16

Distributed Particle Filtering over Sensor Networks for Autonomous Navigation of UAVs

Gerasimos G. Rigatos
Industrial Systems Institute
Greece

1. Introduction

State estimation and control over sensor networks is a problem met in several applications such as surveillance and condition monitoring of large-scale systems, multi-robot systems and cooperating UAVs. In sensor networks the simplest kind of architecture is centralized. Distributed sensors send measurement data to a central processing unit which provides the state estimate for the monitored system. Such an approach has several weaknesses: (i) it lacks fault tolerance: if the central processing unit is subject to a fault then state estimation becomes impossible, (ii) communication overhead often prohibits proper functioning in case of a large number of distributed measurement units. On the other hand decentralized architectures are based on the communication between neighboring measurement units. This assures scalability for the network since the number of messages received or sent by each measurement unit is independent of the total number of measurement units in the system. It has been shown that scalable decentralized state estimation can be achieved for linear Gaussian models, when the measurements are linear functions of the state and the associated process and measurement noise models follow a Gaussian distribution (Nettleton et al. 2003). A solution to decentralized sensor fusion over sensor networks with the use of distributed Kalman Filtering has been proposed in (Olfati-Saber 2006), (Watanabe & Tzafestas 1992), (Olfati-Saber 2005), (Gan & Harris 2001), (Gao et al. 2009). Distributed state estimation in the case of non-Gaussian models has been studied in (Rosencrantz et al. 2003) where decentralized sensor fusion with the use of distributed particle filters has been proposed in several other research works (Mahler 2007), (Makarenko & Durrant-Whyte 2006), (Deming & Perlovsky 2007).

In this paper autonomous navigation of UAVs will be examined and a solution to this problem will be first attempted with the use of the Extended Information Filter and the Unscented Kalman filter (Shima et al. 2007), (Lee et al. 2008), (Lee et al. 2008), (Vercauteren & Wang 2005). Comparatively, autonomous UAV navigation with the use of the Distributed Particle Filter will be studied. This problem belongs to the wider area of multi-source multi-target tracking (Coué et al. 2006), (Hue et al. 2002), (Ing & Coates 2005), (Coué et al. 2003), (Morelande & D. Mušicki 2005). Subproblems to be solved for succeeding autonomous navigation of the UAVs are: (i) implementation of sensor fusion with the use of distributed filtering. In this approach the goal is to consistently combine the local particle distribution with the communicated particle distribution coming from particle filters running on nearby

measurement stations (Caballero et al. 2008). It is assumed that each local measurement station runs its own local filter and communicates information to other measurement stations close to it. The motivation for using particle filters is that they can represent almost arbitrary probability distributions, thus becoming well-suited to accommodate the types of uncertainty and nonlinearities that arise in the distributed estimation (Rigatos 2009a), (Rigatos 2009b) (ii) nonlinear control of the UAVs based on the state estimates provided by the particle filtering algorithm. Various approaches have been proposed for the UAV navigation using nonlinear feedback control (Ren & Beard 2004), (Beard et al. 2002), (Singh & Fuller 2001). The paper proposes flatness-based control for the UAV models. Flatness-based control theory is based on the concept of differential flatness and has been successfully applied to several nonlinear dynamical systems. Flatness-based control for a UAV helicopter-like model has been developed in (Léchevin & Rabbath 2006), assuming that the UAV performs manoeuvres at a constant altitude.

The paper proposes first the Extended Information Filter (EIF) and the Unscented Information Filter (UIF) as possible approaches for fusing the state estimates provided by the local monitoring stations, under the assumption of Gaussian noises. The EIF and UIF estimated state vector is in turn used by a flatness-based controller that makes the UAV follow the desirable trajectory. The Extended Information Filter is a generalization of the Information Filter in which the local filters do not exchange raw measurements but send to an aggregation filter their local information matrices (local inverse covariance matrices) and their associated local information state vectors (products of the local information matrices with the local state vectors) (Shima et al. 2007), (Lee et al. 2008). In the case of the Unscented Information Filter there is no linearization of the UAVs observation equation. However the application of the Information Filter algorithm is possible through an implicit linearization which is performed by approximating the Jacobian matrix of the system's output equation by the product of the inverse of the state vector's covariance matrix (which can be also associated to the Fisher Information matrix) with the cross-correlation covariance matrix between the system's state vector and the system's output (Lee et al. 2008)], (Vercauteren & Wang 2005). Again, the local information matrices and the local information state vectors are transferred to an aggregation filter which produces the global estimation of the system's state vector.

Next, the Distributed Particle Filter (DPF) is proposed for fusing the state estimates provided by the local monitoring stations (local filters). The motivation for using DPF is that it is well-suited to accommodate non-Gaussian measurements. A difficulty in implementing distributed particle filtering is that particles from one particle set (which correspond to a local particle filter) do not have the same support (do not cover the same area and points on the samples space) as particles from another particle set (which are associated with another particle filter) (Ong et al. 2008), (Ong et al. 2006). This can be resolved by transforming the particles sets into Gaussian mixtures, and defining the global probability distribution on the common support set of the probability density functions associated with the local filters. The state vector which is estimated with the use of the DPF is used again by a flatness-based controller to make each UAV follow a desirable flight path.

The structure of the chapter is as follows: in Section 2 the Distributed Extended Kalman Filter (Extended Information Filter) is studied. In Section 3, the Distributed Unscented Kalman Filter (Unscented Information Filter) is analyzed and its use for distributed sensor fusion and state estimation is explained. In Section 4 Distributed Particle Filtering for sensor fusion-based state estimation will be analyzed. In Section 5 nonlinear control will be

proposed for succeeding trajectory tracking by the UAVs. In Section 6 simulation experiments will be provided about UAVs autonomous navigation using the proposed distributed particle filtering algorithm. The test case will be concerned with m helicopter models monitored by n different ground stations. By fusing the measurements from the distributed observation units with the use of the Extended Information Filter and the proposed Particle Filter algorithm, state estimates of the UAVs are obtained. These in turn are used by local nonlinear controllers for succeeding trajectory tracking. Finally in Section 7 concluding remarks will be provided.

2. Distributed Extended Kalman Filtering

2.1 Extended Kalman Filtering at local processing units

The distributed Extended Kalman Filter, also know as Extended Information Filter, performs fusion of the state estimates which are provided by local Extended Kalman Filters. Thus, the functioning of the local Extended Kalman Filters should be analyzed first. The following nonlinear state model is considered (Rigatos & Tzafestas 2007):

$$\begin{aligned} x(k+1) &= \phi(x(k)) + L(k)u(k) + w(k) \\ z(k) &= \gamma(x(k)) + v(k) \end{aligned} \tag{1}$$

where $x \in R^{m\times 1}$ is the system's state vector and $z \in R^{p\times 1}$ is the system's output, while $w(k)$ and $v(k)$ are uncorrelated, zero-mean, Gaussian zero-mean noise processes with co- variance matrices $Q(k)$ and $R(k)$ respectively. The operators $\phi(x)$ and $\gamma(x)$ are $\phi(x) = [\phi_1(x), \phi_2(x), \cdots, \phi_m(x)]^T$, and $\gamma(x) = [\gamma_1(x), \gamma_2(x), \cdots, \gamma_p(x)]^T$, respectively. It is assumed that ϕ and γ are sufficiently smooth in x so that each one has a valid series Taylor expansion.
Following a linearization procedure, ϕ is expanded into Taylor series about $\hat{x}$:

$$\phi(x(k)) = \phi(\hat{x}(k)) + J_\phi(\hat{x}(k))[x(k) - \hat{x}(k)] + \cdots \tag{2}$$

where $J_\phi(x)$ is the Jacobian of ϕ calculated at $\hat{x}(k)$:

$$J_\phi(x) = \frac{\partial \phi}{\partial x}\Big|_{x=\hat{x}(k)} = \begin{pmatrix} \frac{\partial \phi_1}{\partial x_1} & \frac{\partial \phi_1}{\partial x_2} & \cdots & \frac{\partial \phi_1}{\partial x_m} \\ \frac{\partial \phi_2}{\partial x_1} & \frac{\partial \phi_2}{\partial x_2} & \cdots & \frac{\partial \phi_2}{\partial x_m} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \phi_m}{\partial x_1} & \frac{\partial \phi_m}{\partial x_2} & \cdots & \frac{\partial \phi_m}{\partial x_m} \end{pmatrix} \tag{3}$$

Likewise, γ is expanded about $\hat{x}^-(k)$

$$\gamma(x(k)) = \gamma(\hat{x}^-(k)) + J_\gamma[x(k) - \hat{x}^-(k)] + \cdots \tag{4}$$

where $\hat{x}^-(k)$ is the estimation of the state vector $x(k)$ before measurement at the k-th instant to be receivec and $\hat{x}(k)$ is the updated estimation of the state vector after measurement at the k-th instant has been received. The Jacobian $J_\gamma(x)$ is

$$J_\gamma(x)=\frac{\partial \gamma}{\partial x}|_{x=\hat{x}^-(k)}=\begin{pmatrix}\frac{\partial \gamma_1}{\partial x_1} & \frac{\partial \gamma_1}{\partial x_2} & \cdots & \frac{\partial \gamma_1}{\partial x_m}\\ \frac{\partial \gamma_2}{\partial x_1} & \frac{\partial \gamma_2}{\partial x_2} & \cdots & \frac{\partial \gamma_2}{\partial x_m}\\ \vdots & \vdots & \vdots & \vdots\\ \frac{\partial \gamma_p}{\partial x_1} & \frac{\partial \gamma_p}{\partial x_2} & \cdots & \frac{\partial \gamma_p}{\partial x_m}\end{pmatrix} \tag{5}$$

The resulting expressions create first order approximations of ϕ and γ. Thus the linearized version of the system is obtained:

$$\begin{aligned} x(k+1) &= \phi(\hat{x}(k)) + J_\phi(\hat{x}(k))[x(k)-\hat{x}(k)] + w(k)\\ z(k) &= \gamma(\hat{x}^-(k)) + J_\gamma(\hat{x}^-(k))[x(k)-\hat{x}^-(k)] + v(k)\end{aligned} \tag{6}$$

Now, the EKF recursion is as follows: First the time update is considered: by $\hat{x}(k)$ the estimation of the state vector at instant k is denoted. Given initial conditions $\hat{x}(0)$ and $P^-(0)$ the recursion proceeds as:

- *Measurement update*. Acquire $z(k)$ and compute:

$$\begin{aligned} K(k) &= P^-(k)J_\gamma^T(\hat{x}^-(k))\cdot[J_\gamma(\hat{x}^-(k))P^-(k)J_\gamma^T(\hat{x}^-(k)) + R(k)]^{-1}\\ \hat{x}(k) &= \hat{x}^-(k) + K(k)[z(k)-\gamma(\hat{x}^-(k))]\\ P(k) &= P^-(k) - K(k)J_\gamma(\hat{x}^-(k))P^-(k)\end{aligned} \tag{7}$$

- *Time update*. Compute:

$$\begin{aligned} P^-(k+1) &= J_\phi(\hat{x}(k))P(k)J_\phi^T(\hat{x}(k)) + Q(k)\\ \hat{x}^-(k+1) &= \phi(\hat{x}(k)) + L(k)u(k)\end{aligned} \tag{8}$$

The schematic diagram of the EKF loop is given in Fig. 1.

2.2 Calculation of local estimations in terms of EIF information contributions

Again the discrete-time nonlinear system of Eq. (1) is considered. The Extended Information Filter (EIF) performs fusion of the local state vector estimates which are provided by the local Extended Kalman Filters, using the *Information matrix* and the *Information state vector* (Lee et al. 2008), (Lee et al. 2008), (Vercauteren & Wang 2005), (Manyika & H. Durrant-Whyte 1994). The Information Matrix is the inverse of the state vector covariance matrix, and can be also associated to the Fisher Information matrix." (Rigatos & Zhang 2009). The Information state vector is the product between the Information matrix and the local state vector estimate

$$\begin{aligned} Y(k) &= P^{-1}(k) = I(K)\\ \hat{y}(k) &= P^{--1}\hat{x}(k) = Y(k)\hat{x}(k)\end{aligned} \tag{9}$$

The update equation for the Information Matrix and the Information state vector are given by

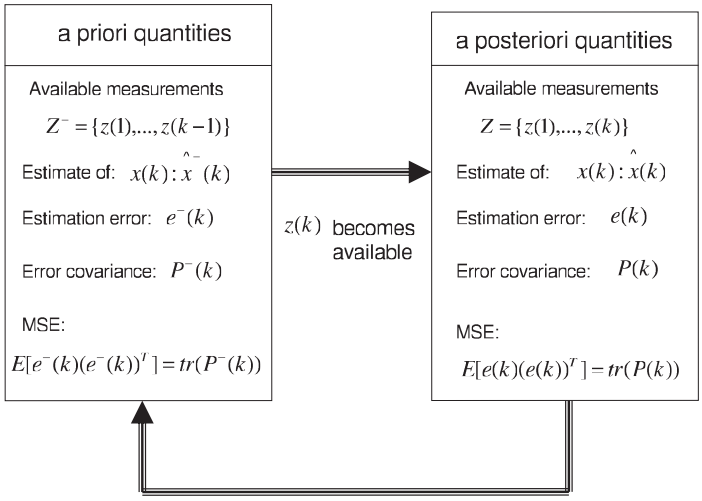

Fig. 1. Schematic diagram of the EKF loop

$$\begin{aligned} Y(k) &= P^-(k)^{-1} + J_\gamma^T(k)R^{-1}(k)J_\gamma(k) \\ &= Y^-(k) + I(k) \end{aligned} \tag{10}$$

$$\begin{aligned} \hat{y}(k) &= \hat{y}^-(k) + J_\gamma^T R(k)^{-1}[z(k) - \gamma(x(k)) + J_\gamma \hat{x}^-(k)] \\ &= \hat{y}^-(k) + i(k) \end{aligned} \tag{11}$$

where

$$\begin{aligned} &I(k) = J_\gamma^T(k)R^(k) - 1J_\gamma(k) \text{ is the associated information matrix and} \\ &i(k) = J_\gamma^T R^(k) - 1[(z(k) - \gamma(x(k))) + J_\gamma \hat{x}^-(k)] \text{ is the information state contribution} \end{aligned} \tag{12}$$

The predicted information state vector and Information matrix are obtained from

$$\begin{aligned} &\hat{y}^-(k) = P^-(k)^{-1}\hat{x}^-(k) \\ &Y^-(k) = P^-(k)^{-1} = [J_\phi(k)P^-(k)J_\phi(k)^T + Q(k)]^{-1} \end{aligned} \tag{13}$$

The Extended Information Filter is next formulated for the case that multiple local sensor measurements and local estimates are used to increase the accuracy and reliability of the estimation. It is assumed that an observation vector $z^i(k)$ is available for N different sensor sites i = 1, 2, ⋯ ,N and each sensor observes a common state according to the local observation model, expressed by

$$z^i(k) = \gamma(x(k)) + v^i(k),\ i = 1,2,\cdots,N \tag{14}$$

where the local noise vector $v^i(k) \sim N(0,R^i)$ is assumed to be white Gaussian and uncorrelated between sensors. The variance of a composite observation noise vector v_k is expressed in terms of the block diagonal matrix

$$R(k) = diag[R(k)^1, \cdots, R^N(k)]^T \tag{15}$$

The information contribution can be expressed by a linear combination of each local information state contribution i^i and the associated information matrix I^i at the i-th sensor site

$$\begin{aligned} i(k) &= \sum_{i=1}^{N} J_\gamma^{iT}(k) R^i(k)^{-1} [z^i(k) - \gamma^k(x(k)) + J_\gamma^i(k) \hat{x}^-(k)] \\ I(k) &= \sum_{i=1}^{N} J_\gamma^{iT}(k) R^i(k)^{-1} J_\gamma^i(k) \end{aligned} \tag{16}$$

Using Eq. (16) the update equations for fusing the local state estimates become

$$\begin{aligned} \hat{y}(k) &= \hat{y}^-(k) + \sum_{i=1}^{N} J_\gamma^{iT}(k) R^i(k)^{-1} [z^i(k) - \gamma^k(x(k)) + J_\gamma^i(k) \hat{x}^-(k)] \\ Y(k) &= Y^-(k) + \sum_{i=1}^{N} J_\gamma^{iT}(k) R^i(k)^{-1} J_\gamma^i(k) \end{aligned} \tag{17}$$

It is noted that in the Extended Information Filter an aggregation (master) fusion filter produces a global estimate by using the local sensor information provided by each local filter.

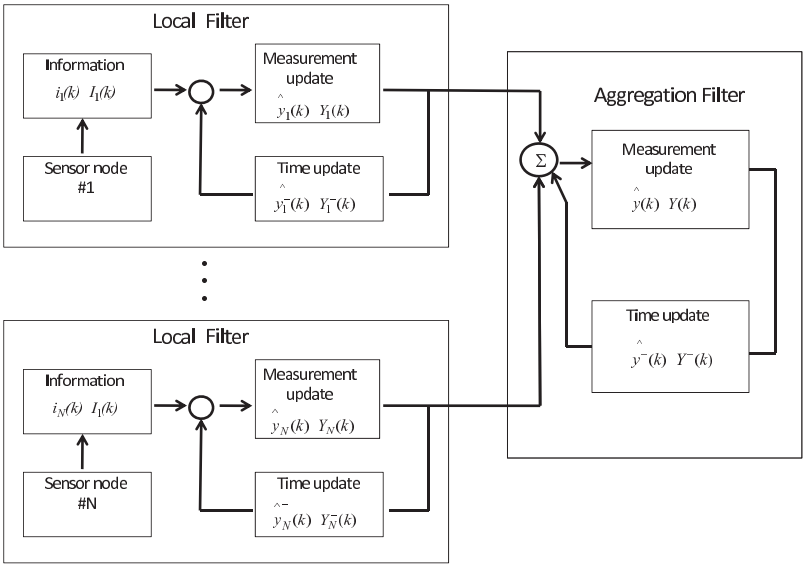

Fig. 2. Fusion of the distributed state estimates with the use of the Extended Information Filter

As in the case of the Extended Kalman Filter the local filters which constitute the Extended Information Filter can be written in terms of *time update* and a *measurement update* equation.
Measurement update: Acquire $z(k)$ and compute

$$\begin{aligned} Y(k) &= P^-(k)^{-1} + J_\gamma^T(k) R(k)^{-1} J_\gamma(k) \\ \text{or } Y(k) &= Y^-(k) + I(k) \text{ where } I(k) = J_\gamma^T(k) R^{-1}(k) J_\gamma(k) \end{aligned} \tag{18}$$

$$\hat{y}(k) = \hat{y}^{-}(k) + J_{\gamma}^{T}(k)R(k)^{-1}[z(k) - \gamma(\hat{x}(k)) + J_{\gamma}\hat{x}^{-}(k)]$$
$$\text{or } \hat{y}(k) = \hat{y}^{-}(k) + i(k) \tag{19}$$

Time update: Compute

$$Y^{-}(k+1) = P^{-}(k+1)^{-1} = [J_{\phi}(k)P(k)J_{\phi}(k)^{T} + Q(k)]^{-1} \tag{20}$$

$$y^{-}(k+1) = P^{-}(k+1)^{-1}\hat{x}^{-}(k+1) \tag{21}$$

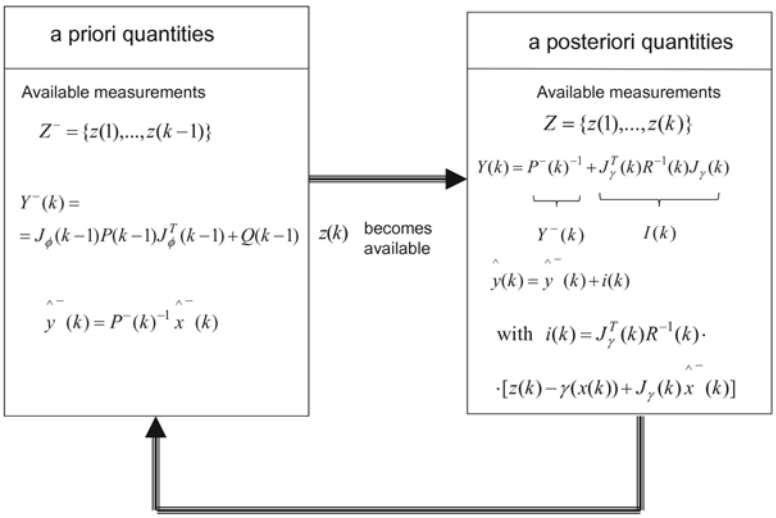

Fig. 3. Schematic diagram of the Extended Information Filter loop

2.3 Extended Information Filtering for state estimates fusion

In the Extended Information Filter each one of the local filters operates independently, processing its own local measurements. It is assumed that there is no sharing of measurements between the local filters and that the aggregation filter (Fig. 2) does not have direct access to the raw measurements feeding each local filter. The outputs of the local filters are treated as measurements which are fed into the aggregation fusion filter (Lee et al. 2008), (Lee et al. 2008), (Vercauteren &Wang 2005). Then each local filter is expressed by its respective error covariance and estimate in terms of information contributions given in Eq.(13)

$$P_i^{-1}(k) = P_i^{-}(k)^{-1} + J_{\gamma}^{T}(k)R(k)^{-1}J_{\gamma}(k)$$
$$\hat{x}_i(k) = P_i(k)(P_i^{-}(k)^{-1}\hat{x}_i^{-}(k) + J_{\gamma}^{T}(k)R(k)^{-1}\cdot[z^{i}(k) - \gamma^{k}(x(k)) + J_{\gamma}^{i}(k)\hat{x}_i^{-}(k)]) \tag{22}$$

It is noted that the local estimates are suboptimal and also conditionally independent given their own measurements. The global estimate and the associated error covariance for the aggregate fusion filter can be rewritten in terms of the computed estimates and covariances from the local filters using the relations

$$J_\gamma^T(k)R(k)^{-1}J_\gamma(k) = P_i^{-1}(k) - P_i^-(k)^{-1}$$
$$J_\gamma^T(k)R(k)^{-1}[z^i(k) - \gamma^k(x(k)) + J_\gamma^i(k)\hat{x}_i^-(k)] = P_i(k)^{-1}\hat{x}_i(k) - P_i^-(k)^{-1}\hat{x}_i^-(k) \quad (23)$$

For the general case of N local filters $i = 1, \cdots, N$, the distributed filtering architecture is described by the following equations

$$P(k)^{-1} = P^-(k)^{-1} + \sum_{i=1}^{N}[P_i(k)^{-1} - P_i^-(k)^{-1}]$$
$$\hat{x}(k) = P(k)[P^-(k)^{-1}\hat{x}^-(k) + \sum_{i=1}^{N}(P_i(k)^{-1}\hat{x}_i(k) - P_i^-(k)^{-1}\hat{x}_i^-(k))] \quad (24)$$

It is noted that the global state update equation in the above distributed filter can be written in terms of the information state vector and of the information matrix

$$\hat{y}(k) = \hat{y}^-(k) + \sum_{i=1}^{N}(\hat{y}_i(k) - \hat{y}_i^-(k))$$
$$\hat{Y}(k) = \hat{Y}^-(k) + \sum_{i=1}^{N}(\hat{Y}_i(k) - \hat{Y}_i^-(k)) \quad (25)$$

The local filters provide their own local estimates and repeat the cycle at step $k + 1$. In turn the global filter can predict its global estimate and repeat the cycle at the next time step $k + 1$ when the new state $\hat{x}(k + 1)$ and the new global covariance matrix $P(k + 1)$ are calculated. From Eq. (24) it can be seen that if a local filter (processing station) fails, then the local covariance matrices and the local state estimates provided by the rest of the filters will enable an accurate computation of the system's state vector.

3. Distributed Sigma-Point Kalman Filtering

3.1 Unscented Kalman Filtering at local processing units

It is also possible to estimate the state vectors of the distributed UAVs which constitute the multi-UAV system through the fusion of the estimates provided by local Sigma-Point Kalman Filters. This can be succeeded using the Distributed Sigma-Point Kalman Filter, also known as *Unscented Information Filter* (UIF) (Lee et al. 2008), (Lee et al. 2008). First, the functioning of the local Sigma-Point Kalman Filters will be explained. Each local Sigma-Point Kalman Filter generates an estimation of the UAV's state vector by fusing measurement from distributed sensors (e.g. IMU and GPS). Sigma-Point Kalman Filtering is proposed (Julier et al. 2000), (Julier et al. 2004), (Särrkä 2007). The Sigma-Point Kalman Filter overcomes the flaws of Extended Kalman Filtering. Unlike EKF no analytical Jacobians of the system equations need to be calculated as in the case for the EKF. This makes the sigma-point approach suitable for application in "black-box" models where analytical expressions of the system dynamics are either not available or not in a form which allows for easy linearization. This is achieved through a different approach for calculating the posterior 1st and 2nd order statistics of a random variable that undergoes a nonlinear transformation. The state distribution is represented again by a Gaussian Random Variable but is now specified using a minimal set of deterministically chosen weighted sample points. The basic sigma-point approach can be described as follows:

1. A set of weighted samples (sigma-points) are deterministically calculated using the mean and square-root decomposition of the covariance matrix of the system's state vector. As a minimal requirement the sigma-point set must completely capture the first and second order moments of the prior random variable. Higher order moments can be captured at the cost of using more sigma-points.
2. The sigma-points are propagated through the true nonlinear function using functional evaluations alone, i.e. no analytical derivatives are used, in order to generate a posterior sigma-point set.
3. The posterior statistics are calculated (approximated) using tractable functions of the propagated sigma-points and weights. Typically, these take on the form of a simple weighted sample mean and covariance calculations of the posterior sigma points.

It is noted that the sigma-point approach differs substantially from general stochastic sampling techniques, such as Monte-Carlo integration (e.g Particle Filtering methods) which require significantly more sample points in an attempt to propagate an accurate (possibly non-Gaussian) distribution of the state. The deceptively simple sigma-point approach results in posterior approximations that are accurate to the third order for Gaussian inputs for all nonlinearities. For non-Gaussian inputs, approximations are accurate to at least the second-order, with the accuracy of third and higher-order moments determined by the specific choice of weights and scaling factors.

The Unscented Kalman Filter (UKF) is a special case of Sigma-Point Kalman Filters. The UKF is a discrete time filtering algorithm which uses the unscented transform for computing approximate solutions to the filtering problem of the form

$$\begin{aligned} x(k+1) &= \phi(x(k)) + L(k)U(k) + w(k) \\ y(k) &= \gamma(x(k)) + v(k) \end{aligned} \tag{26}$$

where $x(k) \in R^n$ is the system's state vector, $y(k) \in R^m$ is the measurement, $w(k) \in R^n$ is a Gaussian process noise $w(k) \sim N(0, Q(k))$, and $v(k) \in R^m$ is a Gaussian measurement noise $v(k) \sim N(0, R(k))$. The mean and covariance of the initial state $x(0)$ are $m(0)$ and $P(0)$, respectively.

Some basic operations performed in the UKF algorithm (*Unscented Transform*) are summarized as follows:

1. Denoting the current state mean as $\hat{x}$, a set of $2n+1$ sigma points is taken from the columns of the $n \times n$ matrix $\sqrt{(n+\lambda)P_{xx}}$ as follows:

$$\begin{aligned} x^0 &= \hat{x} \\ x^i &= \hat{x} + [\sqrt{(n+\lambda)P_{xx}}]_i, \; i = 1, \cdots, n \\ x^i &= \hat{x} - [\sqrt{(n+\lambda)P_{xx}}]_i, \; i = n+1, \cdots, 2n \end{aligned} \tag{27}$$

and the associate weights are computed:

$$\begin{aligned} W_0^{(m)} &= \frac{\lambda}{(n+\lambda)} \qquad W_0^{(c)} = \frac{\lambda}{(n+\lambda) + (1-\alpha^2+b)} \\ W_i^{(m)} &= \frac{1}{2(n+\lambda)}, \; i = 1, \cdots, 2n \qquad W_i^{(c)} = \frac{1}{2(n+\lambda)} \end{aligned} \tag{28}$$

where $i = 1, 2, \cdots, 2n$ and $\lambda = \alpha^2(n + \kappa) - n$ is a scaling parameter, while α, β and κ are constant parameters. Matrix P_{xx} is the covariance matrix of the state x.

2. Transform each of the sigma points as

$$z^i = h(x^i)\ i = 0, \cdots, 2n \tag{29}$$

3. Mean and covariance estimates for z can be computed as

$$\begin{aligned} \hat{z} &\simeq \sum_{i=0}^{2n} W_i^{(m)} z^i \\ P_{zz} &= \sum_{i=0}^{2n} W_i^{(c)} (z^i - \hat{z})(z^i - \hat{z})^T \end{aligned} \tag{30}$$

4. The cross-covariance of x and z is estimated as

$$P_{xz} = \simeq \sum_{i=0}^{2n} W_i^{(c)} (x^i - \hat{x})(z^i - \hat{z})^T \tag{31}$$

The matrix square root of positive definite matrix P_{xx} means a matrix $A = \sqrt{P_{xx}}$ such that $P_{xx} = AA^T$ and a possible way for calculation is SVD.

Next the basic stages of the *Unscented Kalman Filter* are given:

As in the case of the Extended Kalman Filter and the Particle Filter, the Unscented Kalman Filter also consists of prediction stage (time update) and correction stage (measurement update) (Julier et al. 2004), (Särrkä 2007).

Time update: Compute the predicted state mean $\hat{x}^-(k)$ and the predicted covariance $P_{xx}^-(k)$ as

$$\begin{aligned} [\hat{x}^-(k), P_{xx}^-(k)] &= UT(f_d, \hat{x}(k-1), P_{xx}(k-1)) \\ P_{xx}^-(k) &= P_{xx}(k-1) + Q(k-1) \end{aligned} \tag{32}$$

Measurement update: Obtain the new output measurement z_k and compute the predicted mean $\hat{z}(k)$ and covariance of the measurement $P_{zz}(k)$, and the cross covariance of the state and measurement $P_{xz}(k)$

$$\begin{aligned} [\hat{z}(k), P_{zz}(k), P_{xz}(k)] &= UT(h_d, \hat{x}^-(k), P_{xx}^-(k)) \\ P_{zz}(k) &= P_{zz}(k) + R(k) \end{aligned} \tag{33}$$

Then compute the filter gain $K(k)$, the state mean $\hat{x}(k)$ and the covariance $P_{xx}(k)$, conditional to the measurement $y(k)$

$$\begin{aligned} K(k) &= P_{xz}(k) P_{zz}^{-1}(k) \\ \hat{x}(k) &= \hat{x}^-(k) + K(k)[z(k) - \hat{z}(k)] \\ P_{xx}(k) &= P_{xx}^-(k) - K(k) P_{zz}(k) K(k)^T \end{aligned} \tag{34}$$

The filter starts from the initial mean $m(0)$ and covariance $P_{xx}(0)$. The stages of state vector estimation with the use of the Unscented Kalman Filter algorithm are depicted in Fig. 6.

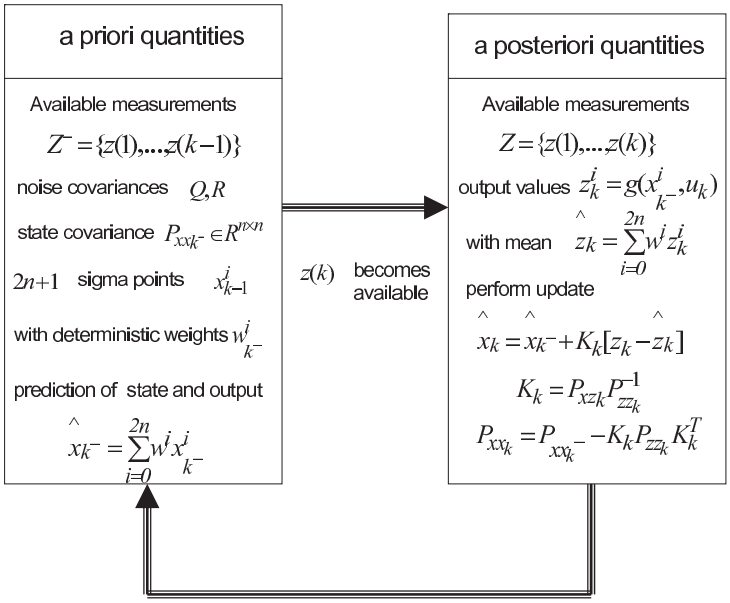

Fig. 4. Schematic diagram of the Unscented Kalman Filter loop

3.2 Unscented Information Filtering

The Unscented Information Filter (UIF) performs fusion of the state vector estimates which are provided by local Unscented Kalman Filters, by weighting these estimates with local Information matrices (inverse of the local state vector covariance matrices which are again recursively computed) (Lee et al. 2008), (Lee et al. 2008), (Vercauteren &Wang 2005). The Unscented Information Filter is derived by introducing a linear error propagation based on the unscented transformation into the Extended Information Filtering structure. First, an augmented state vector $x_\alpha^-(k)$ is considered, along with the process noise vector, and the associated covariance matrix is introduced

$$\hat{x}_\alpha^-(k) = \begin{pmatrix} \hat{x}^-(k) \\ \hat{w}^-(k) \end{pmatrix}, P^{\alpha-}(k) = \begin{pmatrix} P^-(k) & 0 \\ 0 & Q^-(k) \end{pmatrix} \tag{35}$$

As in the case of local (lumped) Unscented Kalman Filters, a set of weighted sigma points $X_\alpha^{i-}(k)$ is generated as

$$\begin{aligned} X^-_{\alpha,0}(k) &= \hat{x}_\alpha^-(k) \\ X^-_{\alpha,i}(k) &= \hat{x}_\alpha^-(k) + [\sqrt{(n_\alpha + \lambda)} P^-_\alpha(k-1)]_i, i = 1,\cdots,n \\ X^-_{\alpha,i}(k) &= \hat{x}_\alpha^-(k) + [\sqrt{(n_\alpha + \lambda)} P^-_\alpha(k-1)]_i, i = n+1,\cdots,2n \end{aligned} \tag{36}$$

where $\lambda = \alpha^2(n_\alpha + \kappa) - n_\alpha$ is a scaling, while $0 \le \alpha \le 1$ and κ are constant parameters. The corresponding weights for the mean and covariance are defined as in the case of the lumped Unscented Kalman Filter

$$W_0^{(m)} = \frac{\lambda}{n_\alpha + \lambda} \qquad W_0^{(c)} = \frac{\lambda}{(n_\alpha + \lambda) + (1 - \alpha^2 + \beta)}$$
$$W_i^{(m)} = \frac{1}{2(n_\alpha + \lambda)}, i = 1, \cdots, 2n_\alpha \qquad W_i^{(C)} = \frac{1}{2(n_\alpha + \lambda)}, i = 1, \cdots, 2n_\alpha \tag{37}$$

where β is again a constant parameter. The equations of the prediction stage (measurement update) of the information filter, i.e. the calculation of the information matrix and the information state vector of Eq. (13) now become

$$\hat{y}^-(k) = Y^-(k) \sum_{i=0}^{2n_\alpha} W_i^m X_i^x(k)$$
$$Y^-(k) = P^-(k)^{-1} \tag{38}$$

where X_i^x are the predicted state vectors when using the sigma point vectors X_i^w in the state equation $X_i^x(k+1) = \phi(X_i^{w-}(k)) + L(k)U(k)$. The predicted state covariance matrix is computed as

$$P^-(k) = \sum_{i=0}^{2n_\alpha} W_i^{(c)} [X_i^x(k) - \hat{x}^-(k)][X_i^x(k) - \hat{x}^-(k)]^T \tag{39}$$

As noted, the equations of the Extended Information Filter (EIF) are based on the linearized dynamic model of the system and on the inverse of the covariance matrix of the state vector. However, in the equations of the Unscented Kalman Filter (UKF) there is no linearization of the system dynamics, thus the UKF cannot be included directly into the EIF equations. Instead, it is assumed that the nonlinear measurement equation of the system given in Eq. (1) can be mapped into a linear function of its statistical mean and covariance, which makes possible to use the information update equations of the EIF. Denoting $Y_i(k) = \gamma(X_i^x(k))$ (i.e. the output of the system calculated through the propagation of the *i*-th sigma point X^i through the system's nonlinear equation) the observation covariance and its cross-covariance are approximated by

$$P_{YY}^-(k) = E[(z(k) - \hat{z}^-(k))(z(k) - \hat{z}^-(k))^T]$$
$$\cong J_\gamma(k) P^-(k) J_\gamma(k)^T \tag{40}$$

$$P_{XY}^-(k) = E[(x(k) - \hat{x}^-(k))(z(k) - \hat{z}^-(k))^T]$$
$$\cong P^-(k) J_\gamma(k)^T \tag{41}$$

where $z(k) = \gamma(x(k))$ and $J_\gamma(k)$ is the Jacobian of the output equation $\gamma(x(k))$. Next, multiplying the predicted covariance and its inverse term on the right side of the information matrix Eq. (12) and replacing $P(k)J_\gamma(k)^T$ with $P_{XY}^-(k)$ (k) gives the following representation of the information matrix (Lee et al. 2008), (Lee et al. 2008), (Vercauteren &Wang 2005)

$$\begin{aligned} I(k) &= J_\gamma(k)^T R(k)^{-1} J_\gamma(k) \\ &= P^-(k)^{-1} P^-(k) J_\gamma(k)^T R(k)^{-1} J_\gamma^-(k) P^-(k)^T (P^-(k)^{-1})^T \\ &= P^-(k)^{-1} P_{XY}(k) R(k)^{-1} P_{XY}(k)^T (P^-(k)^{-1})^T \end{aligned} \tag{42}$$

where $P^{-}(k)^{-1}$ is calculated according to Eq. (39) and the cross-correlation matrix $P_{XY}(k)$ is calculated from

$$P_{XY}^{-}(k) = \sum_{i=0}^{2n_\alpha} W_i^{(c)} [X_i^x(k) - \hat{x}^{-}(k)][Y_i(k) - \hat{z}^{-}(k)]^T \tag{43}$$

where $Y_i(k) = \gamma(X_i^x(k))$ and the predicted measurement vector $\hat{z}^{-}(k)$ is obtained by $\hat{z}^{-}(k) = \sum_{i=0}^{2n} W_i^{(m)} Y_i(k)$. Similarly, the information state vector i_k can be rewritten as

$$\begin{aligned} i(k) &= J_\gamma^T(k) R(k)^{-1} [z(k) - \gamma(x(k)) + J_\gamma^T(k) \hat{x}^{-}(k)] \\ &= P^{-}(k)^{-1} P^{-}(k) J_\gamma^T(k) R(k)^{-1} \cdot \\ &\cdot [z(k) - \gamma(x(k)) + J_\gamma^T(k) \hat{x}^{-}(k) (P^{-}(k))^T (P^{-}(k)^{-1})^T \hat{x}^{-}(k)] \\ &= P^{-}(k)^{-1} P_{XY}^{-}(k) R(k)^{-1} [z(k) - \gamma(x(k)) + P_{XY}^{-}(k) (P^{-}(k)^{-1})^T \hat{x}^{-}(k)] \end{aligned} \tag{44}$$

To complete the analogy to the information contribution equations of the EIF a "measurement" matrix $H^T(k)$ is defined as

$$H(k)^T = P^{-}(k)^{-1} P_{XY}^{-}(k) \tag{45}$$

In terms of the measurement matrix $H(k)$ the information contributions equations are written as

$$\begin{aligned} i(k) &= H^T(k) R(k)^{-1} [z(k) - \gamma(x(k)) + H(k) \hat{x}^{-}(k)] \\ I(k) &= H^T(k) R(k)^{-1} H(k) \end{aligned} \tag{46}$$

The above procedure leads to an implicit linearization in which the nonlinear measurement equation of the system given in Eq. (1) is approximated by the statistical error variance and its mean

$$z(k) = h(x(k)) \simeq H(k) x(k) + \overline{u}(k) \tag{47}$$

where $\overline{u}(k) = \gamma(\hat{x}^{-}(k)) - H(k) \hat{x}^{-}(k)$ is a measurement residual term. (47).

3.3 Calculation of local estimations in terms of UIF information contributions

Next, the local estimations provided by distributed (local) Unscented Kalmans filters will be expressed in terms of the information contributions (information matrix I and information state vector i) of the Unscented Information Filter, which were defined in Eq. (46) (Lee et al. 2008), (Lee et al. 2008), (Vercauteren &Wang 2005). It is assumed that the observation vector $\overline{z}_i(k+1)$ is available from N different sensors, and that each sensor observes a common state according to the local observation model, expressed by

$$\overline{z}_i(k) = H_i(k) x(k) + \overline{u}_i(k) + v_i(k) \tag{48}$$

where the noise vector $v_i(k)$ is taken to be white Gaussian and uncorrelated between sensors. The variance of the composite observation noise vector v_k of all sensors is written in terms of the block diagonal matrix $R(k) = diag[R_1(k)^T, \cdots, R_N(k)^T]^T$. Then one can define the local information matrix $I_i(k)$ and the local information state vector $i^i(k)$ at the i-th sensor, as follows

$$i_i(k) = H_i^T(k)R_i(k)^{-1}[z_i(k) - \gamma_i(x(k)) + H_i(k)\hat{x}^-(k)]$$
$$I_i(k) = H_i^T(k)R_i(k)^{-1}H_i(k) \tag{49}$$

Since the information contribution terms have group diagonal structure in terms of the innovation and measurement matrix, the update equations for the multiple state estimation and data fusion are written as a linear combination of the local information contribution terms

$$\hat{y}(k) = \hat{y}^-(k) + \sum_{i=1}^{N} i_i(k)$$
$$Y(k) = Y^-(k) + \sum_{i=1}^{N} I_i(k) \tag{50}$$

Then using Eq. (38) one can find the mean state vector for the multiple sensor estimation problem.

As in the case of the Unscented Kalman Filter, the Unscented Information Filter running at the *i*-th measurement processing unit can be written in terms of *measurement update* and *time update* equations:

Measurement update: Acquire measurement $z(k)$ and compute

$$Y(k) = P^-(k)^{-1} + H^T(k)R^{-1}(k)H(k)$$
$$\text{or } Y(k) = Y^-(k) + I(k) \text{ where } I(k) = H^T(k)R^{-1}(k)H(k) \tag{51}$$

$$\hat{y}(k) = \hat{y}^-(k) + H^T(k)R^{-1}(k)[z(k) - \gamma(\hat{x}(k)) + H(k)\hat{x}^-(k)]$$
$$\text{or } \hat{y}(k) = \hat{y}^-(k) + i(k) \tag{52}$$

Time update: Compute

$$Y^-(k+1) = (P^-(k+1))^{-1}$$
$$\text{where } P^-(k+1) = \sum_{i=0}^{2n_\alpha} W_i^{(c)}[X_i^x(k+1) - \hat{x}^-(k+1)][X_i^x(k+1) - \hat{x}^-(k+1)]^T \tag{53}$$

$$\hat{y}(k+1) = Y(k+1)\sum_{i=0}^{2n_\alpha} W_i^{(m)} X_i^x(k+1)$$
$$\text{where } X_i^x(k+1) = \phi(X_i^w(k)) + L(k)U(k) \tag{54}$$

3.4 Distributed Unscented Information Filtering for state estimates fusion

It has been shown that the update of the aggregate state vector of the Unscented Information Filter architecture can be expressed in terms of the local information matrices I_i and of the local information state vectors i_i, which in turn depend on the local covariance matrices P and cross-covariance matrices P_{XY}. Next, it will be shown that the update of the aggregate state vector can be also expressed in terms of the local state vectors $x_i(k)$ and in terms of the local covariance matrices $P_i(k)$ and cross-covariance matrices $P_{XY}^i(k)$. It is assumed that the local filters do not have access to each other row measurements and that they are allowed to communicate only their information matrices and their local information state vectors. Thus each local filter is expressed by its respective error covariance and estimate in terms of the local information state contribution i_i and its associated information matrix I_i at the *i*-th filter site. Then using Eq. (38) one obtains

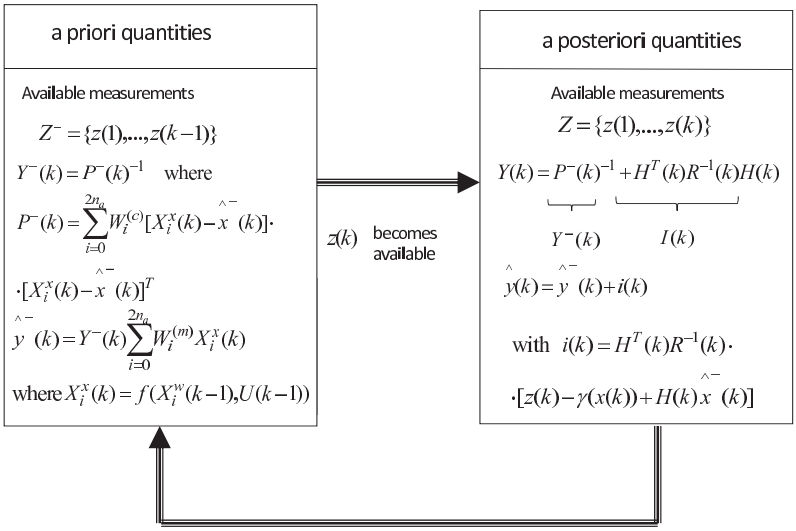

Fig. 5. Schematic diagram of the Unscented Information Filter loop

$$
\begin{aligned}
P_i(k)^{-1} &= P_i^-(k)^{-1} + H_i^T(k)R_i(k)^{-1}H_i(k) \\
\hat{x}_i &= P_i(k)(P_i^-(k)\hat{x}_i^-(k) + H_i^T(k)R_i(k)^{-1}[z_i(k) - \gamma_i(x(k)) + H_i(k)\hat{x}^-(k)])
\end{aligned}
\tag{55}
$$

Using Eq. (55), each local information state contribution i_i and its associated information matrix I_i at the i-th filter are rewritten in terms of the computed estimates and covariances of the local filters

$$
\begin{aligned}
H_i^T(k)R_i(k)^{-1}H_i(k) &= P_i(k)^{-1} - P_i^-(k)^{-1} \\
H_i^T(k)R_i(k)^{-1}[z_i(k) - \gamma_i(x(k)) + H_i(k)\hat{x}^-(k)] &= P_i(k)^{-1}\hat{x}_i(k) - P_i^-(k)^{-1}\hat{x}_i^-(k)
\end{aligned}
\tag{56}
$$

where according to Eq.(45) it holds $H_i(k) = P_i^-(k)^{-1}P_{XY,i}^-(k)$. Next, the aggregate estimates of the distributed unscented information filtering are derived for a number of N local filters $i = 1, \cdots, N$ and sensor measurements, first in terms of covariances (Vercauteren &Wang 2005), (Lee et al. 2008), (Lee et al. 2008)

$$
\begin{aligned}
P(k)^{-1} &= P^-(k)^{-1} + \sum_{i=1}^{N}[P_i(k)^{-1} - P_i^-(k)^{-1}] \\
\hat{x}(k) &= P(k)[P^-(k)^{-1}\hat{x}^-(k) + \sum_{i=1}^{N}(P_i(k)^{-1}\hat{x}_i(k) - P_i^-(k)^{-1}\hat{x}_i^-(k))]
\end{aligned}
\tag{57}
$$

and also in terms of the information state vector and of the information state covariance matrix

$$
\begin{aligned}
\hat{y}(k) &= \hat{y}^-(k) + \sum_{i=1}^{N}(\hat{y}_i(k) - \hat{y}_i^-(k)) \\
Y(k) &= Y^-(k) + \sum_{i=1}^{N}[Y_i(k) - Y_i^-(k)]
\end{aligned}
\tag{58}
$$

State estimation fusion based on the Unscented Information Filter (UIF) is fault tolerant. From Eq. (57) it can be seen that if a local filter (processing station) fails, then the local covariance matrices and local estimates provided by the rest of the filters will enable a reliable calculation of the system's state vector. Moreover, it is and computationally efficient comparing to centralized filters and results in enhanced estimation accuracy.

4. Distributed Particle Filter

4.1 Particle Filtering at local processing units

4.1.1 The particle approximation of probability density functions

One can also estimate the state vector of the UAVs that constitute the multi-UAV system through the fusion of estimates provided by local Particle Filters. This can be succeeded using the Distributed Particle Filter (DPF). First, the functioning of the local Particle Filters will be explained. Each local Particle Filter generates an estimation of the UAV's state vector by fusing measurements from distibuted sensors. Particle Filtering is a method for state estimation that is not dependent on the probability density function of the measurements. In the general case the equations of the optimal filter used for the calculation of the state-vector of a dynamical system do not have an explicit solution. This happens for instance when the process noise and the noise of the output measurement do not follow a Gaussian distribution. In that case approximation through Monte-Carlo methods can be used (Thrun wt al. 2005). A sampling of size N is assumed, i.e. N i.i.d. (independent identically distributed) variables $\xi^1, \xi^2, \cdots, \xi^N$. This sampling follows the p.d.f. $p(x)$ i.e. $\xi^{1:N} \sim p(x)$. Instead of $p(x)$ the function $p(x) \simeq p^N(x) = \frac{1}{N}\sum_{i=1}^{N}\delta_{\xi^i}(x)$ can be used. It is assumed that all points ξ^i have an equal weighted contribution to the approximation of $p(x)$. A more general approach would be if weight factors were assigned to the points ξ^i, which will also satisfy the normality condition $\sum_{i=1}^{N} w^i = 1$. In the latter case

$$p(x) \simeq p^N(x) = \sum_{i=1}^{N} w^i \delta_{\xi^i}(x) \tag{59}$$

If $p(\xi^i)$ is known then the probability $P(x)$ can be approximated using the discrete values of the p.d.f. $p(\xi^i) = w^i$. If sampling over the p.d.f. $p(x)$ is unavailable, then one can use a p.d.f. $\overline{p}(x)$ with similar support set, i.e. $p(x) = 0 \Rightarrow \overline{p}(x) = 0$. Then it holds $E(\phi(x)) = \int \phi(x) p(x) dx = \int \phi(x) \overline{p}(x) \frac{p(x)}{\overline{p}(x)} dx$. If the N samples of $\overline{p}(x)$ are available at the points $\tilde{\xi}^1 \cdots \tilde{\xi}^N$, i.e. $\overline{p}(\tilde{\xi})^i = \delta_{\tilde{\xi}^i}(x)$ and the weight coefficients w^i are defined as $w^i = \frac{p(\tilde{\xi}^i)}{\overline{p}(\tilde{\xi}^i)}$, then it is easily shown that

$$E(\phi(x)) \simeq \sum_{i=1}^{N} w^i \phi(\tilde{\xi}^i), \quad \text{where} \begin{cases} \tilde{\xi}^{1:N} \sim \overline{p}(x) \\ w^i = p(\tilde{x}^i) / \overline{p}(\tilde{x}^i) \end{cases} \tag{60}$$

The meaning of Eq. (60) is as follows: assume that the p.d.f. $p(x)$ is unknown (target distribution), however the p.d.f. $\overline{p}(x)$ (importance law) is available. Then, it is sufficient to sample on $\overline{p}(x)$ and find the associated weight coefficients w^i so as to calculate $E(\phi(x))$.

4.1.2 The prediction stage

As in the case of the Kalman Filter or the Extended Kalman Filter the particles filter consists of the measurement update (correction stage) and the time update (prediction stage) (Rigatos 2009b),(Thrun wt al. 2005). The prediction stage calculates $p(x(k)\,|\,Z^-)$ where Z^- = {$z(1)$, $z(2)$, $\cdots$, $z(n-1)$} according to Eq. (59). It holds that:

$$p(x(k-1)\,|\,Z^-) = \sum_{i=1}^{N} w_{k-1}^{i}\,\delta_{\xi_{k-1}^{i}}(x(k-1)) \tag{61}$$

while from Bayes formula it holds $p(x(k)\,|\,Z^-) = \int p(x(k)\,|\,x(k-1))p(x(k-1)\,|\,Z^-)dx$. Using also Eq. (61) one finally obtains

$$\begin{aligned} &p(x(k)\,|\,Z^-) = \sum_{i=1}^{N} w_{k-1}^{i}\,\delta_{\xi_{k^-}^{i}}(x(k)) \\ &\text{with}\ \ \xi_{k^-}^{i} \sim p(x(k)\,|\,x(k-1) = \xi_{k-1}^{i}) \end{aligned} \tag{62}$$

The meaning of Eq. (62) is as follows: the state equation of the system is executed N times, starting from the N previous values of the state vectors $x(k-1) = \xi_{k-1}^{i}$

$$\begin{aligned} \hat{x}(k+1) &= \phi(\hat{x}(k)) + L(k)u(k) + w(k) \\ z(k) &= \gamma(\hat{x}(k)) + v(k) \end{aligned} \tag{63}$$

Thus estimations of the current value of the state vector $\hat{x}(k)$ are obtained, and consequently the mean value of the state vector will be given from Eq. (62). This means that the value of the state vector which is calculated in the prediction stage is the result of the weighted averaging of the state vectors which were calculated after running the state equation, starting from the N previous values of the state vectors ξ_{k-1}^{i}.

4.1.3 The correction stage

The a-posteriori probability density is found using Eq. (62). Now a new position measurement $z(k)$ is obtained and the objective is to calculate the corrected probability density $p(x(k)\,|\,Z)$, where Z = {$z(1)$, $z(2)$, $\cdots$, $z(k)$}. From Bayes law it holds that $p(x(k)\,|\,Z) = \dfrac{p(Z\,|\,x(k))p(x(k))}{p(Z)}$ which can be also written as

$$p(x(k)\,|\,Z) = \frac{p(z(k)\,|\,x(k))p(x(k)\,|\,Z^-)}{\int p(z(k)\,|\,x(k),Z^-)p(x(k)\,|\,Z^-)dx} \tag{64}$$

Substituting Eq. (62) into Eq. (64) and after intermediate calculations one finally obtains

$$\begin{aligned} &p(x(k)\,|\,Z) = \sum_{i=1}^{N} w_k^i\,\delta_{\xi_{k^-}^{i}}(x(k)) \\ &\text{where} \qquad w_k^i = \frac{w^i_{k^-}\,p(z(k)\,|\,x(k) = \xi_{k^-}^{i})}{\sum_{j=1}^{N} w^j_{k^-}\,p(z(k)\,|\,x(k) = \xi_{k^-}^{j})} \end{aligned} \tag{65}$$