When Does the Marginalized Particle Filter Degenerate?

Jakob Åslund, Fredrik Gustafsson and Gustaf Hendeby

When Does the Marginalized Particle Filter Degenerate?

Jakob Åslund
Department of Electrical Engineering
Linköping University, Sweden
jakob.aslund@liu.se

Fredrik Gustafsson
Department of Electrical Engineering
Linköping University, Sweden
fredrik.gustafsson@liu.se

Gustaf Hendeby
Department of Electrical Engineering
Linköping University, Sweden
gustaf.hendeby@liu.se

***Abstract*—The Particle filter can in theory estimate the state of any nonlinear system, but in practice it suffers from an exponential complexity in terms of the number of particles as the dimension of the state increases. The marginalized particle filter can potentially reduce this problem by improving the estimates, particularly for lower number of particles. However, it turns out that for certain systems, it does not provide any improvement in the accuracy of the estimate. The core cause of degeneracy is linked to when the uncertainty of the linear state conditioned on the nonlinear state is 0. Conditions for determining when this occurs are presented and applied to common constant velocity, constant acceleration and constant jerk models with various sampling methods. Interestingly, some combinations are useful while others should be avoided. These findings are supported using simulated systems.**

***Index Terms*—Marginalized particle filter, Rao-Blackwellized particle filter, Variance reduction, Particle filter**

I. Introduction

The Kalman filter (KF) is a well recognised method for estimating the state of a linear system. When the system contains some mild nonlinearities, the extended Kalman filter (EKF) has proven to be a good method. In order to estimate more complicated nonlinear systems, the particle filter (PF) is a common option [1], [2]. A drawback of the method however is the potentially large computational complexity needed to arrive at a good estimate. In practice, the number of required particles has been found to increase exponentially with the dimension of the state [2]. One attempt at reducing this problem is to marginalize some of the states, resulting in the marginalized particle filter (MPF), also known as the Rao-Blackwellized particle filter (RBPF) [2]–[4]. This is dependent on there existing a conditionally linear substructure in the system, meaning that the state being estimated can be split up as

$$x_k = \begin{pmatrix} x_k^l \\ x_k^n \end{pmatrix},$$

and its update can be described by

State update:

$$x_{k+1}^n = f_k^n(x_k^n) + F_k^n(x_k^n)x_k^l + w_k^n \tag{1a}$$
$$x_{k+1}^l = f_k^l(x_k^n) + F_k^l(x_k^n)x_k^l + w_k^l \tag{1b}$$

Funding from Swedish research council (Scalable Kalman Filters).

Measurement:

$$y_k = h_k(x_k^n) + H_k(x_k^n)x_k^l + e_k, \tag{1c}$$

where y_k is the observed measurement, w_k^n and w_k^l are process noises and e_k is the measurement noise, all assumed zero mean with covariances

$$\operatorname{Cov}\begin{pmatrix} w_k^n \\ w_k^l \end{pmatrix} = \begin{pmatrix} Q_{k,l} & Q_{k,nl} \\ Q_{k,ln} & Q_{k,n} \end{pmatrix}$$
$$\operatorname{Cov}(e_k) = R_k.$$

The dimensions of x_k^n, x_k^l and y_k are n_n, n_l and n_y respectively. Further f_k^n, f_k^l and h_k are arbitrary vector valued functions, and F_k^n, F_k^l and H_k are arbitrary matrix valued functions (all of appropriate dimensions to fit the equations).

Some research has been done regarding the accuracy improvement of the estimate when using an MPF instead of a plain PF. Much of the work has focused on providing an asymptotic estimate of how much the estimate is improved as the number of particles approaches infinity [5]–[8]. These results typically depend on the proposal of x_{k+1}^n being independent of all values $x_{0:k}^l$, meaning that $F_k^n(x_k^n) = \mathbf{0}$, which is not assumed here. In [9], empirical results for the RMSE of estimates are instead provided, combined with an analysis of the computational time of the filters.

However, in [10], we showed that for some common models, such as the constant velocity (CV) model sampled with a constant noise assumption, the MPF provides no improvement at all to the accuracy of the estimate, independent of the number of particles. This is linked to the fact that for those models the covariance $\operatorname{Cov}(x_k^l | x_{0:k}^n, y_{0:k}) \to \mathbf{0}$, as $k \to \infty$. In fact, it is possible to show that once $\operatorname{Cov}(x_k^l | x_{0:k}^n, y_{0:k}) = \mathbf{0}$ occurs, the MPF will behave identically to the PF, with the possible exception of better positioned particles from previous steps.

For this reason, this paper is focusing on determining what circumstances are needed for the MPF to not provide any improvements. Four conditions are presented which determine if the MPF will provide a better accuracy than the PF or not. For the sake of simplicity and ease of reading, these conditions will be formulated for the less general model class seen in (2), though the conditions can fairly easily be extended to work for all systems described by (1).

State update:

$$x_{k+1}^n = f_k^n(x_k^n) + F^n x_k^l + w_k^n \tag{2a}$$
$$x_{k+1}^l = f_k^l(x_k^n) + F^l x_k^l + w_k^l \tag{2b}$$

Measurement:

$$y_k = h_k(x_k^n) + e_k, \tag{2c}$$

where

$$\mathrm{Cov}\begin{pmatrix} w_k^n \\ w_k^l \end{pmatrix} = \begin{pmatrix} Q_n & Q_{nl} \\ Q_{ln} & Q_l \end{pmatrix}$$
$$\mathrm{Cov}(e_k) = R.$$

To get to this expression from (1), let $F_k^n(x_k^n) = F^n$, $F_k^l(x_k^n) = F^l$, $H_k(x_k^n) = 0$, and remove the time dependencies from the covariances $Q_{k,\bullet}$ and R_k.

The rest of this paper is organized as follows. First, the MPF will be introduced in Section II. Then, Section III presents conditions to determine if, for any given system, the MPF will provide an improvement or not. In Section IV, these conditions are applied to setups of constant velocity (CV), constant acceleration (CA) and constant jerk (CJ) models. Finally, in Section V some of these models are simulated to confirm the findings of the previous sections.

II. MPF Theory

The MPF can be viewed as a mixture of a PF and a Kalman filter (KF). The nonlinear states are estimated using a PF where the particles are propagated according to (1a) or (2a) and the weights are updated according to how well the observations fit their predictions in (1c) or (2c). The linear states are instead estimated using a KF, where the time update is provided by (1b) or (2b). The measurement update is provided by (1a) and (1c) or (2a) where the already propagated nonlinear states are viewed as virtual measurements.

The full algorithm for the system seen in (2) can be seen in Algorithm 1

For those interested in the exact expressions, a good description is provided in [3]. One issue with using this algorithm is that it assumes that the noises w_k^n and w_k^l in step 4) and 5) are uncorrelated. One way of solving this is to rewrite (2b) as

$$x_{k+1}^l = f_k^l(x_k^n) + Q_{ln}Q_n^{-1}(x_{k+1}^n - f_k^n(x_k^n)) + \bar{F}^l x_k^l + \bar{w}_k^l, \tag{3}$$

where

$$\mathrm{Cov}(\bar{w}_k^l) = \bar{Q}_l = Q_l - Q_{ln}Q_n^{-1}Q_{nl}$$
$$\bar{F}^l = F^l - Q_{ln}Q_n^{-1}F^n.$$

A full proof of this reformulation can be found in [3].

The root of the issues explored in this work can be found in this reformulation. Primarily, there is a risk that $\bar{Q}_l = \mathbf{0}$, which results in the equivalence to a system with no unknown process noise. In such a case, there is a risk that the KF uncertainty $P_{k|k-1}^{KF} \to \mathbf{0}$. This KF uncertainty is the same as $\mathrm{Cov}(x_k^l | x_{0:k}^n, y_{0:k})$ which, as discussed above, causes the MPF to provide no increase in accuracy compared to a plain PF when it approaches $\mathbf{0}$.

Note that this is not the same as stating that the estimate provided by the filter has no uncertainty in x_k^l. The uncertainty is only $\mathbf{0}$ conditioned on a trajectory $x_{0:k}^n$, which is not known. Instead, each particle will contain its own trajectory, meaning that for each particle there will be a separate estimate of x_k^l. The uncertainty is therefore represented by the spread of the mean term of the particles rather than the internal uncertainty of each particle.

Algorithm 1 The MPF

0 Initialize:

$$x_0^{n,(i)} \sim p(x_0^n)$$
$$x_{0|-1}^{l,(i)} = x_0^l$$
$$P_{0|-1}^{KF} = P_0^l$$

1 PF measurement update using (2c)
2 Resampling based on PF weights
3 PF time update using (2a)
4 KF measurement update using (2a) with $x_{k+1}^{n,(i)}$ as virtual measurements
5 KF time update using (2b)
6 repeat from 1

III. MPF degeneracy

In order to determine exactly when $P_{k|k-1}^{KF}$ will converge to $\mathbf{0}$, and consequently, the MPF will not provide any additional benefits, some insights into the behaviour of the uncertainty will be used to create conditions for when this occurs. Full derivations are lengthy and out of scope of this contribution, so we provide intuitive motivations and the core results as five rules of thumb, split into a primary condition and multiple secondary conditions.

1) If $\mathrm{rank}(Q_n) = \mathrm{rank}(Q)$ holds, all uncertainty present in the state update can be expressed in only w_k^n meaning that $\bar{Q}_l = \mathbf{0}$. If this does not occur $\bar{Q}_l \neq \mathbf{0}$ holds, and there is no possibility for $P_{k|k-1}^{KF} \to \mathbf{0}$. The most common way for this to occur is if Q is full rank.

$$\mathrm{rank}(Q_n) < \mathrm{rank}(Q) \implies P_{k|k-1}^{KF} \not\to \mathbf{0}. \tag{4}$$

This condition means that as long as $\mathrm{rank}(Q_n) < \mathrm{rank}(Q)$ holds, we know that implementing an MPF will result in some accuracy improvement, and do not need to do any further testing. This is good news since in most cases, condition 1) will hold, meaning that the general understanding that MPF will improve the accuracy of the estimate holds in most cases as well, and can be tested with a very simple test.

The following conditions will attempt to determine what happens when condition 1) does not hold. For these rules, $\mathrm{Rank}(Q_n) = \mathrm{Rank}(Q)$ will be assumed implicit.

2) If $\bar{Q}_l = \mathbf{0}$, all uncertainty of x_k^l comes from the initial uncertainty. If further, $|\lambda_i(\bar{F}^l)| \leq 1 \quad \forall i$, the

system defined by (3) is stable meaning that this initial uncertainty will approach $\mathbf{0}$ over time.

$$\left(|\lambda_i(\bar{F})| \leq 1 \quad \forall i\right) \implies P_{k|k-1}^{KF} \to \mathbf{0} \tag{5}$$

This requirement can be equivalently stated as $\bar{F}^l$ being stable.

3) In Algorithm 1 step 4, a measurement update is performed using the update of $x_{k+1}^{n,(i)}$. If the covariance of the noise in that measurement is low rank, there will be some perfect information available from the measurement. If enough such information is present (meaning that the covariance is of low enough rank), it will be possible to get a perfect estimate of $x_{k-1|k-1}^{l,(i)}$. If this is combined with $\bar{Q}_l = \mathbf{0}$, there will be no new information added meaning that $P_{k|k-1}^{KF} \to \mathbf{0}$.

$$\begin{aligned}&(n_n \geq n_l + \operatorname{rank}(Q_n)) \text{ and } \big(\text{the system with state and}\\ &\text{measurement updates (2a) and (2b) is detectable}\big)\\ &\implies P_{k|k-1}^{KF} \to \mathbf{0}.\end{aligned} \tag{6}$$

4) If the rewritten system is unstable, and there are no perfect measurements, the uncertainty caused by the initial uncertainty will increase during the time update, and then decrease during the measurement update, resulting in some non zero steady state.

$$\begin{aligned}&\big(\operatorname{rank}(Q_n) = n_n\big) \text{ and } \big(\exists i \text{ s.t. } |\lambda_i(\bar{F}^l)| > 1\big)\\ &\implies P_{k|k-1}^{KF} \not\to \mathbf{0}\end{aligned} \tag{7}$$

5) By a combination of the reasoning for 3) and 4), if the provided measurement is low rank, but not low rank enough to get a perfect estimate of the whole state, while the system also contains unstable elements, the outcome can not be determined by these conditions.

$$\begin{aligned}&(\operatorname{rank}(Q_n) < n_n) \text{ and } (n_n < n_l + \operatorname{rank}(Q_n)) \text{ and}\\ &\big(\exists i \text{ s.t. } |\lambda_i(\bar{F}^l)| > 1\big) \implies \text{The outcome is undetermined}\end{aligned} \tag{8}$$

As mentioned above, the MPF is not expected to result in any improvement in accuracy when $P_{k|k-1}^{KF} = \mathbf{0}$, meaning that these conditions also tell whether implementing the MPF will improve the estimate or not.

Worth noting is the special case when the user is only interested in estimating the state during a specific time frame during which $P_{k|k-1}^{KF} \neq \mathbf{0}$ despite the uncertainty approaching $\mathbf{0}$ for large enough k. In this special case, an improvement will be had from the MPF despite the conditions presented here implying otherwise. This is primarily of interest when $|\lambda_i(\bar{F})| = 1$ which may cause $P_{k|k-1}^{KF}$ to approach $\mathbf{0}$ very slowly.

IV. Example applications

In this section we will apply the conditions presented above to CV, CA and CJ models derived using different assumptions of the noise when sampling. This is an area where MPF and PF solutions have commonly been applied [2], [9], [11]. Of note here, is that while the dynamics of the system are linear, the measurements can be highly nonlinear, thus motivating the particle based approach.

For the sake of simplicity, all models presented here will be one-dimensional, though the results are easy to generalize. These models have the update matrices

$$F_{CV} = \begin{pmatrix} 1 & T \\ 0 & 1 \end{pmatrix} \tag{9a}$$

$$F_{CA} = \begin{pmatrix} 1 & T & T^2/2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{pmatrix} \tag{9b}$$

$$F_{CJ} = \begin{pmatrix} 1 & T & T^2/2 & T^3/6 \\ 0 & 1 & T & T^2/2 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{9c}$$

respectively, where T is the sampling time. As a further delimitation, it has been decided that the nonlinear state x_k^n contains the position and its $n_n - 1$ first derivatives.

The state transition matrix can then be partitioned as

$$F = \left[\begin{array}{c:c} A^n & F^n \\ \hdashline 0 & \bar{F}^l \end{array}\right],$$

for all cases, where $A^n \in \mathbb{R}^{n_n \times n_n}, A^l \in \mathbb{R}^{n_l \times n_n}, F^n \in \mathbb{R}^{n_n \times n_l}, F^l \in \mathbb{R}^{n_l \times n_l}$. Then the state update, can be written on the form of (2a) and (2b) as

$$x_{k+1}^n = \underbrace{A^n x_k^n}_{f_k^n(x_k^n)} + F^n x_k^l + w_k^n \tag{10a}$$

$$x_{k+1}^l = \underbrace{A^l x_k^n}_{f_k^l(x_k^n)} + F^l x_k^l + w_k^l. \tag{10b}$$

As can be seen in the conditions, the covariance matrix also plays an important role in how $P_{k|k-1}^{KF}$ behaves. This changes based on which sampling method is used for discretizing continuous-time integrators. Therefore, a number of sampling methods and their outcomes are collected in this contribution. For ease of reading, the actual covariance matrices have been moved to the Appendix.

A. Constant noise

The covariance matrices for the models when derived with a constant noise sampling method (Zero order hold, ZOH) are presented in Appendix A

Interestingly, for all cases it holds that no matter the selection of x_n and x_l, $\operatorname{rank}(Q_n) = \operatorname{rank}(Q) = 1$, not fulfilling condition 1). As such, conditions 2), 3), 4) and 5) have to be considered. In Tab. I, $\lambda(\bar{F}^l)$ is presented for different configurations of linear and nonlinear states for the models above. The expected outcome of each situation based on the conditions presented above is also shown in the table.

B. Continuous noise

The covariance matrices obtained if continuous noise is instead used to sample the model are presented in Appendix B.

TABLE I: Eigenvalues of $\bar{F}$ and resulting expected outcomes for different configurations (ZOH)

Nonlinear states	$\begin{pmatrix} p \end{pmatrix}$	$\begin{pmatrix} p & v \end{pmatrix}$	$\begin{pmatrix} p & v & a \end{pmatrix}$
CV	-1	-	-
	$P^{KF}_{k\|k-1} \to \mathbf{0}$ by condition 2)	-	-
CA	$-2 \pm \sqrt{3}$	not necessary	-
	$P^{KF}_{k\|k-1} \not\to \mathbf{0}$ by condition 4)	$P^{KF}_{k\|k-1} \to \mathbf{0}$ by condition 3)	-
CJ	$1,\ -5 \pm 2\sqrt{6}$	values range from $\begin{pmatrix} -0.26 & -3.7 \end{pmatrix}$ to $\begin{pmatrix} -4.7 & -48 \end{pmatrix}$ as a function of T	not necessary
	$P^{KF}_{k\|k-1} \not\to \mathbf{0}$ by condition 4)	undetermined, by condition 5)	$P^{KF}_{k\|k-1} \to \mathbf{0}$ by condition 3)

Here, all the covariance matrices are full rank, meaning that according to condition 1), $P^{KF}_{k|k-1} \not\to \mathbf{0}$. As such, the MPF is expected to improve the estimate for all selections of linear and nonlinear states.

C. Pulse-Noise

Another possible sampling method is to assume that the noise enters the system as a pulse, either at the start of the sampling interval or at the end. If the noise is assumed to enter at the end of the interval, the covariances are as given in Appendix C. For all of these matrices, $\mathrm{rank}(Q_n) = 0$ and $\mathrm{rank}(Q) = 1$, meaning that according to condition 1) $P^{KF}_{k|k-1} \not\to \mathbf{0}$.

If the noise is instead assumed to enter as a pulse at the start of the interval, the noise covariances are presented in Appendix D Here $\mathrm{rank}(Q_n) = \mathrm{rank}(Q)$ will always hold meaning that the other conditions once again have to be examined. This analysis is presented in Tab. II.

V. Simulations

Here, some of the different setups presented above will be simulated. The improvement in accuracy, or lack thereof of the MPF over the plain PF will be compared to the outcome predicted in the previous section. The primary indicator of the accuracy provided by the filters will be the mean squared error MSE of the position. To keep the tests simple, and to allow for an easy comparison with the optimal estimate, the position is measured directly $h_k(x^n_k) = p$, with Gaussian measurement noise, with covariance R. For each of the simulations, a 100 step trajectory was simulated and then estimated by the filters. This was repeated 100 times for each CV setup, 1000 times for the CA setups, and 100 times for the CJ setup. For each number of particles, the states and measurements were simulated anew, which is the cause for the slight variations which can be seen in the KF for different numbers of particles.

The MSE is only counted for step 20 and onward as this is deemed to be where $P^{KF}_{k|k-1}$ has approximately converged to its stationary value meaning that if $P^{KF}_{k|k-1} \to \mathbf{0}$ occurs, the effects of it should be noticeable by this point.

The performance of the MPF versus the PF on CV models was studied in [10], and is presented here again to validate that the more general conditions formulated in this contribution give the same result. The CA and CJ models are significantly more difficult for both the PF and MPF to estimate. For this reason, a method for detecting and removing failed filtering attempts has been implemented. First of all, every simulation where the particles got completely depleted is considered failed. Secondly, outlier detection was implemented to detect other runs that had abnormally high errors. To provide complete information, the number of failed filtering attempts will also be provided in the plots. The parameters used for the simulations can be seen in Tab. III.

TABLE III: Parameter settings for simulations

Model	**CV**	**CA & CJ**
T	1	1
R	1	10
x_0	$\mathbf{0}$	$\mathbf{0}$
P_0	$10\mathbf{I}$	$10^{-3}\mathbf{I}$
σ^2 (see Appendix)	1	0.01

The results for the CV models can be seen in Fig. 1. Notably, for the continuous noise and pulse noise at the end of the time step, in Figs. 1c and 1d, the MPF results in an improved MSE compared to the PF. This corresponds well with the results presented in Section IV. In comparison, when the noise is either zero order hold or a pulse at the start of the sampling interval, there appears to be no meaningful improvement by the MPF. Looking at Tabs. I and II, $P^{KF}_{k|k-1} \to \mathbf{0}$ is expected for both of these models, meaning that the outcome seen here agrees with the previous results.

The tested CA models can be seen in Fig. 3.

For the constant noise model with $x_n = p$, the model using pulse noise at the end of the time step and the continuous noise model, a clear improvement can be seen from the MPF in Figs. 3a, 3c and 3d, both in fewer failed attempts and in a lower MSE. According to the previous section, all of these models should have such an improvement. The constant noise case with $x_n = \begin{pmatrix} p & v \end{pmatrix}^T$, seen in Fig. 3b, seems to have the least clear improvement from the MPF. Out of the simulated models, it is also the one which is not expected to have an improvement from implementing the MPF according to the conditions provided.

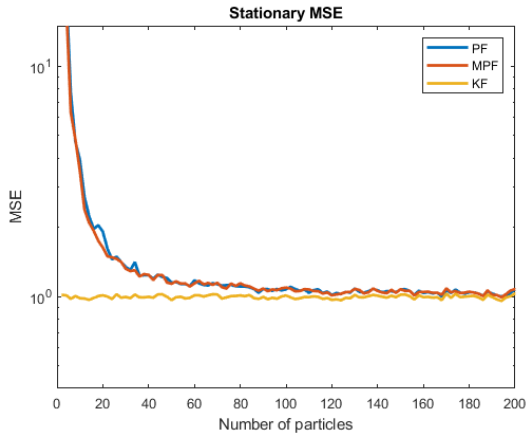
Lastly, one CJ model was implemented, with constant noise and $x_n = p$. Again, a clear improvement can be seen from the MPF, which agrees with the expected results based on the conditions.

VI. Conclusion

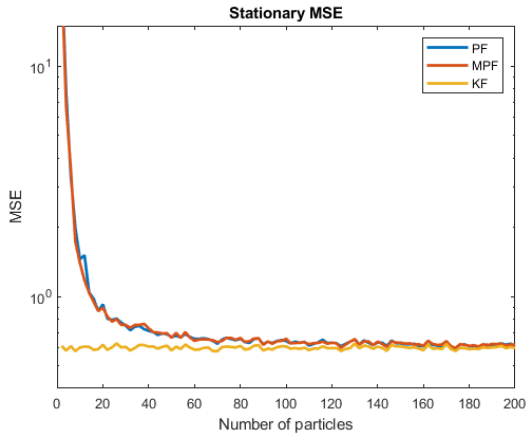
The standard motivation for implementing an MPF is that it will lead to more accurate estimates for the same number of particles than a standard PF would. If this is the case, an MPF should not be implemented if no such improvements can

TABLE II: Eigenvalues of $\bar{F}$ and resulting expected outcomes for different configurations (Pulse-noise)

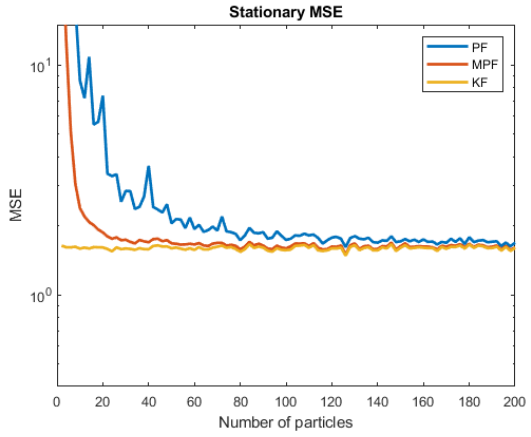
Nonlinear states	(p)	$(p \quad v)$	$(p \quad v \quad a)$
CV	0	-	-
	$P^{KF}_{k\|k-1} \to \mathbf{0}$ by condition 2)	-	-
CA	$0, -1$	not necessary	-
	$P^{KF}_{k\|k-1} \to \mathbf{0}$ by condition 2)	$P^{KF}_{k\|k-1} \to \mathbf{0}$ by condition 3)	-
CJ	$0, -2 \pm \sqrt{3}$	$-\frac{2T^2+9}{T^2+9}$	not necessary
	$P^{KF}_{k\|k-1} \not\to \mathbf{0}$ by condition 4)	undetermined, by condition 5)	$P^{KF}_{k\|k-1} \to \mathbf{0}$ by condition 3)



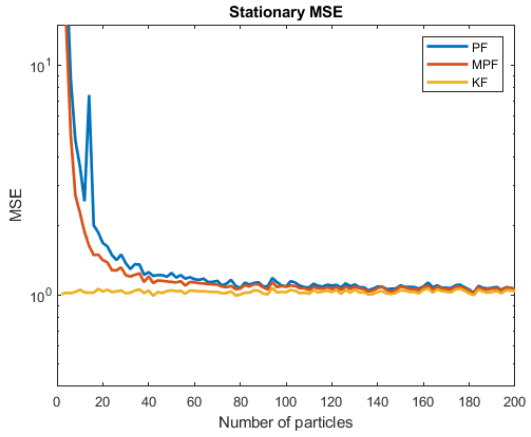
(a) CV model with ZOH



(b) CV model with Pulse noise at the start of the time step



(c) CV model with Pulse noise at the end of the time step



(d) CV model with Continuous noise

Fig. 1: The stationary MSE of the filters as a function of number of particles used, for CV models

be expected. In this contribution, a number of conditions have been provided to determine when said improvements occur. On a positive note for the MPF, we have shown that in any case where the process noise is full rank or where it has a higher rank than the process noise of the nonlinear part, the sought after improvement will be had. The MPF will therefore provide what was intended in most cases.

The rules provided in this article are easy to check, meaning that the user can often determine before implementing the solution, if the work will provide the sought after improvement or not. It is also possible to understand why that is. For instance, would a different sampling time change the outcome? Would changing the noise assumptions have an effect? While the conditions as presented here only work for a subset of all models that can be estimated using an MPF, these conditions can fairly easily be expanded to include all such models. Further, these conditions have been applied to many setups of CA, CV and CJ models to show for which of these the MPF is expected to provide improvements in the accuracy of the estimates. Many of these have also been backed up by simulations which confirm the results.

REFERENCES

[1] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.

[2] F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forssell, J. Jansson, R. Karlsson, and P.-J. Nordlund, "Particle filters for positioning, navigation, and tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 425–437, Feb 2002.

[3] T. Schön, F. Gustafsson, and P.-J. Nordlund, "Marginalized particle filters for mixed linear/nonlinear state-space models," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2279–2289, jul 2005.

[4] G. Hendeby, R. Karlsson, and F. Gustafsson, "The Rao-Blackwellized particle filter: A filter bank implementation," *EURASIP J. Adv. Sig. Proc.*, vol. 2010, 01 2010.

[5] A. Doucet, N. de Freitas, K. P. Murphy, and S. Russell, "Rao-Blackwellised particle filtering for dynamic Bayesian networks," in *UAI '00: Proceedings of the 16th Conference in Uncertainty in Artificial Intelligence, Stanford University, Stanford, California, USA, June 30 - July 3, 2000*, C. Boutilier and M. Goldszmidt, Eds. Morgan Kaufmann, 2000, pp. 176–183.

[6] N. Chopin, "Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference," *The Annals of Statistics*, vol. 32, no. 6, pp. 2385–2411, 2004.

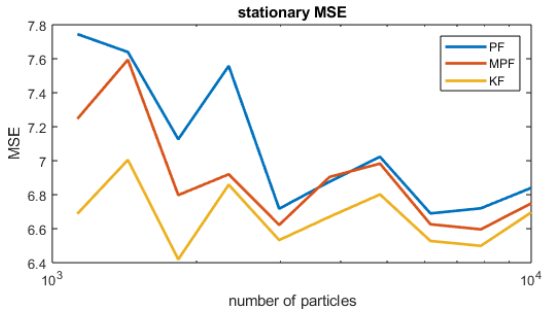
[7] P.-J. Nordlund, "Efficient estimation and detection methods for airborne applications," Ph.D. dissertation, Linköping University, Department of Electrical Engineering, Institute of Technology, 2009.

[8] F. Lindsten, T. B. Schön, and J. Olsson, "An explicit variance reduction expression for the Rao-Blackwellised particle filter," *IFAC Proceedings Volumes*, vol. 44, no. 1, pp. 11 979–11 984, 2011, 18th IFAC World Congress, Milan.

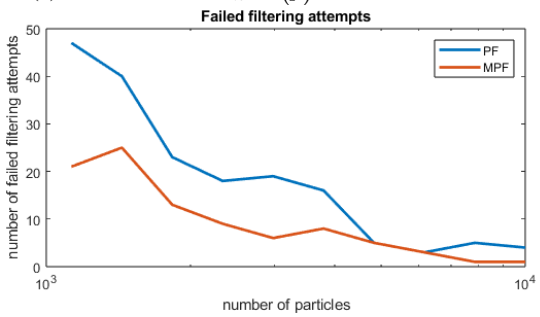
[9] R. Karlsson, T. Schön, and F. Gustafsson, "Complexity analysis of the marginalized particle filter," *IEEE Transactions on Signal Processing*, vol. 53, no. 11, pp. 4408–4411, 2005.

[10] J. Åslund, F. Gustafsson, and G. Hendeby, "On covariance matrix degeneration in marginalized particle filters with constant velocity models," in *25th International Conference on Information Fusion (FUSION),Linköping*, 2022, pp. 1–8.

[11] A. Giremus, J.-Y. Tourneret, and V. Calmettes, "A particle filtering approach for joint detection/estimation of multipath effects on gps measurements," *IEEE Transactions on Signal Processing*, vol. 55, no. 4, pp. 1275–1285, 2007.



(a) CJ model with $x_n = (p)^T$ and constant noise



(b) CJ model with $x_n = (p)^T$ and constant noise

Fig. 2: The stationary MSE and number of failed simulations of the filters as a function of number of particles used for CJ models

APPENDIX

The various covariance matrices are presented here. All of them are scaled by σ^2 which is the variance of the noise.

A. Constant noise

$$Q_{CV} = \sigma^2 \begin{pmatrix} T^4/4 & T^3/2 \\ T^3/2 & T^2 \end{pmatrix} \tag{11a}$$

$$Q_{CA} = \sigma^2 \begin{pmatrix} T^6/36 & T^5/12 & T^4/6 \\ T^5/12 & T^4/4 & T^3/2 \\ T^4/6 & T^3/2 & T^2 \end{pmatrix} \tag{11b}$$

$$Q_{CJ} = \sigma^2 \begin{pmatrix} T^8/576 & T^7/144 & T^6/48 & T^5/24 \\ T^7/144 & T^6/36 & T^5/12 & T^4/6 \\ T^6/48 & T^5/12 & T^4/4 & T^3/2 \\ T^5/24 & T^4/6 & T^3/2 & T^2 \end{pmatrix} \tag{11c}$$

B. Continuous noise

$$Q_{CV} = \sigma^2 \begin{pmatrix} T^3/3 & T^2/2 \\ T^2/2 & T \end{pmatrix} \tag{12a}$$

$$Q_{CA} = \sigma^2 \begin{pmatrix} T^5/20 & T^4/8 & T^3/6 \\ T^4/8 & T^3/3 & T^2/2 \\ T^3/6 & T^2/2 & T \end{pmatrix} \tag{12b}$$

$$Q_{CJ} = \sigma^2 \begin{pmatrix} T^7/252 & T^6/72 & T^5/30 & T^4/24 \\ T^6/72 & T^5/20 & T^4/8 & T^3/6 \\ T^5/30 & T^4/8 & T^3/3 & T^2/2 \\ T^4/24 & T^3/6 & T^2/2 & T \end{pmatrix} \tag{12c}$$

C. Pulse noise, end of time step

$$Q_{CV} = \begin{pmatrix} 0 & 0 \\ 0 & \sigma^2 \end{pmatrix} \tag{13a}$$

$$Q_{CA} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma^2 \end{pmatrix} \tag{13b}$$

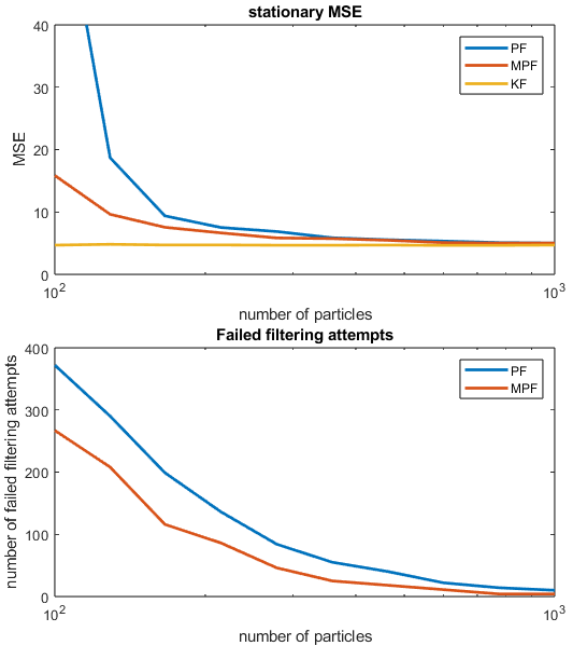
$$Q_{CJ} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{pmatrix} \tag{13c}$$

D. Pulse noise, start of time step

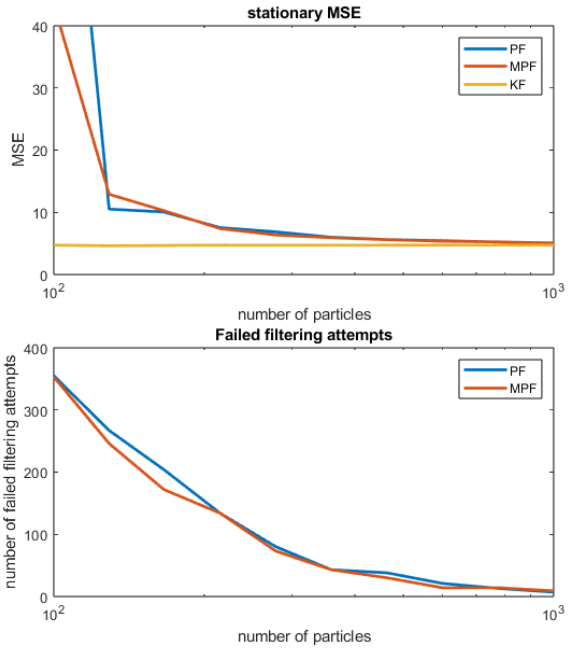
$$Q_{CV} = \sigma^2 \begin{pmatrix} T^2 & T \\ T & 1 \end{pmatrix} \tag{14a}$$

$$Q_{CA} = \sigma^2 \begin{pmatrix} T^4/4 & T^3/2 & T^2/2 \\ T^3/2 & T^2 & T \\ T^2/2 & T & 1 \end{pmatrix} \tag{14b}$$

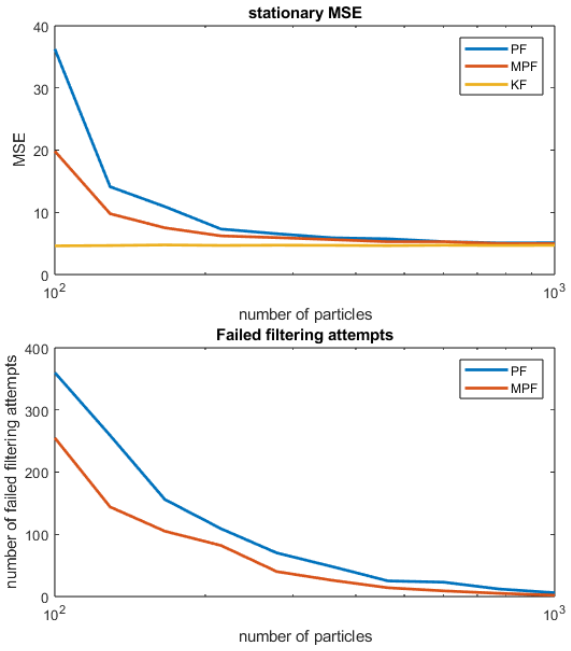
$$Q_{CJ} = \sigma^2 \begin{pmatrix} T^6/36 & T^5/12 & T^4/6 & T^3/6 \\ T^5/12 & T^4/4 & T^3/2 & T^2/2 \\ T^4/6 & T^3/2 & T^2 & T \\ T^3/2 & T^2 & T & 1 \end{pmatrix} \tag{14c}$$



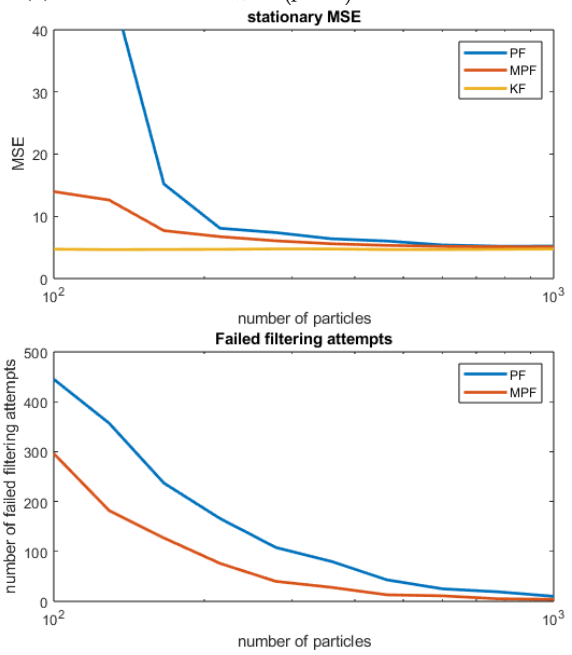
(a) CV model with $x_n = p$ and constant noise



(b) CV model with $x_n = \begin{pmatrix} p & v \end{pmatrix}^T$ and constant noise



(c) CV model with $x_n = p$ and continuous noise



(d) CV model with $x_n = p$ and pulse noise at the end of the time step

Fig. 3: The stationary MSE and failed number of simulations of the filters as a function of number of particles used for CA models