

## STABILITY OF THE DISCRETE TIME FILTER IN TERMS OF THE TAILS OF NOISE DISTRIBUTIONS \*

KARI HEINE<sup>1</sup>

**Abstract.** Recently, it has been pointed out by several authors that the uniform convergence of the stochastic discrete time filter approximations, such as particle filters, is closely related to the stability of the exact filter. This work provides easily verifiable conditions on the signal and observation model that ensure a certain level of stability for the exact filter. Essentially, the conditions are imposed on the tails of the signal and observation noise distributions. For sufficiently light tailed observation noise, the filter is shown to be stable.

### INTRODUCTION

Throughout the recent years, there has been active research on the stability of the stochastic discrete time filter [1–4, 6–9]. Perhaps the most important motivation for this interest in the stability is the analysis of the uniform convergence of various numerical filter approximation schemes, e.g. particle filters. Uniform, in this case, means uniform in time. Because of the recursive nature of the filter, the approximation error that was made earlier will contribute to the computation of the approximation at current time. This propagation of error can be considered as an incorrect initialisation of the filter at the time when the error was made, and thus the convergence analysis can be brought under the framework of filter stability with respect to the initial conditions.

The filtering framework under consideration is as follows. Let  $\mathcal{B}(\mathbb{R}^d)$  denote the Borel  $\sigma$ -algebra on  $\mathbb{R}^d$  and let  $\lambda_d$  denote the Lebesgue measure on  $\mathcal{B}(\mathbb{R}^d)$ . Moreover, let  $\mathcal{P}(\mathcal{B}(\mathbb{R}^d))$  denote the set of probability measures on  $\mathcal{B}(\mathbb{R}^d)$ . We define the signal process  $(X_i)_{i \geq 0}$  such that

$$X_i = f_i(X_{i-1}) + W_i, \tag{1}$$

where  $X_0$  is an independent random variable with distribution  $P_0 \in \mathcal{P}(\mathcal{B}(\mathbb{R}^{d_s}))$ ,  $f_i : \mathbb{R}^{d_s} \rightarrow \mathbb{R}^{d_s}$  is measurable, and  $W_i$  is an independent random variable with distribution  $P_{W_i} \in \mathcal{P}(\mathcal{B}(\mathbb{R}^{d_s}))$ . It is assumed that for all  $i > 0$ ,  $P_0$ ,  $P_{W_i}$  and  $\lambda_{d_s}$  are equivalent. The density of  $P_{W_i}$  with respect to  $\lambda_{d_s}$  is denoted by  $\rho_{W_i}$ .

The observation process  $(Y_i)_{i > 0}$  is of the form

$$Y_i = h_i(X_i) + V_i,$$

where  $h_i : \mathbb{R}^{d_s} \rightarrow \mathbb{R}^{d_m}$  is measurable and  $V_i$  is an independent random variable with distribution  $P_{V_i} \in \mathcal{P}(\mathcal{B}(\mathbb{R}^{d_m}))$ . It is assumed that for all  $i > 0$ ,  $P_{V_i}$  and  $\lambda_{d_m}$  are equivalent, and the density of  $P_{V_i}$  with respect

---

\* The research is funded by the Tampere Graduate School in Information Science and Engineering (TISE).

<sup>1</sup> Institute of Mathematics, Tampere University of Technology, Finland

to  $\lambda_{d_m}$  is denoted by  $\rho_{V_i}$ . The stochastic discrete time filter is then the probability measure valued stochastic process  $(\pi_i^*)_{i \geq 0}$  such that for all bounded and measurable functions  $\varphi : \mathbb{R}^{d_s} \rightarrow \mathbb{R}$ ,

$$\begin{aligned}\pi_0^*(\varphi) &= \mathbf{E}[\varphi(X_0)] \\ \pi_i^*(\varphi) &= \mathbf{E}[\varphi(X_i) \mid \mathcal{Y}_i],\end{aligned}$$

where  $i > 0$  and  $\mathcal{Y}_i$  is the  $\sigma$ -algebra generated by the observations  $(Y_1, Y_2, \dots, Y_i)$ . The filter process can also be expressed as a recursion  $\pi_i^* = Q_i(\pi_{i-1}^*)$ , where  $i > 0$ ,  $\pi_0^* = P_0$  and  $Q_i : \mathcal{P}(\mathcal{B}(\mathbb{R}^{d_s})) \rightarrow \mathcal{P}(\mathcal{B}(\mathbb{R}^{d_s}))$  is specified later in Section 1. We say that the filter  $(\pi_i^*)_{i \geq 0}$  is *stable* if for all  $\pi_0, \bar{\pi}_0 \in \mathcal{P}(\mathcal{B}(\mathbb{R}^{d_s}))$

$$\lim_{i \rightarrow \infty} \|Q_i \circ \dots \circ Q_1(\pi_0) - Q_i \circ \dots \circ Q_1(\bar{\pi}_0)\|_{\text{TV}} = 0, \quad \text{P-a.s.}, \quad (2)$$

where  $\|\cdot\|_{\text{TV}}$  denotes the total variation norm. The Euclidean norm is denoted by  $\|\cdot\|$ . In order to analyse the stability of the filter, we will impose the following assumptions on the filter framework:

- (A1) Functions  $f_i$ ,  $i > 0$  are  $a$ -Lipschitz.
- (A2) Functions  $h_i$ ,  $i > 0$  are bijective such that  $h_i$  and  $h_i^{-1}$  are  $b$ -Lipschitz.
- (A3) There are  $m_1, M_1, \alpha_1, \beta_1 > 0$  such that for all  $i > 0$ ,  $\rho_{W_i}$  satisfies

$$m_1 \exp(-\alpha_1 \|x\|^{\beta_1}) \leq \rho_{W_i}(x) \leq M_1 \exp(-\alpha_1 \|x\|^{\beta_1}).$$

- (A4) There are  $m_2, M_2, \alpha_2, \beta_2 > 0$  such that for all  $i > 0$ ,  $\rho_{V_i}$  satisfies

$$m_2 \exp(-\alpha_2 \|x\|^{\beta_2}) \leq \rho_{V_i}(x) \leq M_2 \exp(-\alpha_2 \|x\|^{\beta_2}).$$

Under these assumptions, the main result of this paper states that for sufficiently large values of  $\alpha_2$  and  $\beta_2$  the filter is stable, i.e., it satisfies (2). Moreover, from the proof of this result one can obtain the rates for the convergence in (2).

The remainder of this paper is organised as follows. Section 1 describes the principle of approximating the filter recursion by a truncated filter and explains the intuition behind the proof of the stability result. Section 2 provides an outline of the proof of the main result.

## 1. FILTER APPROXIMATION BY TRUNCATION

The idea of approximating the exact filter by a filter with compact state space in order to analyse the stability has been proposed in the literature (see, e.g., [7] and [9]). Let  $K_i : \mathbb{R}^{d_s} \times \mathcal{B}(\mathbb{R}^{d_s}) \rightarrow [0, 1]$  where  $i > 0$ , be the Markov kernels of  $(X_i)_{i \geq 0}$  such that for all  $A \in \mathcal{B}(\mathbb{R}^{d_s})$

$$K_i(X_{i-1}, A) = \mathbf{P}(X_i \in A \mid X_{i-1}), \quad \text{P-a.s.},$$

and let  $k_i(x_{i-1}, \cdot)$  denote the density of  $K_i(x_{i-1}, \cdot)$  with respect to  $\lambda_{d_s}$ . In the framework under consideration,  $k_i(x_{i-1}, x_i) = \rho_{W_i}(x_i - f_i(x_{i-1}))$ . We can then define a recursion

$$\pi_i := g_i \cdot \pi_{i-1} K_i,$$

where  $i > 0$ ,  $\pi_0 \in \mathcal{P}(\mathcal{B}(\mathbb{R}^{d_s}))$ ,  $g_i(x) := \rho_{V_i}(Y_i - h_i(x))$ , and  $\cdot$  denotes the *projective product* (see, e.g., [8]). If  $\pi_0 = P_0$ , then  $(\pi_i)_{i \geq 0} = (\pi_i^*)_{i \geq 0}$ , P-a.s., i.e., the recursion yields the filter process. In the stability analysis we are interested in the filter process with arbitrary initial distributions and therefore we allow the slight abuse of terminology and refer to  $(\pi_i)_{i \geq 0}$  with any initial distribution  $\pi_0 \in \mathcal{P}(\mathcal{B}(\mathbb{R}^{d_s}))$  as the *exact filter*.

The *truncated approximation*  $(\pi_i^\Delta)_{i \geq 0}$  of the exact filter  $(\pi_i)_{i \geq 0}$  is defined by the recursion

$$\pi_i^\Delta := g_i^\Delta \cdot \pi_{i-1}^\Delta K_i, \quad (3)$$

where  $\pi_0^\Delta = \pi_0$ ,  $g_i^\Delta := 1_{C_i(\Delta)} g_i$  and for all  $\Delta > 0$ ,

$$C_i(\Delta) := \begin{cases} \{x \in \mathbb{R}^{d_s} \mid \|Y_i - h_i(x)\| \leq \Delta\}, & i > 0 \\ \{x \in \mathbb{R}^{d_s} \mid \|X_0 - x\| \leq b\Delta\}, & i = 0. \end{cases}$$

By defining  $C_i(\infty) := \mathbb{R}^{d_s}$  for all  $i \geq 0$ , we observe that when  $\Delta = \infty$ , (3) yields the exact filter. For all  $i > 0$ , we let  $Q_i^\Delta$  denote the mapping  $\pi_{i-1}^\Delta \mapsto \pi_i^\Delta$ , and  $Q_i := Q_i^\infty$ . Moreover, we define for all  $i \geq j > 0$ ,  $Q_{i,j}^\Delta := Q_i^\Delta \circ \dots \circ Q_j^\Delta$  and if  $i < j$ , the mapping  $Q_{i,j}^\Delta$  is defined to be identity. Then we can define for all  $i \geq j \geq 0$ , and  $\pi_0, \bar{\pi}_0 \in \mathcal{P}(\mathcal{B}(\mathbb{R}^{d_s}))$

$$\begin{aligned} \pi_{i,j}^\Delta &:= Q_{i,j+1}^\Delta(Q_{j,1}(\pi_0)) \\ \bar{\pi}_{i,j}^\Delta &:= Q_{i,j+1}^\Delta(Q_{j,1}(\bar{\pi}_0)). \end{aligned}$$

Accordingly,  $\pi_{i,0}^\Delta = \pi_i^\Delta$  and  $\bar{\pi}_{i,i}^\Delta = \pi_i$ , and similarly  $\bar{\pi}_{i,0}^\Delta = \bar{\pi}_i^\Delta$  and  $\bar{\pi}_{i,i}^\Delta = \bar{\pi}_i$ .

According to the definitions above, the error due to different initialisations can be written by the triangle inequality as

$$\begin{aligned} \|\pi_i - \bar{\pi}_i\|_{\text{TV}} &\leq \|\pi_{i,i} - \pi_{i,0}^\Delta\|_{\text{TV}} + \|\pi_{i,0}^\Delta - \bar{\pi}_{i,0}^\Delta\|_{\text{TV}} + \|\bar{\pi}_{i,0}^\Delta - \bar{\pi}_{i,i}\|_{\text{TV}} \\ &= \|\pi_i - \pi_i^\Delta\|_{\text{TV}} + \|\pi_i^\Delta - \bar{\pi}_i^\Delta\|_{\text{TV}} + \|\bar{\pi}_i^\Delta - \bar{\pi}_i\|_{\text{TV}}. \end{aligned} \quad (4)$$

For illustration, see Figure 1. It would be sufficient for the stability that all terms on the right hand side of (4) converge to zero as  $i \rightarrow \infty$ . However, the truncation does not yield a uniformly convergent approximation of the exact filter as  $\Delta \rightarrow \infty$  and therefore, the first and the last term on the right hand side of (4) do not converge to zero as  $i \rightarrow \infty$ . Uniform, in this case, means uniform in time. For this reason, we write

$$\|\pi_i - \bar{\pi}_i\|_{\text{TV}} \leq \|\pi_i - \pi_i^{\Delta_i}\|_{\text{TV}} + \|\pi_i^{\Delta_i} - \bar{\pi}_i^{\Delta_i}\|_{\text{TV}} + \|\bar{\pi}_i^{\Delta_i} - \bar{\pi}_i\|_{\text{TV}}, \quad (5)$$

i.e., we write the upper bound for the error in terms of truncated filters with different truncation radii  $\Delta_i$ . In this case, the first and the last terms on the right hand side can be made to converge to zero by letting  $(\Delta_i)_{i \geq 0}$  be a sufficiently fast increasing sequence. It should be emphasised that by definition,  $(\pi_i^{\Delta_i})_{i \geq 0}$  satisfies

$$\pi_i^{\Delta_i} = Q_i^{\Delta_i} \circ \dots \circ Q_2^{\Delta_i} \circ Q_1^{\Delta_i}(\pi_0),$$

that is, the same truncation radius  $\Delta_i$  is used up to time  $i$ . We do not allow a filter, where different radii would be used at consecutive steps, i.e.,

$$\pi_i^{\Delta_i} = Q_i^{\Delta_i} \circ \dots \circ Q_2^{\Delta_2} \circ Q_1^{\Delta_1}(\pi_0),$$

where  $\Delta_i \neq \Delta_j$  for  $i \neq j$ . The use of the upper bound in (5) has a downside as well. While the first and the last term converge to zero when  $(\Delta_i)_{i \geq 0}$  is a sufficiently fast increasing sequence, the convergence is not guaranteed for the middle term and it may well become divergent if  $(\Delta_i)_{i \geq 0}$  increases too rapidly. Therefore the fundamental issue in proving the stability is to show that there exists a rate for increasing  $(\Delta_i)_{i \geq 0}$  such that all three terms in (5) converge to zero as  $i \rightarrow \infty$ .

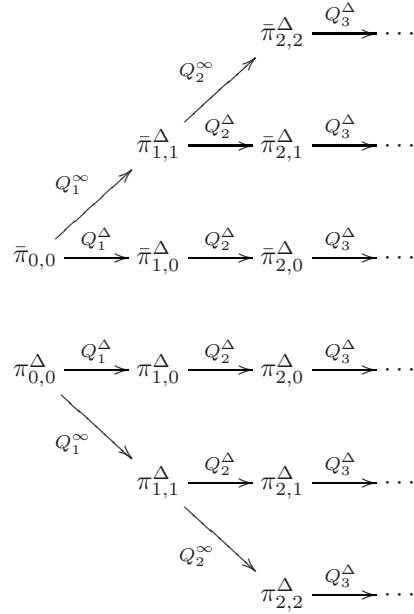


FIGURE 1. An illustration of two exact filters with different initial distributions and their truncated approximations.

## 2. OUTLINE OF THE PROOF OF STABILITY

In this section, we state the main result of this paper accompanied by an outline of the proof and some intermediate results. Detailed proofs are omitted, but they can be found in [5] and they are based on principles originally introduced in [9] and [6]. Let us first consider the bounding of the error due to truncation in  $\mathbf{P}$ -almost sure sense. For this purpose, we have the following proposition.

**Proposition 2.1.** *For all  $\epsilon \in (0, \min(\alpha_1, \alpha_2))$  and  $\Delta^* > 0$ , there exist positive random variables  $c_1 = c_1(\epsilon, \Delta^*)$ ,  $c_2 = c_2(\epsilon)$ ,  $c_3 = c_3(\epsilon)$  such that for all  $\pi_0 \in \mathcal{P}(\mathcal{B}(\mathbb{R}^{d_s}))$ ,  $\Delta \geq \Delta^*$  and  $i > 0$*

$$\|\pi_i - \pi_i^\Delta\|_{\text{TV}} \leq c_1 \sum_{j=1}^i \frac{\exp((-\alpha_2 + \epsilon)\Delta^{\beta_2} + 2\alpha_1((ab+b)\Delta + \xi_{j+1,\epsilon})^{\beta_1})}{1 - \pi_{j-1}(C_{j-1}(\Delta))}, \quad \mathbf{P}\text{-a.s.} \quad (6)$$

where

$$\xi_{i,\epsilon} := (c_2 + (\alpha_1 - \epsilon)^{-1} \ln i)^{1/\beta_1} + (ab+b)(c_3 + (\alpha_2 - \epsilon)^{-1} \ln i)^{1/\beta_2}. \quad (7)$$

According to (6), Proposition 2.1 implies that for

$$\lim_{i \rightarrow \infty} \|\pi_i - \pi_i^\Delta\|_{\text{TV}} = 0, \quad (8)$$

it is sufficient to choose  $(\Delta_i)_{i \geq 0}$  such that the right hand side of (6) converges to zero as  $i \rightarrow \infty$ . For this, in turn, it is necessary to have

$$\lim_{i \rightarrow \infty} (-\alpha_2 + \epsilon)\Delta_i^{\beta_2} + 2\alpha_1\Delta_i^{\beta_1} \left( ab + b + \frac{\xi_{i+1,\epsilon}}{\Delta_i} \right)^{\beta_1} = -\infty. \quad (9)$$

Moreover, it can be shown that for (9) to hold when  $\beta_2 \geq \beta_1$ , it is sufficient that

$$\sup_{i>0} \xi_{i+1,\epsilon}/\Delta_i < \infty,$$

provided that  $\alpha_2$  is sufficiently large<sup>1</sup>. Therefore, it is observed that according to (7), if for all  $i > 1$  we set  $\Delta_i^{\beta_1} = s \ln i$ , where  $s > 0$ , then for all  $\epsilon \in (0, \min(\alpha_1, \alpha_2))$

$$\lim_{i \rightarrow \infty} \frac{\xi_{i+1,\epsilon}}{\Delta_i} = \begin{cases} \frac{ab+b}{s^{1/\beta_1}(\alpha_2-\epsilon)^{1/\beta_2}} + \frac{1}{s^{1/\beta_1}(\alpha_1-\epsilon)^{1/\beta_1}}, & \beta_2 = \beta_1 \\ \frac{1}{s^{1/\beta_1}(\alpha_1-\epsilon)^{1/\beta_1}}, & \beta_2 > \beta_1, \end{cases} \quad (10)$$

and hence  $\xi_{i+1,\epsilon}/\Delta_i$  is bounded. Thus, in order to prove the stability, it remains to prove that (9) is a sufficient condition for (8), provided that the convergence in (9) is fast enough, and that if  $\Delta_i^{\beta_1} = s \ln i$  for all  $i > 1$ , the middle term in (5) also converges to zero, i.e.,

$$\lim_{i \rightarrow \infty} \left\| \pi_i^{\Delta_i} - \bar{\pi}_i^{\Delta_i} \right\|_{\text{TV}} = 0. \quad (11)$$

The proof of the first claim is provided in [5] with the additional condition that if  $\beta_1 = \beta_2$ , then one must have  $\alpha_2 > \kappa_0$ , where  $\kappa_p := \gamma^{-1}(-p)$  for all  $p \geq 0$  and  $\gamma : \mathbb{R}_+ \rightarrow \mathbb{R}$  is defined as

$$\gamma(x) := 2 \left( 2 + \frac{\alpha_1^{1/\beta_1}(ab+b)}{x^{1/\beta_1}} \right)^{\beta_1} - \frac{x}{\alpha_1(ab+b)^{\beta_1}} + 1.$$

The second claim is given by the following proposition whose proof can also be found in [5].

**Proposition 2.2.** *For all  $i > 1$ , let  $\Delta_i^{\beta_1} = s \ln i$ . If  $s^{-1} > \alpha_1(ab+b)^{\beta_1}$ , then for all  $\epsilon > 0$ , there exists  $c_4 = c_4(s) > 0$  and a positive random variable  $c_5 = c_5(\epsilon)$  such that for all  $i > 1$*

$$\left\| \pi_i^{\Delta_i} - \bar{\pi}_i^{\Delta_i} \right\|_{\text{TV}} \leq c_5 \exp((-m_1/2M_1 + \epsilon)i^{c_4}), \quad \text{P-a.s.}, \quad (12)$$

As a conclusion, we have the main result

**Theorem 2.3.** *If one of the following two conditions holds:*

- (1)  $\beta_1 = \beta_2$  and  $\alpha_2 > \kappa_0$ ,
- (2)  $\beta_1 < \beta_2$ ,

then  $(\pi_i^*)_{i \geq 0}$  is stable.

It should be pointed out that the detailed proof of Theorem 2.3 also provides rates at which the initialisation error converges to zero (see [5]). These rates are given by the following corollary.

**Corollary 2.4.** (1) *If  $\beta_1 = \beta_2$ , then for all  $\pi_0, \bar{\pi}_0 \in \mathcal{P}(\mathcal{B}(\mathbb{R}^{d_s}))$ ,  $p > 0$  and  $\alpha_2 > \kappa_p$  there exists a random variable  $c_6 = c_6(\pi_0, \bar{\pi}_0, p)$  such that for all  $i > 0$ ,*

$$\|\pi_i - \bar{\pi}_i\|_{\text{TV}} \leq c_6 i^{-p}, \quad \text{P-a.s.}$$

- (2) *If  $\beta_1 < \beta_2$ , then for all  $\pi_0, \bar{\pi}_0 \in \mathcal{P}(\mathcal{B}(\mathbb{R}^{d_s}))$ , there exist  $c_7 > 0$  and a random variable  $c_8 = c_8(\pi_0, \bar{\pi}_0, c_7) > 0$  such that for all  $i > 0$ ,*

$$\|\pi_i - \bar{\pi}_i\|_{\text{TV}} \leq c_8 i \exp\left(-c_7(\ln i)^{\beta_2/\beta_1}\right), \quad \text{P-a.s.}$$

<sup>1</sup>In the case  $\beta_1 = \beta_2$ , the boundedness of the ratio is also necessary.

The author wishes to thank Dan Crisan who is an important contributor to this work.

## REFERENCES

- [1] R. Atar and O. Zeitouni. Exponential stability for nonlinear filtering. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 33(6):697–725, 1997.
- [2] R. Atar and O. Zeitouni. Lyapunov exponents for finite state nonlinear filtering. *SIAM Journal of Control and Optimization*, 35(1):36–55, 1997.
- [3] A. Budhiraja and D. Ocone. Exponential stability of discrete time filters for bounded observation noise. *Systems & Control Letters*, 30(4):149–207, 1997.
- [4] A. Budhiraja and D. Ocone. Exponential stability in discrete-time filtering for non-ergodic signals. *Stochastic Processes and Their Applications*, 82(2):245–257, 1999.
- [5] D. Crisan and K. Heine. Stability of the discrete time filter in terms of the tails of noise distributions. *Submitted to the London Mathematical Society*, 2006.
- [6] P. Del Moral and A. Guionnet. On the stability of interacting processes with applications to filtering and genetic algorithms. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 37(2):155–194, 2001.
- [7] F. LeGland and N. Oudjane. A robustification approach to stability and to uniform particle approximation of nonlinear filters: the example of pseudo-mixing signals. *Stochastic Processes and their Applications*, 106(2):279–316, 2003.
- [8] F. LeGland and N. Oudjane. Stability and uniform approximation of nonlinear filters using the Hilbert metric, and applications to particle filters. *The Annals of Applied Probability*, 14(1):144–187, 2004.
- [9] N. Oudjane and S. Rubenthaler. Stability and uniform particle approximation of nonlinear filters in case of non ergodic signals. *Stochastic Analysis and Applications*, 23(3):421–448, 2005.