WEIGHT MOMENT CONDITIONS FOR $L^1$ CONVERGENCE OF PARTICLE FILTERS FOR UNBOUNDED TEST FUNCTIONS

Isambi S. Mbalawata
Lappeenranta University of Technology
P.O. Box 20, FI-53851 Lappeenranta, Finland

Simo Särkkä
Aalto University
P.O.Box 12200, FI-00076 Aalto, Finland

ABSTRACT

Particle filters are important approximation methods for solving probabilistic optimal filtering problems on nonlinear non-Gaussian dynamical systems. In this paper, we derive novel moment conditions for importance weights of sequential Monte Carlo based particle filters, which ensure the $L^1$ convergence of particle filter approximations of unbounded test functions. This paper extends the particle filter convergence results of Hu & Schönh & Ljung (2008) and Mbalawata & Särkkä (2014) by allowing for a general class of potentially unbounded importance weights with unbounded importance distributions. The result shows that provided that the seventh order moment is finite, then a particle filter for unbounded importance weights, moment conditions

1. INTRODUCTION

Dynamic state estimation problems are of great interest in many real life applications such as navigation, target tracking, brain imaging, spread of infectious diseases, biological processes, telecommunication, audio signal processing, stochastic optimal control, machine learning, and physical systems [1]. In these problems, the state vector at time $t$ is represented by $x_t \in \mathbb{R}^n$ and satisfies the stochastic equation

$$ x_t \sim f(x_t | x_{t-1}), \quad (1) $$

where $f(x_t | x_{t-1})$ is the transition probability density of the corresponding Markovian stochastic process modeling the dynamics of the system. At each time step $t$ we get a measurement $y_t \in \mathbb{R}^m$ from the measurement model

$$ y_t \sim g(y_t | x_t), \quad (2) $$

This work was supported by the projects 266940 and 273475, and the Finnish Centre of Excellence on Inverse Problems Research of the Academy of Finland.

Index Terms— Particle filter convergence, unbounded importance weights, moment conditions

where $g(y_t | x_t)$ is a conditional probability density modeling the distribution of measurements.

The Bayesian approach to dynamic state estimation problem involves the construction of the probability density function of $x_t$, given $y_{1:t} \triangleq (y_1, y_2, \ldots, y_t)$ [2]. This problem is known as the Bayesian filtering problem. If we denote the probability density function of $x_t$ given $y_{1:t}$ by $p(x_t | y_{1:t})$, then the construction of $p(x_t | y_{1:t})$ can be recursively done using the Bayesian filtering equations [1]

$$ p(x_t | y_{1:t-1}) = \int f(x_t | x_{t-1}) p(x_{t-1} | y_{1:t-1}) dx_{t-1}, $$
$$ p(x_t | y_{1:t}) = \frac{g(y_t | x_t) p(x_t | y_{1:t-1})}{\int g(y_t | x_t) p(x_t | y_{1:t-1}) dx_t}, \quad (3) $$

where the first step is often referred to as the prediction step (or time update) and the second step is the measurement update step (or correction step).

In this paper, we aim to analyze the theoretical convergence properties of particle filters. In such theoretical analysis, it is convenient to rewrite the Bayesian filtering equations in terms of probability measures as follows (see., e.g. [3–6]). Let $\nu$ be a measure and $\phi$ be a measurable function. Then we denote

$$ (\nu, \phi) \triangleq \int \phi \, d\nu, \quad \text{and} \quad f \phi(x) \triangleq \int f(z \mid x) \phi(z). \quad (4) $$

Let $\pi_{t \mid t-1}$ denote the measure corresponding to the probability density $p(x_t \mid y_{1:t-1})$ and $\pi_{t \mid t}$ the measure corresponding to the density $p(x_t \mid y_{1:t})$, then, using notations (4), the Bayesian filtering equations (3) can be written as

$$ (\pi_{t \mid t-1}, \phi_t) = (\pi_{t-1 \mid t-1}, f \phi_t), \quad (5) $$
$$ (\pi_{t \mid t}, \phi_t) = \frac{(\pi_{t \mid t-1}, \phi_t g)}{(\pi_{t \mid t-1}, g)}. $$

2. PARTICLE FILTERING

In the most practical cases, especially in nonlinear or non-Gaussian models, the closed form solution of (3) or (5) is intractable. Thus, several approximate methods have been proposed and the most used classes of approximate methods are...
Gaussian approximation based extended/non-linear Kalman filters (e.g., [1,2]), and sequential Monte Carlo based particle filters (e.g., [1,7,8]). In this paper we study particle filters, where the main idea is to approximate \( \pi_{t|t} \) by a weighted set of Monte Carlo samples \( \{(x^i_t, w^i_t) : i = 1, \ldots, N\} \), and, based on these samples, we can approximate the statistics of the distribution via (weighted) sample averages.

Given a set of assumptions, it is sometimes possible to show that a particle filter converges to the exact filtering distribution, when the number of particles \( N \) tends to infinity [9]. Typically, a particle filter is said to converge if the expectations of a suitable class of test functions \( \phi(.) \) converges in this limit in some suitable topology:

\[
\lim_{N \to \infty} \left( \frac{1}{N} \sum_{i=1}^{N} w^i_t \phi(x^i_t) \right) = \mathbb{E}[\phi(x_t) \mid y_{1:t}]. \tag{6}
\]

General convergence results for particle filters for test bounded functions have been given, for example, in references [3,9–15] while for unbounded test functions, results can be found in [4–6].

3. MODIFIED PARTICLE FILTER

The \( L^4 \)-convergence of particle filter in the paper [4] required the modification of standard bootstrap filter algorithm to cope with unbounded test functions. The convergence results were obtained by computing the bounds for the conditional expectation of the fourth power of error (\( L^4 \)) in the test function estimates.

These results of [4] were extended in the paper [6] to the case of more general importance distributions \( q(x_t \mid x_{t-1}, y_{1:t}) \). The results of [6] showed that with general importance distributions the (modified) particle filter converges provided that the importance weights are bounded.

The modified particle filter algorithm as presented in [6] is given in Algorithm 1. The modified particle filter is constructed such that we always have

\[
(\pi_{t|t-1}, w_t) \approx (\tilde{\pi}_{t|t-1}^N, \tilde{w}_t) = \frac{1}{N} \sum_{i=1}^{N} w^{(i)}_t \geq \gamma_t > 0, \tag{8}
\]

where \( \gamma_t > 0 \) is a chosen threshold [4,6].

In this paper, we extend the \( L^4 \) particle filter convergence proof of [6] to the case of (potentially) unbounded importance weights. We use the same techniques and some assumptions from [6], but impose a weaker assumption that the seventh order moment is finite.

4. CONVERGENCE RESULT

The convergence proof of Algorithm 1 with bounded importance weights is found in the paper [4] for bootstrap type of importance distributions and in the paper [6] for general importance distributions. Here we follow a similar path as in the proof in [6], but modify it such that we can replace the assumption on the boundedness of the important weights with a moment condition.

To guarantee the convergence, we impose the following assumptions.

Assumption 4.1. For any given \( y_{1:t}, x \) we have \( (\pi_{s|s-1}, g_s) > \gamma_s > 0, \) where \( s = 1, \ldots, t \).

Assumption 4.2. The dynamical model density \( f \) and measurement model density \( g \) are bounded, that is, there exists constants \( c_f \) and \( c_g \) such that \( \|f\| \leq c_f \) and \( \|g\| \leq c_g \), where \( \| \cdot \| \) denotes the supremum norm.

Assumption 4.3. The test function of interest \( \phi(\cdot) \) satisfies \( \sup_{x_s} |\phi(x_s)| g(y_s \mid x_s) \leq C(y_{1:s}) \).

Assumption 4.4. For any potentially unbounded importance weights \( w_t(x_t \mid x_{t-1}) \) defined as

\[
w_t(x_t \mid x_{t-1}) = \frac{g(y_t \mid x_t) f(x_t \mid x_{t-1})}{q(x_t \mid x_{t-1}, y_t)}, \tag{9}
\]

the seventh order moment \( \mathbb{E}[(w_t(x_t, x_{t-1}))^7 \mid x_{t-1}] \) is finite, where the expectation is over \( q(.) \).

---

**Algorithm 1** General Modified Particle Filter

1. Initialize the particles, \( \{x_0^{(i)}\}_{i=1}^{N} \sim \pi_0(dx_0) \)
2. Predict the particles by drawing independent samples according to
   \[
   \bar{x}_t^{(i)} \sim \frac{1}{N} \sum_{j=1}^{N} \alpha_j^{(i)} q(x_t \mid x_{t-1}^{(j)}, y_t),
   \]
   where \( \alpha^i = (\alpha_1^i, \alpha_2^i, \ldots, \alpha_N^i) \) are the weights such that
   \( \alpha_j^i \geq 0, \sum_{j=1}^{N} \alpha_j^i = 1, \sum_{i=1}^{N} \alpha_j^i = 1, \)
   and
   \[
   \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_j^{(i)} q(x_t \mid x_{t-1}^{(j)}, y_t) = \frac{1}{N} \sum_{j=1}^{N} q(x_t \mid x_{t-1}^{(j)}, y_t) \tag{7}
   \]
3. If \((1/N) \sum_{i=1}^{N} \bar{w}_t \geq \gamma_t\), proceed to step 4 otherwise return to step 2. Note that \( \bar{w}_t \) is the value computed at \( \bar{x}_t^{(i)} \).
4. Rename \( \bar{x}_t^{(i)} = \bar{x}_t^{(i)} \), and compute the importance weights \( \{w_t^{(i)}\}_{i=1}^{N} \) at \( \bar{x}_t^{(i)} \), and then normalize them.
5. Resample, \( x_t^{(i)} \sim \pi_{t|t}^N(dx_t) = \sum_{i=1}^{N} \bar{w}_t^{(i)} \delta_{\bar{x}_t^{(i)}}(dx_t) \)
6. Set \( t = t + 1 \) and repeat from step 2.
We now present the following convergence theorem, which shows the bound for error of the fourth moment conditional mean.

**Theorem 4.5.** Consider the modified particle filter in Algorithm 1 and suppose that Assumptions 4.1–4.4 are satisfied. Then

i. For sufficiently large \( N \), the algorithm will not run into an infinite loop in steps 2-3.

ii. Let \( L_i^4(g) \) be the class of functions satisfying Assumption 4.3. For any \( \phi \in L_i^4(g) \), there exists a constant \( c \) independent of \( N \) such that

\[
E \left[ \left( (\pi_{ti}^N, \phi) - (\pi_{ti}, \phi) \right)^4 \right] \leq c \frac{\|\phi\|^4_{L_4}}{N^2},
\]

where

\[
\|\phi\|_{L_4} = \max \left\{ 1, (\pi_{si}, |\phi|^4)^{1/4}, s = 0, 1, \ldots, t \right\}.
\]

Proof. The proofs for initialization and resampling steps are the same as in [4]. Therefore, here, we only prove the convergence of the (combined) prediction and update steps as in [6].

That is, we prove the convergence of

\[
(\pi_{ti}^N, \phi) - (\pi_{ti}, \phi) = \left( \tilde{\pi}_{ti}^N, \phi \right) - \left( \tilde{\pi}_{ti}, \phi \right),
\]

where \( \tilde{\pi}_{ti}^N \) is the projection of \( \pi_{ti}^N \) onto \( \pi_{ti} \) and \( \tilde{\pi}_{ti} \) is the projection of \( \pi_{ti}^N \) onto \( \pi_{ti} \).

This is attained by finding the bounds for the following terms:

\[
E \left[ \left( (\pi_{ti}^N, \phi) - (\pi_{ti}, \phi) \right)^4 \right] \quad \text{and} \quad E[(\pi_{ti}^N, |\phi|^4)].
\]

As in [6], it is enough to find the bounds for the following terms:

\[
E \left[ \left( (\tilde{\pi}_{ti}^N, \phi) - (\tilde{\pi}_{ti}, \phi) \right)^4 \right] \quad \text{and} \quad E[(\tilde{\pi}_{ti}^N, |\phi|^4)].
\]

and

\[
E \left[ \left( (\pi_{ti}^N, 1) - (\pi_{ti}, 1) \right)^4 \right] \quad \text{and} \quad E[(\pi_{ti}, |\phi|^4)].
\]

We only study the boundedness of (13). The bounds for (14) are obtained by setting \( \phi = 1 \) in (13). We denote \( \mathcal{F}_{t-1} \) as the \( \sigma \)-algebra generated by \( x_{t-1} \). We write \( (\tilde{\pi}_{ti}^N, \phi) - (\tilde{\pi}_{ti}, \phi) \) as \( \Pi_1 + \Pi_2 + \Pi_3 \), where

\[
\Pi_1 = (\tilde{\pi}_{ti}^N, \phi) - \frac{1}{N} \sum_{i=1}^{N} E[\phi(x_{ti}^i) w(x_{ti}^i, x_{t-1}) | \mathcal{F}_{t-1}],
\]

\[
\Pi_2 = \frac{1}{N} \sum_{i=1}^{N} E[\phi(x_{ti}^i) w(x_{ti}^i, x_{t-1}) | \mathcal{F}_{t-1}]
\]

\[
- \frac{1}{N} \sum_{i=1}^{N} (\pi_{ti}^N, \phi) - (\pi_{ti}, \phi) + (\pi_{ti}^N, \phi) - (\pi_{ti}, \phi),
\]

\[
\Pi_3 = \frac{1}{N} \sum_{i=1}^{N} (\pi_{ti}^N, \phi) - (\pi_{ti}, \phi).
\]

Let \( x_{ti}^i \sim (\pi_{ti}^N, \phi) \), then

\[
E[\phi(x_{ti}^i) w(x_{ti}^i, x_{t-1}) | \mathcal{F}_{t-1}] = (\pi_{ti}^N, \phi) \chi_{\mathcal{F}_{t-1}}.
\]

We next compute the bounds for \( E[(\pi_{ti}^N, \phi) w(x_{ti}^i, x_{t-1}) | \mathcal{F}_{t-1}] \) as in [6]. For \( E[(\pi_{ti}^N, \phi) w(x_{ti}^i, x_{t-1}) | \mathcal{F}_{t-1}] \) and Equation (18) to get

\[
E[(\pi_{ti}^N, \phi) w(x_{ti}^i, x_{t-1}) | \mathcal{F}_{t-1}]
\]

\[
\leq \frac{2^4}{N^4} \sum_{i=1}^{N} \left[ \phi(x_{ti}^i) w(x_{ti}^i, x_{t-1}) \right]^4 | \mathcal{F}_{t-1}
\]

\[
+ \frac{2^4}{N^4} \sum_{i=1}^{N} \left[ \phi(x_{ti}^i) w(x_{ti}^i, x_{t-1}) \right]^2 | \mathcal{F}_{t-1}
\]

\[
\leq \frac{2^4}{N^4(1-\epsilon)^2} \sum_{i=1}^{N} \left[ \phi(x_{ti}^i) w(x_{ti}^i, x_{t-1}) \right]^4 | \mathcal{F}_{t-1}
\]

\[
+ \frac{2^4}{N^4(1-\epsilon)^2} \sum_{i=1}^{N} \left[ \phi(x_{ti}^i) w(x_{ti}^i, x_{t-1}) \right]^2 | \mathcal{F}_{t-1}
\]

From Assumption 4.4, we can deduce the following.

**Lemma 4.6.** Provided that \( E[(u_{ti}(x_{ti}^i, x_{t-1}))^7 | x_{t-1}] \) is bounded, then \( E[(u_{ti}(x_{ti}^i, x_{t-1}))^5 | x_{t-1}] \) is bounded also.

Proof.

\[
E[(w_{ti}(x_{ti}^i, x_{t-1}))^7 | \mathcal{F}_{t-1}]
\]

\[
\leq \sup_{x_{ti}} \left( E[(u_{ti}(x_{ti}^i, x_{t-1}))^7 | x_{t-1}] \right) \leq \sup_{x_{ti}} C_w \leq C_w.
\]

**Remark 4.7.** If the seventh moment is finite then the lower moments are finite too.

Proof. Results are easily obtained from Hölder’s and Jensen’s inequalities.

With Lemma 4.6 and the Cauchy–Schwarz inequality, we get

\[
E \left[ \phi(x_{ti}^i) w(x_{ti}^i, x_{t-1}) | \mathcal{F}_{t-1} \right]
\]

\[
\leq \sqrt{E \left[ \phi^8(x_{ti}^i) w(x_{ti}^i, x_{t-1}) | \mathcal{F}_{t-1} \right]} \times \sqrt{E \left[ w^7(x_{ti}^i, x_{t-1}) | \mathcal{F}_{t-1} \right]}
\]

\[
\leq C_w^{7/2} \sqrt{E \left[ \phi^8(x_{ti}^i) w(x_{ti}^i, x_{t-1}) | \mathcal{F}_{t-1} \right]}.
\]

\[
E \left[ \phi(x_{ti}^i) w(x_{ti}^i, x_{t-1}) | \mathcal{F}_{t-1} \right]
\]

\[
\leq \sqrt{E \left[ \phi^4(x_{ti}^i) w(x_{ti}^i, x_{t-1}) | \mathcal{F}_{t-1} \right]} \times \sqrt{E \left[ w^4(x_{ti}^i, x_{t-1}) | \mathcal{F}_{t-1} \right]}
\]

\[
\leq C_w^{3/2} \sqrt{E \left[ \phi^4(x_{ti}^i) w(x_{ti}^i, x_{t-1}) | \mathcal{F}_{t-1} \right]}.
\]
Thus

\[ E[|\Pi_1|^4 | F_{t-1}] \]

\[ \leq \frac{2^4 C_{w}^7}{N^4(1 - \epsilon)^2} \sum_{i=1}^{N} \sqrt{E[|\phi^8(x_i^t) w(x_i^t, x_{i-1}^t)| | F_{t-1}]} \]

\[ + \frac{2^4 C_{w}^3}{N^4(1 - \epsilon)^2} \left( \sum_{i=1}^{N} \sqrt{E[|\phi^4(x_i^t) w(x_i^t, x_{i-1}^t)| | F_{t-1}]} \right)^2. \]

But

\[ \sum_{i=1}^{N} \sqrt{E[|\phi^8(x_i^t) w(x_i^t, x_{i-1}^t)| | F_{t-1}]} \leq N + \sum_{i=1}^{N} E[|\phi^8(x_i^t) w(x_i^t, x_{i-1}^t)| | F_{t-1}] \]

\[ \leq N + \sum_{i=1}^{N} E[|\phi^4(x_i^t) w(x_i^t, x_{i-1}^t)| | F_{t-1}] \]

\[ \left( \sum_{i=1}^{N} \sqrt{E[|\phi^4(x_i^t) w(x_i^t, x_{i-1}^t)| | F_{t-1}]} \right)^2 \leq (1 + N) \sum_{i=1}^{N} E[|\phi^4(x_i^t) w(x_i^t, x_{i-1}^t)| | F_{t-1}]. \]

Then

\[ E[|\Pi_1|^4 | F_{t-1}] \]

\[ \leq \frac{2^4 C_{w}^7}{N^2(1 - \epsilon)^2} + \frac{2^4 C_{w}^7}{N^2(1 - \epsilon)^2} \sum_{i=1}^{N} E[|\phi^8(x_i^t) w(x_i^t, x_{i-1}^t)| | F_{t-1}] \]

\[ + \frac{2^4 C_{w}^3}{N^2(1 - \epsilon)^2} \sum_{i=1}^{N} E[|\phi^4(x_i^t) w(x_i^t, x_{i-1}^t)| | F_{t-1}] \]

\[ \leq \frac{2^4 C_{w}^7}{N^4(1 - \epsilon)^2} + \frac{2^4 C_{w}^3}{N^4(1 - \epsilon)^2} \sum_{i=1}^{N} E[|\phi^4(x_i^t) w(x_i^t, x_{i-1}^t)| | F_{t-1}] \]

\[ = \hat{C}_{\Pi_1} \frac{|\phi|^4_{t-1,4} N^2}{N^2}. \]

The computation of bounds for \( E[|\Pi_2|^4] \) and \( E[|\Pi_3|^4] \) is the same as in [6]. Therefore

\[ E\left[ |\Pi_2|^4 | F_{t-1} \right] \leq \hat{C}_{\Pi_2} \frac{|\phi|^4_{t-1,4}}{N^2}. \]

\[ E\left[ |\Pi_3|^4 | F_{t-1} \right] \leq \hat{C}_{\Pi_3} \frac{|\phi|^4_{t-1,4}}{N^2}. \]

By combining Equations (19), (20), and (21) via Minkowski’s inequality, we get

\[ E \left[ |(\hat{\pi}_t|^2, \phi) - (\hat{\pi}_t|^2, \phi)|^4 \right]^\frac{1}{4} \leq \hat{C}_{\Pi}^{1/4} \frac{|\phi|^4_{t-1,4}}{N^{1/2}}, \]

which implies

\[ E \left[ |(\hat{\pi}_t|^2, \phi) - (\hat{\pi}_t|^2, \phi)|^4 \right] \leq \hat{C}_{\Pi} \frac{|\phi|^4_{t-1,4}}{N^2}. \] (22)

From [6], the bound for \( E[|\hat{\pi}_t|^2]| | F_{t-1}| \) is

\[ E \left[ |(\hat{\pi}_t|^2, \phi)|^4 \right] \leq M_{\Pi^2} |\phi|^4_{t-1,4}. \] (23)

Note that if we set \( \phi = 1 \) in (22) and (23), we get bounds for (14). Hence the remaining task is to find the bounds for (12), which is done exactly the same way as in [6]. Thus

\[ E \left[ |(\hat{\pi}_t|^2, \phi) - (\pi_t|^2, \phi)|^4 \right] \leq C_{\Pi^2} |\phi|^4_{t-1,4}, \]

\[ E \left[ |(\hat{\pi}_t|^2, \phi)|^4 \right] \leq M_{\Pi^2} |\phi|^4_{t-1,4}, \]

which complete the proof of Theorem 4.5.

5. NUMERICAL EXAMPLE

A relevant question is now what is the actual benefit of the current extension in practical particle filtering models. The clear benefit is that it extends the class of allowed importance distributions to the class which does not ensure that the importance weights are uniformly bounded. For example, the weights might become infinite in isolated points provided that the required expectations of them remain bounded.

However, to get an idea what kind of importance weights have this kind of property, consider

\[ v(x) = |x|^{-1/2} \exp(-|x|). \] (24)

Clearly this function is everywhere positive, but it also contains an infinite value at \( x = 0 \), and hence it is not bounded. However, its integral is finite, which can be seen by computing its integral by reducing it into the definition of the Gamma function:

\[ \int_{-\infty}^{\infty} |x|^{-1/2} \exp(-|x|) dx = 2 \sqrt{\pi}. \] (25)

The function defined by (24) is thus an example of a positive function which is unbounded, but has a bounded integral (see Figure 1). It is now easy to see that it is possible to construct models for which the importance weights are point-wise unbounded but still satisfy Assumption 4.4. Examples of practical models which lead to this kind of importance weights will be considered in future work.

6. CONCLUSION

In this paper, we have extended the \( L^2 \) particle filter convergence proof of [6] to the case of potentially unbounded importance weights, by replacing the boundedness condition with
Fig. 1. Example of a point-wise unbounded function with a finite integral over $(-\infty, \infty)$.

...finiteness of conditional weight moments. Our proof shows that provided that the seventh order moment is finite, then a particle filter for unbounded test functions with unbounded importance weights are ensured to converge.

REFERENCES


