

ASSESSING ROBUSTNESS OF PARTICLE FILTERING BY THE KOLMOGOROV-SMIRNOV STATISTICS

Pau Closas[†], Mónica F. Bugallo* and Petar M. Djurić*

[†] Dept. of Signal Theory and Communications, Universitat Politècnica de Catalunya
Campus Nord UPC, 08034 Barcelona, Spain.

* Dept. of Electrical and Computer Engineering, Stony Brook University
Stony Brook, New York 11779, USA.

ABSTRACT

One of the most criticized aspects of particle filtering algorithms is their dependence on model assumptions. However, a rigorous study of the effect of modeling errors on the performance of such algorithms is still missing. In this paper, the problem of using an inaccurate discrete state-space model is considered and a systematic methodology for studying the effects on its performance is proposed. The methodology is based on the use of the Kolmogorov-Smirnov statistic, which in this case is a distance metric between the posterior characterization when respectively correct and incorrect model assumptions are made. An example with functional and distributional inaccuracies is studied.

Index Terms— Robustness, error analysis, Monte Carlo methods, filtering.

1. INTRODUCTION

The discrete state-space (DSS) approach is adopted to deal with the non-linear filtering problems, as it describes the evolution of states and its observation model. The problem consists of the recursive estimation of the states $\mathbf{x}_k \in \mathbb{R}^{n_x}$ given the measurements $\mathbf{y}_k \in \mathbb{R}^{n_y}$ at time instant $k \in \mathbb{N}$, where n_x and n_y are the dimensions of the state and measurement vectors, respectively. The state equation models the evolution of the state vector (on which the measurements depend on) as a discrete-time stochastic model, where in general

$$\mathbf{x}_k \sim p(\mathbf{x}_k | \mathbf{x}_{k-1}) \quad \text{for } k \geq 1 \quad (1)$$

and where $p(\mathbf{x}_k | \mathbf{x}_{k-1})$ is referred to as the transitional prior. The relationship between the measurements and the state is

[†]This work has been financed by the Spanish/Catalan Science and Technology Commissions and FEDER funds from the European Commission: TEC2006-06481/TCM, TEC2004-04526, TIC2003-05482, CONSOLIDER CSD2008-00010 (COMONSENS) and 2005SGR-00639.

*This work has been supported by the National Science Foundation under Award CCF-0515246 and the Office of Naval Research under Award N00014-06-1-0012. The third author also gratefully acknowledges the support from the award Universidad Carlos III de Madrid-Banco de Santander Chair of Excellence.

generically modeled by the probability distribution

$$\mathbf{y}_k \sim p(\mathbf{y}_k | \mathbf{x}_k) \quad \text{for } k \geq 1 \quad (2)$$

referred to as the likelihood function. The initial *a priori* probability density function of the state vector, defined as $p(\mathbf{x}_0)$, is assumed to be also known.

The presented DSS model is the basis of Bayesian filtering, where point estimates of the states are obtained from their posterior probability density function (PDF). Closed-form solutions to the filtering problem modeled in equations (1) and (2) are only possible in limited situations where the DSS model satisfies certain conditions. That is the case of linear/Gaussian DSS models which can be optimally processed using the Kalman Filter (KF). However, one usually has to resort to sub-optimal methods to deal with general DSS models. Particle filters (PFs) provide a powerful tool to deal with non-linear/non-Gaussian DSS models [1, 2]. In brief, PFs characterize the posterior distribution by a set of N_s random samples taken from an *importance density function*, $\mathbf{x}_k^i \sim \pi(\mathbf{x}_k | \mathbf{x}_{k-1}^i, \mathbf{y}_{1:k})$, with associated importance weights w_k^i . In general,

$$\begin{aligned} \mathbf{x}_k^i &\sim \pi(\mathbf{x}_k | \mathbf{x}_{k-1}^i, \mathbf{y}_k) \\ w_k^i &= w_{k-1}^i \frac{p(\mathbf{y}_k | \mathbf{x}_k^i) p(\mathbf{x}_k^i | \mathbf{x}_{k-1}^i)}{\pi(\mathbf{x}_k^i | \mathbf{x}_{k-1}^i, \mathbf{y}_{1:k})} \frac{1}{t} \end{aligned} \quad (3)$$

where t is a normalization constant.

For a set of generated particles, $\{\mathbf{x}_k^i, w_k^i\}_{i=1}^{N_s}$, the characterization of the marginal posterior PDF is given by

$$\hat{p}(\mathbf{x}_k | \mathbf{y}_{1:k}) = \sum_{i=1}^{N_s} w_k^i \delta(\mathbf{x}_k - \mathbf{x}_k^i) \quad (4)$$

where $\delta(\cdot)$ is the Dirac's delta function. This approximation converges a.s. to the true posterior as $N_s \rightarrow \infty$ if the support of the chosen importance density includes the support of the posterior. Resampling, consisting in replacing particles with low importance weights and replicating those with high importance weights, is performed after state estimation and when significant degeneracy of the particles is observed.

This paper aims at studying the effect of model inaccuracies in the performance of PFs. In Section 2, motivation is provided and a qualitative definition of Bayesian robustness is introduced. The idea behind the presented study is to compare posterior characterizations of two PFs: one with correct DSS model assumptions and another with an inaccurate model. However, PFs exhibit special difficulties when one aims at comparing their posterior approximations, as they may not share the same support. To overcome the latter, the Kolmogorov-Smirnov statistic is considered and presented in Section 3. In Section 4 the proposed methodology is applied to a practical example, and in Section 5 concluding remarks are provided.

2. ROBUSTNESS TO MODEL INACCURACIES

It is important to remark that all methods (whether optimal or sub-optimal) used in Bayesian filtering assume a perfect knowledge of the underlying DSS model. Although it seems apparent that in most cases the considered model may differ from the actual one in some sense, few studies have been devoted in the literature to the evaluation of such effect on Bayesian filtering. For instance, model distributions elicited from empirical data can be erroneous, yielding to a filtering error whose magnitude is yet to be evaluated. Basic research on robust statistics has been focused on the parameter estimation problem [3], while Bayesian robustness has been studied mostly to identify wrongly elicited models [4]. Surprisingly, little (or no) attention has been given to the study of the effect of model inaccuracies on the output of a filtering algorithm, e.g. PFs, except to that of linear filters [5].

The situation where one uses incorrect model can be expressed by modifying the DSS model to

$$\begin{aligned} \mathbf{x}_0 &\sim \tilde{p}(\mathbf{x}_0) \\ \mathbf{x}_k &\sim \tilde{p}(\mathbf{x}_k|\mathbf{x}_{k-1}) \quad \text{for } k \geq 1 \\ \mathbf{y}_k &\sim \tilde{p}(\mathbf{y}_k|\mathbf{x}_k) \quad \text{for } k \geq 1 \end{aligned} \quad (5)$$

where the tilde denotes that a certain degree of uncertainty may be present in the corresponding distribution.

We now extend the commonly adopted Hampel's definition for qualitative robustness [6] to its Bayesian counterpart:

Definition 2.1 (*Qualitative Bayesian Robustness*) Consider a DSS model and let the states \mathbf{x}_k evolve according to a distribution $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ denoted by \mathcal{P}_k and let the observations \mathbf{y}_k be i.i.d with distribution $p(\mathbf{y}_k|\mathbf{x}_k)$, which we denote by \mathcal{L}_k . Consider a functional T_n that produces the desired statistics of the model relying on the posterior PDF, whose distribution is denoted by $\mathfrak{L}_{\mathcal{M}}(T_n)$, where \mathcal{M} denotes the underlying model defined by \mathcal{P}_k and \mathcal{L}_k .

T_n is called robust at $\mathcal{M}_k = \tilde{\mathcal{M}}_k$ if the distribution of T_n is equicontinuous at $\tilde{\mathcal{M}}_k$, that is, if we take a suitable distance d_* , in the space \mathfrak{M} of probability measures, and assume that

for all $\epsilon > 0$ there exists a $\delta_p, \delta_l > 0$ such that,

$$\left. \begin{aligned} d_*(\tilde{\mathcal{P}}_k, \mathcal{P}_k) &\leq \delta_p \\ d_*(\tilde{\mathcal{L}}_k, \mathcal{L}_k) &\leq \delta_l \end{aligned} \right\} \Rightarrow d_*(\mathfrak{L}_{\tilde{\mathcal{M}}}(T_n), \mathfrak{L}_{\mathcal{M}}(T_n)) \leq \epsilon. \quad (6)$$

In other words, if a bounded change in the distributions that define the DSS model is seen as a bounded change in the distribution of the estimates, then the claim is that the estimator is robust.

In this paper, we are interested in studying the effect of incorrect model in (5) on the posterior PDF as obtained by the PF. In the sequel, we express by $\hat{p}_1(\mathbf{x}_k|\mathbf{y}_{1:k})$ the obtained posterior when the correct model is used and with $\hat{p}_2(\mathbf{x}_k|\mathbf{y}_{1:k})$ the posterior computed with the incorrect assumptions. We are interested in using a distance metric that quantifies the difference between distributions. In particular, we have two posterior characterizations which correspond to two PFs that differ in some sense (due to inaccurate prior/likelihood models) and whose supports may be different because of resampling and/or considering different number of particles. Thus, it seems appropriate to compare the cumulative distribution functions (CDFs) of the unknown states instead of their probability mass functions. In that case, the requirement of identical supports as in the case of the computation of the Kullback-Leibler distance is not needed. The Kolmogorov-Smirnov (K-S) statistic provides a suitable metric for comparison of CDFs, and it has been extensively used in the literature [7, 8, 9].

For the sake of simplicity, we consider one-dimensional¹ posterior distributions, i.e., $n_x = 1$. The reason for such simplification is that methods for computing the K-S statistic for multi-dimensional distributions are much more computationally intensive than those available for one-dimensional distributions [8, 10].

3. TWO-SAMPLES KOLMOGOROV-SMIRNOV TEST

The K-S test is a nonparametric test of equality of one-dimensional probability distributions [7]. It is commonly used in statistics to quantify the distance between an empirical distribution function and a reference CDF. In our case we are interested in the two-samples K-S test, where the distance is between two empirical distributions [11].

To form the statistic $D(\cdot, \cdot)$ from the sample distribution functions $\hat{F}_1(x_k)$ and $\hat{F}_2(x_k)$, we compute their maximum absolute difference over all the values of x

$$D\left(\hat{F}_1(x_k), \hat{F}_2(x_k)\right) = \sup_x \left| \hat{F}_1(x_k) - \hat{F}_2(x_k) \right| \quad (7)$$

which is known as the two-sample K-S statistic D . Note that $\hat{F}_j(x_k)$ is the CDF that corresponds to the posterior distribution $\hat{p}_j(\mathbf{x}_k|\mathbf{y}_{1:k})$. The K-S test uses metric in (7) to decide

¹where x_k denotes scalar, as opposed to the vectorial notation \mathbf{x}_k .

between the null hypothesis (\mathcal{H}_0) that both distributions characterize the same posterior PDF or the alternate hypothesis (\mathcal{H}_1) that they approximate different PDFs, or

$$\begin{aligned}\mathcal{H}_0 &: \hat{p}_1(x_k|\mathbf{y}_{1:k}) = \hat{p}_2(x_k|\mathbf{y}_{1:k}) \\ \mathcal{H}_1 &: \hat{p}_1(x_k|\mathbf{y}_{1:k}) \neq \hat{p}_2(x_k|\mathbf{y}_{1:k}).\end{aligned}\quad (8)$$

The K-S test is then implemented by defining a critical distance value $d_\alpha(n_1, n_2)$ such that if $D > d_\alpha(n_1, n_2)$, the hypothesis \mathcal{H}_0 is rejected. Basically, $d_\alpha(n_1, n_2)$ is the highest tolerable distance in order not to reject \mathcal{H}_0 given the level of significance α , where

$$Pr \left\{ D \left(\hat{F}_1(x_k), \hat{F}_2(x_k) \right) > d_\alpha(n_1, n_2) \right\} = \alpha, \quad (9)$$

with $0 < \alpha \leq 1$ indicating how strict the test for rejecting the null hypothesis is. The usefulness of the K-S statistic comes from the fact that the distribution of the statistic $D(\hat{F}_1(x_k), \hat{F}_2(x_k))$ is known. We can compute $Pr \{D > d_\alpha\}$ either exactly for small sample sizes or by using approximations for large sizes [8, 12]. For example, one can show that

$$\begin{aligned}Pr \{D > d_\alpha\} &= Q_{KS} \left((\sqrt{n_e} + 0.12 + 0.11/\sqrt{n_e})d_\alpha \right) \\ Q_{KS}(\lambda) &= 2 \sum_{i=1}^{\infty} (-1)^{i-1} e^{-2i^2\lambda^2},\end{aligned}\quad (10)$$

where $Q_{KS}(\lambda)$ is a monotonic function with limiting values $Q_{KS}(0) = 1$ and $Q_{KS}(\infty) = 0$. In the two-samples case, we have that $n_e = \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$, where n_j is the number of particles used to characterize $\hat{p}_j(x_k|\mathbf{y}_{1:k})$. The approximation in (10) is valid for $n_e \geq 4$ (see [8] and the references therein).

Therefore, after a significance level α is chosen, one can obtain the corresponding critical value. These values can either be computed as in (10) or be found in tables. It is important to note that these values are independent of the underlying distributions.

4. A CASE STUDY

As an example, we considered a one-dimensional nonlinear time series whose *true* underlying DSS model was defined by:

$$\begin{aligned}x_k &= \frac{x_{k-1}}{2} + 25 \frac{x_{k-1}}{1 + x_{k-1}^2} + 8 \cos(1.2k) + u_k \\ y_k &= \frac{x_k^2}{\beta} + v_k,\end{aligned}\quad (11)$$

where $\beta = 20$, $u_k \sim \mathcal{N}(0, \sigma_u^2)$ and $v_k \sim \mathcal{N}(0, \sigma_v^2)$. In the simulations we set the unit measurement noise variance, $\sigma_v^2 = 10$ and the initial distribution $p(x_0) = \mathcal{N}(0, 1)$. The DSS model in (11) is often used for testing in the particle filtering literature [2].

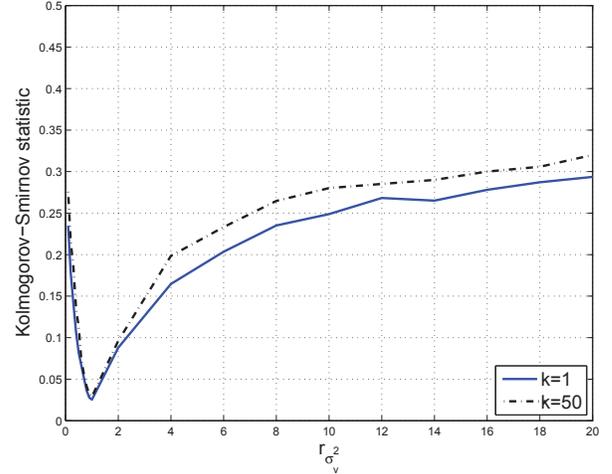


Fig. 1. K-S statistic as a function of $r_{\sigma_v^2}$.

The PF under study was the so-called Bootstrap Filter (BF) [2], where the chosen importance density function is the transitional prior, and the weights are proportional to the likelihood. The posterior PDF obtained by the BF was denoted by $\hat{p}_1(x_k|\mathbf{y}_{1:k})$. We also introduced an incorrect assumption in the DSS model and used a second BF. The resulting (probably erroneous) posterior approximation was denoted by $\hat{p}_2(x_k|\mathbf{y}_{1:k})$. The K-S statistics were averaged over 100 Monte Carlo runs and 10^4 particles were considered for both PFs.

First, we studied the modeling error in the variance of the measurement noise. In other words, we wrongly assumed that v_k was distributed according to $\mathcal{N}(0, \tilde{\sigma}_v^2)$. For the sake of clarity, we defined the ratio between the measurement variances as $r_{\sigma_v^2} = \frac{\tilde{\sigma}_v^2}{\sigma_v^2}$. Figure 1 shows the evaluated K-S statistic as a function of $r_{\sigma_v^2}$ for $k = 1$ and after $k = 50$ instants. From the results it is apparent that, as the inaccuracy increased the distance between the “correct” and wrong posterior increased too. The minimum was attained for $r_{\sigma_v^2} = 1$, as it was expected. From Figure 1 we see that the discrepancy did not significantly increase over time.

Secondly, we studied the effect of assuming erroneous state variances. Similarly, we considered that $u_k \sim \mathcal{N}(0, \tilde{\sigma}_u^2)$ and we defined the ratio $r_{\sigma_u^2} = \frac{\tilde{\sigma}_u^2}{\sigma_u^2}$ between the wrong and correct variances. Figure 2 shows the K-S statistic in that case. After comparing these results with those of Figure 1, we can conclude that, for this setup the performance of the filter was less affected by incorrect assumptions of σ_u^2 than by σ_v^2 . It is also evident that the errors tended to produce increased discrepancy between the correct posterior and the one obtained by using wrong assumptions.

Finally, we consider inaccuracies in the coefficient β in (11). Figure 3 shows the results of the K-S statistic as a function of $r_\beta = \frac{\tilde{\beta}}{\beta}$. In this case, it is clear that the performance

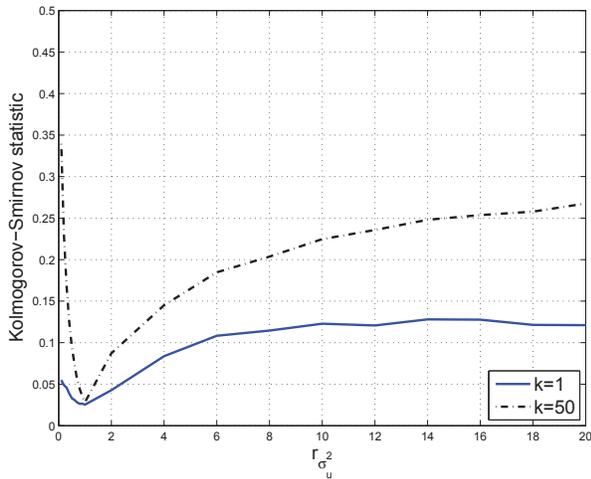


Fig. 2. K-S statistic as a function of $r_{\sigma_u^2}$.

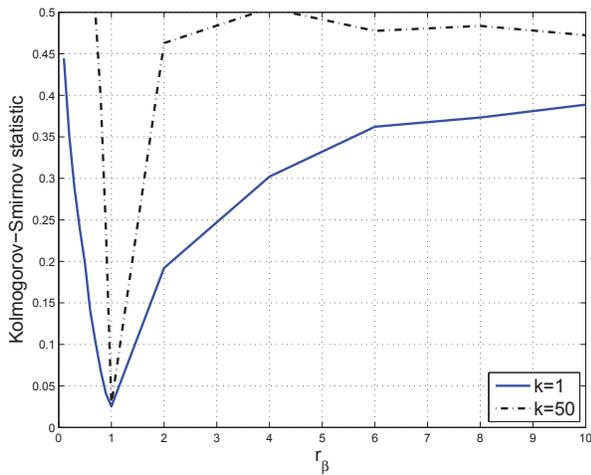


Fig. 3. K-S statistic as a function of r_{β} .

of the filter was the most sensitive to the misspecifications of the model parameters.

5. CONCLUSIONS

A qualitative Bayesian robustness was defined as a bounded change in the distribution of an estimator due to bounded changes in the distributions that describe a DSS model. Following this definition, the paper presents a general methodology for assessing the robustness of particle filtering algorithms that is based on the Kolmogorov-Smirnov statistic. A particular setup was considered in order to study the degradation in performance of particle filtering methods due to inaccurate DSS models. By using the Kolmogorov-Smirnov statistic, an assessment of the impact of different mismodeling errors was computed. The mismodeling was introduced

in the form of wrong parameters of the prior distributions and errors in the likelihood function. The proposed test can point to the parameters whose choice may be critical in the performance of the particle filter.

6. REFERENCES

- [1] P. M. Djurić, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. F. Bugallo, and J. Míguez, "Particle Filtering," *IEEE Signal Processing Mag.*, vol. 20, no. 5, pp. 19–38, September 2003.
- [2] A. Doucet, N. de Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*, Springer, 2001.
- [3] Peter J. Huber, *Robust Statistics*, John Wiley & Sons, 2 edition, 1981.
- [4] J. O. Berger, B. Betrò, E. Moreno, L. R. Pericchi, F. Ruggeri, G. Salinetti, and L. Wasserman, Eds., *Bayesian Robustness*, vol. 29 of *Lecture Notes—Monograph*, Institute of Mathematical Statistics, Hayward, CA, 1996.
- [5] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, vol. 64 of *Mathematics in Science and Engineering*, Academic Press, 1970.
- [6] F. R. Hampel, "A general definition of qualitative robustness," *The Annals of Mathematical Statistics*, vol. 42, pp. 1887–1896, 1971.
- [7] P. Bickel and K. Doksum, *Mathematical Statistics: Basic Ideas and Selected Topics*, vol. 1, Prentice Hall, 2nd edition, 2001.
- [8] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, New York, NY, USA, 2nd edition, 1992.
- [9] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, 4 edition, 2001.
- [10] G. Fasano and A. Franceschini, "A multidimensional version of the Kolmogorov-Smirnov test," *Monthly Notices Royal Astronomy Society*, vol. 225, pp. 155–170, 1987.
- [11] Ian T. Young, "Proof without prejudice: use of the Kolmogorov-Smirnov test for the analysis of histograms from flow systems and other sources," *Journal of Histochemistry and Cytochemistry*, vol. 25, no. 7, pp. 935–941, July 1977.
- [12] Frank J. Massey Jr., "The Kolmogorov-Smirnov Test for Goodness of Fit," *Journal of the American Statistical Association*, vol. 46, no. 253, pp. 68–78, March 1951.