On Covariance Matrix Degeneration in Marginalized Particle Filters with Constant Velocity Models

Jakob Åslund, Fredrik Gustafsson and Gustaf Hendeby

LINKÖPING UNIVERSITY

On Covariance Matrix Degeneration in Marginalized Particle Filters with Constant Velocity Models

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*—Marginalization enables the particle filter to be applied to high-dimensional problems by invoking the Kalman filter to estimate a larger part of the state vector. The marginalized (a.k.a. Rao-Blackwellized) particle filter (MPF) has found many use cases in tracking and navigation applications. These are characterized by having position and its derivatives as states. Here, we take a closer look at the MPF for the constant velocity motion model, which well represents the basic properties of most motion models used in this context. In particular, how the Kalman filter (KF)-part depends on how the continuous time state noise is sampled in the discrete time model. We find that for many of the most common sampling approaches, the KF-part of the MPF degenerates, meaning that the covariance approaches 0. Further, we show that for those same sampling approaches there is no performance increase by switching from a particle filter to an MPF in this situation.**

***Index Terms*—Marginalized particle filter, Rao-Blackwellized particle filter, Discretization, Constant velocity model**

I. Introduction

The particle filter (PF) as proposed in [1] is known to solve all nonlinear non-Gaussian filtering problems, in theory. The catch is that the number of particles that are required grows quickly with state dimension. For instance, part of the error can grow exponentially with the dimension in a simple linear problem [2]. The main purpose of the marginalized particle filter (MPF) is to split the nonlinear filtering problem into two simpler ones, and in that way keep the state dimension low enough for the PF to be feasible. The resulting filter is the MPF, which has been known for quite some time under different names, marginalized particle filter [3], Rao-Blackwellized particle filter [4], mixture Kalman filter [5], or without a specific name at all [6], and is commonly applied to tracking problems, see [7]–[9]. While the general case is quite explored, more detailed results can be obtained for particular sub-classes of nonlinear filtering. Here, we focus on applications where the state consists of position and its derivative, which includes more or less all navigation and tracking applications. The principles of the MPF for navigation and tracking are described in [10], [11].

When the MPF is applied to a constant velocity (CV) model with a measurement that depends on the position only, we get an interesting special case of some fundamental importance. In this case, the particle filter focuses on estimating the position, and each particle represents a trajectory. Seen from the KF point of view, the trajectory is given, and the task is to estimate the velocity vector from given positions.

The MPF applied to CV-models has previously been studied in [12] with regard to computational complexity. We instead focus on the uncertainty of the velocity and provide new theory in the form of exact calculations of the covariance and Kalman gain of the CV model for four discretization approaches. One would expect that the KF would keep some kind of uncertainty of the velocity. However, we show that, depending on the covariance of the noise to the CV model, the KF for each particle may degenerate and either directly or quickly converge towards zero uncertainty. In these cases, it is only the spread of the various particles that contributes to the covariance of the velocity.

This paper will first introduce the background theory of the MPF with specific focus on the simplified form which is used in this paper. We will then show how this form can be applied to a CV model in Sec. III. After this, the KF-part will be analysed for four different approaches to state noise discretization, with calculations of the stationary covariance and Kalman gain in Sec. IV. Lastly, in Sec. V some illustrative simulations are shown together with an analysis of the results from both the simulations as well as the theory.

II. The Marginalized Particle Filter

The main idea of the MPF is to only estimate a part of the state x_k using a particle filter, and estimate the rest of the state using a linear KF conditioned on the part of the state estimated by the PF. The goal is to reduce the number of particles needed by reducing the dimension of the states estimated by the PF. The part estimated by the PF is referred to as the nonlinear state, and the part estimated by the KF is referred to as the linear state. These are marked with the indices n and l respectively

$$x_k = \begin{pmatrix} x_k^n \\ x_k^l \end{pmatrix}. \tag{1a}$$

The requirement for this to work is that the system contains a conditionally linear Gaussian sub-structure given by

Funding from Swedish research council (Scalable Kalman Filters).

$$x^n_{k+1} = f^n_k(x^n_k) + F^n_k(x^n_k)x^l_k + G^n_k(x^n_k)w^n_k, \tag{1b}$$

$$x^l_{k+1} = f^l_k(x^n_k) + F^l_k(x^n_k)x^l_k + G^l_k(x^n_k)w^l_k, \tag{1c}$$

$$y_k = h_k(x^n_k) + H_k(x^n_k)x^l_k + e_k, \tag{1d}$$

where f^n_k, f^l_k and h_k are functions, and $F^n_k(x^n_k)$, $F^l_k(x^n_k)$, $H_k(x^n_k)$, $G^n_k(x^n_k)$ and $G^l_k(x^n_k)$ are matrices of appropriate size. The vectors w^n_k, w^l_k and e_k are noise, where the distribution of e_k is known and

$$w_k = \begin{pmatrix} w^n_k \\ w^l_k \end{pmatrix} \sim \mathcal{N}(0, Q_k), \quad Q_k = \begin{pmatrix} Q^n_k & Q^{ln}_k \\ (Q^{ln}_k)^T & Q^l_k \end{pmatrix}. \tag{1e}$$

Furthermore, x^l_0 is assumed Gaussian, $x^l_0 \sim \mathcal{N}(\bar{x}_0, \bar{P}_0)$. The density of x^n_0 is assumed known. Conditioned on the sequence $x^n_1, \ldots, x^n_k$ written as $x^n_{1:k}$, this model structure is linear in x^l_k with Gaussian prior, process noise and measurement noise, respectively, so the KF theory applies.

The main idea with the MPF is to realize that for each particle sequence $\{x^{n,(i)}_{1:k}\}^N_{i=1}$, the model (1) is a linear state-space model where the KF applies. Here, it should be noted that (1b) can be re-written as a measurement model which is linear in x^l_k. Conversely, if the KF works, we get the mean and covariance of x^l_k, and the particle filter can be applied with extra state noise and measurement noise contributions from the KF estimation error. The details, in a notation similar to what is used here, are given in [10], [11], [13]. Note that, the KF and its estimate are dependent on the nonlinear state, meaning that it is necessary to apply a separate KF for each particle.

A. Simplified MPF Model

We will derive the explicit KF updates for a simplified model given by

$$x^n_{k+1} = f^n_k(x^n_k) + F^n x^l_k + G^n w^n_k, \tag{2a}$$

$$x^l_{k+1} = f^l_k(x^n_k) + F^l x^l_k + G^l w^l_k, \tag{2b}$$

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

where F^n, F^l, G^n and G^l are matrices of appropriate size. Many applications, including the ones described in [10], fit into this model structure. The main reasons for the simplifications is first to make each KF have the same covariance update for all particle sequences $\{x^{n,(i)}_{1:k}\}^N_{i=1}$, and second to make the KF time-invariant (removal of some of the time dependencies).

The standard model for a time-invariant KF can be written as

$$\bar{x}_{k+1} = F\bar{x}_k + G_u u_k + \bar{w}_k, \tag{3a}$$

$$\bar{y}_k = H\bar{x}_k + \bar{e}_k, \tag{3b}$$

where u is a known input and G_u is a matrix. The model (2), conditional on $\{x^{n,(i)}_{1:k+1}\}^N_{i=1}$, fits this model if rewritten as

$$\underbrace{x^l_{k+1}}_{\bar{x}_{k+1}} = \underbrace{F^l}_{F} \underbrace{x^l_k}_{\bar{x}_k} + \underbrace{f^l_k(x^n_k)}_{G_u u_k} + \underbrace{G^l w^l_k}_{\bar{w}_k}, \tag{4a}$$

$$\underbrace{x^n_{k+1} - f^n_k(x^n_k)}_{\bar{y}_k} = \underbrace{F^n}_{H} \underbrace{x^l_k}_{\bar{x}_k} + \underbrace{G^n w^n_k}_{\bar{e}_k}. \tag{4b}$$

We focus on the KF part of the MPF, where as has already been noted, we can consider the sequences $\{x^{n,(i)}_{1:k+1}\}^N_{i=1}$ as given. For each particle, the KF estimates $\hat{v}^{l,(i)}_{k|k}$ with covariance $P^{l,(i)}_{k|k}$. These are then combined to form the overall estimates

$$\hat{v}_{k|k} = \sum_{i=1}^N \omega^{(i)}_k \hat{v}^{(i)}_{k|k}, \tag{5a}$$

$$\mathrm{Cov}(\hat{v}_{k|k}) = \sum_{i=1}^N \omega^{(i)}_k \Big(\big(\hat{v}^{(i)}_{k|k} - \hat{v}_{k|k}\big)\big(\hat{v}^{(i)}_{k|k} - \hat{v}_{k|k}\big)^T + P^{l,(i)}_{k|k} \Big), \tag{5b}$$

where $\omega^{(i)}_k$ is the probability (weight) for sequence $x^{(i)}_{1:k}$. All quantities in the KF derivations should be conditioned on a particular sequence, but to avoid unnecessary notation this will be implicit. All KFs will obey the same Riccati equation, and thus have the same Kalman gain and covariance matrix. Thus, there is no need to add a particle index i to these quantities, only to the individual state estimate. With this knowledge, (5b) can be simplified as

$$\mathrm{Cov}(\hat{v}_{k|k}) = \underbrace{\sum_{i=1}^N \omega^{(\mathrm{i})}_k \big(\hat{v}^{(i)}_{k|k} - \hat{v}_{k|k}\big)\big(\hat{v}^{(i)}_{k|k} - \hat{v}_{k|k}\big)^T}_{P_{\mathrm{SM}}} + \underbrace{P^l_{k|k}}_{P_{\mathrm{KF}}}, \tag{6}$$

where P_{KF} is the covariance calculated by the internal KF, and P_{SM} is the uncertainty based on the spread of the mean, calculated by the PF-part of the filter.

III. Constant Velocity Model

The constant velocity (CV) model in n dimensions with position p and velocity v is derived from the continuous time model

$$\dot{p}_t = v_t + w^c_t, \tag{7}$$

where w^c_t represents, with a little abuse of notation, continuous time noise. For the state $x^T_k = ((x^{(n)}_k)^T, (x^{(l)}_k)^T) = (p^T_k, v^T_k)$, its discretized version becomes

$$x_{k+1} = \begin{pmatrix} I_n & TI_n \\ 0 & I_n \end{pmatrix} x_k + \begin{pmatrix} w^{(p)}_k \\ w^{(v)}_k \end{pmatrix}, \tag{8a}$$

$$y_k = h(p_k) + e_k, \tag{8b}$$

where T is the step time of the discretization. The function to obtain the observation y_k from the position, may contain any kind of nonlinearity, meaning that it is not possible to apply a standard KF to this problem. The MPF applied to this model results in a PF that treats (8b) as the measurement and the first line of (8a) as the dynamics for position. The KF, on the other hand, does not use (8b) at all, and interprets the first line of (8a) as the measurement, and the second line of (8a) as the dynamics. Reformulated as in (4), the model (8) becomes

$$v_{k+1} = v_k + w^{(v)}_k, \tag{9a}$$

$$\bar{y}_k = p_{k+1} - p_k = Tv_k + w^{(p)}_k. \tag{9b}$$

The KF solution to (9) is straightforwardly derived. The main complication is that the time and measurement noise are correlated, so the corresponding updates cannot be separated as they usually are. The noise covariance can be partitioned as

$$Q = \mathrm{Cov}\big(w_k^p,\ w_k^v\big) = \begin{pmatrix} Q^p & (Q^{vp})^T \\ Q^{vp} & Q^v \end{pmatrix}, \tag{10}$$

which is symmetrical since it is a covariance matrix.

We will first derive the KF for the model (9) for a general Q. To simplify the calculations later on we will first rewrite (9b) as

$$z_{k+1}^{(i)} = \frac{\bar{y}_k^{(i)}}{T} = \frac{p_{k+1}^{(i)} - p_k^{(i)}}{T} = v_k^{(i)} + \frac{w_k^{(p)}}{T}. \tag{11}$$

This equation will from here on be used as the measurement equation for the KF. It can be seen as measuring the average velocity of the particle during the time step. The correlation between the updated state and the measurement can be written as

$$\mathrm{Cov}\begin{pmatrix} v_{k+1}^{(i)} \\ z_{k+1}^{(i)} \end{pmatrix} = \begin{pmatrix} P_{k|k}^l + Q^v & P_{k|k}^l + Q^{vp}/T \\ P_{k|k}^l + Q^{pv}/T & P_{k|k}^l + Q^p/T^2 \end{pmatrix}. \tag{12}$$

The covariance is now used to estimate the expectation and covariance of the state at the following time step. Since this is a conditional Gaussian distribution, this is done according to the equations

$$\hat{v}_{k+1|k+1}^{(i)} = \hat{v}_{k|k}^{(i)} + P_{vz}P_{zz}^{-1}(z_{k+1}^{(i)} - \hat{v}_{k|k}^{(i)}), \tag{13a}$$
$$P_{k+1|k+1}^l = P_{vv} - P_{vz}P_{zz}^{-1}P_{zv}, \tag{13b}$$

where

$$P_{vv} = P_{k|k}^l + Q^v, \tag{14a}$$
$$P_{vz} = P_{k|k}^l + Q^{vp}/T = P_{zv}^T, \tag{14b}$$
$$P_{zz} = P_{k|k}^l + Q^p/T^2. \tag{14c}$$

Finally, we note that (13) has the same form as a state update with a Kalman gain K_k according to

$$K_k = P_{vz}P_{zz}^{-1}. \tag{15}$$

IV. KF FOR SPECIFIC NOISE SAMPLINGS

Here, the covariance and Kalman gain for the various discretized state noises Q surveyed in the Appendix will be explored in detail, one by one.

A. Zero order hold

Using zero order hold (ZOH) sampling it is assumed that the noise is constant within each time step. For a CV-model this means that we have a constant but unknown acceleration for each time step. As shown in App. A, the covariance of the noise when arriving at (8) using a ZOH discretization is

$$Q = \begin{pmatrix} qT^4/4 & qT^3/2 \\ qT^3/2 & qT^2 \end{pmatrix}, \tag{16}$$

where q is the covariance matrix of the continuous noise. It is worth to note that for this paper, q could be either a scalar, or an $n \times n$ matrix, depending on the dimensions of the observed system. Unless otherwise stated, all calculations made here work for any covariance q. In order to simplify the notation, we will here use $P_{k|k}^l = P_k$ and $qT^2/4 = M$. With this notation, (14) becomes

$$P_{vv} = P_k + 4M, \tag{17a}$$
$$P_{vz} = P_k + 2M, = P_{zv} \tag{17b}$$
$$P_{zz} = P_k + M. \tag{17c}$$

Combining (17) and (13), gives the covariance update

$$\begin{aligned} P_{k+1} &= P_k + 4M - (P_k + 2M)(P_k + M)^{-1}(P_k + 2M) \\ &= P_k + 4M - (P_k + 3M + M(P_k + M)^{-1}M) \\ &= (P_k^{-1} + M^{-1})^{-1}. \end{aligned} \tag{18}$$

From (18), we see that we can write the update on the innovation form $\mathcal{I}_k = P_k^{-1} = (P_{k|k}^l)^{-1}$ as

$$\mathcal{I}_{k+1} = \mathcal{I}_k + M^{-1}. \tag{19}$$

We can also write $\mathcal{I}_k$ as

$$\mathcal{I}_k = \mathcal{I}_0 + kM^{-1}. \tag{20}$$

Using the Woodbury matrix inversion [14], we get

$$P_k = (\mathcal{I}_0 + kM^{-1})^{-1} = \frac{1}{k}M - \frac{1}{k}M(P_0 + \frac{1}{k}M)^{-1}\frac{1}{k}M, \tag{21}$$

which clearly approaches 0 when k increases. Therefore, the stationary covariance must be 0. Finally, we look at the stationary Kalman gain K. According to (15) and (17) we get

$$\begin{aligned} K_k &= (P_k + 2M)(P_k + M)^{-1} = \\ &= I_n + M(P_k + M)^{-1} \to 2I_n, \text{as } k \to \infty. \end{aligned} \tag{22}$$

A possible way of testing how reasonable this is, is to consider the scalar case where we know the position of an object at the start and end of a time step x_k and x_{k+1}. We know that the acceleration a_k (which is the noise in this case) is constant and we want to find the new velocity v_{k+1}. We can now calculate this using simple physics equations.

$$\begin{cases} x_{k+1} = x_k + Tv_k + \frac{T^2}{2}a_k, \\ v_{k+1} = v_k + Ta_k, \end{cases} \tag{23a}$$

which has the solution

$$\begin{cases} Ta_k = 2\frac{x_{k+1} - x_k}{T} - 2v_k, \\ v_{k+1} = v_k + 2(\frac{x_{k+1} - x_k}{T} - v_k). \end{cases} \tag{23b}$$

If we now let $\frac{x_{k+1} - x_k}{T} = z_{k+1}$ we easily see that we here as well end up with a Kalman gain of $2I_n$.

B. Impulse at the Start of Time Step

In this section, it is assumed that the noise arrives as an impulse at the start of each time step, meaning that the velocity is assumed to shift immediately after the previous measurement has been received. This gives the noise covariance

$$Q = \begin{pmatrix} T^2 q & Tq \\ Tq & q \end{pmatrix}, \tag{24}$$

as shown in App. B. The covariance update can then be calculated from (13b) as

$$\begin{aligned} P^l_{k+1|k+1} &= (P^l_{k|k} + q) - (P^l_{k|k} + q)(P^l_{k|k} + q)^{-1}(P^l_{k|k} + q) \\ &= (P^l_{k|k} + q) - (P^l_{k|k} + q) \\ &= 0. \end{aligned} \tag{25}$$

Similarly, the Kalman gain from (15) becomes

$$K_k = (P^l_{k|k} + q)(P^l_{k|k} + q)^{-1} = I_n. \tag{26}$$

This result can be explained fairly easily. Since the velocity is only assumed to shift at the beginning of each time step, the object will be travelling at a constant velocity throughout the time step. Our measurement (11) will return the average velocity of the particle during the time step without any additional measurement error. Since the particle is however assumed to be travelling at a constant velocity during this time, the average velocity measured will be exactly the sought after velocity $v^{l,(i)}_{k+1}$. Therefore, using this measurement as the estimate will result in no uncertainty.

C. Impulse at the end of the time step

Here, as in Sec. IV-B, the noise is assumed to come as an impulse, but at the end of each time step instead, meaning that the velocity will shift just before each measurement. As shown in App. C, we then arrive at

$$Q = \begin{pmatrix} 0 & 0 \\ 0 & q \end{pmatrix}. \tag{27}$$

From this, the covariance update becomes

$$P^l_{k+1|k+1} = (P^l_{k|k} + q) - (P^l_{k|k} + 0)(P^l_{k|k} + 0)^{-1}(P^l_{k|k} + 0) \tag{28a}$$

$$= P^l_{k|k} + q - P^l_{k|k} = q, \tag{28b}$$

and the Kalman gain is calculated as

$$K_k = (P^l_{k|k} + 0)(P^l_{k|k} + 0)^{-1} = I_n. \tag{29}$$

The results here can be explained much in the same way as the results in Sec. IV-B. The only difference is that in this case, the noise causing the velocity to change, will occur at the end of each time step, meaning that the average velocity measured will be that of the previous time step. As such, we will always at least have the covariance q since we do not know anything about the most recent noise. At the same time, the exact velocity at the previous time step is the best available estimate of the current velocity, which is why the Kalman gain is I_n.

D. Continuous noise

Another fairly common noise assumption, is that the noise is continuous white noise. For this application this would mean that the acceleration at any time is independent of the acceleration at any other time. This noise and its discretization is examined in App. D and results in

$$Q = \begin{pmatrix} qT^3/3 & qT^2/2 \\ qT^2/2 & qT \end{pmatrix}. \tag{30}$$

Similarly to what is done in Sec. IV-A, we here make use of $P_k = P^l_{k|k}$ and $M = qT/6$ to simplify the notation. Note that, M here is slightly different from what was used before. With this notation, (14) becomes

$$P_{vv} = P_k + 6M, \tag{31a}$$

$$P_{vz} = P_k + 3M, \tag{31b}$$

$$P_{zz} = P_k + 2M. \tag{31c}$$

With the values from (31), the general covariance update (13b) becomes

$$\begin{aligned} P_{k+1} &= (P_k + 6M) - (P_k + 3M)(P_k + 2M)^{-1}(P_k + 3M) \\ &= P_k + 6M - (P_k + 3M + M + M(P_k + 2M)^{-1}M) \\ &= \frac{3}{2}M + \frac{1}{4}(P_k^{-1} + \frac{1}{2}M^{-1})^{-1}. \end{aligned} \tag{32}$$

We now look at what happens when this covariance reaches stationarity $P_{k+1} = P_k = P$. At that point, (32) becomes

$$\begin{aligned} &P = \frac{3}{2}M + \frac{1}{4}(P^{-1} + \frac{1}{2}M^{-1})^{-1} \Leftrightarrow \\ &P = 3MP^{-1}M. \end{aligned} \tag{33}$$

Unfortunately, this equation can not be solved analytically for general values of P and M. One special case which can be solved however, is the case where the covariance of the noise q and the initial noise $\bar{P}_0$ are diagonal matrices, which means that

$$M = qT/6 = \begin{pmatrix} q_1 T/6 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & q_n T/6 \end{pmatrix}. \tag{34}$$

A one-dimensional system, with a scalar q is a special case of this. In this case, it follows from (32) that P_k must also be diagonal, meaning that the stationary P must be diagonal

$$P = \begin{pmatrix} p_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & p_n \end{pmatrix}. \tag{35}$$

With this assumption, all matrices in (33) become diagonal, meaning we can deal with each direction separately. This leads to

$$p_i = 3m_i^2/p_i \Leftrightarrow p_i = \sqrt{3}m_i = \frac{T}{\sqrt{12}}q_i \Leftrightarrow \tag{36}$$

$$P = \frac{T}{\sqrt{12}}q. \tag{37}$$

Finally, the Kalman gain at stationarity is also examined. Since all matrices from (15) are assumed to be diagonal, each direction can again be dealt with separately.

$$K = \begin{pmatrix} k_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & k_n \end{pmatrix}, \tag{38a}$$

$$\begin{aligned} k_i &= \frac{p_i + 3m_i}{p_i + 2m_i} \\ &= \frac{\frac{T}{\sqrt{12}}q_i + \frac{T}{2}q_i}{\frac{T}{\sqrt{12}}q_i + \frac{T}{3}q_i} \\ &= 3 - \sqrt{3}. \end{aligned} \tag{38b}$$

E. Summary

We here summarize the different stationary covariances as well as the Kalman gain for each of the discretization methods in Tab. I. Note that, all methods result in diagonal stationary Kalman gains, and multiple methods end up with diagonal elements that are larger than 1. On an intuitive level, this seems wrong. For instance, it would mean that if an element of the state was initially estimated to be 0 and it was later measured as 2, the correct posterior estimate would be larger than 2. However, this is a result of the correlation of the noise. Since what has been measured is the average velocity, and the initial velocity was believed to be smaller than 2, the final velocity is expected to be larger than 2 in order to arrive at the correct average.

TABLE I: Stationary covariance and Kalman gain for each discretization method.

Discretization method	**Stationary covariance**	**Stationary Kalman gain**
ZOH	0	$2I_n$
Noise as an impulse at the start of the time step	0	I_n
Noise as an impulse at the end of the time step	q	I_n
Continuous noise (only for diagonal noise matrix q)	$\frac{T}{\sqrt{12}}q$	$(3-\sqrt{3})I_n$

V. Numerical Illustrations

In this section we provide some simulations to illustrate the results of the previous sections. To simplify the simulations, the position is here assumed to be scalar. Further, we make the measurement from (8b) linear of the form

$$h_k(p_k) = p_k. \tag{39}$$

The main advantage of using a linear measurement is that a KF can now be applied to the system to get the optimal estimates as a benchmark. The measurement noise was considered Gaussian,

$$e_k \sim \mathcal{N}(0, 1). \tag{40}$$

Finally, for simplicity, both the process noise q and the time step T are 1. The initial position x_0^n and the velocity x_0^l are given by

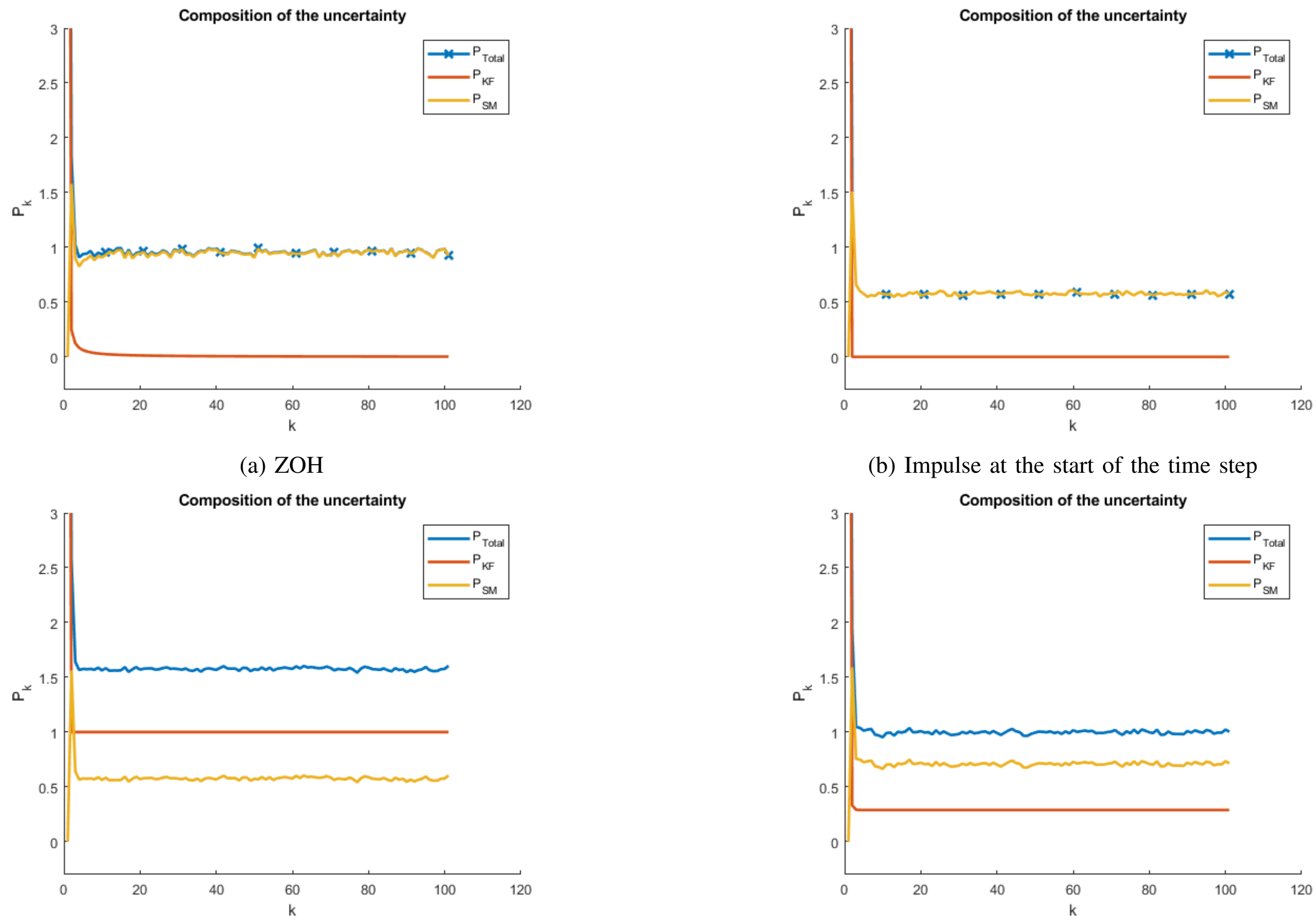
$$\begin{pmatrix} x_0^n \\ x_0^l \end{pmatrix} \sim \mathcal{N}(\mathbf{0}, P_0), \quad P_0 = \begin{pmatrix} 10 & 0 \\ 0 & 10 \end{pmatrix}. \tag{41}$$

Here, two simulations are shown. In the first one, seen in Fig. 1, the MPF was applied to 100 simulated trajectories of the system described above, and the average uncertainty per time-step is shown for the different components of the filter. For each filter, 100 particles were used. In the second simulation, seen in Fig. 2, the performance of the filter as a function of the number of particles was evaluated. The system was simulated and estimated 100 times for each number of particles. For each simulation, the mean squared error (MSE) where the filter had reached a steady error was recorded. Based on the results in Fig. 1, this was done by looking at the MSE for time steps from 20 to 100. The mean of the errors for the simulations with the same number of particles are given. As a comparison, the performance of the KF and PF is also included. The PF used the same number of particles as the MPF.

In Fig. 1, we can see how both P_{KF} and P_{SM} behave as a function of the iteration of the filter. When using the ZOH and noise as an impulse at the end of each time step discretization methods, P_{KF} approaches 0. When instead assuming continuous noise or noise as an impulse at the end of each time step, P_{KF} approaches a value larger than 0, which confirms the theory in Sec. IV for this particular problem. It can be further seen that, when P_{KF} approaches 0 all uncertainty in the system must be expressed by P_{SM}, which is not true in the case when P_{KF} does not approach 0.

One result of this can be seen in Fig. 2, where the performance of the MPF is studied as a function of the number of particles. We note that, both the ZOH and the noise as an impulse at the start of the time step discretizations lead to the PF and the MPF performing similarly for all numbers of particles. In contrast, both the noise as an impulse at the end of the time step and continuous noise discretizations lead to the MPF having a noticeably lower MSE than the PF for fewer particles. This difference in performance is significantly larger when the noise is assumed to come as an impulse at the end of the time step. These results can be connected to the results shown in Fig. 1.

According to (6), the total covariance for the MPF estimate of the linearized state can be described as the sum of two parts P_{KF} and P_{SM}. Since there exists a minimum covariance of $\hat{v}_{k|k}$ (provided by the standard KF if the system is linear), if P_{KF} is small, P_{SM} must become larger. This is equivalent to the PF-part of the filter having to deal with a larger part of the total uncertainty of the system, which will require more particles, and therefore, more computational power. What happens in (a) and (b) is the extreme case, where $P_{\text{KF}} = 0$, meaning that all uncertainty will be modelled by the spread of the particles, which means that nothing is gained from the marginalization. In comparison, (c) and (d) both have some part of the uncertainty represented by P_{KF}, meaning that less of it has to be represented by the spread of the particles. As such, a lower number of particles is needed



(a) ZOH

(b) Impulse at the start of the time step

(c) Impulse at the end of the time step

(d) Continuous noise

Fig. 1: The composition of the covariance between the KF and the spread of the particles. For all these simulations 100 particles were used.

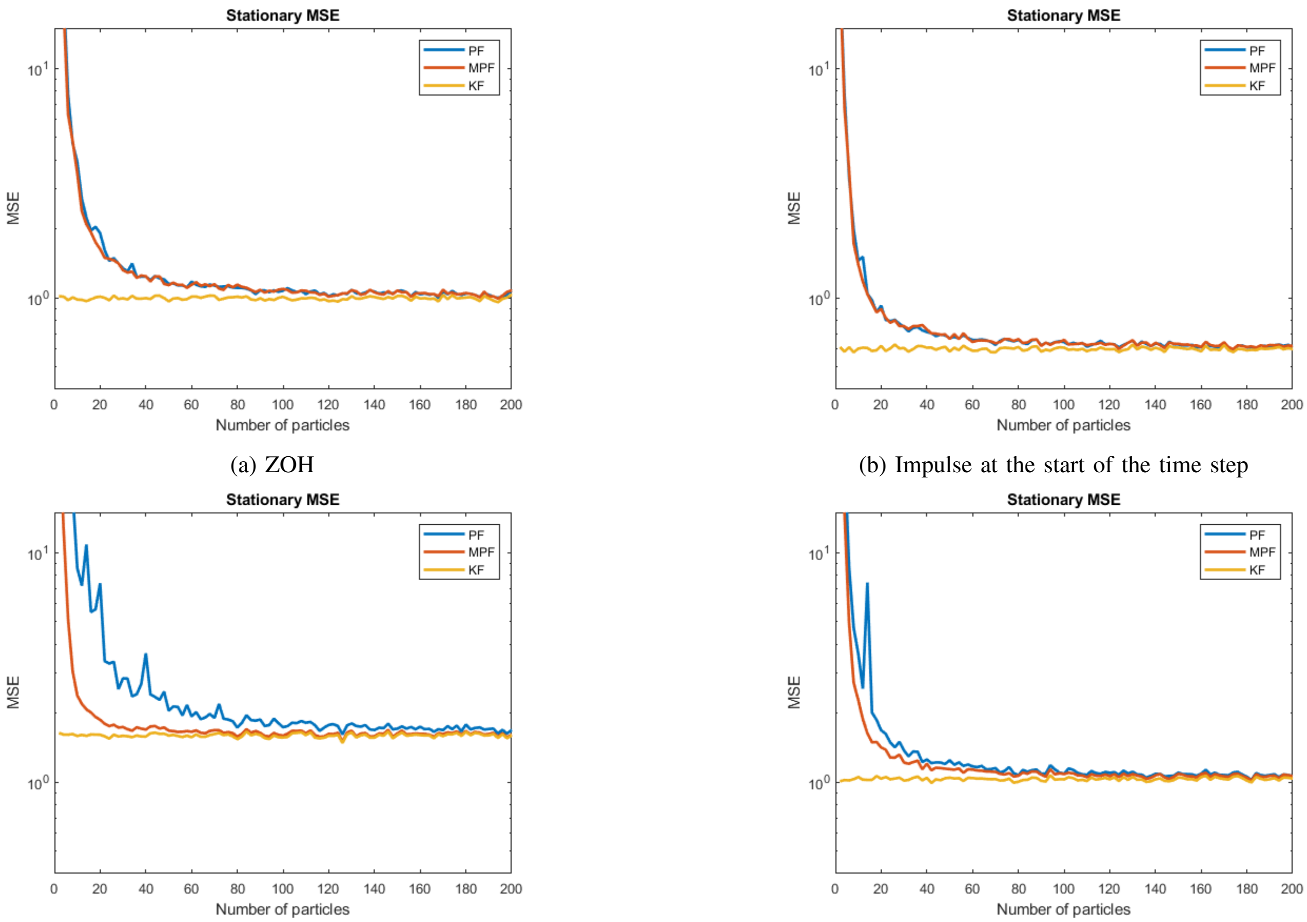
to achieve a good estimate compared to a PF without any marginalization. Further, (c) has the largest part of the total uncertainty represented by P_{KF}, which also leads to it having the largest increase in performance for the MPF compared to the PF.

Important to note however, is that, the uncertainty of P_{KF} being 0 means that the velocity is fully known, conditioned on a known trajectory, which each particle has. Since the trajectory of the actual object is not known, we do not know the actual velocity of the object either. Further, the uncertainty of P_{KF} being 0 is not the same as the velocity estimate not being needed. In both (1b) and (2a), the nonlinear state-update is dependent on the linear state. In this case, it means that the filter still needs to have an estimate of the velocity in order to know how to correctly propagate the position of the particles. The part of the KF which is not needed is P_{KF}, since it is possible to perfectly know the velocity given the position. Since the point of the MPF is to replace multiple particles in the velocity dimension with one particle which contains the possible spread of the velocity, and the spread now only being a point, the MPF will in this situation give identical results to the standard PF.

One interesting observation is that, for all discretization methods included here, the uncertainty reaches a stationary point after only a few iterations, after which the only change appears to be random noise. An effect of this is that it may be reasonable to only run a stationary KF with a pre-calculated P and K for the MPF, in order to reduce the computational requirements. This is however not further evaluated in this work.

VI. Conclusion

In this work, we have calculated the covariance update for the linear part of the MPF for various discretizations of the CV model. For many of the methods used, the covariance converged towards 0. The result of this was that for these discretization methods no gain in performance is made by marginalizing the velocity. This was caused by the KF not providing any additional information to the filter. While this has here only been shown to be the case for this specific application, it is likely that this issue occurs in some other



(a) ZOH

(b) Impulse at the start of the time step

(c) Impulse at the end of the time step

(d) Continuous noise

Fig. 2: The stationary MSE of the filters as a function of number of particles used

applications of the MPF as well. This seems like an interesting path for further research.

For the discretizations where P_{KF} did not converge towards 0, there is some level of performance gain. These benefits can still be very small, as is the case when assuming continuous noise, which can be seen in Fig. 2d.

References

[1] N. Gordon, D. Salmond, and A. Smith, "A novel approach to nonlinear/non-Gaussian Bayesian state estimation," in *IEE Proceedings on Radar and Signal Processing*, vol. 140, 1993, pp. 107–113.

[2] P. Bui Quang, C. Musso, and F. Gland, "An insight into the issue of dimensionality in particle filtering," in *13th Conference on Information Fusion, Fusion 2010*, 08 2010, pp. 1 – 8.

[3] C. Andrieu and A. Doucet, "Particle filtering for partially observed Gaussian state space models," *Journal of the Royal Statistical Society*, vol. 64, no. 4, pp. 827–836, 2002.

[4] A. Doucet, S. J. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.

[5] R. Chen and J. S. Liu, "Mixture Kalman filters," *Journal of the Royal Statistical Society*, vol. 62, no. 3, pp. 493–508, 2000.

[6] A. Doucet, N. Gordon, and V. Krishnamurthy, "Particle filters for state estimation of jump Markov linear systems," *IEEE Transactions on Signal Processing*, vol. 49, no. 3, pp. 613–624, 2001.

[7] 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, pp. 2279 – 2289, 08 2005.

[8] P.-J. Nordlund and F. Gustafsson, "Marginalized particle filter for accurate and reliable terrain-aided navigation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, pp. 1385 – 1399, 11 2009.

[9] S. Särkkä, A. Vehtari, and J. Lampinen, "Rao-Blackwellized particle filter for multiple target tracking," *Information Fusion*, vol. 8, no. 1, pp. 2–15, 2007, special Issue on the Seventh International Conference on Information Fusion-Part II. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1566253505000874

[10] 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, 2002.

[11] F. Gustafsson, "Particle filter theory and practice with positioning applications," *IEEE Aerospace and Electronic Systems Magazine*, vol. 25, no. 7, pp. 53–82, 2010.

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

[13] 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.

[14] A. W. Max, "Inverting modified matrices," in *Memorandum Rept. 42, Statistical Research Group*. Princeton Univ., 1950, p. 4.

[15] X. Rong Li and V. Jilkov, "Survey of maneuvering target tracking. Part i. Dynamic models," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1333–1364, 2003.

Appendix

Here the assumptions made in Sec. IV are explained and Q for each of them is derived. For the ZOH and continuous noise cases [15] provide similar models. However, in the ZOH case, there is no mention of what assumptions are made of the noise during discretization. In the case of the continuous noise, while similar expressions are presented, the notation and assumptions used are fairly different. For this reason, all models used are derived in this appendix.

For all of these discretizations the starting assumption is the time-continuous model

$$\dot{x}(t) = Ax(t) + B\nu(t), \tag{42a}$$

$$A = \begin{pmatrix} 0 & 0 \\ 0 & I_n \end{pmatrix}, \tag{42b}$$

$$B = \begin{pmatrix} 0 \\ I_n \end{pmatrix}, \tag{42c}$$

where $x(t)$ is the continuous state and $\nu(t)$ is the continuous time noise. This should be turned into the time-discrete model

$$x_{k+1} = Fx_k + w_k, \tag{43a}$$

which will be discretized with a sample time T. All discretizations used are on the form

$$F = e^{AT}, \tag{44a}$$

$$w_k = \int_0^T e^{A(T-\tau)} B\nu(kT+\tau)d\tau, \tag{44b}$$

arrived at from considering

$$\frac{d}{dt}e^{-At}x(t) = e^{-At}\dot{x}(t) - Ae^{-At}x(t) = e^{-At}B\nu(t). \tag{44c}$$

Since A and T are assumed to be known, F will therefore always be the same

$$F = \begin{pmatrix} I_n & TI_n \\ 0 & I_n \end{pmatrix}. \tag{45}$$

The difference between the different methods lies in how $\nu(t)$ is assumed to behave.

A. ZOH

In ZOH, we assume that the noise $\nu(t)$ is constant for each time step

$$\nu(t) = \nu_k \quad , kT \leq t < (k+1)T, \tag{46}$$

where ν_k is assumed to have covariance q. Evaluating (44b) yields

$$\begin{aligned} w_k &= \int_0^T e^{A(T-\tau)} d\tau B\nu_k \\ &= A^{-1}(e^{AT} - I_n)B\nu_k \\ &= \begin{pmatrix} \frac{T^2}{2} I_n \\ TI_n \end{pmatrix} \nu_k. \end{aligned} \tag{47}$$

From this, calculating the covariance becomes

$$Q = \mathrm{Cov}(w_k) = E\big[w_k w_k^T\big] = \begin{pmatrix} \frac{T^4}{4}q & \frac{T^3}{2}q \\ \frac{T^3}{2}q & T^2 q \end{pmatrix}. \tag{48}$$

B. Noise as an impulse at the start of the time step

Here, it is assumed that the noise can be described as

$$\nu(t) = \sum_k \delta(kT + \epsilon)\nu_k, \tag{49}$$

where ϵ is a sufficiently small positive number and ν_k is assumed to be noise with covariance q. Using this approach, the noise is discretized as

$$\begin{aligned} w_k &= \int_0^T e^{A(T-\tau)} B\delta(\tau - \epsilon)\nu_k d\tau \\ &= e^{A(T-\epsilon)}B\nu_k = \begin{pmatrix} (T-\epsilon)I_n \\ I_n \end{pmatrix} \nu_k \to \begin{pmatrix} TI_n \\ I_n \end{pmatrix} \nu_k. \end{aligned} \tag{50}$$

The simplification in (50) occurs when $\epsilon \to 0^+$. The covariance for this discretization method is given by

$$Q = E\big[w_k w_k^T\big] = \begin{pmatrix} T^2 q & Tq \\ Tq & q \end{pmatrix}. \tag{51}$$

C. Noise as an impulse at the end of the time step

Similarly to Sec. B, we here describe the noise as

$$\nu(t) = \sum_k \delta((k+1)T - \epsilon)\nu_k, \tag{52}$$

with the same assumptions on ϵ and ν_k. With this description, the noise is discretized as

$$\begin{aligned} w_k &= \int_0^T e^{A(T-\tau)} B\delta\big(\tau - (T-\epsilon)\big)\nu_k d\tau \\ &= e^{A\epsilon}B\nu_k = \begin{pmatrix} \epsilon I_n \\ I_n \end{pmatrix} \nu_k \to \begin{pmatrix} 0 \\ I_n \end{pmatrix} \nu_k, \end{aligned} \tag{53}$$

where once again, $\epsilon \to 0^+$. With this description, the covariance becomes

$$Q = E\big[w_k w_k^T\big] = \begin{pmatrix} 0 & 0 \\ 0 & q \end{pmatrix}. \tag{54}$$

D. Continuous noise

Continuous noise assumes that $\nu(t)$ is continuous white noise meaning

$$\mathrm{Cov}\big(\nu(t), \nu(t+\tau)\big) = \delta(\tau)q. \tag{55}$$

For continuous noise, we get

$$w_k = \int_0^T e^{A(T-\tau)} B\nu(\tau)d\tau. \tag{56}$$

While this can not be simplified in the same way as the previous cases, the covariance can still be calculated as

$$\begin{aligned} Q &= E\big[w_k w_k^T\big] \\ &= \int_0^T e^{A(T-\tau)} BqB^T \big(e^{A(T-\tau)}\big)^T d\tau \\ &= \begin{pmatrix} \frac{T^3}{3}q & \frac{T^2}{2}q \\ \frac{T^2}{2}q & Tq \end{pmatrix}. \end{aligned} \tag{57}$$