Particle Filters with Random Resampling Times

D. Crisan,* O. Obanubi†
Department of Mathematics, Imperial College London
180 Queen's Gate, London, SW7 2AZ, United Kingdom
{d.crisan,oobanubi}@imperial.ac.uk

Abstract

Particle filters are numerical methods for approximating the solution of the filtering problem which use systems of weighted particles that (typically) evolve according to the law of the signal process. These methods involve a corrective/resampling procedure which eliminates the particles that become redundant and multiplies the ones that contribute most to the resulting approximation. The correction is applied at instances in time called resampling/correction times. Practitioners normally use certain overall characteristics of the approximating system of particles (such as the effective sample size of the system) to determine when to correct the system. As a result, the resampling times are random. However, in the continuous time framework, all existing convergence results apply only to particle filters with deterministic correction times. In this paper, we analyse (continuous time) particle filters where resampling takes place at times that form a sequence of (predictable) stopping times. We prove that, under very general conditions imposed on the sequence of resampling times, the corresponding particle filters converge. The conditions are verified when the resampling times are chosen in accordance to effective sample size of the system of particles, the coefficient of variation of the particles' weights and, respectively, the (soft) maximum of the particles' weights. We also deduce central-limit theorem type results for the approximating particle system with random resampling times.

Keywords: Stochastic partial differential equation. Filtering. Zakai equation. Particle filters. Sequential Monte-Carlo Methods. Resampling. Resampling times. Random times. Effective Sample Size. Coefficient of variation. Soft Maximum. Central Limit Theorem.
Subject Classification: 60F05 (60F25, 60G35, 60G40, 60H35, 93E11).

* The work of D. Crisan was partially supported by the EPSRC Grant No: EP/H0005500/1.
† The work of O.Obanubi was supported by an EPSRC Doctoral Training Award.

1 Introduction

The filtering problem involves the estimation of the current state of an evolving dynamical system based on partial observation. The evolution of the dynamical system is customarily modelled by a stochastic process $X = \{X_t,\ t \geq 0\}$ called the *signal* process, where the temporal parameter t runs over the positive half line $[0, \infty)$. The signal process X can not be measured directly. However, a partial measurement of the signal can be obtained. This measurement is modelled by another continuous time process $Y = \{Y_t,\ t \geq 0\}$ which is called the *observation* process. The observation process is a function of X and a measurement noise. The measurement noise is modelled by a stochastic process $W = \{W_t,\ t \geq 0\}$. Hence,

$$Y_t = f_t(X_t, W_t) \qquad t \in [0, \infty).$$

Let $\mathcal{Y} = \{\mathcal{Y}_t,\ t \geq 0\}$ be the filtration generated by the observation process Y; namely, $\mathcal{Y}_t = \sigma\left(Y_s,\ s \in [0,t]\right)$, for $t \geq 0$. Then the filtering problem consists in computing π_t, the conditional distribution of X_t given $\mathcal{Y}_t$. The process $\pi = \{\pi_t,\ t \geq 0\}$ is a $\mathcal{Y}_t$-adapted probability measure valued process, so that

$$\mathbb{E}\left[\varphi(X_t) \mid \mathcal{Y}_t\right] = \int \varphi(x)\pi_t(dx),$$

for all statistics φ for which both terms of the above identity make sense. Generally speaking the filtering problem can not be solved analytically: An explicit formula cannot be obtained for the conditional distribution π_t. Only in specific cases such as the Kalman-Bucy filter and the Benes filter (see, eg Chapter 6 in [1]) is this not true. Numerical methods, of which particle filters are an example, are thus employed to obtain approximations to the solution of the filtering problem.

Particle filters[1] are numerical methods that produce an approximation of π_t using empirical distributions of systems of evolving weighted particles. They are currently one of the most successful methods used to approximate the solution of the filtering problem (see [6] or Chapter VIII in [4] for an overview). The particles evolve according to the law of the signal process X and carry a weight proportional with the likelihood of their recent position/trajectory given the observation data. As time progresses, some of the weights diminish and so the corresponding particles essentially contribute less to the approximation process. In order to counter this phenomenon known as sample degeneracy, a correction procedure is introduced at particular times to cull the redundant particles and multiply the particles that contribute more significantly to the approximation process. This correction procedure is known as resampling and it was first introduced in the papers by Gordon, Salmon and Ewing [9], Gordon, Salmond and Smith [10], Kitagawa [13]. These resampling/correction times are chosen in an adaptive manner and are usually determined by certain overall characteristics of the approximating particle system. One such characteristic (for which the results from below apply) is the effective sample size of the approximating particle system.

In the last fifteen years we have witnessed a rapid development of the theory of particle filters. The discrete time framework has been extensively studied and a multitude of convergence and stability results have been proved. A

[1] These methods are also known under the name of Sequential Monte Carlo Methods in the Statistics and the Engineering literature

comprehensive description of these developments in the wider context of approximations of Feynman–Kac formulae can be found in Del Moral [18]. Results concerning particle filters for the continuous time filtering problem are far fewer that their discrete counterparts. For an up-to-date overview of these results see Chapter VIII in [4].

1.1 Contribution of the paper

This paper studies particle filters that use a standard resampling procedure in a continuous time setting. It investigates the convergence of these approximations to the solution of the filtering problem. In contrast with existing results in the published literature which cover only particle filters with deterministic correction times[2], we assume that the approximating particle system has *random* resampling times. This is the current practice in the area: For example, the resampling times can be chosen to be the times at which the effective sample size of the systems falls below a desired threshold. The paper produces the theoretical justification for such practice, hence it addresses this gap in the literature.

We prove in this paper that the empirical distribution of the system converges to π_t (see Theorem 1) if resampling times occur at times that form a sequence of (predictable) stopping times and satisfy a mild integrability assumption. The proof of the result is a lot harder than the proof of the convergence of particle filters with deterministic stopping time. The difficulty stems from the fact that the randomness of the resampling times prevents us from using the standard approach based on the dual of the conditional distribution process. We circumvented this by developing first an abstract convergence criterion for measure-valued processes, see Theorem 8. This result may be of interest independently of the current application. The conditions under which Theorem 1 holds are verified for the case when the resampling times are chosen in accordance to the effective sample size of the system of particles, or equivalently the coefficient of variation of the particles' weights and, respectively, in accordance to the (soft) maximum of the particles' weights. We emphasize that we do not require that the resampling times converge. This is particularly important as, usually, the times depend on the approximation itself and so we cannot assume a priori that they converge.

We also analyze the fluctuations of adaptive particle filters. Under an additional integrability condition and after assuming that resampling times converge (as a result of the convergence of the approximations), we show a central-limit theorem type result is obtained for the approximating system, see Theorem 2. The conditions are again checked for the effective sample size and, respectively, the maximum of the particles' weights criteria. The central-limit theorem will enable us to perfom a comparative analysis of adaptive particle filters. This and other issues related to the implementation of particle filters with random resampling times (time discretization of the particles' motion and particles' weights, computational effort, etc) will be discussed in sequel to this paper.[3]

[2] A similar result has been proved in the discrete time setting. See [5] for more details and the next Section for a comparison with the results presented here.

[3] D. Crisan, O. Obanubi, *Threshold Inferences and Numerical Results Concerning Random Time Resampling for the Effective Sample Size.*

The following is a summary of the contents of the paper:

In the next section, the filtering framework and the filtering problem are formally introduced and defined. Some background and preliminary results of stochastic filtering theory will also be covered. Key among these results is the Zakai equation, a linear equation which describes the evolution of an unnormalized version of the conditional distribution of the signal. The Zakai equation, as will be seen throughout this paper, plays a fundamental role in allowing approximations of the solution of the filtering problem to be obtained. This is because it provides us with an indirect and relatively easier method, due to its linear form, of obtaining convergence results for the normalized conditional distribution of X. We also state the main results of the paper and the conditions under which they hold.

In Section 3, the class of particle approximation is introduced and discussed. Details of the approximating particle system and how resampling times are determined are given. We discuss the suitability of choosing the resampling times to be determined by various measures of sample degeneracy: the essential sample size, the coefficient of variation, entropy and the maximum weight.

The next two sections contain the proofs for the main results of the paper. In section 4, the evolution equations of the approximating measures are derived and used to show the almost sure convergence of the approximations to the true solutions under certain conditions. In section 5 we get central limit theorem type results. The error between the approximations and the true solutions are recalibrated and shown to form a tight sequence and their limit in distribution found.

The paper is concluded with an Appendix that collates a number of useful lemmatas and results used throught the paper.

1.2 Notation

In the following we will use the following notation:

- $\mathbb{R}^d$ - the d-dimensional Euclidean space
- $\overline{\mathbb{R}^d}$ - the one-point compactification of $\mathbb{R}^d$ formed by adding a single point at infinity to $\mathbb{R}^d$
- $B(\mathbb{R}^d)$ - the space of bounded Borel measurable functions from $\mathbb{R}^d$ to $\mathbb{R}$
- $\mathcal{C}_b(\mathbb{R}^d)$ - the space of bounded continuous functions on $\mathbb{R}^d$
- $\mathcal{C}_b^m(\mathbb{R}^d)$ - the space of bounded continuous functions on $\mathbb{R}^d$ with bounded derivatives up to order $m \in \mathbb{N}$.
- $\mathcal{C}_0^m\left(\mathbb{R}^d\right)$ - the space of continuous functions on $\mathbb{R}^d$, vanishing at infinity with continuous partial derivatives up to order $m \in \mathbb{N}$
- $\mathcal{C}_0^\infty\left(\mathbb{R}^d\right)$-the space of smooth functions on $\mathbb{R}^d$ vanishing at infinity
- $\|\cdot\|_\infty$ - the supremum norm; for $\varphi : \mathbb{R}^d \to \mathbb{R}$, $\|\varphi\|_\infty = \sup_{x\in\mathbb{R}^d} |\varphi(x)|$

- $\|\cdot\|_{m,\infty}$ - the norm such that for $m \in \mathbb{N}$ and a function φ on $\mathbb{R}^d$

$$\|\varphi\|_{m,\infty} = \sum_{|\alpha|\leq m} \sup_{x\in\mathbb{R}^d} |D_\alpha \varphi(x)|,$$

where $\alpha = (\alpha^1, \ldots, \alpha^d)$ is a multi-index and $D_\alpha \varphi = (\partial_1)^{\alpha^1} \cdots (\partial_d)^{\alpha^d} \varphi$

- $\mathcal{M}_F(\mathbb{R}^d)$ - the set of finite measures on $\mathbb{R}^d$
- $\mathcal{M}_F(\overline{\mathbb{R}^d})$ - the set of finite measures on $\overline{\mathbb{R}^d}$
- $D_{\mathcal{M}_F(\mathbb{R}^d)}[0,T]$ - the space of càdlàg functions $f : [0,T] \mapsto \mathcal{M}_F(\mathbb{R}^d)$
- $D_{\mathcal{M}_F(\mathbb{R}^d)}[0,\infty)$ - the space of càdlàg functions $f : [0,\infty) \mapsto \mathcal{M}_F(\mathbb{R}^d)$

2 Filtering problem and Related Results

2.1 The Filtering Framework

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a complete filtered probability space which satisfies the usual conditions. Within $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ we consider an $\mathcal{F}_t$-adapted d-dimensional signal process $X = \{X_t : t \geq 0\}$ which solves the stochastic differential equation:

$$X_t^i = X_0^i + \int_0^t f^i(X_s)\,\mathrm{d}s + \sum_{j=1}^p \int_0^t \sigma^{ij}(X_s)\,\mathrm{d}V_s^j \quad i = 1, \ldots, d, \tag{2.1}$$

where $V = (V^j)_{j=1}^p$ is a p-dimensional Brownian motion. We assume that $f = (f^i)_{i=1}^d : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma = (\sigma^{ij})_{i=1,\ldots,d,j=1,\ldots,p} : \mathbb{R}^d \to \mathbb{R}^{d\times p}$ are globally Lipschitz. Let A be the infinitesimal generator associated with X, that is

$$A = \sum_{i=1}^d f^i \frac{\partial}{\partial x_i} + \sum_{i,j=1}^d a^{ij} \frac{\partial^2}{\partial x_i \partial x_j}, \tag{2.2}$$

and $a^{ij} = \frac{1}{2}\sum_{k=1}^p \sigma^{ik}\sigma^{jk} = \frac{1}{2}(\sigma\sigma^\top)^{ij}$ for all $i, j = 1, \ldots, d$. We denote by $\mathcal{D}(A)$ the domain of A. Next, let W be a standard $\mathcal{F}_t$-adapted m-dimensional Brownian motion defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ independent of X, and let Y be the process satisfying the following evolution equation

$$Y_t = \int_0^t h(X_s)\,\mathrm{d}s + W_t, \tag{2.3}$$

where $h = (h_i)_{i=1}^m : \mathbb{R}^d \to \mathbb{R}^m$ is globally Lipschitz. Let $\{\mathcal{Y}_t : t \geq 0\}$ be the usual augmentation with null sets of the filtration associated with the process Y. The filtering problem consists of determining the conditional distribution process π_t of the signal X_t given the filtration $\mathcal{Y}_t$, that is

$$\pi_t(\varphi) = \mathbb{E}\left[\varphi(X_t)|\mathcal{Y}_t\right],\ \varphi \in B(\mathbb{R}^d). \tag{2.4}$$

Then we have

$$\pi_t(\varphi) = \frac{\rho_t(\varphi)}{\rho_t(\mathbf{1})},\ \varphi \in B(\mathbb{R}^d), \tag{2.5}$$

where $\rho_t(\varphi)$ is an $\mathcal{Y}_t$-adapted measure-valued process which satisfies the Zakai equation:

$$\rho_t(\varphi) = \pi_0(\varphi) + \int_0^t \rho_s(A\varphi)\,\mathrm{d}s + \sum_{i=1}^m \int_0^t \rho_s(\varphi h^i)\,\mathrm{d}Y_s^i,\ \ \varphi \in \mathcal{D}(A). \tag{2.6}$$

Formula (2.5) is called the Kallianpur-Striebel's formula and the process $\rho = \{\rho_t, t \geq 0\}$ is called the unnormalized conditional distribution of the signal. The Zakai equation can also be written in *mild* form (see, for example, [25]):

$$\rho_t(\varphi) = \rho_0(P_t\varphi) + \int_0^t \rho_r(P_{t-r}\varphi h^\top)\,\mathrm{d}Y_r, \quad \tilde{\mathbb{P}} - a.s.\ \forall t \geq 0. \tag{2.7}$$

$(P_r)_{r\geq 0}$ is the Markov $\mathcal{C}_0$-semigroups of contractions whose infinitesimal generator is the operator A as defined in (2.2). The mild form of the Zakai equation holds true for any $\varphi \in C_b(\mathbb{R}^d)$. We will assume throughout the paper that the coefficients σ and f in (2.1) are bounded and continuously differentiable with bounded partial derivatives and h in (2.3) is bounded and Lipschitz.

In the following we will analyse a particle filter with multinomial resampling times $(T_k^n)_{k\geq 0}$ that form a strictly increasing sequence of predictable stopping times.[4] We give details of the particle filter in the following section. We denote by $\pi^n = \{\pi_t^n,\ t \geq 0\}$ the process consisting of the empirical distribution of the particle system and by N_t^n, the number of resampling instances that occur before time t. The convergence of π^n is stated in the following

Theorem 1. *If there exists $p > 1$ such that for all $t > 0$, we have*

$$\sup_{n>0} \mathbb{E}[(N_t^n)^p] < \infty \tag{2.8}$$

then, for any $r < p$, there exists a constant $\alpha = \alpha(T, r)$, independent of n, such that for any $\varphi \in \mathcal{C}_b^1(\mathbb{R}^d)$, we have

$$\sup_{t\in[0,T]} \mathbb{E}\left[(\pi_t^n(\varphi) - \pi_t(\varphi))^r\right] \leq \frac{\alpha}{n^p} \|\varphi\|_{1,\infty}^r. \tag{2.9}$$

A few remarks regarding Theorem 1 are in order. Once again, we emphasize that we do not assume that the resampling times converge as typically they will depend on π_n and therefore their convergence cannot be a priori assumed.

Condition (2.8) implies that $\lim_{k\to\infty} T_k^n = \infty$. In particular, that there are only a finite number of resampling times in any finite interval. It is trivially satisfied for any sequence of *deterministic* times that converge to ∞. Therefore, Theorem 1 generalizes existing convergence results for (non-adaptive) particle filters. In addition, condition (2.8) is

[4] That is for each T_k^n, there exists an announcing sequence of stopping times $(T_k^{n,m})_{m\geq 1}$ such that $T_k^{n,m}$ is increasing, $T_k^{n,m} < T_k^n$ on $\{T_k^n > 0\}$, for all m, and $\lim_{m\to\infty} T_k^{n,m} = T_k^n$. See, for example, Chapter III Section 2 in [20] for more on predictable stopping times.

satisfied for the case when the resampling times are chosen in accordance to the effective sample size of the system of particles, or, equivalently, the coefficient of variation of the particles' weights. This is the most popular resampling criteria among the practitioners. It is also satisfied for the case when the resampling times are chosen in accordance to the (soft) maximum of the particles' weights.

In [5], a similar convergence result has been proved for the particle filters in a discrete-time framework. The result in [5] relies on a ingenious coupling argument. The particle filter with random resampling times $(T_k^n)_{k\geq 0}$ is coupled with one with resampling times $(\bar{T}_k)_{k\geq 0}$ where $(\bar{T}_k)_{k\geq 0}$ can be any deterministic times or times that could depend on the observation process only (and not on the current state of the particle filter). The authors show that, as n increases, T_k^n are exponentially close to $\bar{T}_k$. Since time runs discreetly, they must be equal with high probability and the convergence result follows by analyzing the particle filter with observation dependent resampling times. Of course this argument cannot be applied in a continuous-time framework. In continuous time, the corresponding equivalent of T_k^n and $\bar{T}_k$ can be different no matter how close they are. It would be interesting to see if the argument presented here can be adapted to cover the discrete framework. Condition (2.8) is trivially satisfied when time runs discretely: obviously $N_t^n \leq t$, with the maximum achieved when one resamples at any time instance. An adaptation of the proof presented here would, perhaps, solve the additional constraint imposed in [5] on the T_k^n's that involve the use of certain randomized criteria thresholds (see section 5.2 in [5]).

Condition (2.8) offers a control on the number of resampling times in any finite interval. Heuristically speaking, there must not be "too many of them". In the case when condition (2.8) is not satisfied, the convergence might be slower or, even worse, the particle filter might diverge as the number of particles increases. Theorem 1 is also valid under following alternative to condition (2.8), see [19] for details: There exists $p > 1$ such that for all $t > 0$, we have

$$\sup_{n>0} \sum_{k=1}^{\infty} \mathbb{P}(T_k^n \leq t)^{1/p} < \infty.$$

In addition, to the convergence of π_n we also study its fluctuations around the limiting measure π. In particular, if $\bar{U}^n = \{\bar{U}_t^n,\ t \geq 0\}$ is the measure-valued process defined as $\bar{U}^n = \sqrt{n}(\pi^n - \pi)$, then we deduce the following central limit type theorem.

Theorem 2. *Assume that for any $k \geq 0$, $\lim_{n\to\infty} T_k^n = T_k$, where $(T_k)_{k\geq 0}$ is a strictly increasing sequence of $\mathcal{Y}_t$-adapted predictable stopping times. If (2.8) is satisfied and also there exists $p > 1$ such that for all $t > 0$, we have*

$$\lim_{\delta\to 0} \sup_{n>0} \mathbb{E}[\sup_{s\in[0,t]} \mathbb{E}[(N_{s+\delta}^n - N_s^n)^p | \mathcal{F}_s]] = 0, \tag{2.10}$$

then there exists a measure valued process $\bar{U} = \{\bar{U}_t,\ t \geq 0\}$ such that $\bar{U}^n$ converges in distribution to $\bar{U}$.

Observe that Theorem 2 requires the convergence of the resampling times in order to hold true. If the criteria for choosing the resampling times are functions of π^n then their convergence can be deduced from Theorem 1. Condition (2.10) (the tightness condition) plays the central role in controlling the oscillation of the paths of the converging processes. Heuristically, it says that the resampling times cannot accumulate locally. Condition (2.10) is, again, satisfied for the case when the resampling times are chosen in accordance to the effective sample size of the system of particles, or, equivalently, the coefficient of variation of the particles' weights and for the case when the resampling times are chosen in accordance to the (soft) maximum of the particles' weights.

3 The Approximating Particle System

The particle system consists initially of n particles each with weight $1/n$ and position $v_j^n(0)$, $j = 1, \ldots, n$. The positions of the particles are chosen to be independent, identically distributed (i.i.d.) random variables with common distribution π_0 which is the law of X_0 the signal at time 0. Hence, the approximating measure at time 0 is

$$\pi_0^n = \frac{1}{n} \sum_{j=1}^{n} \delta_{v_j^n(0)}.$$

Let $\{T_k^n\}_{k \in \mathbb{N}}$ be a strictly increasing sequence of predictable stopping times. For ease of notation, we write T_k instead of T_k^n unless when necessary to emphasise the dependency of the predictable stopping times on the sample size, n.

During the random time intervals $[T_k, T_{k+1})$, the particles move with the same law as the signal X; that is for any stopping time $T \in [T_k, T_{k+1})$

$$v_j^n(T) = v_j^n(T_k) + \int_{T_k}^{T} f(v_j^n(s)) \, \mathrm{d}s + \int_{T_k}^{T} \sigma(v_j^n(s)) \, \mathrm{d}V_s^{(j)}, \quad j = 1, \ldots, n, \tag{3.1}$$

where $(V^{(j)})_{j=1}^n$ are mutually independent $\mathcal{F}_t$-adapted p-dimensional Brownian motions which are independent of Y and independent of all other random variables in the system. Each particle is assigned a normalized weights $\bar{a}_j^n(T)$, $j = 1, \ldots, n$, for arbitrary stopping time $T \in [T_k, T_{k+1})$ given by

$$\bar{a}_j^n(T) := \frac{a_j^n(T)}{\sum_{k=1}^{n} a_k^n(T)}$$

where

$$a_j^n(T) = \exp\left(\int_{T_k}^{T} h(v_j^n(s))^\top \, \mathrm{d}Y_s - \frac{1}{2} \int_{T_k}^{T} \|h(v_j^n(s))\|^2 \, \mathrm{d}s \right). \tag{3.2}$$

For $T \in [T_k, T_{k+1})$, define

$$\pi_T^n = \sum_{j=1}^{n} \bar{a}_j^n(T) \delta_{v_j^n(T)}.$$

At the end of the (random) interval $[T_k, T_{k+1})$, the correction procedure is implemented, the particles are re-indexed and their weight reinitialized to 1.

We will now address the questions concerning when and how to resample. A measure or indicator of the extent of sample degeneracy is required to inform us of when to resample. The effective sample size (see [7], [11], [14]) of the system is the most popular measure for the sample degeneracy of our approximating particle system. The effective sample size (or ESS or n_{eff}) cannot be calculated analytically and instead an estimate, given by

$$\hat{n}_{eff} = \frac{1}{\sum_{j=1}^{n}(\bar{a}_j^n(T))^2}, \tag{3.3}$$

is used. The interpretation of the effective sample size is that any inference based on a weighted sample of size n will be approximately as accurate as one based on an independent sample whose size is the effective sample size. An application of the Cauchy-Schwarz inequality leads us to the (intuitive) conclusion that $\hat{n}_{eff} \leq n$. That is, the effective sample size cannot be larger than the actual sample size. Since the worst case scenario one can have is all the weight concentrated on one particle, it follows that the lower bound for n_{eff} is 1.

Resampling occurs when the ESS falls below a selected threshold n_{thres} and clearly from above, we must have $n_{thres} \leq n$. We will, however, only consider the cases where $n_{thres} \in (0, n)$, excluding the trivial cases where $n_{thres} = 0$ (i.e. resampling never occurs: this is the Monte Carlo method) and $n_{thres} = n$ (i.e. resampling occurs continuously). The threshold is set to be of the form $\lambda_{thres} n$ where $\lambda_{thres} \in (0, 1)$.

For $1 \leq k \in \mathbb{N}$ define

$$T_k := \inf\{t \geq T_{k-1} : n_{eff} \leq \lambda_{thres} n\}. \tag{3.4}$$

Consequently, T_k is the first time after the previous resampling time, T_{k-1}, that the ESS falls below the chosen threshold. Put differently, T_k is the k^{th}-resampling time and $[T_{k-1}, T_k)$ is the time interval between the the $(k-1)^{th}$ and k^{th} resampling times where the system particles evolve according to the prescribed signal law.

Before discussing how resampling is actually carried out, we give a heuristic argument to highlight the motivation behind choosing the random resampling times to be predictable stopping times. In between resampling the newly acquired information is stored in the particle weights. As long as the information remains limited or inaccurate, the weights will remain roughly equal (in particular the ESS will be close to n). In this case resampling doesn't make sense as it introduces additional randomness in the system and wouldn't compensate for this by significantly improving the system. However as soon as the information becomes 'reasonable' (thus allowing us to be able to better distinguish between particles in the 'right' and 'wrong' regions) and the weights subsequently become sufficiently uneven, resampling is then desirable to keep the particles in the 'right' region. By resampling at random times rather than, say, at regular (deterministic) time intervals, we resample only when the information is 'reasonable' enough (as determined by the ESS) and don't unnecessarily introduce redundant randomness into the system. Put differently, the resampling procedure is adjusted to the information being received and is not a priori fixed.

The predictability of the resampling times based on the ESS is immediate. Recall that, by definition, a stopping time T is predictable if there exists an announcing sequence of stopping times $(T^m)_{m\geq 1}$ such that T^m is increasing, $T^m < T$ on the set $\{T > 0\}$, and $\lim_{m\to\infty} T^m = T$. So, if for example , n_{thres} is set to $n/3$, an announcing sequence can be $(T_k^m)_{m\geq 1}$ where

$$T_k^m := \inf\left\{t > T_{k-1} : \hat{n}_{eff} < n/(3-1/m)\right\}.$$

The predictability property of the resampling time is used to give a useful characterization of the σ-algebra of events that occur up to but not including the stoping time itself. In particular, for any predictable stopping time T and any announcing sequence, (T_r), of T we have (see Chapter III Theorem 6 of [20])

$$\mathcal{F}_{T-} =: \bigvee_{r\geq 1} \mathcal{F}_{T_r} = \sigma\left\{\bigcup_{r\geq 1} \mathcal{F}_{T_r}\right\}.$$

Heuristically, this means that we can express the information immediately prior to any predictable stopping time in terms of the information generated by events leading up to it.

We will now discuss on how the resampling is performed. Recall that at the end of the interval $[T_{k-1}, T_k)$ the resampling (or correction) procedure is implemented, the particles are re-indexed and their weights reinitialized to 1. During the implementation, each particle is replaced by a random number of particles (possibly zero) with each offspring inheriting the spatial position of their parents. The question which then arises is how to replace the parent particles with offspring particles. Posed differently, what should the offspring distribution of the parent particles be? A possible answer to this question is the following:
Let $O_T^{(j)}, j = 1, \ldots, n$ be the random variables representing the number of offspring produced by the j^{th} parent particle during resampling. Let $o_j^{n,T} \in \{1, \ldots, n\}, j = 1, \ldots, n$ be the particular values of these random variables. Then one possible offspring distribution is the multinomial distribution with (respective) probabilities taken to be the normalized weight of the j^{th} parent particle, that is, $\bar{a}_j^n(T)$, so that we have

$$\mathbb{P}\left(O_T^{(1)} = o_1^{n,T}, \ldots, O_T^{(n)} = o_n^{n,T}\right) = \frac{n!}{\prod_{j=1}^n o_j^{n,T}!} \prod_{j=1}^n \left(\bar{a}_j^n(T)\right)^{o_j^{n,T}}. \tag{3.5}$$

The multinomial sampling algorithm essentially states that, at correction or resampling times, we should sample n-times (with replacement) from the population of particles with positions $v_j^n(T), j = 1, \ldots, n$ according to the probability distribution given by the corresponding normalized weights $\bar{a}_j^n(T), j = 1, \ldots, n$. $o_j^n \equiv o_j^{n,T}$ therefore is the number of times the particle with position $v_j^n(T)$ is chosen.
After carrying out the correction procedure, the unnormalized weights of the particles are re-initialized to 1. A particle filter with this choice of offspring distribution is called a Bootstrap Filter or the Sampling Importance Resampling algorithm (SIR algorithm). It can be traced back to the papers by Gordon, Salmon and Ewing [9], Gordon, Salmond and Smith [10], Kitagawa [13]. The bootstrap filter is popular among practitioners because it is quick and easy to implement and amenable to parallelisation.

3.1 Predictable Stopping Times for Other Measures of Sample Degeneracy

Before proceeding further, we will give examples of predictable stopping times with other measures of sample degeneracy. Firstly, *the coefficient of variation* or CV (see [14]) where

$$CV := \left(\frac{1}{n} \sum_{j=1}^{n} \left(n\bar{a}_j^n(t)^2 - 1 \right)^2 \right)^{\frac{1}{2}} \tag{3.6}$$

is closely rated to the effective sample size. Indeed,

$$CV = \left(\frac{n}{n_{eff}} - 1 \right)^{\frac{1}{2}}.$$

We observe that CV is $\sqrt{n-1}$, its maximum value, when all the normalized weights of the particles save one are zero - the worst case of sample degeneracy. It is 0, its minimum value, when all the normalized are equal (that is $\bar{a}_j^n(t) = \frac{1}{n}, j = 1\ldots, n$). The greater the value of CV, the greater the extent of sample degeneracy. Consequently the resampling times, in the case of CV, are determined when $CV > \alpha$, where $\alpha \in \left(0, \sqrt{n-1}\right)$.
This is equivalent to resampling when $n_{eff} < \bar{\alpha}$ where $\bar{\alpha} := \frac{n}{(\alpha^2+1)} \in (1, n)$. Consequently predictable stopping times for the coefficient of variation are defined similarly to how they are defined for effective sampling size and all results deduced for predictable stopping times for effective sampling size also apply to predictable stopping times that use the CV as measure of nondegeneracy.

The *entropy* of the approximating particle system at time t is defined as:

$$E_t = -\sum_{j=i}^{n} \bar{a}_j^n(t) \log \bar{a}_j^n(t) \tag{3.7}$$

where the convention $\lim_{x\to 0+} x \log x = 0$ is used. Observe that

$$0 \le E_t \le \log n \tag{3.8}$$

and that the value of the system's entropy decreases as sample degeneracy worsens so that we resample when the entropy of the system is less than or equal to a constant, β say, with $\beta \in [0, \log n]$. Hence the predictable stopping times for entropy measure is defined for $1 \le k \in \mathbb{N}$ by

$$T_k := \inf\{t \ge T_{k-1} : E_t \le \beta\}. \tag{3.9}$$

Another measure of sample degeneracy is the *maximum of the unnormalized weights*[5]. Recall from (3.2) the unnormalized weights take the form $\exp(w_j)$, where

$$w_j := \int_{T_k}^{T} h(v_j^n(s))^\top \,\mathrm{d}Y_s - \frac{1}{2}\int_{T_k}^{T} \|h(v_j^n(s))\|^2 \,\mathrm{d}s, \quad j = 1, \ldots, n.$$

[5] The same analysis and results also apply to the minimum of the unnormalized weights.

An obvious choice of a measure of sample degeneracy is the function

$$f(w_1, \ldots, w_n) = \max_{1 \le j \le n} w_j \tag{3.10}$$

that is, the maximum of the log of the un-normalized weights. This function however isn't easy to deal with and therefore a proxy called the soft maximum is employed. The soft maximum is defined as

$$f(w_1, \ldots, w_n, r) = \frac{\log \sum_{j=1}^{n} \exp(r\, w_j)}{r}, \tag{3.11}$$

where $r \in \mathbb{N}$. In particular

$$\lim_{r \to \infty} \frac{\log \sum_{j=1}^{n} \exp(r\, w_j)}{r} = \max_{1 \le j \le n} w_j. \tag{3.12}$$

For the purpose of our analysis, the parameter r plays no rôle, so we will focus, for convenience, on the case where $r = 1$ in (3.11) that is, $f(W_1, \ldots, W_n, 1)$ where

$$f(W_1, \ldots, W_n, 1) = \log \sum_{j=1}^{n} \exp(W_j) = \log \sum_{j=1}^{n} a_j^n(t). \tag{3.13}$$

When using the soft maximum we resample when $f(W_1, \ldots, W_n, 1) \ge \alpha,\ \alpha \in [0, \infty)$. We thus define the k^{th} predictable stopping time as

$$\begin{aligned} T_k &:= \inf\left\{ t \ge T_{k-1} : \frac{1}{n} \sum_{j=1}^{n} a_j^n(t) \ge \frac{\exp(\alpha)}{n} \right\} \\ &= \inf\left\{ t \ge 0 : \xi_t^{n,\infty} \ge \left(\frac{\exp(\alpha)}{n} \right)^k \right\}, \end{aligned}$$

where $\xi^{n,\infty} = \{\xi_t^{n,\infty} : t \ge 0\}$ is the process defined as

$$\xi_t^{n,\infty} := \prod_{i=1}^{\infty} \frac{1}{n} \sum_{j=1}^{n} a_j^{n,i}(t) \tag{3.14}$$

with

$$a_j^{n,i}(t) := \exp\left(\int_{T_{i-1} \wedge t}^{T_i \wedge t} h(v_j^n(s))^\top \, \mathrm{d}Y_s - \frac{1}{2} \int_{T_{i-1} \wedge t}^{T_i \wedge t} \|h(v_j^n(s))\|^2 \, \mathrm{d}s \right). \tag{3.15}$$

4 Convergence Results

The convergence of the approximating process $\pi^n = \{\pi_t^n : t \ge 0\}$ relies on the convergence of an unnormalized version of π^n to the solution of the Zakai equation. To this end we introduce the measure-valued process $\rho^n = \{\rho_t^n : t \ge 0\}$ to be defined by

$$\rho_t^n := \xi_t^{n,\infty} \pi_t^n, \quad t \ge 0,$$

where $\xi^{n,\infty} = \{\xi_t^{n,\infty} : t \geq 0\}$ is the process defined in (3.14). In the following, for ease of notation we write ξ_t^n to denote $\xi_t^{n,\infty}$. In particular if $T \in [\, T_{k-1}, T_k \,)$,

$$\rho_T^n = \frac{\xi_{T_{k-1}}^n}{n} \sum_{j=1}^n a_j^n(T)\delta_{v_j^n(T)}.$$

In the following, we will use the bounds

$$\max_{j=1,\dots,n} \sup_{n\geq 0} \mathbb{E}[\sup_{s\in[0,t]} (a_j^n(s))^p] \quad \leq \quad c_0^{t,p}, \tag{4.1}$$

$$\sup_{n\geq 0} \mathbb{E}\,[\sup_{s\in[0,t]} (\xi_s^n)^p] \quad \leq \quad c_1^{t,p}, \tag{4.2}$$

$$\max_{j=1,\dots,n} \sup_{n\geq 0} \mathbb{E}[\sup_{s\in[0,t]} (\xi_s^n \bar{a}_j^n(s))^p] \quad \leq \quad c_2^{t,p}, \tag{4.3}$$

where $c_0^{t,p}$, $c_1^{t,p}$ and $c_2^{t,p}$ are constants which depend on $\max_{i=1,\dots,m} \|h_i\|_{0,\infty}$. In (4.3), the weights $\bar{a}_j^n(s)$ are the normalized versions of $a_j^n(s)$ as defined in (3.15). The proof of these estimates is standard, see Lemma 31 in the Appendix.

Lemma 3. *We have the following control on the mass of the measure-valued process ρ_n: For any $T \geq 0$ and $p \geq 0$, we have*

$$\sup_{n\geq 0} \sup_{t\in[0,T]} \mathbb{E}\left[(\rho_t^n(\mathbf{1}))^p\right] < \infty. \tag{4.4}$$

Proof. Since π^n is a probability measure-valued process, we have that $\rho_t^n(\mathbf{1}) = \xi_t^n$ and (4.4) is immediate from (4.2). □

Proposition 4. *The measure-valued process ρ^n satisfies the following evolution equation*

$$\rho_t^n(\varphi) = \pi_0^n(\varphi) + \int_0^t \rho_s^n(A\varphi)\,\mathrm{d}s + \bar{S}_t^{n,\varphi} + \bar{M}_t^{n,\varphi} + \sum_{k=1}^m \int_0^t \rho_s^n(h_k\varphi)\,\mathrm{d}Y_s^k, \tag{4.5}$$

for any $\varphi \in \mathcal{C}_b^2(\mathbb{R}^d)$. In (4.5)*, $\bar{S}^{n,\varphi} = \{\bar{S}_t^{n,\varphi} : t \geq 0\}$ is the $\mathcal{F}_t$-adapted martingale*

$$\bar{S}_t^{n,\varphi} := \frac{1}{n} \sum_{k=1}^{\infty} \sum_{j=1}^n \int_{T_{k-1}\wedge t}^{T_k\wedge t} \xi_{T_{k-1}}^n a_j^n(s)((\nabla\varphi)^\top\sigma)(v_j^n(s))\,\mathrm{d}V_s^{(j)}$$

and $\bar{M}^{n,\varphi} = \{\bar{M}_t^{n,\varphi} : t \geq 0\}$ is the $\mathcal{F}_t$-adapted martingale

$$\bar{M}_t^{n,\varphi} := \sum_{k=1}^{\infty} \mathbf{1}_{[0,t]}(T_k)(\rho_{T_k}^n(\varphi) - \rho_{T_k-}^n(\varphi)).$$

Proof. Observe that for any $t \geq 0$ and $\varphi \in \mathcal{C}_b^2(\mathbb{R}^d)$,

$$\rho_t^n(\varphi) = \pi_0^n(\varphi) + \sum_{k=1}^{\infty} \mathbf{1}_{[0,t]}(T_k)(\rho_{T_k}^n(\varphi) - \rho_{T_{k-}}^n(\varphi)) + \sum_{k=1}^{\infty} \rho_{T_{k-}\wedge t}^n(\varphi) - \rho_{T_{k-1}\wedge t}^n(\varphi)$$

and

$$\rho_{T_{k-}\wedge t}^n(\varphi) - \rho_{T_{k-1}\wedge t}^n(\varphi) = \int_{T_{k-1}\wedge t}^{T_k\wedge t} \frac{\xi_{T_{k-1}}^n}{n} \sum_{j=1}^{n} \mathrm{d}\left(a_j^n(s)\varphi(v_j^n(s)\right)), \quad k \in \mathbb{N}.$$

The proof then follows by a straightforward application of Itô's formula. The fact that $\bar{M}^{n,\varphi}$ is an $\mathcal{F}_t$-adapted martingale is proved in the appendix (See Proposition 33). ☐

We remark that one can deduce the corresponding result for the evolution equation for π^n. We do not state it here as it plays no rôle in what follows. The mild version of (4.5) is given by

$$\rho_s^n(\varphi) = \rho_0^n(P_t\varphi) + \int_0^t \rho_r^n(P_{t-r}\varphi h^\top)\,\mathrm{d}Y_r + \bar{M}_t^{n,\varphi} + \bar{S}_t^{n,\varphi} \tag{4.6}$$

where

$$\bar{S}_t^{n,\varphi} := \frac{1}{n}\sum_{k=1}^{\infty}\sum_{j=1}^{n}\int_{T_{k-1}\wedge\, t}^{T_k\wedge\, t} \xi_{T_{k-1}}^n a_j^n(r)\left((\nabla(P_{t-r}\varphi))^\top\sigma\right)(v_j^n)\,\mathrm{d}V_r^j$$

and

$$\bar{M}_t^{n,\varphi} := \sum_{k=1}^{\infty} \mathbf{1}_{[0,t]}(T_k)\left(\rho_{T_k}^n(P_{t-T_k}\varphi) - \rho_{T_{k-}}^n(P_{t-T_k}\varphi)\right).$$

Note that (4.6) holds true for any $\varphi \in \mathcal{C}_b^1(\mathbb{R}^d)$. The error between the approximate measure and the target measure is thus given by

$$\delta\rho_t^n(\varphi) = \delta\rho_0^n(P_t\varphi) + \int_0^t \delta\rho_r^n(P_{t-r}\varphi h^\top)\,\mathrm{d}Y_r + \bar{S}_t^{n,\varphi} + \bar{M}_t^{n,\varphi}, \tag{4.7}$$

where $\delta\rho_n = \rho_t^n - \rho_t$. Being able to control the terms on the right hand side of (4.7) therefore is key to obtaining relevant bounds on the error terms which hopefully will lead to information about the rate of convergence of the approximations.

Lemma 5. *For any $T' \geq 0$ and any $p \geq 1$ there exists a constant, $\beta_{T',p}$ independent of n such that, for any $\varphi \in \mathcal{C}_b^1(\mathbb{R}^d)$, we have*

$$\mathbb{E}\left[\left(\sup_{t\in[0,T']}|\bar{S}_t^{n,\varphi}|\right)^{2p}\right] \leq \frac{\beta_{T',p}}{n^p}\|\nabla\varphi\|_\infty^{2p}. \tag{4.8}$$

Proof. For $t \geq 0$ we note that

$$\bar{S}_t^{n,\varphi} \quad = \quad \sum_{j=1}^{n}\int_0^t \xi_r^n \bar{a}_j^n(r)((\nabla(P_{t-r}\varphi))^\top\sigma)(v_j^n(r))\,\mathrm{d}V_r^j. \tag{4.9}$$

By the Burkholder-Davis-Gundy and Jensen inequalities it follows for $p \geq 1$ and $T' \geq 0$ that

$$
\begin{aligned}
\mathbb{E}\left[\left(\sup_{t\in[0,T']}|\bar{S}^{n,\varphi}_t|\right)^{2p}\right] &\leq C_p\,\mathbb{E}[\langle \bar{S}^{n,\varphi}\rangle^p_{T'}] \\
&= C_p\,\mathbb{E}\bigg[\sum_{j=1}^n \int_0^{T'} (\xi^n_r \bar{a}^n_j(r))^2 \left((\nabla P_{T'-r}\varphi)^\top \sigma\sigma^\top (\nabla P_{T'-r}\varphi)\right)(v^n_j(r))\,\mathrm{d}r\bigg]^p \\
&\leq C_p n^{p-1} T'^{p-1}\tilde{d}\sum_{j=1}^n \int_0^{T'} \mathbb{E}[(\xi^n_r)^{4p}]^{\frac{1}{2}}\,\mathbb{E}[(\bar{a}^n_j(r))^{4p}]^{\frac{1}{2}}\,\|\nabla P_{T'-r}\varphi\|^{2p}_\infty\,\mathrm{d}r \qquad (4.10)
\end{aligned}
$$

and since there exists $C_{T'}$ such that for any $\varphi \in \mathcal{C}^1_b(\mathbb{R}^d)$ we have $\|\nabla P_{T'-r}\varphi\|_\infty \leq C_{T'}\|\nabla\varphi\|_\infty$ (see Remark 4.5 in [19]) we get that (4.8) holds true with

$$\beta_{T',p} = C_p T'^p \tilde{d}(c_2^{T',4p})^{\frac{1}{2}} C^{2p}_{T'} d^{2p}.$$

□

Lemma 6. *For any $k \in \mathbb{N}$ and $\varphi \in \mathcal{C}_b(\mathbb{R}^d)$ there exist a constant C_k independent of n such that*

$$\mathbb{E}[(\pi^n_0(P_t\varphi) - \pi_0(P_t\varphi))^{2k}] \leq \frac{C_k\|\varphi\|^{2k}_\infty}{n^k}. \qquad (4.11)$$

Proof. Let $\zeta_j \equiv P_t\varphi(v^n_j(0)) - \pi_0(P_t\varphi)$ so that

$$\frac{1}{n}\sum_{j=1}^n \zeta_j = \frac{1}{n}\sum_{j=1}^n (P_t\varphi(v^n_j(0)) - \pi_0(P_t\varphi)) \equiv \pi^n_0(P_t\varphi) - \pi_0(P_t\varphi)$$

Note that $\zeta_j, j = 1,\ldots,n$, are independent identically distributed random variables with mean 0. The bound (4.11) then follows from

$$
\begin{aligned}
\mathbb{E}[(\pi^n_0(P_t\varphi) - \pi_0(P_t\varphi))^{2k}] &= \mathbb{E}\left[\left(\frac{1}{n}\sum_{j=1}^n \zeta_j\right)^{2k}\right] \\
&= \frac{1}{n^{2k}}\mathbb{E}\Bigg[\sum_{\substack{\alpha_1,\ldots,\alpha_n \\ \alpha_j\neq 1}} \binom{2k}{\alpha_1,\ldots,\alpha_n}[\zeta_1^{\alpha_1}\cdots\zeta_n^{\alpha_n}]\Bigg] \\
&\leq \frac{1}{n^{2k}} \sum_{\substack{\alpha_1,\ldots,\alpha_n \\ \alpha_j\neq 1}} \binom{2k}{\alpha_1,\ldots,\alpha_n}\mathbb{E}[|\zeta_1|^{\alpha_1}\cdots|\zeta_n|^{\alpha_n}] \\
&\leq \frac{C_k\|\varphi\|^{2k}_\infty}{n^k} \qquad (4.12)
\end{aligned}
$$

where the sum is taken all over the multi-indices $(\alpha_1,\ldots,\alpha_n) \in \mathbb{N}^n$ satisfying $\alpha_j \neq 1$ for $j = 1,\ldots,n$ subject to the condition that $\sum_{j=1}^n \alpha_j = 2k$ and $\binom{2k}{\alpha_1,\ldots,\alpha_n} := \frac{2k!}{\alpha_1!\ldots,\alpha_n!}$. The inequality (4.12) follows since the largest coefficient that can be obtained from the preceding inequality is of order n^k. □

Lemma 7. *Assume that there exists $p>1$ such that for all $t>0$, condition* (2.8) *holds true. Then for any $r<p$ and any $\varphi\in\mathcal{C}_b(\mathbb{R}^d)$ there exists a constant $C_{T',4}$ such that*

$$\mathbb{E}\bigg[\sup_{t\in[0,T']}|\bar{M}_t^{n,\varphi}|^{2r}\bigg]\leq\frac{C_{T',2r}\|\varphi\|_\infty^{2r}}{n^r}. \tag{4.13}$$

Proof. Observe that by Burkholder-Davis-Gundy inequality and Fatou's lemma that

$$\begin{aligned}\mathbb{E}\bigg[\sup_{t\in[0,T']}|\bar{M}_t^{n,\varphi}|^{2r}\bigg] &\leq C\lim_m\mathbb{E}\bigg[\bigg(\sum_{k=1}^m\mathbf{1}_{[0,T']}(T_k)\mathbb{E}[(\rho_{T_k}^n(\varphi)-\rho_{T_{k-}}^n(\varphi))^2|\mathcal{F}_{T_k-}]\bigg)^r\bigg]\\ &= C\lim_m\mathbb{E}\bigg[\bigg(\sum_{k=1}^m\mathbf{1}_{[0,T']}(T_k)(\xi_{T_k}^n)^2\mathbb{E}[(\pi_{T_k}^n(\varphi)-\pi_{T_{k-}}^n(\varphi))^2|\mathcal{F}_{T_k-}]\bigg)^r\bigg].\end{aligned} \tag{4.14}$$

Now observe that

$$\pi_{T_k}^n(\varphi)-\pi_{T_k-}^n(\varphi)=\frac{1}{n}\sum_{j=1}^n(\varphi(v_{\alpha_j}^n(T_k))-\mathbb{E}[\pi_{T_k}^n(\varphi)|\mathcal{F}_{T_k-}])=:\frac{1}{n}\sum_{j=1}^n\zeta_{\alpha_j} \tag{4.15}$$

where we have used the fact that $\pi_{T_k}^n(\varphi)=\dfrac{1}{n}\sum_{j=1}^n\varphi(v_{\alpha_j}^n(T_k))$ where $(\alpha_j)_{j=1}^n$ is a random index of $1,\dots,n$ so that $v_{\alpha_j}^n(T_k)=v_{j'}^n(T_k)$ for some $j'\in\{1,\dots,n\}$ with probability $\bar{a}_{j'}^{n,T_k}$. Since

$$\mathbb{E}[\varphi(v_{\alpha_j}^n(T_k))|\mathcal{F}_{T_k-}]=\sum_{j'=1}^n\bar{a}_{j'}^{n,T_k}\varphi(v_{j'}^n(T_k))=\mathbb{E}[\pi_{T_k}^n(\varphi)|\mathcal{F}_{T_k-}],$$

$$\mathbb{E}\left[\zeta_{\alpha_j}\right]=\mathbb{E}\left[\mathbb{E}\left[\zeta_{\alpha_j}|\mathcal{F}_{T_{k-}}\right]\right]=0 \tag{4.16}$$

and hence,

$$\mathbb{E}\left[\pi_{T_k}^n(\varphi)-\mathbb{E}[\pi_{T_k}^n(\varphi)|\mathcal{F}_{T_k-}]\right]=0.$$

By the conditional independence property of sampling with replacement it follows from (4.16) (and similar to (4.12)) that

$$\begin{aligned}\mathbb{E}\left[(\pi_{T_k}^n(\varphi)-\pi_{T_k-}^n(\varphi))^2|\mathcal{F}_{T_k-}\right] &= \mathbb{E}\bigg[\Big(\frac{1}{n}\sum_{j=1}^n\zeta_{\alpha_j}\Big)^2\Big|\mathcal{F}_{T_k-}\bigg]\\ &= \frac{1}{n^2}\mathbb{E}\bigg[\sum_{j=1}^n\zeta_{\alpha_j}^2\Big|\mathcal{F}_{T_{k-}}\bigg]\leq\frac{C_2\|\varphi\|_\infty^2}{n}.\end{aligned} \tag{4.17}$$

Hence, by Hölder's and Jensen's inequalities,

$$\begin{aligned}\lim_m\mathbb{E}\bigg[\bigg(\sum_{k=1}^m\mathbf{1}_{[0,T']}(T_k)\mathbb{E}[(\rho_{T_k}^n(\varphi)-\rho_{T_{k-}}^n(\varphi))^2|\mathcal{F}_{T_k-}]\bigg)^r\bigg] &\leq \frac{C'\|\varphi\|_\infty^{2r}}{n^r}\mathbb{E}\left[\bigg(\sup_{s\in[0,t]}\xi_s^n\bigg)^{2r}(N_t^n)^r\right]\\ &\leq \frac{C'\|\varphi\|_\infty^{2r}}{n^r}\mathbb{E}\left[(N_t^n)^p\right]^{r/p}\mathbb{E}\left[\bigg(\sup_{s\in[0,t]}\xi_s^n\bigg)^{2r(p-r)/p}\right]^{p/p-r}.\end{aligned}$$

(4.13) now follows as a result of the condition on the family of stopping time $\{T_k\}_k$ and (4.2). $\square$

The convergence results of the paper rely on a general convergence criteria.

Let $U := \{U_t : t \geq 0\}$ be a continuous $\mathbb{R}^d$-valued semi-martingale with the property that, for any $t > 0$ there exists a constant C_t such that

$$\mathbb{E}\left[\sup_{r\in[0,t]}\left|\sum_{k=1}^{m}\int_0^r \xi_s^k \mathrm{d}U_s^k\right|^p\right] \leq C_t \sum_{k=1}^{m}\mathbb{E}\left[\int_0^t |\xi_s^k|^p \mathrm{d}s\right] \tag{4.18}$$

for any progressively measurable $\mathcal{F}_t$-adapted $\mathbb{R}^d$-valued process $\xi := \{\xi_t : t \geq 0\}$. Let $\mu^n := \{\mu_t^n : t \geq 0\}$ be a measure-valued process such that for any $\varphi \in \mathcal{C}_b^1(\mathbb{R}^d)$ we have

$$\mu_t^n(\varphi) = \mu_0^n(a_t(\varphi)) + m_t^{n,\varphi} + \sum_{k=1}^{m}\int_0^t \mu_s^n(a_{s,t}^k(\varphi))\,\mathrm{d}U_s^k \tag{4.19}$$

where $m^{n,\varphi} := \{m_t^{n,\varphi} : t \geq 0\}$ is a martingale and $a_t, a_{s,t}^k : \mathcal{C}_b^1(\mathbb{R}^d) \to \mathcal{C}_b^1(\mathbb{R}^d)$ are bounded linear operators with bounds c and C_k, $k = 1,\dots,m$, respectively. That is, $\|a_t(\varphi)\|_{1,\infty} \leq c\|\varphi\|_{1,\infty}$ and $\|a_{s,t}^k(\varphi)\|_{1,\infty} \leq C_k\|\varphi\|_{1,\infty}$, $k = 1,\dots,m$.

In the following, the notation $|\nu|$ denotes the total variation of a measure ν.

Theorem 8. *If for any $T' > 0$ there exist constants γ_1, γ_2 such that for $t \in [0,T']$ and $p \geq 2$*

$$\mathbb{E}\left[|m_t^{n,\varphi}|^p\right] \leq \frac{\gamma_1}{n^{p/2}}\|\varphi\|_{1,\infty}^p; \quad \textit{and} \quad \mathbb{E}\left[|\mu_0^n(a_t(\varphi))|^p\right] < \frac{\gamma_2}{n^{p/2}}\|\varphi\|_{1,\infty}^p, \tag{4.20}$$

and

$$d := \sup_{t\in[0,T']} \mathbb{E}\left[(|\mu_t^n|(\mathbf{1}))^p\right] < \infty, \tag{4.21}$$

then for any $t \in [0,T']$

$$\|\mu_t^n(\varphi)\|_p^p := \mathbb{E}\left[|\mu_t^n(\varphi)|^p\right] \leq \frac{\alpha}{n^{p/2}}\|\varphi\|_{1,\infty}^p \tag{4.22}$$

where $\alpha = \alpha(t)$ is a constant independent of n.

Proof. Observe, by a combination of Jensen's and Burkholder-Davis-Gundy inequalities, that

$$\begin{aligned}
\|\mu_t^n(\varphi)\|_p &\leq \|\mu_0^n(a_t(\varphi))\|_p + \|m_t^{n,\varphi}\|_p + \left[\mathbb{E}\left|\sum_{k=1}^{m}\int_0^t \mu_s^n(a_{s,t}^k(\varphi))\,\mathrm{d}U_s^k\right|^p\right]^{\frac{1}{p}} \\
&\leq 2\left(\frac{\gamma}{n^{p/2}}\|\varphi\|_{1,\infty}^p\right)^{\frac{1}{p}} + \left[Km^{p-1}t^{\frac{p}{2}-1}\sum_{k=1}^{m}\int_0^t \mathbb{E}\left[|\mu_s^n(a_{s,t}^k(\varphi))|^p\right]\mathrm{d}s\right]^{\frac{1}{p}},
\end{aligned}$$

where $\gamma := \max\{\gamma_1, \gamma_2\}$.
Let

$$A^k_{s,t} := \int_0^t \mathbb{E}\Big[|\mu^n_s(a^k_{s,t}(\varphi))|^p\Big]\,\mathrm{d}s = \int_0^t \|\mu^n_s(a^k_{s,t}(\varphi))\|^p_p\,\mathrm{d}s$$

then

$$\|\mu^n_t(\varphi)\|^p_p \le 2^{p-1}2^p\frac{\gamma}{n^{p/2}}\|\varphi\|^p_{1,\infty} + 2^{p-1}Km^{p-1}t^{\frac{p}{2}-1}\sum_{k=1}^m A^k_{s,t}. \tag{4.23}$$

It thus follows using the fact that for $k = 1,\ldots,m$, $a^k_{s,t}$ is a bounded operator and appealing to (4.21) that

$$\begin{aligned}\|\mu^n_t(\varphi)\|^p_p &\le 2^{p-1}2^p\frac{\gamma}{n^{p/2}}\|\varphi\|^p_{1,\infty} + 2^{p-1}Km^{p-1}t^{\frac{p}{2}-1}\sum_{k=1}^m C^p_k\|\varphi\|^p_{1,\infty}dt \\ &\le 2^{p-1}2^p\frac{\gamma}{n^{p/2}}\|\varphi\|^p_{1,\infty} + 2^{p-1}Km^pC^p\|\varphi\|^p_{1,\infty}dt^{\frac{p}{2}} \end{aligned}\tag{4.24}$$

where $C := \max\{C_1,\ldots,C_m\}$. From (4.24) it follows for $k = 1,\ldots,m$, that

$$\begin{aligned}A^k_{s,t} &\le 2^{p-1}2^p\frac{\gamma}{n^{p/2}}C^p_k\|\varphi\|^p_{1,\infty}t + 2^{p-1}Km^pC^pC^p_k\|\varphi\|^p_{1,\infty}d\frac{t^{\frac{p}{2}+1}}{(p/2+1)} \\ &\le 2^{p-1}2^p\frac{\gamma}{n^{p/2}}C^p\|\varphi\|^p_{1,\infty}t + 2^{p-1}Km^pC^{2p}\|\varphi\|^p_{1,\infty}d\frac{t^{\frac{p}{2}+1}}{(p/2+1)}.\end{aligned}$$

(4.23) therefore becomes

$$\|\mu^n_t(\varphi)\|^p_p \le \delta\left[1+\beta_pt^\kappa\right] + \beta^2_pd\|\varphi\|^p_{1,\infty}\frac{t^{2\kappa}}{(\kappa+1)} \tag{4.25}$$

where $\delta = 2^{p-1}2^p\dfrac{\gamma}{n^{p/2}}\|\varphi\|^p_{1,\infty}$; $\beta_p = 2^{p-1}Km^pC^p$ and; $\kappa = p/2$. In turn, (4.25) now gives us that

$$A^k_{s,t} \le \delta C^pt + \delta\beta_pC^p\frac{t^{\kappa+1}}{(\kappa+1)} + \beta^2_pC^pd\|\varphi\|^p_{1,\infty}\frac{t^{2\kappa+1}}{(2\kappa+1)(\kappa+1)}$$

hence

$$\|\mu^n_t(\varphi)\|^p_p \le \delta\left[1+\beta_pt^\kappa+\beta^2_p\frac{t^{2\kappa}}{(\kappa+1)}\right] + \beta^3_pd\|\varphi\|^p_{1,\infty}\frac{t^{3\kappa}}{(2\kappa+1)(\kappa+1)}. \tag{4.26}$$

So in general with $\|\mu^n_t(\varphi)\|^{p,(k)}_p$ denoting the k^{th}-iteration, it follows by an induction argument that

$$\begin{aligned}\|\mu^n_t(\varphi)\|^{p,(k)}_p &\equiv \|\mu^n_t(\varphi)\|^p_p \\ &\le \delta\left[1+\beta_pt^\kappa+\beta^2_p\frac{t^{2\kappa}}{(\kappa+1)}+\cdots+\beta^{k-1}_p\frac{t^{(k-1)\kappa}}{((k-2)\kappa+1)\ldots(\kappa+1)}\right] \\ &\quad +\beta^k_pd\|\varphi\|^p_{1,\infty}\frac{t^{k\kappa}}{((k-1)\kappa+1)((k-2)\kappa+1)\ldots(\kappa+1)}\end{aligned}$$

and as $k \to \infty$,

$$\|\mu_t^n(\varphi)\|_p^p \leq \delta\left(1+\beta_p t^\kappa \sum_{j=0}^{\infty} \beta_p^j \frac{t^{j\kappa}}{\kappa^j j!}\right) = \delta\left(1+\beta_p t^\kappa e^{\frac{\beta_p t^\kappa}{\kappa}}\right) = 2^{2p-1}\left(1+\beta_p t^\kappa e^{\frac{\beta_p t^\kappa}{\kappa}}\right)\frac{\gamma}{n^{p/2}}\|\varphi\|_{1,\infty}^p.$$

Hence (4.22) follows with $\alpha := 2^{2p-1}\left(1+\beta_p t^\kappa e^{\frac{\beta_p t^\kappa}{\kappa}}\right)\gamma$. $\square$

Remark 9. *If the norms in the bounds* (4.20) *can be replaced with* $\|\cdot\|_\infty$ *that is*

$$\mathbb{E}\left[|m_t^{n,\varphi}|^p\right] \leq \frac{\gamma_1}{n^{p/2}}\|\varphi\|_\infty^p; \quad \textit{and} \quad \mathbb{E}\left[|\mu_0^n(a_t(\varphi))|^p\right] < \frac{\gamma_2}{n^{p/2}}\|\varphi\|_\infty^p, \tag{4.27}$$

then by defining a_t *and* $a_{s,t}^k$ *to be bounded linear operators on* $\mathcal{C}_b(\mathbb{R}^d)$*, we can reach a similar conclusion as in Theorem 8 for* $\varphi \in \mathcal{C}_b(\mathbb{R}^d)$*.*

We also have the following almost sure convergence result

Corollary 10. *Under the same conditions as Theorem 8 but with the exception that* $p > 2$ *there exists a positive random variable,* $c_{\mu,\varepsilon}$*, almost surely finite, such that*

$$|\mu_t^n(\varphi)| \leq \frac{c_{\mu,\varepsilon}}{n^\varepsilon} \tag{4.28}$$

where $\varepsilon \in \left(0, \frac{1}{2} - \frac{1}{p}\right)$*.*
That is

$$\mu_t^n(\varphi) \to 0 \quad \tilde{\mathbb{P}} - \textit{a.s.} \tag{4.29}$$

as $n \to \infty$*.*

Proof. Recall from (4.22) in Theorem 8 that

$$\|\mu_t^n(\varphi)\|_p^p := \mathbb{E}\left[|\mu_t^n(\varphi)|^p\right] \leq \frac{\alpha}{n^{p/2}}\|\varphi\|_{1,\infty}^p \tag{4.30}$$

where $\alpha = \alpha(t)$ is a constant independent of n.
By Fatou's Lemma it follows therefore that for $\varepsilon \in \left(0, \frac{1}{2} - \frac{1}{p}\right)$

$$\mathbb{E}\left[\sum_{n=1}^{\infty} n^{\varepsilon p}|\mu_t^n(\varphi)|^p\right] \leq \alpha\|\varphi\|_{1,\infty}^p \sum_{n=1}^{\infty} \frac{1}{n^{p-2\varepsilon p}} < \infty.$$

Let $c_{\mu,\varepsilon} := \left[\sum_{n=1}^{\infty} n^{\varepsilon p}|\mu_t^n(\varphi)|^p\right]^{\frac{1}{p}}$. Then $c_{\mu,\varepsilon}$ is integrable and $c_{\mu,\varepsilon}$ is finite a.s. (4.28) now follows since

$$n^{2\varepsilon p}|\mu_t^n(\varphi)|^{2p} \le \sum_{n=1}^{\infty} n^{2\varepsilon p}\left(|\mu_t^n(\varphi)|\right)^{2p}.$$

□

In the following we will use the notation $\delta_{\rho_t^n(\varphi)}$ to denote $(\rho_t^n - \rho_t)(\varphi)$ and $\delta_{\pi_t^n(\varphi)}$ to denote $(\pi_t^n - \pi_t)(\varphi)$. Theorem 1 is proved in the following:

Theorem 11. *Assume that there exists $p > 1$ such that for all $t > 0$, condition* (2.8) *holds true. Then for any $T' \ge 0$ and for any $r < p$, there exists a constant $\alpha = \alpha(T')$, independent of n such that for any $\varphi \in \mathcal{C}_b^1(\mathbb{R}^d)$, we have*

$$\mathbb{E}\left[((\rho_t^n - \rho_t)(\varphi))^{2r}\right] \le \frac{\alpha}{n^r}\|\varphi\|_{1,\infty}^{2r}, \qquad t \in [0, T'] \tag{4.31}$$

and

$$\mathbb{E}\left[((\pi_t^n - \pi_t)(\varphi))^{r}\right] \le \frac{\alpha}{n^{\frac{r}{2}}}\|\varphi\|_{1,\infty}^{r}, \qquad t \in [0, T']. \tag{4.32}$$

Proof. (4.31) follows from Theorem 8 by setting:

$$\mu_t^n(\cdot) := (\rho_t^n - \rho_t)(\cdot); \quad \mu_0^n(a_t(\cdot)) := (\rho_0^n - \rho_0)(P_t\cdot);$$

$$m_t^{n,\varphi} := \bar{S}_t^{n,\varphi} + \bar{M}_t^{n,\varphi}; \quad \mu_r^n(a_{r,t}^k(\cdot)) \equiv (\rho_r^n - \rho_r)(P_{t-r}\cdot h^\top), \quad 0 < r \le t;$$

and appealing to Lemma 5, Lemma 6 and Lemma 7. Note that the observation process Y satisfies condition (4.18). The control on the total mass (4.21) follows from Lemma 3 and the fact that $\sup_{t\in[0,T']}\mathbb{E}\left[(\rho_t(\mathbf{1}))^p\right] < \infty$, see Proposition 4.23 in [1].

In order to prove (4.32), note that $\pi_t^n(\varphi)\rho_t^n(\mathbf{1}) = \xi_t^n\pi_t^n(\varphi) = \rho_t^n(\varphi)$. It follows thus that

$$\delta_{\pi_t^n(\varphi)} = \frac{1}{\rho_t(\mathbf{1})}\delta_{\rho_t^n(\varphi)} - \frac{\pi_t^n(\varphi)}{\rho_t(\mathbf{1})}\delta_{\rho_t^n(\mathbf{1})}. \tag{4.33}$$

One can show that $u_t := \sqrt{\mathbb{E}\left[(\rho_t(\mathbf{1}))^{-p}\right]} < \infty$ (see for example Exercise 9.16 in [1]).
By a combination of Jensen's and Cauchy-Schwartz inequalities

$$\mathbb{E}\left[|\delta_{\pi_t^n}(\varphi)|^p\right] \le 2\left(u_t\sqrt{\mathbb{E}\left[\delta_{\rho_t^n}^{2p}(\varphi)\right]} + u_t\|\varphi\|_{0,\infty}^2\sqrt{\mathbb{E}\left[\delta_{\rho_t^n}^{2p}(\mathbf{1})\right]}\right).$$

(4.32) now follows from (4.31). □

Remark 12. *Further to the results in Theorem 11. It can be shown that for $0 < \varepsilon < \frac{1}{2} - \frac{1}{2p}$ and $t \in [0, T']$,*

$$|\delta_{\rho_t^n}(\varphi)| \le \frac{c_{t,\,\varepsilon}}{n^\varepsilon} \tag{4.34}$$

where $c_{t,\,\varepsilon}$ is a positive, almost surely finite random variable and hence (as in Corollary 10),

$$\rho_t^n(\varphi) \to \rho_t(\varphi) \quad \mathbb{P} - \text{a.s.}. \tag{4.35}$$

Hence also $\pi_t^n(\varphi) \to \pi_t(\varphi)$ $\mathbb{P}$-a.s.

Remark 13. *Under the same conditions as in Theorem 11 with the exception that $\varphi \in \mathcal{C}_b^3(\mathbb{R}^d)$, one can prove (see [19]) the following stronger results:*

$$\mathbb{E}\left[\sup_{t\in[0,T']} \delta^{2r}_{\rho_t^n}(\varphi)\right] \le \frac{\beta_{T'}}{n^r}\|\varphi\|^{2r}_{3,\infty}, \qquad t \in [0, T'], \tag{4.36}$$

$$\mathbb{E}\left[\sup_{t\in[0,T']} \delta^{r}_{\pi_t^n}(\varphi)\right] \le \frac{\tilde{\beta}_{T'}}{n^{\frac{r}{2}}}\|\varphi\|^{r}_{3,\infty}, \qquad t \in [0, T'].$$

5 Examples of convergent adaptive particle filters

We show now that the results of the preceding section are valid for the case where the predictable stopping times are determined by the effective sample size. We will show that for any $p > 1$, condition (2.8) holds true. By applications of Itô's formula we have:

Proposition 14. *For $t \ge 0$ let $\bar{S}_t := \sum_{j=1}^n \bar{a}_j^n(t)^2$ so that $\tilde{S}_t := \bar{S}_t^{-1} \equiv \hat{n}_{eff}$. Then for $t \in [T_{k'}, T_{k'+1}), 0 \le k' \in \mathbb{N}$,*

$$\tilde{S}_t \quad = \quad n \exp\left(\sum_{k=1}^m \int_{T_{k'}}^t \tilde{\eta}_s^{n,k}\,\mathrm{d}Y_s^k + \sum_{k=1}^m \int_{T_{k'}}^t \tilde{\zeta}_s^{n,k}\,\mathrm{d}s - \frac{1}{2}\sum_{k=1}^m \int_{T_{k'}}^t (\tilde{\eta}_s^{n,k})^2\,\mathrm{d}s\right)$$

where

$$\tilde{\eta}_t^{n,k} := 2\left(\pi_t^n\left(h^k\right) - \tilde{\pi}_t^n\left(h^k\right)\right), \tag{5.1}$$

and

$$\tilde{\zeta}_t^{n,k} := \pi_t^n\left(h^k\right)^2 + 2\pi_t^n\left(h^k\right)\tilde{\pi}_t^n\left(h^k\right) - \tilde{\pi}_t^n\left((h^k)^2\right) + 4\tilde{\pi}_t^n\left(h^k\right)^2, \tag{5.2}$$

with

$$\tilde{a}_j^n(t)^2 := \bar{S}_t^{-1}\bar{a}_j^n(t)^2 = \frac{\bar{a}_j^n(t)^2}{\sum_{j=1}^n \bar{a}_j^n(t)^2} \quad \text{and} \quad \tilde{\pi}_t^n := \sum_{j=1}^n \tilde{a}_j^n(t)^2 \delta_{v_j^n(t)}. \tag{5.3}$$

So after the k'^{th} stopping time, it follows that

$$\tilde{S}_t = \tilde{S}_{T_{k'}} + \sum_{k=1}^{m} \int_{T_{k'}}^{t} \tilde{S}_s \tilde{\eta}_s^{n,k} \, \mathrm{d}Y_s^k + \sum_{k=1}^{m} \int_{T_{k'}}^{t} \tilde{S}_s \tilde{\zeta}_s^{n,k} \, \mathrm{d}s,$$

where $t > T_{k'}$. In particular for $t \in [T_{k'}, T_{k'+1})$,

$$\tilde{S}_t \;=\; n \exp\Big(\sum_{k=1}^{m} \int_{T_{k'}}^{t} \tilde{\eta}_s^{n,k} \, \mathrm{d}Y_s^k + \sum_{k=1}^{m} \int_{T_{k'}}^{t} \tilde{\zeta}_s^{n,k} \, \mathrm{d}s - \frac{1}{2} \sum_{k=1}^{m} \int_{T_{k'}}^{t} (\tilde{\eta}_s^{n,k})^2 \, \mathrm{d}s \Big)$$

Remark 15. *For $t \in [T_{k'}, T_{k'+1})$, we can also write*

$$\tilde{S}_t \;=\; n \exp(-\alpha_t^{n,k'}) \tag{5.4}$$

where

$$\alpha_t^{n,k'} := \int_{T_{k'}}^{t} \beta_s^{n,k} \, \mathrm{d}Y_s^k + \int_{T_{k'}}^{t} \gamma_s^{n,k} \, \mathrm{d}s \tag{5.5}$$

is a semimartingale with

$$\beta_s^{n,k} = -\sum_{k=1}^{m} \tilde{\eta}_s^{n,k}; \quad \gamma_s^{n,k} = -\sum_{k=1}^{m} \left(\tilde{\zeta}_s^{n,k} - \frac{1}{2} \left(\tilde{\eta}_s^{n,k} \right)^2 \right), \quad s \in [T_{k'}, T_{k'+1})$$

where we use the convention that

$$\beta_s^{n,k} \, \mathrm{d}Y_s^k := \sum_{k=1}^{m} \tilde{\eta}_s^{n,k} \, \mathrm{d}Y_s^k.$$

Recall from (3.4) that the family of predictable stopping times $\{T_k\}_{k \in \mathbb{N}}$ determined by the ESS is defined for $1 \leq k' \in \mathbb{N}$ by

$$T_{k'} := \inf\{t \geq T_{k'-1} : n_{eff} \leq \lambda_{thres} n\} \tag{5.6}$$

where $\lambda_{thres} \in (0, 1)$.
Observe that by (5.4), (5.6) can be rewritten as

$$T_{k'} = \inf\{t \geq T_{k'-1} : \alpha_t^{n,k'-1} \geq \log \lambda^{-1}\}. \tag{5.7}$$

Lemma 16. *Let $t \in [T_{k-1'}, T_{k'})$, $k' \geq 1$ an integer. Then*

$$T_{k'} := \inf\{t \geq T_{k'-1} : n_{eff} \leq \lambda n\}$$

can be equivalently rewritten as

$$T_{k'} := \inf\{t \geq 0 : \alpha_t^n \geq k' \log \lambda^{-1}\} \tag{5.8}$$

where for ease of notation, $\lambda \equiv \lambda_{thres}$ and using the notation of Remark 15, the semi-martingale α_t^n is defined by

$$\alpha_t^n = \alpha_{T_1}^{n,0} + \alpha_{T_2}^{n,1} + \cdots + \alpha_{T_{k'-1}}^{n,k'-2} + \alpha_t^{n,k'-1}. \tag{5.9}$$

That is,

$$\begin{aligned}
\alpha_t^n &= \int_0^{T_1} \beta_s^{n,k}\,\mathrm{d}Y_s^k + \int_0^{T_1} \gamma_s^{n,k}\,\mathrm{d}s + \cdots + \int_{T_{k'-2}}^{T_{k-1'}} \beta_s^{n,k}\,\mathrm{d}Y_s^k + \int_{T_{k'-2}}^{T_{k'-1}} \gamma_s^{n,k}\,\mathrm{d}s + \int_{T_{k'-1}}^{t} \beta_s^{n,k}\,\mathrm{d}Y_s^k \\
&\quad + \int_{T_{k'-1}}^{t} \gamma_s^{n,k}\,\mathrm{d}s \\
&= \int_0^t \beta_s^n\,\mathrm{d}Y_s + \int_0^t \gamma_s^n\,\mathrm{d}s
\end{aligned} \tag{5.10}$$

where for $s \in [T_p, T_{p+1}), p = 0, \ldots, k-1'$,

$$\beta_s^n := \beta_s^{n,p};\ Y_s := Y_s^p;\ \gamma_s^n := \gamma_s^{n,p}.$$

Proof*.* The result follows by the definition of α_t^n and noting that $\alpha_t^{n,k'-1} \geq \log\lambda^{-1}$ and for $0 \leq r \leq k'-2, \alpha_{T_{r+1}}^{n,r} = \log\lambda^{-1}$. □

Proposition 17. *We will show that for any* $p > 1$*, condition* (2.8) *holds true, i.e.,*

$$\sup_{n>0} \mathbb{E}[(N_t^n)^p] < \infty \tag{5.11}$$

Proof*.* From (5.8)

$$\{T_k \leq t\} \equiv \left\{ \sup_{s\in[0,t]} \alpha_s^n \geq k\log\lambda^{-1} \right\} \tag{5.12}$$

therefore

$$N_t^n = \max\{k \geq 0;\ \ T_k \leq t\} = \max\{k \geq 0;\ \ \sup_{s\in[0,t]} \alpha_s^n \geq k\log\lambda^{-1}\} \leq \frac{\sup_{s\in[0,t]} \alpha_s^n}{\log\lambda^{-1}}$$

By Jensen's and the Burkholder-Davis-Gundy inequalities, the fact that $\sum_{j=1}^n \bar{a}_j^n(t) = 1 = \sum_{j=1}^n \tilde{a}_j^n(t)$ and $h(v_j^n(t))$ being bounded for any $t \geq 0$, it follows from (5.1) and (5.2) that $\mathbb{E}[(\beta_s^n)^p] < \infty$ and $\mathbb{E}[(\gamma_s^n)^p] < \infty$ and

$$\begin{aligned}
\mathbb{E}\left[\left(\sup_{s\in[0,t]} \alpha_s^n\right)^p\right] &\leq \mathbb{E}\left[\sup_{s\in[0,t]}\left(\int_0^s \beta_r^n\,\mathrm{d}Y_r + \int_0^s \gamma_r^n\,\mathrm{d}r\right)^p\right] \\
&\leq 2^p\,\mathbb{E}\left[\sup_{s\in[0,t]}\left|\int_0^s \beta_r^n\,\mathrm{d}Y_r\right|^p\right] + 2^p\,\mathbb{E}\left[\sup_{s\in[0,t]}\left|\int_0^s \gamma_r^n\,\mathrm{d}r\right|^p\right] \\
&\leq C
\end{aligned}$$

where C is a constant independent of the number of particles, n. The claim is proved. □

The following result is now immediate

Proposition 18. *The approximating measure of the signal converges to the true measure in a L^p-sense when the predictable stopping resampling times are determined by the effective sample size.*

Moreover we can also conclude that

Proposition 19. *The approximating measure of the signal converges to the true measure in a L^p-sense when the predictable stopping resampling times are determined by the soft maximum.*

Proof. Observe that for any $t \geq 0$ there exist a $M \in \mathbb{N}$ such that

$$\begin{aligned} \xi_t^{n,\infty} &= \xi_t^{n,M} \\ &:= \prod_{i=1}^{M} \frac{1}{n} \sum_{j=1}^{n} a_j^{n,i}(t) \\ &= \frac{1}{n^M} \sum_{(j_1,\ldots,j_M)\in J_M} \exp\left(\sum_{i=1}^{M} \left(\int_{T_{i-1}\wedge t}^{T_i\wedge t} h(v_{j_i}^n(s))^\top \, \mathrm{d}Y_s - \frac{1}{2} \int_{T_{i-1}\wedge t}^{T_i\wedge t} \|h(v_{j_i}^n(s))\|^2 \, \mathrm{d}s \right) \right) \end{aligned} \tag{5.13}$$

$$= \frac{1}{n^M} \sum_{(j_1,\ldots,j_M)\in J_M} \exp\left(\int_0^{T_M\wedge t} h(v_{j_1,\ldots,j_M}^n(s))^\top \, \mathrm{d}Y_s - \frac{1}{2} \int_0^{T_M\wedge t} \|h(v_{j_1,\ldots,j_M}^n(s))\|^2 \, \mathrm{d}s \right)$$

where J_M is the set of multi-indices defined by

$$J_M := \{(j_1,\ldots,j_M) : j_i \in \{1,\ldots,n\}, i = 1,\ldots,M\}$$

and

$$v_{j_1,\ldots,j_M}^n = v_{j_i}^n \quad \text{if} \quad s \in [T_{i-1}\wedge t, T_i \wedge t),\ i = 1,\ldots,M.$$

The claim now follows by using a similar approach to the one employed for the ESS. □

6 A central Limit Theorem

In the section we will assume throughout that the conditions stated in Theorem 2 hold true. That is we assume that for any $k \geq 0$, $\lim_{n\to\infty} T_k^n = T_k$, where $(T_k)_{k\geq 0}$ is a strictly increasing sequence of $\mathcal{Y}_t$-adapted predictable stopping times. We also assume that (2.8) is satisfied and also there exists $p > 2$ such that for all $t > 0$, condition (2.10) holds true. That is

$$\lim_{\delta\to 0} \sup_{n>0} \mathbb{E}[\sup_{s\in[0,t]} \mathbb{E}[(N_{s+\delta}^n - N_s^n)^p | \mathcal{F}_s]] = 0,$$

Let $\{U^n\}_{n\in\mathbb{N}}$ be the family of measure-valued processes defined as

$$U_t^n := \sqrt{n}(\rho_t^n - \rho_t), \quad t \geq 0. \tag{6.1}$$

We will show that U^n converges in distribution to a certain process U identified as the unique solution of a certain evolution equation. Both U^n and U are viewed as processes with values in the space $\mathcal{M}_F(\mathbb{R}^d)$ endowed with a vague topology, that is, the weak*-topology on $\mathcal{C}_0(\mathbb{R}^d)$. It is possible to obtain the same results by endowing $\mathcal{M}_F(\mathbb{R}^d)$ with the weak topology (see Remark 27 for details).

6.1 The Tightness of the sequence U^n

Let $\{\mathbb{P}_n\} \subset \mathcal{P}\left(D_{\mathcal{M}_F(\mathbb{R}^d)}[0,T]\right)$ be the family of associated probability distributions of $\{U^n\}$. Let $(\iota_k)_{k\geq 0}$ be a sequence of functions defined as $\iota_0 \equiv \mathbf{1}$ and $(\iota_k)_{k\geq 1}$, a dense sequence in $\mathcal{C}_0^\infty\left(\mathbb{R}^d\right)$. Then by Theorem 2.1 in [21] it follows that to show that $\{\mathbb{P}_n\}$ is tight if the probability distributions of the sequence $U_t^n(\iota_k)$ is tight. To show this we make use of the following theorem (see Theorems 8.8 + 8.6 in [8]):

Theorem 20 (**Kurtz's criteria of relative compactness**). *Let (E,d) be a separable and complete metric space and let $\{X^n\}_{n\in\mathbb{N}}$ be a sequence of processes with sample paths in $D_E[0,\infty)$. Suppose that for every $\eta > 0$ and rational t, there exists a compact set $\Gamma_{\eta,t}$ such that*

$$\sup_n \mathbb{P}(X_t^n \notin \Gamma_{\eta,t}) \leq \eta. \tag{6.2}$$

Then $\{X^n\}_{n\in\mathbb{N}}$ is relatively compact if for each $T' > 0$, there exists $\beta > 0$ and a family $\{\gamma^n(\delta) : 0 < \delta < 1\}$ of non-negative random variables satisfying

$$\mathbb{E}[(1 \wedge d(X_{t+u}^n, X_t^n))^\beta|\mathcal{F}_t] \leq \mathbb{E}[\gamma^n(\delta)|\mathcal{F}_t] \tag{6.3}$$

for $0 \leq t \leq T'$, $0 \leq u \leq \delta$ and

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \mathbb{E}[\gamma^n(\delta)] = 0. \tag{6.4}$$

We have the following

Theorem 21. *Provided that h in* (2.3) *and the coefficients a^{ij} and f^i, $1 \leq i,j \leq d$, in* (2.2) *belong to $\mathcal{C}_b^3(\mathbb{R}^d)$ then the sequence $\{U^n\}_{n\in\mathbb{N}}$ is relatively compact.*

Proof. It suffices to show that $\{U^n(\iota_k)\}_n$ is relatively compact for any $k \geq 0$. To show this we use Theorem 20 with $E = \mathbb{R}$, d the Euclidean metric and $\{X^n\} = \{U^n(\iota_k)\}$. First note that (6.2) holds as a consequence of the fact that $\mathbb{E}[\sup_t(U^n(\iota_k))^2]$ is bounded above by a constant independent of n which is is an immediate consequence of the bound (4.31) (or (4.36)).
We now obtain a suitable family of random variables $\{\gamma^n(\delta) : 0 < \delta < 1\}$ that satisfies (6.3). Firstly observe that, for any $0 \leq a \leq b$

$$U_b^n(\iota_k) - U_a^n(\iota_k) = q_{a,b}^{n,1} + q_{a,b}^{n,2} + \sqrt{n}(\bar{S}_b^{n,\iota_k} - \bar{S}_a^{n,\iota_k}) + \sqrt{n}(\bar{M}_b^{n,\iota_k} - \bar{M}_a^{n,\iota_k}), \tag{6.5}$$

where

$$q^{n,1}_{a,b} = \int_a^b \sqrt{n}(\rho^n_s(A\iota_k) - \rho_s(A\iota_k))\,\mathrm{d}s \tag{6.6}$$

$$q^{n,2}_{a,b} = \sum_{l=1}^m \int_a^b \sqrt{n}\left(\rho^n_s(h_l\iota_k) - \rho_s(h_l\iota_k)\right)\,\mathrm{d}Y^l_s. \tag{6.7}$$

We find L_{2r}-bounds for each of the terms in (6.5). Observe that, from (4.36), we have that

$$\mathbb{E}[\sup_{u\in[0,\delta]} |q^{n,1}_{t,t+u}|^r|\mathcal{F}_t] \leq \delta^r n^{r/2}\mathbb{E}[\sup_{s\in[0,T'+\delta]}(\rho^n_s(A\iota_k) - \rho_s(A\iota_k))^r|\mathcal{F}_t] \tag{6.8}$$

$$\mathbb{E}[\sup_{u\in[0,\delta]} |q^{n,2}_{t,t+u}|^r|\mathcal{F}_t] \leq c\delta^{r/2} n^{r/2}\sum_{l=1}^m\mathbb{E}[\sup_{s\in[0,T'+\delta]}(\rho^n_s(h_l\iota_k) - \rho_s(h_l\iota_k))^r|\mathcal{F}_t] \tag{6.9}$$

Also similar with the proof of the bound (4.8), we have

$$\mathbb{E}[\sup_{u\in[0,\delta]}(\sqrt{n}|\bar{S}^{n,\iota_k}_{t+u} - \bar{S}^{n,\iota_k}_t|)^r|\mathcal{F}_t] \leq c\delta^{r/2}\|\nabla i_k\|^r_\infty\mathbb{E}[\sup_{s\in[0,T'+\delta]}(\rho^n_s(1))^r|\mathcal{F}_t]. \tag{6.10}$$

Finally, similar with the proof of the bound (4.13) we have that for any $r < p$ and any $\varphi \in \mathcal{C}_b(\mathbb{R}^d)$

$$\mathbb{E}[\sup_{u\in[0,\delta]}\sqrt{n}|\bar{M}^{n,\iota_k}_{t+u} - \bar{M}^{n,\iota_k}_t|^r|\mathcal{F}_t] \leq C_{T',2r}\|\varphi\|^r_\infty\mathbb{E}[(N^n_{t+\delta} - N^n_t)^{r/2}\sup_{s\in[0,T'+\delta]}(\xi^n_s(1))^r|\mathcal{F}_t]. \tag{6.11}$$

Hence, following from (6.8),(6.9),(6.10) and (6.11), we get that (6.3) holds true with

$$\begin{aligned}\gamma^n(\delta) &= \delta^r n^{r/2}\sup_{s\in[0,T'+\delta]}(\rho^n_s(A\iota_k) - \rho_s(A\iota_k))^r + c\delta^{r/2}n^{r/2}\sum_{l=1}^m\sup_{s\in[0,T'+\delta]}(\rho^n_s(h_l\iota_k) - \rho_s(h_l\iota_k))^r \\ &\quad + c\delta^{r/2}\|\nabla i_k\|^r_\infty\sup_{s\in[0,T'+\delta]}(\rho^n_s(1))^r + C_{T',2r}\|\varphi\|^r_\infty\sup_{t\in[0,t'+\delta]}[\mathbb{E}[(N^n_{t+\delta} - N^n_t)^{r/2}\sup_{s\in[0,T'+\delta]}(\xi^n_s(1))^r|\mathcal{F}_t]\end{aligned}$$

which in turn satisfy (6.4) following the bounds (4.2), (4.36) and the condition (2.10). □

6.2 The process U

We start by introducing the process $\tilde{\rho} = \{\tilde{\rho}_t,\ \ t \geq 0\}$.

$$\begin{aligned}\tilde{\rho}_t(\varphi) &= \tilde{\rho}_0(\varphi) + \int_0^t \Bigg[\rho_s(\mathbf{1})\rho_s(A\varphi) - \sum_{r'=1}^m\left(\rho_s(\mathbf{1})\rho_s(h^{r'}\varphi) - \rho_s(h^{r'})\rho_s(\varphi)\right)\pi_s(h^{r'}) \\ &\quad + \sum_{k'=1}^m \pi_s(\varphi)\rho_s(h^{k'})^2 + 2\sum_{k'=1}^m\left(\rho_s(h^{k'})\rho_s(h^{k'}\varphi) - \rho_s(h^{k'})^2\pi_s(\varphi)\right)\Bigg]\,\mathrm{d}s \\ &\quad + \sum_{k'=1}^m\int_0^t\left[\rho_s(\mathbf{1})\rho_s(h^{k'}\varphi) + \rho_s(h^{k'})\rho_s(\varphi)\right]\,\mathrm{d}Y^{k'}_s\end{aligned}$$

The following lemma is proved in the same manner as the results in the previous section:

Lemma 22. *Let $\tilde{\rho}^n = \{\tilde{\rho}^n_t,\ \ t \geq 0\}$ be the sequence of processes*

$$\tilde{\rho}^n_t := (\xi^n_t)^2\, n \sum_{j=1}^{n} \bar{a}^n_j(t)^2 \delta_{v^n_j(t))}.$$

Then $\tilde{\rho}^n_t(\varphi) \to \tilde{\rho}_t(\varphi)$ $\mathbb{P}$-a.s. for any $\varphi \in C^2_b(\mathbb{R}^d)$ and $t \geq 0$.

Let N^φ be a $(\mathcal{F}_t \vee \mathcal{Y})$-adapted square-integrable martingale given by

$$N^\varphi_t = \int_0^t \int_{\mathbb{R}^d} \sqrt{\tilde{\rho}_s\left((\nabla\varphi)^\top \sigma\sigma^\top (\nabla\varphi)\right)}\, B(\mathrm{d}x, \mathrm{d}s) + \sum_{k=1}^{\infty} \mathbf{1}_{[0,t]}(T_k)\rho_{T_k}(\mathbf{1})\sqrt{\pi_{T_k}(\varphi^2) - \pi_{T_k-}(\varphi)^2}\,\Upsilon_k \tag{6.12}$$

where $B(\mathrm{d}x, \mathrm{d}s)$ is a Brownian sheet or space-time white noise, $\{\Upsilon_k\}_{k\in\mathbb{N}}$ is a sequence of i.i.d standard normal random variables mutually independent given the sigma algebra $\mathcal{Y}$.

Proposition 23. *If $U := \{U_t : t \geq 0\}$ is a $D_{\mathcal{M}_F(\mathbb{R}^d)}[0,\infty)$-valued process such that for $\varphi \in \mathcal{C}^2_0(\mathbb{R}^d)$*

$$U_t(\varphi) = U_0(\varphi) + \int_0^t U_s(A\varphi)\,\mathrm{d}s + N^\varphi_t + \sum_{k=1}^{m} \int_0^t U_s(h^k\varphi)\,\mathrm{d}Y^k_s \tag{6.13}$$

then U is pathwise unique.

Proof. Let U and $\tilde{U}$ be two solutions of (6.13). Define $\hat{U} := U - \tilde{U}$. Then $\hat{U}_t$ satisfies the following equation

$$\hat{U}_t(\varphi) = \hat{U}_0(\varphi) + \int_0^t \hat{U}_s(A\varphi)\,\mathrm{d}s + \sum_{k=1}^{m} \int_0^t \hat{U}_s(h^k\varphi)\,\mathrm{d}Y^k_s \tag{6.14}$$

and it follows similarly to Theorem 2.21(i) and Remark 3.4 in [17] or Lemma 4.2 in [16] that $\hat{U} = 0$. □

Remark 24. *That a solution to* (6.13) *exists will be shown in Theorem 25 where it will be shown that $\{U^n\}_n$ converges in distribution to U.*

6.3 Convergence in distribution

Theorem 25. *$\{U^n\}_n$ converges in distribution to a unique $D_{\mathcal{M}_F(\mathbb{R}^d)}[0,\infty)$-valued process, $U := \{U_t : t \geq 0\}$, such that for $\varphi \in \mathcal{C}^2_0(\mathbb{R}^d)$*

$$U_t(\varphi) = U_0(\varphi) + \int_0^t U_s(A\varphi)\,\mathrm{d}s + N^\varphi_t + \sum_{k=1}^{m} \int_0^t U_s(h^k\varphi)\,\mathrm{d}Y^k_s. \tag{6.15}$$

where N^φ is an $(\mathcal{F}_t \vee \mathcal{Y})$-adapted martingale with quadratic variation

$$\langle N^\varphi_\cdot \rangle_t = \int_0^t \tilde{\rho}_s\left((\nabla\varphi)^\top \sigma\sigma^\top (\nabla\varphi)\right)\,\mathrm{d}s + \sum_{k=1}^{\infty} \mathbb{E}\left[\mathbf{1}_{[0,\,t]}(T_k)\rho_{T_k}(\mathbf{1})^2 \left[\pi_{T_k}(\varphi^2) - \pi_{T_k-}(\varphi)^2 | \mathcal{F}_{T_k-}\right]\right]. \tag{6.16}$$

Proof. By Proposition 5.3.20 in [12] and its extension to stochastic partial differential equations and infinitely dimensional stochastic differential equations (see [15] and [23]) it follows that the Yamada-Watanabe results applies to equation (6.15), in other words, pathwise uniqueness implies uniqueness in law for (6.15). Hence the solution of (6.15) is unique in distribution.

Let $\{U^{n_r}\}_r$ be any convergent subsequence of $\{U^n\}_n$. We denote the limit of $\{U^{n_r}\}_r$ by U and show that it is a solution of (6.15). The result then follows by the uniqueness in law of the solution of (6.15) since this then implies that the original sequence $\{U^n\}$ converges to the unique solution of of (6.15). Define

$$N_t^{\varphi} = U_t(\varphi) - U_0(\varphi) - \int_0^t U_r(A\varphi)\,\mathrm{d}r - \sum_{k=1}^{m}\int_0^t U_s(h^k\varphi)\,\mathrm{d}Y_r^k,$$

To prove the result it suffices to show that N^{φ} is an $(\mathcal{F}_t \vee \mathcal{Y})$-adapted martingale with quadratic variation given by (6.16). For this it is enough to show that for all $d, d' \geq 0$, $0 \leq t_1 \leq t_2 \cdots \leq t_d \leq s$, $0 \leq t'_1 \leq t'_2 \cdots \leq t'_{d'}$, continuous bounded functions $\alpha_1, \ldots, \alpha_d$ on $\mathcal{M}_F(\mathbb{R}^d)$ and continuous bounded functions $\alpha'_1, \ldots, \alpha'_{d'}$ on $\mathbb{R}^m$ we have

$$\mathbb{E}\left[(N_t^{\varphi} - N_s^{\varphi})\prod_{i=1}^{d}\alpha_i(U_{t_i})\prod_{j=1}^{d'}\alpha'_j(Y_{t'_j})\right] = 0 \tag{6.17}$$

and

$$\begin{aligned}\mathbb{E}\Bigg[\Bigg((N_t^{\varphi} - N_s^{\varphi})^2 &- \int_s^t \tilde{\rho}_s\left((\nabla\varphi)^{\top}\sigma\sigma^{\top}(\nabla\varphi)\right)\,\mathrm{d}s \\ &- \sum_{k=1}^{\infty}\mathbf{1}_{(s,\,t]}\,(T_k)\rho_{T_k}(\mathbf{1})^2\left[\pi_{T_k}(\varphi^2) - \pi_{T_k-}(\varphi)^2\right]\Bigg)\prod_{i=1}^{d}\alpha_i(U_{t_i})\prod_{j=1}^{d'}\alpha'_j(Y_{t'_j})\Bigg] = 0.\end{aligned} \tag{6.18}$$

Both (6.17) and (6.18) follow by using the argument in Theorem 4.8.2 of [8]. The approach used is identical with the proofs of Theorem 4.11 in [3] and Theorem 5.3 in [2] and we omit it here. ☐

Corollary 26. *Let $\bar{U}^n : \{\bar{U}_t^n : t \geq 0\}$ be the process defined as $\bar{U}_t^n := \sqrt{n}(\pi_t^n - \pi_t),\ \ t \geq 0$. Then $\{\bar{U}^n\}_n$ converges in distribution to the measure-valued process $\bar{U} : \{\bar{U}_t : t \geq 0\}$ defined as*

$$\bar{U}_t = \frac{1}{\rho_t(\mathbf{1})}\left(U_t - U_t(\mathbf{1})\pi_t\right), \quad t \geq 0, \tag{6.19}$$

where U satisfies (6.15)*.*

Proof. Follows from (4.33) and the fact that $\rho_t^n(\varphi) \overset{a.s}{\to} \rho_t(\varphi)$ and $\pi_t^n(\varphi) \overset{a.s}{\to} \pi_t(\varphi)$. ☐

Remark 27. *It is possible to obtain the tightness and convergence in distribution of the processes U^n when the space $\mathcal{M}_F(\mathbb{R}^d)$ is endowed with the weak topology in the sense that a sequence of finite measures $\{\mu_n\}_{n\in\mathbb{N}}$ in $\mathcal{M}_F(\mathbb{R}^d)$ converges to $\mu \in \mathcal{M}_F(\mathbb{R}^d)$ if and only if $\mu_n(\varphi)$ converges to $\mu(\varphi)$ for all $\varphi \in \mathcal{C}_b(\mathbb{R}^d)$. To do this one introduces the metric*

$$d_{\mathcal{M}} : \mathcal{M}_F(\mathbb{R}^d) \times \mathcal{M}_F(\mathbb{R}^d) \to [0,\infty), \qquad d_{\mathcal{M}}(\mu,\nu) = \sum_{i=0}^{\infty} \frac{|\mu(\varphi) - \nu(\varphi)|}{2^i \|\varphi\|_\infty}$$

where $\varphi_0 \equiv 1$ and $\{\varphi_i\}_{i\geq 0}$ is a sequence of functions dense in $\mathcal{C}_k(\mathbb{R}^d)$, the space of continuous functions with compact support on $\mathbb{R}^d$. Then $d_{\mathcal{M}}$ generates the weak topology on $\mathcal{M}_F(\mathbb{R}^d)$. The main obstacle to obtaining the tightness and convergence in distribution results under this new metric is that $D_{\mathcal{M}_F(\mathbb{R}^d)}[0,\infty)$ is not complete under $d_{\mathcal{M}}$ since the underlying space $\mathcal{M}_F(\mathbb{R}^d)$ is separable but not complete under $d_{\mathcal{M}}$. This inconvenience is catered for by using the same approach presented in Section 5 of [2]: The space $D_{\mathcal{M}_F(\mathbb{R}^d)}[0,\infty)$ is embedded into the compact and separable space $D_{\mathcal{M}_F\left(\overline{\mathbb{R}^d}\right)}[0,\infty)$ by defining a map or projection $\mathcal{P}$ such that

$$\mu \in \mathcal{M}_F(\overline{\mathbb{R}^d}) \xrightarrow{\mathcal{P}} \mu|_{\mathbb{R}^d} \in \mathcal{M}_F(\mathbb{R}^d).$$

Here $\overline{\mathbb{R}^d}$ is the one point compactification $\mathbb{R}^d$. Note that $\mathcal{P}(\mathcal{M}_F(\mathbb{R}^d)) = \mathcal{M}_F(\mathbb{R}^d)$. The family of measures $\{\mathbb{P}_n\}_n$ can therefore now be viewed as measures over $D_{\mathcal{M}_F\left(\overline{\mathbb{R}^d}\right)}[0,\infty)$ (and $\{U^n\}_{n\in\mathbb{N}}$ consequently can be seen as processes with sample paths in $D_{\mathcal{M}_F\left(\overline{\mathbb{R}^d}\right)}[0,\infty)$). By employing the strategy outlined above, we can show that $\{U^n\}_{n\in\mathbb{N}}$ converges in distribution to U where U has sample paths in $D_{\mathcal{M}_F\left(\overline{\mathbb{R}^d}\right)}[0,\infty)$. Finally since the weak topology on $\mathcal{M}_F(\mathbb{R}^d)$ coincides with the trace topology from $\mathcal{M}_F(\overline{\mathbb{R}^d})$ to $\mathcal{M}_F(\mathbb{R}^d)$, it is enough to show that U only takes values in the space $\mathcal{M}_F(\mathbb{R}^d)$ (i.e. U is indeed a $D_{\mathcal{M}_F(\mathbb{R}^d)}[0,\infty)$-valued random variable). To do this we have to show that $\mathcal{P}(U) = U$. In other words U doesn't 'put' any 'mass at ∞'. To show this, we need to prove that for arbitrary t, there exists a sequence of compact sets $\{K_p\}_{p\geq 0} \in \mathbb{R}^d$ (possibly depending on t) which exhaust $\mathbb{R}^d$ such that for all $\varepsilon > 0$,

$$\lim_{p\to\infty} \mathbb{P}\left[\sup_{s\in[0,t]} \Big(U_s(\mathbf{1}_{K_p^c})\Big)\right] = 0,$$

where K_p^c denotes the compliment of K_p. Consequently it follows that by using the approach described above, we obtain identical results to the ones obtained under the vague topology for the weak topology.

6.4 Tightness and Convergence of Resampling Times in the case of the Effective Sample Size

6.4.1 Tightness Condition

Theorem 28. *Let $\{T_k\}_{k\in\mathbb{N}}$ be the resampling times as determined by the ESS. Then $\{T_k\}_{k\in\mathbb{N}}$ satisfies the tightness condition* (2.10)*. In particular for all $t>0$, we have*

$$\lim_{\delta\to 0}\sup_{n>0}\mathbb{E}[\sup_{s\in[0,t]}\mathbb{E}[(N^n_{s+\delta}-N^n_s)^2|\mathcal{F}_s]]=0,$$

Proof. Since

$$t<T^n_k\le t+\delta \Leftrightarrow \alpha^{n*}_t<k\le \alpha^{n*}_{t+\delta}$$

where $\alpha^{n*}_t := \left(\log\lambda^{-1}\right)^{-1}\sup_{s\in[0,t]}\alpha^n_s$, we get that

$$\mathbb{E}[(N^n_{s+\delta}-N^n_s)^2|\mathcal{F}_s]=\mathbb{E}[([\alpha^{n*}_{s+\delta}]-[\alpha^{n*}_s])^2|\mathcal{F}_s]$$

By the tail sum theorem (see Theorems 4.3.11 and 4.3.12 in [24]) we have that for any random variable X taking integer values,

$$\begin{aligned}\mathbb{E}\left[X^2\right] &= \mathbb{E}\left[X(X-1)\right]+\mathbb{E}\left[X\right]\\ &= 2\sum_{r=1}^{\infty}(r-1)\,\mathbb{P}(X\ge r)+\sum_{r=1}^{\infty}\mathbb{P}(X\ge r).\end{aligned}$$

Hence,

$$\mathbb{E}\left[\left([\alpha^{n*}_{t+\delta}]-[\alpha^{n*}_t]\right)^2|\mathcal{F}_s\right]=2\sum_{r=2}^{\infty}(r-1)\,\mathbb{P}\left[[\alpha^{n*}_{t+\delta}]-[\alpha^{n*}_t]\ge r|\mathcal{F}_s\right]+\sum_{r=1}^{\infty}\mathbb{P}\left[[\alpha^{n*}_{t+\delta}]-[\alpha^{n*}_t]\ge r|\mathcal{F}_s\right]. \qquad (6.20)$$

We will use the convention that K_p, where $p\in\mathbb{N}$ is a constant independent of n. Since for $r>1$

$$\{[\alpha^{n*}_{t+\delta}]-[\alpha^{n*}_t]\ge r\}\subset\{\sup_{s\in[0,\delta]}(\alpha^n_{t+s}-\alpha^n_t)\ge r-1\}$$

it follows by the conditional Markov inequality that

$$
\begin{aligned}
2\sum_{r=2}^{\infty}(r-1)\,\mathbb{P}\left[[\alpha_{t+\delta}^{n*}]-[\alpha_t^{n*}]\geq r|\mathcal{F}_t\right] &\leq 2\sum_{r=2}^{\infty}(r-1)\,\mathbb{P}\left[\sup_{s\in[0,\delta]}(\alpha_{t+s}^n-\alpha_t^n)\geq (r-1)|\mathcal{F}_t\right] \\
&\leq 2\sum_{r=2}^{\infty}(r-1)\frac{\mathbb{E}\left[\left(\sup_{s\in[0,\delta]}(\alpha_{t+s}^n-\alpha_t^n)\right)^3|\mathcal{F}_t\right]}{(r-1)^3} \\
&= 2\mathbb{E}\left[\left(\sup_{s\in[0,\delta]}(\alpha_{t+s}^n-\alpha_t^n)\right)^3|\mathcal{F}_t\right]\sum_{r=2}^{\infty}\frac{1}{(r-1)^2} \\
&= \frac{\pi^2}{3}\mathbb{E}\left[\left(\sup_{s\in[0,\delta]}(\alpha_{t+s}^n-\alpha_t^n)\right)^3|\mathcal{F}_t\right] \\
&\leq K_1\delta^{\frac{3}{2}}. \qquad (6.21)
\end{aligned}
$$

Note that

$$
\sum_{r=1}^{\infty}\mathbb{P}\left[[\alpha_{t+\delta}^{n*}]-[\alpha_t^{n*}]\geq r|\mathcal{F}_t\right]=\mathbb{P}\left[[\alpha_{t+\delta}^{n*}]-[\alpha_t^{n*}]\geq 1|\mathcal{F}_t\right]+\sum_{r=2}^{\infty}\mathbb{P}\left[[\alpha_{t+\delta}^{n*}]-[\alpha_t^{n*}]\geq r|\mathcal{F}_t\right]. \qquad (6.22)
$$

Similar to (6.21)

$$
\begin{aligned}
\sum_{r=2}^{\infty}\mathbb{P}\left[[\alpha_{t+\delta}^{n*}]-[\alpha_t^{n*}]\geq r|\mathcal{F}_t\right] &\leq \mathbb{E}\left[\left(\sup_{s\in[0,\delta]}(\alpha_{t+s}^n-\alpha_t^n)\right)^2|\mathcal{F}_t\right]\sum_{r=2}^{\infty}\frac{1}{(r-1)^2}. \\
&\leq K_2\delta. \qquad (6.23)
\end{aligned}
$$

We will now show that as $\delta\to 0$

$$
\mathbb{P}\left[[\alpha_{t+\delta}^{n*}]-[\alpha_t^{n*}]\geq 1|\mathcal{F}_t\right]\leq c(\delta)\to 0
$$

where $c(\delta)$ is a constant depending only on δ.
For $\varepsilon>0$ let

$$
A^{\varepsilon}=\bigcup_{r\in\mathbb{N}}[r-1,r-\frac{\varepsilon}{r^4}] \quad \text{and} \quad (A^{\varepsilon})^c:=[0,\infty)\backslash A^{\varepsilon}=\bigcup_{r\in\mathbb{N}}(r-\frac{\varepsilon}{r^4},r\ )
$$

then with $B:=\{[\alpha_{t+\delta}^{n*}]-[\alpha_t^{n*}]\geq 1\}$,

$$
\begin{aligned}
\mathbb{P}(B|\mathcal{F}_t) &= \mathbb{P}(B\cap\{\alpha_t^{n*}\in A^{\varepsilon}\}|\mathcal{F}_t)+\mathbb{P}(B\cap\{\alpha_t^{n*}\in(A^{\varepsilon})^c\}|\mathcal{F}_t) \\
&\leq \mathbb{P}(B\cap\{\alpha_t^{n*}\in A^{\varepsilon}\}|\mathcal{F}_t)+\mathbb{P}(\{\alpha_t^{n*}\in(A^{\varepsilon})^c\}|\mathcal{F}_t)
\end{aligned}
$$

Note then that since $\alpha_t^n\overset{a.s.}{\to}\alpha_t$ and $\alpha_t^{n*}\overset{a.s.}{\to}\alpha_t^*$ as $n\to\infty$

$$
\mathbb{P}\left(\{\alpha_t^{n*}\in(A^{\varepsilon})^c\}|\mathcal{F}_t\right)\to\mathbb{P}\left(\{\alpha_t^{*}\in(A^{\varepsilon})^c\}|\mathcal{F}_t\right). \qquad (6.24)
$$

Also as $\varepsilon \to 0$,

$$(A^{\varepsilon})^c \searrow \phi \Rightarrow \mathbb{P}((A^{\varepsilon})^c|\mathcal{F}_t) \searrow 0$$

and hence,

$$f(\varepsilon) := \mathbb{P}(\{\alpha_t^* \in (A^{\varepsilon})\}|\mathcal{F}_t)^c \to 0. \tag{6.25}$$

To control $\mathbb{P}(B \cap \{\alpha_t^{n*} \in A^{\varepsilon}\})$ we make use of the conditional Markov inequality and the fact (see Proposition C.1 in the appendix in [19]) that

$$\{[\alpha_{t+\delta}^{n*}] - [\alpha_t^{n*}] \geq 1\} \subseteq \{\sup_{s\in[0,\delta]} (\alpha_{t+s}^n - \alpha_t^n) \geq 1 - \{\alpha^{n*}\}\}$$

where $\{\cdot\}$ denotes the fractional part function, to get that

$$\begin{aligned}
\mathbb{P}(B \cap \{\alpha_t^{n*} \in A^{\varepsilon}\}|\mathcal{F}_t) &= \mathbb{E}\left[\mathbf{1}_{A^{\varepsilon}}(\alpha_t^{n*})\mathbf{1}_B|\mathcal{F}_t\right] \\
&= \mathbb{E}\left[\mathbf{1}_{A^{\varepsilon}}(\alpha_t^{n*})\mathbf{1}_{\{[\alpha_{t+\delta}^{n*}]-[\alpha_t^{n*}]\geq 1\}}|\mathcal{F}_t\right] \\
&= \mathbf{1}_{A^{\varepsilon}}(\alpha_t^{n*})\mathbb{E}\left[\mathbf{1}_{\{[\alpha_{t+\delta}^{n*}]-[\alpha_t^{n*}]\geq 1\}}|\mathcal{F}_t\right] \\
&\leq \mathbf{1}_{A^{\varepsilon}}(\alpha_t^{n*})\mathbb{P}\left[\sup_{s\in[0,\delta]}\left(\alpha_{t+s}^n - \alpha_t^n\right) \geq 1 - \{\alpha_t^{n*}\}\middle|\mathcal{F}_t\right] \\
&\leq \frac{\mathbf{1}_{A^{\varepsilon}}(\alpha_t^{n*})\mathbb{E}\left[\sqrt{\sup_{s\in[0,\delta]}\left(\alpha_{t+s}^n - \alpha_t^n\right)}\middle|\mathcal{F}_t\right]}{\sqrt{1-\{\alpha_t^{n*}\}}}.
\end{aligned}$$

By observing that

$$\begin{aligned}
\sum_{r=1}^{\infty}(r-1)^2\mathbf{1}_{\left[r-1,\, r-\frac{\varepsilon}{r^4}\right]}(\alpha_t^{n*}) &\leq \sum_{r=1}^{\infty}(r-1)^2\mathbf{1}_{[r-1\,,\,r\,)}(\alpha_t^{n*}) \\
&= [\alpha_t^{n*}]^2 \\
&\leq (\alpha_t^{n*})^2
\end{aligned} \tag{6.26}$$

and noting similarly that

$$\sum_{r=1}^{\infty}(r-1)\mathbf{1}_{\left[r-1,\, r-\frac{\varepsilon}{r^4}\right]}(\alpha_t^{n*}) \leq (\alpha_t^{n*}) \tag{6.27}$$

and

$$\sum_{r=1}^{\infty}\mathbf{1}_{\left[r-1,\, r-\frac{\varepsilon}{r^4}\right]}(\alpha_t^{n*}) \leq 1, \tag{6.28}$$

it follows thus that

$$
\begin{aligned}
&\frac{\mathbb{E}\left[\mathbf{1}_{A^{\varepsilon}}(\alpha_t^{n*})\sqrt{\sup_{s\in[0,\delta]}\left(\alpha_{t+s}^n-\alpha_t^n\right)}\middle|\mathcal{F}_t\right]}{\sqrt{1-\{\alpha_t^{n*}\}}}\\
\leq\ & \frac{1}{\sqrt{\varepsilon}}\sum_{r=1}^{\infty} r^2\,\mathbb{E}\left[\mathbf{1}_{[r-1,r-\frac{\varepsilon}{r^4})}(\alpha_t^{n*})\sqrt{\sup_{s\in[0,\delta]}\left(\alpha_{t+s}^n-\alpha_t^n\right)}|\mathcal{F}_t\right]\\
=\ & \frac{1}{\sqrt{\varepsilon}}\sum_{r=1}^{\infty}\left((r-1)^2+2\,(r-1)+1\right)\mathbb{E}\left[\mathbf{1}_{[r-1,r-\frac{\varepsilon}{r^4})}(\alpha_t^{n*})\sqrt{\sup_{s\in[0,\delta]}\left(\alpha_{t+s}^n-\alpha_t^n\right)}|\mathcal{F}_t\right]\\
\leq\ & \frac{1}{\sqrt{\varepsilon}}\left\{\mathbb{E}\left[(\alpha_t^{n*})^2\sqrt{\sup_{s\in[0,\delta]}\left(\alpha_{t+s}^n-\alpha_t^n\right)}|\mathcal{F}_t\right]+2\,\mathbb{E}\left[(\alpha_t^{n*})\sqrt{\sup_{s\in[0,\delta]}\left(\alpha_{t+s}^n-\alpha_t^n\right)}|\mathcal{F}_t\right]\right.\\
&\left.\qquad\qquad\qquad\qquad+\,\mathbb{E}\left[\sqrt{\sup_{s\in[0,\delta]}\left(\alpha_{t+s}^n-\alpha_t^n\right)}|\mathcal{F}_t\right]\right\}\\
\leq\ & \frac{1}{\sqrt{\varepsilon}}\mathbb{E}\left[\sup_{s\in[0,\delta]}\left(\alpha_{t+s}^n-\alpha_t^n\right)|\mathcal{F}_t\right]^{\frac{1}{2}}\left[\mathbb{E}\left[(\alpha_t^{n*}|\mathcal{F}_t)^4\right]^{\frac{1}{2}}+\mathbb{E}\left[(\alpha_t^{n*})^2\,|\mathcal{F}_t\right]^{\frac{1}{2}}+1\right]\\
\leq\ & \frac{1}{\sqrt{\varepsilon}}K_3\delta^{\frac{1}{2}}.
\end{aligned}
\tag{6.29}
$$

Hence it follows from (6.24), (6.25) and (6.29) and choosing $\varepsilon=\delta^{\frac{1}{2}}$ that

$$
\begin{aligned}
\mathbb{P}\left[[\alpha_{t+\delta}^{n*}]-[\alpha_t^{n*}]\geq 1|\mathcal{F}_t\right] &\leq \frac{1}{\sqrt{\varepsilon}}K_3\delta^{\frac{1}{2}}+f(\varepsilon)\\
&\leq K_3\delta^{\frac{3}{8}}+f(\delta^{\frac{1}{2}})
\end{aligned}
\tag{6.30}
$$

and so from (6.21), (6.22), (6.23) and (6.30),

$$
\begin{aligned}
\mathbb{E}\left[\left([\alpha_{t+\delta}^{n*}]-[\alpha_t^{n*}]\right)^2|\mathcal{F}_t\right] &= 2\sum_{r=2}^{\infty}(r-1)\,\mathbb{P}\left[[\alpha_{t+\delta}^{n*}]-[\alpha_t^{n*}]\geq r|\mathcal{F}_t\right]+\sum_{r=1}^{\infty}\mathbb{P}\left[[\alpha_{t+\delta}^{n*}]-[\alpha_t^{n*}]\geq r|\mathcal{F}_t\right]\\
&\leq K_1\delta^{\frac{3}{2}}+K_2\delta+K_3\delta^{\frac{3}{8}}+f(\delta^{\frac{1}{2}})\\
&\leq K_4\delta^{\frac{3}{8}}+f(\delta^{\frac{1}{2}}).
\end{aligned}
$$

(2.10) now holds as $\delta\to 0$. □

6.4.2 Convergence of Resampling Times

Recall that the predictable stopping times as determined by the ESS is defined by

$$
T_{k+1}^{n,\lambda}\equiv T_{k+1}^n=\inf\{t\geq T_k:\alpha_t^{n,k}\geq\log\lambda^{-1}\}.
\tag{6.31}
$$

The notation $T_k^{n,\lambda}$ is used the emphasise the dependence on the threshold λ.
Now define

$$\alpha_t^k := \sum_{r=1}^m \int_0^t \tilde{\eta}_s^r \, \mathrm{d}Y_s^r + \sum_{r=1}^m \int_0^t \tilde{\zeta}_s^r \, \mathrm{d}s - \frac{1}{2} \sum_{r=1}^m \int_0^t (\tilde{\eta}_s^r)^2 \, \mathrm{d}s$$

where $\tilde{\eta}_s^r := \lim_n \tilde{\eta}_s^{n,r}$ and $\tilde{\zeta}_s^r := \lim_n \tilde{\zeta}_s^{n,r}$ with $\tilde{\eta}_s^{n,r}$ and $\tilde{\zeta}_s^{n,r}$ as in (5.1) and (5.2) so that

$$\tilde{\eta}_s^r = 2 \left(\pi_t \left(h^r \right) - \tilde{\pi}_t \left(h^r \right) \right),$$

and

$$\tilde{\zeta}_s^r = \pi_t \left(h^r \right)^2 + 2\pi_t \left(h^r \right) \tilde{\pi}_t \left(h^r \right) - \tilde{\pi}_t \left((h^r)^2 \right) + 4\tilde{\pi}_t \left(h^k \right)^2.$$

We observe that $\alpha_t^{n,k} \overset{a.s.}{\to} \alpha_t^k$. Ignoring '$k$' in $\alpha_t^{n,k}$ and α_t^k, recall that

$$T_k^{n,\lambda} \equiv T_k^n = \inf\{t \geq 0 : \alpha_t^n \geq k \log \lambda^{-1}\}. \tag{6.32}$$

Now let

$$T_k^\lambda := \inf\{t \geq 0 : \alpha_t \geq k \log \lambda^{-1}\} \tag{6.33}$$

where α_t is defined (and obtained) in a similar manner to α_t^n in (5.10). Then clearly for $k \in \mathbb{N}$ T_k^λ is a stopping time.

Theorem 29. $T_k^{n,\lambda} \overset{a.s.}{\to} T_k^\lambda$

Proof. As already indicated in Remark 13 we can show for any $p > 0$, $T > 0$ and $\varphi \in \mathcal{C}_b^3(\mathbb{R}^d)$ that there exists constants $c_{p,t}$ and $\bar{c}_{p,t}$ such that

$$\mathbb{E}\left[\sup_{t\in[0,T']} \delta_{\rho_t^n}^{2p}(\varphi) \right] \leq \frac{c_{p,t}}{n^p} \quad \text{and} \quad \mathbb{E}\left[\sup_{t\in[0,T']} \delta_{\pi_t^n}^{2p}(\varphi) \right] \leq \frac{\bar{c}_{p,t}}{n^p}.$$

Similarly we can show that

$$\mathbb{E}\left[\sup_{t\in[0,T']} \delta_{\tilde{\pi}_t^n}^{2p}(\varphi) \right] \leq \frac{\tilde{c}_{p,t}}{n^p},$$

where $\tilde{\pi}$ is as defined in (5.3).
It follows thus that for any $p > 0$ and $T > 0$,

$$\mathbb{E}\left[\sup_{t\in[0,T]} |\alpha_t^n - \alpha_t|^{2p} \right] \leq \frac{c_{p,T}}{n^p}$$

which, by choosing $p > 1$ and appealing to the proof of Lemma 5, implies that

$$\sup_{t\in[0,T]} |\alpha_t^n - \alpha_t| \to 0.$$

Hence for any $\varepsilon' > 0$ there exist $N \in \mathbb{N}$ such that for $n > N$,

$$\sup_{t\in[0,T]} |\alpha_t^n - \alpha_t| < \varepsilon'.$$

Now let $\varepsilon > 0$ be such that $(\lambda + \varepsilon) \in (0, 1)$ and let

$$T_k^{\lambda-\varepsilon} := \inf\{t \geq 0 : \alpha_t \geq k \log(\lambda + \varepsilon)^{-1}\}$$

so that for $t \in [0, T_k^{\lambda-\varepsilon}]$, $\alpha_t \leq k \log(\lambda + \varepsilon)^{-1}$.
Also let $\varepsilon' = \frac{1}{2} \min \left(k \left(\log(\lambda - \varepsilon)^{-1} - \log \lambda^{-1}\right), k \left(\log \lambda^{-1} - \log(\lambda + \varepsilon)^{-1}\right)\right) > 0$ then for $n > N$ and $t \in [0, T_k^{\lambda-\varepsilon}]$:

$$\alpha_t^n \quad = \quad \alpha_t^n - \alpha_t + \alpha_t < k \log \lambda^{-1}$$

and hence by (6.32) we have that $T_k^{\lambda-\varepsilon} \leq T_k^{n,\lambda}$.
By a similar, symmetric argument we conclude also that $T_k^{n,\lambda} \leq T_k^{\lambda+\varepsilon}$.
So

$$T_k^{\lambda-\varepsilon} \leq T_k^{n,\lambda} \leq T_k^{\lambda+\varepsilon}$$

and hence for $k \in \mathbb{N}$, $T_k^{n,\lambda} \overset{a.s.}{\to} T_k^{\lambda}$ as $n \to \infty$. □

Remark 30. *That the (soft) maximum also satisfies the tightness condition* (2.10) *and the convergence of resampling times (as determined by it) follows using an approach similar to case of the effective sample size.*

7 Appendix

We include below a number of useful lemmas and results use throughout the paper.

Lemma 31. *For any $t \geq 0$ and $p \geq 1$ there exists constants $c_0^{t,p}$, $c_1^{t,p}$ and $c_2^{t,p}$ which depend on $\max_{i=1,\dots,m} \|h_i\|_{0,\infty}$ such that the bounds* (4.1)*,* (4.2) *and* (4.3) *hold true.*

Proof. Let $\tilde{P}$ be a measure absolutely continuous with respect to P given by

$$\left.\frac{d\tilde{P}}{dP}\right|_{\mathcal{F}_t} = \frac{1}{\tilde{Z}_t},$$

where $\tilde{Z} = \{\tilde{Z}_t : t \geq 0\}$ is the process given by

$$\tilde{Z}_t = \exp\left(\int_0^t h(X(s))^\top \,\mathrm{d}Y_s - \frac{1}{2}\int_{T_k}^T \|h(X(s))\|^2 \,\mathrm{d}s\right) \quad t \geq 0. \tag{7.1}$$

Under $\tilde{P}$ the process Y is a Brownian motion and both the unnormalized weights and the process ξ_n are $\mathcal{F}_t$-adapted martingale. The proof of the estimates (under $\tilde{P}$) then follows by a standard argument using Burkholder-Davis-Gundy's inequality (see for example Exercise 9.10 in [1] and [19]). Finally observe that, for example

$$\mathbb{E}[\sup_{s\in[0,t]}(a_j^n(s))^p] = \tilde{\mathbb{E}}[\sup_{s\in[0,t]}(a_j^n(s))^p\tilde{Z}_t] \leq \sqrt{\tilde{\mathbb{E}}[\sup_{s\in[0,t]}(a_j^n(s))^{2p}]\tilde{\mathbb{E}}[\tilde{Z}_t^2]},$$

which give us the corresponding bounds under P after observing that $\tilde{Z}_t$ has finite second moment under $\tilde{P}$. □

Lemma 32. *Let S and T be stopping times such that $S < T$ and T is predictable then $\mathcal{F}_S \subset \mathcal{F}_{T-}$.*

Proof. Let (T_n) be an announcing sequence for T. Then $(S \vee T_n)_n$ is also an announcing sequence for T. In particular, $S \vee T_n \nearrow S \vee T = T$, hence $\mathcal{F}_S \subset \mathcal{F}_{S\vee T} \subset \bigvee_n \mathcal{F}_{S\vee T_n} = \mathcal{F}_{T-}$. □

Proposition 33. *The process $\bar{M}^{n,\varphi} := \{\bar{M}_t^{n,\varphi} : t \geq 0\}$ defined by*

$$\begin{aligned}\bar{M}_t^{n,\varphi} &:= \sum_{k=1}^{\infty} \mathbf{1}_{[0,t]}(T_k)(\rho_{T_k}^n(\varphi) - \rho_{T_k-}^n(\varphi)) \\ &\equiv \sum_{k=1}^{\infty} \mathbf{1}_{[0,t]}(T_k)\xi_{T_k}^{n,\infty}(\pi_{T_k}^n(\varphi) - \pi_{T_k-}^n(\varphi))\end{aligned} \tag{7.2}$$

where $\varphi \in \mathcal{C}_b(\mathbb{R}^d)$ is an $\mathcal{F}_t$-adapted martingale provided there exists $p > 1$ such that condition (2.8) *holds true.*

Proof. Let $\bar{\eta}(t) := \{\bar{\eta}_m(t) : m \in \mathbb{N}\}$ and $\bar{M}_t^{n,\varphi} \equiv \bar{\eta}_\infty(t)$ where

$$\bar{\eta}_m(t) := \sum_{k=1}^{m} \mathbf{1}_{[0,t]}(T_k)\xi_{T_k}^{n,\infty}(\pi_{T_k}^n(\varphi) - \pi_{T_k-}^n(\varphi)),$$

and

$$\bar{\eta}_\infty(t) := \sum_{k=1}^{\infty} \mathbf{1}_{[0,t]}(T_k)\xi_{T_k}^{n,\infty}(\pi_{T_k}^n(\varphi) - \pi_{T_k-}^n(\varphi)).$$

We will show that $t \mapsto \bar{\eta}_m(t)$ is a $\mathcal{F}_t$-martingale that is

$$\mathbb{E}[\bar{\eta}_m(t)|\mathcal{F}_s] = \bar{\eta}_m(s), \ \ \forall s \leq t.$$

By linearity it suffices to show for any $k \in \mathbb{N}$ that

$$\mathbb{E}[\mathbf{1}_{[0,t]}(T_k)\xi_{T_k}^n(\pi_{T_k}^n(\varphi) - \pi_{T_k-}^n(\varphi))|\mathcal{F}_s] = \mathbf{1}_{[0,s]}(T_k)\xi_{T_k}^n(\pi_{T_k}^n(\varphi) - \pi_{T_k-}^n(\varphi)) \tag{7.3}$$

and noting that

$$\mathbf{1}_{[0,s]}(T_k)\xi_{T_k}^n(\pi_{T_k}^n(\varphi) - \pi_{T_k-}^n(\varphi))$$

is $\mathcal{F}_{T_k \wedge s}$-measurable and hence $\mathcal{F}_s$-measurable. It follows that to obtain (7.3), it now remains to show that

$$\mathbb{E}[\mathbf{1}_A \mathbf{1}_{(s,t]}(T_k)\xi^n_{T_k}(\pi^n_{T_k}(\varphi) - \pi^n_{T_k-}(\varphi))] = 0, \quad \forall A \in \mathcal{F}_s.$$

To this extent note that

$$\begin{aligned}
&\mathbb{E}\left[\mathbf{1}_A \mathbf{1}_{(s,t]}(T_k)\xi^n_{T_k}(\pi^n_{T_k}(\varphi) - \pi^n_{T_k-}(\varphi))\right] \\
&\quad = \mathbb{E}[(\mathbf{1}_A \mathbf{1}_{(s,\infty)} - \mathbf{1}_A \mathbf{1}_{(t,\infty)})(T_k)\xi^n_{T_k}(\pi^n_{T_k}(\varphi) - \pi^n_{T_k-}(\varphi))] \\
&\quad = \mathbb{E}\left[(\mathbf{1}_A \mathbf{1}_{(s,\infty)} - \mathbf{1}_A \mathbf{1}_{(t,\infty)})(T_k)\xi^n_{T_k}\, \mathbb{E}[(\pi^n_{T_k}(\varphi) - \pi^n_{T_k-}(\varphi))|\mathcal{F}_{T_k-}]\right] \\
&\quad = 0
\end{aligned}$$

since $\mathbf{1}_A \mathbf{1}_{(s,\infty)}(T_k)$ and $\mathbf{1}_A \mathbf{1}_{(t,\infty)}(T_k)$ corresponds to the $\mathcal{F}_{T_{k-}}$-measurable sets $A \cap \{s < T_k\}$ and $A \cap \{t < T_k\}$ respectively (see Theorem 7 on pp. 106 of [20]).

We will now show that $m \mapsto \bar{\eta}_m(t)$, where $t \geq 0$, is a $\mathcal{F}_{T_m \wedge t}$-adapted martingale. That is, for any $m \in \mathbb{N}$,

$$\mathbb{E}[\bar{\eta}_{m+1}|\mathcal{F}_{T_m \wedge t}] = \bar{\eta}_m(t).$$

Since for $1 \leq k \leq m$

$$\mathbf{1}_{[0,t]}(T_k)\xi^n_{T_k}(\pi^n_{T_k}(\varphi) - \pi^n_{T_{k-}}(\varphi))$$

is $\mathcal{F}_{T_m \wedge t}$-measurable, the result will then follow if

$$\mathbb{E}\left[\mathbf{1}_A \mathbf{1}_{[0,t]}(T_{m+1})\xi^n_{T_{m+1}}(\pi^n_{T_{m+1}}(\varphi) - \pi^n_{T_{m+1-}}(\varphi))\right] = 0, \;\; \forall A \in \mathcal{F}_{T_m \wedge t}.$$

By Lemma 32 and the tower property

$$\begin{aligned}
&\mathbb{E}\left[\mathbf{1}_A \mathbf{1}_{[0,t]}(T_{m+1})\xi^n_{T_k}(\pi^n_{T_{m+1}}(\varphi) - \pi^n_{T_{m+1-}}(\varphi))\right] \\
&\quad = \mathbb{E}\left[\mathbf{1}_A \mathbf{1}_{[0,t]}(T_{m+1})\xi^n_{T_{m+1}}\, \mathbb{E}[(\pi^n_{T_{m+1}}(\varphi) - \pi^n_{T_{m+1-}}(\varphi))|\mathcal{F}_{T_{m+1-}}]\right] \\
&\quad = 0
\end{aligned}$$

and so, $(\bar{\eta}_m(t))_m$ is a $\mathcal{F}_{T_m \wedge t}$-adapted martingale.
We now proceed to show that $\bar{\eta}_\infty(t) \equiv M^{n,\varphi}_t$ exists almost surely, is finite, integrable and $(\bar{\eta}_\infty(t))_t$ is an $\mathcal{F}_t$-adapted martingale provided condition (2.8) is satisfied. Observe that

$$\sup\{|\bar{\eta}_m(t)|, m \geq 1, |\bar{\eta}_\infty(t)\}| \leq \|\varphi\|_{0,\infty} \sum_{k=1}^{\infty} \mathbf{1}_{[0,t]}\,(T_k)\xi^{n,\infty}_{T_k} \leq \|\varphi\|_{0,\infty} N^n_t \sup_{s\in[0,t]} \xi^{n,\infty}_s$$

Hence $\bar{\eta}(t) := \{\bar{\eta}_m(t) : m \in \mathbb{N}\}$ is bounded in $\mathcal{L}^r$ for any $r < p$ (using condition (2.8) and Hölder's inquality) which implies that $\bar{\eta}(t)$ is bounded in $\mathcal{L}^1$ and is a uniformly integrable martingale.

The boundedness of $\bar{\eta}(t)$ in $\mathcal{L}^1$ implies the almost sure existence and finiteness of $\bar{\eta}_\infty(t)$(see II.49, Theorem 49.1 in [22]). By II.50, Theorem 50.1 in [22], we then have that $\bar{\eta}_m(t) \to \bar{\eta}_\infty(t)$ in $\mathcal{L}^1$, that is, $\mathbb{E}[|\bar{\eta}_m(t) - \bar{\eta}_\infty(t)|] \to 0$. Furthermore $\forall s \leq t$,

$$\begin{aligned}
\mathbb{E}\left[\Big|\mathbb{E}[\bar{\eta}_m(t)|\mathcal{F}_s] - \mathbb{E}[\bar{\eta}_\infty(t)|\mathcal{F}_s]\Big|\right] &\leq \mathbb{E}\left[\mathbb{E}\left[|\bar{\eta}_m(t) - \bar{\eta}_\infty(t)|\Big|\mathcal{F}_s\right]\right] \\
&= \mathbb{E}[|\bar{\eta}_m(t) - \bar{\eta}_\infty(t)|] \\
&\to 0,
\end{aligned}$$

that is, $\mathbb{E}[\bar{\eta}_m(t)|\mathcal{F}_s] \to \mathbb{E}[\bar{\eta}_\infty(t)|\mathcal{F}_s]$ in $\mathcal{L}^1$ and since $\mathbb{E}[\bar{\eta}_m(t)|\mathcal{F}_s] = \bar{\eta}_m(s) \to \bar{\eta}_\infty(s)$ in $\mathcal{L}^1$ it follows that

$$\mathbb{E}[\bar{\eta}_\infty(t)|\mathcal{F}_s] = \bar{\eta}_\infty(s), \quad \forall s \leq t.$$

Hence $(\bar{\eta}_\infty(t))_t$ is an $\mathcal{F}_t$-adapted martingale. We now proceed to show that $\bar{\eta}(t) := \{\bar{\eta}_m(t) : m \in \mathbb{N}\}$ is indeed bounded in $\mathcal{L}^2$. We need to first show that $\bar{\eta}_m(t) \in \mathcal{L}^2$, for all $m \in \mathbb{N}$ and by II.53, Theorem 53.3 in [22], the boundedness property follows if and only if

$$\sum_{m=1}^{\infty} \mathbb{E}[(\bar{\eta}_m(t) - \bar{\eta}_{m-1}(t))^2] < \infty. \tag{7.4}$$

Observe that for any integer $k \geq 2$,

$$\bar{\eta}_k(t) - \bar{\eta}_{k-1}(t) = \mathbf{1}_{[0,t]}(T_k)\xi^n_{T_k}(\pi^n_{T_k}(\varphi) - \pi^n_{T_{k-}}(\varphi))$$

with $\bar{\eta}_0(t) := 0$ so that

$$\begin{aligned}
\mathbb{E}[(\bar{\eta}_k(t) - \bar{\eta}_{k-1}(t))^2] &= \mathbb{E}\left[\mathbf{1}_{[0,t]}(T_k)\xi^n_{T_k}(\pi^n_{T_k}(\varphi) - \pi^n_{T_{k-}}(\varphi))^2\right] \\
&= \mathbb{E}\left[\mathbf{1}_{[0,t]}(T_k)\xi^n_{T_k}\,\mathbb{E}[(\pi^n_{T_k}(\varphi) - \pi^n_{T_{k-}}(\varphi))^2|\mathcal{F}_{T_{k-}}]\right] \\
&= \mathbb{E}\left[\mathbf{1}_{[0,t]}(T_k)\xi^n_{T_k}\,\mathbb{E}\left[(\pi^n_{T_k}(\varphi) - \mathbb{E}[\pi^n_{T_k}(\varphi)|\mathcal{F}_{T_{k-}}])^2\Big|\mathcal{F}_{T_{k-}}\right]\right].
\end{aligned} \tag{7.5}$$

Now,

$$\pi^n_{T_k}(\varphi) - \mathbb{E}[\pi^n_{T_k}(\varphi)|\mathcal{F}_{T_{k-}}] = \frac{1}{n}\sum_{j=1}^{n}\Big(\varphi(v^n_{\alpha_j}) - \mathbb{E}[\pi^n_{T_k}(\varphi)|\mathcal{F}_{T_{k-}}]\Big)$$

where $\alpha_j, j = 1, \ldots, n$ is a random index of $1, \ldots, n$ and $v^n_{\alpha_j}(T_k) = v^n_l(T_k)$ for some $l \in \{1, \ldots, n\}$ with probability $\bar{a}^{n,T_k}_l$ so that

$$\mathbb{E}\left[\varphi(v_{\alpha_j}(T_k))|\mathcal{F}_{T_{k-}}\right] = \sum_{l=1}^{n}\bar{a}^{n,T_k}_l\varphi(v^n_l(T_k)) = \mathbb{E}[\pi^n_{T_k}(\varphi)|\mathcal{F}_{T_{k-}}]$$

that is,

$$\mathbb{E}\Big[\Big(\varphi(v_{\alpha_j}(T_k)) - \mathbb{E}[\pi^n_{T_k}(\varphi)|\mathcal{F}_{T_{k-}}]\Big)\Big] = 0$$

and so

$$\begin{aligned}
\mathbb{E}\Big[\Big(\pi^n_{T_k}(\varphi) - \mathbb{E}[\pi^n_{T_k}(\varphi)|\mathcal{F}_{T_{k-}}]\Big)^2\Big|\mathcal{F}_{T_{k-}}\Big] &= \mathbb{E}\left[\frac{1}{n^2}\left(\sum_{j=1}^n\Big(\varphi(v^n_{\alpha_j}) - \mathbb{E}[\pi^n_{T_k}|\mathcal{F}_{T_{k-}}]\Big)\right)^2\Big|\mathcal{F}_{T_{k-}}\right] \\
&= \frac{1}{n^2}\sum_{j=1}^n \mathbb{E}\Big[(\varphi(v^n_{\alpha_j}) - \mathbb{E}[\pi^n_{T_k}|\mathcal{F}_{T_{k-}}])^2\Big|\mathcal{F}_{T_k}\Big] \\
&= \frac{1}{n^2}\sum_{j=1}^n\left(\mathbb{E}[(\varphi(v^n_{\alpha_j})^2|\mathcal{F}_{T_{k-}}] - \Big(\mathbb{E}[\pi^n_{T_k}|\mathcal{F}_{T_{k-}}]\Big)^2\right) \\
&\leq \frac{1}{n^2}\sum_{j=1}^n \mathbb{E}[(\varphi(v^n_{\alpha_j})^2|\mathcal{F}_{T_{k-}}] \\
&\leq \frac{\|\varphi\|^2}{n}. \qquad (7.6)
\end{aligned}$$

Using (7.6) and (7.5) we get

$$\begin{aligned}
\sum_{k=1}^{\infty}\mathbb{E}[(\bar{\eta}_k(t) - \bar{\eta}_{k-1}(t))^2] &= \mathbb{E}\left[\mathbf{1}_{[0,t]}(T_k)\xi^n_{T_k}\,\mathbb{E}\Big[(\pi^n_{T_k}(\varphi) - \mathbb{E}[\pi^n_{T_k}(\varphi)|\mathcal{F}_{T_{k-}}])^2\Big|\mathcal{F}_{T_{k-}}\Big]\right] \\
&\leq \frac{\|\varphi\|^2}{n}\sum_{k=1}^{\infty}\mathbb{E}[\mathbf{1}_{[0,t]}(T_k)\xi^n_{T_k}] \leq \frac{\|\varphi\|^2}{n}\mathbb{E}[N^n_t \sup_{s\in[0,t]}\xi^{n,\infty}_s]
\end{aligned}$$

Therefore $\bar{\eta}(t)$ is a martingale bounded in $\mathcal{L}^2$ again by using condition (2.8) and Hölder's inquality. □

References

[1] A. Bain and D. Crisan. *Fundamentals of Stochastic Filtering*, volume 60 of *Stochastic modelling and Applied Probability*. Springer, 2008.

[2] D. Crisan. Superprocesses in a Brownian environment. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 460(2041): 243–270, 2004. ISSN 1364-5021. Stochastic analysis with applications to mathematical finance.

[3] D. Crisan and T. Lyons. Nonlinear filtering and measure-valued processes. *Probab. Theory Related Fields*, 109(2):217–244, 1997. ISSN 0178-8051.

[4] D. Crisan and B. Rozovsky, editors. *The Oxford Handbook of Nonlinear Filtering*. OUP, 2011.

[5] P. Del Moral, A. Doucet, and A. Jasra. On adaptive sequential monte carlo methods. *Institut National de Recherche en Informatique et en Automatique (INRIA)*, (6438), May 2008.

[6] A. Doucet, N. de Freitas, and N. Gordon, editors. *Sequential Monte Carlo methods in practice*. Statistics for Engineering and Information Science. Springer-Verlag, New York, 2001. ISBN 0-387-95146-6.

[7] A. Doucet, S. Godsill, and C. Andrieu. On sequential monte carlo sampling methods for bayesian filtering. *Statistics and Computing*, 10(3):197–208, 2000. ISSN 0960-3174. doi: 10.1023/A:1008935410038.

[8] S.N. Ethier and T.G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. ISBN 0-471-08186-8. Characterization and convergence.

[9] N.J. Gordon, D.J. Salmond, and C Ewing. Bayesian state estimation for tracking and guidance using the bootstrap filter. *Journal of Guidance, Control, and Dynamics*, 18(6):1434–1443, 1995.

[10] N.J. Gordon, D.J. Salmond, and A.F.M. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. *Radar and Signal Processing, IEE Proceedings F*, 140(2):107 –113, Apr 1993. ISSN 0956-375X.

[11] S.L. Jun. Metropolized independent sampling with comparisons to rejection sampling and importance sampling. *Statistics and Computing*, 6:113–119., 1996.

[12] I. Karatzas and S.E. Shreve. *Brownian motion and Stochastic Calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[13] G. Kitagawa. Monte carlo filter and smoother for non-gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics*, 5(1):1–25, 1996. ISSN 10618600. URL `http://www.jstor.org/stable/1390750`.

[14] A. Kong, J.S. Liu, and W.H. Wong. Sequential imputations and bayesian missing data problems. *Journal of the American Statistical Association*, 89(425):278–288, 1994. ISSN 01621459. URL `http://www.jstor.org/stable/2291224`.

[15] T.G. Kurtz. The Yamada-Watanabe-Engelbert theorem for general stochastic equations and inequalities. *Electron. J. Probab.*, 12:951–965, 2007. ISSN 1083-6489.

[16] T.G. Kurtz and J. Xiong. A stochastic evolution equation arising from the fluctuations of a class of interacting particle systems. *Commun. Math. Sci.*, 2(3):325–358, 2004. ISSN 1539-6746. URL `http://projecteuclid.org/getRecord?id=euclid.cms/1109868725`.

[17] V.M. Lucic and A.J. Heunis. On uniqueness of solutions for the stochastic differential equations of nonlinear filtering. *Ann. Appl. Probab.*, 11(1):182–209, 2001. ISSN 1050-5164. doi: 10.1214/aoap/998926990. URL `http://dx.doi.org/10.1214/aoap/998926990`.

[18] P. Del Moral. *Feynman-Kac formulae. Genealogical and Interacting Particle Systems with Applications*. Springer, New York, 2004.

[19] O. Obanubi. *Particle Filters with Random Resampling Times*. PhD thesis, Imperial College London, 2011.

[20] P.E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2005. ISBN 3-540-00313-4. Second edition. Version 2.1, Corrected third printing.

[21] S. Roelly-Coppoletta. A criterion of convergence of measure-valued processes: application to measure branching processes. *Stochastics*, 17(1-2):43–65, 1986. ISSN 0090-9491.

[22] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 1*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. ISBN 0-521-77594-9. Foundations, Reprint of the second (1994) edition.

[23] M. Röckner, B. Schmuland, and X. Zhang. Yamada-watanabe theorem for stochastic evolution equations in. *Condensed Matter Physics*, 11(2(54)):247–259, 2008.

[24] D. Stirzaker. *Elementary probability*. Cambridge University Press, Cambridge, second edition, 2003. ISBN 0-521-83344-2; 0-521-53428-3.

[25] J. Szpirglas. Sur l'équivalence d'équations différentielles stochastiques à valeurs mesures intervenant dans le filtrage markovien non linéaire. *Ann. Inst. H. Poincaré Sect. B (N.S.)*, 14(1):33–59, 1978. ISSN 0020-2347.