi_n^\mu - \pi_n^\gamma\|_{TV}] = 0 \qquad \lim_{n\to\infty} E[\|\pi_n^\nu - \pi_n^\gamma\|_{TV}] = 0$$

therefore, if the limits exist a.s., they must be zero. Via Fatou's lemma, we have that $\underline{\lim}_{n\to\infty} \|\pi_n^\mu - \pi_n^\nu\|_{TV} = 0$, and via triangle inequality

$$\overline{\lim_{n\to\infty}} \|\pi_n^\mu - \pi_n^\nu\|_{TV} \le \overline{\lim} \|\pi_n^\mu - \pi_n^\gamma\|_{TV} + \overline{\lim_{n\to\infty}} \|\pi_n^\mu - \pi_n^\gamma\|_{TV} = 0 \quad P^\mu \ a.s.$$

□

2.3.4 Relative Entropy Merging

We will now show that the relative entropy merging of the filter is essentially equivalent to merging in total variation in expectation. Via Lemma 2.6 and 2.9, it is clear that the filter and predictor admit Radon-Nikodym derivatives. Therefore, working with $D(\pi_n^\mu \| \pi_n^\nu)$ and $D(\pi_{n-}^\mu \| \pi_{n-}^\nu)$ is well defined.

It has been established in [12] that the relative entropy of the filter is a decreasing sequence, but the analysis is in continuous time and it is worth recreating the results here in discrete time. To this end, we will extensively use the chain rule for relative entropy [21, Theorem 5.3.1]:

Lemma 2.12. *For joint measures P, Q on random variables X, Y we have*

$$D(P(X,Y)\|Q(X,Y)) = D(P(X)\|Q(X)) + D(P(Y|X)\|Q(Y|X))$$

Note for two sigma fields $\mathcal{F}$ and $\mathcal{G}$ and two joint measures P and Q on $\mathcal{F} \vee \mathcal{G}$ one could also express this relationship as

$$D(P|_{\mathcal{F}\vee\mathcal{G}}\|Q|_{\mathcal{F}\vee\mathcal{G}}) = D(P|_{\mathcal{F}}\|Q|_{\mathcal{F}}) + D(P|_{\mathcal{G}}|\mathcal{F}\|Q|_{\mathcal{G}}|\mathcal{F})$$

we will use either notation where it is most convenient. We now use the chain rule to establish the monotonicity and convergence of the respective relative entropy sequences.

Lemma 2.13.

$$E^{\mu}[D(\pi_n^{\mu}\|\pi_n^{\nu})] \leq E^{\mu}[D(\pi_{n-}^{\mu}\|\pi_{n-}^{\nu})]$$

Proof. Using the chain rule (Lemma 2.12) we arrive at the following:

$$\begin{aligned}
&D(P^{\mu}(X_n, Y_n|Y_{[0,n-1]})\|P^{\nu}(X_n, Y_n|Y_{[0,n-1]})) \\
=&D(P^{\mu}(X_n|Y_{[0,n-1]})\|P^{\nu}(X_n|Y_{[0,n-1]})) + D(P^{\mu}(Y_n|Y_{[0,n-1]}, X_n)\|P^{\nu}(Y_n|Y_{[0,n-1]}, X_n)) \\
=&E^{\mu}[D(\pi_{n-}^{\mu}\|\pi_{n-}^{\nu})] + D(P^{\mu}(Y_n|Y_{[0,n-1]}, X_n)\|P^{\nu}(Y_n|Y_{[0,n-1]}, X_n)) \\
=&E^{\mu}[D(\pi_{n-}^{\mu}\|\pi_{n-}^{\nu})] \qquad (2.14)
\end{aligned}$$

As was discussed in the proof of Theorem 2.1, Y_n conditioned on X_n is independent of past $Y_{[0,n-1]}$ values and initial measure ν or μ since $Y_n \sim Q(dy_n|x_n)$, therefore

$D(P^{\mu}(Y_n|Y_{[0,n-1]},X_n)\|P^{\nu}(Y_n|Y_{[0,n-1]},X_n)) = 0$. If we apply the chain rule the other way we have

$$\begin{aligned}
&D(P^{\mu}(X_n,Y_n|Y_{[0,n-1]})\|P^{\nu}(X_n,Y_n|Y_{[0,n-1]}))\\
=&D(P^{\mu}(X_n|Y_{[0,n]})\|P^{\nu}(X_n|Y_{[0,n]})) + D(P^{\mu}(Y_n|Y_{[0,n-1]})\|P^{\nu}(Y_n|Y_{[0,n-1]}))\\
=&E^{\mu}[D(\pi_n^{\mu}\|\pi_n^{\nu})] + D(P^{\mu}(Y_n|Y_{[0,n-1]})\|P^{\nu}(Y_n|Y_{[0,n-1]})) \qquad (2.15)
\end{aligned}$$

Since relative entropy is always greater than zero, we can equate (2.14) and (2.15) and arrive at our conclusion, that the relative entropy of the one step predictor is is greater than the non-linear filter. □

Lemma 2.14.

$$E^{\mu}[D(\pi_{n+1-}^{\mu}\|\pi_{n+1-}^{\nu})] \leq E^{\mu}[D(\pi_n^{\mu}\|\pi_n^{\nu})]$$

Proof. Using the chain rule in a similar fashion we have

$$\begin{aligned}
&D(P^{\mu}(X_n,X_{n+1}|Y_{[0,n]})\|P^{\nu}(X_n,X_{n+1}|Y_{[0,n]}))\\
=&D(P^{\mu}(X_n|Y_{[0,n]})\|P^{\nu}(X_n|Y_{[0,n]})) + D(P^{\mu}(X_{n+1}|Y_{[0,n]},X_n)\|P^{\nu}(X_{n+1}|Y_{[0,n]},X_n))\\
=&E^{\mu}[D(\pi_n^{\mu}\|\pi_n^{\nu})] + D(P^{\mu}(X_{n+1}|Y_{[0,n]},X_n)\|P^{\nu}(X_{n+1}|Y_{[0,n]},X_n))\\
=&E^{\mu}[D(\pi_n^{\mu}\|\pi_n^{\nu})] \qquad (2.16)
\end{aligned}$$

Now, Y_n is a noisy measurement of X_n, and $\{X_n\}_{n=0}^{\infty}$ is a Markov chain, therefore X_{n+1} conditioned on X_n is independent of $Y_{[0,n]}$ and the initial measure, therefore the second term above is zero yielding (2.16). Applying the chain rule the other way we have

$$
\begin{aligned}
&D(P^\mu(X_n, X_{n+1}|Y_{[0,n]})\|P^\nu(X_n, X_{n+1}|Y_{[0,n]})) \\
=&D(P^\mu(X_{n+1}|Y_{[0,n]})\|P^\nu(X_{n+1}|Y_{[0,n]})) + D(P^\mu(X_n|X_{n+1}, Y_{[0,n]})\|P^\nu(X_n|X_{n+1}, Y_{[0,n]})) \\
=&E^\mu[D(\pi^\mu_{n+1-}\|\pi^\nu_{n+1-})] + D(P^\mu(X_n|X_{n+1}, Y_{[0,n]})\|P^\nu(X_n|X_{n+1}, Y_{[0,n]})) \qquad (2.17)
\end{aligned}
$$

relative entropy is always non-negative, therefore we equate (2.16) and (2.17) to arrive at our conclusion. □

Corollary 2.2. *The relative entropy of the one step predictor and the non-linear filter are monotonically decreasing sequences bounded below by zero, and therefore admit limits.*

Proof. By a simply application of Lemma 2.13 and 2.14 we have

$$E^\mu[D(\pi^\mu_{n+1-}\|\pi^\nu_{n+1-})] \leq E^\mu[D(\pi^\mu_n\|\pi^\nu_n)] \leq E^\mu[D(\pi^\mu_{n-}\|\pi^\nu_{n-})]$$

therefore the one step predictor is a monotonically decreasing sequence bounded below by zero, and admits a limit. Similarly we have

$$E^\mu[D(\pi^\mu_{n+1}\|\pi^\nu_{n+1})] \leq E^\mu[D(\pi^\mu_{n+1-}\|\pi^\nu_{n+1-})] \leq E^\mu[D(\pi^\mu_n\|\pi^\nu_n)]$$

so the non-linear filter also exhibits this property. □

In the literature it has been remarked that relative entropy merging of the filter is equivalent to total variation merging in expectation. See for example [10, Remark 4.2] or [38, Remark 5.9]. In [32] it is shown that relative entropy is a non-increasing sequence, but not that the limit of this sequence is zero. The following result establishes this.

Lemma 2.15. *Assume there exists some finite n such that $E^\mu[D(\pi_n^\mu \| \pi_n^\mu)] < \infty$ and some m such that $E^\mu[D(P^\mu|_{\mathcal{F}^{\mathcal{Y}}_{0,m}} \| (P^\nu|_{\mathcal{F}^{\mathcal{Y}}_{0,m}})] < \infty$. Then the filter is stable in relative entropy if and only if it is stable in total variation in expectation.*

Proof. First assume the filter is stable in relative entropy. Since the square root function is continuous and convex, we have

$$0 = \lim_{n\to\infty} \sqrt{\frac{2}{\log(e)} E^\mu[D(\pi_n^\mu \| D(\pi_n^\nu)]} \geq \lim_{n\to\infty} E^\mu \left[\sqrt{\frac{2}{\log(e)} D(\pi_n^\mu \| D(\pi_n^\nu)} \right]$$

where we have applied Jensen's inequality. We then apply Pinsker's inequality (1.7) and we have $\lim_{n\to\infty} E^\mu[\|\pi_n^\mu - \pi_n^\nu\|_{TV}] = 0$.

For the converse direction, by the chain rule, it is clear that

$$\begin{aligned} E^\mu[D(\pi_n^\mu \| \pi_n^\nu)] &= D(P^\mu|_{\mathcal{F}_n^{\mathcal{X}}} | \mathcal{F}^{\mathcal{Y}}_{0,n} \| P^\nu|_{\mathcal{F}_n^{\mathcal{X}}} | \mathcal{F}^{\mathcal{Y}}_{0,n}) \\ &= D(P^\mu|_{\mathcal{F}_n^{\mathcal{X}} \vee \mathcal{F}^{\mathcal{Y}}_{0,n}} \| P^\nu|_{\mathcal{F}_n^{\mathcal{X}} \vee \mathcal{F}^{\mathcal{Y}}_{0,n}}) - D(P^\mu|_{\mathcal{F}^{\mathcal{Y}}_{0,n}} \| (P^\nu|_{\mathcal{F}^{\mathcal{Y}}_{0,n}}) \end{aligned}$$

by the Markov Property we have $X_{[0,n-1]}, Y_{[0,n-1]}$ and $X_{[n+1,\infty)}, Y_{[n+1,\infty)}$ are conditionally independent given X_n, Y_n therefore we have:

$$D(P^\mu|_{\mathcal{F}_n^{\mathcal{X}} \vee \mathcal{F}^{\mathcal{Y}}_{0,n}} \| P^\nu|_{\mathcal{F}_n^{\mathcal{X}} \vee \mathcal{F}^{\mathcal{Y}}_{0,n}}) = D(P^\mu|_{\mathcal{F}^{\mathcal{X}}_{n,\infty} \vee \mathcal{F}^{\mathcal{Y}}_{0,\infty}} \| P^\nu|_{\mathcal{F}^{\mathcal{X}}_{n,\infty} \vee \mathcal{F}^{\mathcal{Y}}_{0,\infty}})$$

Then $\mathcal{F}^{\mathcal{X}}_{n,\infty} \vee \mathcal{F}^{\mathcal{Y}}_{0,\infty}$ is a decreasing sequence of sigma fields. By [2, Theorem 2] we have that if the relative entropy is ever finite, the limit of the relative entropy restricted to these sigma fields is the relative entropy restricted to the intersection of the decreasing fields, that is

$$\lim_{n\to\infty} D(P^\mu|_{\mathcal{F}^{\mathcal{X}}_{n,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}}\|P^\nu|_{\mathcal{F}^{\mathcal{X}}_{n,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}}) = D(P^\mu|_{\bigcap_{n\geq 0}\mathcal{F}^{\mathcal{X}}_{n,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}}\|P^\nu|_{\bigcap_{n\geq 0}\mathcal{F}^{\mathcal{X}}_{n,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}})$$

Likewise, $\mathcal{F}^{\mathcal{Y}}_{0,n}$ is an increasing sequence of sigma fields, therefore by [2, Theorem 3] we have that if the relative entropy is ever finite, the relative entropy restricted to these sigma fields is the relative entropy over the limit field, that is

$$\lim_{n\to\infty} D(P^\mu|_{\mathcal{F}^{\mathcal{Y}}_{0,n}}\|P^\nu|_{\mathcal{F}^{\mathcal{Y}}_{0,n}}) = D(P^\mu|_{\mathcal{F}^{\mathcal{Y}}_{0,\infty}}\|P^\nu|_{\mathcal{F}^{\mathcal{Y}}_{0,\infty}})$$

Therefore,

$$\lim_{n\to\infty} E^\mu[D(\pi^\mu_n\|\pi^\nu_n)] = D(P^\mu|_{\bigcap_{n\geq 0}\mathcal{F}^{\mathcal{X}}_{n,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}}\|P^\nu|_{\bigcap_{n\geq 0}\mathcal{F}^{\mathcal{X}}_{n,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}}) - D(P^\mu|_{\mathcal{F}^{\mathcal{Y}}_{0,\infty}}\|P^\nu|_{\mathcal{F}^{\mathcal{Y}}_{0,\infty}})$$

By Lemma 2.3 we have

$$\frac{dP^\mu|_{\bigcap_{n\geq 0}\mathcal{F}^{\mathcal{X}}_{n,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}}}{dP^\nu|_{\bigcap_{n\geq 0}\mathcal{F}^{\mathcal{X}}_{n,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}}} = E^\nu\left[\frac{d\mu}{d\nu}(X_0)\middle|\bigcap_{n\geq 0}\mathcal{F}^{\mathcal{X}}_{n,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}\right] = f_1$$
$$\frac{dP^\mu|_{\mathcal{F}^{\mathcal{Y}}_{0,\infty}}}{dP^\nu|_{\mathcal{F}^{\mathcal{Y}}_{0,\infty}}} = E^\nu\left[\frac{d\mu}{d\nu}(X_0)\middle|\mathcal{F}^{\mathcal{Y}}_{0,\infty}\right] = f_2$$

Note that f_1 is $\bigcap_{n\geq 0}\mathcal{F}^{\mathcal{X}}_{n,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}$ measurable, while f_2 is $\mathcal{F}^{\mathcal{Y}}_{0,\infty}$ measurable, and $\mathcal{F}^{\mathcal{Y}}_{0,\infty}\subset\bigcap_{n\geq 0}\mathcal{F}^{\mathcal{X}}_{n,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}$. By Lemma 2.8, we have that if the filter merges in total variation in expectation, then for a set of state and observation sequences $\omega = (x_i, y_i)_{i=0}^\infty \in A \subset \mathcal{F}^{\mathcal{X}}_{0,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}$ with $P^\nu(A) = 1$, we have $f_1(\omega) = f_2(\omega)$. Yet this then means over the set A of P^ν measure 1, $f_1 = f_2$ is $\mathcal{F}^{\mathcal{Y}}_{0,\infty}$ measurable. We then have

$$
\begin{aligned}
&D(P^\mu|_{\bigcap_{n\geq 0}\mathcal{F}^{\mathcal{X}}_{n,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}}\|P^\nu|_{\bigcap_{n\geq 0}\mathcal{F}^{\mathcal{X}}_{n,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}}) - D(P^\mu|_{\mathcal{F}^{\mathcal{Y}}_{0,\infty}}\|(P^\nu|_{\mathcal{F}^{\mathcal{Y}}_{0,\infty}})\\
&= E^\mu[\log(f_1)] - E^\mu[\log(f_2)] = E^\nu[f_1\log(f_1)] - E^\nu[f_2\log(f_2)]\\
&= \int_\Omega f_1(\omega)\log(f_1(\omega))P^\nu|_{\bigcap_{n\geq 0}\mathcal{F}^{\mathcal{X}}_{n,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}}(d\omega) - \int_\Omega f_2(\omega)\log(f_2(\omega))P^\nu|_{\mathcal{F}^{\mathcal{Y}}_{0,\infty}}(d\omega)\\
&= \int_A f_1(\omega)\log(f_1(\omega))P^\nu|_{\bigcap_{n\geq 0}\mathcal{F}^{\mathcal{X}}_{n,\infty}\vee\mathcal{F}^{\mathcal{Y}}_{0,\infty}}(d\omega) - \int_A f_2(\omega)\log(f_2(\omega))P^\nu|_{\mathcal{F}^{\mathcal{Y}}_{0,\infty}}(d\omega)\\
&= \int_A f_1(\omega)\log(f_1(\omega))P^\nu|_{\mathcal{F}^{\mathcal{Y}}_{0,\infty}}(d\omega) - \int_A f_2(\omega)\log(f_2(\omega))P^\nu|_{\mathcal{F}^{\mathcal{Y}}_{0,\infty}}(d\omega)\\
&= 0
\end{aligned}
$$

Therefore, if the relative entropy of the filter is ever finite, then total variation merging in expectation is equivalent to merging in relative entropy. □

2.4 Examples of Observable POMP

We will now study some POMP which are observable. To characterize these observation channels, rather than view Y_n as related to X_n via the stochastic kernel Q we will consider an i.i.d. noise process $\{Z_n\}_{n=0}^\infty$ independent of $\{X_n\}_{n=0}^\infty$ taking values in the measurable space $(\mathcal{Z}, \mathcal{B}(\mathcal{Z}))$ with distribution $\zeta \in \mathcal{P}(\mathcal{Z})$. We then have a measurement function $h : \mathcal{X} \times \mathcal{Z} \to \mathcal{Y}$ and each measurement is realized as $Y_n = h(X_n, Z_n)$. For a fixed $X_n = x$ we will denote $h(x, Z_n) = h_x(Z_n)$. We denote the push forward measure of ζ under h_x as $h_x\zeta \in \mathcal{P}(\mathcal{Z})$ for each $x \in \mathcal{X}$. That is, for any $A \in \mathcal{B}(\mathcal{Y})$ we have $h_x\zeta(A) = \zeta(h_x^{-1}(A))$. This of course defines the measurement kernel as before, $Y_n \sim Q(dy|X_n) = h_{X_n}\zeta(dy)$ and in fact any measurement channel Q can be expressed via some function h and noise distribution ζ and vice versa. However, in some cases it is more transparent to characterize the measurement channel via the noise process and a measurement function rather than the measurement kernel.

2.4.1 Affine Measurement Channel

Consider $\mathcal{X}, \mathcal{Z}$ as compact subsets of $\mathbb{R}$. and let $y = h(x, z) = a(z)x + b(z)$ for some functions a, b where the image of $\mathcal{Z}$ under a and b is compact (this ensures that $\mathcal{Y}$ is compact). Note that for a fixed choice of z, this is an affine function of x. We will show sufficient conditions for one step observability. Since $\mathcal{X}$ is compact, the set of polynomials is dense in the set of continuous and bounded functions. Therefore, without loss of generality we assume f is a polynomial. Consider then the mapping

$$S : \mathbb{R}^{\mathbb{R}} \to \mathbb{R}^{\mathbb{R}} \qquad S(g)(\cdot) \mapsto \int_Z g(h(\cdot, z))\zeta(dz)$$

Let $\mathbb{R}[x]_n$ represent the polynomials on the real line up to degree n. If we take a polynomial $g(y) = \sum_{i=0}^{n} \alpha_i y^i$ and apply S we have

$$\begin{aligned} S(g)(x) &= \int_{\mathcal{Z}} \sum_{i=0}^{n} \alpha_i (x(a(z)) + b(z))^i \zeta(dz) \\ &= \sum_{i=0}^{n} \alpha_i \int_Z (x(a(z)) + b(z))^i \zeta(dz) \\ &= \sum_{i=0}^{k} \alpha_i \int_{\mathcal{Z}} \sum_{k=0}^{i} \binom{i}{k} (xa(z))^k b(z)^{i-k} \zeta(dz) \end{aligned}$$

where we have applied binomial theorem to expand the exponent terms. X is independent of Z, therefore inside the integral it acts as a constant and we can move the

x outside the integral.

$$= \sum_{i=0}^{k} \alpha_i \sum_{k=0}^{i} x^k \binom{i}{k} \int_{\mathcal{Z}} a(z)^k b(z)^{i-k} \zeta(dz)$$
$$= \sum_{k=0}^{n} x^k \sum_{i=k}^{n} \alpha_i \binom{i}{k} E(a(Z)^k b(Z)^{i-k})$$

we see that this is a polynomial in x therefore $S(g)$ is invariant on $\mathbb{R}[x]_n$, that is if g is polynomial of degree n then $S(g)$ is a polynomial of degree n. Furthermore, the coefficients of $S(g)(x) = \sum_{i=0}^{n} \beta_i x^i$ can be related to the coefficients of $g(x) = \sum_{i=0}^{n} \alpha_i x^i$ by a linear transformation. Define $N(i,k) = \binom{i}{k} E(a(Z)^k b(Z)^{i-k})$ then by recursive application of binomial theorem we have

$$\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} N(0,0) & N(1,0) & \cdots & N(n,0) \\ 0 & N(1,1) & \cdots & N(n,1) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & N(n,n) \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

if we want to generate any polynomial, we require this matrix to be invertible, and since it is upper triangular this amounts to none of the diagonal entries being zero, that is $E[a(z)^n] \neq 0 \; \forall n \in \mathbb{N}$. Furthermore, we want g to be bounded so we must have $N(n,k) < \infty \; \forall n \in \mathbb{N}, k \in \{0, \cdots, i\}$.

2.4.2 Threshold Channel

Consider $\mathcal{X}$ as a compact subset of $\mathbb{R}$, $\mathcal{Z} = \mathbb{R}$. Let $y = h(x,z) = 1_{x>z}x + 1_{x\leq z}z$ and assume that ζ admits a density q with respect to Lebesgue. We have

$$\int_{\mathcal{Z}} g(h(x,z))\zeta(dz) = \int_{-\infty}^{x} g(x)q(z)dz + \int_{x}^{\infty} g(z)q(z)dz$$

again, we can approximate any continuous and bounded function f on $\mathcal{X}$ as polynomial, so we assume f is differentiable. We have

$$\begin{aligned} f(x) &= \int_{-\infty}^{x} g(x)q(z)dz + \int_{x}^{\infty} g(z)q(z)dz \\ f'(x) &= g(x)q(x) + \int_{-\infty}^{x} g'(x)q(z)dz - g(x)q(x) = g'(x)\zeta(Z \leq x) \end{aligned}$$

Since $\mathcal{X}$ is compact there exists some $x_{\min} \in \mathbb{R}$ such that $x_{\min} < \mathcal{X}$. We require for some $\epsilon > 0$ that $\zeta(Z < x_{min}) > \epsilon$. This condition says every $x \in \mathcal{X}$ has some positive probability of being observed through $h(x,z)$ and we will not always get pure noise. Then we have

$$\begin{aligned} g'(x) &= 1_{\mathcal{X}}(x)\frac{f'(x)}{\zeta(Z \leq x)} \\ g(x) &= c + \int_{-\infty}^{x} 1_{\mathcal{X}}(u)\frac{f'(u)}{\zeta(Z \leq u)}du \end{aligned}$$

for some constant c. Therefore, we only need to define g over $\mathcal{X}$. Furthermore, we require g to be bounded, which is implied if g' is bounded since g is only defined over a compact space.

2.4.3 Direct Observation

Consider now the case when $y = h(x)$ for some invertible function h. This can be written as $y = h(x) + z$ where $\zeta \sim \delta_0$, that is a point mass at zero. We then have for any measurable bounded function g

$$\int_Z g(h(x) + z)\zeta(dz) = g(h(x))$$

Then for any continuous and bounded function f, define $g = f \circ h^{-1}$ and we have $f(x) = \int_Z g(h(x) + z)\zeta(dz)$ and the measurement channel is one step observable.

2.4.4 Finite State and Measurement Space

Consider $\mathcal{X} = \{a_1, \cdots, a_n\}$ and $\mathcal{Y} = \{b_1, \cdots, b_m\}$ as finite spaces. We note that for such a setup, there is already a sufficient and necessary condition provided in [40, Theorem V.2]. However, we examine this case to show that our definition is equivalent to the sufficient direction of this theorem, which is van Handel's notion of observability [37].

The transition kernel T is then an $n \times n$ matrix, while the measurement kernel Q is then an $n \times m$. The POMP is one step observable if and only if Q is rank n or higher. For $k > 1$ step observability, we can directly compute the conditional distribution of $Y_{[0,k]}|X_0$ via Q and T, which is a $n \times m^k$ matrix. If this matrix is rank n or higher, than the POMDP is k step observable.

However, the size of this matrix grows exponentially in k. We can instead compute a simpler sufficient condition for observability using the marginal measures $Y_j|X_0, j \in \{0, \cdots k\}$ rather than the joint measure $Y_{[0,k]}|X_0$. Each measure $Y_j|X_0$ is an $n \times m$

matrix, therefore all the marginal measures have complexity $n \times k(m)$ which grows linearly in k rather than exponentially. Consider the class of functions $\mathcal{G}^n = \{g : \mathcal{Y}^n \to \mathbb{R}\}$ and a subclass $\mathcal{G}^n_{LC} = \{g(y_{[1,n]}) = \sum_{i=1}^n g_i(y_i) | g_i \in \mathcal{G}^1\}$. That is, a linear combination of functions of the individual y_i values.

Lemma 2.16. *Assume $|\mathcal{X}| = n$ and $|\mathcal{Y}| = m$. If the $n \times n(m)$ matrix*

$$M = \begin{pmatrix} Q & TQ & \cdots & T^{n-1}Q \end{pmatrix}$$

is rank n or higher, then the measurement channel is n step observable. Furthermore, appending more blocks of the form T^kQ for $k \geq n$ will not increase the rank of M.

Proof. Consider some continuous and bounded function f on the state space. Since the state space is finite, any function is a vector in $\mathbb{R}^n$. Similarly, any function $g \in \mathcal{G}^n_{LC}$ is a vector in $\mathbb{R}^{nm}$, $\alpha = (g_1(b_1), \cdots, g_1(b_m), \cdots, g_n(b_1), \cdots, g_n(b_m))$. Solving the equation

$$f(x) = \int_{\mathcal{Y}^n} g(y_{[0,n-1]}) P^\mu(dy_{[0,n-1]})$$

is solving the matrix equation

$$\begin{pmatrix} f(a_1) \\ \vdots \\ f(a_n) \end{pmatrix} = \begin{pmatrix} Q & TQ & \cdots & T^{n-1}Q \end{pmatrix} \alpha$$

If M is rank n, then for any function f we can find a vector α and corresponding function g to make the system observable.

Consider if M is not rank n and if we append another block T^nQ to M. By the Cayley-Hamilton theorem, T^n is a linear combination of lower powers of T, e.g. $T^n = \sum_{i=0}^{n} \alpha_i T^i$ for some coefficients α_i. Therefore this additional block is a linear combination of the previous blocks, and adds no dimension to the matrix M. □

If the conditions of this lemma fail, i.e. M is not rank n, that means integrating g over the marginal measures cannot generate any f function. Yet the product of the marginal measures is not the the joint measure since the $Y_i|X_1$ are not independent. Hence, working with the marginal measures only is not enough to determine observability as also noted by van Handel in [37, Remark 13] in a slightly different setup. Consider the following example

Example

Consider if $\mathcal{X} = \{1, 2, 3, 4\}$ and $Y = 1_{x \leq 2}$. This can be realized as

$$Q = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \end{pmatrix}$$

Consider the following transition kernel,

$$T = \begin{pmatrix} 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}$$

Notice that the odd and even rows are identical. If we consider the marginal measures of $Y_1|X_1, \cdots, Y_4|X_1$ we have the matrix

$$\begin{pmatrix} Q & \cdots & T^3Q \end{pmatrix} =$$
$$\begin{pmatrix} 0 & 1 & 0.75 & 0.25 & 0.5625 & 0.4375 & 0.609375 & 0.390625 \\ 0 & 1 & 0.50 & 0.50 & 0.6250 & 0.3750 & 0.593750 & 0.406250 \\ 1 & 0 & 0.75 & 0.25 & 0.5625 & 0.4375 & 0.609375 & 0.390625 \\ 1 & 0 & 0.50 & 0.50 & 0.6250 & 0.3750 & 0.593750 & 0.406250 \end{pmatrix}$$

Which is only rank 3, not rank 4. Therefore, we cannot use the marginal measures to determine observability.

However, if we consider the joint measure of $(Y_1, Y_2)|X_1$ we have the matrix

$$A' = \begin{pmatrix} 0 & 0 & \frac{3}{4} & \frac{1}{4} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{3}{4} & \frac{1}{4} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{pmatrix}$$

Where row i is conditioned on $x = i$ and the columns are ordered in binary y_2y_1, e.g. $P(y_1 = 1, y_2 = 0|x_1 = 2)$ is row 2 column 3. This matrix is full rank, hence the system is N step observable with $N = 2$, even though the marginal measures failed to be full rank.

Recall van Handel defines the term "observability" as every distinct prior resulting in a unique measure on $\mathcal{Y}^{\mathbb{Z}_+}$. Therefore, for a finite system our notion of N step observability is a sufficient condition for this notion.

2.5 Summary

In conclusion, the main properties of a POMP that result in filter stability are the Markovian dependency structure, the observability of the measurement channel, as well as the merging in total variation of the predictive measures $P^{\mu}(Y_{[n,n+k]}|Y_{[0,n-1]})$ and $P^{\nu}(Y_{[n,n+k]}|Y_{[0,n-1]})$ as $n \to \infty$. The dependacy structure of the POMP is inherent, and the merging of the predictive measures on the Y_n follows from absolute continuity of the priors $\mu \ll \nu$ and the theorem of Blackwell and Dubins [5]. Observability then is the contentious property of a POMP which must be investigated to achieve stability. As we show in Section 2.4, this property holds for a variety of example systems.

Observability however only results in the weak merging of the predictor almost surely, and from here different assumptions are required to extend to more stringent notions of stability for the filter.

Via our results on the structure of the Radon Nikodym derivative in Section 2.3.2, we show in Corollary 2.1 that total variation merging in expectation is equivalent for the predictor and filter. Under a mild finiteness assumption of the relative entropies, total variation merging in expectation is also equivalent to relative entropy merging.

Chapter 3

Filter Stability in Controlled Environments

3.1 Introduction

The filter can also be examined in a POMDP, where the filter is now the conditional measure on X_n given knowledge of the measurements $Y_{[0,n]}$ and control actions $U_{[0,n-1]}$, or equivalently the measurements and knowledge of the control policy which generates the actions.

Except for the well-established theory on the stability of Kalman filters, in the literature there is not much discussion of filter stability in controlled environments. Much of the filter stability literature focuses on the nature of the measurement channel, or the ergodic properties of the transition kernel in a POMP. With significant complexity already arising from these areas, the dependency structure of a POMDP further complicates matters and has not been studied in detail. Furthermore, those results which rely on the transition kernel to bring about stability run into significant issues in a POMDP due to the time varying nature and past dependency of the transition kernel.

In the control literature, the effects of control policies on future conditional probabilities is well recognized. This concept is often referred to as the dual effect of control, in that the control action is implemented to minimize the cost but also affects the conditional distributions given the measurements (i.e. the filter recursion). Vaguely speaking, dual effect is with regard to the conditional estimation error being independent of the applied control policy [1]. A system has the no dual effect if the expected error of the filter will be the same conditioned on the measurements regardless of the control actions taken, even though there is a subtle distinction between policy independence and action independence [42]. In the problems we study, the dual effect is present and as such, the dependency structure of the POMDP is more complicated than that of the POMP.

As will be discussed later in Chapter 5, in the theory of optimal partially observed stochastic control, an optimal policy can without any loss use the non-linear filter realization as a sufficient statistic. Thus, the control admits, in some sense, a separation structure where first one computes the filter, and then one computes the control as a function of the filter. Therefore, we see that the filter is vital as it is the source of information for the optimal controller.

Many of the results for POMPs will carry over to POMDPs, but there is an important distinction to make. A POMDP is in general *not* a Markov chain under an arbitrary policy. Therefore, we cannot appeal to the Markov property to consider the future and the past as independent conditioned on the present. This has a number of implications in modifying the results from the POMP setup.

For one step observability, the channel $Y_n|X_n$ is unaffected by the control policy, therefore the POMP results carry over easily. However, for $N > 1$ step observability the channel $Y_{[n,n+N-1]}|X_n$ now depends on the control actions and thus may be time varying and dependent on the past measurements. In Section 3.3 we examine where these issues lie in greater detail. Furthermore, in considering the total variation merging and relative entropy merging certain proofs must be redone without appealing to the Markov property.

3.2 Predictor Stability

Theorem 3.1. *Let $\mu \ll \nu$ and let the POMDP be one-step observable (Definition 2.1 i), then the predictor is universally stable in weak merging $P^{\mu,\lambda}$ a.s. That is $\pi_{n-}^{\mu,\gamma}$ and $\pi_{n-}^{\nu,\gamma}$ merge weakly $P^{\mu,\gamma}$ a.s. for any control policy $\gamma \in \Gamma$.*

Proof. We will first consider a generalized result which we can return to when considering $N > 1$ step observability. Consider three stochastic processes $A = \{A_n\}_{n=0}^{\infty}, B = \{B_n\}_{n=0}^{\infty}, C = \{C_n\}_{n=0}^{\infty}$ defined on the same measurable space $(\Omega, \mathcal{F})$ mapping to spaces $\mathcal{A}, \mathcal{B}, \mathcal{C}$ with their respective Borel sigma fields. We can think of A and C as the "observed" processes and B as some "hidden" process whose realizations are not known by an observer. For possible measures P and Q on $(\Omega, \mathcal{F})$ we have the following definitions:

Definition 3.1. *For a measure P we say the process C only depends on the process A through B if for every $n \in \mathbb{N}$ we have P almost surely*

$$P(C_n \in \cdot | A_n = a, B_n = b) = P(C_n \in \cdot | B_n = b)$$

Definition 3.2. *For a measure P we say the channel $C|B$ is time homogeneous if* *for every $n \in \mathbb{N}$, P almost surely*

$$P(C_n \in \cdot | B_n = b) = P(C_0 \in \cdot | B_0 = b)$$

Definition 3.3. *For two measures P and Q we say the channel $C|B$ is measure equivalent if for all $n \in \mathbb{N}$ we have, P, Q almost surely:*

$$P(C_n \in \cdot | B_n = b) = Q(C_n \in \cdot | B_n = b)$$

Definition 3.4. *For a measure P, the channel $C|B$ is observable if for every continuous and bounded function $f : \mathcal{B} \to \mathbb{R}$ we can find a measurable and bounded function $g : \mathcal{C} \to \mathbb{R}$ such that*

$$\sup_{b \in \mathcal{B}} \left| f(b) - \int_{\mathcal{C}} g(c_0) P(dc_0 | B_0 = b) \right| < \epsilon$$

Lemma 3.1. *Let A,B,C be stochastic processes as above and assume measures P, Q satisfy Definitions 3.1-3.4. Assume that*

$$\lim_{n \to \infty} \| P(C_n | A_n) - Q(C_n | A_n) \|_{TV} = 0 \;\; P \; a.s.$$

then we have that $P(B_n|A_n)$ and $Q(B_n|A_n)$ merge weakly P a.s..

Proof. Consider any continuous and bounded function $f : \mathcal{B} \to \mathbb{R}$. Pick any $\epsilon > 0$, by observability (Definition 3.4) we can find a measurable and bounded function

$g : \mathcal{C} \to \mathbb{R}$ such that

$$\tilde{f}(b) = \int_{\mathcal{C}} g(c_0) P(dc_0 | B_0 = b) \qquad \| f - \tilde{f} \|_\infty < \frac{\epsilon}{3}$$

now consider

$$\begin{aligned}
&\left| \int f(b_n) P(db_n | a_n) - \int f(b_n) Q(db_n | a_n) \right| \\
\le &\left| \int \tilde{f}(b_n) P(db_n | a_n) - \int \tilde{f}(b_n) Q(db_n | a_n) \right| + \left| \int (f - \tilde{f})(b_n) P(db_n | a_n) \right| \\
+ &\left| \int (f - \tilde{f})(b_n) Q(db_n | a_n) \right| \\
\le &\left| \int \tilde{f}(b_n) P(db_n | a_n) - \int \tilde{f}(b_n) Q(db_n | a_n) \right| + 2 \| f - \tilde{f} \|_\infty \\
\le &\left| \int \tilde{f}(b_n) P(db_n | a_n) - \int \tilde{f}(b_n) Q(db_n | a_n) \right| + \frac{2}{3} \epsilon \qquad (3.1)
\end{aligned}$$

we then have

$$\begin{aligned}
&\left| \int_{\mathcal{B}} \tilde{f}(b_n) P(db_n | a_n) - \int_{\mathcal{B}} \tilde{f}(b_n) Q(db_n | a_n) \right| \\
= &\left| \int_{\mathcal{B}} \int_{\mathcal{C}} g(c_0) P(dc_0 | B_0 = b) P(db_n | a_n) - \int_{\mathcal{B}} \int_{\mathcal{C}} g(c_0) P(dc_0 | B_0 = b) Q(db_n | a_n) \right|
\end{aligned}$$

by measure equivalence (Definition 3.3), we can replace $P(dc_0 | B_0 = b)$ with $Q(dc_0 | B_0 = b)$ in the second term. By time homogeneity (Definition 3.2), we can replace $P(dc_0 | B_0 = b)$ with $P(dc_n | B_n = b)$ and the same for Q. We then have

$$\left| \int_{\mathcal{B}} \int_{\mathcal{C}} g(c_n) P(dc_n | B_n = b) P(db_n | a_n) - \int_{\mathcal{B}} \int_{\mathcal{C}} g(c_n) Q(dc_n | B_n = b) Q(db_n | a_n) \right|$$

By assumption, C only depends on A through B (Definition 3.1), so we can write $P(dc_n | B_n = b) = P(dc_n | B_n = b, A_n = a)$ for any a. We finally apply chain rule for conditional probability and we have

$$\left| \int_{\mathcal{C}} g(c_n)P(dc_n|a_n) - \int_{\mathcal{C}} g(c_n)Q(dc_n|a_n) \right| \leq \|P(C_n|A_n) - Q(C_n|A_n)\|_{TV} \tag{3.2}$$

By assumption, this goes to zero P a.s. as $n \to \infty$ so we can find an N where $\forall\, n > N$ we have (3.2) is less than $\frac{\epsilon}{3}$ and therefore (3.1) is less than ϵ. □

We can then apply this result to our one step observable POMDP. Denote the processes $A_n = Y_{[0,n-1]}$, $B_n = X_n$ and $C_n = Y_n$. We must then check that for the measures $P^{\mu,\gamma}$ and $P^{\nu,\gamma}$ these processes satisfy Definitions 3.1 to 3.3. Despite the POMDP not being a Markov chain, for one step observability the measurement Y_n is still fully determined by X_n and the channel Q, therefore

$$P^{\mu,\gamma}(Y_n|Y_{[0,n-1]}, X_n) = Q(Y_n|X_n) = P^{\nu,\gamma}(Y_n|X_n, Y_{[0,n-1]})$$

therefore the process Y_n only depends on $Y_{[0,n-1]}$ through X_n (Definition 3.1). Furthermore the channel between $Y_n|X_n$ is Q regardless of the time index or the initial measure, therefore the channel is time homogeneous (Definition 3.2) and measure equivalent (Definition 3.3). The process is observable (Definition 3.4) by assumption. The assumption $\mu \ll \nu$ implies $P^{\mu,\gamma}|_{\mathcal{F}^{\mathcal{Y}}_{0,\infty}} \ll P^{\nu,\gamma}|_{\mathcal{F}^{\mathcal{Y}}_{0,\infty}}$. $\{Y_n\}_{n=0}^{\infty}$ is a fully observed stochastic process, therefore by the theorem of Blackwell and Dubins [5] we have that

$$\|P^{\mu,\gamma}(Y_{[n,n+N-1]} \in \cdot|Y_{[0,n-1]}) - P^{\nu,\gamma}(Y_{[n,n+N-1]} \in \cdot|Y_{[0,n-1]})\|_{TV} \to 0 \tag{3.3}$$

Therefore we satisfy all the conditions of Lemma 3.1 and the weak merging of the predictor follows. □

The total variation results for the predictor in a POMP carry over easily to a POMDP with a slight modification to the assumption.

Assumption 3.1. *Assume $T(\cdot|x,u)$ is absolutely continuous with respect to a σ-finite measure λ for every $x \in \mathcal{X}, u \in \mathcal{U}$ and denote the resulting pdf as $t(\cdot|x,u)$. Further, assume the family $\{t(\cdot|x,u)\}_{x\in\mathcal{X},u\in\mathcal{U}}$ is uniformly bounded and equicontinuous. That is for every $x' \in \mathcal{X}$ and every $\epsilon > 0$, we can find a $\delta > 0$ such that if $d(y,x') < \delta$ we have that $|t(y|x,u) - t(x'|x,u)| < \epsilon$ for every $x \in \mathcal{X}, u \in \mathcal{U}$.*

Theorem 3.2. *Let Assumption 3.1 hold. If the predictor is universally stable in weak merging almost surely then the predictor is also universally stable in total variation almost surely.*

Proof. The proof is similar to the uncontrolled case Theorem. 2.2. □

3.3 Policy dependence of N-Step Observability in a POMDP

We will now discuss why Lemma 3.1 does not hold for $N > 1$ step observability and a control policy with memory. For a control policy γ a potential definition of of $N > 1$ observability is for every $f \in C_b(\mathcal{X})$ and every $\epsilon > 0$ we can find a measurable and bounded function g such that

$$\|f(\cdot) - \int_{\mathcal{Y}^N} g(y_{[1,N]}) P^{\mu,\gamma}(dy_{[1,N]}|X_1 = \cdot)\|_{TV} < \epsilon$$

then denote the stochastic processes $A_n = Y_{[0,n-1]}$, $B_n = X_n$ and $C_n = Y_{[n,n+N-1]}$. Recall to apply the general process Lemma 3.1 we must satisfy three properties:

1. $Y_{[n,n+N-1]}$ can only depend on $Y_{[0,n-1]}$ through X_n:

$$P^{\mu,\gamma}(Y_{[n,n+N-1]} \in \cdot|X_n, Y_{[0,n-1]}) = P^{\mu,\gamma}(Y_{[n,n+N-1]} \in \cdot|X_n) \tag{3.4}$$

2. The channel $Y_{[n,n+N-1]}|X_n$ is measure equivalent for $P^{\mu,\gamma}$ and $P^{\nu,\gamma}$:

$$P^{\mu,\gamma}(Y_{[n,n+N-1]} \in \cdot|X_n) = P^{\nu,\gamma}(Y_{[n,n+N-1]} \in \cdot|X_n) \tag{3.5}$$

3. The channel $Y_{[n,n+N-1]}|X_n$ is time homogeneous:

$$P^{\mu,\gamma}(Y_{[n,n+N-1]} \in \cdot|X_n) = P^{\mu,\gamma}(Y_{[0,N-1]} \in \cdot|X_0) \ \ \forall\ n \in \mathbb{N} \tag{3.6}$$

If we take the LHS of (3.4) we have

$$\begin{aligned}
&P^{\mu,\gamma}(Y_{[n,n+N-1]} \in \cdot|X_n, Y_{[0,n-1]}) \\
&= \int_{\mathcal{U}^{N-1}} P^{\mu,\gamma}(Y_{[n,n+N-1]} \in \cdot|X_n, Y_{[0,n-1]}, U_{[n,n+N-2]}) P^{\mu,\gamma}(du_{[n,n+N-2]}|X_n, Y_{[0,n-1]})
\end{aligned}$$

by chain rule of conditional probability. Now it is true $Y_{[n,n+N-1]}|X_n, U_{[n,n+N-2]}$ is independent of $Y_{[0,n-1]}$ so we can stop conditioning on the past measurements in the inner argument. However, in the outer conditional measure the $U_{[n,n+N-2]}$ may still depend on the past and we have

$$\int_{\mathcal{U}^{N-1}} P^{\mu,\gamma}(Y_{[n,n+N-1]} \in G|X_n, U_{[n,n+N-2]}) P^{\mu,\gamma}(du_{[n,n+N-2]}|X_n, Y_{[0,n-1]}) \tag{3.7}$$

if we take the RHS of (3.4) we have

$$\begin{aligned} &P^{\mu,\gamma}(Y_{[n,n+N-1]} \in G|X_n) \\ &= \int_{\mathcal{U}^{N-1}} P^{\mu,\gamma}(Y_{[n,n+N-1]} \in G|X_n, U_{[n,n+N-2]})P^{\mu,\gamma}(du_{[n,n+N-2]}|X_n) \end{aligned} \tag{3.8}$$

these two equations are not equal for a general control policy, therefore the process fails Definition 3.1, $Y_{[n,n+N-1]}$ does not depend on $Y_{[0,n-1]}$ through X_n. As a result, a definition of $N > 1$ step observability is incompatible with the proof technique outlined in Lemma 3.1 and we cannot utilize $N > 1$ step observability in a controlled environment to prove stability.

3.4 Filter Stability

3.4.1 Weak Merging

To achieve weak merging of the filter, the proof is the same as in the POMP case. The measurement channel Q is unaffected by control actions, so assumption 2.2 applies directly.

Theorem 3.3. *Let Assumption 2.2 hold, if the predictor is universally stable in weak merging almost surely, then the filter is universally stable in weak merging in expectation.*

Proof. The proof is similar to the uncontrolled case Theorem 2.3. □

3.4.2 Total Variation Merging

So far, the stability results in this chapter have carried over easily since they do not appeal to the Markov property of the POMP. However, Lemma 2.10, Lemma 2.14

and Lemma 2.15 all utilize the Markov property in some way, and therefore must be re-analysed in a control environment.

Lemma 2.5, 2.6, and 2.9 carry over directly from the uncontrolled case to the controlled case. Therefore we have that for a POMDP

$$\frac{d\pi_{n-}^{\mu,\gamma}}{d\pi_{n-}^{\nu,\gamma}}(x) = \frac{E^{\nu,\gamma}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,n-1]}, X_n = x]}{E^{\nu,\gamma}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,n-1]}]} \qquad P^{\mu,\gamma}\ a.s. \tag{3.9}$$

$$\frac{d\pi_{n}^{\mu,\gamma}}{d\pi_{n}^{\nu,\gamma}}(x) = \frac{E^{\nu,\gamma}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,n]}, X_n = x]}{E^{\nu,\gamma}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,n]}]} \qquad P^{\mu,\gamma}\ a.s. \tag{3.10}$$

Lemma 3.2. *Assume $\mu \ll \nu$ then we can express*

$$\|\pi_{n-}^{\mu,\gamma} - \pi_{n-}^{\nu,\gamma}\|_{TV} = \frac{E^{\nu,\gamma}\left[\left|E^{\nu,\gamma}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,\infty)}, X_{[n,\infty)}] - E^{\nu,\gamma}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,n-1]}]\right|\,\middle|\, Y_{[0,n-1]}\right]}{E^{\nu,\gamma}\left[\frac{d\mu}{d\nu}(X_0)\middle| Y_{[0,n-1]}\right]} \tag{3.11}$$

$$\|\pi_{n}^{\mu,\gamma} - \pi_{n}^{\nu,\gamma}\|_{TV} = \frac{E^{\nu,\gamma}\left[\left|E^{\nu,\gamma}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,\infty)}, X_{[n,\infty)}] - E^{\nu,\gamma}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,n]}]\right|\,\middle|\, Y_{[0,n]}\right]}{E^{\nu,\gamma}\left[\frac{d\mu}{d\nu}(X_0)\middle| Y_{[0,n]}\right]} \tag{3.12}$$

Proof. By Lemma 2.5 we can write

$$\|\pi_{n-}^{\mu,\gamma} - \pi_{n-}^{\nu,\gamma}\|_{TV} = \frac{E^{\nu,\gamma}\left[|E^{\nu,\gamma}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,n-1]}, X_n] - E^{\nu,\gamma}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,n-1]}]|\,\middle|\, Y_{[0,n-1]}\right]}{E^{\nu,\gamma}\left[\frac{d\mu}{d\nu}(X_0)\middle| Y_{[0,n-1]}\right]}$$

we no longer have the Markov property, therefore $X_{[n+1,\infty)}, Y_{[n+1,\infty)}|X_n, Y_n$ is not independent of $X_{[0,n-1]}, Y_{[0,n-1]}$. That is the future conditioned on the present is not independent of the past. However, this is not the direction of dependence we require in this proof. What we need is $X_0|Y_{[0,n-1]}, X_n$ is independent of $X_{[n+1,\infty)}, Y_{[n,\infty)}$ which still holds without the Markov property.

Consider a visual depiction in Figure 3.1. Say we are conditioning on X_2, Y_0, Y_1.

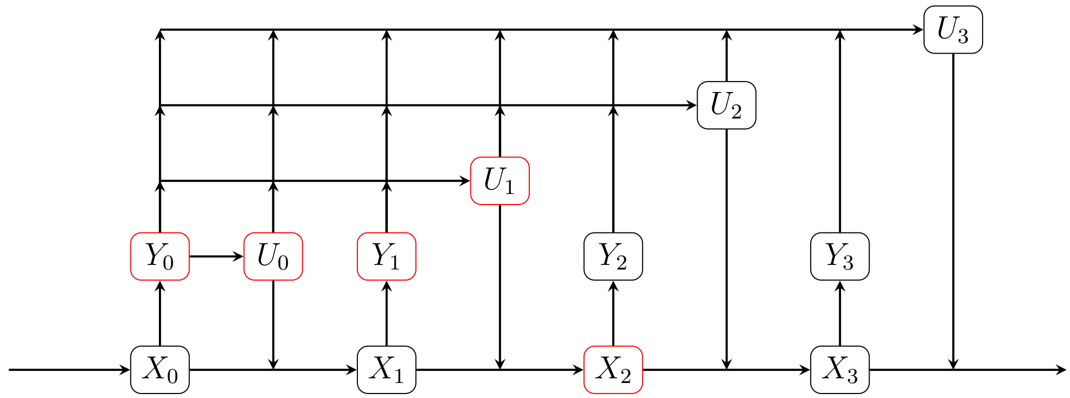

Figure 3.1: Diagram of Dependence in a POMDP. When X_0 is conditioned on $X_2, Y_{[0,1]}$ then $X_{[3,\infty)}, Y_{[2,\infty)}$ do not add any new information to the conditioning

Since the control policy is known to us, we also know the realizations of U_0 and U_1. If we colour in these nodes on the diagram we see that they cordon off X_0 from $X_{[3,\infty)}, Y_{[2,\infty)}$. This is, if we follow back any line from a node $X_n, n > 2$ or $Y_n, n > 1$ to X_0 we must go through one of the red nodes, and hence these future nodes do not add anything useful to the conditioning. Therefore we can state

$$E^{\nu,\gamma}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,n-1]}, X_n] = E^{\nu,\gamma}[\frac{d\mu}{d\nu}(X_0)|Y_{[0,\infty)}, X_{[n,\infty)}]$$

the rest of the proof follows with a similar approach for the filter. □

Lemma 2.8 and 2.11 then carry over directly.

Corollary 3.1. *For a POMDP, the filter is universally stable in total variation in expectation if and only if the predictor is universally stable in total variation in expectation.*

Proof. The proof is the similar to the uncontrolled case Corollary 2.1. □

Theorem 3.4. *For a POMDP, if the filter is universally stable in total variation in expectation then it is universally stable in total variation almost surely.*

Proof. The proof is similar to the uncontrolled case Theorem 2.4 □

3.4.3 Relative Entropy Merging

For relative entropy, Lemma 2.13 carries over directly, since the measurement channel $Y_n|X_n$ is the same for a POMDP and a POMP. However, Lemma 2.14 must be re-proven since it appeals to the Markov property.

Lemma 3.3.

$$E^{\mu,\gamma}[D(\pi_{n+1-}^{\mu,\gamma}\|\pi_{n+1-}^{\nu})] \leq E^{\mu,\gamma}[D(\pi_n^{\mu,\gamma}\|\pi_n^{\nu,\gamma})]$$

Proof. Using chain rule we have

$$\begin{aligned}
&D(P^{\mu,\gamma}(X_n, X_{n+1}|Y_{[0,n]})\|P^{\nu,\gamma}(X_n, X_{n+1}|Y_{[0,n]})) \\
=&D(P^{\mu,\gamma}(X_n|Y_{[0,n]})\|P^{\nu,\gamma}(X_n|Y_{[0,n]})) + D(P^{\mu,\gamma}(X_{n+1}|Y_{[0,n]}, X_n)\|P^{\nu,\gamma}(X_{n+1}|Y_{[0,n]}, X_n)) \\
=&E^{\mu,\gamma}[D(\pi_n^{\mu,\gamma}\|\pi_n^{\nu,\gamma})] + D(P^{\mu,\gamma}(X_{n+1}|Y_{[0,n]}, X_n)\|P^{\nu,\gamma}(X_{n+1}|Y_{[0,n]}, X_n)) \\
=&E^{\mu,\gamma}[D(\pi_n^{\mu,\gamma}\|\pi_n^{\nu,\gamma})] \qquad (3.13)
\end{aligned}$$

Now consider for any set $A \in \mathcal{B}(\mathcal{X})$ we have

$$\begin{aligned}
P^{\mu,\gamma}(X_{n+1} \in A|Y_{[0,n]}, X_n) &= \int_{\mathcal{U}} P^{\mu,\gamma}(X_{n+1} \in G|Y_{[0,n]}, X_n, U_n)P^{\mu,\gamma}(du_n|Y_{[0,n]}, X_n) \\
&= T(A|X_n, \gamma_n(Y_{[0,n]})) \\
&= P^{\nu,\gamma}(X_{n+1} \in A|Y_{[0,n]}, X_n)
\end{aligned}$$

therefore the channel $X_{n+1}|Y_{[0,n]}, X_n$ is measure equivalent for $P^{\mu,\gamma}$ and $P^{\nu,\gamma}$ therefore the second term above is zero yielding (3.13). Applying chain rule the other way we have

$$\begin{aligned}
&D(P^{\mu,\gamma}(X_n, X_{n+1}|Y_{[0,n]})\|P^{\nu,\gamma}(X_n, X_{n+1}|Y_{[0,n]})) \\
=&D(P^{\mu,\gamma}(X_{n+1}|Y_{[0,n]})\|P^{\nu,\gamma}(X_{n+1}|Y_{[0,n]})) + D(P^{\mu,\gamma}(X_n|X_{n+1}, Y_{[0,n]})\|P^{\nu,\gamma}(X_n|X_{n+1}, Y_{[0,n]})) \\
=&E^{\mu,\gamma}[D(\pi^{\mu,\gamma}_{n+1-}\|\pi^{\nu,\gamma}_{n+1-})] + D(P^{\mu,\gamma}(X_n|X_{n+1}, Y_{[0,n]})\|P^{\nu,\gamma}(X_n|X_{n+1}, Y_{[0,n]})) \qquad (3.14)
\end{aligned}$$

relative entropy is always non-negative, therefore we equate (3.13) and (3.14) to arrive at our conclusion. □

We then have the the relative entropy of the predictor and filter are non-increasing sequences bounded bellow by zero, and hence admit limits. It remains to show that this limit is zero. Lemma 2.15 cannot be applied directly since it appeals to the Markov property and therefore must be re-proven.

Lemma 3.4. *Assume there exists some finite n such that $E^{\mu,\gamma}[D(\pi^{\mu,\gamma}_n\|\pi^{\mu,\gamma}_n)] < \infty$ and some m such that $E^{\mu,\gamma}[D(P^{\mu,\gamma}|_{\mathcal{F}^{\mathcal{Y}}_{0,m}}\|(P^{\nu,\gamma}|_{\mathcal{F}^{\mathcal{Y}}_{0,m}})] < \infty$. Then the filter is universally stable in relative entropy if and only if it is universally stable in total variation in expectation.*

Proof. The entire proof of Lemma 2.15 carries over except for the following equation

$$D(P^{\mu,\gamma}|_{\mathcal{F}^{\mathcal{X}}_n \vee \mathcal{F}^{\mathcal{Y}}_{0,n}}\|P^{\nu,\gamma}|_{\mathcal{F}^{\mathcal{X}}_n \vee \mathcal{F}^{\mathcal{Y}}_{0,n}}) = D(P^{\mu,\gamma}|_{\mathcal{F}^{\mathcal{X}}_{n,\infty} \vee \mathcal{F}^{\mathcal{Y}}_{0,\infty}}\|P^{\nu,\gamma}|_{\mathcal{F}^{\mathcal{X}}_{n,\infty} \vee \mathcal{F}^{\mathcal{Y}}_{0,\infty}}) \qquad (3.15)$$

We will switch from sigma field notation to random variable notation for clarity. If we take $D(P^{\mu,\gamma}(X_{[n,\infty)},Y_{[0,\infty)})\|P^{\nu,\gamma}(X_{[n,\infty)},Y_{[0,\infty)})$ and apply chain rule we have

$$D(P^{\mu}(X_{[n,\infty)},Y_{[0,\infty]})\|P^{\nu}(X_{[n,\infty)},Y_{[0,\infty]}))=D(P^{\mu,\gamma}(X_n,Y_{[0,n]})\|P^{\nu,\gamma}(X_n,Y_{[0,n]}))$$
$$+D(P^{\mu,\gamma}((X,Y)_{[n+1,\infty)}|X_n,Y_{[0,n]})\|P^{\nu,\gamma}((X,Y)_{[n+1,\infty)}|X_n,Y_{[0,n]}))$$

as was shown in Lemma 3.3, the channel $X_{n+1}|X_n,Y_{[0,n]}$ is measure equivalent and the same holds for $(X,Y)_{[n+1,\infty)}|X_n,Y_{[0,n]}$. Therefore the second term above is zero and we have our desired result.

□

3.5 Summary

Much like a POMP, filter stability for a POMDP can arise via observability of the measurement channel. The measurement channel $Y_n|X_n \sim Q(dy|x)$ is not effected by control actions, hence one step observability applies directly to POMDPs. However, due to the dual effect of the control, the dependency structure of the POMDP is more complicated and the notion of $N>1$ step observability is only useful for filter stability in control problems when the DM cannot effect the transition kernel of the underlying Markov process. If this is the case, the POMDP to behaves like a POMDP and the stability results carry over.

Total variation merging and relative entropy merging are achieved with similar results to the POMP case, but the proofs are modified to avoid using the Markov property. As we will later see, the total variation merging of the predictor will prove useful when studying the robustness of the average cost problem to incorrect priors in Section 5.6

Chapter 4

Stability Via Ergodic Transition Kernels

4.1 Introduction

In this chapter, we consider a kernel "ergodic" if it brings measures together over time, that is $\|T^n(\mu) - T^n(\nu)\| \to 0$ as $n \to \infty$ for any measures μ and ν. We will look at different properties of the transition kernel T and measurement kernel Q that result in this ergodicity and how this can lead to filter stability.

In the previous chapters, stability follows from observability and Blackwell's martingale convergence results [5]. The measurements provide sufficient information over time about the underlying process, and inform the observer of the state realization. However, these results are asymptotic and do not provide a rate of convergence. In a control application, a system designer may want guarantees on rates of convergence to achieve sufficient merging in finite time.

A definition of stability with a rate attached to it is exponential stability.

Definition 4.1. *A POMP is said to be exponentially stable in total variation in* *expectation if there exists a coefficient* $0 < \alpha < 1$ *such that for any* $\mu \ll \nu$ *we have*

$$E^{\mu}[\|\pi^{\mu}_{n+1} - \pi^{\nu}_{n+1}\|_{TV}] \leq \alpha E^{\mu}[\|\pi^{\mu}_{n} - \pi^{\nu}_{n}\|_{TV}] \quad n \in \{0, 1, \cdots\}$$

With this definition we can guarantee how fast the filter is merging and for any $\epsilon > 0$ we can find provide a specific finite N such that

$$E^{\mu}[\|\pi^{\mu}_{n} - \pi^{\nu}_{n}\|_{TV}] \leq \epsilon \;\; \forall n > N$$

Let us first turn our attention to POMPs. Recall the filter update operator $\pi^{\mu}_{n+1} = \phi(\pi_n, y_{n+1})$ which is the composition of the transition kernel $T(\cdot)$ and the Bayesian update operator $\psi(\cdot, y_{n+1})$. In this chapter we will study when ϕ is a contraction in expectation, that is

$$E^{\mu}[\|\phi(\pi^{\mu}_{n+1}, y_{n+1}) - \phi(\pi^{\nu}_{n+1}, y_{n+1})\|_{TV}] \leq \alpha \|\pi^{\mu}_{n} - \pi^{\nu}_{n}\|_{TV}$$

for some $\alpha < 1$.

A common approach in the literature is the study ϕ directly as the aggregate of T and ψ and show that it is a contraction using the Hilbert metric [20]. However, this approach relies on a very restrictive assumption called the *mixing condition* that is not applicable to many systems one would want to analyse.

In our approach, we study the contraction properties of T and Q separately and then combine them. We arrive at a result that does not require the mixing condition. Our approach studies the Dobrushin coefficient of kernel operators, rather than the mixing condition and is much less restrictive. T is a linear operator and under mild conditions (i.e. a non-zero Dobrushin coefficient), a contraction. On the other hand, ψ is a non-linear operator and not a contraction since the expected posterior distributions may have larger total variation than their priors under a Bayesian update.

T has a well known contraction coefficient that is determined by its Dobrushin coefficient, and we derive an upper bound expansion coefficient for ψ that is determined by the Dobrushin coefficient of Q. If the product of the contraction coefficient and the expansion coefficient is less than one, then the composed operator is a contraction and we have our desired result.

However, there is a consequence to this approach. In deriving the upper bound expansion coefficient for ψ we apply the triangle inequality, and this step loses any benefit from the informative nature of Q. Q then becomes adversarial to T in that T is a contraction, while ψ is expansive. Counter-intuitively, our result prioritizes measurement channels Q that are un-informative, where $Q(dy|x)$ and $Q(dy|x')$ are similar distributions for $x, x' \in \mathcal{X}$. Such measurement channels have high Dobrushin coefficients and thus ψ as a lower expansion coefficient.

4.2 Hilbert Metric and Its Limitations

One common approach in the literature to achieve exponential stability is to utilize the Hilbert metric.

Definition 4.2. *[20, Definition 3.1] Two non-negative measures μ and ν on a measurable space $(S, \mathcal{F})$ are called comparable if $\exists\ 0 < a \leq b$ such that $\forall A \in \mathcal{F}$*

$$a\mu(A) \leq \nu(A) \leq b\mu(A)$$

Definition 4.3. *[20, Definition 3.2] A kernel $K : \mathcal{S}_1 \to \mathcal{P}(\mathcal{S}_2)$ is called mixing if there exists a finite non-negative measure $\lambda \in \mathcal{P}(\mathcal{S}_2)$ and $0 < \epsilon \leq 1$ such that* $\forall\ A \in \mathcal{B}(\mathcal{S}_2), s \in \mathcal{S}_1$

$$\epsilon\lambda(A) \leq K(s, A) \leq \frac{1}{\epsilon}\lambda(A)$$

Mixing is a very strong assumption on a kernel. For a kernel on a finite probability space (which is a stochastic matrix) the kernel is mixing if and only if each column of the matrix is fully zero or fully non-zero. For example

$$\begin{pmatrix} 0 & 0.25 & 0.75 \\ 0.25 & 0.25 & 0.5 \\ 0 & 0.1 & 0.9 \end{pmatrix}$$

is not a mixing kernel.

For a kernel $K : \mathcal{S}_1 \to \mathcal{P}(\mathcal{S}_2)$ which is non-degenerate with dominating measure λ and likelihood function $k(s_1, s_2)$, the kernel is mixing if and only if there exists two enveloping functions $f_1, f_2 \in L^1(\lambda)$ such that

$$0 < a \leq \frac{f_1(s_2)}{f_2(s_2)} \leq b < \infty\ \forall s_2 \in \mathcal{S}_2$$
$$f_1(s_2) \leq k(s_1, s_2) \leq f_2(s_2)\ \forall s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2$$

For example, if $K : \mathbb{R} \to \mathcal{P}(\mathbb{R})$ where $K(dx'|x) \sim N(f(x), \sigma)$ where $\|f\|_\infty < \infty$ then K is not a mixing kernel.

Definition 4.4. *[20, Definition 3.3] Let μ, ν be two non-negative finite measures. We* *define their Hilbert metric as*

$$h(\mu,\nu)=\begin{cases}\log\left(\frac{\sup_{A|\nu(A)>0}\frac{\mu(A)}{\nu(A)}}{\inf_{A|\nu(A)>0}\frac{\mu(A)}{\nu(A)}}\right) & \textit{if } \mu,\nu \textit{ are comparable}\\ 0 & \textit{if } \mu=\nu=0\\ \infty & \textit{else}\end{cases}$$

We see that the Hilbert metric is only meaningful when μ and ν are comparable. Yet comparablility implies mutual absolute continuity (i.e. $\mu \ll \nu$ and $\nu \ll \mu$) and therefore the Radon Nikodym derivatives $\frac{d\mu}{d\nu}$ and $\frac{d\nu}{d\mu}$ exist. Furthermore, these derivatives are bounded from above and bellow away from zero and

$$h(\mu,\nu)=\log\left(\left\|\frac{d\mu}{d\nu}\right\|_\infty\left\|\frac{d\nu}{d\mu}\right\|_\infty\right)$$

when the measures are comparable. The Hilbert metric is a projective distance, meaning if we scale either of the measures by a constant it will not change the Hilbert metric. This makes the metric very useful when studying the Bayesian update operator ψ since the denominator in a Bayesian update is a non-linear scaling operator, while the numerator is a linear operator.

Theorem 4.1. *[20, Corollary 4.2] Assume the measurement channel is non-degenerate. Let $\bar{\phi}$ represent the un-normalized filter update,*

$$\bar{\phi}(\mu,y)(dx)=g(x,y)T(\mu)(dx)$$

which is a kernel mapping to the space of non-negative finite measures and not necessarily the space of probability measures. If $\bar{\phi}$ is a mixing Kernel with coefficient

$\epsilon > 0\ \forall y \in \mathcal{Y}$ then

$$\|\pi^{\mu}_{n+m} - \pi^{\nu}_{n+m}\|_{TV} \leq \left(\frac{2}{\log(3)\epsilon^2}\right)\left(\frac{1-\epsilon^2}{1+\epsilon^2}\right)^{m-1} \|\pi^{\mu}_{n} - \pi^{\nu}_{n}\|_{TV} \tag{4.1}$$

Note that if T is a mixing kernel with coefficient ϵ, then $\bar{\phi}$ is as well but this can also be achieved without T begin mixing, see [20, Example 3.10]. However, there are a few undesirable points that follow from the Hilbert metric approach.

- Requiring $\bar{\phi}$ to be a mixing kernel is a very restrictive assumption.
- Equation (4.1) is not a strict contraction due the scaling term of $\frac{2}{\log(3)\epsilon^2}$ which can get quite large for $\epsilon \ll 1$.
- The "exponential coefficient" $\frac{1-\epsilon^2}{1+\epsilon^2}$ is close to 1 for most reasonable values of $\epsilon \ll 1$ and hence has a very slow rate of decay.

4.3 A New Approach Via Dobrushin Coefficients

Therefore, we see that an approach utilizing the Hilbert metric can yield exponential filter stability, but has undesirable properties. We will take a different approach by looking instead at the Dobrushin coefficient of a kernel operator.

Definition 4.5. *[16, Equation 1.16] For a kernel operator $K : S_1 \to \mathcal{P}(S_2)$ we define the Dobrushin coefficient as:*

$$\delta(K) = \inf \sum_{i=1}^{n} \min(K(x, A_i), K(y, A_i)) \tag{4.2}$$

where the infimum is over all $x, y \in S_1$ and all partitions $\{A_i\}_{i=1}^{n}$ of S_2.

Note this definition holds for continuous or finite/countable spaces S_1 and S_2. Note that $0 \leq \delta(K) \leq 1$ for any kernel operator. We then have for two probability measures $\pi, \pi' \in \mathcal{P}(S_1)$ [16]:

$$\|K(\pi) - K(\pi')\|_{TV} \leq (1 - \delta(K))\|\pi - \pi'\|_{TV}$$

As was discussed in Chapter 1, the filter update operator ϕ is a composition of the transition kernel T and the Bayesian update operator ψ. The transition operator T is a contraction mapping with coefficient $(1 - \delta(T))$, which potentially could be 1. Assume that it is less than 1, then without the Bayes' update the transition operator would being measures together with each successive application. However, the Bayes' operator is in general not a contraction, and can in fact increase the total variation distance between posteriors compared to the priors. We are therefore not guaranteed that the composition of the two operators is a contraction. However, if we have an upper bound on

$$\frac{E^{\mu}[\|\psi(\mu, y) - \psi(\nu, y)\|_{TV}]}{\|\mu - \nu\|_{TV}}$$

then if $\delta(T)$ is sufficiently large, the possible expansion property of ψ is dominated by the contraction property of T and the composed operator ϕ is itself a contraction in expectation.

Lemma 4.1. *Consider a true prior μ and a false prior ν with $\mu \ll \nu$. Assume that the measurement channel is non-degenerate, then we have that*

$$E^{\mu}[\|\psi(\mu, y) - \psi(\nu, y)\|_{TV}] \leq (2 - \delta(Q))\|\mu - \nu\|_{TV}$$

Proof. Let us now take a closer look at the operator ψ. For a general probability measure π define the normalizing constant:

$$N^{\pi}(y) = \int_{\mathcal{X}} g(x,y)\pi(dx)$$

and we have

$$\begin{aligned}
&\|\psi(\mu,y)-\psi(\nu,y)\|_{TV} = \sup_{\|f\|_\infty\le 1}\left|\int_{\mathcal{X}} \frac{f(x)g(x,y)}{N^{\mu}(y)}\mu(dx) - \int_{\mathcal{X}} \frac{f(x)g(x,y)}{N^{\nu}(y)}\nu(dx)\right| \\
&= \sup_{\|f\|_\infty\le 1}\left|\int_{\mathcal{X}} \frac{f(x)g(x,y)}{N^{\mu}(y)}\mu(dx) - \int_{\mathcal{X}} \frac{f(x)g(x,y)}{N^{\mu}(y)}\nu(dx) + \int_{\mathcal{X}} \frac{f(x)g(x,y)}{N^{\mu}(y)}\nu(dx)\right. \\
&\qquad\qquad \left. - \int_{\mathcal{X}} \frac{f(x)g(x,y)}{N^{\nu}(y)}\nu(dx)\right| \\
&\le \sup_{\|f\|_\infty\le 1} \frac{1}{N^{\mu}(y)}\left|\int_X f(x)g(x,y)(\mu-\nu)(dx)\right| \\
&+ \sup_{\|f\|_\infty\le 1}\left|\frac{1}{N^{\mu}(y)} - \frac{1}{N^{\nu}(y)}\right|\left|\int_{\mathcal{X}} f(x)g(x,y)\nu(dx)\right| \\
&\le \sup_{\|f\|_\infty\le 1} \frac{1}{N^{\mu}(y)}\left|\int_X f(x)g(x,y)\left(\frac{d\mu}{d\nu}(x)-1\right)\nu(dx)\right| + \left|\frac{N^{\nu}(y)-N^{\mu}(y)}{N^{\mu}(y)N^{\nu}(y)}\right| N^{\nu}(y) \\
&\le \left(\frac{1}{N^{\mu}(y)}\right)\left(|N^{\mu}(y)-N^{\nu}(y)| + \int_X g(x,y)\left|1-\frac{d\mu}{d\nu}(x)\right|\nu(dx)\right)
\end{aligned}$$

we then have

$$\begin{aligned}
&E^{\mu}[\|\psi(\mu,y)-\psi(\nu,y)\|_{TV}] = \int_X\int_Y \|\psi(\mu,y)-\psi(\nu,y)\|_{TV} Q(dy|x)\mu(dx) \\
&= \int_X\int_Y \|\psi(\mu,y)-\psi(\nu,y)\|_{TV} g(x,y)\lambda(dy)\mu(dx) \\
&= \int_Y \|\psi(\mu,y)-\psi(\nu,y)\|_{TV}\left(\int_X g(x,y)\mu(dx)\right)\lambda(dy) \\
&= \int_Y \|\psi(\mu,y)-\psi(\nu,y)\|_{TV} N^{\mu}(y)\lambda(dy)
\end{aligned}$$

$$
\begin{aligned}
&\leq \int_{\mathcal{Y}} \left(|N^\mu(y) - N^\nu(y)| + \int_X g(x,y) \left| 1 - \frac{d\mu}{d\nu}(x) \right| \nu(dx) \right) \lambda(dy) \\
&\leq \int_{\mathcal{Y}} |N^\mu(y) - N^\nu(y)| \lambda(dy) + \int_{\mathcal{Y}} \int_X g(x,y) \left| 1 - \frac{d\mu}{d\nu}(x) \right| \nu(dx)\lambda(dy) \\
&= \int_{\mathcal{Y}} \left| \int_{\mathcal{X}} g(x,y)(\mu - \nu)(dx) \right| \lambda(dy) + \int_X \left| 1 - \frac{d\mu}{d\nu}(x) \right| \left(\int_{\mathcal{Y}} g(x,y)\lambda(dy) \right) \nu(dx)
\end{aligned}
$$

Let us examine these two terms separately. For the second term, $g(x,y)$ is a probability density function for a fixed x, therefore it integrates to 1 over λ and we have

$$
\int_X \left| 1 - \frac{d\mu}{d\nu}(x) \right| \left(\int_{\mathcal{Y}} g(x,y)\lambda(dy) \right) \nu(dx) = \int_X \left| 1 - \frac{d\mu}{d\nu}(x) \right| \nu(dx) = \|\mu - \nu\|_{TV}
$$

for the first term, define the sets

$$
\begin{aligned}
S^+ &= \{y | \int_X g(x,y)(\mu - \nu)(dx) > 0\} \\
S^- &= \{y | \int_X g(x,y)(\mu - \nu)(dx) \leq 0\}
\end{aligned}
$$

then we have

$$
\begin{aligned}
&\int_{\mathcal{Y}} \left| \int_{\mathcal{X}} g(x,y)(\mu - \nu)(dx) \right| \lambda(dy) \\
&= \int_{S^+} \int_{\mathcal{X}} g(x,y)(\mu - \nu)(dx)\lambda(dy) - \int_{S^-} \int_{\mathcal{X}} g(x,y)(\mu - \nu)(dx)\lambda(dy) \\
&= \int_{\mathcal{Y}} (1_{S^+}(y) - 1_{S^-}(y))\, g(x,y)(\mu - \nu)(dx)\lambda(dy)
\end{aligned}
$$

We then have that $1_S^+(y) - 1_{S^-}(y)$ is a measurable function of y with infinity norm equal to 1, and in fact it achieves the supremum overall such functions. That is

$$\begin{aligned}\int_{\mathcal{Y}} \left(1_{S^+}(y) - 1_{S^-}(y)\right) g(x,y)(\mu-\nu)(dx)\lambda(dy) &= \sup_{\|f\|_\infty \le 1} \left| \int_{\mathcal{Y}} f(y) g(x,y)(\mu-\nu)(dx)\lambda(dy) \right| \\ &= \|Q(\mu) - Q(\nu)\|_{TV} \\ &\le (1-\delta(Q))\|\mu-\nu\|_{TV}\end{aligned}$$

putting these together,

$$\begin{aligned}E^\mu[\|\psi(\mu,y) - \psi(\nu,y)\|_{TV}] &\le (1-\delta(Q))\|\mu-\nu\|_{TV} + \|\mu-\nu\|_{TV} \\ &= (2-\delta(Q))\|\mu-\nu\|_{TV}\end{aligned}$$

□

Theorem 4.2. *Assume that $\mu \ll \nu$ and that the measurement channel is non-degenerate. Then we have*

$$E^\mu[\|\pi^\mu_{n+1} - \pi^\nu_{n+1}\|_{TV}] \le (1-\delta(T))(2-\delta(Q))E^\mu[\|\pi^\mu_n - \pi^\nu_n\|_{TV}]$$

Proof.

$$\begin{aligned}&E^\mu[\|\pi^\mu_{n+1} - \pi^\nu_{n+1}\|_{TV}] = E^\mu[\|\phi(\pi^\mu_n, y_{n+1}) - \phi(\pi^\nu_n, y_{n+1})\|_{TV}] \\ &= E^\mu[\|\psi(T(\pi^\mu_n), y_{n+1}) - \psi(T(\pi^\nu_n), y_{n+1})\|_{TV}] \\ &= \int_{\mathcal{Y}^{n+2}} \|\psi(T(\pi^\mu_n), y_{n+1}) - \psi(T(\pi^\nu_n), y_{n+1})\|_{TV} P^\mu(dy_{[0,n+1]}) \\ &= \int_{\mathcal{Y}^{n+1}} \int_{\mathcal{X}} \int_{\mathcal{Y}} \|\psi(T(\pi^\mu_n), y_{n+1}) - \psi(T(\pi^\nu_n), y_{n+1})\|_{TV} P^\mu(dy_{n+1}|x_{n+1}, y_{[0,n]}) P^\mu(dx_{n+1}|y_{[0,n]}) P^\mu(dy_{[0,n]}) \\ &= \int_{\mathcal{Y}^{n+1}} \int_{\mathcal{X}} \int_{\mathcal{Y}} \|\psi(T(\pi^\mu_n), y_{n+1}) - \psi(T(\pi^\nu_n), y_{n+1})\|_{TV} Q(dy_{n+1}|x_{n+1}) T(\pi^\mu_n)(dx) P^\mu(dy_{[0,n]})\end{aligned}$$

$$= \int_{\mathcal{Y}^{n+1}} E^{T(\pi_n^\mu)}[\|\psi(T(\pi_n^\mu), y_{n+1}) - \psi(T(\pi_n^\nu), y_{n+1})\|_{TV}] P^\mu(dy_{[0,n]})$$
$$\leq (2 - \delta(Q)) \int_{\mathcal{Y}^{n+1}} \|T(\pi_n^\mu) - T(\pi_n^\nu)\|_{TV} P^\mu(dy_{[0,n]})$$
$$\leq (2 - \delta(Q))(1 - \delta(T)) \int_{\mathcal{Y}^{n+1}} \|\pi_n^\mu - \pi_n^\nu\|_{TV} P^\mu(dy_{[0,n]})$$
$$= (2 - \delta(Q))(1 - \delta(T)) E^\mu[\|\pi_n^\mu - \pi_n^\nu\|_{TV}]$$

□

Corollary 4.1. *Assume $\mu \ll \nu$ and that the measurement channel is non-degenerate. If we have*

$$\alpha = (1 - \delta(T))(2 - \delta(Q)) \leq 1 \tag{4.3}$$

then the filter is exponentially stable in total variation in expectation with coefficient α and

$$E^\mu[\|\pi_n^\mu - \pi_n^\nu\|_{TV}] \leq (2 - \delta(Q)) (\alpha^n) \|\mu - \nu\|_{TV}$$

Furthermore, if $\delta(T) > \frac{1}{2}$ then $\alpha < 1$ and the POMP is exponentially stable regardless of the measurement kernel Q.

Proof. By recursive application of Theorem 4.2 we have

$$E^\mu[\|\pi_n^\mu - \pi_n^\nu\|_{TV}] \leq \alpha^n E^\mu[\|\pi_0^\mu - \pi_0^\nu\|_{TV}]$$

π_0^μ is then the Bayesian update of μ under the first observation Y_0, therefore we apply

Lemma 4.1 and we have

$$\alpha^n E^\mu[\|\pi_0^\mu - \pi_0^\nu\|_{TV}] = \alpha^n E^\mu[\|\psi(\mu, y_0) - \psi(\nu, y_0)\|_{TV}] \leq (2 - \delta(Q))(\alpha^n)\|\mu - \nu\|_{TV}$$

Finally, recall that for any kernel K we have $0 \leq \delta(K) \leq 1$ therefore if we have $\delta(T) > \frac{1}{2}$

$$\alpha = (1 - \delta(T))(2 - \delta(Q)) < \frac{1}{2}(2 - \delta(Q)) \leq \frac{2}{2} = 1$$

□

Note that this result is sufficient, but certainly not necessary. In what seems like a counter-intuitive result, this result prioritizes measurement channels Q that are un-informative as opposed to those that are informative. For example a completely independent observation Y will have $\delta(Q) = 1$ and direct observation will have $\delta(Q) = 0$. However, the idea of this result is that T is sufficiently ergodic in that without any Bayes' update, the mapping T is a contraction and would bring measures together. We then want a transition kernel Q that does not "change" this ergodic property, and a completely independent observation will result in $\psi(\mu) = \mu$, and hence will not conflict with the transition kernel.

For example, consider a finite system and direct observation. That is y is an invertible deterministic function of x, $Y = h(X)$. Then we have

$$\begin{aligned}
\|\psi(\mu, y) - \psi(\nu, y)\| &= \sup_{\|f\|_\infty \leq 1} \left| \sum_{x \in \mathcal{X}} f(x) g(x, y) \left(\frac{\mu(x)}{N^\mu(x)} - \frac{\nu(x)}{N^\nu(y)} \right) \right| \\
&= \sup_{\|f\|_\infty \leq 1} \left| \sum_{x \in \mathcal{X}} f(x) 1_{h^{-1}(y)}(x) \left(\frac{\mu(x)}{\mu(h^{-1}(y))} - \frac{\nu(x)}{\nu(h^{-1}(y))} \right) \right|
\end{aligned}$$

$$
\begin{aligned}
&= \sup_{\|f\|_\infty \le 1} \left| f(h^{-1}(y)) \left(\frac{\mu(h^{-1}(y))}{\mu(h^{-1}(y))} - \frac{\nu(h^{-1}(y))}{\nu(h^{-1}(y))} \right) \right| \\
&= 0
\end{aligned}
$$

However, if we add and subtract $\frac{\mu(h^{-1}(y))}{\nu(h^{-1}(y))}$ in the first line and apply the triangle inequality we have:

$$
\begin{aligned}
&\left(\frac{1}{\mu(h^{-1}(y))} \right) \sup_{\|f\|_\infty \le 1} \left| f(h^{-1}(y))(\mu(h^{-1}(y)) - \nu(h^{-1}(y)) \right| \\
&+ |f(h^{-1}(y))\nu(h^{-1}(y))| \left| \frac{\mu(h^{-1}(y)) - \nu(h^{-1}(y))}{\nu(h^{-1}(y))} \right| \\
&= \left(\frac{1}{\mu(h^{-1}(y))} \right) \left| (\mu(h^{-1}(y)) - \nu(h^{-1}(y)) \right| + \left| \mu(h^{-1}(y)) - \nu(h^{-1}(y)) \right| \\
&\neq 0
\end{aligned}
$$

this is the same approach taken in the proof of Lemma 4.1. We see that the triangle inequality results in a loose bound that ignores the informative nature of the measurement channel, and thus Theorem 4.2 relies on the ergodic properties of the transition kernel to achieve exponential filter stability.

4.4 Controlled System

These results easily extend to a POMDP. In a controlled environment the measurement channel Q is unchanged, however the transition kernel $T(dx'|x, u)$ is different for each control action u. If we define

$$
\tilde{\delta}(T) = \inf_{u \in \mathcal{U}} \delta(T(\cdot|\cdot, u))
$$

then the result for a POMDP follows easily.

Corollary 4.2. *Assume $\mu \ll \nu$ and that the measurement channel is non-degenerate. If we have*

$$\alpha = (1 - \tilde{\delta}(T))(2 - \delta(Q)) < 1$$

then the filter is universally exponentially stable with coefficient α.

Therefore, in order to guarantee exponential stability in a control environment we first check the expansion coefficient of the Bayesian update operator $(2 - \delta(Q))$. Then, we find the Dobrushin coefficient of $T(\cdot|\cdot, u)$ for every difference control action u. If under each control action $T(\cdot|\cdot, u)$ has a high enough Dobrushin coefficient, then for every control action the filter update operator is a contraction in total variation in expectation. It then does not matter what control policy is implemented, since each control action results in a transition kernel with a sufficient high Dobrushin coefficient, and thus we have universal exponential stability.

4.5 Summary

To achieve an exponential rate of filter stability, observability of the measurement channel is not enough. It is more useful to study the contraction properties of the composed filter update operator ϕ which is the composition of the transition kernel T and the Bayesian update operator ψ. One such approach common in the literature is to utilize the Hilbert metric, which is a projective distance hence we can study ϕ without the non-linear normalizing term that arises in the Bayesian update. However, this approach requires the update kernel ϕ to have a mixing coefficient [20] which greatly reduces the types of kernels which can be studied with this approach.

We instead focus on the Dobrushin coefficient, which is a much more general measure of ergodicity for a kernel. It is well known that the transition kernel T is a contraction with coefficient $1 - \delta(T)$ [16]. We further show in Lemma 4.1 that the Bayesian update operator, while not being a contraction, has an upper bound $2-\delta(Q)$ on how expansive it is based on the Dobrushin coefficient of Q. If the product of these coefficient is less than one, than the composed operator is a contraction and we achieve exponential filter stability. Furthermore, if $\delta(T) > \frac{1}{2}$ than we have exponential stability regardless of the measurement kernel.

Stochastic kernels where the conditional measures $K(dy|x) \approx K(dy|x')$ for different x, x' have high Dobrushin coefficients, while kernels with very diverse conditional measures have lower Dobrushin coefficients. Therefore, our result prioritizes transition kernels T and Q with high Dobrushin coefficients. This is in many ways counter-intuitive, as an informative measurement channel Q would have a low Dobrushin coefficient and thus be undesirable. Yet this is because our results rely primarily on the transition kernel bringing measures together quickly, and an informative kernel has the possibility of interfering with that process. On the other hand, an uninformative kernel will not greatly change the measure in the Bayesian update, and thus the ergodic transition kernel T will dominate the composition in the filter update operator.

Chapter 5

Filter Stability and Robustness in POMDPs

5.1 Introduction

In this chapter we will utilize our stability results to consider the expected cost incurred by an incorrectly designed control policy. As noted in Chapter 1, we consider three different control problems: the single stage cost problem, the infinite horizon discount cost problem, and the infinite horizon average cost problem. In each case, the DM makes control actions based on its observations and incurs a cost $J(\mu, \gamma)$ which is a function of the initial measure μ of the POMDP and the control policy γ implemented.

An optimal control policy γ^{μ} for a given prior μ is the control policy which achieves the infimum of the expected cost over all admissible control policies, $J(\mu, \gamma^{\mu}) = \inf_{\gamma\in\Gamma} J(\mu, \gamma) \equiv J^{*}(\mu)$. In this work we will not be interested in the existence of such optimal policies, which is not always guaranteed. Throughout this chapter we will assume that

- Optimal policies exist for the priors being discussed
- Optimal policies are stationary functions of the filter realization, that is there

exists a mapping $\Phi : \mathcal{P}(\mathcal{X}) \to \mathbb{U}$ where

$$\gamma_n^{\mu}(y_{[0,n]}) = \Phi(\pi_n^{\mu,\gamma^{\mu}})$$

These assumptions are justified for the kinds of problems we study in this chapter. Under weak continuity and measurable selection conditions, for both finite and infinite horizon problems we can establish the existence of optimal policies. For infinite horizon discounted cost problems, by studying the measurable selection criteria we can use dynamic programming to establish optimal solutions for finite horizon discounted problems [23]. By taking the limit as the horizon goes to infinity, we see that the optimal policy is a stationary mapping of the filter realization that solves a fixed point equation

$$v(\mu) = \min_{u \in \mathcal{U}} \left(E^{\mu}[c(X,u)] + \beta E^{\mu}(v(\phi(\mu,u,Y_1)) | U_0 = u] \right) \;\; \forall \mu \in \mathcal{P}(\mathcal{X})$$

Under suitable regularity conditions, such as the weak Feller property of the controlled transition kernel, the compactness of the state space, and the continuity of the cost function, the existence of an optimal policy for the average cost problem has been established by the convex analytic method of Borkar [7], also see [23]. Through recent results in filtering theory which ensures conditions under which the filter process is weak Feller [25],[19], we can conclude that an optimal solution exists. However, we note that in these studies an optimal policy, while being a stationary mapping of the filter, may be a randomized mapping. Nonetheless, the existence question has been more or less settled and in this chapter, we will assume that an optimal policy exists as a stationary function of the filter realization.

Consider a setup where a DM thinks the initial measure is ν, when in reality the measure is μ. The DM will then implement the control policy γ^ν and incur a cost $J(\mu, \gamma^\nu)$, but if the DM had known the correct prior the cost could have been $J^*(\mu)$. The question of robustness as we are interested studies the difference between these two terms

$$J(\mu, \gamma^\nu) - J^*(\mu) \tag{5.1}$$

and quantifies the cost the DM incurs for implementing the incorrect policy compared to the optimal policy.

Robustness is studied in [26] for the single stage cost problem and the infinite horizon discounted cost problem. The paper provides conditions for when (5.1) goes to 0 as $\nu \to \mu$ in either weak convergence or total variation distance. We note two useful results:

Proposition 5.1. *[26, Proposition 3.1] Assume the cost function c is bounded, non-negative, and measurable. Let γ^ν be the optimal control policy designed with respect to a prior ν. Then we have*

$$J(\mu, \gamma^\nu) - J^*(\mu) \leq 2\|c\|_\infty \|\mu - \nu\|_{TV}$$

Theorem 5.1. *[26, Proposition 3.2] Assume the cost function c is bounded, non-negative, and measurable. Let γ^ν be the optimal control policy designed with respect to a prior ν. Then we have*

$$J^*_\beta(\mu, \gamma^\nu) - J^*_\beta(\mu) \leq 2\frac{\|c\|_\infty}{1-\beta}\|\mu - \nu\|_{TV}$$

While the average cost case is not directly studied in [26], we can adapt the technique to achieve a similar result for the average cost problem

Theorem 5.2. *Assume the cost function c is non-negative, bounded, and measurable. Then for two priors μ and ν with optimal policies γ^μ and γ^ν we have*

$$J_\infty(\mu, \gamma^\nu) - J^*_\infty(\mu) \leq 2\|c\|_\infty \|\mu - \nu\|_{TV}$$

Proof. Consider the robustness difference

$$\begin{aligned} J_\infty(\mu, \gamma^\nu) - J^*_\infty(\mu) &= J^*_\infty(\nu) - J^*_\infty(\mu) + J_\infty(\mu, \gamma^\nu) - J^*_\infty(\nu) \\ &\leq |J^*_\infty(\nu) - J^*_\infty(\mu)| + J_\infty(\mu, \gamma^\nu) - J_\infty(\nu, \gamma^\nu) \end{aligned} \tag{5.2}$$

Consider now the difference $|J^*_\infty(\mu) - J^*_\infty(\nu)|$. If $J^*_\infty(\mu) < J^*_\infty(\nu)$ then the absolute value of their difference is the larger value $J^*_\infty(\nu)$ minus the smaller value $J^*_\infty(\mu)$. Then $J^*_\infty(\nu) = J_\infty(\nu, \gamma^\nu)$ since γ^ν is the optimally policy which achieves the infimum over all policies. If we replace γ^ν with γ^μ we have $J_\infty(\mu, \gamma^\mu) > J^*_\infty(\nu)$ and the difference is even greater. Therefore,

$$|J^*_\infty(\mu) - J^*_\infty(\nu)| \leq J_\infty(\nu, \gamma^\mu) - J_\infty(\mu, \gamma^\mu)$$

on the other hand, if we have if $J^*_\infty(\mu) > J^*_\infty(\nu)$ by the same arguement

$$|J^*_\infty(\mu) - J^*_\infty(\nu)| \leq J_\infty(\mu, \gamma^\nu) - J_\infty(\nu, \gamma^\nu)$$

therefore

$$|J_\infty^*(\mu) - J_\infty^*(\nu)| \le \max\left(J_\infty(\mu,\gamma^\nu) - J_\infty(\nu,\gamma^\nu), J_\infty(\nu,\gamma^\mu) - J_\infty(\mu,\gamma^\mu)\right)$$

and we can ultimately determine the robustness bound by studying the expected cost operator $J_\infty(\cdot,\gamma^\nu)$ under different priors but the same control policy.

By same control policy, we mean γ^ν is the optimal control policy designed with respect to the prior ν. This means the DM sees observations $y_{[0,n]}$, computes the filter believing the prior is ν, and then employs the control actions. When we study $J(\mu,\gamma^\nu)$ and $J(\nu,\gamma^\nu)$ we study two different DM's experiencing different measurement realizations, but both believing the prior is ν. This does not mean that both chains experience some kind of coupled control actions, but only that the mapping from measurements to control actions is the same for each POMDP with different priors.

Note that for two non-negative bounded sequences $0 < a_n < m < \infty$ and $0 < b_n < k < \infty$ we have that the difference of their limsup's is less than the limsup of the difference

$$\limsup_{n\to\infty} a_n - \limsup_{n\to\infty} b_n \le \limsup_{n\to\infty}(a_n - b_n)$$

therefore since c is non-negative and bounded we have

$$\begin{aligned}
&J_\infty(\mu,\gamma^\nu) - J_\infty(\nu,\gamma^\nu)\\
&= \limsup_{T\to\infty} \frac{1}{T}\left(\sum_{i=0}^{T-1} E^{\mu,\gamma^\nu}[c(X_i,U_i)]\right) - \limsup_{T\to\infty} \frac{1}{T}\left(\sum_{i=0}^{T-1} E^{\nu,\gamma^\nu}[c(X_i,U_i)]\right)\\
&\le \limsup_{T\to\infty} \frac{1}{T}\sum_{i=0}^{T-1}\left(E^{\mu,\gamma^\nu}[c(X_i,U_i)] - E^{\nu,\gamma^\nu}[c(X_i,U_i)]\right)
\end{aligned}$$

$$\leq \limsup_{T\to\infty} \frac{\|c\|_\infty}{T} \left(\sum_{i=0}^{T-1} \|P^{\mu,\gamma^\nu}((X_i, U_i) \in \cdot) - P^{\nu,\gamma^\nu}((X_i, U_i) \in \cdot)\|_{TV} \right)$$

then we see that

$$\begin{aligned}
&\|P^{\mu,\gamma^\nu}((X_i, U_i) \in \cdot) - P^{\nu,\gamma^\nu}((X_i, U_i) \in \cdot)\|_{TV} \\
&= \sup_{\|f\|_\infty \leq 1} \left| \int_{\mathcal{X}\times\mathcal{U}} f(x,u) P^{\mu,\gamma^\nu}(dx_i, du_i) - \int_{\mathcal{X}\times\mathcal{U}} f(x,u) P^{\nu,\gamma^\nu}(dx_i, du_i) \right| \\
&= \sup_{\|f\|_\infty \leq 1} \left| \int_{\mathcal{X}} \int_{\mathcal{X}\times\mathcal{U}} f(x,u) P^{\mu,\gamma^\nu}(dx_i, du_i | X_0) \mu(dx_0) - \int_{\mathcal{X}} \int_{\mathcal{X}\times\mathcal{U}} f(x,u) P^{\nu,\gamma^\nu}(dx_i, du_i | X_0) \nu(dx_0) \right|
\end{aligned} \tag{5.3}$$

As was discussed, both POMDP use the same control policy, which maps measurements to control actions in the same way. Once we fix the realization of $X_0 = x$ in either case, the distribution on Y_0 is the same, hence the distribution on U_0 is the same, and hence X_1, Y_1, U_1 and so on. Therefore in (5.3) the two inner integrals are the same function of x and we can upper bound by

$$\sup_{\|\tilde{f}\|_\infty \leq 1} \left| \int_X \tilde{f}(x_0) \mu(dx_0) - \int_{\mathcal{X}} \tilde{f}(x_0) \nu(dx_0) \right| = \|\mu - \nu\|_{TV}$$

therefore

$$J_\infty(\mu, \gamma^\nu) - J_\infty(\nu, \gamma^\nu) \leq \limsup_{T\to\infty} \frac{\|c\|_\infty}{T} \sum_{i=0}^{T-1} \|\mu - \nu\|_{TV} = \|c\|_\infty \|\mu - \nu\|_{TV}$$

therefore when the policy is the same and the priors different, $J(\mu, \gamma^\nu) - J(\nu, \gamma^\nu)$ is upper bound by the norm of the cost function and the total variation difference of the priors. Both terms in (5.2) have this bound and the overall bound is multiplied by a factor of 2. □

If the filter is stable, then an incorrectly initialized filter will merge with the correctly initialized filter. Since the optimal policy is a stationary function of the filter, these merging filters act like new priors for the control problem and we can apply the robustness results from [26] to bound the difference (5.1).

5.2 Incorrect Initialization Costs

Consider the infinite horizon discounted cost problem. Fix a prior μ and a control policy γ^ν which is optimal with respect to a different prior ν. If we break the conditioning at time n we have

$$\begin{aligned} J_\beta(\mu,\gamma^\nu) &= E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{\infty}\beta^i c(x_i,u_i)\right] \\ &= E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right] + E^{\mu,\gamma^\nu}\left[E^{\mu,\gamma^\nu}\left[\sum_{i=n}^{\infty}\beta^i c(x_i,u_i)|Y_{[0,n-1]}\right]\right] \\ &= E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right] + (\beta^n)E^{\mu,\gamma^\nu}\left[E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{\infty}\beta^i c(x_{n+i},u_{n+i})|Y_{[0,n-1]}\right]\right] \end{aligned} \tag{5.4}$$

Consider the 0^{th} time stage in the problem. X_0 is distributed according to μ, an observation Y_0 is made, and the control action $u_0 = \gamma_0^\mu(\pi_0^{\nu,\gamma^\nu})$ is a function of the filter realization believing the prior is ν. The filter realization is the Bayesian update of the prior ν under the measurement Y_0.

Now consider the n^{th} time stage. Conditioned on $Y_{[0,n-1]}$ the distribution of X_n is $\pi_{n-}^{\mu,\gamma^\nu}$ (the true predictor). The control action $u_n = \gamma_n^\nu(\pi_n^{\nu,\gamma^\nu}) = \Phi(\pi_n^{\nu,\gamma^\nu})$ is a stationary function of the false filter realization. Yet the false filter at time n is the false predictor $\pi_{n-}^{\nu,\gamma^\nu}$ put through a Bayesian update under the measurement Y_n. Note that throughout we assume that $\nu \ll \mu$, and this ensures that the observer who believes the prior is ν will not witness observations that have probability 0 under P^{ν,γ^ν} and thus the filter will be well defined.

Therefore, the optimal policy under prior ν at time n is the same as the optimal policy under the prior $\nu' = \pi_{n-}^{\nu,\gamma^\nu}$ at time 0 since it is a stationary function of the filter realization. We see that the true predictor at time n acts as a new prior for a restarted control problem, and the new control policy is optimal with respect to the false filter $\nu' = \pi_{n-}^{\nu,\gamma^\nu}$. We then have

$$E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{\infty}\beta^i c(x_{i+n},u_{i+n})|Y_{[0,n-1]}\right] = J_\beta(\pi_{n-}^{\mu,\gamma^\nu},\gamma^{\nu'})$$

and therefore

$$J_\beta(\mu,\gamma^\nu) = E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right] + (\beta^n)E^{\mu,\gamma^\nu}\left[J_\beta(\pi_{n-}^{\mu,\gamma^\nu},\gamma^{\nu'})\right] \tag{5.5}$$

If we instead apply the correctly designed policy γ^μ and let $\mu' = \pi_{n-}^{\mu,\gamma^\mu}$ be the correctly initialized predictor we have

$$\begin{aligned} J_\beta^*(\mu) &= E^{\mu,\gamma^\mu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right] + (\beta^n)E^{\mu,\gamma^\mu}\left[J_\beta(\pi_{n-}^{\mu,\gamma^\mu},\gamma^{\mu'})\right] &(5.6)\\ &= E^{\mu,\gamma^\mu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right] + (\beta^n)E^{\mu,\gamma^\mu}\left[J_\beta^*(\pi_{n-}^{\mu,\gamma^\mu})\right] &(5.7)\end{aligned}$$

combining the two

$$\begin{aligned}
&J_\beta(\mu,\gamma^\nu) - J^*_\beta(\mu)\\
=&E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right] - E^{\mu,\gamma^\mu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right]\\
+&\beta^n\left(E^{\mu,\gamma^\nu}\left[J_\beta(\pi^{\mu,\gamma^\nu}_{n-},\gamma^{\nu'})\right] - E^{\mu,\gamma^\mu}\left[J^*_\beta(\pi^{\mu,\gamma^\mu}_{n-})\right]\right)\\
=&E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right] - E^{\mu,\gamma^\mu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right]\\
+&\beta^n\left(E^{\mu,\gamma^\nu}\left[J_\beta(\pi^{\mu,\gamma^\nu}_{n-},\gamma^{\nu'}) + J^*_\beta(\pi^{\mu,\gamma^\nu}_{n-}) - J^*_\beta(\pi^{\mu,\gamma^\nu}_{n-})\right] - E^{\mu,\gamma^\mu}\left[J^*_\beta(\pi^{\mu,\gamma^\mu}_{n-})\right]\right)\\
=&E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right] - E^{\mu,\gamma^\mu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right] \qquad (5.8)\\
+&\beta^n\left(E^{\mu,\gamma^\nu}\left[J^*_\beta(\pi^{\mu,\gamma^\nu}_{n-})\right] - E^{\mu,\gamma^\mu}\left[J^*_\beta(\pi^{\mu,\gamma^\mu}_{n-})\right]\right) \qquad (5.9)\\
+&\beta^n\left(E^{\mu,\gamma^\nu}\left[J_\beta(\pi^{\mu,\gamma^\nu}_{n-},\gamma^{\nu'}) - J^*_\beta(\pi^{\mu,\gamma^\nu}_{n-})\right]\right) \qquad (5.10)
\end{aligned}$$

Therefore, we see that there are in general three costs associated with applying an incorrectly designed control policy to a control system. The first cost (5.8) can be thought of as the "past mistakes" cost. At time $[0, n-1]$ the control policy γ^ν makes different control decisions than the optimal policy γ^μ. As such, the costs incurred from time $[0, n-1]$ will be different for the optimal control policy and the incorrectly designed policy.

The second cost (5.9) is the "strategic measure cost". It is analogous to the DM operating under it's own false assumptions from time $[0, n-1]$ and then at time n the DM "wakes up" and realizes that the true prior is in fact μ, not ν. As such, from time n onwards the DM will make optimal control decisions, however it cannot change the past control decisions $U_{[0,n-1]}$. As such, P^{μ,γ^ν} and P^{ν,γ^μ} have different strategic measures on $(X,Y)_{[n,\infty)}$ and thus incur different costs over the future time stages even if from time n onwards the DM makes optimal control decisions.

This mirrors the dynamic programming approach taken to compute optimal control policies. When choosing a control action at a given time stage n, there are two costs to consider. The first cost is the current cost: the expected cost that will be incurred at time stage n due to a control action. The second cost is the cost to go: the expected costs to be incurred in the future based on the control action being made. In an optimal control policy, these two costs are balanced. If a policy is greedy and makes the control action at every time stage that minimizes the current cost, it may result in a poor strategic measure on the future states and thus incur a higher cost to go. Similarly, if a policy is too forward thinking and only focuses on the cost to go, then the policy may incur very high cost in the current time stage that outweigh the benefits of lower expected future costs.

The optimal policy γ^μ, by nature of being the infimum over all polices, has the lowest overall cost therefore

$$E^{\mu,\gamma^\mu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right] + \beta^n E^{\mu,\gamma^\mu}\left[J^*_\beta(\pi_{n-}^{\mu,\gamma^\mu})\right] \le E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right] + \beta^n E^{\mu,\gamma^\nu}\left[J^*_\beta(\pi_{n-}^{\mu,\gamma^\nu})\right]$$

the sum of the current costs for time $[0,n-1]$ and the cost to go is less for the optimal policy (LHS) than a policy which implements the incorrectly designed policy for the first $[0,n-1]$ time stages and then makes optimal decisions from time n onwards (RHS). If we re-arrange this

$$\begin{aligned} 0 \le &\left(E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right] - E^{\mu,\gamma^\mu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right]\right) \\ &+\beta^n\left(E^{\mu,\gamma^\nu}\left[J^*_\beta(\pi_{n-}^{\mu,\gamma^\nu})\right] - E^{\mu,\gamma^\mu}\left[J^*_\beta(\pi_{n-}^{\mu,\gamma^\mu})\right]\right) \end{aligned}$$

we see that the sum of the past mistakes cost and the strategic measure cost is greater than zero, but that does not mean that each cost individually is greater than 0. For example, $E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right]$ may be less than $E^{\mu,\gamma^\mu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right]$, but then the incorrect control policy will pay a higher price in the strategic measure cost and the sum of the two costs will be greater than zero.

The third cost (5.10) is the "approximation cost." The DM does not actually wake up at time stage n and realize that the true prior is μ rather than ν. Instead, under predictor stability the falsely initialized predictor $\pi_{n-}^{\nu,\gamma^\nu}$ and the true predictor $\pi_{n-}^{\mu,\gamma^\nu}$ merge as time goes on which has a similar effect. However the true and false predictors are still slightly different and hence the DM incurs the approximation cost. The results in [26] studying the robustness of optimal costs under merging priors paired with predictor stability easily show that the approximation cost (5.10) goes to zero as the predictors merge.

These three costs form the fundamentals of studying the losses a DM incurs for using an incorrectly designed policy. We can study different control problems that effect these three costs in different ways, and achieve a notion of robustness: that the DM does not suffer greatly for using a poorly designed policy under filter stability.

The past mistakes cost essentially cannot be avoided, unless the problem is designed in such a way that this cost is unimportant (see for example the section on trained control problems or the infinite horizon average cost problem). It takes time for the predictors to merge, and as such the DM's mistakes in the first few time steps simply cannot be avoided until the DM has had time to learn the correct predictor realization.

The second cost similarly has less to do with the filter stability, and more to do with the ergodic nature of the filter update process. If every optimal cost is the same, that is $J^*_\beta(\mu) = J^*_\beta(\nu)$ for any prior μ or ν, then this cost goes to zero. For the discounted cost problem, this is unlikely, but for the average cost problem this could be possible if the probability measure valued Markov chain defined by the filter update admits a unique invariant measure.

The approximation cost is directly related to filter convergence. As we see in [26], via robustness if $\pi^{\mu,\gamma^\nu}_{n-}$ and $\pi^{\nu,\gamma^\nu}_{n-}$ merge in total variation we have $J_\beta(\pi^{\mu,\gamma^\nu}_{n-}, \gamma^{\pi^{\nu,\gamma^\nu}_{n-}}) \to J^*_\beta(\pi^{\mu,\gamma^\nu}_{n-})$.

5.3 Trained Control Problems

A control problem where the DM has sufficient time to learn the proper filter realization before making any control actions will be called a *trained control problem.* The simplest control problem we can consider is a trained one shot control problem. Say the DM is given a POMP where they cannot influence the development of the state process and only observe the measurements. They are given a fixed number N and are allowed to observe the measurements $Y_{[0,N]}$ and then at time N, they must make a control action U. The DM's objective is to minimize $E^\mu[c(X_N, U)]$. A diagram of the problem is presented in Figure 5.1.

Now assume the DM falsely believes the prior is ν, when it is in fact μ and we will study the robustness difference (5.1). In this problem, there is no past mistakes cost since the DM does not make control actions or incur costs for the first N time stages. Similarly, there is no strategic measure cost since the system evolves uncontrolled up until time N, therefore the DM cannot effect the measure on $Y_{[0,N]}$ that determines the filter realization. Therefore, the only penalty the DM suffers for using the incorrectly

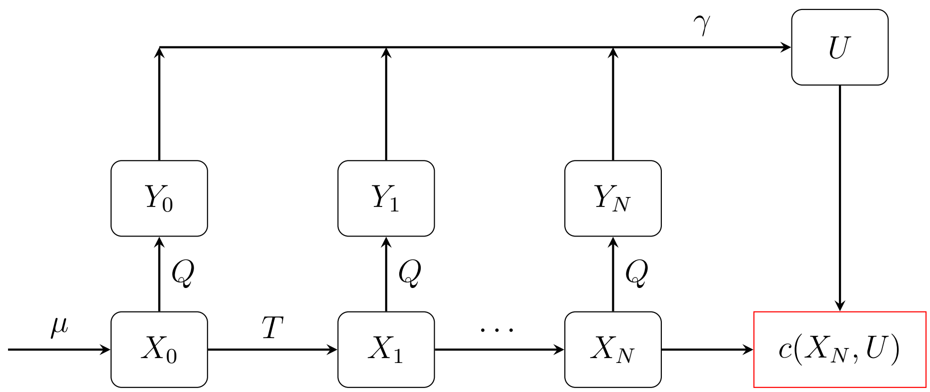

Figure 5.1: Trained One Shot Control Problem

designed policy in this problem is the approximation cost, and this can be made sufficiently small by filter stability (note exponential stability of the predictor and filter are equivalent).

Theorem 5.3. *Assume the cost function c is bounded, non-negative, and measurable and assume the filter is exponentially stable in total variation in expectation with coefficient α. Then*

$$E^{\mu}[c(X_N, \gamma^{\nu}(Y_{[0,N]}))] - E^{\mu}[c(X_N, \gamma^{\mu}(Y_{[0,N]}))] \leq 4\|c\|_{\infty}\alpha^N$$

Proof.

$$\begin{aligned}
&E^{\mu}[c(X_N, \gamma^{\nu}(Y_{[0,N-1]}))] - E^{\mu}[c(X_N, \gamma^{\mu}(Y_{[0,N-1]}))] \\
=&E^{\mu}\left[E^{\mu}[c(X_N, \gamma^{\nu}(Y_{[0,N]}))|Y_{[0,N-1]}] - E^{\mu}[c(X_N, \gamma^{\mu}(Y_{[0,N]}))|Y_{[0,N-1]}]\right] \\
=&E^{\mu}\left[J(\pi^{\mu}_{N-}, \gamma^{\pi^{\nu}_{N-}}) - J^*(\pi^{\mu}_{N-})\right]
\end{aligned}$$

by Proposition 5.1 we have

$$J(\pi^{\mu}_{N-}, \gamma^{\pi^{\nu}_{N-}}) - J^*(\pi^{\mu}_{N-}) \leq 2\|c\|_{\infty}\|\pi^{\mu}_{N-} - \pi^{\nu}_{N-}\|_{TV}$$

and by exponential stability we have

$$E^{\mu}[\|\pi^{\mu}_{N-} - \pi^{\nu}_{N-}\|_{TV}] \leq \alpha^N \|\mu - \nu\|_{TV} \leq 2\alpha^N$$

□

This bound is loose and could be replaced with the tighter bound

$$2\|\mu - \nu\|_{TV}\|c\|_{\infty}\alpha^N$$

However, the benefit of the bounds presented in the theorem is that it is independent of the priors μ and ν. Note that $\|\mu - \nu\|_{TV} \leq 2$, and the idea of the problems we study is that the DM has the wrong prior ν and does not know the true prior μ. Therefore, the DM would not know the total variation distance between the false prior and the true prior, so we think it is best to provide an upper bound the covers the worst case and therefore applies regardless of the true and false priors.

We can also consider a trained infinite horizon discounted cost problem. Here the DM again makes measurements $Y_{[0,N]}$ about the development of a POMP without incurring any cost or affecting the development of the process. However, at turn N instead of acting in a one shot control problem, the DM starts an infinite horizon discounted cost problem from time N onwards inuring costs and effecting the

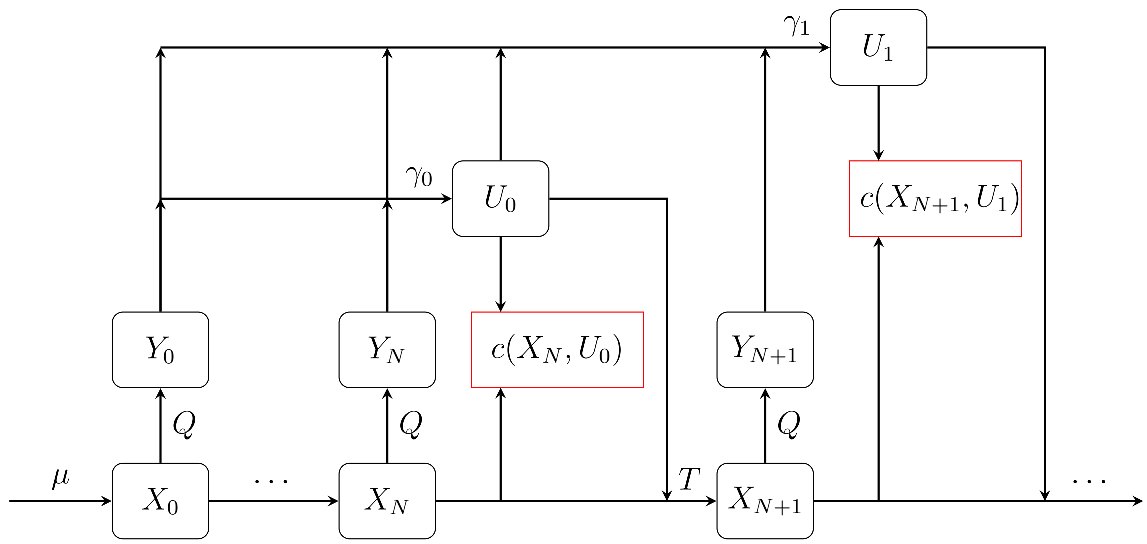

Figure 5.2: Trained Infinite Horizon Discounted Cost Problem

development of the process. The DM wants to minimize

$$E^{\mu}[E^{\mu,\gamma}[\sum_{i=0}^{\infty}\beta^i c(X_{N+i}, U_i)|Y_{[0,N-1]}]]$$

A diagram of the problem is seen in Figure 5.2.

Corollary 5.1. *Assume the cost function c is bounded, non-negative, and measurable and assume the filter is exponentially stable in total variation in expectation with coefficient α. Then we have*

$$E^{\mu}[E^{\mu,\gamma^{\nu}}[\sum_{i=0}^{\infty}\beta^i c(X_{N+i}, U_i)|Y_{[0,N-1]}]] - E^{\mu}[E^{\mu,\gamma^{\mu}}[\sum_{i=0}^{\infty}\beta^i c(X_{N+i}, U_i)|Y_{[0,N-1]}]] \leq 4\frac{\|c\|_{\infty}}{1-\beta}\alpha^N$$

Proof. The proof is the same as Theorem 5.3 except that instead of Proposition 5.1 we apply Theorem 5.1 and have a factor of $\frac{1}{1-\beta}$ in our bound. $\square$

5.4 Cost Free Learning Control Problems

Now we turn our attention to a slightly different control problem, a *cost free learning control problem*. In a trained control problem, the DM simply observes for the first N time stages and does not effect the development of the process. In a cost free learning problem, the DM does not incur costs for the first N stages, however they must still make control actions at time $[0, N-1]$ that effect the development of the process and therefore incur a strategic measure cost in addition to an approximation cost in the robustness problem.

Theorem 5.4. *Assume the cost function c is bounded, non-negative, and measurable and assume the filter is universally exponentially stable in total variation in expectation with coefficient α. Let*

$$|J| = \sup_{\mu,\nu\in\mathcal{P}(\mathcal{X})} J^*(\mu) - J^*(\nu)$$

$$|J_\beta| = \sup_{\mu,\nu\in\mathcal{P}(\mathcal{X})} J^*_\beta(\mu) - J^*_\beta(\nu)$$

then we have

$$E^{\mu,\gamma^\nu}[c(X_N,U_N)] - E^{\mu,\gamma^\mu}[c(X_N,U_N)] \leq 4\|c\|_\infty \alpha^N + |J|$$

$$E^{\mu,\gamma^\nu}[\sum_{i=0}^{\infty}\beta^i c(X_{N+i},U_{N+i})] - E^{\mu,\gamma^\nu}[\sum_{i=0}^{\infty}\beta^i c(X_{N+i},U_{N+i})] \leq 4\frac{\|c\|_\infty}{1-\beta}\alpha^N + |J_\beta|$$

Proof. First for the single stage problem.

$$\begin{aligned}
&E^{\mu,\gamma^\nu}[c(X_N,U_N)] - E^{\mu,\gamma^\mu}[c(X_N,U_N)] \\
=&E^{\mu,\gamma^\nu}\left[E^{\mu,\gamma^\nu}[c(X_N,\gamma^\nu(Y_{[0,N]}))|Y_{[0,N-1]}]\right] - E^{\mu,\gamma^\mu}\left[E^{\mu,\gamma^\mu}[c(X_N,\gamma^\mu(Y_{[0,N]}))|Y_{[0,N-1]}]\right]
\end{aligned}$$

$$=E^{\mu,\gamma^\nu}[J(\pi_{N-}^{\mu,\gamma^\nu},\gamma^{\pi_{n-}^{\nu,\gamma^\nu}})]-E^{\mu,\gamma^\mu}[J^*(\pi_{N-}^{\mu,\gamma^\mu})]$$

$$=E^{\mu,\gamma^\nu}[J(\pi_{N-}^{\mu,\gamma^\nu},\gamma^{\pi_{n-}^{\nu,\gamma^\nu}})\pm J^*(\pi_{N-}^{\mu,\gamma^\nu})]-E^{\mu,\gamma^\mu}[J^*(\pi_{N-}^{\mu,\gamma^\mu})]$$

$$=E^{\mu,\gamma^\nu}[J^*(\pi_{N-}^{\mu,\gamma^\nu})]-E^{\mu,\gamma^\mu}[J^*(\pi_{N-}^{\mu,\gamma^\mu})] \tag{5.11}$$

$$+E^{\mu,\gamma^\nu}[J(\pi_{N-}^{\mu,\gamma^\nu},\gamma^{\pi_{n-}^{\nu,\gamma^\nu}})-J^*(\pi_{N-}^{\mu,\gamma^\nu})] \tag{5.12}$$

Term (5.11) is then upper bound by $|J|$ while term (5.12) is exactly the expression in Theorem 5.3 and the result follows.

For the discounted case, the result is the same except for the approximation cost we use Corollary 5.1. □

5.5 On-Line Learning Control Problem

When the DM has the wrong prior, but must start controlling and incurring costs at the very first time stage we have an *on-line learning control problem.* Consider an infinite horizon discounted cost problem. If we assume a non-negative, bounded, and measurable cost function we have $J^*_\beta(\mu)\geq 0$ for every μ and $J_\beta(\mu,\gamma)\leq\frac{\|c\|_\infty}{1-\beta}$ for every μ,γ. Therefore we have a trivial robustness bound

$$J_\beta(\mu,\gamma^\nu)-J^*_\beta(\mu)\leq\frac{\|c\|_\infty}{1-\beta}$$

if we are to get a meaningful bound it must be tighter than this.

Theorem 5.5. *Assume the cost function c is bounded, non-negative, and measurable* *and assume the filter is universally exponentially stable in total variation in expectation with coefficient* α. *Let*

$$|J_\beta| = \sup_{\mu,\nu\in\mathcal{P}(\mathcal{X})} J^*_\beta(\mu) - J^*_\beta(\nu)$$

$$\rho = \left(\frac{\|c\|_\infty}{1-\beta} - |J_\beta|\right)\left(\frac{\|c\|_\infty}{1-\beta}\right)^{-1} \tag{5.13}$$

$$f(n) = \beta^n(\rho - 4\alpha^n) \tag{5.14}$$

$$n^* = \frac{\ln\left(\left(\frac{\rho}{4}\right)\left(\frac{\ln(\beta)}{\ln(\alpha)+\ln(\beta)}\right)\right)}{\ln\alpha} \tag{5.15}$$

if $\rho > 0, \alpha > 0, 1 > \beta > 0$ *then for any priors* $\mu \ll \nu$ *we have*

$$J_\beta(\mu,\gamma^\nu) - J^*_\beta(\mu) \leq \frac{\|c\|_\infty}{1-\beta}\left(1 - \max(f(\lfloor n^*\rfloor), f(\lceil n^*\rceil)\right)$$

Proof. Pick any $n \in \mathbb{N}$. Starting from expressions (5.8), (5.9), and (5.10) we will consider the three costs. The past mistakes cost is upper bound by

$$E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right] - E^{\mu,\gamma^\mu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right] \leq \|c\|_\infty \sum_{i=0}^{n-1}\beta^i = \|c\|_\infty\left(\frac{1-\beta^n}{1-\beta}\right)$$

the strategic measure cost is upper bound by

$$\beta^n\left(E^{\mu,\gamma^\nu}\left[J^*_\beta(\pi^{\mu,\gamma^\nu}_{n-})\right] - E^{\mu,\gamma^\mu}\left[J^*_\beta(\pi^{\mu,\gamma^\mu}_{n-})\right]\right) \leq \beta^n|J_\beta|$$

and the approximation cost from Corollary 5.1

$$\beta^n\left(E^{\mu,\gamma^\nu}\left[J_\beta(\pi^{\mu,\gamma^\nu}_{n-}, \gamma^{\pi^{\nu,\gamma^\nu}_{n-}}) - J^*_\beta(\pi^{\mu,\gamma^\nu}_{n-})\right]\right) \leq 4\frac{\|c\|_\infty}{1-\beta}(\alpha\beta)^n$$

putting these together

$$\begin{aligned}
J_\beta(\mu,\gamma^\nu) - J^*_\beta(\mu) &\le \|c\|_\infty \left(\frac{1-\beta^n}{1-\beta}\right) + \beta^n |J_\beta| + 4\frac{\|c\|_\infty}{1-\beta}(\alpha\beta)^n \\
&= \frac{\|c\|_\infty}{1-\beta} + \beta^n \left(|J_\beta| + 4\frac{\|c\|_\infty}{1-\beta}\alpha^n - \frac{\|c\|_\infty}{1-\beta}\right) \\
&= \frac{\|c\|_\infty}{1-\beta}\left(1 + \beta^n(4\alpha^n - \rho))\right) \\
&= \frac{\|c\|_\infty}{1-\beta}(1 - f(n))
\end{aligned}$$

This of course holds for any n. If $\rho > 0$ and $\alpha < 1$ then there will exist n such that $f(n) > 0$ and hence we will get an improvement over the trivial bound. Taking the derivative in n

$$\frac{d}{dn} f(n) = \beta^n \left(\rho \ln(\beta) - 4\ln(\alpha\beta)\alpha^n\right)$$

then there exists n^* where

$$\rho \ln(\beta) - 4\ln(\alpha\beta)\alpha^{n^*} = 0$$

for all n less than this value, the derivative is positive and for all n greater than this value, the derivative is negative hence this value is the global maximizer.

$$\begin{aligned}
\alpha^{n^*} &= \left(\frac{\rho}{4}\right)\left(\frac{\ln(\beta)}{\ln(\alpha)+\ln(\beta)}\right) \\
n^* &= \frac{\ln\left(\left(\frac{\rho}{4}\right)\left(\frac{\ln(\beta)}{\ln(\alpha)+\ln(\beta)}\right)\right)}{\ln\alpha}
\end{aligned}$$

The maximum among the natural numbers $n \in \mathbb{N}$ will occur at the ceiling or floor of

n^*. □

The value of ρ is the percent improvement of $|J_\beta|$ over the trivial bound, and the maximum of $f(n)$ is the percent improvement of the robustness bound over the trivial bound. The higher the maximum, the better the percent improvement.

The maximum of f is larger when β is close to 1, as this puts more weight on later values and allows for more learning. Also when α is close to 0 as this increases the rate of filer merging, and when ρ is close to 1 as this means $|J_\beta|$ is small relative to the trivial bound. A reasonable set of values is $\beta = 0.9, \alpha = 0.8, \rho = 0.25$. This will result in a percent of improvement of 2.68% over the trivial bound. If we use more contrived numbers such as $\beta = 0.95, \alpha = 0.7, \rho = 1$ then we have an improvement of 53% over the trivial bound.

This result may be considered a "prior independent" bound since the bound on the robustness distance does not depend on the actual priors μ and ν being considered. This is useful when we have no real knowledge of how "close" (in total variation distance) μ and ν are, thus if they are far apart the bound will not change. As previously discussed in the introduction to this chapter, [26, Proposition 3.2] provides a prior dependent bound for the discounted cost problem

$$J_\beta(\mu, \gamma^\nu) - J_\beta^*(\mu) \leq 2\frac{\|c\|_\infty}{1-\beta}\|\mu - \nu\|_{TV}$$

When $\|\mu - \nu\|_{TV}$ is small, i.e. our priors start close together to begin with, this bound is actually better than our prior independent bound. However, when they are far apart it may be that

$$2\|\mu - \nu\|_{TV} > 1 - \max(f(\lfloor n^* \rfloor), f(\lceil n^* \rceil))$$

and our prior independent bound is better. However, we can also use filter stability to derive a new prior dependent bound which improves upon [26]

Theorem 5.6. *Assume the cost function c is bounded, non-negative, and measurable and assume the filter is universally exponentially stable in total variation in expectation with coefficient α. Fix two priors μ and ν with $\mu \ll \nu$. Let*

$$\sigma = (|J_\beta|)\left(\frac{\|c\|}{1-\beta}\right)^{-1}$$
$$\upsilon = 1 - \frac{\sigma}{2\|\mu-\nu\|_{TV}}$$
$$f = \beta^n(\upsilon - \alpha^n)$$
$$n^* = \frac{\ln\left(\upsilon \frac{\ln(\beta}{\ln(\alpha)+\ln(\beta)}\right)}{\ln(\alpha)}$$

If $\upsilon > 0, \alpha > 0, 1 > \beta > 0$ then we have

$$J_\beta(\mu,\gamma^\nu) - J^*_\beta(\mu) \leq 2\frac{\|c\|_\infty}{1-\beta}\|\mu-\nu\|_{TV}\left(1 - \max(f(\lfloor n^*\rfloor), f(\lceil n^*\rceil)\right)$$

Proof. As before, we will consider the past mistakes cost, strategic measure cost, and the approximation cost separately to come to our conclusion. However, this time we will make use of the actual priors being considered, and not utilize prior independent bounds.

First the past mistakes cost. For notational convenience let

$$J_N(\mu,\gamma) = E^{\mu,\gamma}\left[\sum_{i=0}^{N-1}\beta^i c(X_i,U_i)\right]$$

we then have

$$\begin{aligned}
&E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right]-E^{\mu,\gamma^\mu}\left[\sum_{i=0}^{n-1}\beta^i c(x_i,u_i)\right]\\
=&J_N(\mu,\gamma^\nu)-J_N(\mu,\gamma^\mu)+J_N(\nu,\gamma^\nu)-J_N(\nu,\gamma^\nu)\\
\leq&|J_N(\nu,\gamma^\nu)-J_N(\mu,\gamma^\mu)|+J_N(\mu,\gamma^\nu)-J_N(\nu,\gamma^\nu)
\end{aligned}$$

This follows the same structure as the proof of Theorem 5.2. We can show that

$$|J_N(\nu,\gamma^\nu)-J_N(\mu,\gamma^\mu)|\leq\max\left(J_N(\mu,\gamma^\nu)-J_N(\nu,\gamma^\nu),J_N(\nu,\gamma^\mu)-J_N(\mu,\gamma^\mu)\right)$$

we then have

$$J_N(\mu,\gamma^\nu)-J_N(\nu,\gamma^\nu)\leq\|c\|_\infty\sum_{i=0}^{n-1}\beta^i\|P^{\mu,\gamma^\nu}(X_i,U_i\in\cdot)-P^{\nu,\gamma^\nu}(X_i,U_i\in\cdot)\|_{TV}$$

since the two POMDPs are employing the same policy, we have

$$\|P^{\mu,\gamma^\nu}(X_i,U_i\in\cdot)-P^{\nu,\gamma^\nu}(X_i,U_i\in\cdot)\|_{TV}\leq\|\mu-\nu\|_{TV}\ \forall i\in\mathbb{N}$$

and therefore the past mistakes cost is upper bound by

$$J_N(\mu,\gamma^\nu)-J_N(\mu,\gamma^\mu)\leq 2\|c\|_\infty\left(\frac{1-\beta^n}{1-\beta}\right)\|\mu-\nu\|_{TV}$$

the strategic measure cost is upper bound by

$$\beta^n\left(E^{\mu,\gamma^\nu}\left[J^*_\beta(\pi_{n-}^{\mu,\gamma^\nu})\right]-E^{\mu,\gamma^\mu}\left[J^*_\beta(\pi_{n-}^{\mu,\gamma^\mu})\right]\right)\leq\beta^n|J_\beta|$$

and the approximation cost from a modified version of Corollary 5.1

$$\beta^n \left(E^{\mu,\gamma^\nu} \left[J_\beta(\pi_{n-}^{\mu,\gamma^\nu}, \gamma^{\pi_{n-}^{\nu,\gamma^\nu}}) - J_\beta^*(\pi_{n-}^{\mu,\gamma^\nu}) \right] \right) \leq 2\frac{\|c\|_\infty}{1-\beta}(\alpha\beta)^n \|\mu - \nu\|_{TV}$$

putting these together

$$\begin{aligned}
J_\beta^*(\mu,\gamma^\nu) - J_\infty^*(\mu) &\leq 2\|c\|_\infty \frac{1-\beta^n}{1-\beta}\|\mu-\nu\|_{TV} + \beta^n |J_\beta| + 2\frac{\|c\|_\infty}{1-\beta}(\alpha\beta)^n \|\mu-\nu\|_{TV} \\
&= \frac{2\|c\|_\infty}{1-\beta}\|\mu-\nu\|_{TV}\left(1-\beta^n + \beta^n|J_\beta| \left(\frac{2\|c\|_\infty\|\mu-\nu\|_{TV}}{1-\beta}\right)^{-1} + (\alpha\beta)^n\right) \\
&= \frac{2\|c\|_\infty}{1-\beta}\|\mu-\nu\|_{TV}\left(1-\beta^n + \frac{\beta^n}{2\|\mu-\nu\|_{TV}}|J_\beta| \left(\frac{\|c\|_\infty}{1-\beta}\right)^{-1} + (\alpha\beta)^n\right) \\
&= \frac{2\|c\|_\infty}{1-\beta}\|\mu-\nu\|_{TV}\left(1-\beta^n\left(1-\frac{\sigma}{2\|\mu-\nu\|_{TV}} - \alpha^n\right)\right) \\
&= \frac{2\|c\|_\infty}{1-\beta}\|\mu-\nu\|_{TV}\left(1-\beta^n\left(\upsilon - \alpha^n\right)\right) \\
&= \frac{2\|c\|_\infty}{1-\beta}\|\mu-\nu\|_{TV}\left(1-f(n)\right)
\end{aligned}$$

The first term in this exression is the previous bound derived in [26]. If we have $\sigma \leq 2\|\mu-\nu\|_{TV}$ then we will have $\upsilon > 0$ and there will exist n where $f(n) > 0$ and we will get an improvement over this bound. Taking the derivative in n

$$\frac{d}{dn}f(n) = \frac{d}{dn}\left(\beta^n(\upsilon) - \beta^n\alpha^n\right) = \beta^n(\ln(\beta)(\upsilon) - \ln(\alpha\beta)(\alpha)^n)$$

then there will exist n^* where

$$\ln(\beta)(\upsilon) - \ln(\alpha\beta)(\alpha)^n = 0$$

for all n less than this value the derivative is positive, and for all n greater than this value the derivative is negative hence this value is the global maximizer.

$$\alpha^{n^*} = \frac{(\upsilon)\ln(\beta)}{\ln(\alpha)+\ln(\beta)}$$
$$n^* = \frac{\ln\left(\frac{(\upsilon)\ln(\beta)}{\ln(\alpha)+\ln(\beta)}\right)}{\ln(\alpha)}$$

□

Note that σ is the ratio of the maximum distance of the optimal cost operator $|J_\beta|$ and the trivial bound $\frac{\|c\|_\infty}{1-\beta}$. If this ratio is less than double the original total variation distance $2\|\mu-\nu\|_{TV}$ (note total variation distance is between 0 and 2) then our Theorem 5.6 results in an improved bound over that in [26]. However, this policy dependent bound may be worse than our policy independent bound in Theorem 5.5.

5.6 Average Cost Problem

We now turn our attention to the infinite horizon average cost problem. Consider first a trained control problem

Theorem 5.7. *Assume the cost function c is bounded, non-negative, and measurable and assume the filter is exponentially stable in total variation in expectation with coefficient α. Then*

$$E^{\mu}\left[\limsup_{T\to\infty} E^{\mu,\gamma^{\nu}}\left[\frac{1}{T}\sum_{i=0}^{T-1} c(X_{N+i},U_{[0,N+i]}|Y_{[0,N-1]}\right]\right.$$
$$\left. - \limsup_{T\to\infty} E^{\mu,\gamma^{\mu}}\left[\frac{1}{T}\sum_{i=0}^{T-1} c(X_{N+i},U_{[0,N+i]})|Y_{[0,N-1]}\right]\right] \leq 4\|c\|_\infty \alpha^N$$

Proof. The distribution on X_N given $Y_{[0,N-1]}$ is the true predictor π^{μ}_{N-} which acts as a new prior for a control policy. Since the optimal policy γ^{ν} is a stationary function of the false filter realization and γ^{μ} is a stationary function of the true filter we have

$$E^{\mu}\left[\limsup_{T\to\infty} E^{\mu,\gamma^{\nu}}\left[\frac{1}{T}\sum_{i=0}^{T-1} c(X_{N+i}, U_{[0,N+i]})|Y_{[0,N-1]}\right]\right.$$
$$\left. - \limsup_{T\to\infty} E^{\mu,\gamma^{\mu}}\left[\frac{1}{T}\sum_{i=0}^{T-1} c(X_{N+i}, U_{[0,N+i]})|Y_{[0,N-1]}\right]\right]$$
$$\leq E^{\mu}[J_{\infty}(\pi^{\mu}_{N-}, \gamma^{\pi^{\nu}_{N-}}) - J^{*}_{\infty}(\pi^{\mu}_{N-})]$$

by Theorem 5.2 we have

$$J_{\infty}(\pi^{\mu}_{N-}, \gamma^{\pi^{\nu}_{n-}}) - J^{*}_{\infty}(\pi^{\mu}_{N-}) \leq 2\|c\|_{\infty}\|\pi^{\mu}_{N-} - \pi^{\nu}_{N-}\|_{TV}$$

and by exponential filter (equivalently predictor) stability

$$E^{\mu}[\|\pi^{\mu}_{N-} - \pi^{\nu}_{N-}\|_{TV}] \leq \alpha^{N}\|\mu - \nu\| \leq 2\alpha^{N}$$

□

Now consider the online learning problem.

Theorem 5.8. *Assume the cost function c is bounded, non-negative, and measurable and assume the filter (equivalently the predictor) is universally stable in total variation in expectation. Let*

$$|J_{\infty}| = \sup_{\mu_1\in\mathcal{P}(\mathcal{X})} J^{*}_{\infty}(\mu_1) - \inf_{\mu_2\in\mathcal{P}(\mathcal{X})} J^{*}_{\infty}(\mu_2)$$

then we have

$$J_\infty(\mu, \gamma^\nu) - J^*_\infty(\mu) \leq |J_\infty|$$

Proof. Fix some finite n, we have

$$\begin{aligned}
J_\infty(\mu, \gamma^\nu) &= \limsup_{T\to\infty} \frac{1}{T}\left(\sum_{i=0}^{n-1} E^{\mu,\gamma^\nu}[c(X_i, U_i)] + \sum_{i=n}^{T-1} E^{\mu,\gamma^\nu}[c(X_i, U_i)]\right) \\
&\leq \limsup_{T\to\infty} \frac{1}{T}\sum_{i=0}^{n-1} E^{\mu,\gamma^\nu}[c(X_i, U_i)] + \limsup_{T\to\infty} \frac{1}{T} E^{\mu,\gamma^\nu}\left[\sum_{i=n}^{T-1} c(X_i, U_i)\right] \\
&\leq \limsup_{T\to\infty} \frac{n\|c\|_\infty}{T} + \limsup_{T\to\infty} \frac{1}{T} E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{T-n-1} c(X_{n+i}, U_{n+i})\right] \\
&= \limsup_{T\to\infty} \frac{1}{T} E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{T-n-1} c(X_{n+i}, U_{n+i})\right]
\end{aligned}$$

therefore, we see that no matter what decision the DM makes in the first n time stages, since n is finite and c bounded this cost will eventually be dominated by the denominator as $T \to \infty$ and there will be no past mistakes cost associated with this robustness problem. We then claim that

$$\limsup_{T\to\infty} \frac{1}{T} E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{T-n-1} c(X_{n+i}, U_{n+i})\right] = \limsup_{T\to\infty} \frac{1}{T-n} E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{T-n-1} c(X_{n+i}, U_{n+i})\right] \tag{5.16}$$

All terms in the two limsup expressions are positive and bounded since c is a non-negative bounded function, therefore we have that

- The difference of the limsups is less than or equal to the limsup of the difference.
- The limsup of a product is less than or equal to the product of the limsups.

Using these results,

$$
\begin{aligned}
&\limsup_{T\to\infty} \frac{1}{T-n} E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{T-n-1} c(X_{n+i},U_{n+i})\right] - \limsup_{T\to\infty} \frac{1}{T} E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{T-n-1} c(X_{n+i},U_{n+i})\right] \\
\leq &\limsup_{T\to\infty} \left(\frac{1}{T-n} - \frac{1}{T}\right) E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{T-n-1} c(X_{n+i},U_{n+i})\right] \\
= &\limsup_{T\to\infty} \left(\frac{n}{T(T-n)}\right) E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{T-n-1} c(X_{n+i},U_{n+i})\right] \\
\leq &\left(\limsup_{T\to\infty} \frac{n}{T}\right)\left(\limsup_{T\to\infty} \frac{1}{T-n} E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{T-n-1} c(X_{n+i},U_{n+i})\right]\right) \\
&= 0
\end{aligned}
$$

We then apply iterated expectations and Fatou's lemma and we have

$$
\begin{aligned}
&\limsup_{T\to\infty} \frac{1}{T-n} E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{T-n-1} c(X_{n+i},U_{n+i})\right] \\
= &\limsup_{T\to\infty} \frac{1}{T-n} E^{\mu,\gamma^\nu}\left[E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{T-n-1} c(X_{n+i},U_{n+i})|Y_{[0,n-1]}\right]\right] \\
\leq &E^{\mu,\gamma^\nu}\left[\limsup_{T\to\infty} \frac{1}{T-n} E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{T-n-1} c(X_{n+i},U_{n+i})|Y_{[0,n-1]}\right]\right]
\end{aligned}
$$

As was shown for the discounted problem, the optimal control policy is a time invariant function of the filter realization so the predictor at time n acts as a new prior for the control problem. We then have

$$
E^{\mu,\gamma^\nu}\left[\limsup_{T\to\infty} \frac{1}{T-n} E^{\mu,\gamma^\nu}\left[\sum_{i=0}^{T-n-1} c(X_{n+i},U_{n+i})|Y_{[0,n-1]}\right]\right] = E^{\mu,\gamma^\nu}\left[J_\infty(\pi_{n-}^{\mu,\gamma^\nu}, \gamma^{\pi_{n-}^{\nu,\gamma^\nu}})\right]
$$

With this established, we can now move on to our robustness problem.

$$\begin{aligned}
J_\infty(\mu,\gamma^\nu) - J^*_\infty(\mu) &\leq E^{\mu,\gamma^\nu}[J_\infty(\pi^{\mu,\gamma^\nu}_{n-},\gamma^{\pi^{\nu,\gamma^\nu}_{n-}})] - \inf_{\tilde{\mu}\in\mathcal{P}(\mathcal{X})} J^*_\infty(\tilde{\mu}) \\
&= E^{\mu,\gamma^\nu}[J_\infty(\pi^{\mu,\gamma^\nu}_{n-},\gamma^{\pi^{\nu,\gamma^\nu}_{n-}}) + J^*_\infty(\pi^{\mu,\gamma^\nu}_{n-}) - J^*_\infty(\pi^{\mu,\gamma^\nu}_{n-})] - \inf_{\tilde{\mu}\in\mathcal{P}(\mathcal{X})} J^*_\infty(\tilde{\mu}) \\
&= E^{\mu,\gamma^\nu}[J_\infty(\pi^{\mu,\gamma^\mu}_{n-},\gamma^{\pi^{\nu,\gamma^\nu}_{n-}}) - J^*_\infty(\pi^{\mu,\gamma^\nu}_{n-})] + E^{\mu,\gamma^\nu}[J^*_\infty(\pi^{\mu,\gamma^\nu}_{n-})] - \inf_{\tilde{\mu}\in\mathcal{P}(\mathcal{X})} J^*_\infty(\tilde{\mu}) \\
&\leq E^{\mu,\gamma^\nu}[J_\infty(\pi^{\mu,\gamma^\nu}_{n-},\gamma^{\pi^{\nu,\gamma^\nu}_{n-}}) - J^*_\infty(\pi^{\mu,\gamma^\nu}_{n-})] + \sup_{\tilde{\mu}\in\mathcal{P}(\mathcal{X})} J^*_\infty(\tilde{\mu}) - \inf_{\tilde{\mu}\in\mathcal{P}(\mathcal{X})} J^*_\infty(\tilde{\mu}) \\
&= E^{\mu,\gamma^\nu}[J_\infty(\pi^{\mu,\gamma^\nu}_{n-},\gamma^{\pi^{\nu,\gamma^\nu}_{n-}}) - J^*_\infty(\pi^{\mu,\gamma^\nu}_{n-})] + |J_\infty|
\end{aligned}$$

Then by Theorem 5.2 we have

$$J_\infty(\mu,\gamma^\nu) - J^*_\infty(\mu) \leq 2\|c\|_\infty E^{\mu,\gamma^\nu}[\|\pi^{\mu,\gamma^\nu}_{n-} - \pi^{\nu,\gamma^\nu}_{n-}\|_{TV}] + |J_\infty| \tag{5.17}$$

and this result holds for any n since our choice of n was arbitrary. By assumption, the predictor is universally stable in total variation in expectation (equivalently the filter, see Corollary (3.1)). Therefore, pick any $\epsilon > 0$. There exists an N such that for all $n > N$ we have

$$E^{\mu,\gamma^\nu}[\|\pi^{\mu,\gamma^\nu}_{n-} - \pi^{\nu,\gamma^\nu}_{n-}\|_{TV}] \leq \frac{\epsilon}{2\|c\|_\infty}$$

and therefore since (5.17) holds for every n

$$J_\infty(\mu,\gamma^\nu) - J^*_\infty(\mu) \leq |J_\infty| + \epsilon$$

for any $\epsilon > 0$ yielding our result. □

Note that this result is significantly improved over the discounted cost problem. The discounted problem puts more weight on earlier costs than later costs, while the average cost problems weighs them all equally. The first consequence of this is the past mistakes cost is 0 in the average cost problem. Secondly, the rate of the filter merging does not matter. We do not need exponential stability, but only general total variation stability in expectation. Third, the approximation cost is 0 and the robustness difference is only the strategic measure cost.

5.7 Strategic Measure Cost

A large gap in these results is the strategic measure cost $|J_\beta|$ or $|J_\infty|$ as these make the robustness bounds significantly larger for the on-line learning problems. If $|J_\beta| = 0$ then the discounted cost results would be improved as $\rho = 1$ would be maximized, but this is not as drastic an improvement as the average cost problem. If it is the case that $|J_\infty| = 0$, then the robustness bound becomes zero for the average cost problem. That is, the DM suffers *no penalty* for using the incorrectly designed control policy.

In order for $|J_\infty| = 0$, we require the optimal cost operator $J_\infty^*(\mu)$ to be the same for all priors μ. No matter the starting prior, under optimal control the DM incurs the same cost. This property has less to do with filter stability, and more with the ergodic nature of the filter as a stochastic process.

Filter stability guarantees that the true filter and the false filter will merge as $n \to \infty$, however it does not guarantee what limit distribution they will merge towards. However, if the filter as a stochastic process admits a unique invariant measure on the space of probability measures, then any initialized filter will merge to this invariant measure. This implies filter stability, but is indeed stronger as we know the limit distribution beforehand.

This result remains future work as the uniqueness of an invariant measure (i.e. unique ergodicity) of the filter process is still an open problem. With filter stability alone, the DM can "wake up" at a certain point in the process and start making optimal control decisions given the environment the DM finds itself in. Yet if the past control decisions up to that point have placed a poor distribution on the predictor realization $\pi_{n-}^{\mu,\gamma^\nu}$, without guarantees on the difference of the optimal cost operator $|J_\infty|$ the DM may not be able to overcome its past mistakes, despite being optimal for all future time stages.

Additionally, the bounds for the strategic measure difference could be made tighter, but in a policy dependent sense that cannot be computed directly with current knowledge in the literature. When we consider the strategic measure difference

$$E^{\mu,\gamma^\nu}[J_\infty^*(\pi_{n-}^{\mu,\gamma^\mu})] - J_\infty^*(\mu) \tag{5.18}$$

we choose to upper bound this difference with the term $|J_\infty|$ which is independent of μ and easily computable. However, if we define the reachable set of a prior μ as all filter realizations that are achievable under a given prior

$$\mathcal{R}(\mu) = \{\rho \in \mathcal{P}(\mathcal{X}) | \exists n \in \mathbb{N}, \exists y_{[0,n]} \in \mathcal{Y}^{n+1} \ s.t. \ P^{\mu,\gamma}(X_{n+1} \in \cdot | Y_{[0,n]} = y_{[0,n]}) = \rho\}$$

then we can present the tighter bound of

$$J_\infty(\mu, \gamma^\nu) - J_\infty^*(\mu) \leq \sup_{\rho \in \mathcal{R}(\mu)} J_\infty^*(\rho) - J_\infty^*(\mu)$$

However, this bound is problematic in that is is dependent on the prior μ and furthermore we do not know how to characterize the reachable set $\mathcal{R}(\mu)$. In future studies of unique ergodicity of the filter process, one should also study reachability of the filter process as well to get a better picture of this tighter bound.

5.8 Summary

In this chapter we study the problem of a DM designing an optimal control policy with respect to the wrong prior distribution ν when the true distribution is μ. The cost incurred by the DM is then $J(\mu, \gamma^{\nu})$ when the minimum cost that could be incurred with the correctly designed policy is $J^*(\mu)$. When we investigate robustness, we consider the difference between these two terms

$$J(\mu, \gamma^{\nu}) - J^*(\mu)$$

For the single stage cost problem, the infinite horizon average cost problem, and the infinite horizon average cost problem, when optimal solutions are known to exist the filter realization is a sufficient statistic for the optimal policy [23],[7]. Thus, filter stability is closely tied to the DM having accurate information about the state and thus selecting optimal control decisions.

We show there are three costs associated with a DM employing a falsely initialized policy: the past mistakes cost, the strategic measure cost, and the approximation cost. The first two cost are analogous to the DM operating under its false assumptions for an initial number of time stages N, and then "waking up" for later time stages and employing the optimal policy from that time on. The poor decisions the DM makes at times $[0, N]$ incur a high cost at these times, but also put an undesirable measure on the future states. It may be possible that even if the DM starts making optimal control decisions from time N onwards, it cannot correct the poor position the DM finds itself in based on its past control actions. The third cost arises since the DM does not actually wake up at time N, but rather filter stability brings the true filter and the false filter together which essentially has the same effect.

When the DM is given time to learn the filter before any control actions are made, the only cost incurred by the filter is the approximation cost which is related to the total variation distance of the true and false filter. In such problems, the DM's learning window is finite and thus exponential stability is required over asymptotic stability which may not merge fast enough.

If the DM must start controlling and incurring cost from the start of the problem with the incorrect prior, the DM faces an on-line learning problem. For the infinite horizon discounted cost problem, the later costs are weighed less than the current costs. Therefore, the rate of the filter merging is significant and again exponential stability is necessary. An upper bound on the robustness difference is provided in Theorem 5.5 which depends on exponential rate of filter stability α, the discount factor β of the problem, and the maximum distance of optimal cost under different priors $|J_\beta|$.

For the average cost problem, since all costs are weighed equally the speed of filter stability is not as important and asymptotic stability, the result appearing in Chapter 3, is sufficient. Additionally, the robustness cost is entirely determined by the maximal distance of the optimal cost operator under different priors, $|J_\infty|$. An avenue of future work studies this maximal distance, which is closely tied to the ergodic nature of the filter as a stochastic process.

Chapter 6

Conclusions

6.1 Summary

In this thesis we investigate the stability of the non-linear filter in controlled and control-free environments. As an application of filter stability, we study the robustness problem of a DM employing a control policy designed with respect to an incorrect prior. We show how filter stability can bound the penalty the DM incurs for this mistake.

In Chapter 2, we study filter stability in control-free systems through the lens of observability. Stability arises via the informative nature of the measurements in relation to the hidden state process. With successive measurements, the observer gains a better picture of the state process and corrects the initialization error asymptotically. We compare different notions of stability under weak stability, total variation, and relative entropy.

In Chapter 3, we study filter stability in controlled environments. Stability again follows from observability, but the affect of the control actions on the conditional distributions adds complications. In general, the one stage observation channel $Y_n|X_n$

is unaffected by control, thus one step observability still applies. However, the multi-stage channel $Y_{[n,n+N]}|X_n$ is influenced by past control actions, and hence N step observability only applies when the control actions do not influence the transition kernel. Aside form this distinction, results are identical to Chapter 2.

In Chapter 4, we study filter stability not through observability, but through the ergodic properties of the filter update kernel. We show that the transition kernel T is a contraction with a bounded determined by its Dobrushin coefficient, and the Bayesian update ψ has an upper expansion bound based on the Dobrushin coefficient of the measurement kernel Q. If the product of contraction coefficient of T and bound on the expansion coefficient of ψ is less than one, their composed operator is a contraction in total variation in expectation and thus the filter merges at an exponential rate.

In Chapter 5, we apply the filter stability results to the problem of a controller implementing a control policy designed with respect to the wrong prior. For problems with a finite time window for learning or the infinite horizon discounted cost problem, exponential filter stability leads to bounds on the robustness difference. In the average cost problem, asymptotic stability is sufficient to provide a bound that only depends of the maximal difference of the optimal cost operator under different priors. If this can be shown to be 0, then the filter incurs no cost for using the incorrectly designed policy.

6.2 Future Work

For the filter stability problem, there are a number of directions which could be studied. Results could be extended from discrete stochastic processes as studied here to continuous time stochastic processes by reformulating our observability definition for stochastic differential equations. Additionally, the filter as utilized here is not finitely computable in practice. Many applications use approximate filters such as the particle filter or quantization to sacrifice accuracy in the filter estimate for simpler computation. It would be interesting to study filter stability in conjunction with these approximate filters.

For robustness, the gap remains of bounding the maximal distance of the optimal cost operator under different priors $|J_\infty|$. If the filter update kernel is shown to admit an unique invariant distribution, then all optimal costs under different priors are the same and $|J_\infty| = 0$. Future work could investigate conditions for the existence of a unique invariant measure for the filter update kernel in POMDP, which is currently an open problem.

Bibliography

[1] Y. Bar-Shalom and E. Tse. Dual effect certainty equivalence and separation in stochastic control. *IEEE Transactions on Automatic Control*, 19:494–500, October 1974.

[2] A.R. Barron. Limits of information, Markov chains, and projections. Sorrento, Italy, 1999. Presented at the 2000 IEEE Int. Symp. on Inform. Theory. 8 page original submission.

[3] T. Başar and P. Bernhard. *H-infinity optimal control and related minimax design problems: a dynamic game approach*. Birkhäuser, 1995.

[4] P. Baxendale, P. Chigansky, and R. Liptser. Asymptotic stability of the wonham filter: ergodic and nonergodic signals. *SIAM journal on control and optimization*, 43(2):643–669, 2004.

[5] D. Blackwell and L. Dubins. Merging of opinions with increasing information. *The Annals of Mathematical Statistics*, 33(3):882–886, 1962.

[6] V.I. Bogachev. *Measure Theory*, volume 1. Springer Science & Business Media, 2007.

[7] V.S. Borkar. Convex analytic methods in Markov decision processes. In *Handbook of Markov decision processes*, pages 347–375. Springer, 2002.

[8] A. Budhiraja and D. Ocone. Exponential stability in discrete-time filtering for non-ergodic signals. *Stochastic processes and their applications*, 82(2):245–257, 1999.

[9] P. Chigansky and R. Liptser. On a role of predictor in the filtering stability. *Electron. Comm. Probab*, 11:129–140, 2006.

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

[11] P. Chigansky and R. van Handel. A complete solution to Blackwell's unique ergodicity problem for hidden Markov chains. *The Annals of Applied Probability*, 20(6):2318–2345, 2010.

[12] J.M.C. Clark, D.L. Ocone, and C. Coumarbatch. Relative entropy and error bounds for filtering of Markov processes. *Mathematics of Control, Signals and Systems*, 12(4):346–360, 1999.

[13] D. Crisan and K. Heine. Stability of the discrete time filter in terms of the tails of noise distributions. *Journal of the London Mathematical Society*, 78(2):441–458, 2008.

[14] I. Csiszár. Information-type measures of difference of probability distributions and indirect observation. *studia scientiarum Mathematicarum Hungarica*, 2:229–318, 1967.

[15] A. D'Aristotile, P. Diaconis, and D. Freedman. On merging of probabilities. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 363–380, 1988.

[16] R.L. Dobrushin. Central limit theorem for nonstationary Markov chains. i. *Theory of Probability & Its Applications*, 1(1):65–80, 1956.

[17] R. Douc, E. Gassiat, B. Landelle, and E. Moulines. Forgetting of the initial distribution for nonergodic hidden Markov chains. *The Annals of Applied Probability*, 20(5):1638–1662, 2010.

[18] R.M. Dudley. *Real Analysis and Probability*. Cambridge University Press, Cambridge, 2nd edition, 2002.

[19] E.A. Feinbergand, P.O. Kasyanov, and M.Z. Zgurovsky. Partially observable total-cost Markov decision processes with weakly continuous transition probabilities. *Mathematics of Operations Research*, 41(2):656–681, 2016.

[20] F. Le Gland and N. Oudjane. Stability and uniform approximation of nonlinear filters using the Hilbert metric and application to particle filters. *The Annals of Applied Probability*, 14(1):144–187, 2004.

[21] R.M. Gray. *Entropy and information theory*. Springer Science & Business Media, 2011.

[22] O. Hernández-Lerma and J.B. Lasserre. *Markov chains and invariant probabilities*. Birkhäuser-Verlag, Basel, 2003.

[23] O. Hernández-Lerma and J.B. Lasserre. *Discrete-time Markov control processes: basic optimality criteria*, volume 30. Springer Science & Business Media, 2012.

[24] O. Hernández-Lerma and J.B. Lasserre. *Markov chains and invariant probabilities*, volume 211. Birkhäuser, 2012.

[25] A.D. Kara, N. Saldi, and S. Yüksel. Weak Feller property of non-linear filters. *arXiv:1812.05509*, 2018. to appear in Systems and Control Letters.

[26] A.D. Kara and S. Yüksel. Robustness to incorrect priors in partially observed stochastic control. *SIAM Journal on Control and Optimization*, 57(3):1929–1964, 2019.

[27] S. Kullback. A lower bound for discrimination information in terms of variation. *IEEE Transactions on Information Theory*, 13(1):126–127, 1967.

[28] H. Kunita. Asymptotic behavior of the nonlinear filtering errors of Markov processes. *Journal of Multivariate Analysis*, 1(4):365–393, 1971.

[29] H.J. Kushner. *Introduction to Stochastic Control Theory.* Holt, Rinehart, and Winston, New York, 1972.

[30] T. Linder and S. Yüksel. On optimal zero-delay coding of vector Markov sources. *IEEE Trans. Information Theory*, 60(10):5975–5991, 2014.

[31] G. B. Di Masi and L. Stettner. Ergodicity of hidden markov models. *Mathematics of Control, Signals and Systems*, 17(4):269–296, 2005.

[32] D. Ocone and E. Pardoux. Asymptotic stability of the optimal filter with respect to its initial condition. *SIAM Journal on Control and Optimization*, 34(1):226–243, 1996.

[33] D.L. Ocone. Asymptotic stability of beneš filters. *Stochastic analysis and applications*, 17(6):1053–1074, 1999.

[34] M.S. Pinsker. Information and information stability of random variables and processes. 1960.

[35] W. Rudin. *Real and Complex Analysis*. Tata McGraw-Hill Education, 2006.

[36] R. van Handel. Discrete time nonlinear filters with informative observations are stable. *Electronic Communications in Probability*, 13:562–575, 2008.

[37] R. van Handel. Observability and nonlinear filtering. *Probability theory and related fields*, 145(1-2):35–74, 2009.

[38] R. van Handel. The stability of conditional Markov processes and Markov chains in random environments. *The Annals of Probability*, 37(5):1876–1925, 2009.

[39] R. van Handel. Uniform observability of hidden Markov models and filter stability for unstable signals. *The Annals of Applied Probability*, 19(3):1172–1199, 2009.

[40] R. van Handel. Nonlinear filtering and systems theory. In *Proceedings of the 19th International Symposium on Mathematical Theory of Networks and Systems (MTNS semi-plenary paper)*, 2010.

[41] P. Whittle. A risk-sensitive maximum principle: The case of imperfect state observation. *IEEE Transactions on Automatic Control*, 36(7):793–801, 1991.

[42] S. Yüksel. A note on the separation of optimal quantization and control policies in networked control. *SIAM Journal on Control and Optimization*, 57(1):773–782, 2019.