The Scaled Unscented Transformation

Simon J. Julier

Abstract

This paper describes the *scaled unscented transformation*, a new method of applying the unscented transform to a nonlinear system. A set of samples are deterministically chosen which match the mean and covariance of a (not necessarily Gaussian) probability distribution. Each point in the set is scaled by a user-specified constant. A method is derived which preserves the second order accuracy in mean and covariance, giving performance as good as second order truncated filter.

1 Introduction

One of the most fundamental tasks in filtering and estimation is to calculate the statistics of a random variable that has undergone a transformation. When the system models are non-linear no general closed-form solutions exist [11] and many approximations have been proposed [1–3, 12, 13].

In [7] and [9] 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 are required. In [7] we presented a symmetric sigma point solution that used $2n + 1$ points to match the mean and covariance of an n-dimensional random variable. With this set of points, the unscented transform guarantees the *same* performance as the truncated second order filter, with the *same* order of calculations as an extended Kalman filter (EKF) but *without* the need to calculate any approximations or derivatives [10]. 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 [5] or the first four non-zero moments of a Gaussian distribution [8].

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. Therefore, the specified information (such as mean and covariance) are captured with the cost of introducing errors due to non-local sampling. The originally unscented transformation overcame these difficulties using two methods. The first is to specify a point at the mean of the prior with a weight that can be positive, zero or negative. When this weight is negative, it causes the sigma points to be scaled into the origin. As a result, the first two moments are preserved and this is reflected by second order performance in the mean. However, although the mean is successfully calculated, non-positive weights can lead to a predicted covariance that is non-positive semidefinite. The second strategy was to adopt a *modified form* of the unscented transformation that added a positive semidefinite correction term. Although the point scaling approach has been successfully demonstrated in a number of systems [4], the theoretical development of the modified form has a number of shortcomings. First, the approach was introduced 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 in the symmetric unscented transformation, 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 for scaling sigma points. This approach introduces an *auxillary random variable* that is related to the original system equations. It has a number of desirable properties. First, as with the modified unscented transform the set of sigma points can be scaled but their first two moments are preserved. Second, it provides a general framework within which the conventional unscented and modified forms are limiting values. Third, the approach is proved to work with *any* sigma point set, and not just the symmetric set. Fourth, the method 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. Finally, at no extra cost the method can incorporate known higher order information (such as the kurtosis). Given its superior implementation properties, this form of the algorithm is superior whenever the unscented transformation is used to propagate the first two moments of a distribution. We have already used this method to develop minimal sigma point filters [6] as well as demonstrate a filter which propagates the skew (third order moments) [5].

The structure of this paper is as follows. In Section 2 we introduce the problem statement and the unscented transformation. The scaled unscented transformation is considered in Section 3. We present two complementary approaches. The first uses an auxillary random variable that introduces a modified form of the process model. This is shown to be equivalent to the second method, that uses the conventional unscented transformation with an additional postprocessing step. We also show that it is possible to incorporate additional higher order information through an extra term. Conclusions are drawn in Section 5.

2 Background

2.1 Problem Statement

Let $\mathbf{x}$ be a 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 refer to the Taylor Series expansion of this equation. 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 Series. Taking expectations, it can be shown that

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

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

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 $\mathcal{S} = \{W_i, \boldsymbol{\mathcal{X}}_i\}$ (such that $\sum_{i=0}^{p} W_i = 1$) are chosen to reflect certain properties of $\mathbf{x}$. In other words, they obey a condition of the form $\mathbf{g}\left[\mathcal{S}, p_x(\mathbf{x})\right] = \mathbf{0}$ where $\mathbf{g}\left[\cdot,\cdot\right]$ specifies *what* information from $\mathbf{x}$ is to be matched by $\mathcal{S}$ [8].

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

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

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

and other properties (such as the third order moments or *skew*) can be calculated accordingly [5].

In [7] we presented the following set of points that obey both the mean and covariance constraints:

$$\begin{aligned}
\boldsymbol{\mathcal{X}}_0(k \mid k) &= \hat{\mathbf{x}}(k \mid k) & W_0 &= \kappa/(n+\kappa) \\
\boldsymbol{\mathcal{X}}_i(k \mid k) &= \hat{\mathbf{x}}(k \mid k) + \left(\sqrt{(n+\kappa)\mathbf{P}(k \mid k)}\right)_i & W_i &= 1/\{2(n+\kappa)\} \\
\boldsymbol{\mathcal{X}}_{i+n}(k \mid k) &= \hat{\mathbf{x}}(k \mid k) - \left(\sqrt{(n+\kappa)\mathbf{P}(k \mid k)}\right)_i & W_{i+n} &= 1/\{2(n+\kappa)\}
\end{aligned} \tag{7}$$

where $\kappa \in \Re$, $\left(\sqrt{(n+\kappa)\mathbf{P}(k \mid k)}\right)_i$ is the ith row or column of the matrix square root of $(n+\kappa)\mathbf{P}(k \mid k)$ and W_i is the weight that is associated with the ith point. The distance of the ith point, $|\boldsymbol{\mathcal{X}}_i - \bar{\mathbf{x}}| = \alpha\sqrt{(n+\kappa)}$. The value of κ had a direct effect on the scaling of the points. When $\kappa = 0$, the distance of the ith sigma point from $\bar{\mathbf{x}}$ is proportional to $\sqrt{n}$. When $\kappa > 0$ the points are scaled further from $\bar{\mathbf{x}}$ and when $\kappa < 0$ the points are scaled towards the origin. If, for example, $\kappa = 3 - n$ the effect of n is cancelled out. However, when $\kappa < 0$ it is possible that the covariance, calculated by Equation 6, is non-positive semidefinite. The *modified form* of the the unscented transformation calculates the covariance according to

$$\mathbf{P}_{yy}^{MOD} = \mathbf{P}_{yy} + \{\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}\} \{\boldsymbol{\mathcal{Y}}_0 - \bar{\mathbf{y}}\}^T. \tag{8}$$

The justification for this form is that

$$\begin{aligned}
\lim_{(n+\kappa)\to 0} \bar{\mathbf{y}} &= \mathbf{f}[\bar{\mathbf{x}}] + \frac{1}{2}\boldsymbol{\nabla}^2\mathbf{f}\,\mathbf{P}_{xx} \\
\lim_{(n+\kappa)\to 0} \mathbf{P}_{yy}^{MOD} &= \boldsymbol{\nabla}\mathbf{f}\,\mathbf{P}_{xx}(\boldsymbol{\nabla}\mathbf{f})^T.
\end{aligned}$$

In other words, in the limit as $n+\kappa$ tends to zero, both the mean and covariance are calculated correctly to the second order.

The next section introduces the scaled unscented methods.

3 Sigma Point Scaling Methods

The scaled unscented transformation is a generalisation of the unmodified and modified unscented transformations. It is a method that scales an arbitrary sigma point set but ensures that the mean and covariance are maintained correctly. Implicitly, the idea is to replace the existing set of sigma points by a transformed set $\boldsymbol{\mathcal{X}}'_i = \boldsymbol{\mathcal{X}}_0 + \alpha(\boldsymbol{\mathcal{X}}_i - \boldsymbol{\mathcal{X}}_0)$, where α is a positive scaling parameter. The next two subsections discuss two alternative implementations of this form. The first introduces an *auxillary random variable* that used a modified form of the transition model and is related to $\mathbf{y}$. In many respects, this form offers the clearest explanation of how the scaled unscented transformation works. Subsection 3.2 shows that the same performance can be achieved using the conventional unscented transform with a simple post-processing step. Finally, Subsection 3.3 describes how some fourth order information can be incorporated into the calculated covariance.

3.1 The Auxillary Random Variable

The *auxillary random variable* $\mathbf{z}$ is related to $\mathbf{x}$ through the nonlinear equation $\mathbf{z} = \mathbf{g}[\mathbf{x}, \bar{\mathbf{x}}, \alpha, \mu]$ where

$$\mathbf{g}[\mathbf{x}, \bar{\mathbf{x}}, \alpha, \mu] = \frac{\mathbf{f}[\bar{\mathbf{x}} + \alpha(\mathbf{x} - \bar{\mathbf{x}})] - \mathbf{f}[\bar{\mathbf{x}}]}{\mu} + \mathbf{f}[\bar{\mathbf{x}}]. \tag{9}$$

α is the point scaling parameter which was introduced above and μ is a normalisation term which scales the transformed point about $\mathbf{f}[\bar{\mathbf{x}}]$. Taking a Taylor Series expansion of $\mathbf{g}[\cdot,\cdot,\cdot,\cdot]$ about $\bar{\mathbf{x}}$,

$$\mathbf{g}[\mathbf{x}, \bar{\mathbf{x}}, \alpha, \mu] = \mathbf{f}[\bar{\mathbf{x}}] + \boldsymbol{\nabla}\mathbf{f}\frac{\alpha}{\mu}\boldsymbol{\delta}\mathbf{x} + \frac{1}{2}\boldsymbol{\nabla}^2\mathbf{f}\frac{\alpha^2}{\mu}\boldsymbol{\delta}\mathbf{x}^2 + \frac{1}{3!}\boldsymbol{\nabla}^3\mathbf{f}\frac{\alpha^3}{\mu}\boldsymbol{\delta}\mathbf{x}^3 + \cdots \tag{10}$$

Comparing this with Equation 2, $\mathbf{g}[\cdot,\cdot,\cdot,\cdot]$ has a similar structure to $\mathbf{f}[\cdot]$ and that α and μ scale the first and higher order terms of the series. Taking expectations, the mean and covariance of $\mathbf{z}$ are

$$\bar{\mathbf{z}} = \mathbf{f}[\bar{\mathbf{x}}] + \frac{1}{2}\boldsymbol{\nabla}^2\mathbf{f}\,\frac{\alpha^2}{\mu}\mathbf{P}_{xx} + \frac{1}{6}\boldsymbol{\nabla}^3\mathbf{f}\,\frac{\alpha^3}{\mu}\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^3\right] + \cdots \tag{11}$$

$$\mathbf{P}_{zz} = \frac{\alpha^2}{\mu^2}\boldsymbol{\nabla}\mathbf{f}\,\mathbf{P}_{xx}(\boldsymbol{\nabla}\mathbf{f})^T + \frac{\alpha^3}{\mu^2}\frac{1}{2}\boldsymbol{\nabla}\mathbf{f}\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^3\right]\left(\boldsymbol{\nabla}^2\mathbf{f}\right)^T \tag{12}$$

$$+ \frac{\alpha^4}{\mu^2}\frac{1}{2}\boldsymbol{\nabla}^2\mathbf{f}\left(\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^4\right] - \mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^2\mathbf{P}_{yy}\right] - \mathrm{E}\left[\mathbf{P}_{yy}\boldsymbol{\delta}\mathbf{x}^2\right] + \mathbf{P}_{yy}^2\right)(\boldsymbol{\nabla}^2\mathbf{f})^T \tag{13}$$

$$+ \frac{\alpha^4}{\mu^2}\frac{1}{3!}\boldsymbol{\nabla}^3\mathbf{f}\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^4\right](\boldsymbol{\nabla}\mathbf{f})^T + \cdots \tag{14}$$

From Equation 5, if $\mu = \alpha^2$, the values of $\bar{\mathbf{y}}$ and $\bar{\mathbf{z}}$ agree with one another up to the second order and the third and higher order moments scale geometrically with a common factor of α.Similarly, let $\mathbf{P}^*_{zz} = \mu\mathbf{P}_{zz}$. From Equation 4, $\mathbf{P}^*_{zz}$ agrees with $\mathbf{P}_{yy}$ up to the second order and higher order terms scale with α. Therefore, α can be chosen to incorporate appropriate higher order information. For example, to offset the dimension-related scaling that occurs in Equation 7, $\alpha = 1/n$.

Therefore, the *auxillary form of the unscented transformation* is

$$\boldsymbol{\mathcal{Z}}_i = \mathbf{g}\left[\boldsymbol{\mathcal{X}}_i, \bar{\mathbf{x}}, \alpha, \mu\right] \tag{15}$$

$$\bar{\mathbf{z}} = \sum_{i=0}^{p} W_i \boldsymbol{\mathcal{Z}}_i \tag{16}$$

$$\mathbf{P}_{zz} = \mu \sum_{i=0}^{p} W_i \left\{\boldsymbol{\mathcal{Z}}'_i - \bar{\mathbf{z}}\right\}\left\{\boldsymbol{\mathcal{Z}}'_i - \bar{\mathbf{z}}\right\}^T \tag{17}$$

3.2 The Scaled Unscented Transform

The scaled unscented transform yields the same results as the auxillary random variable, but without the need to modify the process and observation model. Rather, an initial set of points are chosen and a transformation is applied to these points. The mean and covariance are calculated as normal and then a final term is added back.

Suppose a set of sigma points $\mathcal{S}$ have been constructed using a sigma point selection algorithm. These have mean $\bar{\mathbf{x}}$ and covariance $\mathbf{P}_{xx}$. We first construct a set of auxillary sigma points, $\mathcal{S}' = \{i = 0, 1, \ldots, p : \boldsymbol{\mathcal{X}}'_i, W'_i\}$. These have the same mean and covariance as $\mathcal{S}$, but the sigma points obey the conditions

$$\boldsymbol{\mathcal{X}}'_i = \boldsymbol{\mathcal{X}}_0 + \alpha(\boldsymbol{\mathcal{X}}_i - \boldsymbol{\mathcal{X}}_0) \tag{18}$$

where $\boldsymbol{\mathcal{X}}.$ are the sigma points which would be used to calculate the auxillary random variable. The weights on $\mathcal{S}'$ are related to $\mathcal{S}$ through

$$W'_i = \begin{cases} W_0/\alpha^2 + (1 - 1/\alpha^2) & i = 0 \\ W_i/\alpha^2 & i \neq 0 \end{cases} \tag{19}$$

The proof can be found in the Appendix. This set of points still obeys the same mean and covariance condition and, for sigma point selection algorithms that are parameterised by the value of W_0, $\mathcal{S} = \mathcal{S}[W_0]$, $\mathcal{S}\prime = \mathcal{S}[W\prime_0]$.

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

$$\boldsymbol{\mathcal{Y}}'_i = \mathbf{f}\left[\boldsymbol{\mathcal{X}}'_i\right]. \tag{20}$$

$$\bar{\mathbf{y}}' = \sum_{i=0}^{p} W'_i \boldsymbol{\mathcal{Y}}'_i. \tag{21}$$

$$\mathbf{P}'_{yy} = \sum_{i=0}^{p} W'_i \left\{\boldsymbol{\mathcal{Y}}'_i - \bar{\mathbf{y}}\right\}\left\{\boldsymbol{\mathcal{Y}}'_i - \bar{\mathbf{y}}\right\}^T + (1 - \alpha^2)\left\{\boldsymbol{\mathcal{Y}}'_0 - \bar{\mathbf{y}}\right\}\left\{\boldsymbol{\mathcal{Y}}'_0 - \bar{\mathbf{y}}\right\}^T. \tag{22}$$

In the Appendix we prove that $\bar{\mathbf{y}}' = \bar{\mathbf{z}}$ and $\mathbf{P}'_{yy} = \mu \mathbf{P}_{zz}$ when $\mu = \alpha^2$. Comparing this form with Equations 5 and 6, we see that both forms are virtually the same apart from the extra term applied to $\mathbf{P}'_{yy}$. The only difference is that an extra correction term, is included. When $\alpha = 1$, this gives the conventional unscented. When $\alpha = 0$, this form gives the modified covariance of Equation 8.

3.3 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), it is possible to include extra terms that capture some of the higher order behaviour of the system in the covariance equation. From Equation 4, the fourth order contribution to the covariance is

$$\begin{aligned}\mathbf{A} = &\frac{1}{2}\boldsymbol{\nabla}^2\mathbf{f}\left(\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^4\right] - \mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^2\mathbf{P}_{yy}\right] - \mathrm{E}\left[\mathbf{P}_{yy}\boldsymbol{\delta}\mathbf{x}^2\right] + \mathbf{P}^2_{yy}\right)(\boldsymbol{\nabla}^2\mathbf{f})^T \\ &+ \frac{1}{3!}\boldsymbol{\nabla}^3\mathbf{f}\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^4\right](\boldsymbol{\nabla}\mathbf{f})^T.\end{aligned} \tag{23}$$

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,

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

Comparing this with Equation 23, it can be seen that the first term in this expression equals one contribution to the fourth order term for the covariance. Therefore, by 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 22 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 the special case where the distribution is Gaussian, $\mathrm{E}\left[\boldsymbol{\delta}\mathbf{x}^4\right] = 3\mathbf{P}^2_{xx}$. Therefore, the actual fourth order term is

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

whereas the approximated term is

$$\mathbf{A} = \boldsymbol{\nabla}^2\mathbf{f}\,\mathbf{P}^2_{yy}(\boldsymbol{\nabla}^2\mathbf{f})^T.$$

The relevance of this extra information will be demonstrated below.

4 Example — Not Completed Yet

This section demonstrates the use of the scaled unscented transformation in a high order nonlinear system. The system is a navigation system for a high speed vehicle that uses INS, GPS and a laser scanner. The system state vector consists of six states and ten process noises, giving a total state space of 16. For example, the process model for this vehicle is:

$$\begin{aligned}
X_F(k) &= X_F(k-1) + V\cos[\psi(k-1) + \delta(k-1) - \alpha_f(k-1)]\Delta T \\
Y_F(k) &= Y_F(k-1) + V\sin[\psi(k-1) + \delta(k-1) - \alpha_f(k-1)]\Delta T \\
\psi(k) &= \psi(k-1) + V(k-1)\rho_F\Delta T \\
R(k) &= R(k-1) \\
\alpha_f(k) &= \alpha_f(k-1) \\
\alpha_r(k) &= \alpha_r(k-1) \\
\omega(k) &= \omega(k-1) + \dot{\omega}(k-1)\Delta T \\
\rho_F &= \frac{\sin[\delta - \alpha_f] + \cos[\delta - \alpha_f]\tan\alpha_r}{B} \\
\omega(k) &= \omega(k-1) + \dot{\omega}(k-1)\Delta T.
\end{aligned}$$

5 Conclusions

This paper has presented the scaled unscented transform — a new parameterisation of the unscented transform which introduces the additional scaling parameters α and μ. We have shown that the properties of this algorithm are superior to those of the conventional unscented transform in all respects. It is able to retain second order accuracy in both the mean and covariance, but without the problem of the sigma point position "explosion" and without the need to use negative weights. Given its superior properties, we believe this algorithm is superior to conventional unscented for all systems whose dimensionality is greater than three.

A The Relationship Between the Scaled and Unscaled Unscented Transforms

This Appendix shows that the *any* scaling strategy of the form of Equation 9 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.

Theorem 1. *The weights of the $\mathcal{S}'$ are related to those of $\mathcal{S}$ by Equation 19.*

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

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

$$\sum_{i=1}^{p} W_i(\boldsymbol{\mathcal{X}}_i - \bar{\mathbf{x}})(\boldsymbol{\mathcal{X}}_i - \bar{\mathbf{x}})^T = \mathbf{P}_{xx} \tag{A.2}$$

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

$$\sum_{i=0}^{p} W'_i = 1 \tag{A.3}$$

$$\sum_{i=1}^{p} W'_i(\boldsymbol{\mathcal{X}}'_i - \bar{\mathbf{x}})(\boldsymbol{\mathcal{X}}'_i - \bar{\mathbf{x}})^T = \mathbf{P}_{xx} \tag{A.4}$$

Comparing Equations A.2 with A.4 and substituting from Equation 18, it can be seen that $W_i = W'_i\alpha^2$ for $i > 0$. W_0 is found from Equations A.1 and A.3,

$$\begin{aligned} 1 = \sum_{i=0}^{p} W_i &= W_0 + \sum_{i=1}^{p} W_i \\ &= W_0 + \alpha^2 \sum_{i=1}^{p} W'_i \\ &= W_0 + \alpha^2(1 - W'_0) \end{aligned} \tag{A.5}$$

□

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 15,

$$\begin{aligned} \boldsymbol{\mathcal{Z}}_i &= \mathbf{g}\left[\boldsymbol{\mathcal{X}}_i, \bar{\mathbf{x}}, \alpha, \mu\right] \\ &= \left(1 - \frac{1}{\mu}\right)\mathbf{f}\left[\boldsymbol{\mathcal{X}}'_0\right] + \frac{1}{\mu}\mathbf{f}\left[\boldsymbol{\mathcal{X}}'_i\right] \\ &= \left(1 - \frac{1}{\mu}\right)\boldsymbol{\mathcal{Z}}'_0 + \frac{1}{\mu}\boldsymbol{\mathcal{Y}}'_i. \end{aligned} \tag{A.6}$$

Theorem 2. *Let*

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

Then

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

Proof. Substituting from Equation A.6 and using the fact that $\sum_{i=0}^{p} W_i = 1$,

$$\begin{aligned}\bar{\mathbf{z}} &= \left(1 - \frac{1}{\mu}\right) \boldsymbol{\mathcal{Z}}'_0 + \frac{1}{\mu}\sum_{i=0}^{p} W_i \boldsymbol{\mathcal{Z}}'_i \\ &= \left(\frac{\mu - 1}{\mu}\right) \boldsymbol{\mathcal{Z}}'_0 + \frac{1-\alpha^2}{\mu}\boldsymbol{\mathcal{Z}}'_0 + \frac{\alpha^2}{\mu}\sum_{i=0}^{p} W'_i \boldsymbol{\mathcal{Z}}'_i \\ &= \frac{\mu - \alpha^2}{\mu}\boldsymbol{\mathcal{Z}}'_0 + \frac{\alpha^2}{\mu}\bar{\mathbf{y}}'\end{aligned} \tag{A.7}$$

□

Theorem 3. *Let*

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

Then

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

Proof. Substituting from Equations A.6 and A.7,

$$\boldsymbol{\mathcal{Z}}_i - \bar{\mathbf{z}} = \frac{1}{\mu}(\boldsymbol{\mathcal{Z}}'_i - \bar{\mathbf{y}}') + \frac{(\alpha^2 - 1)}{\mu}(\boldsymbol{\mathcal{Z}}'_0 - \bar{\mathbf{y}}') \tag{A.8}$$

Therefore,

$$\begin{aligned}\mathbf{P}_{zz} &= \mu \sum_{i=0}^{p} W_i \left\{\frac{1}{\mu}(\boldsymbol{\mathcal{Z}}'_i - \bar{\mathbf{y}}') + \frac{(\alpha^2 - 1)}{\mu}(\boldsymbol{\mathcal{Z}}'_0 - \bar{\mathbf{y}}')\right\}\left\{\frac{1}{\mu}(\boldsymbol{\mathcal{Z}}'_i - \bar{\mathbf{y}}') + \frac{(\alpha^2 - 1)}{\mu}(\boldsymbol{\mathcal{Z}}'_0 - \bar{\mathbf{y}}')\right\}^T \\ &= \frac{1}{\mu}\sum_{i=0}^{p} W_i \left\{(\boldsymbol{\mathcal{Z}}'_i - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Z}}'_i - \bar{\mathbf{y}}')^T + (\alpha^2 - 1)(\boldsymbol{\mathcal{Z}}'_i - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Z}}'_0 - \bar{\mathbf{y}}')^T + \right. \\ &\qquad \left. (\alpha^2 - 1)(\boldsymbol{\mathcal{Z}}'_0 - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Z}}'_i - \bar{\mathbf{y}}')^T + (\alpha^2 - 1)^2(\boldsymbol{\mathcal{Z}}'_0 - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Z}}'_0 - \bar{\mathbf{y}}')^T\right\}\end{aligned} \tag{A.9}$$

From Equation 19,

$$\sum_{i=0}^{p} W_i(\boldsymbol{\mathcal{Z}}'_i - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Z}}'_i - \bar{\mathbf{y}}')^T = \alpha^2\mathbf{P}_{yy} + (1-\alpha^2)(\boldsymbol{\mathcal{Z}}'_0 - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Z}}'_0 - \bar{\mathbf{y}}')^T \tag{A.10}$$

$$\sum_{i=0}^{p} W_i(\boldsymbol{\mathcal{Z}}'_i - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Z}}'_0 - \bar{\mathbf{y}}')^T = (1-\alpha^2)(\boldsymbol{\mathcal{Z}}'_0 - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Z}}'_0 - \bar{\mathbf{y}}')^T \tag{A.11}$$

$$\sum_{i=0}^{p} W_i(\boldsymbol{\mathcal{Z}}'_0 - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Z}}'_0 - \bar{\mathbf{y}}')^T = (\boldsymbol{\mathcal{Z}}'_0 - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Z}}'_0 - \bar{\mathbf{y}}')^T \tag{A.12}$$

Substituting Equations A.10 to A.12 into Equation A.9,

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

□

Remark 1. *When* $\mu = \alpha^2$,

$$\begin{aligned} \bar{\mathbf{z}} &= \bar{\mathbf{y}}' \\ \mathbf{P}_{zz} &= \mathbf{P}'_{yy} + (1-\alpha^2)(\boldsymbol{\mathcal{Z}}'_0 - \bar{\mathbf{y}}')(\boldsymbol{\mathcal{Z}}'_0 - \bar{\mathbf{y}}')^T. \end{aligned}$$

References

[1] F. E. Daum. New Exact Nonlinear Filters. In J. C. Spall, editor, *Bayesian Analysis of Time Series and Dynamic Models*, chapter 8, pages 199–226. Marcel Drekker, Inc., 1988.

[2] N. J. Gordon, D. J. Salmond and A. F. M. Smith. Novel Approach to Nonlinear/non-Gaussian Bayesian State Estimation. *IEE Proceedings-F*, 140(2):107–113, April 1993.

[3] A. H. Jazwinski. *Stochastic Processes and Filtering Theory.* Academic Press, 1970.

[4] S. J. Julier. *Comprehensive Process Models for High-Speed Navigation.* PhD thesis, University of Oxford, 1997.

[5] 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, Florida.* SPIE, 1998. Signal and Data Processing of Small Targets.

[6] S. J. Julier. The Simplex Unscented Transform. In Preparation, 1999.

[7] S. J. Julier and H. F. Durrant-Whyte. Navigation and Parameter Estimation of High Speed Road Vehicles. In *Robotics and Automation Conference, Nagoya, Japan.*, pages 101–105, 1995.

[8] S. J. Julier and J. K. Uhlmann. A Consistent, Debiased Method for Converting Between Polar and Cartesian Coordinate Systems. In *The Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls, Orlando, Florida.* SPIE, 1997. Acquisition, Tracking and Pointing XI.

[9] 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, Florida.* SPIE, 1997. Multi Sensor Fusion, Tracking and Resource Management II.

[10] 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*, Accepted for publication as a Technical Note. 1996.

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

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

[13] H. W. Sorenson and A. R. Stubberud. Non-linear Filtering by Approximation of the *a posteriori* Density. *International Journal of Control*, 8(1):33–51, 1968.