INTRINSIC METHODS IN FILTER STABILITY

P. CHIGANSKY, R. LIPTSER, AND R. VAN HANDEL

Abstract. The purpose of this article is to survey some *intrinsic* methods for studying the stability of the nonlinear filter. By 'intrinsic' we mean methods which directly exploit the fundamental representation of the filter as a conditional expectation through classical probabilistic techniques such as change of measure, martingale convergence, coupling, etc. Beside their conceptual appeal and the additional insight gained into the filter stability problem, these methods allow one to establish stability of the filter under weaker conditions compared to other methods, e.g., to go beyond strong mixing signals, to reveal connections between filter stability and classical notions of observability, and to discover links to martingale convergence and information theory.

1. Inroduction

Consider a pair of random sequences $(X, Y) = (X_n, Y_n)_{n\in\mathbb{Z}_+}$, where the signal component X_n takes values in a Polish space[1] $\mathbb{S}$ and the observation component Y_n takes values in $\mathbb{R}^p$ for some $p \geq 1$. The classical filtering problem is to compute the conditional distribution

$$\pi_n(\cdot) = \mathbb{P}(X_n \in \cdot\, |\mathcal{F}^Y_{0,n}), \tag{1.1}$$

where $\mathcal{F}^Y_{k,n}$ stands for the σ-algebra of events generated by Y_m, $k \leq m \leq n$ (similarly, we will use below the σ-algebra $\mathcal{F}^X_{k,n}$ generated by X_m, $k \leq m \leq n$). Once π_n is found, the optimal mean square estimate of $f(X_n)$ can be calculated as

$$\mathbb{E}(f(X_n)|\mathcal{F}^Y_{0,n}) = \int f(x)\,\pi_n(dx)$$

for any function f with $\mathbb{E}|f(X_n)|^2 < \infty$. If both X and (X, Y) are Markov processes, π_n satisfies a recursive *filtering* equation. Specifically, let Λ and ν denote the transition probability and the initial distribution of X, i.e., for $A \in \mathcal{B}(\mathbb{S})$

$$\begin{aligned} &\nu(A) = \mathbb{P}(X_0 \in A), \\ &\Lambda(X_{n-1}, A) = \mathbb{P}(X_n \in A|\mathcal{F}^X_{0,n-1}) \qquad \mathbb{P}\text{-a.s.}, \end{aligned} \tag{1.2}$$

and assume that Y is a sequence of conditionally independent random variables given $\mathcal{F}^X_{0,\infty} := \bigvee_{n\geq 0} \mathcal{F}^X_{0,n}$ with

$$\begin{aligned} &\mathbb{P}(Y_0 = 0) = 1, \\ &\mathbb{P}(Y_n \in A|\mathcal{F}^X_{0,\infty}) = \int_A g(X_n, y)\,\psi(dy), \qquad n \geq 1 \end{aligned} \tag{1.3}$$

where $g(x, y)$ is a probability density with respect to the σ-finite measure ψ (the deterministic choice of Y_0 is only a matter of convenience; it means that all the

[1] Typical choices are a finite or countable set, a subset of $\mathbb{R}^q$ for some $q \geq 1$, or $\mathbb{R}^q$ itself.

information about X_0 is contained in its a priori distribution ν). For such a model π_n satisfies the recursive equation (see, e.g., Proposition 3.2.5 in [7])

$$\pi_n(dx) = \frac{g(x, Y_n) \int \Lambda(u, dx)\, \pi_{n-1}(du)}{\int g(x, Y_n) \int \Lambda(u, dx)\, \pi_{n-1}(du)}, \tag{1.4}$$

subject to $\pi_0 = \nu$. Suppose (1.4) can be solved starting from a probability distribution $\bar{\nu}$ different from ν and denote by $\bar{\pi}_n$ the resulting sequence of random measures. A typical question of stability is under which conditions (in terms of the model ingredients Λ, g, etc.), the distance between π_n and $\bar{\pi}_n$ vanishes as $n \to \infty$:

$$\lim_{n\to\infty} \mathbb{E}\|\pi_n - \bar{\pi}_n\| = 0, \tag{1.5}$$

where $\|\cdot\|$ denotes the total variation norm.[2] If the filter (1.4) is started from two initial conditions $\bar{\nu}$ and $\tilde{\nu}$, both different from ν, the distance between the corresponding solutions satisfies $\|\bar{\pi}_n - \tilde{\pi}_n\| \le \|\bar{\pi}_n - \pi_n\| + \|\pi_n - \tilde{\pi}_n\|$ and it therefore suffices to consider the case where $\tilde{\nu} = \nu$ is the true initial distribution.

Depending on the way the filtering equation is thought of, different tools can be used to solve this problem. For example, (1.4) can be seen as the iteration of a positive random operator acting on nonnegative measures and consequently filter stability can be treated using the appropriate tools from the theory of positive operators, namely the Birkhoff contraction inequality for the Hilbert projective metric (see, e.g., [2], [5], [22, 23], [20]). Equation (1.4) can also be considered as a random dynamical system with a special projective structure, so that the stability problem can be related to the Lyapunov exponents of the bilinear Zakai equation for the unnormalized conditional law ([1], [6], [11, 10]). In the continuous time case, when both signal and the observation processes are sufficiently regular diffusions, the filtering equation corresponds to certain stochastic PDE and can be analyzed using PDE tools ([30, 31]). These approaches are reviewed elsewhere in this volume.

In contrast to the above techniques, which essentially study the filtering *recursion* (1.4), this article aims to survey results which rely fundamentally on the probabilistic representation (1.1) of the filtering process π_n as a *conditional expectation*. As will shortly become evident, this 'intrinsic' approach is particularly transparent when we impose the following condition:

$$\nu \ll \bar{\nu}. \tag{A}$$

Though this assumption can be weakened in certain cases, we will generally restrict ourselves to this setting in the following for sake of simplicity (further details on the relevance of this assumption can be found in section 5).

To give the reader an idea about the methods we have in mind, let us begin our investigation of the filter stability problem assuming only (A). Remarkably, significant insight can be gained already at this level of generality. Let $\bar{\mathbb{P}}$ be the probability measure on $\mathcal{F}$, such that under $\bar{\mathbb{P}}$ the process (X, Y) has the same transition law as under $\mathbb{P}$, but $X_0 \sim \bar{\nu}$. Then (A) implies that $\mathbb{P} \ll \bar{\mathbb{P}}$ with

$$\frac{d\mathbb{P}}{d\bar{\mathbb{P}}}(X, Y) = \frac{d\nu}{d\bar{\nu}}(X_0) \qquad \bar{\mathbb{P}}\text{-a.s.} \tag{1.6}$$

The random measures $\pi_n(\cdot)$ and $\bar{\pi}_n(\cdot)$ obtained by the recursion (1.4) are regular versions of the conditional probabilities $\mathbb{P}\big(X_n \in \cdot\,|\mathcal{F}^Y_{0,n}\big)$ and $\bar{\mathbb{P}}\big(X_n \in \cdot\,|\mathcal{F}^Y_{0,n}\big)$, respectively. Since $\mathbb{P} \ll \bar{\mathbb{P}}$, $\bar{\pi}_n(\cdot)$ is well defined on a set of full $\mathbb{P}$-probability,

[2] We will denote the Euclidean norm on $\mathbb{R}^p$ as $|\cdot|$.

which means that (1.4) can be solved subject to $\bar\nu$ when the actual observations are drawn from $\mathbb{P}$. Using (1.6) and the Bayes formula, we find that there is a set of full $\mathbb{P}$-probability on which for any bounded measurable f

$$\begin{aligned}\int_{\mathbb{S}} f(x)\,d\pi_n(x) = \mathbb{E}\big(f(X_n)\big|\mathcal{F}^Y_{0,n}\big) &= \frac{\bar{\mathbb{E}}\big(f(X_n)\frac{d\nu}{d\bar\nu}(X_0)\big|\mathcal{F}^Y_{0,n}\big)}{\bar{\mathbb{E}}\big(\frac{d\nu}{d\bar\nu}(X_0)\big|\mathcal{F}^Y_{0,n}\big)} \\ &= \bar{\mathbb{E}}\left(f(X_n)\,\frac{\bar{\mathbb{E}}\big(\frac{d\nu}{d\bar\nu}(X_0)\big|\mathcal{F}^Y_{0,n}\vee\sigma\{X_n\}\big)}{\bar{\mathbb{E}}\big(\frac{d\nu}{d\bar\nu}(X_0)\big|\mathcal{F}^Y_{0,n}\big)}\right|\left.\mathcal{F}^Y_{0,n}\right) \\ &= \int_{\mathbb{S}} f(x)\,\frac{\bar{\mathbb{E}}\big(\frac{d\nu}{d\bar\nu}(X_0)\big|\mathcal{F}^Y_{0,n},X_n=x\big)}{\bar{\mathbb{E}}\big(\frac{d\nu}{d\bar\nu}(X_0)\big|\mathcal{F}^Y_{0,n}\big)}\,d\bar\pi_n(x)\end{aligned}$$

(note that the denominator is strictly positive $\mathbb{P}$-a.s.) Thus evidently $\pi_n \ll \bar\pi_n$ $\mathbb{P}$-a.s. and the corresponding Radon-Nikodym derivative is given by ([13])

$$\frac{d\pi_n}{d\bar\pi_n}(x) = \frac{\bar{\mathbb{E}}\big(\frac{d\nu}{d\bar\nu}(X_0)\big|\mathcal{F}^Y_{0,n},X_n=x\big)}{\bar{\mathbb{E}}\big(\frac{d\nu}{d\bar\nu}(X_0)\big|\mathcal{F}^Y_{0,n}\big)} \qquad \mathbb{P}\text{-a.s.} \tag{1.7}$$

Therefore, we obtain $\mathbb{P}$-a.s.

$$\begin{aligned}\|\pi_n-\bar\pi_n\| &= \int_{\mathbb{S}}\left|\frac{d\pi_n}{d\bar\pi_n}(x)-1\right|d\bar\pi_n(x) \\ &= \frac{\bar{\mathbb{E}}\Big(\big|\bar{\mathbb{E}}\big(\frac{d\nu}{d\bar\nu}(X_0)\big|\mathcal{F}^Y_{0,n}\vee\sigma\{X_n\}\big)-\bar{\mathbb{E}}\big(\frac{d\nu}{d\bar\nu}(X_0)\big|\mathcal{F}^Y_{0,n}\big)\big|\,\Big|\mathcal{F}^Y_{0,n}\Big)}{\bar{\mathbb{E}}\big(\frac{d\nu}{d\bar\nu}(X_0)\big|\mathcal{F}^Y_{0,n}\big)}.\end{aligned} \tag{1.8}$$

By the Markov property of (X,Y), $\mathcal{F}^Y_{0,n-1}\vee\mathcal{F}^X_{0,n-1}$ and $\mathcal{F}^Y_{n+1,\infty}\vee\mathcal{F}^X_{n+1,\infty}$ are conditionally independent given $\sigma\{X_n,Y_n\}$, which implies that

$$\bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar\nu}(X_0)\Big|\mathcal{F}^Y_{0,n}\vee\sigma\{X_n\}\Big) = \bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar\nu}(X_0)\Big|\mathcal{F}^Y_{0,\infty}\vee\mathcal{F}^X_{n,\infty}\Big).$$

Combined with (1.8), this implies that

$$\begin{aligned}\mathbb{E}\|\pi_n-\bar\pi_n\| = \bar{\mathbb{E}}\left(\bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar\nu}(X_0)\Big|\mathcal{F}^Y_{0,n}\Big)\,\|\pi_n-\bar\pi_n\|\right) = \\ \bar{\mathbb{E}}\Big|\bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar\nu}(X_0)\Big|\mathcal{F}^Y_{0,\infty}\vee\mathcal{F}^X_{n,\infty}\Big)-\bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar\nu}(X_0)\Big|\mathcal{F}^Y_{0,n}\Big)\Big|.\end{aligned} \tag{1.9}$$

The conditional expectations in the latter expression are nonnegative uniformly integrable martingales with respect to the decreasing and increasing filtrations $\mathcal{F}^Y_{0,\infty}\vee\mathcal{F}^X_{n,\infty}$ and $\mathcal{F}^Y_{0,n}$, respectively. Hence both converge in $L^1(\bar{\mathbb{P}})$, and thus

$$\lim_{n\to\infty}\mathbb{E}\|\pi_n-\bar\pi_n\| = \bar{\mathbb{E}}\left|\bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar\nu}(X_0)\Big|\bigcap_{n\ge 0}\mathcal{F}^Y_{0,\infty}\vee\mathcal{F}^X_{n,\infty}\Big)-\bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar\nu}(X_0)\Big|\mathcal{F}^Y_{0,\infty}\Big)\right|.$$

This suggests that the filter is stable if

$$\bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar\nu}(X_0)\Big|\bigcap_{n\ge 0}\mathcal{F}^Y_{0,\infty}\vee\mathcal{F}^X_{n,\infty}\Big) = \bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar\nu}(X_0)\Big|\mathcal{F}^Y_{0,\infty}\Big) \quad \bar{\mathbb{P}}\text{-a.s.} \tag{1.10}$$

In particular, under assumption (A), (1.5) holds if and only if (1.10) holds.

It is tempting to interchange the supremum and the intersection operations on the σ-algebras in the left hand side of (1.10)

$$\bigcap_{n\geq 0} \mathcal{F}^Y_{0,\infty} \vee \mathcal{F}^X_{n,\infty} \stackrel{?}{=} \mathcal{F}^Y_{0,\infty} \vee \bigcap_{n\geq 0} \mathcal{F}^X_{n,\infty}, \tag{1.11}$$

as this would imply stability for a large class of signals, namely those with a.s. empty tail field $\bigcap_{n\geq 0} \mathcal{F}^X_{n,\infty}$. Unfortunately, the exchange of intersection and supremum need not be permitted if no further constraints on the model are imposed, as the illuminating example below shows (see [19], [16], [3, 8] for related discussions). This subtle problem was not recognized in the pioneering work of H. Kunita [21], where the relation (1.11) was taken for granted, and was subsequently inherited by a number of contributions that are based on [21]. The insight gained in [3] from the 'intrinsic' perspective on the filter stability problem revealed this as a serious gap in [21], which to date has not yet been completely resolved. The validity of (1.11) was recently verified in [36] under slightly stronger assumptions than imposed in [21]; see a sketch of the ideas in Section 3 below.

Example 1.1. Let X be a Markov chain on $\mathbb{S} = \{1, 2, 3, 4\}$ with transition matrix

$$\Lambda = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 1/2 & 0 & 0 & 1/2 \end{pmatrix}.$$

Let $Y_n = \mathbf{1}_{\{X_n \in \{1,3\}\}}$, $n \geq 1$. If one observes Y and X_k for some $k \geq 1$, then the whole trajectory of X is revealed. Indeed, Y reveals exactly when the transitions of X occur, and the knowledge of a single value of X_k then pins down which one of the two possible trajectories of X occurs given the known transition times. Hence $\mathcal{F}^Y_{0,\infty} \vee \sigma\{X_k\} = \mathcal{F}^Y_{0,\infty} \vee \mathcal{F}^X_{0,\infty} = \mathcal{F}^Y_{0,\infty} \vee \mathcal{F}^X_{n,\infty}$ for any $n \geq 1$ and therefore

$$\bigcap_{n\geq 0} \mathcal{F}^Y_{0,\infty} \vee \mathcal{F}^X_{n,\infty} = \mathcal{F}^Y_{0,\infty} \vee \mathcal{F}^X_{0,\infty}. \tag{1.12}$$

Recall that a finite state Markov chain is ergodic, i.e., irreducible and aperiodic, if and only if its transition matrix is primitive of order m (the entries of Λ^m are positive for some integer $m \geq 1$). An ergodic chain has almost surely trivial tail σ-algebra. The transition matrix defined above is primitive of order 3 and hence $\bigcap_{n\geq 0} \mathcal{F}^X_{n,\infty}$ is empty $\bar{\mathbb{P}}$-a.s. Therefore $\mathcal{F}^Y_{0,\infty} \vee \bigcap_{n\geq 0} \mathcal{F}^X_{n,\infty} = \mathcal{F}^Y_{0,\infty}$ $\bar{\mathbb{P}}$-a.s.

However, observing Y alone does not eliminate the uncertainty about X_0 (and thus about the whole trajectory of X):

$$\bar{\mathbb{P}}(X_n = 1 | \mathcal{F}^Y_{0,\infty}) = \frac{\bar{\nu}(1)}{\bar{\nu}(1) + \bar{\nu}(3)} \mathbf{1}_{\{Y_n = 1\}} \neq \mathbf{1}_{\{X_n = 1\}},$$

which means that $\mathcal{F}^Y_{0,\infty}$ is strictly smaller than $\mathcal{F}^Y_{0,\infty} \vee \mathcal{F}^X_{0,\infty}$ and by (1.12) than $\bigcap_{n\geq 0} \mathcal{F}^Y_{0,\infty} \vee \mathcal{F}^X_{n,\infty}$. Therefore, both (1.10) and (1.11) fail. In this case, equation (1.4) shows that $\|\bar{\pi}_n - \pi_n\| \geq C$ for all $n \geq 0$, where C is a positive constant depending only on ν and $\bar{\nu}$. Thus the filter is not stable.

Evidently, contrary to intuition, ergodicity of the signal (i.e., triviality of the tail σ-field) alone may not be enough to guarantee filter stability. What additional ingredient is needed? We will consider several possibilities below, including:

- In section 2, we will show that the filter is stable regardless of the observation structure if the signal possesses a strong mixing property. This assumption is stronger than ergodicity and holds, for example, if the signal has a uniformly positive transition density.
- In section 3, we will show that ergodicity of the signal is already sufficient for filter stability if the observations are nondegenerate, i.e., $g(x, y) > 0$.
- In section 4, we will show that the filter may be stable even when the signal is not ergodic provided that the observations are 'good enough'.

These results indicate that stability of the filter emerges as an interaction between the ergodic properties of the signal and the structure of the observations, a complete understanding of which is still lacking. For example, necessary and sufficient conditions for stability are unknown in the general setting, and the existing sufficient conditions are often difficult to verify in terms of the filtering model. Moreover, the difficult quantitative question of how the rate of stability of the filter is affected by the ergodic properties of the signal and the quality of the observations remains largely open. Despite the abundance of open questions, however, the various results reviewed in this paper indicate that significant insight can be obtained by employing an intrinsic analysis of the filter stability problem.

The remainder of the paper consists of four sections: sections 2, 3 and 4 each describes a particular intrinsic argument for (1.5) (and sometimes its stronger/weaker forms) to hold, while section 5 explores when (1.5) cannot hold.

In Section 2, inspired by (1.9), we explore the connection between the stability of the filter and *smoothing* conditional expectations of X_0 given $\mathcal{F}^Y_{0,\infty} \vee \mathcal{F}^X_{n,\infty}$ under $\bar{\mathbb{P}}$. The main outcome is that the so called *mixing condition*, which is often imposed on the signal transition law by other methods, can be relaxed regardless of the observation noise density ([3], [8]).

Section 3 studies the (filtering) conditional distribution of X_n given $\mathcal{F}^Y_{0,n}$ as the marginal of the law induced on the space of signal trajectories by conditioning on $\mathcal{F}^Y_{0,n}$. The latter is well known to correspond to a time inhomogeneous Markov process on the signal state space, whose transition probability is controlled by the observation path. This fact places at our disposal a number of tools from the theory of Markov processes, including coupling ([18], [35]). If we condition on $\mathcal{F}^Y_{0,\infty}$ instead, the 'conditioned signal' approach can be used ([36]) to verify (1.11).

Section 4 deals with the stabilizing role of the observations. It turns out that estimates of particular functions are stable, in the sense that

$$\lim_{n\to\infty} \mathbb{E}|\pi_n(f) - \bar{\pi}_n(f)| = 0 \tag{1.13}$$

for certain f, even in cases where (1.5) may fail. This is possible when f is *observable* in an appropriate sense [13], [9], [32, 33, 34]. When every function is observable, it follows that the filter is stable in the sense that (1.13) holds for all f. Remarkably, these results do not rely on any ergodic property of the signal as in the earlier sections, but emerge instead when the observations are 'sufficiently informative'.

Finally, we will show in section 5 that there are some inherent limitations to when (1.5) can hold. In particular, we will discuss how far the assumption (A) can be weakened, and we will argue that some form of absolute continuity is in fact necessary for the filter to forget its initial condition in the sense of (1.5).

A notable omission from this article is the pioneering approach to the filter stability problem, due to Ocone and Pardoux [27], who deduce stability of the nonlinear filter from the results of Kunita [21] by weak convergence arguments. Though this is very much an 'intrinsic' approach, it unfortunately appeals directly to the argument in [21] where (1.11) is taken for granted, and this gap is therefore inherited. Nonetheless this approach remains of significant interest, particularly as some of the machinery is of use in applications to Monte Carlo approximations of nonlinear filters (see [14]). The approach of Ocone and Pardoux, and its relation with the work of Kunita and the gap therein, is discussed elsewhere in this volume.

In order to keep the presentation as transparent as possible, we do not formulate the results in the most general form possible and only sketch the proofs, emphasizing the key ideas. We refer the reader to the original articles for the details of the proofs and for the (important!) technicalities.

2. Stability via smoothing

The formula (1.8) suggests that the filter is stable only if the conditional expectation of X_0 given $\mathcal{F}^Y_{0,n} \vee \sigma\{X_n\}$ ceases to depend on X_n as $n \to \infty$. The *smoothing* problem of computing the conditional distribution of X_0 given $\mathcal{F}^Y_{0,n} \vee \sigma\{X_n\}$ leads to a linear equation, whose long time behavior can be efficiently studied for strongly mixing signals.

Consider the signal/observation model (1.2) and (1.3), where the signal transition probability $\Lambda(u, dy)$ is assumed to have a density $\lambda(x, y)$ with respect to some σ-finite measure $\varphi(dy)$, i.e.,

$$\Lambda(x, dy) = \lambda(x, y)\, \varphi(dy) \qquad \forall\, x \in \mathbb{S}. \tag{2.1}$$

Suppose $\bar\nu$ has a density with respect to φ. Then the regular conditional probability $\bar{\mathbb{P}}(X_0 \in \cdot\,|\mathcal{F}^Y_{0,n} \vee \sigma\{X_n\})$ also has a density $q_n(u; x)$:

$$\bar{\mathbb{P}}(X_0 \in A|\mathcal{F}^Y_{0,n} \vee \sigma\{X_n\}) = \int_A q_n(u; X_n)\, \varphi(du) \qquad \forall A \in \mathcal{B}(\mathbb{S}) \qquad \bar{\mathbb{P}}\text{-a.s.}$$

A simple calculation shows that q_n satisfies the recursion (see Lemma 3.1 in [8])

$$\begin{aligned} q_1(u; x) &= \frac{\lambda(u, x)\frac{d\bar\nu}{d\varphi}(u)}{\int_{\mathbb{S}} \lambda(v, x)\, \bar\nu(dv)}, \\ q_n(u; x) &= \frac{\int_{\mathbb{S}} \lambda(z, x)\, q_{n-1}(u; z)\, \bar\pi_{n-1}(dz)}{\int_{\mathbb{S}} \lambda(v, x)\, \bar\pi_{n-1}(dv)}, \qquad n \geq 2. \end{aligned} \tag{2.2}$$

Define

$$\hat q_n(u) = \sup_{x \in \mathbb{S}} q_n(u; x), \qquad \check q_n(u) = \inf_{x \in \mathbb{S}} q_n(u; x).$$

Our goal is to show that $\Delta_n(u) := \hat q_n(u) - \check q_n(u)$ converges to zero as $n \to \infty$, i.e., that $q_n(u; x)$ ceases to depend on its second argument. Let

$$\alpha_{n-1}(u; z) := \frac{\hat q_{n-1}(u) - q_{n-1}(u; z)}{\hat q_{n-1}(u) - \check q_{n-1}(u)}$$

(with the convention $0/0 = 0$). Then by (2.2), for any $x, x' \in \mathbb{S}$ and $n \geq 2$,

$$\begin{aligned} q_n(u; x) - q_n(u; x') = \Delta_{n-1}(u)\bigg(1 - \int_{\mathbb{S}} \bigg\{ &\frac{\lambda(z, x)}{\int_{\mathbb{S}} \lambda(v, x)\bar\pi_{n-1}(dv)} \alpha_{n-1}(u; z) \\ &+ \frac{\lambda(z, x')}{\int_{\mathbb{S}} \lambda(v, x')\bar\pi_{n-1}(dv)} \big(1 - \alpha_{n-1}(u; z)\big) \bigg\} \bar\pi_{n-1}(dz) \bigg). \end{aligned}$$

Assume that the transition density is uniformly bounded, i.e. $\lambda(x,u)\le \lambda^*<\infty$ for some constant λ^*. Since $\alpha_n\in[0,1]$,

$$q_n(u;x)-q_n(u;x')\le \Delta_{n-1}(u)\Big(1-\frac{1}{\lambda^*}\int_{\mathbb{S}}\big\{\lambda(z,x)\wedge\lambda(z,x')\big\}\bar\pi_{n-1}(dz)\Big),$$

and by the arbitrariness of x and x',

$$\Delta_n(u)\le \Delta_{n-1}(u)\Big(1-\frac{1}{\lambda^*}\int_{\mathbb{S}}\inf_{x\in\mathbb{S}}\lambda(z,x)\,\bar\pi_{n-1}(dz)\Big). \tag{2.3}$$

Notice that

$$\begin{aligned}\bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar\nu}(X_0)\Big|\mathcal{F}^Y_{0,n}\vee\sigma\{X_n\}\Big)&=\int_{\mathbb{S}}\frac{d\nu}{d\bar\nu}(u)\,q_n(u;X_n)\,\varphi(du)\\ \bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar\nu}(X_0)\Big|\mathcal{F}^Y_{0,n}\Big)&=\int_{\mathbb{S}}\int_{\mathbb{S}}\frac{d\nu}{d\bar\nu}(u)\,q_n(u;x)\,\varphi(du)\,\bar\pi_n(dx)\end{aligned}$$

and assume $\frac{d\nu}{d\bar\nu}(u)\ge\varepsilon>0$ for a constant $\varepsilon>0$. Then by (1.8)

$$\begin{aligned}\|\pi_n-\bar\pi_n\|&=\frac{\int_{\mathbb{S}}\Big|\int_{\mathbb{S}}\int_{\mathbb{S}}\frac{d\nu}{d\bar\nu}(u)\,\big(q_n(u;x')-q_n(u;x)\big)\varphi(du)\bar\pi_n(dx)\Big|\,\bar\pi_n(dx')}{\bar{\mathbb{E}}\big(\frac{d\nu}{d\bar\nu}(X_0)\big|\mathcal{F}^Y_{0,n}\big)}\\ &\le\frac{1}{\varepsilon}\int_{\mathbb{S}}\int_{\mathbb{S}}\int_{\mathbb{S}}\frac{d\nu}{d\bar\nu}(u)\,\big|q_n(u;x')-q_n(u;x)\big|\,\varphi(du)\,\bar\pi_n(dx)\,\bar\pi_n(dx')\\ &\le\frac{1}{\varepsilon}\int_{\mathbb{S}}\frac{d\nu}{d\bar\nu}(u)\,\Delta_n(u)\,\varphi(du),\end{aligned}\tag{2.4}$$

and

$$\mathbb{E}\|\pi_n-\bar\pi_n\|\le\int_{\mathbb{S}}\frac{d\nu}{d\bar\nu}(u)\,\bar{\mathbb{E}}\Delta_n(u)\,\varphi(du). \tag{2.5}$$

Now, if a constant $\lambda^*>0$ can be found such that $\lambda(x,u)\ge\lambda_*$ for all $x,u\in\mathbb{S}$, then by the first equation in (2.2)

$$\begin{aligned}\int_{\mathbb{S}}\frac{d\nu}{d\bar\nu}(u)\,\Delta_1(u)\,\varphi(du)&\le\int_{\mathbb{S}}\frac{d\nu}{d\bar\nu}(u)\,\sup_{x\in\mathbb{S}}q_1(u;x)\,\varphi(du)\\ &\le\int_{\mathbb{S}}\frac{d\nu}{d\bar\nu}(u)\,\sup_{x\in\mathbb{S}}\frac{\lambda(u,x)\frac{d\bar\nu}{d\varphi}(u)}{\int_{\mathbb{S}}\lambda(v,x)\,\bar\nu(dv)}\,\varphi(du)\le\lambda^*/\lambda_*.\end{aligned}$$

The latter and (2.3)-(2.5) give the the following bounds

Theorem 2.1. *Assume*

$$0<\lambda_*<\lambda(x,u)\le\lambda^*<\infty, \tag{2.6}$$

and $\nu\ll\bar\nu$, *then*

$$\mathbb{E}\|\pi_n-\bar\pi_n\|\le\frac{\lambda^*}{\lambda_*}\Big(1-\frac{\lambda_*}{\lambda^*}\Big)^{n-1}. \tag{2.7}$$

If in addition, $\frac{d\nu}{d\bar\nu}(x)\ge\varepsilon>0$ *with a constant* $\varepsilon>0$, *then*

$$\|\pi_n-\bar\pi_n\|\le\frac{1}{\varepsilon}\frac{\lambda^*}{\lambda_*}\Big(1-\frac{\lambda_*}{\lambda^*}\Big)^{n-1}\qquad\mathbb{P}\text{-}a.s. \tag{2.8}$$

The condition (2.6) forces the transition density of the signal to be bounded away from zero uniformly over $\mathbb{S}$. This can be somewhat relaxed, at the expense of giving up the time uniformity of the bound (2.8). Suppose X is an aperiodic irreducible Markov chain with the unique invariant measure μ and that

$$\lambda_\diamond := \int_{\mathbb{S}} \inf_{x\in\mathbb{S}} \lambda(u,x)\,\mu(du) > 0. \tag{2.9}$$

In this case $\mu(dx)$ has a density $m(x)$ with respect to φ, satisfying $\lambda_\diamond \le m(u) \le \lambda^*$. If we assume that $\frac{d\bar\nu}{d\varphi}(x) \ge \varepsilon > 0$, then by (2.2)

$$\begin{aligned}\int_{\mathbb{S}} \frac{d\nu}{d\bar\nu}(u)\,\Delta_1(u)\,\varphi(du) &\le \int_{\mathbb{S}} \frac{d\nu}{d\bar\nu}(u)\,\sup_{x\in\mathbb{S}} q_1(u;x)\,\varphi(du)\\ &\le \frac{(\lambda^*)^2}{\varepsilon}\int_{\mathbb{S}} \frac{d\nu}{d\bar\nu}(u)\left(\int_{\mathbb{S}} \inf_{x\in\mathbb{S}} \lambda(v,x) m(v)\varphi(dv)\right)^{-1} \bar\nu(du) \le \frac{(\lambda^*)^2}{\varepsilon\lambda_\diamond}.\end{aligned}$$

This, combined with (2.3) and (2.4), gives

$$\|\pi_n - \bar\pi_n\| \le \frac{(\lambda^*)^2}{\varepsilon^2\lambda_\diamond}\prod_{m=1}^{n-1}\left(1 - \frac{1}{\lambda^*}\int_{\mathbb{S}} \inf_{x\in\mathbb{S}} \lambda(z,x)\,\bar\pi_m(dz)\right). \tag{2.10}$$

Finally, (2.9) implies (Theorem 2.1 in [8]) that the chain X is geometrically ergodic (i.e., its marginal distribution converges to μ in total variation geometrically fast), and that $\bar\pi_n$ satisfies the law of large numbers under $\bar{\mathbb{P}}$ (Theorem 2.2 in [8]):

$$\lim_{n\to\infty}\frac{1}{n}\sum_{m=1}^{n-1}\int_{\mathbb{S}} \inf_{x\in\mathbb{S}} \lambda(u,x)\,\bar\pi_m(du) = \int_{\mathbb{S}} \inf_{x\in\mathbb{S}} \lambda(u,x)\,\mu(du) \qquad \bar{\mathbb{P}}\text{-a.s.}$$

Since $\mathbb{P} \ll \bar{\mathbb{P}}$ the latter convergence holds $\mathbb{P}$-a.s. as well and (2.10) gives the following asymptotic bound.

Theorem 2.2 (essentially Theorem 1.1 in [8])**.** *Assume that $\frac{d\nu}{d\varphi}$ and $\frac{d\bar\nu}{d\varphi}$ are bounded away from zero and infinity uniformly over $\mathbb{S}$. Suppose that X is irreducible and aperiodic with the unique invariant measure μ, satisfying the following conditions*

$$\begin{aligned}&\lambda(x,u) \le \lambda^* < \infty\\ &\lambda_\diamond := \int_{\mathbb{S}} \inf_{x\in\mathbb{S}} \lambda(u,x)\,\mu(du) > 0.\end{aligned} \tag{2.11}$$

Then

$$\varlimsup_{n\to\infty}\frac{1}{n}\log\|\pi_n - \bar\pi_n\| \le -\frac{\lambda_\diamond}{\lambda^*} \qquad \mathbb{P}\text{-}a.s.$$

Remark 2.3. The statement of Theorem 2.2 remains true under weaker assumptions on $\frac{d\nu}{d\bar\nu}$ (see [8]) since only an asymptotic bound is obtained.

The condition (2.11) is significantly weaker than (2.6): for example, in the case of an ergodic chain X which takes a finite number of values, (2.6) requires that all the entries of the transition matrix are positive, while (2.11) holds when there is at least one row with strictly positive entries. For instance, (2.11) holds true and the filter becomes stable, if the transition matrix of the signal chain from Example 1.1

is perturbed with an $\varepsilon \in (0,1)$ in a single row:

$$\Lambda = \begin{pmatrix} 1/2(1-\varepsilon) & 1/2(1-\varepsilon) & \varepsilon/2 & \varepsilon/2 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 1/2 & 0 & 0 & 1/2 \end{pmatrix}. \tag{2.12}$$

However, both condition (2.6) and condition (2.11) imply that X has a strong mixing property with geometric rate.

The more serious drawback of both types of mixing, (2.6) and (2.11), is that they are not well suited to noncompact $\mathbb{S}$. For instance, the condition (2.6) is equivalent to requiring that for any $x_1, x_2 \in \mathbb{S}$, $\Lambda(x_1, \cdot)$ has a density with respect to $\Lambda(x_2, \cdot)$, which is uniformly bounded from zero and infinity over the pairs (x_1, x_2). Indeed, if the latter is true, one can just take $\varphi(\cdot) = \Lambda(x, \cdot)$ for an $x \in \mathbb{S}$. Conversely, (2.6) means that there exists a σ-finite measure φ such that

$$\lambda_* \varphi(A) \le \Lambda(x, A) \le \lambda^* \varphi(A)$$

for all measurable A and all $x \in S$. If such measure exists then for any $x_1, x_2 \in \mathbb{S}$

$$(\lambda_*/\lambda^*)\, \Lambda(x_2, A) \le \Lambda(x_1, A) \le (\lambda^*/\lambda_*)\, \Lambda(x_2, A),$$

and hence

$$\lambda_*/\lambda^* \le \frac{d\Lambda(x_1, \cdot)}{d\Lambda(x_2, \cdot)} \le \lambda^*/\lambda_* \quad \forall\, x_1, x_2 \in \mathbb{S}$$

The latter property is relatively easy to check when $\mathbb{S}$ is compact. For noncompact $\mathbb{S}$, such a choice of φ is sometimes impossible.

Example 2.4 (taken from [23]). Let $\mathbb{S} = \mathbb{R}$. Suppose X is generated by the recursion $X_n = h(X_{n-1}) + Z_n$, where h is a bounded function and Z is a sequence of i.i.d. random variables independent of X. Assume Z_1 has a Laplacian distribution, i.e.,

$$\mathbb{P}(Z_1 \le x) = \int_{-\infty}^{x} \frac{1}{2} e^{-|u|} du, \qquad x \in \mathbb{R}.$$

The Lebesgue measure is obviously not the right choice for φ, since the corresponding transition density $\lambda(x, u) = \frac{1}{2} e^{-|u - h(x)|}$ violates the lower bound. However if one chooses $\varphi(du) = \frac{1}{2} e^{-|u|} du$, the density

$$\lambda(x, u) = e^{-|u - h(x)| + |u|}$$

is bounded between $\lambda_* := e^{-\|h\|_\infty}$ and $\lambda^* := e^{\|h\|_\infty}$ and (2.6) becomes applicable. If, however, Z_1 has standard Gaussian distribution, there is no φ which would guarantee (2.6), since

$$\frac{d\Lambda(x_1, \cdot)}{d\Lambda(x_2, \cdot)}(u) = \frac{e^{-(u - h(x_1))^2/2}}{e^{-(u - h(x_2))^2/2}} = e^{(h(x_1) - h(x_2))u - \frac{1}{2}(h^2(x_1) - h^2(x_2))}$$

is not bounded for $h(x_1) \neq h(x_2)$.

Theorem 2.1 can be proved using different methods, including those based on the Birkhoff (as in [2]) or Dobrushin (as in [15]) contraction inequalities. However, we are not aware of an alternative proof of the result stated in Theorem 2.2 and to the best of our knowledge, the condition (2.11) is the weakest known ergodic property of the signal which implies filter stability *without* further constraints on the observation model (note, in particular, that no assumptions are needed on the observation density $g(x, y)$ in order for Theorems 2.1 and 2.2 to hold).

2.1. **Continuous time.** Equation (1.8) remains valid when the time parameter is continuous, and thus the approach based on analysis of the smoothing equation is applicable to continuous time models as well.

Let $X = (X_t)_{t\in\mathbb{R}_+}$ be a Markov chain with values in $\mathbb{S} = \{a_1, \dots, a_d\}$, transition rates λ_{ij} and initial distribution ν. The real-valued observation process is given by

$$Y_t = \int_0^t h(X_s)ds + \sigma B_t, \quad t \in \mathbb{R}_+, \tag{2.13}$$

with $h : \mathbb{S} \mapsto \mathbb{R}$, $\sigma > 0$ is a constant (noise intensity), and B is a Brownian motion independent of X. The nonlinear filter in this case is finite dimensional and the vector of the conditional probabilities $\pi_t(i) = \mathbb{P}(X_t = a_i|\mathcal{F}_t^Y)$ solves the Shiryaev-Wonham stochastic differential equation (see Chapter 9 in [24])

$$d\pi_t = \Lambda^\top \pi_t dt + \sigma^{-2}\big(\operatorname{diag}(\pi_t) - \pi_t\pi_t^\top\big)h(dY_t - h^\top \pi_t dt), \quad \pi_0 = \nu, \tag{2.14}$$

where $\operatorname{diag}(x)$, $x \in \mathbb{R}^d$ stands for the diagonal matrix with entries x_i, Λ is the matrix of transition rates, and h is a vector[3] with entries $h(a_i)$, $i = 1, \dots, d$. Denote by $\bar\pi_t$ the strong solution of (2.14) subject to $\bar\nu \in \mathcal{P}(\mathbb{S}) = \mathcal{S}^{d-1}$.

Recall that X is ergodic, i.e., irreducible and aperiodic, if and only if the matrix exponential $\exp(\Lambda)$ has strictly positive entries, or, equivalently, if all the entries of Λ communicate. An ergodic chain has a unique invariant measure μ and $\lim_{t\to\infty} \mathbb{P}(X_t = a_i) = \mu_i > 0$, $i = 1, \dots, d$.

Theorem 2.5. *Assume that X is ergodic and $\sigma > 0$. Then for any $\nu, \bar\nu \in \mathcal{S}^{d-1}$ the following stability properties hold:*

$$\gamma := \lim_{t\to\infty} \frac{1}{t} \log \|\pi_t - \bar\pi_t\| < 0 \qquad \mathbb{P}\text{-a.s.}, \tag{2.15}$$

$$\lim_{t\to\infty} \frac{1}{t} \log \|\pi_t - \bar\pi_t\| \le -\sum_{i=1}^d \mu_i \min_{j\ne i} \lambda_{ij} \qquad \mathbb{P}\text{-a.s.}, \tag{2.16}$$

and

$$\|\pi_t - \bar\pi_t\| \le C \exp\big\{ -2t \min_{i\ne j} \sqrt{\lambda_{ij}\lambda_{ji}}\big\}, \tag{2.17}$$

where $C := 2 \wedge \max_k \left(\frac{1}{\nu_k} \vee \frac{1}{\bar\nu_k}\right)\|\nu - \bar\nu\|$.

The inequality (2.15) appeared in Theorem 4.1 [3] and its proof uses the Birkhoff contraction inequality following [1]. It says that the filter is actually exponentially stable if the observation noise is nondegenerate and the signal is ergodic. If $\sigma = 0$, then the filtering equation looks different from (2.14) and can be unstable as in the Example 1.1 (see Section 3 in [3]). The existence of the limit in (2.15) and (2.16) follows from Oseledec's Multiplicative Ergodic Theorem (see [1]). Both (2.16) and the time uniform bound (2.17) were derived in [3] (Theorems 4.2 and 4.3) by the same arguments used in the proof of Theorem (2.2) above (the asymptotic version of (2.17) appeared before in [1]). Notice that (2.16) remains nontrivial as long as Λ has at least one row with nonzero entries, unlike (2.17) which requires that none of the transition rates vanish. The particular value of the constant in (2.17), taken from Proposition 3.5 [12] and Corollary 2.3.2 in [35], makes precise the dependence

[3]Functions and measures on finite $\mathbb{S}$ are identified with vectors in $\mathbb{R}^d$ and $\mathcal{S}^{d-1} = \{x \in \mathbb{R}^d : \sum_{i=1}^d x_i = 1, x_i \ge 0\}$, respectively.

on the initial conditions. Other bounds, which shed light on the dependence of γ on the noise intensity σ, etc., can be found in [16], [1], [10], [11].

3. Conditioned signal

3.1. **Finite horizon conditioning.**

The ideas outlined in this section are based on the following simple consequence of the Markov property of (X,Y):

$$\mathbb{P}\big(X_m \in A|\mathcal{F}^X_{0,m-1} \vee \mathcal{F}^Y_{0,n}\big) = \mathbb{P}\big(X_m \in A|\sigma\{X_{m-1}\} \vee \mathcal{F}^Y_{0,n}\big) \qquad \mathbb{P}\text{-a.s.} \tag{3.1}$$

For simplicity of notation, we shall consider hereafter the coordinate processes (X,Y) on the canonical space $(\Omega,\mathcal{F})$ (i.e., Ω is the space of semi-infinite sequences of points in $\mathbb{S}\times\mathbb{R}^p$). Denote by $\mathbb{P}^Y_n$ the regular conditional probability measure, induced on the restriction of the signal paths to the time interval $[0,n]$ by conditioning on $\mathcal{F}^Y_{0,n}$:

$$\mathbb{P}^Y_n(\Gamma) = \mathbb{P}\big((X_0,\dots,X_n)\in\Gamma\big|\mathcal{F}^Y_{0,n}\big), \qquad \Gamma\in\mathcal{B}(\mathbb{S}^{n+1}).$$

Then π_n is nothing but the law of X_n under $\mathbb{P}^Y_n$:

$$\pi_n(f) = \mathbb{E}\big(f(X_n)\big|\mathcal{F}^Y_{0,n}\big) = \mathbb{E}^Y_n f(X_n),$$

where $\mathbb{E}^Y_n$ stands for the expectation with respect to $\mathbb{P}^Y_n$.

The property (3.1) means that the coordinate process X under $\mathbb{P}^Y_n$, referred to hereafter as the *conditioned signal*, is Markov:

$$\mathbb{P}^Y_n(X_m\in A|\mathcal{F}^X_{0,m-1}) = \mathbb{P}^Y_n(X_m\in A|\sigma\{X_{m-1}\}) \qquad \mathbb{P}^Y_n\text{-a.s.}$$

and a simple calculation should convince the reader that for the model (1.2) with (2.1) and (1.3), the transition probability of X under $\mathbb{P}^Y_n$

$$\Lambda^Y_{m|n}(X_{m-1},A) := \mathbb{P}^Y_n\big(X_m\in A|\sigma\{X_{m-1}\}\big), \qquad A\in\mathcal{B}(\mathbb{S}),$$

has a density $\lambda^Y_{m|n}(u,x)$ with respect to $\varphi(dx)$ satisfying the following backward recursion (see, e.g., Proposition 3.3.2 in [7]):

$$\begin{aligned}
\lambda^Y_{m|n}(u,x) &= \frac{\lambda(u,x)\,g(x,Y_m)\,Q_m(x)}{\int_{\mathbb{S}}\lambda(u,v)\,g(v,Y_m)\,Q_m(v)\,\varphi(dv)}, \quad m=1,\dots,n\\
Q_m(x) &= \int_{\mathbb{S}}\lambda(x,z)\,g(z,Y_{m+1})\,Q_{m+1}(z)\,\varphi(dz)\\
Q_n(x) &\equiv 1.
\end{aligned} \tag{3.2}$$

The conditioned signal X is a time inhomogeneous Markov process whose transition density at time m depends on $\mathcal{F}^Y_{m,n}$, i.e., on the future of the observed path. Notice that $\lambda^Y_{m|n}(u,x)$ is independent of the initial distribution ν of X_0, while the law of X_0 under $\mathbb{P}^Y_n$ depends on all the ingredients of the model, including ν (and the whole observation path). Let us also stress that $\mathbb{P}^Y_n$ is not the restriction of $\mathbb{P}^Y_{n+1}$ to the first n coordinates, or in other words, increasing the time horizon changes the conditional measure completely. This should not come as a surprise, since when Y_{n+1} is observed the conditional law of the whole $X_0,\dots,X_n$ changes.

As before we shall rely on the auxiliary probability measure $\bar{\mathbb{P}}$, under which (X,Y) has the same law as under $\mathbb{P}$ but $X_0\sim\bar\nu$. Conditioning on $\mathcal{F}^Y_{0,n}$ under $\bar{\mathbb{P}}$ induces a regular probability measure on the signal paths restricted to $[0,n]$ which will be denoted as $\bar{\mathbb{P}}^Y_n$:

$$\bar{\mathbb{P}}^Y_n(\Gamma) = \bar{\mathbb{P}}\big((X_0,\dots,X_n)\in\Gamma\big|\mathcal{F}^Y_{0,n}\big), \qquad \Gamma\in\mathcal{B}(\mathbb{S}^{n+1}).$$

Clearly, the conditioned signal has the same transition law under $\mathbb{P}^Y_n$ and $\bar{\mathbb{P}}^Y_n$, but different initial distributions.

As was mentioned above, the filtering conditional distribution is nothing but the restriction of $\mathbb{P}^Y_n$ to the last coordinate and hence

$$\|\pi_n - \bar{\pi}_n\| = \|\mathbb{P}^Y_n(X_n \in \cdot\,) - \bar{\mathbb{P}}^Y_n(X_n \in \cdot\,)\|. \tag{3.3}$$

This interpretation relates the filter stability problem to the mixing properties of the conditioned signal, which in turn places the ergodic theory of Markov processes at our disposal. In particular, the following fact is well known (see, e.g., [25]).

Proposition 3.1. *Suppose that $\xi = (\xi_n)_{n\ge 0}$ is an inhomogeneous Markov chain with values in a Polish space $\mathbb{S}$ and transition probabilities $K_n(x,\cdot)$ under the probability measures $\mathbb{P}$ and $\bar{\mathbb{P}}$, such that ξ_0 has distribution ν under $\mathbb{P}$ and $\bar{\nu}$ under $\bar{\mathbb{P}}$. Assume that there exists a sequence of σ-finite measures μ_n such that*

$$\varepsilon\mu_n(A) \le K_n(x,A) \le \frac{1}{\varepsilon}\mu_n(A) \qquad \forall A \in \mathcal{B}(\mathbb{S}), \tag{3.4}$$

for some fixed $\varepsilon > 0$. Then

$$\|\mathbb{P}(\xi_n \in \cdot) - \bar{\mathbb{P}}(\xi_n \in \cdot)\| \le 2(1-\varepsilon)^n, \qquad n \ge 0. \tag{3.5}$$

If the mixing condition (2.6) is satisfied, (3.2) implies

$$\Lambda^Y_{m|n}(u,A) \ge \frac{\lambda_*}{\lambda^*}\frac{\int_A g(x,Y_m)\,Q_m(x)\,\varphi(dx)}{\int_{\mathbb{S}} g(v,Y_m)\,Q_m(v)\,\varphi(dv)} =: \frac{\lambda_*}{\lambda^*}\,\mu_n(A),$$

and, similarly, $\Lambda^Y_{m|n}(u,A) \le (\lambda^*/\lambda_*)\,\mu_n(A)$ for any $A \in \mathcal{B}(\mathbb{S})$. Thus (3.4) holds with $\varepsilon := \lambda_*/\lambda^*$ and (3.3) recovers the statement of Theorem 2.1 (even without the assumption $\nu \ll \bar{\nu}$, as long as $\bar{\pi}_n$ is well defined).

Proposition 3.1 can be verified, e.g., by constructing an appropriate coupling using Nummelin's splitting technique ([26]). In the filtering context under consideration, this coupling method can be pushed further to get finer results which go beyond signals with compact state space (see [18]). In particular, the filter can be shown to be stable for linear models which are driven by noises with unimodal probability densities (see Section 5.2 [18]). However, the essential limitation of this approach stems from the time inhomogeneity of the conditioned signal. Unfortunately, the ergodic theory of inhomogeneous Markov processes is not as rich as in the homogeneous case (however, this drawback can be mitigated by conditioning on the infinite time horizon as in the following section).

The property (3.1) remains valid when the time parameter is continuous and, with some caution, the conditioned signal measure can be explicitly constructed. For example, for finite state signals, the conditioned measure corresponds to a finite state Markov chain with time-varying rates which depend on the observation trajectory. Consequently, the bound (2.17) can be derived via coupling of the conditioned chain (see Section 2.3.2 in [35]).

Another variation on the same theme, which combines both the formula (1.8) and the conditioned signal representation, is to look at the conditioned signal backwards in time, i.e., to consider the process $\tilde{X}_m := X_{n-m}$, $m = 0,\dots,n$. Note that as X_n is a Markov process, $\tilde{X}_m$ is also Markov. Now suppose that for every $x \in \mathbb{S}$, we can construct a stochastic process $\tilde{X}_n(x)$ on the same probability space such that the law of $(\tilde{X}_m(x))_{m\le n}$ under $\bar{\mathbb{E}}^Y_n$ coincides with the law of $(\tilde{X}_m)_{m\le n}$ under $\bar{\mathbb{E}}^Y_n(\,\cdot\,|\tilde{X}_0 = x)$. This point of view is particularly fruitful when the signal process is obtained from a stochastic differential equation, so that $\tilde{X}_m(x)$ can be obtained from the stochastic flow generated by this equation. In this setting

$$\bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar{\nu}}(X_0)\Big|\mathcal{F}^Y_{0,n}, X_n = x\Big) = \bar{\mathbb{E}}^Y_n\Big(\frac{d\nu}{d\bar{\nu}}(\tilde{X}_n)\Big|\tilde{X}_0 = x\Big) = \bar{\mathbb{E}}^Y_n\left\{\frac{d\nu}{d\bar{\nu}}(\tilde{X}_n(x))\right\},$$

In these terms (1.8) reads:

$$\|\pi_n - \bar{\pi}_n\| = \frac{\int_{\mathbb{S}}\Big|\bar{\mathbb{E}}^Y_n\left\{\frac{d\nu}{d\bar{\nu}}(\tilde{X}_n(y))\right\} - \int_{\mathbb{S}}\bar{\mathbb{E}}^Y_n\left\{\frac{d\nu}{d\bar{\nu}}(\tilde{X}_n(x))\right\}\bar{\pi}_n(dx)\Big|\,\bar{\pi}_n(dy)}{\int_{\mathbb{S}}\bar{\mathbb{E}}^Y_n\left\{\frac{d\nu}{d\bar{\nu}}(\tilde{X}_n(x))\right\}\bar{\pi}_n(dx)}$$

$$\le \frac{\int_{\mathbb{S}}\int_{\mathbb{S}}\bar{\mathbb{E}}^Y_n\Big|\frac{d\nu}{d\bar{\nu}}(\tilde{X}_n(y)) - \frac{d\nu}{d\bar{\nu}}(\tilde{X}_n(x))\Big|\,\bar{\pi}_n(dy)\,\bar{\pi}_n(dx)}{\int_{\mathbb{S}}\bar{\mathbb{E}}^Y_n\left\{\frac{d\nu}{d\bar{\nu}}(\tilde{X}_n(x))\right\}\bar{\pi}_n(dx)}.$$

If $\mathbb{S} = \mathbb{R}^q$ and one assumes that $\frac{d\nu}{d\bar{\nu}}(x)$ is a Lipschitz function which is bounded away from zero by $\varepsilon > 0$, then the latter implies

$$\|\pi_n - \bar{\pi}_n\| \le \frac{1}{\varepsilon}\Big\|\frac{d\nu}{d\bar{\nu}}\Big\|_{\mathrm{Lip}}\iint_{\mathbb{S}\times\mathbb{S}}\bar{\mathbb{E}}^Y_n\big|\tilde{X}_n(y) - \tilde{X}_n(x)\big|\,\bar{\pi}_n(dy)\,\bar{\pi}_n(dx).$$

This translates the filter stability problem to the contraction analysis of the stochastic flow generated by the backward conditioned signal.

For example, for the particular type of continuous time models studied by W. Stannat in the papers [30, 31], one can verify the bounds (see Chapter 4, [35]):

$$\bar{\mathbb{E}}^Y_t|\tilde{X}_t(y) - \tilde{X}_t(x)| \le e^{-\kappa t}|x - y|,$$

with a constant $\kappa > 0$, expressed explicitly in terms of the model ingredients, and

$$\iint_{\mathbb{S}\times\mathbb{S}}|x - y|\,\bar{\pi}_t(dy)\,\bar{\pi}_t(dx) \le \text{const}.$$

This proves the uniform exponential stability of the filter with the rate $\kappa > 0$, thus establishing by probabilistic techniques the stability results obtained by W. Stannat using PDE techniques [30, 31]. It should be stressed that this is one of the few cases where time uniform exponential pathwise filter stability is known for (possibly non-ergodic) signals on noncompact domains.

3.2. **Infinite horizon conditioning.** The relation (3.1) still holds if conditioning on the observations is done on the *infinite* horizon, namely:

$$\mathbb{P}\big(X_m \in A\big|\mathcal{F}^X_{0,m-1}\vee\mathcal{F}^Y_{0,\infty}\big) = \mathbb{P}\big(X_m \in A\big|\sigma\{X_{m-1}\}\vee\mathcal{F}^Y_{0,\infty}\big) \qquad \mathbb{P}\text{-a.s.} \qquad (3.6)$$

This means that the measure $\mathbb{P}^Y_\infty$, induced on the signal path space by conditioning on $\mathcal{F}^Y_{0,\infty}$

$$\mathbb{P}^Y_\infty(\Gamma) := \mathbb{P}(X \in \Gamma|\mathcal{F}^Y_{0,\infty}), \qquad \Gamma \in \mathcal{B}(\mathbb{S}^\infty)$$

is Markov, i.e.

$$\mathbb{P}^Y_\infty(X_m \in A|\mathcal{F}^X_{0,m-1}) = \mathbb{P}^Y_\infty(X_m \in A|\sigma\{X_{m-1}\}) \qquad \mathbb{P}^Y_\infty\text{-a.s.}$$

As before we refer the coordinate process on $\mathbb{S}^\infty$ under $\mathbb{P}^Y_\infty$ as conditioned signal. Though the transition probability in this case can no longer be expressed in a convenient closed form such as (3.2), an advantage of such infinite horizon conditioning is that a *single* conditioned signal process is obtained, rather than a family of processes whose transition law changes when the horizon increases. More importantly, if the signal process is stationary then a form of stationarity (in the sense of Markov chains in random environments) is inherited by the conditioned signal on the infinite time horizon, which is not the case if one conditions on a finite time horizon. This stationarity property brings into relevance the ergodic theory of Markov chains in random environments, which bears much resemblance to the homogeneous case but is not applicable to general time inhomogeneous Markov chains (see [36]).

Unlike in the previous subsection, the filtering distributions π_n and $\bar{\pi}_n$ cannot be obtained as marginals of the conditional measure $\mathbb{P}^Y_\infty$, and we must therefore make a different connection with the filter stability problem. Somewhat surprisingly, the mysterious relation (1.11) can be restated in terms of the conditioned process, so that we can attempt to establish stability of the filter directly through (1.9). The connection with (1.11) is established using the following general fact (see Lemma 4.II.1 in [37]). Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be sub σ-algebras of $\mathcal{F}$ and let $\mathbb{P}_{\mathcal{G}_1}(\cdot)$ be a regular version of the conditional probability given $\mathcal{G}_1$. If $\mathcal{G}_2$ is countably generated, then

$$\mathbb{P}(\cdot|\mathcal{G}_1 \vee \mathcal{G}_2) = \mathbb{P}_{\mathcal{G}_1}(\cdot|\mathcal{G}_2) \qquad \mathbb{P}\text{-a.s.} \tag{3.7}$$

Since X_n takes values in a Polish space, $\mathcal{F}^X_{n,\infty}$ is countably generated. In the context of the filtering problem (3.7), this means that for any $A \in \mathcal{B}(\mathbb{S}^\infty)$

$$\bar{\mathbb{P}}(A|\mathcal{F}^Y_{0,\infty} \vee \mathcal{F}^X_{n,\infty}) = \bar{\mathbb{P}}^Y_\infty(A|\mathcal{F}^X_{n,\infty}) \qquad \bar{\mathbb{P}}\text{-a.s.}$$

Applying the martingale convergence theorem twice, we obtain

$$\begin{aligned}\bar{\mathbb{P}}\big(A\big|\textstyle\bigcap_{n\ge 0}\mathcal{F}^Y_{0,\infty} \vee \mathcal{F}^X_{n,\infty}\big) &= \lim_{n\to\infty} \bar{\mathbb{P}}(A|\mathcal{F}^Y_{0,\infty} \vee \mathcal{F}^X_{n,\infty}) = \\ &= \lim_{n\to\infty} \bar{\mathbb{P}}^Y_\infty(A|\mathcal{F}^X_{n,\infty}) = \bar{\mathbb{P}}^Y_\infty(A|\textstyle\bigcap_{n\ge 0}\mathcal{F}^X_{n,\infty}) \qquad \bar{\mathbb{P}}\text{-a.s.}\end{aligned} \tag{3.8}$$

One might be tempted to conclude from (3.7) that

$$\bar{\mathbb{P}}(A|\mathcal{F}^Y_{0,\infty} \vee \textstyle\bigcap_{n\ge 0}\mathcal{F}^X_{n,\infty}) \overset{?}{=} \bar{\mathbb{P}}^Y_\infty(A|\textstyle\bigcap_{n\ge 0}\mathcal{F}^X_{n,\infty}),$$

so that (1.11) would follow from (3.8). However, it is well known that the tail σ-algebra $\bigcap_{n\ge 0}\mathcal{F}^X_{n,\infty}$ is *not* countably generated, so this argument is not correct.[4]

Nonetheless this general approach can be rescued due to the following observation: it already suffices to show that the tail σ-algebra $\bigcap_{n\ge 0}\mathcal{F}^X_{n,\infty}$ is $\bar{\mathbb{P}}^Y_\infty$-trivial for $\bar{\mathbb{P}}$-a.e. observation path. Indeed, in this case

$$\bar{\mathbb{P}}^Y_\infty(A|\textstyle\bigcap_{n\ge 0}\mathcal{F}^X_{n,\infty}) = \bar{\mathbb{P}}^Y_\infty(A) = \bar{\mathbb{P}}(A|\mathcal{F}^Y_{0,\infty}) \qquad \bar{\mathbb{P}}\text{-a.s.},$$

so that we can conclude directly from (3.8) and (1.9) that the filter is stable. Of course, the problem remains to establish that $\bigcap_{n\ge 0}\mathcal{F}^X_{n,\infty}$ is indeed $\bar{\mathbb{P}}^Y_\infty$-trivial. This can be done in the framework of ergodic theory of Markov chains in random environments, which leads to the following result.

Theorem 3.2 (Corollary 5.5, [36]). *Suppose that the observation density $g(x,y)$ in* (1.3) *is strictly positive and that the signal is positive Harris recurrent and aperiodic. Then* (1.5) *holds for every $\nu, \bar{\nu}$.*

Note that the nondegeneracy assumption rules out the problem in Example 1.1.

[4] The countable generation requirement can be weakened somewhat, see Lemma 4.II.1 in [37]; in general, however, verifying the weaker requirement appears to be a very hard problem.

Remark 3.3. It is tempting to assume that the triviality of $\bigcap_{n\geq 0}\mathcal{F}^X_{n,\infty}$ $\bar{\mathbb{P}}$-a.s. already implies its triviality $\bar{\mathbb{P}}^Y_\infty$-a.s. regardless of any other ingredients of the model (e.g., the observation structure). After all, it is elementary that $\bar{\mathbb{P}}(A)=0$ or $\bar{\mathbb{P}}(A)=1$ implies $\bar{\mathbb{P}}(A|\mathcal{F}^Y_{0,\infty})=\bar{\mathbb{P}}(A)$ $\bar{\mathbb{P}}$-a.s. However, as $\bigcap_{n\geq 0}\mathcal{F}^X_{n,\infty}$ is not countably generated, it may be impossible to choose a regular conditional probability $\bar{\mathbb{P}}^Y_\infty(\cdot)$ such that $\bar{\mathbb{P}}^Y_\infty(A)=\bar{\mathbb{P}}(A)$ for all $A\in\bigcap_{n\geq 0}\mathcal{F}^X_{n,\infty}$ simultaneously on a set of full $\bar{\mathbb{P}}$-probability (i.e., as the number of sets A is uncountable one may not be able to eliminate the dependence of the $\bar{\mathbb{P}}$-null sets $\{\omega:\bar{\mathbb{P}}(A|\mathcal{F}^Y_{0,\infty})(\omega)\neq\bar{\mathbb{P}}(A)\}$ on A). Example 1.1 shows that this is a real problem in models which are by no means pathological.

4. Observability

In the previous sections, the stability of the filter was essentially inherited from the ergodic properties of the signal. On the other hand, it is evident that the observations may also have a stabilizing effect on the filter: a trivial example is the case where $Y_n=X_n$, so that $\pi_n=\bar{\pi}_n$, $n\geq 1$ for any $\nu,\bar{\nu}$ *regardless* of the properties of the signal. The aim of this section is to outline two approaches which provide a link between the quality of the observations and the stability of the filter.

4.1. **An information theoretic bound.** The first result of this kind appeared in [13] and is based on the connection with the information theoretic notion of *relative entropy.* Recall the definition of the relative entropy between two probability measures $\mathbb{P}$ and $\mathbb{Q}$:

$$D\big(\mathbb{P}\,||\,\mathbb{Q}\big)=\begin{cases}\displaystyle\int\log\frac{d\mathbb{P}}{d\mathbb{Q}}\,d\mathbb{P}, & \mathbb{P}\ll\mathbb{Q},\\ \infty, & \mathbb{P}\not\ll\mathbb{Q}.\end{cases}$$

The relative entropy is a pseudo-distance in the sense that it is nonnegative and vanishes if and only if the measures are identical. Note that as

$$\frac{d\mathbb{P}_{|\mathcal{G}}}{d\mathbb{Q}_{|\mathcal{G}}}=\mathbb{E}_{\mathbb{Q}}\left(\frac{d\mathbb{P}}{d\mathbb{Q}}\,\middle|\,\mathcal{G}\right),$$

where $\mathbb{P}_{|\mathcal{G}}$ and $\mathbb{Q}_{|\mathcal{G}}$ stand for the restrictions of $\mathbb{P}$ and $\mathbb{Q}$ to the σ-algebra $\mathcal{G}$, it follows easily from Jensen's inequality that

$$D\big(\mathbb{P}_{|\mathcal{G}}\,||\,\mathbb{Q}_{|\mathcal{G}}\big)\leq D\big(\mathbb{P}\,||\,\mathbb{Q}\big).\tag{4.1}$$

To develop the result of [13], we work in the following continuous time setting. Suppose that $X=(X_t)_{t\in\mathbb{R}_+}$ is a Markov process and

$$Y_t=\int_0^t h(X_s)\,ds+B_t,\tag{4.2}$$

where B is a p-dimensional Brownian motion independent of X and $h:\mathbb{S}\to\mathbb{R}^p$ is a function such that

$$\mathbb{E}\left(\int_0^T|h(X_s)|^2ds\right)<\infty,\qquad\bar{\mathbb{E}}\left(\int_0^T|h(X_s)|^2ds\right)<\infty\qquad\forall\,T>0.$$

Classical filtering theory tells us (see [24]) that $dY_t=\pi_t(h)\,dt+dV_t$, where V_t is the innovation Brownian motion under $\mathbb{P}$. Similarly $dY_t=\bar{\pi}_t(h)\,dt+d\bar{V}_t$, where $\bar{V}_t$

is the innovation Brownian motion under $\bar{\mathbb{P}}$. Hence, the Girsanov theorem shows that the laws of Y under $\mathbb{P}$ and $\bar{\mathbb{P}}$ are equivalent with

$$\begin{aligned}\frac{d\mathbb{P}_{|\mathcal{F}^Y_{0,t}}}{d\bar{\mathbb{P}}_{|\mathcal{F}^Y_{0,t}}} &= \exp\left\{\int_0^t \big(\pi_s(h)-\bar{\pi}_s(h)\big)\cdot dY_s - \frac{1}{2}\int_0^t \big(|\pi_s(h)|^2 - |\bar{\pi}_s(h)|^2\big)ds\right\} \\ &= \exp\left\{\int_0^t \big(\pi_s(h)-\bar{\pi}_s(h)\big)\cdot dV_s + \frac{1}{2}\int_0^t |\pi_s(h)-\bar{\pi}_s(h)|^2 ds\right\},\end{aligned}$$

and thus

$$D\Big(\mathbb{P}_{|\mathcal{F}^Y_{0,t}}\,||\,\bar{\mathbb{P}}_{|\mathcal{F}^Y_{0,t}}\Big) = \frac{1}{2}\,\mathbb{E}\left(\int_0^t |\pi_s(h)-\bar{\pi}_s(h)|^2 ds\right).$$

Since $\frac{d\mathbb{P}}{d\bar{\mathbb{P}}} = \frac{d\nu}{d\bar{\nu}}(X_0)$, the inequality (4.1) gives

Theorem 4.1 (Theorem 3.1, [13]). *Suppose that $D(\nu||\bar{\nu})<\infty$. Then*

$$\frac{1}{2}\,\mathbb{E}\left(\int_0^\infty |\pi_t(h)-\bar{\pi}_t(h)|^2\,ds\right) \le D(\nu||\bar{\nu}) < \infty. \tag{4.3}$$

Remark 4.2. There is in fact a deeper connection between the relative entropy and the filter stability problem in the general setting, which is developed in [13] also. Let us briefly sketch an alternative proof of this result. It is easily established using (1.7) and the Markov property of (X,Y) that

$$D(\pi_n||\bar{\pi}_n) = \mathbb{E}\left(\log\,\bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar{\nu}}(X_0)\Big|\mathcal{F}^Y_{0,\infty}\vee\mathcal{F}^X_{n,\infty}\Big)\,\bigg|\,\mathcal{F}^Y_{0,n}\right) - \log\,\bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar{\nu}}(X_0)\Big|\mathcal{F}^Y_{0,n}\Big).$$

Applying Jensen's inequality and the Bayes formula to this expression, it is not difficult to establish explicitly that

$$\mathbb{E}\big(D(\pi_n||\bar{\pi}_n)\big|\mathcal{F}^Y_{0,m}\big) \le D(\pi_m||\bar{\pi}_m) \quad \mathbb{P}\text{-a.s.} \qquad \text{for all } m\le n,$$

i.e., $D(\pi_n||\bar{\pi}_n)$ is an $\mathcal{F}^Y_{0,n}$-supermartingale under $\mathbb{P}$. Thus the relative entropy is a type of 'Lyapunov function' for the filter stability problem, and in particular the quantity $\mathbb{E}[D(\pi_n||\bar{\pi}_n)]$ is nonincreasing. Unfortunately, showing that the relative entropy actually decreases to zero as $n\to\infty$ appears to be not much easier than verifying filter stability in the total variation distance (see, e.g., Theorem 4.2 in [29]).

4.2. **Observability.** The information theoretic bound in theorem 4.1 establishes that the filtered estimate of the observation function h is stable in a weak sense virtually without any assumptions on the signal: in particular, neither compactness of the signal state space nor ergodicity of the signal was assumed! Note, however, that (4.3) does not guarantee the convergence of $\|\pi_t-\bar{\pi}_t\|$, and this may in fact very well fail. This raises an interesting possibility: perhaps there are other functions f for which $|\pi_t(f)-\bar{\pi}_t(f)|$ converges to zero as $t\to\infty$ regardless of whether the filter is stable? It turns out that this question has a nice affirmative answer which naturally leads to the notion of *observability* for nonlinear filtering models.

The basic idea is particularly transparent in discrete time for a model whose observations are defined in a slightly different manner from (1.3): we assume that Y_n is a noisy observation of X_{n-1}, rather than of X_n. In other words, the signal is observed with one time step delay. To be precise, Y still forms a sequence of independent random variables when conditioned on X, where (cf. (1.3))

$$\mathbb{P}(Y_n\in A|\mathcal{F}^X_{0,\infty}) = \int_A g(X_{n-1},y)\,\psi(dy). \tag{4.4}$$

The filtering equation in this case is the recursion (cf. (1.4))

$$\pi_n(dx) = \frac{\int_{\mathbb{S}} \Lambda(u,dx)\, g(u,Y_n)\, \pi_{n-1}(du)}{\int_{\mathbb{S}} g(u,Y_n)\, \pi_{n-1}(du)}, \qquad n \ge 1, \tag{4.5}$$

whose solution is denoted by π_n when the equation is initialized by ν and by $\bar\pi_n$ when it is started from $\bar\nu$.

Remark 4.3. Though the following results are more naturally formulated in the modified setting (4.4), some additional work allows one to consider the setting of (1.3) as well; see, e.g., [34]. Moreover, as will be discussed briefly below, these ideas can also be developed in the continuous time setting where the difference between (4.4) and (1.3) disappears. For simplicity, however, *we will operate in modified setting* (4.4) *throughout the remainder of this section.*

To develop stability results in this setting, we first make a brief detour. Instead of considering the filters π_n and $\bar\pi_n$, let us turn our attention for the moment to the one step predictors of the observation process:

$$\eta_{n|n-1}(f) := \mathbb{E}\big(f(Y_n)|\mathcal{F}^Y_{0,n-1}\big) = \int_{\mathbb{R}^p}\int_{\mathbb{S}} f(y)\, g(u,y)\, \pi_{n-1}(du)\, \psi(dy),$$

$$\bar\eta_{n|n-1}(f) := \bar{\mathbb{E}}\big(f(Y_n)|\mathcal{F}^Y_{0,n-1}\big) = \int_{\mathbb{R}^p}\int_{\mathbb{S}} f(y)\, g(u,y)\, \bar\pi_{n-1}(du)\, \psi(dy).$$

It turns out that these predictors are always stable in the following sense.

Proposition 4.4 (Theorem 2.1, [9])**.** *If $\nu \ll \bar\nu$, then for any bounded function f*

$$\lim_{n\to\infty} \mathbb{E}|\eta_{n|n-1}(f) - \bar\eta_{n|n-1}(f)| = 0. \tag{4.6}$$

The proof uses the ideas similar to those presented in the Introduction. By the Bayes formula

$$\eta_{n|n-1}(f) = \mathbb{E}\left(f(Y_n)\big|\mathcal{F}^Y_{0,n-1}\right) = \frac{\bar{\mathbb{E}}\left(\frac{d\nu}{d\bar\nu}(X_0) f(Y_n)\big|\mathcal{F}^Y_{0,n-1}\right)}{\bar{\mathbb{E}}\left(\frac{d\nu}{d\bar\nu}(X_0)\big|\mathcal{F}^Y_{0,n-1}\right)} \qquad \mathbb{P}\text{-a.s.}$$

and hence under $\mathbb{P}$

$$\eta_{n|n-1}(f) - \bar\eta_{n|n-1}(f) = \frac{\bar{\mathbb{E}}\left(\frac{d\nu}{d\bar\nu}(X_0) f(Y_n)\big|\mathcal{F}^Y_{0,n-1}\right) - \bar{\mathbb{E}}\left(f(Y_n)\big|\mathcal{F}^Y_{0,n-1}\right)\bar{\mathbb{E}}\left(\frac{d\nu}{d\bar\nu}(X_0)\big|\mathcal{F}^Y_{0,n-1}\right)}{\bar{\mathbb{E}}\left(\frac{d\nu}{d\bar\nu}(X_0)\big|\mathcal{F}^Y_{0,n-1}\right)}.$$

Simple manipulations with conditional expectations give

$$\begin{aligned}
\mathbb{E}|\eta_{n|n-1}(f) - \bar\eta_{n|n-1}(f)| &= \bar{\mathbb{E}}\left[\bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar\nu}(X_0)\Big|\mathcal{F}^Y_{0,n-1}\Big)\, |\eta_{n|n-1}(f) - \bar\eta_{n|n-1}(f)|\right] = \\
\bar{\mathbb{E}}\left[\bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar\nu}(X_0) f(Y_n)\Big|\mathcal{F}^Y_{0,n-1}\Big) - \bar{\mathbb{E}}\left(f(Y_n)\big|\mathcal{F}^Y_{0,n-1}\right)\bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar\nu}(X_0)\Big|\mathcal{F}^Y_{0,n-1}\Big)\right] &= \\
\bar{\mathbb{E}}\left(\left\{\bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar\nu}(X_0)\Big|\mathcal{F}^Y_{0,n}\Big) - \bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar\nu}(X_0)\Big|\mathcal{F}^Y_{0,n-1}\Big)\right\} f(Y_n)\right)&.
\end{aligned}$$

But as f is bounded (by a constant C, say), we find that

$$\mathbb{E}|\eta_{n|n-1}(f) - \bar\eta_{n|n-1}(f)| \le C\,\bar{\mathbb{E}}\left|\bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar\nu}(X_0)\Big|\mathcal{F}^Y_{0,n}\Big) - \bar{\mathbb{E}}\Big(\frac{d\nu}{d\bar\nu}(X_0)\Big|\mathcal{F}^Y_{0,n-1}\Big)\right|,$$

which converges to zero as $n\to\infty$ by the martingale convergence theorem.

Proposition 4.4 shows that the one step predictive estimates of the observation process are stable, but we are ultimately interested in the stability of the filter. To make the connection with the latter problem, let us now consider (in analogy with (4.2)) the *additive* noise observation scenario, i.e., we assume that

$$Y_n = h(X_{n-1}) + \xi_n, \qquad n \geq 1,$$

where $\xi = (\xi_n)_{n\geq 1}$ is an i.i.d. sequence of $\mathbb{R}^p$-valued random variables independent of X and $h : \mathbb{S} \to \mathbb{R}^p$ is a given observation function. In this case Proposition 4.4 can be used to prove the following result:

Theorem 4.5 (Variant of Proposition 3.3, [9])**.** *Suppose that*

(a_1) *h is bounded.*

(a_2) *$|\mathbb{E}e^{ik\cdot\xi}| > 0$ for all $k \in \mathbb{R}^p$.*

Then for any continuous function f and $\nu \ll \bar{\nu}$

$$\lim_{n\to\infty} \mathbb{E}|\pi_n(f \circ h) - \bar{\pi}_n(f \circ h)| = 0. \tag{4.7}$$

Indeed, in this case

$$\eta_{n|n-1}\big(e^{ik\cdot}\big) = \mathbb{E}\big(e^{ik\cdot Y_n}|\mathcal{F}^Y_{0,n-1}\big) = \pi_{n-1}\big(e^{ik\cdot h(\cdot)}\big)\,\mathbb{E}e^{ik\cdot\xi_1},$$
$$\bar{\eta}_{n|n-1}\big(e^{ik\cdot}\big) = \bar{\mathbb{E}}\big(e^{ik\cdot Y_n}|\mathcal{F}^Y_{0,n-1}\big) = \bar{\pi}_{n-1}\big(e^{ik\cdot h(\cdot)}\big)\,\mathbb{E}e^{ik\cdot\xi_1},$$

and as $\nu \ll \bar{\nu}$, we obtain by Proposition 4.4 and assumption (a_2)

$$\mathbb{E}\left|\pi_{n-1}\big(e^{ik\cdot h(\cdot)}\big) - \bar{\pi}_{n-1}\big(e^{ik\cdot h(\cdot)}\big)\right| = \frac{\mathbb{E}\left|\eta_{n|n-1}\big(e^{ik\cdot}\big) - \bar{\eta}_{n|n-1}\big(e^{ik\cdot}\big)\right|}{|\mathbb{E}e^{ik\cdot\xi_1}|} \xrightarrow{n\to\infty} 0.$$

We therefore find that for any finite order trigonometric polynomial T

$$\lim_{n\to\infty} \mathbb{E}|\pi_n(T \circ h) - \bar{\pi}_n(T \circ h)| = 0. \tag{4.8}$$

Now note that as a consequence of the Weierstrass approximation theorem, any continuous function can be approximated uniformly on compact sets by trigonometric polynomials. As h is bounded it takes values in a compact set. Therefore, given a continuous function f, there is a sequence of trigonometric polynomials T_ℓ such that $\|f \circ h - T_\ell \circ h\|_\infty \leq \ell^{-1}$. But then

$$\mathbb{E}|\pi_n(f \circ h) - \bar{\pi}_n(f \circ h)| \leq \mathbb{E}|\pi_n(T_\ell \circ h) - \bar{\pi}_n(T_\ell \circ h)| + 2\ell^{-1}$$

for all n, so $\limsup_{n\to\infty} \mathbb{E}|\pi_n(f \circ h) - \bar{\pi}_n(f \circ h)| \leq 2\ell^{-1}$. But ℓ was arbitrary, so letting $\ell \to \infty$ completes the proof.

It follows immediately from Theorem 4.5 that the stability of $\pi_n(h)$

$$\lim_{n\to\infty} \mathbb{E}|\pi_n(h) - \bar{\pi}_n(h)| = 0 \tag{4.9}$$

is recovered by choosing $f(x) = x$. This resembles the result (4.3). Moreover, as in the case of Theorem 4.1, ergodicity of the signal process is not assumed. However, there are essential differences between the two results. On the one hand, (4.3) places minimal assumptions on the observations, whereas (4.9) relies on the restrictive assumption that the observation function h is bounded. On the other hand, (4.3) only provides stability of the observation function h itself, while (4.7) provides stability also for functions of h. The latter class of functions can be quite large: for example, when h is invertible as in the following corollary, we can even conclude stability of the filter in a weak (as opposed to total variation) sense.

Corollary 4.6. *Let $\mathbb{S} \subset \mathbb{R}^p$ and let $h : \mathbb{S} \to \mathbb{R}^p$ be bounded. Suppose there is a continuous function $h^{-1} : \mathbb{R}^p \to \mathbb{R}^p$ such that $h^{-1}(h(x)) = x$ for all $x \in \mathbb{S}$. Then*

$$\lim_{n\to\infty} \mathbb{E}|\pi_n(g) - \bar{\pi}_n(g)| = 0$$

for every continuous function g and $\nu \ll \bar{\nu}$, provided $|\mathbb{E}e^{ik\cdot\xi}| > 0$ for all $k \in \mathbb{R}^p$.

The proof is immediate from Theorem 4.5.

Remark 4.7. The assumptions of the previous corollary require that $\mathbb{S}$ is a bounded subset of $\mathbb{R}^p$; indeed, as $h(\mathbb{S})$ is bounded and h^{-1} is continuous, $\mathbb{S} = h^{-1}(h(\mathbb{S}))$ must be bounded. The result is therefore the most natural when $\mathbb{S}$ is compact, in which case the boundedness of h is not restrictive. When $\mathbb{S}$ is not compact, it may still be the case that the signal is outside a compact set with uniformly small probability, i.e., that the sequence $(X_k)_{k\geq 0}$ is tight or uniformly integrable. In this case, it is straightforward to localize the proof by truncating to a compact set, see [9], and one can relax the boundedness of h. Though this is perhaps not surprising, it should be noted that localization is often not so straightfoward in other methods. For example, we showed that the assumption of Theorem 2.1 is most natural when $\mathbb{S}$ is compact; however, the localization of that result is highly nontrivial [20].

When h is not invertible, Theorem 4.5 yields only 'partial' stability, i.e., stability of the estimates of particular functions. However, with a little more work we can establish the stability of a much larger class of functions than we have investigated so far, which opens the possibility of proving stability of the filter (in the spirit of the previous corollary) even when h is not invertible. To this end, let us begin by noting that the arguments for (4.6) apply to predictors of a more general form (for an even more general statement, see the classic paper [4]):

Proposition 4.8. *For an integer $m \geq 1$, let $f_1, \ldots, f_m$ be continuous bounded functions and $k_1, \ldots, k_m$ distinct positive integers. Then if $\nu \ll \bar{\nu}$*

$$\lim_{n\to\infty} \mathbb{E}\big|\mathbb{E}\big(f_1(Y_{n+k_1}) \cdot\cdot f_m(Y_{n+k_m})\big|\mathcal{F}^Y_{0,n}\big) - \bar{\mathbb{E}}\big(f_1(Y_{n+k_1}) \cdot\cdot f_m(Y_{n+k_m})\big|\mathcal{F}^Y_{0,n}\big)\big| = 0. \tag{4.10}$$

Now note that by time homogeneity and the Markov property of (X, Y)

$$\mathbb{E}\big(f_1(Y_{n+k_1}) \cdots f_m(Y_{n+k_m})\big|\mathcal{F}^Y_{0,n}\big) = \pi_n(f)$$

with

$$f(x) := \mathbb{E}\big(f_1(Y_{k_1}) \cdots f_m(Y_{k_m})\big|X_0 = x\big), \tag{4.11}$$

and thus (4.10) states that the filtered estimates of such f are always stable. As the number of times m and the functions $f_1, \ldots, f_m$ are arbitrary, this suggests that the class of functions with stable estimates can be quite large. We are interested in characterizing this class of functions in terms of the filtering model.

Such a characterization is indeed possible and can be formulated in terms of *observability* ([32]). Consider the following equivalence relation for probability measures on $\mathbb{S}$: we say that two probability measures ν_1, ν_2 are equivalent $\nu_1 \backsim \nu_2$ if they induce the same law of the observation process, i.e.,

$$\nu_1 \backsim \nu_2 \qquad \text{iff} \qquad \mathbb{P}^{\nu_1}|_{\mathcal{F}^Y_{0,\infty}} = \mathbb{P}^{\nu_2}|_{\mathcal{F}^Y_{0,\infty}},$$

where $\mathbb{P}^{\nu_1}$ and $\mathbb{P}^{\nu_2}$ denote the law of (X, Y) when $X_0 \sim \nu_1$ and $X_0 \sim \nu_2$, respectively. As probability measures on $\mathcal{B}((\mathbb{R}^p)^\infty)$ are determined by their finite dimensional distributions, $\nu_1 \backsim \nu_2$ if and only if for any $m \geq 1$, any bounded and continuous functions $f_1, \ldots, f_m$ and time indices $k_1, \ldots, k_m$

$$\int_{\mathbb{S}} \mathbb{E}\big(f_1(Y_{k_1})\cdot\cdot f_m(Y_{k_m})\big|X_0 = x\big)\,\nu_1(dx) = \int_{\mathbb{S}} \mathbb{E}\big(f_1(Y_{k_1})\cdot\cdot f_m(Y_{k_m})\big|X_0 = x\big)\,\nu_2(dx).$$

In other words, $\nu_1 \backsim \nu_2$ whenever the signed measure $\nu_1 - \nu_2$ is orthogonal to the linear subspace $\mathcal{O}^o$ spanned by the functions of the form (4.11).

Let us now suppose that the Markov process (X, Y) is Feller, so that all functions of the form (4.11) are continuous. Moreover, let us suppose that the signal state space $\mathbb{S}$ is compact. Then an elementary functional analytic argument (Proposition 3.3, [32]) shows that $\mathcal{O}^o \subseteq \mathcal{C}_b(\mathbb{S})$ is dense in the subspace

$$\mathcal{O} = \left\{ f \in \mathcal{C}_b(\mathbb{S}) : \int f\,d\nu_1 = \int f\,d\nu_2 \text{ whenever } \nu_1 \backsim \nu_2 \right\}$$

in the topology of uniform convergence ($\mathcal{C}_b(\mathbb{S})$ is the space of bounded continuous functions on $\mathbb{S}$). In other words, for any $f \in \mathcal{O}$ there is a sequence of functions f_n of the form (4.11) such that $f_n \to f$ uniformly (this argument replaces the application of the Weierstrass theorem in the proof of Theorem 4.5). Since the filtered estimates of functions of the form (4.11) are stable, we obtain

Theorem 4.9 (Variant of Theorem 4.4, [32])**.** *Assume that $\mathbb{S}$ is compact, (X, Y) is Feller, and $\nu \ll \bar{\nu}$. Then for any $f \in \mathcal{O}$*

$$\lim_{n\to\infty} \mathbb{E}|\pi_n(f) - \bar{\pi}_n(f)| = 0. \tag{4.12}$$

By definition, the space $\mathcal{O}$ consists of those functions whose expectation is uniquely determined by the law of the observations. In this sense it is natural to call $\mathcal{O}$ the observable space of the filtering model and $f \in \mathcal{O}$ observable functions. If $\mathcal{O} = \mathcal{C}_b(\mathbb{S})$, the model is referred to as *(fully) observable*, and

Corollary 4.10. *Assume that $\mathbb{S}$ is compact and (X, Y) is Feller. Then the model is (fully) observable if and only if $\nu_1 \backsim \nu_2$ implies $\nu_1 = \nu_2$, i.e., if the law of the observations uniquely determines the initial law of the signal. When this is the case*

$$\lim_{n\to\infty} \mathbb{E}|\pi_n(f) - \bar{\pi}_n(f)| = 0 \quad \textit{for all } f \in \mathcal{C}_b,\ \nu \ll \bar{\nu}.$$

This is a generalization of Corollary 4.6 in the present setting. Note that observability does not require h to be invertible; on the other hand, it is not difficult to establish that a sufficient condition for observability is that h is invertible and $|\mathbb{E}e^{ik\cdot\xi}| > 0$ for all $k \in \mathbb{R}^p$, reproducing essentially the result of Corollary 4.6.

Example 4.11. Let X be a Markov chain on $\mathbb{S} = \{1, 2, 3, 4\}$ with transition matrix

$$\Lambda = \begin{pmatrix} 1/2+\varepsilon & 1/2-\varepsilon & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 1/2 & 0 & 0 & 1/2 \end{pmatrix},$$

where $0 < \varepsilon \leq 1/2$, and let $Y_n = \mathbf{1}_{\{X_{n-1}\in\{1,3\}\}}$, $n \geq 1$. This differs from the model of Example 1.1 in that we have perturbed one of the transition probabilities, and that we have introduced one time step delay in the observation model in keeping with the setting of this section. It is easily verified, however, that when $\varepsilon = 0$ the corresponding filter is unstable exactly as in Example 1.1. In contrast, we now show that when $\varepsilon \neq 0$ the model is observable and hence the filter is stable.

To prove observability, note that for $k \geq 1$

$$\begin{pmatrix} \mathbb{E}(Y_k|X_0=1) \\ \mathbb{E}(Y_k|X_0=2) \\ \mathbb{E}(Y_k|X_0=3) \\ \mathbb{E}(Y_k|X_0=4) \end{pmatrix} = \Lambda^{k-1} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} := f_k.$$

Computing explicitly, we find that

$$f_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad f_2 = \frac{1}{2}\begin{pmatrix} 2\varepsilon+1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \qquad f_3 = \frac{1}{2}\begin{pmatrix} 2\varepsilon^2+\varepsilon+1 \\ 1 \\ 1 \\ \varepsilon+1 \end{pmatrix}.$$

But the vectors f_1, f_2, f_3 and $f_0 = (1\ 1\ 1\ 1)^\top$ span $\mathbb{R}^4$, provided $\varepsilon \neq 0$. Therefore every function $f : \mathbb{S} \to \mathbb{R}$ can be written as a function of the form $f(x) = \mathbf{E}(\alpha_0 + \alpha_1 Y_1 + \alpha_2 Y_2 + \alpha_3 Y_3 | X_0 = x)$ for some $\alpha_0, \ldots, \alpha_3 \in \mathbb{R}$, so that observability, and consequently stability of the filter, follow.

The same theory works out in continuous time, where the nuance of the observation delay disappears ((4.4) vs. (1.3)). For the additive white noise model of Theorem 4.1, one can show that $\nu_1 \backsim \nu_2$ if and only if $(h(X_t))_{t\geq 0}$ has the same law under $\mathbb{P}^{\nu_1}$ and $\mathbb{P}^{\nu_2}$ (see Proposition 5.2 in [32]). Therefore, if $\nu_1 \backsim \nu_2$, then in particular $\mathbb{E}^{\nu_1}(f(h(X_0))) = \mathbb{E}^{\nu_2}(f(h(X_0)))$ for every measurable function f so that

$$\nu_1 \backsim \nu_2 \quad \Longrightarrow \quad \int_{\mathbb{S}} f \circ h \, d\nu_1 = \int_{\mathbb{S}} f \circ h \, d\nu_2.$$

Thus evidently $f \circ h \in \mathcal{O}$ for any f such that $f \circ h \in \mathcal{C}_b(\mathbb{S})$ (Lemma 5.6, [32]). The continuous time counterparts of Theorem 4.5 and of Corollary 4.6 follow directly.

In the general case where h is not invertible, characterizing the observable space $\mathcal{O}$ in terms of the model ingredients remains a nontrivial task. However, in the special case of finite state space signals this can be done explicitly (Section 6, [32]). Consider the model from subsection 2.1 and let $\mathbb{H} = h(\mathbb{S}) := \{h_1, ..., h_r\}$, $r \leq d$ be the set of the distinct observations values. Define $d \times d$ diagonal matrices H_{h_k}, $k = 1, ..., r$ such that $H_{h_k}(i,j) = 1$ whenever $i = j$ and $h(a_i) = h_k$ and such that the remaining entries are zero. Then again identifying functions on $\mathbb{S}$ with vectors in $\mathbb{R}^d$, writing $\mathbf{1}$ for the $d \times 1$ vector with unit entries and, as before, denoting the transition matrix of the chain by Λ, one can easily prove the following result along the same lines as the computation in Example 4.11.

Proposition 4.12 (Lemma 6.4 in [32])**.**

$$\mathcal{O} = \operatorname{span}\left\{ H_{n_0} \Lambda H_{n_1} \Lambda \cdots \Lambda H_{n_m} \mathbf{1} : m \geq 0, n_i \in \mathbb{H} \right\}.$$

In particular, if $\dim(\mathcal{O}) = d$, the filter is stable in the sense of (1.5).

In this finite state setting, one can in fact go one step further and give the complete characterization of filter stability. For this purpose the following notion of *detectability* is introduced: the model is called detectable if $\lim_{t\to\infty} e^{\Lambda^\top t}\mu = 0$ whenever $\mu \perp \mathcal{O}$ (thus every observable model is detectable, but not vice versa). Let us note that detectability, like observability, can be verified algebraically in terms of the model parameters Λ and h.

Theorem 4.13 (Theorem 6.12 in [32])**.** *Assume that the observations are nondegenerate $\sigma > 0$. Then the Shiryaev-Wonham filter* (2.14) *is stable in the sense of* (1.5) *whenever $\nu \ll \bar{\nu}$ if and only if the model is detectable.*

The proof of this result is obtained by combining Theorem 2.5 and the continuous time counterpart of Theorem 4.9. This suggests that at least in the finite state setting, the two main structural assumptions that we have imposed in this article—ergodicity of the signal process and observability—are indeed the fundamental mechanisms that conspire to give stability of the filter.

4.3. **Uniform observability.** A drawback of the observability results above is that they rely on compactness of the state space (or uniform integrability of the signal, as in Remark 4.7, which allows truncation to a compact set). This rules out the interesting possibility that the filter may be stable even in models where the signal itself is unstable (i.e., when the signal diverges to infinity), which is known to hold, e.g., in the linear Gaussian case when the Kalman filter is observable [27]. It turns out the the approach of the previous section can be extended to cover also the unstable case, though the analysis is more subtle in this setting.

The reason that compactness was required above is that both Theorem 4.5 and Theorem 4.9 are proved by showing that a class of continuous functions, obtained from the predictor, is dense in the *uniform* topology. Uniform approximation of continuous functions is a natural problem for functions on a compact state space, and can be tackled using elementary functional analytic arguments. However, when the state space is not compact one obtains approximation uniformly on compact sets, which is insufficient for our purposes when the signal is unstable. Nonetheless a more refined argument, using a uniform approximation property of convolution operators, allows one to resolve this problem in the case of additive observations [33, 34]. This gives, for example, the following counterpart of Corollary 4.6. Here

$$\|\nu_1 - \nu_2\|_{\mathrm{BL}} := \sup_{f\in\mathrm{BL}} \left| \int f\, d\nu_1 - \int f\, d\nu_2 \right|,$$

where BL denotes the class of functions f such that $\|f\|_\infty \le 1$ and $|f(x) - f(z)| \le d(x,z)$ for all x, z (the unit ball in the space of bounded Lipschitz functions).

Proposition 4.14 (Proposition 3.11 in [33])**.** *Suppose that* $Y_n = h(X_{n-1}) + \xi_n$ *with*

(a_1) $h : \mathbb{S} \to \mathbb{R}^p$ *is invertible.*
(a_2) h^{-1} *is uniformly continuous.*
(a_3) $|\mathbb{E}e^{ik\cdot\xi}| > 0$ *for all* $k \in \mathbb{R}^p$.

Then

$$\lim_{n\to\infty} \mathbb{E}\|\pi_n - \bar{\pi}_n\|_{\mathrm{BL}} = 0$$

whenever $\nu \ll \bar{\nu}$.

A remarkable property of this result is that only assumptions on the observation structure are made, while the signal transition kernel can be completely arbitrary. This is opposite in spirit to the conditions given in section 2: there it was shown that the filter is stable regardless of the observation structure if the signal is sufficiently mixing, while we see here that the filter is stable regardless of the signal structure of the observations are sufficiently informative.

A noncompact counterpart to Corollary 4.10 can be obtained if the notion of observability is replaced by the stronger notion of *uniform observability.* Recall that the filtering model is called observable if

$$\mathbb{P}^{\nu_1}|_{\mathcal{F}^Y_{0,\infty}} = \mathbb{P}^{\nu_2}|_{\mathcal{F}^Y_{0,\infty}} \qquad \text{implies} \qquad \nu_1 = \nu_2.$$

Uniform observability is, in a sense, a quantitative counterpart of observability: the model is said to be uniformly observable if for every $\varepsilon > 0$, there exists a $\delta > 0$ (which depends only on ε) such that

$$\|\mathbb{P}^{\nu_1}|_{\mathcal{F}^Y_{0,\infty}} - \mathbb{P}^{\nu_2}|_{\mathcal{F}^Y_{0,\infty}}\| < \delta \quad \text{implies} \quad \|\nu_1 - \nu_2\|_{\mathrm{BL}} < \varepsilon.$$

Using this definition, one obtains the following counterpart of Corollary 4.10.

Theorem 4.15 (Variant of Theorem 3.3, [33])**.** *Suppose that the filtering model is uniformly observable. Then*

$$\lim_{n\to\infty} \mathbb{E}\|\pi_n - \bar{\pi}_n\|_{\mathrm{BL}} = 0 \quad \textit{for all } \nu \ll \bar{\nu}.$$

It can be shown that a known result about the stability of the Kalman filter is a special case of this general result, while in the compact case it turns out that observability and uniform observability are equivalent [33]. In general, however, uniform observability remains difficult to verify for specific filtering models.

Finally, we remark that the results in this section do not provide rates of convergence, while many filter stability results give rise to exponential rates. The key to the stability proofs in this section is the martingale convergence theorem, which does not guarantee a rate of convergence. As the following example shows, exponential stability cannot always be expected to hold without further assumptions.

Example 4.16. For real-valued signal and observations, consider the model $Y_n = X_{n-1} + \xi_n$, where ξ_n are i.i.d. $N(0,1)$ and $X_n = X_0$ for all n. Let $\nu = N(\alpha, \sigma^2)$ and $\bar{\nu} = N(\beta, \sigma^2)$ for some $\alpha, \beta, \sigma \in \mathbb{R}$ (so $\nu \ll \bar{\nu}$). Linear filtering theory shows that π_n is a random Gaussian measure with mean Z_n and variance V_n given by

$$Z_n = \frac{\alpha}{1+\sigma^2 n} + \frac{\sigma^2 n}{1+\sigma^2 n} \cdot \frac{1}{n}\sum_{\ell=1}^{n} Y_\ell, \qquad V_n = \frac{\sigma^2}{1+\sigma^2 n},$$

and similarly for $\bar{\pi}_n, \bar{Z}_n, \bar{V}_n$ where α is replaced by β. Evidently the conditional mean of the filter is stable with rate $\Omega(n^{-1})$, which is not exponential. (The conditional mean is an unbounded function of the signal; however, it is not difficult to show that $\mathbb{E}\|\pi_n - \bar{\pi}_n\|_{\mathrm{BL}} = \Omega(n^{-1})$ also, see Remark 2.8 in [34].)

5. Necessary conditions for stability

Almost all the above results (the exception being the approach of section 3.1) appeal directly to the absolute continuity assumption (A), either through (1.9) or through Proposition 4.4. Indeed, in a sense this assumption lies at the heart of the 'intrinsic' approach to filter stability, as it allows to relate the conditional expectations (1.1) for different initial measures $\nu, \bar{\nu}$ using the Bayes formula.

Assumption (A) was introduced without fanfare in the Introduction. However, the assumption is not as innocent as it may seem: for example, it implies that $\bar{\nu}$ has an atom at every point ν does, i.e., a suitable choice of $\bar{\nu}$ requires some information about the possibly unknown true distribution ν. Moreover, many filter stability results obtained by other methods hold for arbitrary $\nu, \bar{\nu}$. One might therefore wonder whether the assumption (A) is a restriction of the intrinsic method, or whether it has a deeper relevance. In this section, we will outline how the assumption (A) can be weakened in the context of the intrinsic approach, and we will show that the weakened assumption is in fact a necessary condition for stability in the total variation distance. This indicates that some form of absolute continuity is a fundamental ingredient of the filter stability problem. Though this discussion is not of direct practical interest, it sheds some light on the (often hidden) assumptions that are common to all methods of proving filter stability.

5.1. **Well posedness.** The first question one needs to confront in weakening assumption (A) is whether the recursion (1.4) is even well posed. The problem, which was glossed over in the Introduction, is that the denominator in (1.4) may be zero for some observation sequences. This is typically resolved by noting that the denominator of (1.4) is nonzero for $\mathbb{P}$-a.e. observation sequence (this holds by construction as the recursion (1.4) is obtained from the Bayes formula; see, e.g., Remark 3.1.5 in [7]). However, this need no longer hold if the initial distribution ν is replaced by $\bar{\nu}$, as the following example shows.

Example 5.1. Consider the signal X_n on $\mathbb{S} = \{2, 1, -1\}$ with the transition matrix

$$\Lambda = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \\ 1/2 & 0 & 1/2 \end{pmatrix},$$

and initial distribution $\nu(\{1\}) = 1$. Suppose that the observation sequence is

$$Y_n = X_n U_n,$$

where $U = (U_n)_{n\ge1}$ are i.i.d. random variables with uniform distribution over $[0, 1]$. If the filter is started with $\bar{\nu}(\{2\}) = 1$, the formula (1.4) yields

$$\begin{aligned} \bar{\pi}_1(\{2\}) &= \frac{\frac{1}{2}\mathbf{1}_{\{Y_1\in[0,2]\}}}{\frac{1}{2}\mathbf{1}_{\{Y_1\in[0,2]\}} + \mathbf{1}_{\{Y_1\in[0,1]\}}} \\ \bar{\pi}_1(\{1\}) &= \frac{\mathbf{1}_{\{Y_1\in[0,1]\}}}{\frac{1}{2}\mathbf{1}_{\{Y_1\in[0,2]\}} + \mathbf{1}_{\{Y_1\in[0,1]\}}} \\ \bar{\pi}_1(\{-1\}) &= 0. \end{aligned}$$

But $Y_1 = X_1 U_1 = -U_1 < 0$ a.s. and hence the right hand side is ill-posed (0/0).

When is the filtering recursion $\mathbb{P}$-a.s. well posed? We can give a general answer to this question. Note that, by definition, the conditional probability $\mathbb{P}(X_n \in \cdot|\mathcal{F}^Y_{0,n})$ is defined uniquely up to $\mathbb{P}_{|\mathcal{F}^Y_{0,n}}$-a.s. equivalence. Therefore π_n is well defined for $\mathbb{P}_{|\mathcal{F}^Y_{0,n}}$-a.e. observation path, while $\bar{\pi}_n$ is well defined for $\bar{\mathbb{P}}_{|\mathcal{F}^Y_{0,n}}$-a.e. observation path. In order for $\bar{\pi}_n$, $n \ge 0$ to be well defined for $\mathbb{P}$-a.e. observation path, we must require

$$\mathbb{P}_{|\mathcal{F}^Y_{0,n}} \ll \bar{\mathbb{P}}_{|\mathcal{F}^Y_{0,n}}, \qquad n \ge 0. \tag{B}$$

This is obviously satisfied under assumption (A), a fact that we have implicitly used throughout the paper. However, assumption (A) is not necessary for the filtering recursion to be well posed; for example, (B) holds regardless of the choice of $\bar{\nu}$ in the nondegenerate case where $g(x, y) > 0$ for all x, y. Indeed, it is immediately evident from (1.4) that the filtering recursion is always well posed in this case.

5.2. **Absolute continuity.** In the previous sections, we have proved various sufficient conditions for filter stability under assumption (A). However, in most cases, it is enough to impose the weaker assumption

$$\mathbb{P}_{|\mathcal{F}^Y_{0,\infty}} \ll \bar{\mathbb{P}}_{|\mathcal{F}^Y_{0,\infty}}. \tag{C}$$

Let us briefly outline one way to do this. Suppose that we have proved that (1.5) holds under assumption (A). Now suppose that $\nu, \bar{\nu}$ are such that only (C) holds. Defining $\tilde{\nu} := (\nu + \bar{\nu})/2$, we have $\nu \ll \tilde{\nu}$ and $\bar{\nu} \ll \tilde{\nu}$. Therefore, using (A), we have

$$\lim_{n\to\infty} \mathbb{E}\|\pi_n - \tilde{\pi}_n\| = \lim_{n\to\infty} \bar{\mathbb{E}}\|\bar{\pi}_n - \tilde{\pi}_n\| = 0,$$

where $\tilde{\pi}_n$ is defined in the obvious fashion. In particular, $\|\bar{\pi}_n - \tilde{\pi}_n\| \to 0$ in $\bar{\mathbb{P}}$-probability. But then (C) implies that $\|\bar{\pi}_n - \tilde{\pi}_n\| \to 0$ in $\mathbb{P}$-probability, and by dominated convergence $\mathbb{E}\|\bar{\pi}_n - \tilde{\pi}_n\| \to 0$. Therefore

$$\lim_{n\to\infty} \mathbb{E}\|\pi_n - \bar{\pi}_n\| \le \lim_{n\to\infty} \mathbb{E}\|\pi_n - \tilde{\pi}_n\| + \lim_{n\to\infty} \mathbb{E}\|\bar{\pi}_n - \tilde{\pi}_n\| = 0$$

by the triangle inequality, and we find that indeed the result is automatically extended to the weaker setting of assumption (C). Similar considerations apply to the weaker notions of convergence considered in section 4.

Assumption (C), however, is still stronger than the minimal assumption (B) needed for the filtering recursion (and hence the filter stability problem) to be well posed. As we will argue in the next section, assumption (C) cannot be weakened in general any further if the filter is to be stable, at least if we are interested in proving stability in the total variation distance. Indeed, we will prove that *assumption* (C) *is necessary for filter stability in total variation.* Evidently absolute continuity on the infinite time horizon is, in a sense, fundamental to the filter stability problem. Though assumption (C) is not commonly stated in the literature on filter stability, it is typically an implicit consequence of the model assumptions. This insight sheds some light on the minimal requirements needed by any method for proving stability. It also reassures us that little is lost by imposing the convenient assumption (A), which was not entirely obvious at the outset.

Let us note that assumption (C) holds in the following special cases:

(c_1) When $\nu \ll \bar{\nu}$ (assumption (A));
(c_2) When $g(x,y) > 0$ for all x, y and $\mathbb{P}(X_n \in \cdot) \ll \bar{\mathbb{P}}(X_n \in \cdot)$ for some $n \ge 0$;
(c_3) When $g(x,y) > 0$ for all x, y and $\|\mathbb{P}(X_n \in \cdot) - \bar{\mathbb{P}}(X_n \in \cdot)\| \to 0$ as $n \to \infty$.

The case (c_1) is immediate from (1.6). The case (c_2) is not difficult to prove, e.g., as in the proof of Proposition 2.5 in [34]. This is the case, for example, when the signal transition kernel has a strictly positive transition density, or in the linear Gaussian filtering model the signal is *controllable*—a typical assumption in stability results for the Kalman filter. The case (c_3) is proved as Lemma 3.7 in [36]. This is typically the case when the signal process is ergodic.

5.3. **Necessity.** We will finally argue that assumption (C) is necessary for filter stability, at least in the sense of total variation.

Proposition 5.2. *Suppose that assumption* (B) *holds, i.e., that the filter stability problem is well posed, but that* (C) *does not hold. Then* $\liminf_{n\to\infty} \mathbb{E}\|\pi_n - \bar{\pi}_n\| > 0$.

To prove the result, assume that (C) does not hold. Then there is set $A \in \mathcal{F}^Y_{0,\infty}$ such that $\bar{\mathbb{P}}(A) = 0$ and $\mathbb{P}(A) > 0$. But note that $\bar{\mathbb{P}}(A|\mathcal{F}^Y_{0,n}) = 0$ $\bar{\mathbb{P}}$-a.s., and by assumption (B) we also have $\bar{\mathbb{P}}(A|\mathcal{F}^Y_{0,n}) = 0$ $\mathbb{P}$-a.s. In particular,

$$|\mathbb{P}(A|\mathcal{F}^Y_{0,n}) - \bar{\mathbb{P}}(A|\mathcal{F}^Y_{0,n})| = \mathbb{P}(A|\mathcal{F}^Y_{0,n}) \xrightarrow{n\to\infty} I_A \quad \mathbb{P}\text{-a.s.}$$

Using $\|\mu - \mu'\| := 2\sup_B |\mu(B) - \mu'(B)|$, we obtain

$$\|\mathbb{P}((Y_k)_{k>n} \in \cdot|\mathcal{F}^Y_{0,n}) - \bar{\mathbb{P}}((Y_k)_{k>n} \in \cdot|\mathcal{F}^Y_{0,n})\| \ge 2|\mathbb{P}(A|\mathcal{F}^Y_{0,n}) - \bar{\mathbb{P}}(A|\mathcal{F}^Y_{0,n})|,$$

so using Fatou's lemma

$$\liminf_{n\to\infty} \mathbb{E}\|\mathbb{P}((Y_k)_{k>n} \in \cdot|\mathcal{F}^Y_{0,n}) - \bar{\mathbb{P}}((Y_k)_{k>n} \in \cdot|\mathcal{F}^Y_{0,n})\| \geq 2\mathbb{P}(A) > 0.$$

Finally, note that for any $B \in \mathcal{F}^Y_{n+1,\infty}$

$$\mathbb{P}(B|\mathcal{F}^Y_{0,n}) = \pi_n(f_B), \qquad \bar{\mathbb{P}}(B|\mathcal{F}^Y_{0,n}) = \bar{\pi}_n(f_B),$$

where $f_B(x) = \mathbb{P}(B|X_n = x)$. Therefore

$$\|\mathbb{P}((Y_k)_{k>n} \in \cdot|\mathcal{F}^Y_{0,n}) - \bar{\mathbb{P}}((Y_k)_{k>n} \in \cdot|\mathcal{F}^Y_{0,n})\| \leq \|\pi_n - \bar{\pi}_n\|,$$

and the proof is easily completed.

Remark 5.3. The above proof applies only to stability in the total variation norm. In general, it may be the case that (C) can be weakened if one is interested in weaker notions of stability; this is related to the consistency problem in Bayesian statistics [17]. Nonetheless, the necessity of (C) for total variation stability reassures us that our absolute continuity assumptions are not particularly restrictive. In particular, most of the literature to date has been concerned with total variation stability, and we have shown that no approach to the filter stability problem can circumvent the absolute continuity assumption (C) in this setting.

Acknowledgement. The work on this survey was initiated during the visit of P. Ch. in Laboratório Nacional de Computacão Cientifica, Petropolis, Brazil in August 2006 upon the invitation of Prof. Jack Baczynski, whose hospitality is greatly appreciated. The authors also thank a referee for several suggestions that have helped improve the presentation.

References

[1] Atar, Rami; Zeitouni, Ofer Lyapunov exponents for finite state nonlinear filtering. SIAM J. Control Optim. 35 (1997), no. 1, 36–55.

[2] Atar, Rami; Zeitouni, Ofer Exponential stability for nonlinear filtering. Ann. Inst. H. Poincare' Probab. Statist. 33 (1997), no. 6, 697–725.

[3] Baxendale, Peter; Chigansky, Pavel; Liptser, Robert Asymptotic stability of the Wonham filter: ergodic and nonergodic signals. SIAM J. Control Optim. 43 (2004), no. 2, 643–669

[4] Blackwell, D.; Dubins, L. Merging of opinions with increasing information. Ann. Math. Statist. 33 (1962), 882–886.

[5] Budhiraja, A.; Ocone, D. Exponential stability of discrete-time filters for bounded observation noise. Systems Control Lett. 30 (1997), no. 4, 185–193

[6] Budhiraja, A.; Ocone, D. Exponential stability in discrete-time filtering for non-ergodic signals. Stochastic Process. Appl. 82 (1999), no. 2, 245–257.

[7] Cappé, O.; Moulines, E.; Rydén, T. Inference in hidden Markov models. Springer Series in Statistics. Springer, New York, 2005.

[8] Chigansky, Pavel; Liptser, Robert Stability of nonlinear filters in nonmixing case. Ann. Appl. Probab. 14 (2004), no. 4, 2038–2056.

[9] Chigansky, Pavel; Liptser, Robert On a role of predictor in the filtering stability. Electron. Comm. Probab. 11 (2006), 129–140

[10] Chigansky, Pavel An ergodic theorem for filtering with applications to stability. Systems Control Lett. 55 (2006), no. 11, 908–917

[11] Chigansky, Pavel Stability of the nonlinear filter for slowly switching Markov chains. Stochastic Process. Appl. 116 (2006), no. 8, 1185–1194.

[12] Chigansky, Pavel; van Handel, Ramon Model robustness of finite state nonlinear filtering over the infinite time horizon. Ann. Appl. Probab. 17 (2007), no. 2, 688–715.

[13] Clark, J. M. C.; Ocone, D. L.; Coumarbatch, C. Relative entropy and error bounds for filtering of Markov processes. Math. Control Signals Systems 12 (1999), no. 4, 346–360.

[14] Del Moral, P. A uniform convergence theorem for the numerical solving of the nonlinear filtering problem. J. Appl. Probab. 35 (1998), 873–884.

[15] Del Moral, Pierre; Guionnet, Alice On the stability of interacting processes with applications to filtering and genetic algorithms. Ann. Inst. H. Poincare' Probab. Statist. 37 (2001), no. 2, 155–194.

[16] Delyon, Bernard; Zeitouni, Ofer Lyapunov exponents for filtering problems. Applied stochastic analysis, 511–521, Stochastics Monogr., 5, Gordon and Breach, New York, 1991.

[17] Diaconis, P.; Freedman, D. On the consistency of Bayes estimates. Ann. Statist. 14 (1986), 1–67.

[18] Douc, R.; Moulines, E.; Ritov, Y. Forgetting of the initial condition for the filter in general state-space hidden Markov chain: a coupling approach, preprint `arXiv:0712.0285`

[19] Kaijser, Thomas A limit theorem for partially observed Markov chains. Ann. Probability 3 (1975), no. 4, 677–696.

[20] Kleptsyna, Marina L.; Veretennikov, Alexander Yu. On ergodic filters with wrong initial data. C. R. Math. Acad. Sci. Paris 344 (2007), no. 11, 727–731

[21] Kunita, Hiroshi Asymptotic behavior of the nonlinear filtering errors of Markov processes. J. Multivariate Anal. 1 (1971), 365–393.

[22] LeGland, Francois; Oudjane, Nadia A robustification approach to stability and to uniform particle approximation of nonlinear filters: the example of pseudo-mixing signals. Stochastic Process. Appl. 106 (2003), no. 2, 279–316.

[23] Le Gland, Francois; Oudjane, Nadia Stability and uniform approximation of nonlinear filters using the Hilbert metric and application to particle filters. Ann. Appl. Probab. 14 (2004), no. 1, 144–187.

[24] Liptser, Robert S.; Shiryaev, Albert N. Statistics of random processes. Theory (I) and Applications (II). Applications of Mathematics (New York), 5,6. Stochastic Modelling and Applied Probability. Springer-Verlag, Berlin, 2001

[25] Meyn, S. P.; Tweedie, R. L. Markov chains and stochastic stability. Communications and Control Engineering Series. Springer-Verlag London, Ltd., London, 1993

[26] Nummelin, E. A splitting technique for Harris recurrent Markov chains. Z. Wahrsch. Verw. Gebiete 43 (1978), no. 4, 309–318.

[27] Ocone, Daniel; Pardoux, Etienne Asymptotic stability of the optimal filter with respect to its initial condition. SIAM J. Control Optim. 34 (1996), no. 1, 226–243.

[28] Ocone, Daniel Entropy inequalities and entropy dynamics in nonlinear filtering of diffusion processes. (English summary) Stochastic analysis, control, optimization and applications, 477–496, Systems Control Found. Appl., Birkhauser Boston, Boston, MA, 1999.

[29] Ocone, D. L. Asymptotic stability of Benes filters. Stochastic Anal. Appl. 17 (1999), no. 6, 1053–1074.

[30] Stannat, Wilhelm Stability of the filter equation for a time-dependent signal on $\mathbb{R}^d$. Appl. Math. Optim. 52 (2005), no. 1, 39–71.

[31] Stannat, Wilhelm Stability of the optimal filter via pointwise gradient estimates. Stochastic partial differential equations and applications—VII, 281–293, Lect. Notes Pure Appl. Math., 245, Chapman & Hall/CRC, Boca Raton, FL, 2006.

[32] Van Handel, Ramon Observability and nonlinear filtering, to appear in Probab. Th. Rel. Fields, preprint `arXiv:0708.3412`

[33] Van Handel, Ramon Uniform observability of hidden Markov models and filter stability for unstable signals, to appear in Ann. Appl. Probab., preprint `http://arxiv.org/abs/0804.2885`

[34] Van Handel, Ramon Discrete time nonlinear filters with informative observations are stable. Electr. Commun. Probab. 13 (2008), 562–575.

[35] Van Handel, Ramon Filtering, stability, and robustness (Caltech Ph.D thesis) 2007

[36] Van Handel, Ramon The Stability of Conditional Markov Processes and Markov Chains in Random Environments, preprint `arXiv:0801.4366`

[37] von Weizsäcker, Heinrich Exchanging the order of taking suprema and countable intersections of σ-algebras. Ann. Inst. H. Poincare' Sect. B (N.S.) 19 (1983), no. 1, 91–100.

Department of Statistics, The Hebrew University, Mount Scopus, Jerusalem 91905 Israel

E-mail address: pchiga@mscc.huji.ac.il

Department of Electrical Engineering Systems, Tel Aviv University, 69978 Tel Aviv, Israel

E-mail address: liptser@eng.tau.ac.il

Department of Operations Research and Financial Engineering, Princeton University, Princeton, NJ 08544, USA

E-mail address: rvan@princeton.edu