

Ondřej Straka; Miroslav Šimandl
Particle filter with adaptive sample size

Kybernetika, Vol. 47 (2011), No. 3, 385--400

Persistent URL: <http://dml.cz/dmlcz/141592>

Terms of use:

© Institute of Information Theory and Automation AS CR, 2011

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://project.dml.cz>

PARTICLE FILTER WITH ADAPTIVE SAMPLE SIZE

ONDŘEJ STRAKA AND MIROSLAV ŠIMANDL

The paper deals with the particle filter in state estimation of a discrete-time nonlinear non-Gaussian system. The goal of the paper is to design a sample size adaptation technique to guarantee a quality of a filtering estimate produced by the particle filter which is an approximation of the true filtering estimate. The quality is given by a difference between the approximate filtering estimate and the true filtering estimate. The estimate may be a point estimate or a probability density function estimate. The proposed technique adapts the sample size to keep the difference within pre-specified bounds with a pre-specified probability. The particle filter with the proposed sample size adaptation technique is illustrated in a numerical example.

Keywords: stochastic systems, nonlinear filtering, particle filter, sample size, adaptation

1. INTRODUCTION

Recursive state estimation of discrete-time nonlinear stochastic dynamic systems from noisy measurement data is an essential element in many applications involving control and cognition of complex stochastic systems. Over the last four decades it has been a subject of a considerable research interest. General solution to the state estimation problem is described by the Bayesian recursive relations (BRR's). The closed form solution to the BRR's is available for a few special cases only and therefore, usually an approximative solution has to be applied.

Since the nineties, the particle filter (PF) has dominated in the recursive nonlinear state estimation due to its easy implementation in very general settings and cheap and formidable computational power. The PF solves the BRR's using Monte Carlo (MC) methods, particularly using the importance sampling method, and approximates the continuous state space by a swarm of samples (particles) with associated relative weights.

The fundamental paper dealing with the MC solution to the BRR's was published in the 90's [1] where the first efficient PF, called the bootstrap filter, was proposed. Many improvements of the bootstrap filter have been proposed since, see for example [2]. Among these improvements, in particular the design of the sampling probability density function (pdf) [3] as one of the key parameters of the PF has to be mentioned.

Another key parameter of the PF, significantly affecting the estimate quality, is the sample size (i. e. the number of the particles), nonetheless efficient sample size setting has been disregarded for a long time. Vague recommendations of sample

size specification limits applicability and user-friendliness of the PF. The sample size is usually determined ad hoc. Some advances in the sample size setting were achieved in [4] where the time-invariant sample size was considered and the Cramér Rao bound [5] was used as a gauge for quality evaluation of the PF. Sample size adaptation (SSA) techniques were treated for example in [6, 7, 8, 9, 10]. In [3] a survey of the SSA techniques has been provided.

The goal of this paper is to propose a generalization of the SSA technique proposed in [11], which adapted the sample size according to the filtering pdf quality. The generalized SSA technique should enable adaptation with respect to arbitrary estimate quality, i. e. both a point estimate or a filtering pdf estimate. The intention is to adapt the number of samples while keeping bounded a difference between the approximate estimate provided by the PF and the true value of the estimate.

The paper is organized as follows: State estimation by the PF and a brief survey of the SSA techniques are given in Section 2. Then, the proposed SSA technique is presented in Section 3, which consists of the basic idea of the technique, a demonstration of the adaptation with respect to the filtering pdf quality, a discussion of a multidimensional case and computational issues of the technique. Further, an application of the proposed SSA technique is illustrated in a numerical example in Section 4 and finally, Section 5 concludes the paper.

2. STATE ESTIMATION BY THE PARTICLE FILTER

Consider a discrete-time nonlinear stochastic system given by the state equation (1) and measurement equation (2):

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k, \mathbf{e}_k), \quad k = 0, 1, 2, \dots, \quad (1)$$

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{v}_k), \quad k = 0, 1, 2, \dots, \quad (2)$$

where the vectors $\mathbf{x}_k \in \mathbb{R}^{n_x}$ and $\mathbf{z}_k \in \mathbb{R}^{n_z}$ represent a state of the system and a measurement at time k , respectively, $\mathbf{e}_k \in \mathbb{R}^{n_x}$ and $\mathbf{v}_k \in \mathbb{R}^{n_z}$ are state and measurement white noises, mutually independent and independent of the initial state \mathbf{x}_0 , with known pdf's $p(\mathbf{e}_k)$ and $p(\mathbf{v}_k)$, respectively, $\mathbf{f}_k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x}$, $\mathbf{h}_k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_z} \rightarrow \mathbb{R}^{n_z}$ are known vector mappings and the pdf $p(\mathbf{x}_0)$ of the initial state \mathbf{x}_0 is known. The system given by (1) and (2) can alternatively be described by the transition pdf $p(\mathbf{x}_k | \mathbf{x}_{k-1})$ and measurement pdf $p(\mathbf{z}_k | \mathbf{x}_k)$.

The general solution to the state estimation problem in the form of the filtering pdf $p(\mathbf{x}_k | \mathbf{z}^k)$ with $\mathbf{z}^k \triangleq [\mathbf{z}_0^T, \dots, \mathbf{z}_k^T]^T$ is provided by the BRR's:

$$p(\mathbf{x}_k | \mathbf{z}^k) = \frac{p(\mathbf{x}_k | \mathbf{z}^{k-1})p(\mathbf{z}_k | \mathbf{x}_k)}{\int p(\mathbf{x}_k | \mathbf{z}^{k-1})p(\mathbf{z}_k | \mathbf{x}_k) d\mathbf{x}_k}, \quad (3)$$

$$p(\mathbf{x}_k | \mathbf{z}^{k-1}) = \int p(\mathbf{x}_k | \mathbf{x}_{k-1})p(\mathbf{x}_{k-1} | \mathbf{z}^{k-1}) d\mathbf{x}_{k-1}, \quad (4)$$

which produce the filtering pdf $p(\mathbf{x}_k | \mathbf{z}^k)$ and the predictive pdf $p(\mathbf{x}_k | \mathbf{z}^{k-1})$.

The idea of the PF in nonlinear state estimation is to approximate the true filtering pdf $p(\mathbf{x}_k | \mathbf{z}^k)$ by the empirical filtering pdf $r_{N_k}(\mathbf{x}_k | \mathbf{z}^k)$, which is given by

N_k random samples of the state $\{\mathbf{x}_k^{(i)}\}_{i=1}^{N_k}$ and associated weights $\{\mathbf{w}_k(\mathbf{x}_k^{(i)})\}_{i=1}^{N_k}$, $\mathbf{w}_k(\cdot) \geq 0$, $\sum_{i=1}^{N_k} \mathbf{w}_k(\mathbf{x}_k^{(i)}) = 1$ as

$$r_{N_k}(\mathbf{x}_k|\mathbf{z}^k) = \sum_{i=1}^{N_k} \mathbf{w}_k(\mathbf{x}_k^{(i)}) \delta(\mathbf{x}_k - \mathbf{x}_k^{(i)}),$$

where $\delta(\cdot)$ is the Dirac function defined as $\delta(\mathbf{x}) = 0$ for $\mathbf{x} \neq 0$ and $\int \delta(\mathbf{x}) d\mathbf{x} = 1$.

The samples are drawn from a sampling pdf $\pi(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k)$, which is chosen by the user, and the weights are computed to balance the discrepancy between the sampling pdf $\pi(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k)$ and the filtering pdf $p(\mathbf{x}_k|\mathbf{z}^k)$, which is known up to a normalizing constant only. The general algorithm of the PF can be summarized in Alg. 1 as follows:

Alg. 1: particle filter

Initialization:

- Set $k = 0$.
- Draw N_0 samples $\{\mathbf{x}_0^{(i)}\}_{i=1}^{N_0}$ from the prior pdf $p(\mathbf{x}_0|\mathbf{z}^{-1}) = p(\mathbf{x}_0)$.
- Compute the unnormalized weights $\{\tilde{\mathbf{w}}_0(\mathbf{x}_0^{(i)})\}_{i=1}^{N_0}$ as

$$\tilde{\mathbf{w}}_0(\mathbf{x}_0^{(i)}) = p(\mathbf{z}_0|\mathbf{x}_0^{(i)}). \quad (5)$$

Resampling: Generate a new set $\{\mathbf{x}_k^{*(i)}\}_{i=1}^{N_{k+1}}$ by resampling with replacement N_{k+1} times from $\{\mathbf{x}_k^{(i)}\}_{i=1}^{N_{k+1}}$ with probability $Prob(\mathbf{x}_k^{*(i)} = \mathbf{x}_k^{(i)}) = \mathbf{w}_k(\mathbf{x}_k^{(i)})$ and set $\mathbf{w}_k(\mathbf{x}_k^{*(i)}) = \frac{1}{N_{k+1}}$. Replace the original sets $\{\mathbf{x}_k^{(i)}\}_{i=1}^{N_{k+1}}$ and $\{w_k(\mathbf{x}_k^{(i)})\}_{i=1}^{N_{k+1}}$ by the resampled sets $\{\mathbf{x}_k^{*(i)}\}_{i=1}^{N_{k+1}}$ and $\{w_k(\mathbf{x}_k^{*(i)})\}_{i=1}^{N_{k+1}}$ respectively.

Sampling:

- Draw N_k samples $\{\mathbf{x}_k^{(i)}\}_{i=1}^{N_k}$ from the sampling pdf $\pi(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k)$.

Weighting:

- The weights $\{\tilde{\mathbf{w}}_k(\mathbf{x}_k^{(i)})\}_{i=1}^{N_k}$ are calculated using the following relation

$$\tilde{\mathbf{w}}_k(\mathbf{x}_k^{(i)}) = \frac{p(\mathbf{z}_k|\mathbf{x}_k^{(i)})p(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{*(i)})}{\pi(\mathbf{x}_k^{(i)}|\mathbf{x}_{k-1}^{*(i)}, \mathbf{z}_k)}. \quad (6)$$

The weights are normalized, i. e. $\mathbf{w}_k(\mathbf{x}_k^{(i)}) = \tilde{\mathbf{w}}_k(\mathbf{x}_k^{(i)}) / \sum_{j=1}^{N_k} \tilde{\mathbf{w}}_k(\mathbf{x}_k^{(j)})$. The empirical pdf $r_{N_k}(\mathbf{x}_k|\mathbf{z}^k)$ is given by the samples $\{\mathbf{x}_k^{(i)}\}_{i=1}^{N_k}$ and the weights $\{\mathbf{w}_k(\mathbf{x}_k^{(i)})\}_{i=1}^{N_k}$ as

$$r_{N_k}(\mathbf{x}_k|\mathbf{z}^k) = \sum_{i=1}^{N_k} \mathbf{w}_k(\mathbf{x}_k^{(i)}) \delta(\mathbf{x}_k - \mathbf{x}_k^{(i)}).$$

Increase k and iterate to step **Resampling**.

For details see [2].

Note that the algorithm uses a general sampling pdf $\pi(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k)$ based on utilization of the current measurement \mathbf{z}_k . This general sampling pdf covers either the prior sampling pdf with $\pi(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k) = p(\mathbf{x}_k|\mathbf{x}_{k-1})$, or other sampling pdf's, e. g. the optimal sampling pdf or the auxiliary sampling pdf [3].

2.1. Sample size specification

Besides the sampling pdf $\pi(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k)$, also the sample size N_k represents a key parameter of the PF significantly affecting estimate quality. The first papers dealing with the PF usually got along with the rule the more particles the better estimate quality. While this is true as for $N_k \rightarrow \infty$ the empirical pdf $r_{N_k}(\mathbf{x}_k|\mathbf{z}^k)$ converges to the true pdf $p(\mathbf{x}_k|\mathbf{z}^k)$, the limited computational resources lead to an endeavor to keep computational demands of the PF algorithm as low as possible. To save the computational resources, it is possible to fix the estimate quality and change the sample size accordingly as at some time instants the sample size is unnecessarily high to achieve the desired estimate quality. Or the task may request keeping estimate quality at a certain level without regard to computational costs. These are the basic motivating ideas for SSA techniques.

Currently, only few SSA techniques have been designed; their survey can be found in [12]. Basically, there are two groups of SSA techniques. The former one adapts the sample size by means of *estimate quality* evaluation [7, 8, 9, 10], the latter by means of *sample set quality* evaluation [6, 13, 14].

As the paper focuses on adaptation with respect to estimate quality, only representatives of the former group are briefly introduced.

2.1.1. Kullback–Leibler divergence sampling

In [7] a SSA technique called Kullback–Leibler divergence (KLD) sampling was proposed. It adapts the sample size to bound the error between the true pdf and the empirical pdf by ε with probability $1 - \delta$. The error is measured by the KLD. The technique assumes that the true pdf can be represented by a discrete piecewise constant pdf. The drawback of the proposed technique is that the samples of the empirical pdf are assumed to be drawn directly from the true pdf and the relation for sample size calculation utilizes the number of bins with support n as the only information concerning the true pdf. The relation for computing N_k is $N_k = \frac{1}{2\varepsilon} \chi_{n-1, 1-\delta}^2$, where $\chi_{n-1, 1-\delta}^2$ is $1 - \delta$ quantile of the chi-square distribution with $n - 1$ degrees of freedom. The KLD sampling technique can be referred to as an adaptation with respect to complexity of the true pdf, without taking into account the sampling pdf.

2.1.2. KLD sampling improvement

In [8] the KLD sampling technique was elaborated further to grasp the fact that the samples of the empirical pdf are drawn from a sampling pdf which is different from the true pdf. To take this fact into account, the relative accuracy of the estimator

between sampling from the true pdf and the sampling pdf was utilized. The sample size is in [8] adapted according to $N_k = \frac{\text{var}_\pi(\frac{p(\mathbf{x}_k|\mathbf{z}^k)}{\pi(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{z}_k)}\mathbf{x}_k)}{\text{var}_p(\mathbf{x}_k|\mathbf{z}^k)} \frac{1}{2\varepsilon} \chi_{n-1,1-\delta}^2$.

2.1.3. Asymptotic normal approximation

In the same paper [8] another adaptation technique has been proposed checking quality of the estimate not in terms of the true pdf $p(\mathbf{x}_k|\mathbf{z}^k)$ estimate but in terms of a moment of this pdf. The idea of this technique is such that the central limit theorem (CLT) justifies asymptotic normal approximation for the mean of the estimate, i.e., $E_{r_{N_k}}(\mathbf{x}_k|\mathbf{z}^k) \sim \mathcal{N}\{\mathbf{x}_k; E_p(\mathbf{x}_k|\mathbf{z}^k), \text{var}_\pi(\frac{p(\mathbf{x}_k|\mathbf{z}^k)}{\pi(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{z}_k)}\mathbf{x}_k|\mathbf{z}^k)/N_k\}$, with $\mathcal{N}\{\mathbf{x}; \mu, \sigma^2\}$ being normal distribution of \mathbf{x} with mean μ and variance σ^2 . Based on this approximation, the paper proposes one-sided confidence interval for the number of particles to bound the relative error of the mean estimate $Pr ob(|\frac{E_{r_{N_k}}(\mathbf{x}_k|\mathbf{z}^k) - E_p(\mathbf{x}_k|\mathbf{z}^k)}{E_p(\mathbf{x}_k|\mathbf{z}^k)}| \leq \varepsilon) \geq (1 - \delta)$. Then the relation for sample size N_k proposed in the paper is $N_k = \frac{t_{1-\delta/2}^2 \text{var}_\pi(\frac{p(\mathbf{x}_k|\mathbf{z}^k)}{\pi(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{z}_k)}\mathbf{x}_k|\mathbf{z}^k)}{\varepsilon^2 E_p(\mathbf{x}_k|\mathbf{z}^k)^2}$, where $t_{1-\delta/2}$ is $1 - \delta/2$ quantile of the standard normal distribution.

2.1.4. Fixed efficient sample size

The idea proposed in [9] is based on keeping efficient sample size (ESS) fixed and adapting the sample size accordingly. The ESS represents the number of particles drawn from the true pdf $p(\mathbf{x}_k|\mathbf{z}^k)$ necessary to attain the same estimate quality as N_k particles drawn from the sampling pdf $\pi(\mathbf{x}_k)$. The technique adapts the sample size N_k as $N_k = N_k^{\text{ESS}}(1 + \text{var}_\pi(\tilde{\mathbf{w}}_k(\mathbf{x}_k)))$, where N_k^{ESS} is a design parameter representing the fixed efficient sample size.

2.1.5. Fixed empirical density quality

The technique proposed in this paper is a generalization of the SSA technique developed in [11] which was originally focused on keeping fixed quality of the filtering pdf estimate. The quality of the estimate was measured by the inaccuracy [15] between the estimate and the true pdf. The relation for sample size adaptation is presented in Section 3.1.

The above mentioned SSA techniques focus on either a point estimate quality or pdf estimate quality. To design a single technique that enables sample size adaptation with respect to both a point estimate and a pdf estimate is the fundamental motivation of this paper.

3. SAMPLE SIZE ADAPTATION

Frequently, the filtering pdf $p(\mathbf{x}_k|\mathbf{z}^k)$ obtained from a filter is utilized to compute the integral

$$\mathbf{I}(\mathbf{g}) = \int \mathbf{g}(\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{z}^k)d\mathbf{x}_k, \tag{7}$$

where \mathbf{g} is a mapping $\mathbf{g} : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^m$. An important special case of (7) is the filtering mean $\hat{\mathbf{x}}_{k|k}$ given by $\mathbf{g}(\mathbf{x}_k) = \mathbf{x}_k$. An approximation of (7) provided by the PF is given by

$$\begin{aligned} \hat{\mathbf{I}}(\mathbf{g}) &= \int \mathbf{g}(\mathbf{x}_k) r_{N_k}(\mathbf{x}_k | \mathbf{z}^k) d\mathbf{x}_k = \int \mathbf{g}(\mathbf{x}_k) \frac{\frac{1}{N_k} \sum_{i=1}^{N_k} \tilde{\mathbf{w}}(\mathbf{x}_k^{(i)}) \delta(\mathbf{x}_k - \mathbf{x}_k^{(i)})}{\frac{1}{N_k} \sum_{j=1}^{N_k} \tilde{\mathbf{w}}(\mathbf{x}_k^{(j)})} d\mathbf{x}_k \\ &= \frac{\frac{1}{N_k} \sum_{i=1}^{N_k} \tilde{\mathbf{w}}(\mathbf{x}_k^{(i)}) \mathbf{g}(\mathbf{x}_k^{(i)})}{\frac{1}{N_k} \sum_{j=1}^{N_k} \tilde{\mathbf{w}}(\mathbf{x}_k^{(j)})}, \end{aligned} \tag{8}$$

where it can be proved [2] that

$$\lim_{N_k \rightarrow \infty} \hat{\mathbf{I}}(\mathbf{g}) = \mathbf{I}(\mathbf{g}).$$

A reasonable requirement for quality of the approximation (8) is to keep the difference between $\hat{\mathbf{I}}(\mathbf{g})$ and $\mathbf{I}(\mathbf{g})$ bounded. The difference is solely affected by the sampling pdf and the sample size. As the sampling pdf is supposed to be unaltered, keeping the difference bounded naturally causes a variation of the sample size N_k . This is the main idea of the SSA technique proposed in this paper. For simplicity, the following relations will consider a scalar case, i.e. $g : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^1$ and the result will consequently be generalized to a multidimensional case.

The difference between $\hat{I}(g)$ and $I(g)$ is given by

$$\hat{I}(g) - I(g) = \frac{\frac{1}{N_k} \sum_{i=1}^{N_k} \tilde{\mathbf{w}}_k(\mathbf{x}_k^{(i)}) g(\mathbf{x}_k^{(i)})}{\frac{1}{N_k} \sum_{j=1}^{N_k} \tilde{\mathbf{w}}_k(\mathbf{x}_k^{(j)})} - I(g) = \frac{\frac{1}{N_k} \sum_{i=1}^{N_k} \tilde{\mathbf{w}}_k(\mathbf{x}_k^{(i)}) (g(\mathbf{x}_k^{(i)}) - I(g))}{\frac{1}{N_k} \sum_{j=1}^{N_k} \tilde{\mathbf{w}}_k(\mathbf{x}_k^{(j)})} \tag{9}$$

$$= \frac{\frac{1}{N_k} \sum_{i=1}^{N_k} \tilde{\mathbf{w}}_k(\mathbf{x}_k^{(i)}) \gamma(\mathbf{x}_k^{(i)})}{\frac{1}{N_k} \sum_{j=1}^{N_k} \tilde{\mathbf{w}}_k(\mathbf{x}_k^{(j)})} = \frac{\bar{Y}}{\bar{W}} \triangleq R, \tag{10}$$

where $\gamma(\mathbf{x}_k^{(i)}) = g(\mathbf{x}_k^{(i)}) - I(g)$. Let us denote the ratio in (10) by R . Both the numerator and denominator in (10) are given by sample means and according to the central limit theorem (CLT), for $N_k \rightarrow \infty$ they converge to a normal distribution. Note that the CLT can only be applied if means and variances of the terms in the sums are finite. Denote the sample mean in the numerator as $\bar{Y} = \frac{1}{N_k} \sum_{i=1}^{N_k} [\tilde{\mathbf{w}}_k(\mathbf{x}_k^{(i)}) \gamma(\mathbf{x}_k^{(i)})]$ and the sample mean in the denominator as $\bar{W} = \frac{1}{N_k} \sum_{i=1}^{N_k} \tilde{\mathbf{w}}_k(\mathbf{x}_k^{(i)})$, i.e. $Y = \tilde{\mathbf{w}}_k(\mathbf{x}_k) \gamma(\mathbf{x}_k)$ and $W = \tilde{\mathbf{w}}_k(\mathbf{x}_k)$. Then, according to the CLT

$$\begin{aligned} p(\bar{Y}) &\xrightarrow[N_k \rightarrow \infty]{} \mathcal{N}\{\bar{Y} : \mu_{\bar{Y}}, \sigma_{\bar{Y}}^2\}, \\ p(\bar{W}) &\xrightarrow[N_k \rightarrow \infty]{} \mathcal{N}\{\bar{W} : \mu_{\bar{W}}, \sigma_{\bar{W}}^2\}, \end{aligned}$$

where $\mu_{\bar{Y}} = \mu_Y$, $\mu_{\bar{W}} = \mu_W$ $\sigma_{\bar{Y}}^2 = \frac{\sigma_Y^2}{N_k}$ and $\sigma_{\bar{W}}^2 = \frac{\sigma_W^2}{N_k}$ with

$$\mu_Y = 0, \tag{11}$$

$$\mu_W = E_{\pi}(\tilde{\mathbf{w}}_k(\mathbf{x}_k)), \tag{12}$$

$$\begin{aligned} \sigma_Y^2 = & E_{\pi}([\tilde{\mathbf{w}}_k(\mathbf{x}_k)]^2 [g(\mathbf{x}_k)]^2) - 2E_{\pi}([\tilde{\mathbf{w}}_k(\mathbf{x}_k)]^2 g(\mathbf{x}_k)) \frac{E_{\pi}(\tilde{\mathbf{w}}_k(\mathbf{x}_k) g(\mathbf{x}_k))}{E_{\pi}(\tilde{\mathbf{w}}_k(\mathbf{x}_k))} \\ & + E_{\pi}([\tilde{\mathbf{w}}_k(\mathbf{x}_k)]^2) \left[\frac{E_{\pi}(\tilde{\mathbf{w}}_k(\mathbf{x}_k) g(\mathbf{x}_k))}{E_{\pi}(\tilde{\mathbf{w}}_k(\mathbf{x}_k))} \right]^2, \end{aligned} \tag{13}$$

$$\sigma_W^2 = E_{\pi}([\tilde{\mathbf{w}}_k(\mathbf{x}_k)]^2) - E_{\pi}^2(\tilde{\mathbf{w}}_k(\mathbf{x}_k)). \tag{14}$$

As the ratio R is a stochastic variable, its quantile must be used to bound it. A quantile of the ratio R in (10) as a function of N_k can not be computed directly due to intricate distribution of R , nevertheless the Geary-Hinkley transformation to normality [16] can be applied here. It says that, under a certain condition, the random variable R given by a ratio of two possibly correlated, normal random variables \bar{Y} and \bar{W} may be transformed to a standard normal variable T using the transformation

$$T = \frac{\mu_{\bar{W}}R - \mu_{\bar{Y}}}{\sqrt{\sigma_{\bar{W}}^2 R^2 - 2\text{cov}(\bar{Y}, \bar{W})R + \sigma_{\bar{Y}}^2}} = \frac{\mu_W R - \mu_Y}{\sqrt{\frac{\sigma_W^2}{N_k} R^2 - 2\frac{\text{cov}(Y, W)}{N_k} R + \frac{\sigma_Y^2}{N_k}}}. \tag{15}$$

The covariance $\text{cov}(\bar{Y}, \bar{W}) = \frac{\text{cov}(Y, W)}{N_k}$ in (15) can be expressed as

$$\text{cov}(Y, W) = E_{\pi}([\tilde{\mathbf{w}}_k(\mathbf{x}_k)]^2 g(\mathbf{x}_k)) - E_{\pi}([\tilde{\mathbf{w}}_k(\mathbf{x}_k)]^2) \frac{E_{\pi}([\tilde{\mathbf{w}}_k(\mathbf{x}_k)] g(\mathbf{x}_k))}{E_{\pi}([\tilde{\mathbf{w}}_k(\mathbf{x}_k)])}. \tag{16}$$

The transformation (15) holds even for quantiles, i. e.

$$t_{1-\delta/2} = \frac{\mu_W r_{1-\delta/2} - \mu_Y}{\sqrt{\frac{\sigma_W^2}{N_k} r_{1-\delta/2}^2 - 2\frac{\text{cov}(Y, W)}{N_k} r_{1-\delta/2} + \frac{\sigma_Y^2}{N_k}}}, \tag{17}$$

where $t_{1-\delta/2}$ is $1-\delta/2$ quantile of the standard normal distribution and $r_{1-\delta/2}$ is $1-\delta/2$ quantile of the distribution of R . Equation (17) represents a relation between a quantile of the standard normal distribution, a quantile of $R = \hat{I}(g) - I(g)$ and sample size N_k . Therefore, it is possible to introduce the following relation for sample size N_k :

$$N_k = \left\lceil t_{1-\delta/2}^2 \frac{\sigma_W^2 r_{1-\delta/2}^2 - 2\text{cov}(Y, W) r_{1-\delta/2} + \sigma_Y^2}{(\mu_W r_{1-\delta/2} - \mu_Y)^2} \right\rceil. \tag{18}$$

The sample size N_k given by (18) is necessary for the difference $\hat{I}(g) - I(g)$ to be within the interval $(-r_{1-\delta/2}, +r_{1-\delta/2})$ with probability $1-\delta$.

The relation (18) is the central point of the proposed SSA technique. The confidence coefficient $1-\delta$ and the value of $1-\delta/2$ quantile $r_{1-\delta/2}$ are both parameters of the technique and are chosen by the user.

Note that according to [16] the transformation (15) holds at the five percent significance level as long as the coefficient of variation of the denominator \bar{W} , denoted $C_{\bar{W}} = \sigma_{\bar{W}}/\mu_{\bar{W}}$ is less than 0.39 and coefficient of variation of the numerator \bar{Y} is greater than 0.005. If this condition does not apply, the quantile can be computed numerically but not in terms of a function of the sample size N_k . In such a case, it is possible to determine an upper bound for the sample size N_k using Chebychev's inequality

$$Prob(|\hat{I}(g) - I(g)| \geq \varepsilon) \leq \sqrt{\text{var}(\hat{I}(g) - I(g))}/\varepsilon,$$

in the following form

$$N_k = \left\lceil \frac{1}{\varepsilon^2 \delta} \text{var}(\hat{I}(g) - I(g)) \right\rceil. \tag{19}$$

The relation states that if sample size N_k given by (19) is used, then

$$Prob(|\hat{I}(g) - I(g)| \geq \varepsilon) \leq \delta.$$

The term $\text{var}(\hat{I}(g) - I(g))$ in (19) can be computed using the standard delta method for ratio statistics as

$$\begin{aligned} \text{var}_\pi(\hat{I}(g) - I(g)) &\approx \left(\frac{\text{var}_\pi(\tilde{\mathbf{w}}_k(\mathbf{x}_k)g(\mathbf{x}_k))}{\mathbb{E}_\pi^2(\tilde{\mathbf{w}}_k(\mathbf{x}_k))} \right. \\ &\quad \left. - 2 \frac{\mathbb{E}_\pi(\tilde{\mathbf{w}}_k(\mathbf{x}_k)g(\mathbf{x}_k)) \text{cov}_\pi(\tilde{\mathbf{w}}_k(\mathbf{x}_k), \tilde{\mathbf{w}}_k(\mathbf{x}_k)g(\mathbf{x}_k))}{\mathbb{E}_\pi^3(\tilde{\mathbf{w}}_k(\mathbf{x}_k))} + \frac{\mathbb{E}_\pi^2(\tilde{\mathbf{w}}_k(\mathbf{x}_k)g(\mathbf{x}_k)) \text{var}_\pi(\tilde{\mathbf{w}}_k(\mathbf{x}_k))}{\mathbb{E}_\pi^4(\tilde{\mathbf{w}}_k(\mathbf{x}_k))} \right) \end{aligned} \tag{20}$$

which can be further explored using second order raw moments as

$$\begin{aligned} \text{var}_\pi(\hat{I}(g) - I(g)) &\approx \left(\frac{\mathbb{E}_\pi([\tilde{\mathbf{w}}_k(\mathbf{x}_k)]^2[g(\mathbf{x}_k)]^2)}{\mathbb{E}_\pi^2(\tilde{\mathbf{w}}_k(\mathbf{x}_k))} \right. \\ &\quad \left. - 2 \frac{\mathbb{E}_\pi(\tilde{\mathbf{w}}_k(\mathbf{x}_k)g(\mathbf{x}_k)) \mathbb{E}_\pi([\tilde{\mathbf{w}}_k(\mathbf{x}_k)]^2g(\mathbf{x}_k))}{\mathbb{E}_\pi^3(\tilde{\mathbf{w}}_k(\mathbf{x}_k))} + \frac{\mathbb{E}_\pi^2(\tilde{\mathbf{w}}_k(\mathbf{x}_k)g(\mathbf{x}_k)) \mathbb{E}_\pi(\tilde{\mathbf{w}}_k(\mathbf{x}_k))}{\mathbb{E}_\pi^4(\tilde{\mathbf{w}}_k(\mathbf{x}_k))} \right) \\ &= \frac{\sigma_Y^2}{\mathbb{E}_\pi^2(\tilde{\mathbf{w}}_k(\mathbf{x}_k))}. \end{aligned} \tag{21}$$

It must be noted that for the sample size given by (19) no information concerning distribution of $\hat{I}(g) - I(g)$ is used; therefore this condition for the sample size is loose and the relation (18) should be used if possible.

The asymptotic normal approximation proposed in [8] can be seen as a simplification of the relation (18) by considering the denominator W to be a constant (i. e. $\sigma_W^2 = 0$ and $\text{cov}(Y, W) = 0$), setting $\mathbf{g}(\mathbf{x}_k) = \mathbf{x}_k$ and bounding a relative difference $\frac{\hat{I}(\mathbf{g}) - I(\mathbf{g})}{I(\mathbf{g})}$ instead of the absolute error $\hat{I}(\mathbf{g}) - I(\mathbf{g})$ used in this paper.

3.1. Adaptation with respect to empirical pdf quality

The proposed relation (18) for sample size adaptation can also be used to keep quality of the empirical pdf fixed by a special choice of the function $g(\mathbf{x}_k)$ [11].

To measure the empirical filtering pdf $r_{N_k}(\mathbf{x}_k|\mathbf{z}^k)$ quality, which approximates the filtering pdf $p(\mathbf{x}_k|\mathbf{z}^k)$, the Kullback–Leibler (KL) distance given by

$$D(r_{N_k}, p) \triangleq \int r_{N_k}(\mathbf{x}_k|\mathbf{z}^k) \log \frac{r_{N_k}(\mathbf{x}_k|\mathbf{z}^k)}{p(\mathbf{x}_k|\mathbf{z}^k)} d\mathbf{x}_k \tag{22}$$

can be chosen. It is a generic choice of information measure to quantify a discrepancy between two pdf's $r_{N_k}(\mathbf{x}_k|\mathbf{z}^k)$ and $p(\mathbf{x}_k|\mathbf{z}^k)$. The KL distance can also be written as a difference of two components

$$D(r_{N_k}, p) = \underbrace{\int r_{N_k}(\mathbf{x}_k|\mathbf{z}^k) \log \frac{1}{p(\mathbf{x}_k|\mathbf{z}^k)} d\mathbf{x}_k}_{K(r_{N_k}, p)} - \underbrace{\int r_{N_k}(\mathbf{x}_k|\mathbf{z}^k) \log \frac{1}{r_{N_k}(\mathbf{x}_k|\mathbf{z}^k)} d\mathbf{x}_k}_{H(r_{N_k})} \tag{23}$$

where the former $K(r_{N_k}, p)$ is inaccuracy [15] and the latter $H(p)$ is the Shannon differential entropy (SDE). The inaccuracy measures actual discrepancy between the pdf's $r_{N_k}(\mathbf{x}_k|\mathbf{z}^k)$ and $p(\mathbf{x}_k|\mathbf{z}^k)$, while the SDE measures entropy of $r_{N_k}(\mathbf{x}_k|\mathbf{z}^k)$. The inaccuracy is of an opportune form for the comparison of $r_{N_k}(\mathbf{x}_k|\mathbf{z}^k)$ and $p(\mathbf{x}_k|\mathbf{z}^k)$ as the empirical pdf is a mixture of Dirac functions and the relation $\int \delta(x - a)f(x)dx = f(a)$ holds. The SDE component of the KL distance will be dropped as $H(r_{N_k}) = -\infty$. It must be noted that the inaccuracy alone may be negative nevertheless it still provides a measure of a disagreement between the pdf's.

The inaccuracy $K(r_{N_k}, p)$ given by

$$K(r_{N_k}, p) = \int \frac{\frac{1}{N_k} \sum_{i=1}^{N_k} \tilde{\mathbf{w}}_k(\mathbf{x}_k^{(i)}) \delta(\mathbf{x}_k - \mathbf{x}_k^{(i)})}{\frac{1}{N_k} \sum_{j=1}^{N_k} \tilde{\mathbf{w}}_k(\mathbf{x}_k^{(j)})} \log \frac{1}{p(\mathbf{x}_k|\mathbf{z}^k)} d\mathbf{x}_k \tag{24}$$

can be written as

$$K(r_{N_k}, p) = \frac{\frac{1}{N_k} \sum_{i=1}^{N_k} \tilde{\mathbf{w}}_k(\mathbf{x}_k^{(i)}) \log \frac{1}{p(\mathbf{x}_k^{(i)}|\mathbf{z}^k)}}{\frac{1}{N_k} \sum_{j=1}^{N_k} \tilde{\mathbf{w}}_k(\mathbf{x}_k^{(j)})}. \tag{25}$$

Therefore, the inaccuracy (25) can be seen as a special case of $\hat{\mathbf{I}}(\mathbf{g})$ in (8) with

$$g(\mathbf{x}_k) = \log \frac{1}{p(\mathbf{x}_k|\mathbf{z}^k)}. \tag{26}$$

Analogically to $\hat{\mathbf{I}}(\mathbf{g})$ being an approximation of $\mathbf{I}(\mathbf{g})$, the inaccuracy (25) is an approximation of $K(p, p)$ given by

$$K(p, p) = \int p(\mathbf{x}_k|\mathbf{z}^k) \log \frac{1}{p(\mathbf{x}_k|\mathbf{z}^k)} d\mathbf{x}_k,$$

which is equal to the SDE $H(p)$. Therefore the function γ in (10) is equal to

$$\gamma(\mathbf{x}_k) = \log \frac{1}{p(\mathbf{x}_k|\mathbf{z}^k)} - H(p).$$

So by considering $\hat{\mathbf{I}}(\mathbf{g}) = \mathbf{K}(r_{N_k}, p)$ and $\mathbf{I}(\mathbf{g}) = \mathbf{H}(p)$ the adaptation with respect to the empirical pdf quality can be seen as a special case of the SSA technique proposed above.

As the terms σ_Y^2 and $\text{cov}(Y, W)$ in (13) and (16) for $g(\mathbf{x}_k)$ given by (26) depend on the true filtering pdf $p(\mathbf{x}_k|\mathbf{z}^k)$ which is unknown, the following substitution

$$p(\mathbf{x}_k|\mathbf{z}^k) = \frac{\tilde{\mathbf{w}}_k(\mathbf{x}_k)\pi(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k)}{c}, \tag{27}$$

which arises due to the fact that the normalized weight $\mathbf{w}(\mathbf{x}_k)$ is given by

$$\mathbf{w}(\mathbf{x}_k) = \frac{\tilde{\mathbf{w}}_k(\mathbf{x}_k)}{c} = \frac{p(\mathbf{x}_k|\mathbf{z}^k)}{\pi(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k)}, \tag{28}$$

with $c = \mathbf{E}_\pi(\tilde{\mathbf{w}}_k(\mathbf{x}_k))$, will be used. After a few arrangements the terms can be expressed as

$$\begin{aligned} \sigma_Y^2 &= \mathbf{E}_\pi([\tilde{\mathbf{w}}_k(\mathbf{x}_k)]^2[\tilde{g}(\mathbf{x}_k)]^2) - 2\mathbf{E}_\pi([\mathbf{w}(\mathbf{x}_k)]^2\tilde{g}(\mathbf{x}_k)) \frac{\mathbf{E}_\pi(\tilde{\mathbf{w}}_k(\mathbf{x}_k)\tilde{g}(\mathbf{x}_k))}{\mathbf{E}_\pi(\tilde{\mathbf{w}}_k(\mathbf{x}_k))} \\ &\quad + \mathbf{E}_\pi([\tilde{\mathbf{w}}_k(\mathbf{x}_k)]^2) \frac{\mathbf{E}_\pi^2(\tilde{\mathbf{w}}_k(\mathbf{x}_k)\tilde{g}(\mathbf{x}_k))}{\mathbf{E}_\pi^2(\tilde{\mathbf{w}}_k(\mathbf{x}_k))}, \\ \text{cov}(Y, W) &= \log \mathbf{E}_\pi(\tilde{\mathbf{w}}_k(\mathbf{x}_k)) \mathbf{E}_\pi([\tilde{\mathbf{w}}_k(\mathbf{x}_k)]^2) + \mathbf{E}_\pi([\tilde{\mathbf{w}}_k(\mathbf{x}_k)]^2\tilde{g}(\mathbf{x}_k)) \\ &\quad - \mathbf{E}_\pi([\tilde{\mathbf{w}}_k(\mathbf{x}_k)]^2) \frac{\log \mathbf{E}_\pi(\tilde{\mathbf{w}}_k(\mathbf{x}_k)) \mathbf{E}_\pi(\tilde{\mathbf{w}}_k(\mathbf{x}_k)) + \mathbf{E}_\pi(\tilde{\mathbf{w}}_k(\mathbf{x}_k)\tilde{g}(\mathbf{x}_k))}{\mathbf{E}_\pi(\tilde{\mathbf{w}}_k(\mathbf{x}_k))} \end{aligned}$$

with $\tilde{g}(\mathbf{x}_k) = \log \frac{1}{\tilde{\mathbf{w}}_k(\mathbf{x}_k)\pi(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k)}$.

3.2. Multidimensional case

If the mapping \mathbf{g} is multidimensional, i. e. $m > 1$, the relation (18) must be modified accordingly. The relation is derived analogically to the scalar case. In the multidimensional case, the user parameter $\mathbf{r}_{1-\delta/2}$ is given by a vector of dimensions $m \times 1$, which means that estimate quality for each dimension of \mathbf{g} may be bounded by a different value.

Instead of obtaining a single N_k in the scalar case, a set of m sample sizes $\{N_k^i\}_{i=1}^m$ is obtained here using the following relation

$$N_k^i = \left\lceil \left[\left(\frac{[\mathbf{S} \mathbf{1}_{m \times 1} t_{1-\delta/2}]_i}{\mu_W [\mathbf{r}_{1-\delta/2}]_i - [\mu_Y]_i} \right)^2 \right] \right\rceil, \tag{29}$$

where $[\cdot]_i$ stands for the i th component of the vector, $\mathbf{1}_{m \times 1}$ is an m -dimensional column vector of ones and \mathbf{S} is the Cholesky decomposition of the $m \times m$ matrix \mathbf{P} given by

$$\mathbf{P} = \sigma_W^2 \mathbf{r}_{1-\delta/2} \mathbf{r}_{1-\delta/2}^T - \text{cov}(Y, W) \mathbf{r}_{1-\delta/2}^T - \mathbf{r}_{1-\delta/2} \text{cov}(Y, W)^T + \text{var}(Y) \tag{30}$$

such that $\mathbf{S}\mathbf{S}^T = \mathbf{P}$. The matrix \mathbf{S} is an analogy to the denominator in the Geary-Hinkley transformation (15), where the matrix \mathbf{P} is a multidimensional counterpart to the term under the square root in (17).

The sample size used in the PF must be chosen as a maximum of the set $\{N_k^i\}_{i=1}^m$, i. e.

$$N_k = \max_i \{N_k^i\}_{i=1}^m, \quad (31)$$

because the condition for estimate quality must hold for each element of the mapping \mathbf{g} . This implies that for some elements of \mathbf{g} the sample size N_k may be unnecessarily high but such a sample size is required for $Prob\left(|\hat{I}(\mathbf{g}) - I(\mathbf{g})| \leq \mathbf{r}_{1-\delta/2}\right) = \delta$.

3.3. Computational issues

To implement the proposed SSA technique in the PF framework, several issues must be discussed. First of all, let us restate individual terms used in the sample size adaptation (18):

$$E_\pi(\tilde{\mathbf{w}}_k(\mathbf{x}_k)) = \int \tilde{\mathbf{w}}_k(\mathbf{x}_k)\pi(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k)d\mathbf{x}_k \quad (32)$$

$$E_\pi([\tilde{\mathbf{w}}_k(\mathbf{x}_k)]^2) = \int \tilde{\mathbf{w}}_k(\mathbf{x}_k)^2\pi(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k)d\mathbf{x}_k \quad (33)$$

$$E_\pi(\tilde{\mathbf{w}}_k(\mathbf{x}_k)\mathbf{g}(\mathbf{x}_k)) = \int \tilde{\mathbf{w}}_k(\mathbf{x}_k)\mathbf{g}(\mathbf{x}_k)\pi(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k)d\mathbf{x}_k \quad (34)$$

$$E_\pi([\tilde{\mathbf{w}}_k(\mathbf{x}_k)]^2\mathbf{g}(\mathbf{x}_k)) = \int \tilde{\mathbf{w}}_k(\mathbf{x}_k)^2\mathbf{g}(\mathbf{x}_k)\pi(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k)d\mathbf{x}_k \quad (35)$$

$$E_\pi([\tilde{\mathbf{w}}_k(\mathbf{x}_k)]^2[\mathbf{g}(\mathbf{x}_k)]^2) = \int \tilde{\mathbf{w}}_k(\mathbf{x}_k)^2[\mathbf{g}(\mathbf{x}_k)]^2\pi(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k)d\mathbf{x}_k. \quad (36)$$

Value of the integrals (32–36) cannot usually be computed analytically, hence they must be calculated either numerically or using MC integration. In the case of the numerical integration, calculation of $\pi(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k)$ at an arbitrary point for large N_{k-1} can be computationally demanding as N_{k-1} evaluations of local sampling pdf's $\pi(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)}, \mathbf{z}_k)$ is required for each point. A simple solution to this problem is approximation of the sampling pdf $\pi(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k)$ by a piecewise linear function which evaluation at an arbitrary point is modest from the computational point of view.

MC integration of the expectations (32–36) is preferable especially for high dimension of the state \mathbf{x}_k . To calculate the expectations using MC integration, first, generate N_{MC} samples from $\pi(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k)$, second, utilize them to compute N_k according to (18) and finally, the remaining $N_k - N_{MC}$ samples are drawn.

Note that at some cases, which include adaptation with respect to a point estimate, it is also possible to compute the values of σ_W^2 , σ_Y^2 , $\text{cov}(Y, W)$ and μ_W directly from the initial N_{MC} samples and corresponding weights.

The accuracy of the sample size N_k computation based on the above mentioned approach is affected by N_{MC} , i. e. another sample size. To avoid this *deja vu*, it is possible to use a strategy consisting in successive increasing the precision of

the sample size N_k calculation by utilizing the newly generated samples. More specifically, instead of drawing all the remaining $N_k - N_{MC}$ in one batch, only ΔN samples are drawn and used for N_k recalculation. This process is repeated until N_k equals to the number of already drawn samples. This technique attenuates the dependency of the sample size N_k on the sample size N_{MC} .

The PF with the proposed SSA technique will be denoted as the adaptive PF (APF) and its algorithm can be summarized as

Alg. 2: adaptive particle filter

Specification of parameters: Choose the confidence coefficient $1 - \delta$ and length of the interval $r_{1-\delta/2}$.

Initial sample size adaptation: Compute value of the integrals (32–36). and calculate the sample size N_0 according to (18).

Initialization Corresponds to Alg. 1

Sample size adaptation: Compute value of the integrals (32–36). and calculate the sample size N_k according to (18).

Resampling: Corresponds to Alg. 1

Sampling: Corresponds to Alg. 1

Weighting: Corresponds to Alg. 1

Increase k and iterate to step **Sample size adaptation**.

4. NUMERICAL EXAMPLE

To illustrate the proposed SSA technique, a scalar nonlinear non-Gaussian system [17] is considered:

$$\begin{aligned}x_{k+1} &= \varphi_1 x_k + 1 + \sin(\omega \pi k) + e_k \\z_k &= \varphi_2 x_k^2 + v_k\end{aligned}$$

with $p(x_0) = \mathcal{N}\{x_0; 0, 12\}$, $p(e_k) = \mathcal{G}\{e_k; 3, 2\}$, $p(v_k) = \mathcal{N}\{v_k; 0, 1\}$, $\varphi_1 = 0.5$, $\varphi_2 = 0.2$, $\omega = 0.04$, where $\mathcal{G}\{x; a, b\}$ means Gamma distribution with parameters a and b . The state is estimated by the PF for $k = 0, \dots, 29$. The PF considers $p(x_0|z^{-1}) = p(x_0)$ and prior sampling density. In all experiments 1000 MC simulations were performed.

The sample size was adapted first with respect to a point estimate quality, more specifically the filtering mean, and second with respect to empirical filtering pdf quality. The true values (i.e. the filtering mean $\hat{x}_{k|k} = \mathbf{E}(x_k|z^k)$ and the true filtering pdf $p(\mathbf{x}_k|z^k)$) used for comparison of the results were calculated using the point-mass method [18], which computes the BRP's numerically, with a large number of grid points.

For adaptation with respect to the filtering mean the following parameters were chosen $g(x_k) = x_k$, $\delta = 0.1$ and $r_{1-\delta/2} = 0.1$. The results, i. e. the 90 % quantile of the difference $\hat{I}(x_k) - I(x_k)$ calculated using 1000 MC simulations is depicted in Figure 1a and an example of APF sample size evolution for several simulations (dashed), together with an shaded area formed by 95 % sample sizes is given in Figure 1b.

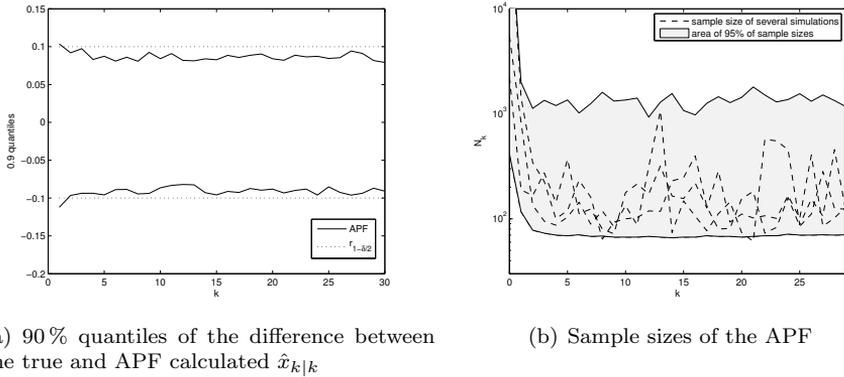


Fig. 1. SSA with respect to the filtering mean quality.

From Figures 1a and 1b it can be seen that the APF guarantees the given level of estimate quality at the cost of dramatic increase of computational costs during the first few steps.

The APF with adaptation with respect to empirical filtering pdf quality considered the confidence coefficient $\delta = 0.01$ and $r_{1-\delta/2} = 1$ in this case. The results obtained by APF were compared to that of the unadapted PF. To compare the APF and an unadapted PF meaningfully, the sample size for the unadapted PF was calculated in the following way. Firstly, an average N_{AV} of all the sample sizes of the APF was calculated and rounded, i. e.

$$N_{AV} = \left\lceil \frac{1}{30} \frac{1}{1000} \sum_{k=0}^{29} \sum_{s=1}^{1000} N_k(s) \right\rceil = 410,$$

where $N_k(s)$ is sample size of the APF in s th simulation at the time instant k . Consequently, the unadapted PF was applied twice with $N = N_{AV}$, and $N = 2 \cdot N_{AV}$. It should be remembered that the unadapted PF uses some information obtained from the APF (N_{AV}) but does not adapt the sample size N in time.

Figure 2c contains 99 % quantiles of the difference $K(r_{N_k}) - H(p)$ calculated from the simulations for the APF (solid) and unadapted PF with $N = N_{AV}$ (dot-dashed), and $N = 2 \cdot N_{AV}$ (dashed). An example of APF sample size evolution for several simulations is given in Figure 2d (dashed), together with an shaded area formed by 95 % sample sizes.

From Figure 2c it is clear that the APF adapts the sample size to keep the difference $K(r_{N_k}) - H(p)$ within the interval $(-r_{0.99}, r_{0.99})$, $r_{0.99} = 1$ with probability

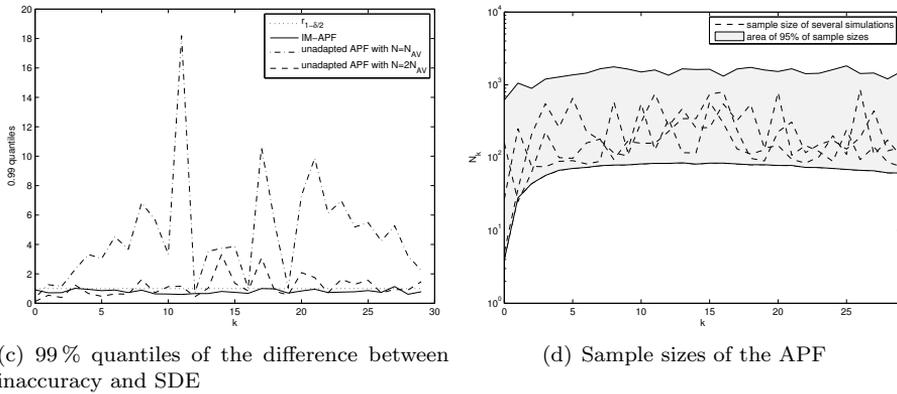


Fig. 2. SSA with respect to the empirical filtering pdf quality.

0.99. The unadapted PF with the same sample size on average as the APF provides much worse results in terms of empirical filtering pdf quality than the APF. If the sample size of the unadapted PF is considered twice as large as the average N_{AV} , then the empirical pdf quality is approximately the same as with the APF. Therefore, in this case the unadapted PF requires more than twice as many samples as the APF to guarantee the same quality of the empirical filtering pdf.

5. CONCLUSION

The paper dealt with the particle filter for nonlinear state estimation problem. A technique adapting sample size of the particle filter has been proposed. The adaptation enables guaranteeing a quality of the approximate estimate provided by the particle filter. As the estimate, either a point estimate, such as filtering mean, or a filtering pdf estimate may be chosen. The proposed relation for sample size guarantees the difference between the approximate estimate and the true value of the estimate to be within a user-specified boundary with a user-specified probability. The paper also discussed implementation issues of the technique and showed a similarity between the proposed technique and the asymptotic normal approximation technique. The adaptive particle filter was illustrated in a numerical example where it was shown that the adaptation guarantees the pre-specified quality of the chosen estimate and that the adaptive particle filter is superior to the unadapted particle filter in terms of efficiency.

The technique proposed in the paper is intended as a module for any PF algorithm which can later be applied in arbitrary problems such as signal processing, image recognition, tracking, navigation and control.

ACKNOWLEDGEMENT

The work was supported by the Ministry of Education, Youth and Sports of the Czech Republic, project No. 1M0572 and by the Czech Science Foundation, project No. P103/11/1353.

(Received June 25, 2010)

REFERENCES

-
- [1] N. Gordon, D. Salmond, and A.F.M. Smith: Novel approach to nonlinear/ non-Gaussian Bayesian state estimation. *IEE Proc. F 140* (1993), 107–113, 1993.
 - [2] A. Doucet, N. de Freitas, and N. Gordon, eds.: *Sequential Monte Carlo Methods in Practice*. Springer, 2001.
 - [3] O. Straka and M. Šimandl: Sampling densities of particle filter: a survey and comparison. In: *Proc. 26th American Control Conference (ACC), AACC, New York 2007*, pp. 4437–4442.
 - [4] M. Šimandl and O. Straka: Nonlinear estimation by particle filters and Cramér–Rao bound. In: *Proc. 15th Triennial World Congress of the IFAC, Barcelona 2002*, pp. 79–84.
 - [5] M. Šimandl, J. Kráľovec, and P. Tichavský: Filtering, predictive and smoothing Cramér–Rao bounds for discrete-time nonlinear dynamic filters. *Automatica 37* (2001), 11, 1703–1716.
 - [6] D. Koller and R. Fratkina: Using learning for approximation in stochastic processes. In: *Proc. 15th Internat. Conference on Machine Learning, Morgan Kaufmann, San Francisco 1998*, pp. 287–295.
 - [7] D. Fox: Adapting the sample size in particle filters through KLD-sampling. *Internat. J. Robotics Research 22* (2003), 985–1003.
 - [8] A. Soto: Self adaptive particle filter. In: *Proc. Internat. Joint Conference on Artificial Intelligence 2005*, pp. 1398–1406.
 - [9] O. Straka and M. Šimandl: Adaptive particle filter based on fixed efficient sample size. In: *Proc. 14th IFAC Symposium on System Identification, Newcastle 2006*.
 - [10] O. Lanz: An information theoretic rule for sample size adaptation in particle filtering. In: *14th Internat. Conference on Image Analysis and Processing (ICIAP 2007)*, pp. 317–322.
 - [11] O. Straka and M. Šimandl: Adaptive particle filter with fixed empirical density quality. In: *Proc. 17th World Congress of the IFAC, 2008*.
 - [12] O. Straka and M. Šimandl: A survey of sample size adaptation techniques for particle filters. In: *Proc. 15th Symposium on System Identification, IFAC, Saint-Malo 2009*.
 - [13] R. Karlsson and F. Gustafsson: Monte Carlo data association for multiple target tracking. In: *IEE Workshop on Target Tracking, Eindhoven 2001*.
 - [14] M. Bolic, S. Hong, and P.M. Djuric: Performance and complexity analysis of adaptive particle filtering for tracking applications. In: *Proc. Conference on Signals, Systems and Computers, 2002*.
 - [15] D.F. Kerridge: Inaccuracy and inference. *J. Roy. Statist. Soc. 23* (1961), 184–194.
 - [16] J. Hayya, D. Armstrong, and N. Gressis: A note on the ratio of two normally distributed variables. *Management Sci. 21* (1975), 11, 338–341.
 - [17] R. van der Merwe and E.A. Wan: Sigma-point particle filters for sequential probabilistic inference in dynamic state-space models. In: *Proc. Internat.l Conference on Acoustics, Speech, and Signal Processing (ICASSP), IEEE, Hong Kong 2003*, p. 4.

- [18] M. Šimandl, J. Královec, and T. Söderström: Advanced point – mass method for nonlinear state estimation. *Automatica* 42 (2006), 7, 1133–1145.

Ondřej Straka, Department of Cybernetics and Research Centre Data – Algorithms – Decision Making, Faculty of Applied Sciences, University of West Bohemia, Univerzitní 8, 306 14 Plzeň. Czech Republic.

e-mail: straka30@kky.zcu.cz

Miroslav Šimandl, Department of Cybernetics and Research Centre Data – Algorithms – Decision Making, Faculty of Applied Sciences, University of West Bohemia, Univerzitní 8, 306 14 Plzeň. Czech Republic.

e-mail: simandl@kky.zcu.cz