15th Triennial World Congress, Barcelona, Spain

NONLINEAR ESTIMATION BY PARTICLE FILTERS AND CRAMÉR-RAO BOUND

Miroslav Šimandl and Ondřej Straka

Department of Cybernetics
Cybernetic System Research Center
Faculty of Applied Sciences, University of West Bohemia
Univerzitní 8, 306 14 Plzeň, Czech Republic
simandl@kky.zcu.cz, straka30@kky.zcu.cz

Abstract: A solution of the Bayesian recursive relations by the particle filter approach is treated. The stress is laid on the sample size setting as the main user design problem. The Cramér-Rao bound was chosen as a tool for setting the sample size for the three basic types of the state estimation, for filtering, prediction and smoothing. The mean square error matrices of particle filter state estimates for different sample sizes and the CR bounds are compared. Quality of the particle filters and their computational load are illustrated in a numerical example.

Keywords: Monte Carlo method, Nonlinear filters, Cramér-Rao bound, Mean-square error, Nonlinear systems.

1. INTRODUCTION

State estimation of discrete-time nonlinear dynamic systems from noisy measurement data has been a subject of considerable research interest for the last three decades. Bayesian approach can be used for design of the general solution of the state estimation problem. The closed form solution of the Bayesian recursive relations (BRR) is available only for a few special cases. Thus, it is necessary to provide some analytical or numerical approximations. The detailed discussion of the development of practical nonlinear estimators is proceeded e.g. in Sorenson (1974), Kulhavý (1996).

In nineties a significant approach to nonlinear filter synthesis using simulation Monte Carlo (MC) methods appeared (Liu and Chen, 1998). Simplicity of the MC approach is the main reason for attractivity of these methods in nonlinear estimation. Concerning on-line estimation, the most important representative of this approach is particle filter (Gordon *et al.*, 1993), (Pitt and Shephard, 1999) which approximates the posterior probability density function (pdf) by weighted random samples. The samples are sometimes called particles which the filter was named after, and can be looked upon as a grid of points covering the state space. Nonetheless the sample size of the filter is insufficiently specified issue and thus should be subject of research interest. There are some indirect recommendations to set out the sufficient sample size e.g in statistical learning theory (Vidyasagar, 2001), however their application in particle filters has not been satisfactorily solved.

An alternative approach to setting sample size for filtering pdf by filtering CR bound was presented in Šimandl and Straka (2001). The aim of the paper is to extend the result for multi-step prediction and smoothing pdf.

The paper is organized as follows: The usage of the particle filter in state estimation is introduced in Section 2, computation of the CR bound in nonlinear estimation is dealt with in Section 3 and usage of the CR bound in determination of the sample size is discussed in Section 4. Further, the numerical illustration of sample size determination procedure is provided in

Section 5 and finally, main results of the paper are summarized in Section 6.

2. STATE ESTIMATION BY PARTICLE FILTER

This section provides usage of the particle filter to the state estimation of a discrete-time nonlinear stochastic system and consequently discussion of sample size as a filter parameter.

The general solution of state estimation problem is provided by the BRR (Kramer and Sorenson, 1988). The particle filter supplies an approximative numerical solution of the BRR. Filtering pdf and predictive pdf of the BRR are approximated by weighted random samples as well as the smoothing pdf.

Consider the discrete time stochastic system:

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

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

where the vectors $\mathbf{x}_k \in \mathbb{R}^n$, $\mathbf{z}_k \in \mathbb{R}^m$ represent the state of the system and the measurement at time k, respectively, $\mathbf{f}_k : \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{h}_k : \mathbb{R}^n \to \mathbb{R}^m$ are known vector functions, $\mathrm{e}_k \in \mathbb{R}^n$, $\mathbf{v}_k \in \mathbb{R}^m$ are state and measurement zero mean white noise sequences with positive definite covariance matrices $\mathbf{Q}_k$ and $\mathbf{R}_k$ respectively, mutually independent and independent of $\mathbf{x}_0$ and the pdf of the initial state $\mathbf{x}_0$ is given by $p(\mathbf{x}_0) = \mathscr{N}(\mathbf{x}_0 : \mathbf{m}_0, \mathbf{M}_0)$.

The recursive relations for filtering pdf, prediction pdf and smoothing pdf have the following form:

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

$$p(\mathbf{x}_l|\mathbf{z}^k) = \int p(\mathbf{x}_{l-1}|\mathbf{z}^k)p(\mathbf{x}_l|\mathbf{x}_{l-1})d\mathbf{x}_{l-1} \tag{4}$$

$$p(\mathbf{x}_k|\mathbf{z}^l) = p(\mathbf{x}_k|\mathbf{z}^k)\int \frac{p(\mathbf{x}_{k+1}|\mathbf{z}^l)}{p(\mathbf{x}_{k+1}|\mathbf{z}^k)} p(\mathbf{x}_{k+1}|\mathbf{x}_k)d\mathbf{x}_{k+1} \tag{5}$$

where $l > k, \mathbf{z}^k = [\mathbf{z}_0^{\mathrm{T}}, \mathbf{z}_1^{\mathrm{T}}, \ldots, \mathbf{z}_k^{\mathrm{T}}]^{\mathrm{T}}$ and $p(\mathbf{x}_0|\mathbf{z}^{-1})$ is prior conditional pdf of the initial state x_0. The relations require knowledge of the transient pdf $p(\mathbf{x}_{k+1}|\mathbf{x}_k)$ and the measurement pdf $p(\mathbf{z}_k|\mathbf{x}_k)$. The pdf's can be found by (1), (2) and are given by:
$p(\mathbf{x}_{k+1}|\mathbf{x}_k) = \mathscr{N}(\mathbf{x}_{k+1} : \mathbf{f}_k(\mathbf{x}_k), \mathbf{Q}_k)$ and
$p(\mathbf{z}_k|\mathbf{x}_k) = \mathscr{N}(\mathbf{z}_k : \mathbf{h}_k(\mathbf{x}_k), \mathbf{R}_k)$.

The idea of the particle filter in nonlinear state estimation is to approximate the filtering, predictive and smoothing pdf $p(\mathbf{x}_k|\mathbf{z}^m)$, $m = 0,1,2\ldots$, by the set $S_{k|m}$ of ν random samples of state at time instant k denoted as $\mathbf{x}_{k|m}^{(i)}, \quad i = 1,\ldots,\nu$ and corresponding weights $w_{k|m}^{(i)}, \quad i = 1,\ldots,\nu$. The sample $\mathbf{x}_{k|m}^{(i)}$ with weight $w_{k|m}^{(i)}$ is called the weighted random sample and denoted as $(\mathbf{x}_{k|m}^{(i)}, w_{k|m}^{(i)})$. Thus, the set $S_{k|m}$ has the following form:
$S_{k|m} = \{(\mathbf{x}_{k|m}^{(1)}, w_{k|m}^{(1)}), \ldots, (\mathbf{x}_{k|m}^{(\nu)}, w_{k|m}^{(\nu)})\}$.

Now, the filtering step given by (3) can be described. Considering the set of weighted samples
$S_{k|k-1} = \{(\mathbf{x}_{k|k-1}^{(1)}, w_{k|k-1}^{(1)}), \ldots, (\mathbf{x}_{k|k-1}^{(\nu)}, w_{k|k-1}^{(\nu)})\}$ representing the approximative pdf $p_A(\mathbf{x}_k|\mathbf{z}^{k-1})$ of the predictive pdf $p(\mathbf{x}_k|\mathbf{z}^{k-1})$ and the new measurement $\mathbf{z}_k$ it is possible to obtain the set $S_{k|k}$ representing the approximative pdf $p_A(\mathbf{x}_k|\mathbf{z}^k)$ of the filtering pdf $p(\mathbf{x}_k|\mathbf{z}^k)$. The samples $\mathbf{x}_{k|k}^{(i)}$ remain the same as samples $\mathbf{x}_{k|k-1}^{(i)}$ and the weights $w_{k|k}^{(i)}$ in $S_{k|k}$ are given by
$w_{k|k}^{(i)} = w_{k|k-1}^{(i)} p_{\mathbf{z}_k|\mathbf{x}_k}(\mathbf{z}_k|\mathbf{x}_{k|k}^{(i)})$.

Due to equality of samples $\mathbf{x}_{k|k}^{(i)}$ for filtering, $\mathbf{x}_{k|k-1}^{(i)}$ for one step prediction and $\mathbf{x}_{k|k+m}^{(i)}$ for smoothing the following notation for samples at time instant k will be introduced: $\mathbf{x}_k^{(i)}$, where $\mathbf{x}_k^{(i)} \triangleq \mathbf{x}_{k|k}^{(i)} = \mathbf{x}_{k|k-1}^{(i)} = \mathbf{x}_{k|k+m}^{(i)}$.

The prediction step related to (4) in BRR between the set $S_{l-1|k}$ approximating the $l-1-k$ predictive (filtering for $l = k+1$) pdf $p(\mathbf{x}_{l-1}|\mathbf{z}^k)$ and the set $S_{l|k}$ approximating the $l-k$ step predictive pdf $p(\mathbf{x}_l|\mathbf{z}^k)$ at the next time instant consists in transforming the samples by the system dynamics. The next ith sample $\mathbf{x}_l^{(i)}$ is randomly generated from the so called importance function, in this case from the transient pdf $p(\mathbf{x}_l|\mathbf{x}_{l-1}^{(i)})$. The weights $w_{l|k}^{(i)}$ belonging to the new samples $\mathbf{x}_l^{(i)}$ are the same as the weights $w_{l-1|k}^{(i)}$ of the last prediction (filtering for $l = k+1$) step. Then the new set $S_{l|k}$ represents the approximative pdf $p_A(\mathbf{x}_l|\mathbf{z}^k)$ of the $l-k$ step predictive pdf $p(\mathbf{x}_l|\mathbf{z}^k)$.

The smoothing step related to (5) consists of updating the previous smoothing (filtering) weights of the samples by information provided by future measurements according to
$w_{k|l}^{(i)} = w_{k|k}^{(i)} \sum_{j=1}^{\nu} w_{k+1|l}^{(j)} p(\mathbf{x}_{k+1|1}^{(j)}|\mathbf{x}_{k|k}^{(i)})$ As already mentioned, the smoothing particles $\mathbf{x}_{k|l}^{(i)}$ are same as filtering ones, i.e. $\mathbf{x}_{k|l}^{(i)} = \mathbf{x}_{k|k}^{(i)}$. As it can be seen, to evaluate approximative pdf $p_A(\mathbf{x}_k|\mathbf{z}^l)$ it is necessary to evaluate all approximative filtering pdf's up to time instant l, i.e. $p_A(\mathbf{x}_l|\mathbf{z}^l)$.

Note that the filter is initialized by the set S_0 of random samples that are equally weighted due to assumed knowledge of the initial pdf $p(\mathbf{x}_0)$ that samples were generated from.

The aforesaid procedure of approximative filtering pdf, prediction pdf and smoothing pdf generation often suffers from samples degeneration that decreases approximation quality. Therefore the algorithm of the particle filter often involves the so called resampling step which is intended to rejuvenate the samples. It removes samples that have very small weights and thus contain “unimportant” information and boosts influence of the important samples with higher weights. The new resampled set of equally weighted samples consists of samples that are drawn from the old set of samples according to the weights.

It is necessary to mention the fact that the presented algorithm is very similar to the well known bootstrap filter. A more sophisticated algorithm should use importance function that involves new measurement into particles generation. It is also advisable to incorporate a MCMC procedure (Markov Chain Monte Carlo) (Godsill *et al.*, 2001) after resampling which improves positioning of the particles in the state space. Nevetherless the presented way of sample size specification can be used for all different types of particle filter.

To summarize, the basic computation scheme of the particle filter follows:

FILTERING PART

0. Initialization: Let $k = 0$. The filter starts from the set S_0 of ν weighted samples $S_{0|-1} = \{(\mathbf{x}_0^{(1)}, w_{0|-1}^{(1)}) \dots (\mathbf{x}_0^{(\nu)}, w_{0|-1}^{(\nu)})\}$ where $\mathbf{x}_0^{(i)} \sim p(\mathbf{x}_{0|\mathbf{z}^{-1}})$; $i = 1,2,\dots,\nu$ and $w_{0|-1}^{(i)} = 1/\nu$; $i = 1,2,\dots,\nu$. This set represents the prior pdf $p(\mathbf{x}_0|\mathbf{z}^{-1})$.

1. Update: The filter updates the weights using the new received measurement $\mathbf{z}_k$ so that

$$w_{k|k}^{(i)} = w_{k|k-1}^{(i)} p(\mathbf{z}_k|\mathbf{x}_k^{(i)}),\ i = 1,2,\dots,\nu$$

Then the weights $w_{k|k}^{(i)}$; $i = 1,2,\dots,\nu$ are scaled to sum to 1. The new set $S_{k|k}$ of weighted samples approximates the posterior pdf $p(\mathbf{x}_k|\mathbf{z}^k)$.

2. Resampling: The resampling step together with the MCMC procedure passes in the case that the Effective Sample Size (ESS) holds the condition $\text{ESS}_k < \nu\,\alpha_k$, where the ESS expresses equivalent number of random draws which describe the pdf with the same accuracy as samples of the particle filter and the $\alpha_k \in (0,1)$ is threshold set by designer. The ESS at time k is approximately given by

$$\text{ESS}_k = \frac{\nu}{1 + \text{var}(w_k^{(i)})}.$$

3. Prediction: The set $S_{k+1|k}$ of weighted samples representing the predictive pdf $p(\mathbf{x}_{k+1}|\mathbf{z}^k)$ is constructed from new samples $\mathbf{x}_{k+1}^{(i)}$ generated from $p(\mathbf{x}_{k+1}|\mathbf{x}_k^{(i)})$ and the corresponding weights $w_{k+1|k}^{(i)} = w_{k|k}^{(i)}$ for $i = 1,2,\dots,\nu$. Now, let $k = k+1$ and continue at the step **1. Update**.

MULTISTEP PREDICTION PART

0. Requirements To obtain approximation of the predictive pdf $p(\mathbf{x}_l|\mathbf{z}^k)$ $(l > k)$, the filtering approximative pdf $p_A(\mathbf{x}_k|\mathbf{z}^k)$ must be evaluated at first. Thus the set of weighted random samples $S_{k|k}$ is required for the evaluation of the samples of the set $S_{l|k}$.

1. Prediction The set $S_{k+1|k}$ of weighted samples representing the predictive pdf $p(\mathbf{x}_{k+1}|\mathbf{z}^k)$ is constructed from new samples $\mathbf{x}_{k+1}^{(i)}$ generated from $p(\mathbf{x}_{k+1}|\mathbf{x}_k^{(i)})$ and the corresponding weights $w_{k+1|k}^{(i)} = w_{k|k}^{(i)}$ for $i = 1,2,\dots,\nu$. This step is repeated for increasing time instant k until the required set $S_{l|k}$ is obtained.

SMOOTHING PART

0. Requirements To obtain approximation of the smoothing pdf $p(\mathbf{x}_k|\mathbf{z}^l)$ $(l > k)$, the approximation of the filtering pdf $p(\mathbf{x}_l|\mathbf{z}^l)$ must be evaluated at first. Thus the set of weighted random samples $S_{l|l}$ is required for the evaluation of the samples of the set $S_{k|l}$.

1. Smoothing The set $S_{l-1|l}$ of weighted samples representing the smoothing pdf $p(\mathbf{x}_{1-1}|\mathbf{z}^l)$ is constructed from the samples $\mathbf{x}_{l-1}^i$ and the corresponding weights $w_{l-1|l}^{(i)} = w_{l-1|l-1}^{(i)} \sum_{j=1}^{\nu} w_{l|l}^{(j)}\, p(\mathbf{x}_l^{(j)}|\mathbf{x}_{l-1}^{(i)})$ for $i = 1,2,\dots,\nu$. This step is repeated until the required set $S_{k|l}$ is obtained.

It is obvious that this basic scheme of the particle filter is relatively simple, nevertheless the crucial parameter ν defining the sample size and strongly affecting approximation quality is not sufficiently specified. A procedure for sample size setting is designed for filtering problem in Šimandl and Straka (2001) where determination of a feasible sample size is based on filtering CR bound.

To extend the result for prediction and smoothing, the predictive and smoothing Cramér-Rao (CR) bound (Šimandl *et al.*, 2001) will be chosen for determination of a feasible sample size of the predictive and smoothing particle filters. Hence, the recursive relations for filtering, predictive and smoothing CR bounds will be presented in the next section.

3. CR BOUND FOR NONLINEAR FILTERING

The CR bound is the standard lower bound in parameter estimation (Kerr, 1989) defined as the inverse of the Fisher Information Matrix (FIM). The idea of the CR bound for random parameters can be applied to the state estimation problem for nonlinear stochastic dynamic systems, (Bobrovsky *et al.*, 1987), (Doerschuk, 1995). This section is based on (Šimandl *et al.*, 2001) where the recursive relation for computation of filtering, multi-step predictive and smoothing CR bounds are derived.

Consider nonlinear systems (1), (2) with Gaussian initial conditions and Gaussian disturbances, where $\mathbf{Q}_k$, $\mathbf{R}_k$ are non-singular matrices. To simplify notations, the nabla operator will be used $\nabla_{\mathbf{x}} = \left[\frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} \cdots \frac{\partial}{\partial x_n}\right]$.

The filtering, predictive and smoothing CR bounds $\mathbf{C}_{k|k}$, $\mathbf{C}_{l|k}$ and $\mathbf{C}_{k|l}$ for state $\mathbf{x}_k$ can be derived from the FIM $\mathbf{J}(\mathbf{x}^k)$ for the complete state history $\mathbf{x}^k \triangleq [\mathbf{x}_0^{\mathrm{T}}, \mathbf{x}_1^{\mathrm{T}}, \ldots, \mathbf{x}_k^{\mathrm{T}}]^{\mathrm{T}}$. The state history $\mathbf{x}^k$ may be interpreted as a vector of parameters of the random measured vector $\mathbf{z}^k$. In this case the FIM $\mathbf{J}(\mathbf{x}^k)$ is defined as:

$$\mathbf{J}(\mathbf{x}^k) \triangleq -\mathsf{E}\{\nabla_{\mathbf{x}^k}[\nabla_{\mathbf{x}^k} \ln p(\mathbf{x}^k, \mathbf{z}^k)]\}, \qquad (6)$$

provided that the derivatives and expectation exist.

Knowledge of the $(k+1)n \times (k+1)n$ FIM matrix is fundamental for derivation of recursive relations for the filtering, the predictive and the smoothing CR bounds $\mathbf{C}_{k|k}$, $\mathbf{C}_{l|k}$ and $\mathbf{C}_{k|1}$. The detailed description of the above outlined derivation is provided in Šimandl *et al.* (2001).

The CR bound sets the limit of cognizability of the state of the stochastic dynamic system and thus it may serve as a gauge for evaluating the filter performance quality.

Let $\hat{\mathbf{x}}_{k|k}$, $\hat{\mathbf{x}}_{l|k}$ and $\hat{\mathbf{x}}_{k|l}$ be arbitrary filtering, predictive and smoothing point estimates of the state. Conditional-mean values generated by the particle filter will be preferred in this paper.

The mean-square error matrices (MSEMs) $\Pi_{k|k}$, $\Pi_{l|k}$ and $\Pi_{k|l}$ are defined as:

$$\Pi_{k|k} \triangleq \mathsf{E}\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})^{\mathrm{T}}\} \qquad (7)$$

$$\Pi_{l|k} \triangleq \mathsf{E}\{(\mathbf{x}_l - \hat{\mathbf{x}}_{l|k})(\mathbf{x}_l - \hat{\mathbf{x}}_{l|k})^{\mathrm{T}}\} \qquad (8)$$

$$\Pi_{k|l} \triangleq \mathsf{E}\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|l})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|l})^{\mathrm{T}}\}. \qquad (9)$$

The MSEMs are bounded by the filtering CR bound $\mathbf{C}_{k|k}$, the predictive CR bound $\mathbf{C}_{l|k}$ and the smoothing CR bound $\mathbf{C}_{k|l}$ as follows:

$$\Pi_{k|k} \geq \mathbf{C}_{k|k}(\mathbf{x}_k) \qquad (10)$$

$$\Pi_{l|k} \geq \mathbf{C}_{l|k}(\mathbf{x}_k) \qquad (11)$$

$$\Pi_{k|l} \geq \mathbf{C}_{k|l}(\mathbf{x}_k). \qquad (12)$$

To compute (6) the following notation for $n \times n$ matrices will be introduced:

$$\mathbf{K}_{i+1}^{i} = \mathsf{E}\{-\nabla_{\mathbf{x}_i}[\nabla_{\mathbf{x}_i} \ln p(\mathbf{x}_{i+1}|\mathbf{x}_i)]^{\mathrm{T}}\} \qquad (13)$$

$$\mathbf{K}_{i+1}^{i,i+1} = \mathsf{E}\{-\nabla_{\mathbf{x}_{i+1}}[\nabla_{\mathbf{x}_i} \ln p(\mathbf{x}_{i+1}|\mathbf{x}_i)]^{\mathrm{T}}\} \qquad (14)$$

$$\mathbf{K}_{i+1}^{i+1} = \mathsf{E}\{-\nabla_{\mathbf{x}_{i+1}}[\nabla_{\mathbf{x}_{i+1}} \ln p(\mathbf{x}_{i+1}|\mathbf{x}_i)]^{\mathrm{T}}\} \qquad (15)$$

$$\mathbf{L}_{i}^{i} = \mathsf{E}\{-\nabla_{\mathbf{x}_i}[\nabla_{\mathbf{x}_i} \ln p(\mathbf{z}_i|\mathbf{x}_i)]^{\mathrm{T}}\} \qquad (16)$$

with $\mathbf{K}_{i+1}^{i+1,i} = [\mathbf{K}_{i+1}^{i,i+1}]^{\mathrm{T}}$, $i = 0, 1, \ldots, k$, and

$$\mathbf{K}_0^0 = \mathsf{E}\{-\nabla_{\mathbf{x}_0}[\nabla_{\mathbf{x}_0} \ln p(\mathbf{x}_0)]^{\mathrm{T}}\}. \qquad (17)$$

Considering the nonlinear system (1), (2), the relations (13) - (17) have the following form:

$$\mathbf{K}_{i+1}^{i} = \mathsf{E}\{[\nabla_{\mathbf{x}_i}\mathbf{f}_i(\mathbf{x}_i)]^{\mathrm{T}}\mathbf{Q}_i^{-1}\nabla_{\mathbf{x}_i}\mathbf{f}_i(\mathbf{x}_i)\} \qquad (18)$$

$$\mathbf{K}_{i+1}^{i,i+1} = -\mathsf{E}\{[\nabla_{\mathbf{x}_i}\mathbf{f}_i(\mathbf{x}_i)]^{\mathrm{T}}\}\mathbf{Q}_i^{-1} \qquad (19)$$

$$\mathbf{K}_{i+1}^{i+1} = \mathbf{Q}_i^{-1} \qquad (20)$$

$$\mathbf{L}_{i}^{i} = \mathsf{E}\{[\nabla_{\mathbf{x}_i}\mathbf{h}_i(\mathbf{x}_i)]^{\mathrm{T}}\mathbf{R}_i^{-1}\nabla_{\mathbf{x}_i}\mathbf{h}_i(\mathbf{x}_i)\}. \qquad (21)$$

$$\mathbf{K}_0^{00} = \mathbf{M}_0^{-1}. \qquad (22)$$

Then the recursive relations for the filtering, the predictive and the smoothing CR bounds are given by:

$$\mathbf{C}_{k|k}^{-1} = \mathbf{C}_{k|k-1}^{-1} + \mathbf{L}_k^k \qquad (23)$$

$$\mathbf{C}_{\ell|k}^{-1} = \mathbf{K}_{\ell}^{\ell} - \mathbf{K}_{\ell}^{\ell,\ell-1}(\mathbf{K}_{\ell}^{\ell-1} + \mathbf{C}_{\ell-1|k}^{-1})^{-1}\mathbf{K}_{\ell}^{\ell-1,\ell} \qquad (24)$$

$$\mathbf{C}_{k|\ell}^{-1} = \mathbf{C}_{k|k}^{-1} + \mathbf{K}_{k+1}^{k} - \mathbf{K}_{k+1}^{k,k+1}(\mathbf{K}_{k+1}^{k+1} + \mathbf{C}_{k+1|\ell}^{-1} - \mathbf{C}_{k+1|k}^{-1})^{-1}\mathbf{K}_{k+1}^{k+1,k} \qquad (25)$$

4. THE USE OF CR BOUND IN PARTICLE FILTER DESIGN

Accuracy of the particle filter is affected by the sample size. Usually the sample size is determined ad hoc by the filter designer. A more technical approach can be based on comparing the MSEMs obtained from the particle filters with different sample sizes and confronting them with the CR bound. This procedure allows to set a feasible sample size.

To compare accuracy of the particle filter with different sample sizes the following criterion utilizing the CR bound was chosen:

$$V_{k|m}^{CR}(\nu) = \frac{1}{N}\sum_{k=0}^{N} \mathrm{tr}(|\Pi_{k|m} - \mathbf{C}_{k|m}|), \qquad (26)$$

where $\mathrm{tr}(\cdot)$ denotes matrix trace and $\Pi_{k|m}$ is one of the MSEMs (7), (8), (9) which depend on ν. The sample size can be determined with respect to value of $V_{k|m}^{CR}(\nu)$ for different sample sizes and also respecting computational demands of the filter. Comparing values of the criterion, one can set up the sample size which is sufficient to ensure satisfactory quality of the filter.

Both the MSEM of a state estimate and the CR bound are computed using the MC simulation method. After carrying M experiments for the system (1), (2) with $k = 0, 1, \ldots, N$, we obtain M realizations of the state trajectory $\{\mathbf{x}_k\}_{k=0}^N$ and of corresponding measurements $\{\mathbf{z}_k\}_{k=0}^N$. The CR bounds are computed by (23),(24) and (26) with the matrices $\mathbf{K}_{k,k}^{k+1}$, $\mathbf{K}_{k,k+1}^{k+1}$ and

$\mathbf{L}_k$ replaced by their estimates; e.g. the estimate of $\mathbf{K}_{k,k+1}^{k+1}$ is computed as:

$$\hat{\mathbf{K}}_{k,k+1}^{k+1} = \frac{1}{M}\sum_{j=1}^{M}\left[-\frac{\partial \mathbf{f}_k(\mathbf{x}_k)}{\partial \mathbf{x}_k}\bigg|_{\mathbf{x}_k=\mathbf{x}_k(j)}\right]^{\mathrm{T}}\mathbf{Q}_k^{-1}, \quad (27)$$

where $\{\mathbf{x}_k(j)\}_{k=0}^{N}$ is a j-th state trajectory, $j = 1, 2, \ldots, M$.

The MSEM $\Pi_{k|m}$ is estimated from M different sequences of filtering, prediction and smoothing estimates $\{\hat{\mathbf{x}}_{k|m}(j)\}_{k=0}^{N}$ which correspond to M state trajectories, thus

$$\hat{\Pi}_{k|m} = \frac{1}{M}\sum_{j=1}^{M}\left[\mathbf{x}_k(j) - \hat{\mathbf{x}}_{k|m}(j)\right]\left[\mathbf{x}_k(j) - \hat{\mathbf{x}}_{k|m}(j)\right]^{\mathrm{T}}. \quad (28)$$

This estimate converges to the $\Pi_{k|m}$ for $M \to \infty$. Thus for a sufficient number of MC runs the estimate $\hat{\Pi}_{k|m}$ is close to the MSEM $\Pi_{k|m}$ which is always greater than or equal to the filtering, prediction and smoothing CR bound. The greater the difference is, the worse is quality of the filter. It is necessary for MC estimation of the CR bound and MSEMs to specify sufficient number of MC runs. The number of MC runs can be chosen such that further increasing of MC runs influences results very slightly.

5. NUMERICAL ILLUSTRATIONS

The proposed procedure of determination the sample size for the particle filter will be presented for a nonlinear stochastic system.

Consider the nonlinear system with two-dimensional state $\mathbf{x}_k \triangleq [x_{1,k}, x_{2,k}]^{\mathrm{T}}$ described by

$$\begin{bmatrix} x_{1,k+1} \\ x_{2,k+1} \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0 & 1.0 \end{bmatrix}\begin{bmatrix} x_{1,k} \\ x_{2,k} \end{bmatrix} + \begin{bmatrix} e_{1,k} \\ e_{2,k} \end{bmatrix} \quad (29)$$

with the state noise $\mathbf{e}_k \triangleq [e_{1,k}, e_{2,k}]^{\mathrm{T}}$ with $\mathscr{N}(\mathbf{e}_k : [0\ 0]^{\mathrm{T}}, \mathbf{Q})$, where $\mathbf{Q} = \begin{bmatrix} 0.1 & 0.05 \\ 0.05 & 0.1 \end{bmatrix}$. The state is observed by the scalar measurement z_k described by the measurement equation:

$$z_k = atan\left(\frac{x_{2,k} - \sin(k)}{x_{1,k} - \cos(k)}\right) + v_k \quad (30)$$

The measurement z_k is influenced by the measurement noise v_k with pdf $v_k \sim \mathscr{N}(v_k : 0, r)$ where $r = 0.025$. The initial state is given by $p(\mathbf{x}_0) = \mathscr{N}(\mathbf{x}_0[27,\ 0]^{\mathrm{T}}, \mathrm{diag}\{9,\ 4\})$. The predictive pdf $p(\mathbf{x}_0|\mathbf{z}^{-1})$ for the filter is same as $p(\mathbf{x}_0)$.

Table 1 presents value of the criterion (26) for different sample sizes. The table together with Figs. 2, 3 and 4 where the traces of the filtering, prediction and smoothing MSEMs for $\nu = 1\,000$ tend to the corresponding CR bound, demonstrate sufficient quality of the filter for this sample size. On the other hand, the experiments also indicate that the same estimation quality given by the criterion (26) requires slightly different sample size for prediction, filtering and smoothing.

The CR bound for this system was evaluated using $M = 10\,000$ MC simulations. The number of simulation used for MSEMs evaluation for $\nu = 50$, 300 was $M = 10\,000$ whereas for $\nu = 1000$ $M = 1000$ simulations were used only. The reason is extreme growth of computational load as documented in Table 2. That is the reason why trace of the MSEM is at some k lower than the CR bound.

Table 1. Value of the criterion V^{CR} with respect to the sample size

ν	50	100	300	1000
$V_{k\|k}^{CR}$	0.91	0.84	0.80	0.42
$V_{k\|k-2}^{CR}$	1.10	0.71	0.66	0.35
$V_{k\|k+1}^{CR}$	1.56	0.86	0.86	0.41

Fig. 1 presents comparison of time development of traces of the filtering, predictive and smoothing CR bounds.

Figs. 2, 3 and 4 illustrate time development of traces of MSEMs of the estimates obtained from the particle filter with different sample sizes compared to time development of CR bound trace.

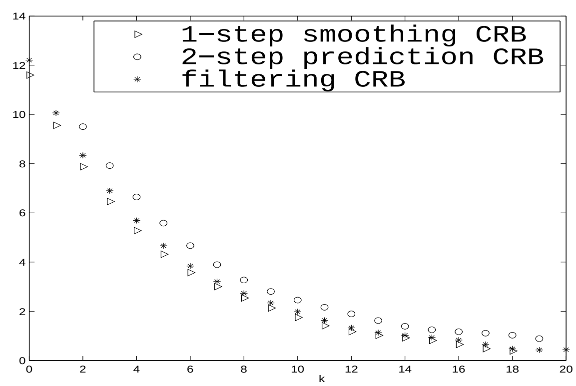

Fig. 1. Time development of traces of the filtering, predictive and smoothing CR bounds

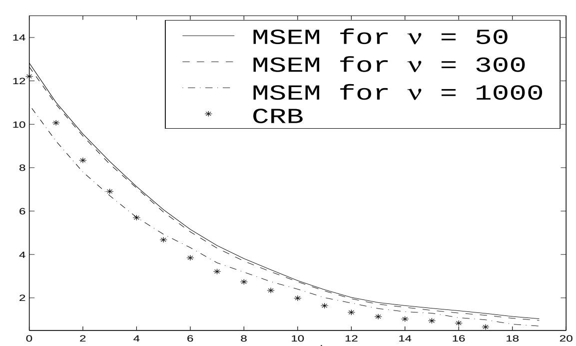

Fig. 2. Time development of traces of the CR bound and MSEMs of estimate produced by filter with different sample size

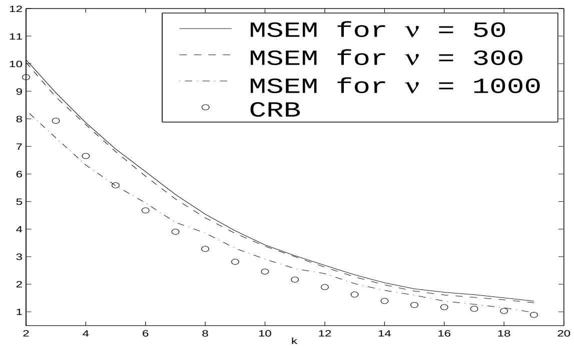

Fig. 3. Time development of traces of the CR bound and MSEMs of estimate produced by 2-step predictor with different sample size

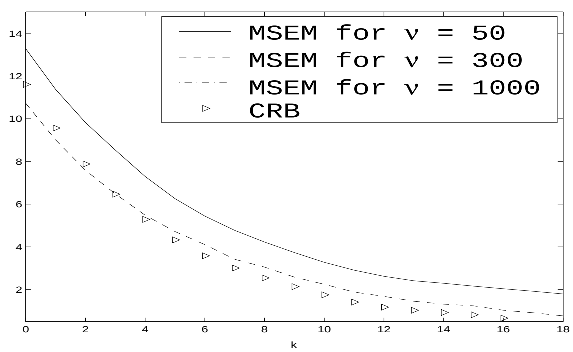

Fig. 4. Time development of traces of the CR bound and MSEMs of estimate produced by 1-step smoother with different sample size

Table 2. Number of FLOPS used for one simulation for different ν

ν	50	300	1000
filtering	$3.66 \cdot 10^6$	$7.04 \cdot 10^6$	$1.03 \cdot 10^8$
2-step prediction	$7.04 \cdot 10^6$	$4.50 \cdot 10^7$	$1.69 \cdot 10^8$
1-step smoothing	$2.70 \cdot 10^7$	$8.62 \cdot 10^7$	$9.94 \cdot 10^8$

6. CONCLUSION

Aspects of particle filter solution of nonlinear estimation were discussed and a new procedure for sample size setting was developed. The procedure was based on comparison of the particle filter MSEMs for filtering, prediction and smoothing with the corresponding CR bound. This allows to take into account computational load and filter quality in sample size setting. To illustrate the designed procedure, a numerical example was presented. The experiments also indicate that the same estimation quality requires different sample size for prediction, filtering and smoothing.

7. ACKNOWLEDGMENT

The work was supported by the Grant Agency of the Czech Republic, project GA ČR 102/01/0021 and the Ministry of Education, Youth and Sports of the Czech Republic, project MSM 2352 00004.

8. REFERENCES

Bobrovsky, B.Z., E. Mayer-Wolf and M. Zakai (1987). Some Classes of Global Cramér-Rao Bounds. *Annals of Statistics* **15**(4), 1421–1438.

Doerschuk, P.C. (1995). Cramer-Rao Bounds for Discrete-Time Nonlinear Filtering Problems. *IEEE Trans. on Automatic Control*.

Godsill, S., and T. Clapp, (2001). Improvement Strategies for Monte Carlo Particle Filters In: *Monte Carlo Methods in Practise*. Springer-Verlag, New York.

Gordon, N., D. Salmond and A.F.M. Smith (1993). Novel approach to nonlinear/non-gaussian Bayesian state estimation. *IEE proceedings-F* **140**, 107–113.

Kerr, T.H. (1989). Status of CR-Like Lower Bounds for Nonlinear Filtering. *IEEE Trans. on Aerospace and Electronic Systems*.

Kramer, S. and H. Sorenson (1988). Recursive Bayesian estimation using piece-wise constant approximations. *Automatica* **24**(6), 789–801.

Kulhavý, R. (1996). *Recursive Nonlinear Estimation*. Springer. London.

Liu, J.S. and R. Chen (1998). Sequential Monte Carlo Methods for Dynamic Systems. *J. Amer. Statist. Assoc.* **93**(443), 1032–1044.

Pitt, M.K. and N. Shephard (1999). Filtering via simulation: auxiliary particle filter. *J. Amer. Statist. Assoc.* **94**, 590–599.

Šimandl, M. and O. Straka (2001). Setting sample size using CR bound. In: *Preprints of 5th IFAC Symposium !Nonlinear Control Systems*. Saint Petersburg.

Šimandl, M., J. Královec and P. Tichavský (2001). Filtering, predictive and smoothing Cramér-Rao bounds for discrete-time nonlinear dynamic filters. *Automatica*.

Sorenson, H.W. (1974). On the Development of Practical Nonlinear Filters. *Inf. Sci.* **7**, 230–270.

Vidyasagar, M. (2001). Randomized algorithms for robust controller synthesis using statistical learning theory: A tutorial overview. *European Journal of Control* **7**(2), 287–30.