ON THE L_P-CONVERGENCE OF A GIRSANOV THEOREM BASED PARTICLE FILTER

Simo Särkkä[*]

Dept. of Electrical Engineering and Automation
Aalto University, Finland

Eric Moulines

Center of Applied Mathematics
École Polytechnique, France

ABSTRACT

We analyze the L_p-convergence of a previously proposed Girsanov theorem based particle filter for discretely observed stochastic differential equation (SDE) models. We prove the convergence of the algorithm with the number of particles tending to infinity by requiring a moment condition and a step-wise initial condition boundedness for the stochastic exponential process giving the likelihood ratio of the SDEs. The practical implications of the condition are illustrated with an Ornstein–Uhlenbeck model and with a non-linear Beneš model.

Index Terms— Girsanov theorem, particle filter, convergence, stochastic differential equation

1. INTRODUCTION

In this article, we analyze the L_p-convergence of the Girsanov theorem based particle filter introduced in [1–3]. The particle filter is concerned with the classical problem [4] of discretely observed stochastic differential equations of the form

$$\begin{aligned} dX(t) &= f(X(t), t)\, dt + L(t)\, dB(t), \\ y_k &\sim \rho_k(y_k \mid x_k), \end{aligned} \tag{1}$$

where $\rho_k(y_k \mid x_k)$ is the conditional probability density (here w.r.t. the Lebesgue measure) of the measurement $y_k \in \mathbb{R}^d$ given the state $X(t_k) = x_k \in \mathrm{R}^n$. We assume that the vector of standard Brownian motions $B(t) \in \mathbb{R}^n$ and that $L(t)$ is invertible, but this can be relaxed [3].

Although there exists a wide range of L_p-convergence results for particle filters (see, e.g., [5–13] and references therein), the main difficulty in applying these results to the present filter is that unlike in many other cases, the importance weights cannot be assumed to be point-wise bounded. Therefore we base our analysis on the recently proposed more general moment conditions on the weights [14, 15].

Furthermore, although here we only consider the convergence of the filtering measures at the measurement times, the particle filter method [1–3] actually produces samples of the full paths of the posterior process. Even though this limits the possible choices of importance processes to those which are absolutely continuous with respect to the dynamic model process, it also enables the possibility estimate the values of functionals of the paths.

[*]Thanks to Academy of Finland for funding.

2. THE PARTICLE FILTER

The particle filter introduced in [1–3] is based on the classical Girsanov theorem [16] which gives an expression for the likelihood ratio between an SDE and its driving Brownian motion. From the theorem, it is also possible to derive an expression for the stochastic exponential process $Z(t)$ giving the likelihood ratio between two SDEs driven by the same Brownian motion (see [3]), which can be used for importance sampling of the SDEs in particle filtering. The resulting particle filter algorithm is the following.

Algorithm 1 (Girsanov theorem based particle filter)**.** *Given a set of Monte Carlo samples $\{x_{k-1}^{(i)} : i = 1, \ldots, N\}$ and the new measurement y_k, a single step of the filter is:*

1. *Simulate N independent realizations of the importance process from $t = t_{k-1}$ to $t = t_k$:*

$$\begin{aligned} dS^{(i)}(t) &= g(S^{(i)}, t)\, dt + L(t)\, dB^{(i)}(t), \\ S^{(i)}(t_{k-1}) &= x_{k-1}^{(i)}. \end{aligned} \tag{2}$$

2. *Simulate the corresponding log-likelihood ratios*

$$\begin{aligned} d\Lambda^{(i)}(t) &= h^T(S^{(i)}(t), t)\, [L^{-1}(t)]^T dB^{(i)}(t) \\ &- \frac{1}{2} h^T(S^{(i)}(t), t)\, \left(L(t)\, L^T(t)\right)^{-1} h(S^{(i)}(t), t)\, dt, \end{aligned} \tag{3}$$

where we have defined

$$h(S, t) = f(S, t) - g(S, t), \tag{4}$$

from $t = t_{k-1}$ to $t = t_k$ with $\Lambda^{(i)}(t_{k-1}) = 0$ and set

$$\begin{aligned} \tilde{x}_k^{(i)} &= S^{(i)}(t_k), \\ z_k^{(i)} &= \exp\left\{\Lambda^{(i)}(t_k)\right\}. \end{aligned} \tag{5}$$

Note that the realizations of Brownian motions must be the same as in simulation of the importance processes.

3. *For each i compute*

$$w_k^{(i)} = z_k^{(i)} \, \rho_k(y_k \mid \tilde{x}_k^{(i)}), \tag{6}$$

and normalize them:

$$\tilde{w}_k^{(i)} = \frac{w_k^{(i)}}{\sum_{j=1}^n w_k^{(j)}}. \tag{7}$$

4. *Resample $\{\tilde{x}_k^{(i)}, \tilde{w}_k^{(i)}\}$ to obtain $\{x_k^{(i)} \,:\, i = 1, \ldots, N\}$.*

3. MEASURE-THEORETICAL INTERPRETATION

Let us denote the set of bounded Borel-measurable functions $\|\phi\|_\infty < \infty$ by $\mathcal{B}(\mathbb{R}^n)$. If Q is the transition kernel of a Markov process we denote $Q(\phi)(x) = \int \phi(y)\, Q(x, dy)$. We denote the transition kernel from $X(t_{k-1}) = x_{k-1}$ to $X(t_k)$ defined by the SDE in (1) as Q_k which is usually intractable to write down explicitly. We write η_k for the conditional (filtering) measure of $X(t_k)$, given the observations $y_1, \ldots, y_k$. The Bayesian filter in a "test function form" can then be written as [8]

$$\alpha_k(\phi) = \eta_{k-1}(Q_k(\phi\, \rho_k)), \quad \eta_k(\phi) = \frac{\alpha_k(\phi)}{\alpha_k(1)}, \tag{8}$$

where α_k is an unnormalized measure and $\rho_k(y_k \mid x_k)$ is considered as a function of x_k. To account for the importance process, it is convenient to rewrite the equations into the following equivalent form (cf. [15])

$$\alpha_k(\phi) = \eta_{k-1}(\Pi_k(\phi\, w_k)), \quad \eta_k(\phi) = \frac{\alpha_k(\phi)}{\alpha_k(1)}, \tag{9}$$

where Π_k is the transition kernel of the Markov process defined by the importance process SDE for the transition from $S(t_{k-1}) = x_{k-1}$ to $S(t_k)$. In the above display we have the weight function

$$w_k(x_{k-1}, x_k) = \rho_k(y_k \mid x_k)\, \frac{dQ_k}{d\Pi_k}(x_{k-1}, x_k), \tag{10}$$

where $dQ_k / d\Pi_k$ is a Radon–Nikodym derivative. The advantage of this formulation is that the particle filter in Algorithm 1 can be seen as a direct Monte Carlo approximation of Equations (9) as follows:

1. The simulation of the importance process in (2) can be seen as drawing N samples from the measure $\eta_{k-1}^N(\Pi_k)$, where η_{k-1}^N is the N-particle approximation of the filtering distribution from the time step $k-1$.

2. Combining with (3) and (5) leads to the approximation

$$\tilde{\alpha}_k^N(\phi) = \sum_{i=1}^N \phi(\tilde{x}_k^{(i)})\, w_k(x_{k-1}^{(i)}, \tilde{x}_k^{(i)}). \tag{11}$$

3. In (6) and (7) we form the approximation

$$\tilde{\eta}_k^N(\phi) = \frac{\tilde{\alpha}_k^N(\phi)}{\tilde{\alpha}_k^N(1)}. \tag{12}$$

4. Resampling step forms a new measure η_k^N from $\tilde{\eta}_k^N$.

4. L_P CONVERGENCE THEORY

The main convergence theorem is the following.

Theorem 2. *Assume that*

1. *The measurement model density is bounded $\rho_k(y_k \mid x_k) \leq D_k < \infty$.*

2. *The transition kernels of the SDEs are Feller.*

3. *The importance weights and the measurement model density satisfy the inequality*

$$\sup_{x_{k-1}} \Pi_k(|w_k|^p)(x_{k-1}) \leq E_k \tag{13}$$

for some constants $E_k < \infty$ for all $k = 1, \ldots, M$, that is, it bounds uniformly for all starting points x_{k-1}.

4. *The resampling algorithm satisfies (e.g. [8]):*

$$\mathrm{E}\left[\left|\tilde{\eta}_k^N(\phi) - \eta_k^N(\phi)\right|^p\right] \leq \frac{\hat{c}_k\, \|\phi\|_\infty^p}{N^{\frac{p}{2}}} \tag{14}$$

for some constant $\hat{c}_k$, independent of N.

Then for some set of constants c_k, for all $k = 1, \ldots, M$, independent of N, for all $\phi \in \mathcal{B}(\mathbb{R}^n)$ we have

$$\mathrm{E}\left[\left|\eta_k^N(\phi) - \eta_k(\phi)\right|^p\right] \leq \frac{c_k\, \|\phi\|_\infty^p}{N^{\frac{p}{2}}}. \tag{15}$$

We start the proof of the above with the following lemma.

Lemma 3. *Let $p \geq 2$, and $\{\xi_i : i = 1, \ldots, N\}$ be conditionally independent random variables given a sigma-algebra $\mathcal{G}$ such that $\mathrm{E}[|\xi_i|^p \mid \mathcal{G}] < \infty$. Then we have*

$$\mathrm{E}\left[\left|\sum_{i=1}^N \xi_i - \sum_{i=1}^N \mathrm{E}\left[\xi_i | \mathcal{G}\right]\right|^p \mid \mathcal{G}\right] \leq C_p \left(\sum_{i=1}^N \mathrm{E}\left[|\xi_i|^p \,|\mathcal{G}\right]^{\frac{2}{p}}\right)^{\frac{p}{2}} \tag{16}$$

where C_p is independent of N.

Proof. Follows from Theorem 2.12 in [17]. □

Lemma 4. *Assume that we have*

$$\mathrm{E}\left[\left|\eta_{k-1}^N(\phi) - \eta_{k-1}(\phi)\right|^p\right] \leq \frac{c_{k-1}\, \|\phi\|_\infty^p}{N^{\frac{p}{2}}} \tag{17}$$

for some constant c_{k-1}, independent of N. Then

$$\mathrm{E}\left[\left|\tilde{\eta}_k^N(\phi) - \eta_k(\phi)\right|^p\right] \leq \frac{\tilde{c}_k\, \|\phi\|_\infty^p}{N^{\frac{p}{2}}} \tag{18}$$

for some constant $\tilde{c}_k$, independent of N.

Proof. By using Minkowski's inequality we get

$$\mathrm{E}\left[\left|\tilde{\eta}_k^N(\phi) - \eta_k(\phi)\right|^p\right]^{\frac{1}{p}} = \mathrm{E}\left[\left|\frac{\tilde{\alpha}_k^N(\phi)}{\tilde{\alpha}_k^N(1)} - \frac{\alpha_k(\phi)}{\alpha_k(1)}\right|^p\right]^{\frac{1}{p}}$$
$$\leq \mathrm{E}\left[\left|\frac{\tilde{\alpha}_k^N(\phi)}{\tilde{\alpha}_k^N(1)} - \frac{\tilde{\alpha}_k^N(\phi)}{\alpha_k(1)}\right|^p\right]^{\frac{1}{p}} + \mathrm{E}\left[\left|\frac{\tilde{\alpha}_k^N(\phi)}{\alpha_k(1)} - \frac{\alpha_k(\phi)}{\alpha_k(1)}\right|^p\right]^{\frac{1}{p}}. \tag{19}$$

For the first term above we get

$$\mathrm{E}\left[\left|\frac{\tilde{\alpha}_k^N(\phi)}{\tilde{\alpha}_k^N(1)} - \frac{\tilde{\alpha}_k^N(\phi)}{\alpha_k(1)}\right|^p\right] \leq \frac{||\phi||_\infty^p}{|\alpha_k(1)|^p}\,\mathrm{E}\left[\left|\alpha_k(1) - \tilde{\alpha}_k^N(1)\right|^p\right]. \tag{20}$$

For the second term we get

$$\mathrm{E}\left[\left|\frac{\tilde{\alpha}_k^N(\phi)}{\alpha_k(1)} - \frac{\alpha_k(\phi)}{\alpha_k(1)}\right|^p\right] = \frac{1}{|\alpha_k(1)|^p}\,\mathrm{E}\left[\left|\tilde{\alpha}_k^N(\phi) - \alpha_k(\phi)\right|^p\right]. \tag{21}$$

We also have

$$\begin{aligned}
&\mathrm{E}\left[\left|\tilde{\alpha}_k^N(\phi) - \alpha_k(\phi)\right|^p\right]^{\frac{1}{p}} \\
&\leq \mathrm{E}\left[\left|\tilde{\alpha}_k^N(\phi) - \mathrm{E}\left[\tilde{\alpha}_k^N(\phi) \mid \mathcal{G}_{k-1}\right]\right|^p \mid \mathcal{G}_{k-1}\right]^{\frac{1}{p}} \\
&\quad + \mathrm{E}\left[\left|\mathrm{E}\left[\tilde{\alpha}_k^N(\phi) \mid \mathcal{G}_{k-1}\right] - \alpha_k(\phi)\right|^p\right]^{\frac{1}{p}},
\end{aligned} \tag{22}$$

where $\mathcal{G}_{k-1}$ denotes the sigma-algebra generated by the particles $\{x_{0:k-1}^{(i)} : i = 1, \ldots, N\}$. For the first term here we get by using Lemma 3 and assumption 3:

$$\begin{aligned}
&\mathrm{E}\left[\left|\tilde{\alpha}_k^N(\phi) - \mathrm{E}[\tilde{\alpha}_k^N(\phi) \mid \mathcal{G}_{k-1}]\right|^p \mid \mathcal{G}_{k-1}\right] \\
&\leq \frac{C_p}{N^p}\left(\sum_{i=1}^N \mathrm{E}\left[\left|\phi(\tilde{x}_k^{(i)})\, z_k^{(i)}\, \rho_k(y_k \mid \tilde{x}_k^{(i)})\right|^p \mid \mathcal{G}_{k-1}\right]^{\frac{2}{p}}\right)^{p/2} \\
&\leq \frac{E_k\, C_p\, \|\phi\|_\infty^p}{N^{p/2}}.
\end{aligned} \tag{23}$$

The second term gives by using the induction assumption

$$\begin{aligned}
&\mathrm{E}\left[\left|\mathrm{E}\left[\tilde{\alpha}_k^N(\phi) \mid \mathcal{G}_{k-1}\right] - \alpha_k(\phi)\right|^p\right] \\
&= \mathrm{E}\left[\left|\eta_{k-1}^N(Q_k(\phi\,\rho_k)) - \eta_{k-1}(Q_k(\phi\,\rho_k))\right|^p\right] \\
&\leq \frac{c_k\, \|Q_k(\phi\,\rho_k)\|_\infty^p}{N^{p/2}} \leq \frac{c_k\, D_k \|\phi\|_\infty^p}{N^{p/2}},
\end{aligned} \tag{24}$$

where we have used $\mathrm{E}[\tilde{\alpha}_k^N(\phi) \mid \mathcal{G}_{k-1}] = \eta_{k-1}^N(Q_k(\phi\,\rho_k))$. The result follows by substituting (23) and (24) into (22), then the result to (20) and (21) (to former with $\phi = 1$), and by finally using (19). ☐

Lemma 5. *Assume that we have*

$$\mathrm{E}\left[\left|\tilde{\eta}_k^N(\phi) - \eta_k(\phi)\right|^p\right] \leq \frac{\tilde{c}_k\, \|\phi\|_\infty^p}{N^{\frac{p}{2}}} \tag{25}$$

for some constant $\tilde{c}_k$, independent of N. Then

$$\mathrm{E}\left[\left|\eta_k^N(\phi) - \eta_k(\phi)\right|^p\right] \leq \frac{c_k\, \|\phi\|_\infty^p}{N^{\frac{p}{2}}} \tag{26}$$

for some constant c_k, independent of N.

Proof. The result follows from $\mathrm{E}\left[\left|\eta_k^N(\phi) - \eta_k(\phi)\right|^p\right]^{\frac{1}{p}} \leq \mathrm{E}\left[\left|\eta_k^N(\phi) - \tilde{\eta}_k^N(\phi)\right|^p\right]^{\frac{1}{p}} + \mathrm{E}\left[\left|\tilde{\eta}_k^N(\phi) - \eta_k(\phi)\right|^p\right]^{\frac{1}{p}}$ together with the assumption (14). ☐

Proof of Theorem 2. The result follows by combining Lemmas 4 and 5 together with a simple induction argument similarly to [15]. ☐

5. ENSURING THE ASSUMPTION 3

Let us now discuss what the condition that $\Pi_k(|w_k|^p)(x_{k-1}) \leq E_k$ uniformly for all starting points x_{k-1} actually means and how it can be checked in practice. If the Lebesgue densities of Π_k and Q_k exist and are π_k and q_k, respectively, then the condition is equivalent to the following being true regardless of x_{k-1}:

$$\int \left[\frac{\rho_k(y_k \mid x_k)\, q_k(x_k \mid x_{k-1})}{\pi_k(x_k \mid x_{k-1})}\right]^p \pi_k(x_k \mid x_{k-1})\, dx_k \leq E_k. \tag{27}$$

This will certainly be true if we can ensure that the unnormalized weights in the brackets above are uniformly bounded in both variables x_k and x_{k-1}. However, we cannot generally ensure that.

One way to proceed is to explicitly check that the condition above is true for the transition densities of the dynamic model and importance process SDEs. However, for non-linear SDEs the computation of the densities is usually intractable (they are solutions of the Fokker–Planck–Kolmogorov partial differential equation). Still, sometimes analytical or numerical analysis is possible.

We have $\Pi_k(|w_k|^p) \leq D_k\, \Pi_k([dQ_k/d\Pi_k]^p)$ and thus we can also attempt to ensure that $\Pi_k([dQ_k/d\Pi_k]^p) \leq \tilde{E}_k$ regardless of the starting point x_{k-1}. It is worth noting that this gives a sufficient condition for the convergence, but $\Pi_k(|w_k|^p) \leq E_k$ might be true even when $\Pi_k([dQ_k/d\Pi_k]^p) \leq \tilde{E}_k$ is not due to appearance of the potentially regularizing function ρ_k. Explicitly written, the latter condition is (recall (4))

$$\begin{aligned}
&\mathrm{E}_{x_{k-1}}\Big[\exp\Big(p\int_{t_{k-1}}^{t_k} h^T(S(t),t)\,[L^{-1}(t)]^T dB(t) \\
&- \frac{p}{2}\int_{t_{k-1}}^{t_k} h^T(S(t),t)\left(L(t)\, L^T(t)\right)^{-1} h(S(t),t)\, dt\Big] \leq \tilde{E}_k
\end{aligned} \tag{28}$$

which is related to so called Novikov's conditions for martingales (with $p = 1$) and the moments of the likelihood ratio considered in [18]. These conditions essentially say that provided that

$$\mathrm{E}_{x_{k-1}} \Big[\exp \Big(c_p \int_{t_{k-1}}^{t_k} h^T(S(t),t) \left(L(t)\, L^T(t)\right)^{-1} \times h(S(t),t)\, dt \Big) \Big] < \infty \tag{29}$$

for a suitably chosen constant c_p, then the moment is bounded. However, these conditions do not say anything about the boundedness in the initial conditions (i.e., x_{k-1}).

We can also put back the measurement model into the condition (28), which leads to the condition

$$\mathrm{E}_{x_{k-1}} \Big[\exp \Big(p \int_{t_{k-1}}^{t_k} h^T(S(t),t) [L^{-1}(t)]^T dB(t) - \frac{p}{2} \int_{t_{k-1}}^{t_k} h^T(S(t),t) \, \left(L(t)\, L^T(t)\right)^{-1} h(S(t),t)\, dt \Big) \times \rho_k^p(y_k \mid S(t_k)) \Big] \le \tilde{E}_k. \tag{30}$$

6. EXAMPLE: ORNSTEIN–UHLENBECK MODEL

In this section we illustrate the condition (27) discussed in the previous section by explicitly analyzing its implications on the following Ornstein–Uhlenbeck model:

$$\begin{aligned} dX(t) &= -a\, X(t)\, dt + q^{1/2}\, dB(t), \\ \rho(y_k \mid x_k) &= \frac{1}{\sqrt{2 \pi R}} \exp \left(- \frac{(y_k - X(t_k))^2}{2\, R} \right), \end{aligned} \tag{31}$$

with an importance distribution of the form

$$dS(t) = -b\, S(t)\, dt + q^{1/2}\, dB(t). \tag{32}$$

In the above displays a, b, q, and R are positive constants. We now obtain that the condition $\Pi_k(|w_k|^p) \le E_k < \infty$ is satisfied if by selecting the ranges of the parameters suitably. Figure 1 shows the ranges of a and b when these conditions are met with $R = 1$ and $R = 1/10$ when the other parameters are $q = 1$, $\Delta t = 1$, and $p = 4$.

7. EXAMPLE: NON-LINEAR BENEŠ MODEL

We now consider the non-linear model

$$\begin{aligned} dX(t) &= \tanh(X(t))\, dt + dB(t) \\ \rho(y_k \mid X(t_k)) &= \frac{1}{\sqrt{2 \pi R}} \exp \left(- \frac{(y_k - \theta(X(t_k)))^2}{2\, R} \right), \end{aligned} \tag{33}$$

where $\theta(\cdot)$ is a non-linear function, with an importance distribution of the form

$$dS(t) = b_k\, dt + dB(t). \tag{34}$$

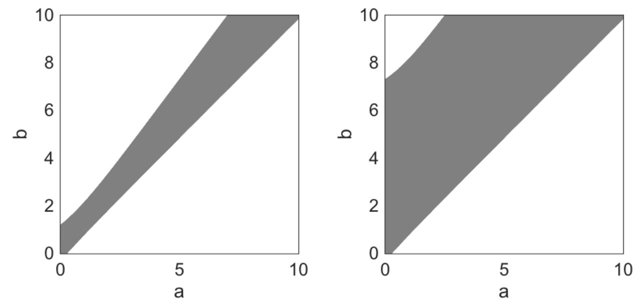

Fig. 1: Ranges of Ornstein–Uhlenbeck model parameters where the condition (27) is met (the gray area) with $R = 1$ (left) and $R = 1/10$ (right). In both the figures the lower right "forbidden" part is the result of the initial condition dependence and the upper left part depends on both p and the variance R of the measurement noise.

The above kind of importance distribution typically arises when we use an extended Kalman filter (EKF), unscented Kalman filter (UKF), or a similar method to form the importance distribution [3].

By using the closed-form transition density for the SDE in (33) [1, 19], it is easy to show that the ratio between the SDE transition densities is bounded both in x_k and x_{k-1} and thus the particle filter converges regardless of the value of b_k. It is also easy to show that the Novikov conditions are also satisfied due to boundedness of the drifts in both of the SDEs.

8. CONCLUSION AND DISCUSSION

In this article we have proved that the Girsanov theorem based particle filter proposed in [1–3] converges in L_p sense provided that a moment condition is satisfied by the likelihood ratio process and if it is bounded with respect to the step-wise initial condition. It is worth noting that the results also imply the almost sure convergence of the empirical filtering measure due to a Borel–Cantelli argument (see, e.g., [15]).

Although we have required that the moments are bounded for any x_{k-1}, in fact they only need to be bounded given $\mathcal{G}_{k-1}$, which might open up chance to relax the initial condition boundedness requirement. In this article we have also completely ignored the discretization error caused by numerical integration of the SDEs, which certainly affects convergence. However, more detailed analysis of the effect of this error is left as a future work.

9. REFERENCES

[1] S. Särkkä, *Recursive Bayesian Inference on Stochastic Differential Equations*, Doctoral dissertation, Helsinki University of Technology, Department of Electrical and Communications Engineering, 2006.

[2] S. Särkkä, "On sequential Monte Carlo sampling of discretely observed stochastic differential equations," in *Proceedings of Nonlinear Statistical Signal Processing Workshop, Cambridge, September 2006*, September 2006.

[3] S. Särkkä and T. Sottinen, "Application of Girsanov theorem to particle filtering of discretely observed continuous-time non-linear systems," *Bayesian Analysis*, vol. 3, no. 3, pp. 555–584, 2008.

[4] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, 1970.

[5] P. Del Moral, "Non-linear filtering: interacting particle resolution," *Markov processes and related fields*, vol. 2, no. 4, pp. 555–581, 1996.

[6] P. Del Moral and L. Miclo, *Branching and interacting particle systems approximations of Feynman-Kac formulae with applications to non-linear filtering*, Springer, 2000.

[7] D. Crisan and A. Doucet, "A survey of convergence results on particle filtering for practitioners," *IEEE Transactions on Signal Processing*, vol. 50, no. 3, pp. 736–746, 2002.

[8] P. Del Moral, *Feynman–Kac Formulae: Genealogical and Interacting Particle Systems with Applications*, Springer, 2004.

[9] O. Cappé, E. Moulines, and T. Rydén, *Inference in Hidden Markov Models*, Springer, 2005.

[10] A. Bain and D. Crisan, *Fundamentals of stochastic filtering*, vol. 60, Springer, 2009.

[11] X.-L. Hu, T. B. Schön, and L. Ljung, "A general convergence result for particle filtering," *IEEE Transactions on Signal Processing*, vol. 59, no. 7, 2011.

[12] P. Del Moral, *Mean field simulation for Monte Carlo integration*, Chapman & Hall/CRC, 2013.

[13] R. Douc, E. Moulines, and D. Stoffer, *Nonlinear time series: Theory, Methods and Applications with R Examples*, CRC Press, 2014.

[14] I. S. Mbalawata, *Adaptive Markov Chain Monte Carlo and Bayesian Filtering for State Space Models*, Doctoral dissertation, Lappeenranta University of Technology, 2014.

[15] I. S. Mbalawata and S. Särkkä, "Moment conditions for convergence of particle filters with unbounded importance weights," *Signal Processing*, vol. 118, pp. 133–138, 2016, arXiv preprint arXiv:1403.6585.

[16] I. V. Girsanov, "On transforming a certain class of stochastic processes by absolutely continuous substitution of measures," *Theory of Probability and its Applications*, vol. 5, pp. 285–301, 1960.

[17] P. Hall and C. C. Heyde, *Martingale limit theory and its application*, Academic press, 1980.

[18] B. Grigelionis and V. Mackevičius, "The finiteness of moments of a stochastic exponential," *Statistics & probability letters*, vol. 64, no. 3, pp. 243–248, 2003.

[19] F. E. Daum, "Exact finite-dimensional nonlinear filters for continuous time processes with discrete time measurements," in *Proceedings of 23rd Conference on Decision and Control*, 1984, pp. 16–22.