The scaled unscented transformation — **Source link** ↗

Simon Julier

Related papers:

- Unscented filtering and nonlinear estimation

- A new method for the nonlinear transformation of means and covariances in filters and estimators

- New extension of the Kalman filter to nonlinear systems

- The unscented Kalman filter for nonlinear estimation

- A new approach for filtering nonlinear systems

The Scaled Unscented Transformation

Simon J. Julier, IDAK Industries,

901 Missouri Blvd., #179

Jefferson City, MO 65109

E-mail:`sjulier@idak.com`

Abstract— **This paper describes a generalisation of the unscented transformation (UT) which allows sigma points to be scaled to an arbitrary dimension. The UT is a method for predicting means and covariances in nonlinear systems. A set of samples are deterministically chosen which match the mean and covariance of a (not necessarily Gaussian-distributed) probability distribution. These samples can be scaled by an arbitrary constant. The method guarantees that the mean and covariance second order accuracy in mean and covariance, giving the same performance as a second order truncated filter but without the need to calculate any Jacobians or Hessians. The impacts of scaling issues are illustrated by considering conversions from polar to Cartesian coordinates with large angular uncertainties.**

Keywords: Kalman filter, non-linear estimation, unscented filtering.

I. Introduction

One of the most fundamental tasks in filtering and estimation is to calculate the statistics of a random variable which has undergone a transformation. The Kalman filter, for example, uses two such transformations to predict the state future state of a system and the measurements which a suite of sensors will make of that system. When the transformation is nonlinear no general closed-form solutions exist [1] and many approximations have been proposed [2–6]. Probably the most widely used estimator for nonlinear systems is the extended Kalman filter (EKF) [7, 8]. The EKF applies the Kalman filter to nonlinear systems by simply linearising all the nonlinear models so that the traditional linear Kalman filter equations can be applied. However, in practice, the EKF has two well-known drawbacks. First, linearisation can produce highly unstable filters if the assumptions of local linearity is violated [9]. Second, the derivation of the Jacobian matrices are nontrivial in most applications and often lead to significant implementation difficulties.

In [10] and [11] we introduced a new approximate method for propagating means and covariances through nonlinear transformations called the *unscented transformation*. A set of weighted *sigma points* are deterministically chosen so that certain properties of these points (such as their first two moments) match those of the prior distribution. Each point undergoes the nonlinear transformation and the properties of the transformed set are calculated. Although this algorithm superficially resembles a Monte Carlo method, *no* random sampling is used and, in consequence, only a small number of points ($2n + 1$ for an n-dimensional space)are required. In subsequent work we have developed other sigma point selection schemes which exploit more information such as the first three moments of an arbitrary distribution [12] or the first four non-zero moments of a Gaussian distribution [13].

However, all of these sigma point solutions share the property that as the dimension of the state space increases, the radius of the sphere that bounds all the sigma points increases as well. Even though the specified information is still captured correctly (i.e., the mean and covariance of the sigma points matches the aprior distribution for all dimensions), it does so at the cost of sampling non-local effects. For many kinds of nonlinearities (such as exponents or trigonometric functions) this can lead to significant difficulties. In [14] we proposed a method for overcoming these difficulties through the use of negative weights and a "modified" form of the algorithm to guarantee positive semidefiniteness. However, the approach was developed from studying the higher order properties of the system and no physical intuition was used. Second, it was only developed to study the problem of point scaling for the specific set introduced in [11] and its applicability to other sigma point sets was not examined.

This paper re-examines the problem of sigma point scaling and introduces a new, general framework. Called the *scaled unscented transformation*, the method allows any set of sigma points to be scaled by an arbitrary scaling factor in such a manner that the first two moments of the set are preserved. It is equivalent to applying the conventional unscented transformation followed by a simple post-processing step. The storage and computational costs are exactly the same as a non-scaled version of the same transformation. The method can also be used to partially incorporate contributions higher order information into the estimates.

The structure of this paper is as follows. The problem is stated in Section II and the unscented transformation is described. Methods for sigma point scaling are examined in Section III and two complementary forms are derived. The first form uses an *auxillary random variable* — the nonlinear transformation is modified but the sigma point set is not. In this form it is easy to prove a number of properties including second order accuracy in mean and covariance predictions and also clearly shows the condition underwhich the predicted covariance is guaranteed to be positive semidefinite. We then derive the *scaled unscented transformation* which has the same properties as the auxillary form but modifies the sigma points themselves rather than the nonlinear transformation. We also show how some higher order information can be incorporated into the scaled transformation [15]. Conclusions are drawn in Section V.

II. Background

A. Problem Statement

Let $\mathbf{x}$ be an n-dimensional random variable with mean $\bar{\mathbf{x}}$ and covariance $\mathbf{P}_{xx}$. A second random variable, $\mathbf{y}$ is related to $\mathbf{x}$ through the nonlinear transformation

$$\mathbf{y} = \mathbf{f}\left[\mathbf{x}\right]. \tag{1}$$

The objective is to calculate the mean $\bar{\mathbf{y}}$ and covariance $\mathbf{P}_{yy}$ of $\mathbf{y}$.

Throughout this paper, we utilise to the Taylor Series expansion of Equation 1. Let $\mathbf{x} = \boldsymbol{\delta}\mathbf{x} + \bar{\mathbf{x}}$ where $\boldsymbol{\delta}\mathbf{x}$ is a zero mean random variable with covariance $\mathbf{P}_{xx}$. Expanding $\mathbf{f}\left[\cdot\right]$ about $\bar{\mathbf{x}}$,

$$\mathbf{f}\left[\mathbf{x}\right] = \mathbf{f}\left[\bar{\mathbf{x}} + \boldsymbol{\delta}\mathbf{x}\right] = \mathbf{f}\left[\bar{\mathbf{x}}\right] + \boldsymbol{\nabla}\mathbf{f}\boldsymbol{\delta}\mathbf{x} + \frac{1}{2}\boldsymbol{\nabla}^2\mathbf{f}\boldsymbol{\delta}\mathbf{x}^2 + \frac{1}{3!}\boldsymbol{\nabla}^3\mathbf{f}\boldsymbol{\delta}\mathbf{x}^3 + \cdots \tag{2}$$

where, for the sake of simplicity, we use the informal notation that $\boldsymbol{\nabla}^i\mathbf{f}\boldsymbol{\delta}\mathbf{x}^i$ is the ith order term in the multidimensional Taylor

Box II.1: The Point Selection Algorithm for the Simplex Unscented Transform.

Series. Taking expectations, it can be shown that

$$\begin{aligned} \bar{\mathbf{y}} &= \mathrm{E}\left[\mathbf{y}\right] \\ &= \mathbf{f}\left[\bar{\mathbf{x}}\right] + \frac{1}{2}\boldsymbol{\nabla}^2\mathbf{f}\,\mathbf{P}_{xx} + \frac{1}{6}\boldsymbol{\nabla}^3\mathbf{f}\,\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^3\right] + \cdots \end{aligned} \tag{3}$$

$$\begin{aligned} \mathbf{P}_{yy} &= \mathrm{E}\left[\left(\mathbf{y} - \bar{\mathbf{y}}\right)\left(\mathbf{y} - \bar{\mathbf{y}}\right)^T\right] \\ &= \boldsymbol{\nabla}\mathbf{f}\,\mathbf{P}_{xx}(\boldsymbol{\nabla}\mathbf{f})^T + \frac{1}{4}\boldsymbol{\nabla}^2\mathbf{f}\,\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^3\right](\boldsymbol{\nabla}\mathbf{f})^T + \frac{1}{2}\boldsymbol{\nabla}\mathbf{f}\,\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^3\right]\left(\boldsymbol{\nabla}^2\mathbf{f}\right)^T \\ &\quad + \frac{1}{2}\boldsymbol{\nabla}^2\mathbf{f}\left(\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^4\right] - \mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^2\mathbf{P}_{xx}\right] - \mathrm{E}\left[\mathbf{P}_{xx}\boldsymbol{\delta}\mathbf{x}^2\right] + \mathbf{P}_{xx}^2\right)(\boldsymbol{\nabla}^2\mathbf{f})^T \\ &\quad + \frac{1}{3!}\boldsymbol{\nabla}^3\mathbf{f}\,\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^4\right](\boldsymbol{\nabla}\mathbf{f})^T + \cdots \end{aligned} \tag{4}$$

The Unscented Transform builds on the principle that it is easier to approximate a probability distribution than it is to approximate an arbitrary nonlinear function. A set of $p+1$ weighted points $\mathcal{S} = \{\mathcal{W}_\rangle, \boldsymbol{\mathcal{X}}_\rangle\}$ (such that $\sum_{i=0}^{p} W_i = 1$) are chosen to reflect certain properties of $\mathbf{x}$ [13]. Once the set has been derived, the prediction method is straightforward. First, each point is instantiated through the nonlinear function, $\boldsymbol{\mathcal{Y}}_i = \mathbf{f}\left[\boldsymbol{\mathcal{X}}_i\right]$. The estimated mean and covariance of $\mathbf{y}$ are then

$$\bar{\mathbf{y}} = \sum_{i=0}^{p} W_i \boldsymbol{\mathcal{Y}}_i \tag{5}$$

$$\mathbf{P}_{yy} = \sum_{i=0}^{p} W_i \{\boldsymbol{\mathcal{Y}}_i - \bar{\mathbf{y}}\} \{\boldsymbol{\mathcal{Y}}_i - \bar{\mathbf{y}}\}^T . \tag{6}$$

The difficulties of higher dimensions are clearly illustrated in the simplex set of sigma points which were derived in [16] and are listed in Box II.1. This set utilises the minimum number of points $(n+1)$ required to match the mean and covariance of an n-dimensional random variable.

The simplex set of points have two related difficulties. The first difficulty is that the distance of the ith point from the origin is a function of $2^{(n-1/2)}$. Therefore, as n increases, the radius of the sphere which bounds the sigma points increases as well. Second, these points are asymmetrically distributed about the origin. Therefore, higher order effects such as the skew become more significant as the dimension increases.

III. Sigma Point Scaling Methods

The sigma point scaling methods attempt to overcome dimensional scaling effects by calculating the transformation of a scaled set of sigma points of the form

$$\boldsymbol{\mathcal{X}}'_i = \boldsymbol{\mathcal{X}}_0 + \alpha(\boldsymbol{\mathcal{X}}_i - \boldsymbol{\mathcal{X}}_0), \tag{7}$$

where α is a positive scaling parameter which can be made arbitrarily small to minimise higher order effects. Any permissible formulation should have the following two properties. First, for all choices of α the predicted covariance should be guaranteed to be positive semidefinite. Second, the second order accuracy in both the mean and covariance are preserved. We now describe two formulations for achieving this objective.

A. The Auxillary Random Variable

The *auxillary random variable* formulation considers a transformation which is related to the problem stated in Equation 1. Specifically, we consider the problem of estimating the mean $\bar{\mathbf{z}}$ and covariance $\mathbf{P}_{zz}$ of the auxillary random variable $\mathbf{z}$. It is related to $\mathbf{x}$ through the nonlinear equation $\mathbf{z} = \mathbf{g}\left[\mathbf{x}, \bar{\mathbf{x}}, \alpha, \mu\right]$ where

$$\mathbf{g}\left[\mathbf{x}, \bar{\mathbf{x}}, \alpha, \mu\right] = \frac{\mathbf{f}\left[\bar{\mathbf{x}} + \alpha(\mathbf{x} - \bar{\mathbf{x}})\right] - \mathbf{f}\left[\bar{\mathbf{x}}\right]}{\mu} + \mathbf{f}\left[\bar{\mathbf{x}}\right]. \tag{8}$$

α is a positive point scaling parameter and μ is a normalisation term which scales the transformed point about $\mathbf{f}\left[\bar{\mathbf{x}}\right]$ to offset the effects of α. Because all sigma points are propagated through the term $\mathbf{f}\left[\bar{\mathbf{x}} + \alpha(\mathbf{x} - \bar{\mathbf{x}})\right]$, the scaling effect of Equation 7 is implicitly achieved. To prove the second order accuracy of this form, we consider the role played by α and μ. Taking a Taylor Series expansion of $\mathbf{g}\left[\cdot, \cdot, \cdot, \cdot\right]$ about $\bar{\mathbf{x}}$,

$$\mathbf{g}\left[\mathbf{x}, \bar{\mathbf{x}}, \alpha, \mu\right] = \mathbf{f}\left[\bar{\mathbf{x}}\right] + \boldsymbol{\nabla}\mathbf{f}\frac{\alpha}{\mu}\boldsymbol{\delta}\mathbf{x} + \frac{1}{2}\boldsymbol{\nabla}^2\mathbf{f}\frac{\alpha^2}{\mu}\boldsymbol{\delta}\mathbf{x}^2 + \frac{1}{3!}\boldsymbol{\nabla}^3\mathbf{f}\frac{\alpha^3}{\mu}\boldsymbol{\delta}\mathbf{x}^3 + \cdots \tag{9}$$

Taking expectations, the mean of $\mathbf{z}$ is

$$\bar{\mathbf{z}} = \mathbf{f}\left[\bar{\mathbf{x}}\right] + \frac{1}{2}\boldsymbol{\nabla}^2\mathbf{f}\frac{\alpha^2}{\mu}\mathbf{P}_{xx} + \frac{1}{6}\boldsymbol{\nabla}^3\mathbf{f}\frac{\alpha^3}{\mu}\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^3\right] + \cdots \tag{10}$$

These terms can be related directly to those of the Taylor Series expansion of $\mathbf{y}$. If $\mu = \alpha^2$, the expressions for $\bar{\mathbf{y}}$ and $\bar{\mathbf{z}}$ agree up to the second order. The ratio of the third and higher order terms of $\bar{\mathbf{z}}$ against $\bar{\mathbf{y}}$ scale geometrically with a commmon ratio of α. Since α only affects the third and higher orders, its value can be chosen so that the scaling effects in the higher order terms are minimised. With a sufficiently small value of α, the same mean can be calculated as with the modified form of the unscented transformation.

A similar result holds for the covariance. Let $\mathbf{P}_{zz}^* = \mu\mathbf{P}_{zz}$. Taking expectations,

$$\begin{aligned} \mathbf{P}_{zz}^* &= \frac{\alpha^2}{\mu}\boldsymbol{\nabla}\mathbf{f}\,\mathbf{P}_{xx}(\boldsymbol{\nabla}\mathbf{f})^T + \frac{\alpha^3}{\mu^2}\frac{1}{2}\boldsymbol{\nabla}\mathbf{f}\,\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^3\right]\left(\boldsymbol{\nabla}^2\mathbf{f}\right)^T \\ &\quad + \frac{\alpha^4}{\mu^2}\frac{1}{2}\boldsymbol{\nabla}^2\mathbf{f}\left(\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^4\right] - \mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^2\mathbf{P}_{yy}\right] - \mathrm{E}\left[\mathbf{P}_{yy}\boldsymbol{\delta}\mathbf{x}^2\right] + \mathbf{P}_{yy}^2\right)(\boldsymbol{\nabla}^2\mathbf{f})^T \\ &\quad + \frac{\alpha^4}{\mu^2}\frac{1}{3!}\boldsymbol{\nabla}^3\mathbf{f}\,\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^4\right](\boldsymbol{\nabla}\mathbf{f})^T + \cdots \end{aligned} \tag{11}$$

When $\mu = \alpha^2$, the expansion of $\mathbf{P}_{zz}^*$ agrees with $\mathbf{P}_{yy}$ up to the second order and the third and higher order terms scale with α.

The *auxillary form of the unscented transformation* simply applies the unscented transformation to the problem of estimating the mean and covariance of the auxillary random variable. Given an n-dimensional random variable $\mathbf{x}$ with mean $\bar{\mathbf{x}}$ and covariance $\mathbf{P}_{xx}$, a set of $p+1$ sigma points are chosen such that the

mean and covariance of those points are $\bar{\mathbf{x}}$ and $\mathbf{P}_{xx}$ respectively. The unscented transformation is then

$$\boldsymbol{\mathcal{Z}}_i = \frac{\mathbf{f}\left[\bar{\mathbf{x}} + \alpha(\boldsymbol{\mathcal{X}}_i - \bar{\mathbf{x}})\right] - \mathbf{f}\left[\bar{\mathbf{x}}\right]}{\alpha^2} + \mathbf{f}\left[\bar{\mathbf{x}}\right] \tag{12}$$

$$\bar{\mathbf{z}} = \sum_{i=0}^{p} W_i \boldsymbol{\mathcal{Z}}_i \tag{13}$$

$$\mathbf{P}_{zz}^* = \alpha^2 \sum_{i=0}^{p} W_i \left\{\boldsymbol{\mathcal{Y}}_i - \bar{\mathbf{z}}\right\} \left\{\boldsymbol{\mathcal{Y}}_i - \bar{\mathbf{z}}\right\}^T \tag{14}$$

From Equations 10 and 11 and given the fact that the sigma points have mean $\bar{\mathbf{x}}$ and covariance $\mathbf{P}_{xx}$, $\bar{\mathbf{z}}$ and $\mathbf{P}_{zz}^*$ are correct to the second order for any value of α. Furthermore, because $\mathbf{P}_{zz}^*$ is calculated from the weighted outer products of vectors, positive semidefiniteness is guaranteed if all of the weights W_i are non-negative. Since the only motive for choosing negative values of W_i is to scale the points, this incentive is removed.

The auxillary form is able to meet the requirements set out at the beginning of this section. However, it requires a change in the fundamental transformation system itself. We now show that it is possible to leave the original problem in place but apply a transformation to the sigma points themselves.

B. The Scaled Unscented Transform

The scaled unscented transform yields the same results as the auxillary form, but without the need to modify the transformation (Equation 1). Rather, an initial set of points are chosen using a normal sigma point selection algorithm. A specific transformation is applied to these points. The mean and covariance are calculated using Equations 5 and 6. A final term is added to offset the initial transformation which was applied to the sigma points.

Suppose a set of sigma points $\mathcal{S}$ have been constructed with mean $\bar{\mathbf{x}}$ and covariance $\mathbf{P}_{xx}$ and a positive scaling parameter α has been chosen. These points are transformed to a new set $\mathcal{S}' = \{\} = \prime, \infty, \ldots, \sqrt{} : \boldsymbol{\mathcal{X}}_\rangle', \mathcal{W}_\rangle'\}$ which has the same mean and covariance as $\mathcal{S}$ but the points now obey the condition of Equation 7. As a result, the weights of this transformed sequence are

$$W_i' = \begin{cases} W_0/\alpha^2 + (1 - 1/\alpha^2) & i = 0 \\ W_i/\alpha^2 & i \neq 0 \end{cases} \tag{15}$$

The proof can be found in the Appendix. Because $\mathcal{S}'$ is, itself, a sigma point set, it is possible with some selection algorithms to implicitly combine the scaling directly with the original sigma point selection.

Given this set of points, the scaled unscented transform calculates its statistics as follows:

$$\boldsymbol{\mathcal{Y}}_i' = \mathbf{f}\left[\boldsymbol{\mathcal{X}}_i'\right]. \tag{16}$$

$$\bar{\mathbf{y}}' = \sum_{i=0}^{p} W_i' \boldsymbol{\mathcal{Y}}_i'. \tag{17}$$

$$\mathbf{P}_{yy}' = \sum_{i=0}^{p} W_i' \left\{\boldsymbol{\mathcal{Y}}_i' - \bar{\mathbf{y}}\right\} \left\{\boldsymbol{\mathcal{Y}}_i' - \bar{\mathbf{y}}\right\}^T + (1 - \alpha^2) \left\{\boldsymbol{\mathcal{Y}}_0' - \bar{\mathbf{y}}\right\} \left\{\boldsymbol{\mathcal{Y}}_0' - \bar{\mathbf{y}}\right\}^T. \tag{18}$$

In the Appendix we prove that, for any sigma point distribution, $\bar{\mathbf{y}}' = \bar{\mathbf{z}}$ and $\mathbf{P}_{yy}' = \mathbf{P}_{zz}^*$ when $\mu = \alpha^2$. This has a number of important consequences. First, the scaled unscented transformation possesses all of the properties of the auxillary form. The predicted mean and covariance are accurate to the second order and $\mathbf{P}_{yy}'$ is guaranteed to be positive semidefinite if all of the *untransformed* weights are non-negative. Second, the numerical costs of this form are the same as with the unscaled unscented transform. Comparing Equation 18 to Equation 6, the only difference is that a term $(1 - \alpha^2)$ is added to the weight on the zeroth sigma point. Finally this form provides a very simple

interpretation for α. When $\alpha = 1$, this gives Equation 6. When $\alpha = 0$, this form gives the modified form of the covariance equation which was used in [14].

C. Incorporating Higher Order Information

Although the sigma points only capture the first two moments of the sigma points (and so the first two moments of the Taylor Series expansion), the scaled unscented can be extended to include partial higher order information of the fourth order term in the Taylor Series expansion of the covariance [15]. The fourth order term of Equation 4 is

$$\mathbf{A} = \frac{1}{4}\boldsymbol{\nabla}^2\mathbf{f}\left(\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^4\right] - \mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^2\mathbf{P}_{yy}\right] - \mathrm{E}\left[\mathbf{P}_{yy}\boldsymbol{\delta}\mathbf{x}^2\right] + \mathbf{P}_{yy}^2\right)(\boldsymbol{\nabla}^2\mathbf{f})^T$$
$$+ \frac{1}{3!}\boldsymbol{\nabla}^3\mathbf{f}\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^4\right](\boldsymbol{\nabla}\mathbf{f})^T. \tag{19}$$

The term $\frac{1}{2}\boldsymbol{\nabla}^2\mathbf{f}\mathbf{P}_{yy}^2(\boldsymbol{\nabla}^2\mathbf{f})^T$ can be calculated from the same set of sigma points which match the mean and covariance. From Equations 2 and 3,

$$\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}} = \frac{1}{2}\boldsymbol{\nabla}^2\mathbf{f}\,\mathbf{P}_{xx} + \frac{1}{6}\boldsymbol{\nabla}^3\mathbf{f}\,\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^3\right] + \cdots.$$

Taking outer products,

$$(\bar{\mathbf{y}} - \boldsymbol{\mathcal{Y}}_0)(\bar{\mathbf{y}} - \boldsymbol{\mathcal{Y}}_0)^T = \frac{1}{4}\boldsymbol{\nabla}^2\mathbf{f}\mathbf{P}_{yy}^2(\boldsymbol{\nabla}^2\mathbf{f})^T + \cdots$$

Therefore, adding extra weighting to the contribution of the zeroth point, further higher order effects can be incorporated at no additional computational cost by rewriting Equation 18 as

$$\mathbf{P}_{yy}' = \sum_{i=0}^{p} W_i' \left\{\boldsymbol{\mathcal{Y}}_i' - \bar{\mathbf{y}}\right\} \left\{\boldsymbol{\mathcal{Y}}_i' - \bar{\mathbf{y}}\right\}^T + (\beta + 1 - \alpha^2) \left\{\boldsymbol{\mathcal{Y}}_0' - \bar{\mathbf{y}}\right\} \left\{\boldsymbol{\mathcal{Y}}_0' - \bar{\mathbf{y}}\right\}^T.$$

In this form, the error in the fourth order term is

$$\Delta\mathbf{A} = \frac{1}{4}\boldsymbol{\nabla}^2\mathbf{f}\left(\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^4\right] - \mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^2\mathbf{P}_{yy}\right] - \mathrm{E}\left[\mathbf{P}_{yy}\boldsymbol{\delta}\mathbf{x}^2\right] + (1 - \beta)\mathbf{P}_{yy}^2\right)(\boldsymbol{\nabla}^2\mathbf{f})^T$$
$$+ \frac{1}{3!}\boldsymbol{\nabla}^3\mathbf{f}\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^4\right](\boldsymbol{\nabla}^2\mathbf{f})^T. \tag{20}$$

In the special case that $\mathbf{x}$ Gaussian-distributed, $\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^4\right] = 3\mathbf{P}_{xx}^2$ and so the error is

$$\Delta\mathbf{A} = (2 - \beta)\boldsymbol{\nabla}^2\mathbf{f}\mathbf{P}_{yy}^2(\boldsymbol{\nabla}^2\mathbf{f})^T + \frac{1}{3!}\boldsymbol{\nabla}^3\mathbf{f}\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^4\right](\boldsymbol{\nabla}\mathbf{f})^T. \tag{21}$$

Under the assumption that no information about $\mathbf{f}\left[\cdot\right]$ is used, this term is minimised when $\beta = 2$.

D. Summary

The scaled unscented transformation can be written as

$$\boldsymbol{\mathcal{X}}_i' = \boldsymbol{\mathcal{X}}_0 + \alpha(\boldsymbol{\mathcal{X}}_i - \boldsymbol{\mathcal{X}}_0) \tag{22}$$

$$\boldsymbol{\mathcal{Y}}_i' = \mathbf{f}\left[\boldsymbol{\mathcal{X}}_i'\right] \tag{23}$$

$$W_i' = \begin{cases} W_0/\alpha^2 + (1/\alpha^2 - 1) & i = 0 \\ W_i/\alpha^2 & i \neq 0 \end{cases} \tag{24}$$

$$\bar{\mathbf{y}} = \sum_{i=1}^{p} W_i' \boldsymbol{\mathcal{Y}}_i' \tag{25}$$

$$\mathbf{P}_{yy} = \sum_{i=0}^{p} W_i' \left\{\boldsymbol{\mathcal{Y}}_i' \prime - \bar{\mathbf{y}}\right\} \left\{\boldsymbol{\mathcal{Y}}_i' \prime - \bar{\mathbf{y}}\right\}^T \tag{26}$$

$$+ (W_0 + 1 + \beta - \alpha^2) \left\{\boldsymbol{\mathcal{Y}}_0' - \bar{\mathbf{y}}\right\} \left\{\boldsymbol{\mathcal{Y}}_0' - \bar{\mathbf{y}}\right\}^T. \tag{27}$$

IV. EXAMPLE

Suppose a mobile robot detects a beacon in its environment using a range-optimised sonar sensor. The sensor returns polar information (range r and bearing θ) and this is to be converted to estimate to Cartesian coordinates. The transformation is:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} \text{ with } \boldsymbol{\nabla}\mathbf{f} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}.$$

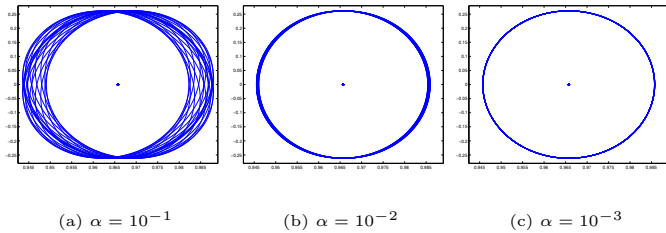(a) $\alpha = 10^{-1}$ (b) $\alpha = 10^{-2}$ (c) $\alpha = 10^{-3}$

Fig. 1. $\bar{\mathbf{y}}$ and $\mathbf{P}_{yy}$ calculated using the simplex unscented algorithm and 20 different orientations of the sigma points. In all of these plots, $\beta = 0$.

The simplex points for a two-dimensional space consists of an isoceles triangle and a further point at the origin. Rotating the points about the origin by an angle ϕ does not affect their first or second moments. However, the rotation affects the third and higher moments which, through the nonlinear transformation, affect the predicted mean and covariance. The effect is illusrated in Figure 1(a) which shows the mean and the 1σ contours for $\bar{\mathbf{y}}$ and $\mathbf{P}_{yy}$ calculated for 20 different orientations of the sigma points. For this plot, α (the scaled unscented scaling parameter defined in Equation 22) has the value 10^{-1}. As can be seen, the value of $\bar{\mathbf{y}}$ is hardly affected by the point orientation. However, the covariance does show significant variation, confirming the statement in [17]. These higher order effects can be greatly reduced by decreasing the value of α. Figures 1(b) and 1(c) show the results of the same trials when $\alpha = 10^{-2}$ and $\alpha = 10^{-3}$. In the last value, the effects of sigma point orientation are minimal (the maximum change in the covariance is less than 0.09%). Therefore, this study has shown that the scaled sigma points directly contributes to the use of the simplex by eliminating higher order orientation dependent effects.

To investigate the accuracy of the simplex method, its results were compared with the *true* results (calculated by a Monte Carlo of 3.5×10^6 randomly drawn samples) and the results calculated through linearisation. The means and covariance ellipses for these three approaches are shown in Figure 2. The linearised estimate contains significant errors. Its mean is biased in the x–direction and, furthermore, its covariance is "too small". Whenever the filter updates with this estimate, it introduces unmodelled biases and correlations that can undermine filter stability. The simplex transform yields better results. When $\beta = 0$, its mean is unbiased. However, its covariance is similar to that calculated by linearisation. This is a direct consequence of the fact that the simplex only captures the first two moments of the mean and covariance correctly. However, the performance of this algorithm can be significantly improved by exploiting the fact that additional higher order information can be readily incorporated into the estimate. Because the distributions are Gaussian, the analysis from [16] shows that the information gained by the scaled unscented weight parameter β (defined in Equation 27) is optimised when $\beta = 2$. This is confirmed in Figure 2 where it can be seen that the resulting estimate is, in fact, slightly conservative.

V. Discussion and Conclusions

This paper has presented and analysied the scaled unscented transform. This new parameterisation permits a set of sigma points to be scaled using an arbitrary scaling parameter α. Second order accuracy is maintained and the algorithm is guaranteed to give a positive semi-definite covariance if all of the weights on the sigma points are non-negative. Furthermore, its
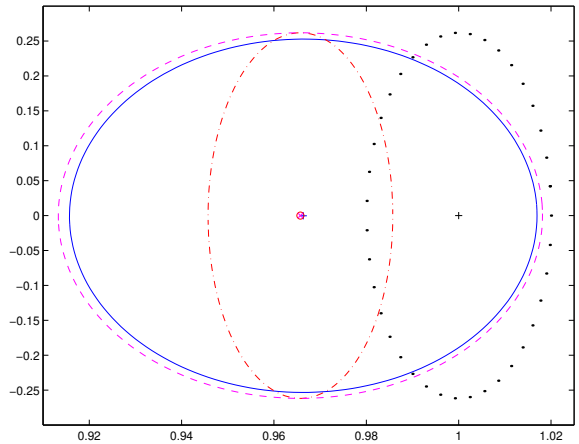

Fig. 2. The means and 1σ contours calculated by different prediction algorithms. The Monte Carlo "true" solution has a mean + at (0.965,0) and its covariance is the solid ellipse. Linearisation yields the estimate + at (1,0) with the dotted ellipse. The simplex transform with $\beta = 0$ gives the mean ○ with the dot-dashed ellipse. The simplex transform with $\beta = 2$ has mean × and covariance the dashed line.

computational costs are exactly the same as those of the original formulation of the uscented transform.

This parameterisation provides a framework which can be applied with any sigma point distribution.

Appendix

This Appendix shows that the *any* scaling strategy of the form of Equation 8 can be written as an application of the straightforward method plus a post-processing term. This means that the equation has exactly the same number of calculations as conventional unscented. We proceed by showing the equivalence of the weights, means and covariances respectively.

Theorem 1: The weights of the $\mathcal{S}'$ are related to those of $\mathcal{S}$ by Equation 15.

Proof: The normalisation and covariance conditions obeyed by $\mathcal{S}$ are

$$\sum_{i=0}^{p} W_i = 1 \tag{28}$$

$$\sum_{i=1}^{p} W_i (\boldsymbol{\mathcal{X}}_i - \bar{\mathbf{x}})(\boldsymbol{\mathcal{X}}_i - \bar{\mathbf{x}})^T = \mathbf{P}_{xx} \tag{29}$$

where the fact that $\boldsymbol{\mathcal{X}}_0 = \bar{\mathbf{x}}$ has been used. The conditions obeyed by $\mathcal{S}'$ are

$$\sum_{i=0}^{p} W_i' = 1 \tag{30}$$

$$\sum_{i=1}^{p} W_i' (\boldsymbol{\mathcal{X}}_i' - \bar{\mathbf{x}})(\boldsymbol{\mathcal{X}}_i' - \bar{\mathbf{x}})^T = \mathbf{P}_{xx} \tag{31}$$

Comparing Equations 29 with 31 and substituting from Equation 7, it can be seen that $W_i = W_i' \alpha^2$ for $i > 0$. W_0 is found from Equations 28 and 30,

$$\begin{aligned}
1 = \sum_{i=0}^{p} W_i &= W_0 + \sum_{i=1}^{p} W_i \\
&= W_0 + \alpha^2 \sum_{i=1}^{p} W_i' \\
&= W_0 + \alpha^2 (1 - W_0')
\end{aligned} \tag{32}$$

∎

Each scaled unscented sigma point is $\boldsymbol{\mathcal{Y}}_i' = \mathbf{f}\left[\boldsymbol{\mathcal{X}}_i'\right]$, whereas $\boldsymbol{\mathcal{Z}}_i$ is given by Equation 8,

$$\begin{aligned}
\boldsymbol{\mathcal{Z}}_i &= \mathbf{g}\left[\boldsymbol{\mathcal{X}}_i, \bar{\mathbf{x}}, \alpha, \mu\right] \\
&= \left(1 - \frac{1}{\mu}\right)\mathbf{f}\left[\boldsymbol{\mathcal{X}}_0'\right] + \frac{1}{\mu}\mathbf{f}\left[\boldsymbol{\mathcal{X}}_i'\right] \\
&= \left(1 - \frac{1}{\mu}\right)\boldsymbol{\mathcal{Y}}_0 + \frac{1}{\mu}\boldsymbol{\mathcal{Y}}_i'.
\end{aligned} \tag{33}$$

Theorem 2: Let

$$\bar{\mathbf{z}} = \sum_{i=0}^{p} W_i \boldsymbol{\mathcal{Z}}_i, \quad \bar{\mathbf{y}}' = \sum_{i=0}^{p} W_i' \boldsymbol{\mathcal{Y}}_i$$

Then

$$\bar{\mathbf{z}} = \frac{\mu - \alpha^2}{\mu}\boldsymbol{\mathcal{Y}}_0 + \frac{\alpha^2}{\mu}\bar{\mathbf{y}}'.$$

Proof: Substituting from Equations 15 and 33 and using the fact that $\sum_{i=0}^{p} W_i = 1$,

$$\begin{aligned}
\bar{\mathbf{z}} &= \left(1 - \frac{1}{\mu}\right)\boldsymbol{\mathcal{Y}}_0 + \frac{1}{\mu}\sum_{i=0}^{p} W_i \boldsymbol{\mathcal{Y}}_i \\
&= \left(\frac{\mu-1}{\mu}\right)\boldsymbol{\mathcal{Y}}_0 + \frac{1-\alpha^2}{\mu}\boldsymbol{\mathcal{Y}}_0 + \frac{\alpha^2}{\mu}\sum_{i=0}^{p} W_i'\boldsymbol{\mathcal{Y}}_i \\
&= \frac{\mu-\alpha^2}{\mu}\boldsymbol{\mathcal{Y}}_0 + \frac{\alpha^2}{\mu}\bar{\mathbf{y}}'
\end{aligned} \tag{34}$$

Theorem 3: Let

$$\mathbf{P}_{zz}^* = \mu \sum_{i=0}^{p} W_i \left(\boldsymbol{\mathcal{Z}}_i - \bar{\mathbf{z}}\right)\left(\boldsymbol{\mathcal{Z}}_i - \bar{\mathbf{z}}\right)^T, \quad \mathbf{P}_{yy}' = \sum_{i=0}^{p} W_i' \left(\boldsymbol{\mathcal{Y}}_i' - \bar{\mathbf{y}}'\right)\left(\boldsymbol{\mathcal{Y}}_i - \bar{\mathbf{y}}'\right)^T.$$

Then

$$\mathbf{P}_{zz}^* = \frac{\alpha^2}{\mu}\left\{\mathbf{P}_{yy}' + (1-\alpha^2)(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')^T\right\}.$$

Proof: Substituting from Equations 33 and 34,

$$\boldsymbol{\mathcal{Z}}_i - \bar{\mathbf{z}} = \frac{1}{\mu}(\boldsymbol{\mathcal{Y}}_i - \bar{\mathbf{y}}') + \frac{(\alpha^2-1)}{\mu}(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}') \tag{35}$$

Therefore,

$$\begin{aligned}
\mathbf{P}_{zz}^* = \mu \sum_{i=0}^{p} W_i &\left\{\frac{1}{\mu}(\boldsymbol{\mathcal{Y}}_i - \bar{\mathbf{y}}') + \frac{(\alpha^2-1)}{\mu}(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')\right\} \\
&\times \left\{\frac{1}{\mu}(\boldsymbol{\mathcal{Y}}_i - \bar{\mathbf{y}}') + \frac{(\alpha^2-1)}{\mu}(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')\right\}^T \\
= \frac{1}{\mu}\sum_{i=0}^{p} W_i &\left\{(\boldsymbol{\mathcal{Y}}_i - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Y}}_i - \bar{\mathbf{y}}')^T + (\alpha^2-1)(\boldsymbol{\mathcal{Y}}_i - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')^T + \right. \\
&\left. (\alpha^2-1)(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Y}}_i - \bar{\mathbf{y}}')^T + (\alpha^2-1)^2(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')^T\right\}
\end{aligned} \tag{36}$$

From Equation 15,

$$\sum_{i=0}^{p} W_i(\boldsymbol{\mathcal{Y}}_i - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Y}}_i - \bar{\mathbf{y}}')^T = \alpha^2 \mathbf{P}_{yy} + (1-\alpha^2)(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')^T$$

$$\tag{37}$$

$$\sum_{i=0}^{p} W_i(\boldsymbol{\mathcal{Y}}_i - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')^T = (1-\alpha^2)(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')^T \tag{38}$$

$$\sum_{i=0}^{p} W_i(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')^T = (\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')^T \tag{39}$$

Substituting Equations 37 to 39 into Equation 36,

$$\mathbf{P}_{zz}^* = \frac{\alpha^2}{\mu}\left(\mathbf{P}_{yy}' + (1-\alpha^2)(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')^T\right).$$

$\blacksquare$

Remark 1: When $\mu = \alpha^2$,

$$\bar{\mathbf{z}} = \bar{\mathbf{y}}'$$
$$\mathbf{P}_{zz}^* = \mathbf{P}_{yy}' + (1-\alpha^2)(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}')^T.$$

REFERENCES

[1] H. J. Kushner, "Dynamical Equations For Optimum Non-linear Filtering," *Journal of Differential Equations*, vol. 3, pp. 179–190, 1967.

[2] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, San Diego, CA, 1970.

[3] P. S. Maybeck, *Stochastic Models, Estimation, and Control*, vol. 2, Academic Press, 1982.

[4] N. J. Gordon, D. J. Salmond and A. F. M. Smith, "Novel Approach to Nonlinear/non-Gaussian Bayesian State Estimation," *IEE Proceedings-F*, vol. 140, no. 2, pp. 107–113, April 1993.

[5] H. W. Sorenson and A. R. Stubberud, "Non-linear Filtering by Approximation of the *a posteriori* Density," *International Journal of Control*, vol. 8, no. 1, pp. 33–51, 1968.

[6] F. E. Daum, "New Exact Nonlinear Filters," in *Bayesian Analysis of Time Series and Dynamic Models*, J. C. Spall, Ed., chapter 8, pp. 199–226. Marcel Drekker, Inc., 1988.

[7] J. K. Uhlmann, "Algorithms for multiple target tracking," *American Scientist*, vol. 80, no. 2, pp. 128–141, 1992.

[8] H. W. Sorenson, Ed., *Kalman filtering: theory and application*, IEEE Press, Piscataway NJ, USA, 1985.

[9] D. Lerro and Y. K. Bar-Shalom, "Tracking with Debiased Consistent Converted Measurements vs. EKF," *IEEE Transactions on Aerospace and Electronics Systems*, vol. AES-29, no. 3, pp. 1015–1022, July 1993.

[10] S. J. Julier and J. K. Uhlmann, "A New Extension of the Kalman Filter to Nonlinear Systems," in *The Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls*, Orlando FL, USA, 1997, SPIE, Multi Sensor Fusion, Tracking and Resource Management II.

[11] S. J. Julier, J. K. Uhlmann and H. F. Durrant-Whyte, "A New Approach for Filtering Nonlinear Systems," in *The Proceedings of the American Control Conference, Seattle, Washington.*, 1995, pp. 1628–1632.

[12] S. J. Julier, "A Skewed Approach to Filtering," in *The Proceedings of AeroSense: The 12th International Symposium on Aerospace/Defense Sensing, Simulation and Controls*, Orlando FL, USA, April 1998, vol. 3373, pp. 54–65, SPIE, Signal and Data Processing of Small Targets.

[13] S. J. Julier and J. K. Uhlmann, "A Consistent, Debiased Method for Converting Between Polar and Cartesian Coordinate Systems," in *The Proceedings of AeroSense: Acquisition, Tracking and Pointing XI.* 1997, vol. 3086, pp. 110–121, SPIE.

[14] S. J. Julier, J. K. Uhlmann and H. F. Durrant-Whyte, "A New Approach for the Nonlinear Transformation of Means and Covariances in Linear Filters," *IEEE Transactions on Automatic Control*, vol. 5, no. 3, pp. 477–482, March 2000.

[15] J. K. Uhlmann, "A real time algorithm for simultaneous map building and localization," 1995.

[16] S. J. Julier and J. K. Uhlmann, "Simultaneous Localisation and Map Building Using Split Covariance Intersection," Submitted to IROS 2001, 2001.

[17] J. K. Uhlmann, *Dynamic Map Building and Localization for Autonomous Vehicles*, Ph.D. thesis, University of Oxford, 1995.