ACC03−IEEE0215

The Higher Order Unscented Filter

Dirk Tenne
Graduate Student

Tarunraj Singh
Associate Professor

Department of Mechanical & Aerospace Engineering,
State University of New York at Buffalo

Abstract

This article proposes a technique for the selection of the σ-set for a probability distribution approximation filter, i.e. the unscented filter. The σ-set is selected so as to capture the higher order input statistics. The Taylor series expansion is used to illustrate that the 3^{rd} order and higher statistics of the nonlinear transformation can be reproduced while the unscented filter captures the statistics up to the 3^{rd} order only. Two benchmark problems are used to corroborate the proposed solution.

1 Introduction

Recently, there has been an increasing interest in the development of techniques for nonlinear estimation, which can approximate the statistics of the process as accurately as possible [1, 2, 3].

The process of state estimation combines a priori knowledge about the state and its transition due to a set of inputs, with a sequence of measurements. Generally, the state estimate is recursively obtained in two stages: (i) the prediction and (ii) the update. The well known optimal state estimator utilizes the analytical probability density function (pdf) to predict the states using the total probability theorem and subsequently uses the measurement to update with Bayes' rule [4]. Since this optimal estimator requires storing the pdf and integrating it, often resulting in impractical algorithms, techniques have been developed to approximate the optimal estimator. A majority of algorithms such as the extended Kalman filter (EKF), divided difference filter (DDF), unscented Kalman filter (UKF) etc., focus on approximating the prediction probability characteristics and use the standard linear minimum mean square error estimator. The EKF uses first order Taylor series, which can be improved by higher order approximations [5] at the cost of computational efficiency. Nørgaard et al. [1] exploited polynomial interpolations of the nonlinear transformations replacing the analytical Jacobian of the EKF with numerically evaluated divided differences (DDF). The Gaussian filter applied by Ito et al. [6] numerically integrates the predictor step and yields, depending on the integration scheme, i.e. central difference filter (CDF), results similar to the DDF. However, Ito states that the Gauss-Hermite rule outperforms the CDF. Emulating a Monte Carlo simulation, Julier and Uhlman [2] perform the prediction of the statistics with a set of carefully chosen sample points.

On the other hand, research has focused on the update equation (essentially Bayes' rule), which is optimal when one has perfect knowledge of the pdf. The particle filter (PF) represents the required pdf's by a finite sum based on samples (or particles) drawn from an assumed distribution [3, 7, 8, 9]. Difficulties arise with the selection of the proposal distribution, which can be simplified by combining the unscented filter (or EKF) with the particle filter as shown by van der Merwe et al. [8]. The UKF generates approximations of the prior distribution with adjustable accuracy. Challa and Bar-Shalom [10] numerically propagate the prior pdf through the Fokker-Planck-Kolmogorov equation, where the negligible tails of the distribution are identified by Chebyshev's inequality. Although this approach solves the optimal estimator, its computational efforts are large compared to any of the approaches, which emulate the Monte Carlo simulations.

The probability distribution approximation filter (PDAF) termed *unscented Filter* by Uhlman and Julier [2] has been developed in the context of state estimation of dynamic systems, for example fast moving robots. Initial application to parameter estimation of this technique has been proposed by van der Merwe and Wan [8]. The unscented filter carefully selects a set of stencils called the σ-set, which exhibits the same statistical properties (mean, moments) to a certain degree as the true distribution of the state. The σ-set is propagated through the state equations and the statistics of the predicted state is the result of a weighted sum, which agrees up to the third order of the Taylor series of the true mean and covariance. The proposed technique modifies the selection of the σ-set to arrive at a design space capable of approximating higher order input moments and thus approaching the true mean and covariance.

2 The Unscented Filter

Consider a nonlinear transformation of the random variable X with mean $\bar{x}$ and variance P_x

$$Y = g(X) \ , \tag{1}$$

where we would like to approximate the statistics of the transformed random variable Y, for example the mean and variance. Note, that this paper improves on the approximation of the statistics of the aformentioned transformation, which can be seen as the prediction algorithm of the state estimation.

The unscented filter selects a σ-set consisting of $2n+1$ points, which are perturbations from the mean by a scaled deviation. The deviations σ_i are defined as the columns of the matrix square root of P_x [2].

$$\sigma = \pm\sqrt{P_x} \tag{2}$$

The σ-set is defined as:

$$\begin{aligned} \zeta_0 &= \bar{X} && (3) \\ \zeta_i &= \zeta_0 + \sqrt{(n+\kappa)}\ \sigma_i \quad \text{for} \quad i = 1 \ldots 2n \ , && (4) \end{aligned}$$

such that the σ-set exhibits the same probabilistic characteristics as the random variable X and κ is a free variable of the unscented filter. The weighting scheme

$$\bar{\zeta} = \sum_{i=0}^{2n} w_i \zeta_i \tag{5}$$

$$P_\zeta = \sum_{i=0}^{2n} w_i (\zeta_i - \bar{\zeta})^2 \tag{6}$$

$$\text{where} \qquad w_o = \frac{\kappa}{n+\kappa} \tag{7}$$

$$w_i = \frac{1}{2(n+\kappa)} \tag{8}$$

has been selected to match the first four central moments of X, i.e.

$$\bar{x} = \bar{\zeta} \quad P_x = P_\zeta \tag{9}$$

$$\mu_x^3 = \mu_\zeta^3 = 0 \quad \mu_x^4 = \mu_\zeta^4 \tag{10}$$

which is true for all Gaussian random variables if κ is selected to satisfy the constraint, $n + \kappa = 3$.

The σ points are transformed by the nonlinear relationship

$$\eta_i = g[\zeta_i, t] \quad \text{for} \quad i = 1 \ldots 2n \tag{11}$$

and the estimate of the mean and variance of Y is obtained by the same weighting scheme:

$$\bar{\eta} = \sum_{i=0}^{2n} w_i \eta_i \tag{12}$$

$$P_\eta = \sum_{i=0}^{2n} w_i (\eta_i - \bar{\eta})^2 \ , \tag{13}$$

where the weights remain the same as for matching the input moments, equations 7 and 8.

3 The Higher Order Unscented Filter

It is a well known fact that the Unscented Filter described by equations 12, 13 exhibits an accuracy in estimating the statistical characteristics up to the second moment. This has been achieved by selecting a σ-set resembling the input statistics. By utilizing the Taylor series expansion it can be shown that the transformed statistics match the covariance up to the third order, i.e. errors are introduced at the fourth and higher orders [2].

Since the higher order terms of the Taylor series consist of a combination of higher order moments, it is essential to correctly estimate these moments to achieve an increase in the performance of the estimator. Appendix A shows the derivation of the Taylor series expansion. The following section introduces an algorithm matching higher order moments of the input.The performance increase is illustrated on a simple example.

The higher order unscented filter (HOUF) remains the same structure as the original unscented filter. The HOUF consists of an augmented σ-set with separate weights. The new σ-set can be constructed as follows:

$$\begin{aligned} \zeta_0 &= \bar{X} && (14) \\ \zeta_i &= \zeta_0 + \sqrt{m+\kappa}\ \sigma_i && (15) \\ & \quad \text{for} \quad \sigma_i = \pm\sigma_1, \pm\sigma_2, \ldots, \pm\sigma_m \ , \end{aligned}$$

where $m = n + q$ is the augmented dimension and q is the number of additional σ points. The statistical properties of the σ-set are defined by the weighted sums:

$$\bar{\zeta} = \sum_{i=0}^{2m} w_i \zeta_i \tag{16}$$

$$\mu_\zeta^j = \sum_{i=0}^{2m} w_i (\zeta_i - \bar{\zeta})^j \tag{17}$$

where $\bar{\zeta}$ and μ_ζ^j represent the mean and the j^{th} central moment.

3.1 2-σ HOUF

The statistics of the σ-set are selected to match the statistics of the input X. This procedure is exemplified for the HOUF with two σ's, σ_1 and σ_2. Expanding equations 16 and 17 up to the eighth central moment, yields the following relationships:

$$\begin{aligned} 1 &= w_0 + 2w_1 + 2w_2 && \text{(18a)} \\ P_x &= 2(m+\kappa)(w_1\sigma_1^2 + w_2\sigma_2^2) && \text{(18b)} \\ \mu_x^3 &= 0 && \text{(18c)} \\ \mu_x^4 &= 2(m+\kappa)^2(w_1\sigma_1^4 + w_2\sigma_2^4) && \text{(18d)} \\ \mu_x^5 &= 0 && \text{(18e)} \\ \mu_x^6 &= 2(m+\kappa)^3(w_1\sigma_1^6 + w_2\sigma_2^6) && \text{(18f)} \\ \mu_x^7 &= 0 && \text{(18g)} \\ \mu_x^8 &= 2(m+\kappa)^4(w_1\sigma_1^8 + w_2\sigma_2^8)\ , && \text{(18h)} \end{aligned}$$

where μ_x^i is the i^{th} central moment of X. With the knowledge of the input moments we are able to solve for the σ points σ_i, the weights w_i and κ. The general solution becomes rather difficult to present and one has to decide on a case by case basis about possible simplifications. The following transformation of the weights and the requirement of the standard unscented filter $m + \kappa = 3$ lead to an amenable solution in the particular case of Gaussian moments.

$$w_1' = 2(m+\kappa)w_1 \qquad w_2' = 2(m+\kappa)w_2 \tag{19}$$

The equations 18 can be rewritten as:

$$\begin{aligned} m+\kappa &= (m+\kappa)w_0 + w_1' + w_2' && \text{(20a)} \\ P_x &= w_1'\sigma_1^2 + w_2'\sigma_2^2 && \text{(20b)} \\ 3P_x^2 &= w_1'\sigma_1^4 + w_2'\sigma_2^4 && \text{(20c)} \\ \frac{5}{3}P_x^3 &= w_1'\sigma_1^6 + w_2'\sigma_2^6 && \text{(20d)} \\ \frac{35}{9}P_x^4 &= w_1'\sigma_1^8 + w_2'\sigma_2^8\ , && \text{(20e)} \end{aligned}$$

yielding a set of 5 equations for the 5 unknown parameters.

In contrast to the unscented filter, which assumes the same weights w_i for all ζ_i for $i > 0$, solving equations 20 requires two independent weights w_1 and w_2. Assuming $w_1 = w_2$, Equation 18, for example, can be written as:

$$\frac{15}{(m+\kappa)^2}P_x^3 = (\sigma_1^2)^3 + (\sigma_2^2)^3\ , \tag{21}$$

which according to Fermats Last Theorem [11] does not have any nonzero integer solution P_x, σ_1^2 and σ_2^2. However, employing an analytical solver, a solution can be obtained in the following form:

$$\begin{aligned} \sigma_1^2 &= \left(\frac{5}{3} \pm \sqrt{\frac{10}{9}}\right) P_x && \text{(22a)} \\ \sigma_2^2 &= \frac{10}{3}P_x - \sigma_1^2 && \text{(22b)} \\ w_1 &= \frac{1}{60}\frac{21\sigma_1^2 - 55P_x}{3\sigma_1^2 - 5P_x} && \text{(22c)} \\ w_2 &= \frac{1}{3}\frac{3P_x(\sigma_1^2 - P_x)}{(3\sigma_1^2 - P_x)(\sigma_2^2 - \sigma_1^2)} && \text{(22d)} \\ w_0 &= 1 - 2(w_1 + w_2) && \text{(22e)} \end{aligned}$$

For example, consider the nonlinear transformation:

$$Z = |X|\ , \tag{23}$$

where X is a zero mean Gaussian random variable with variance P_x. It is a well known fact that the transformed random variable Y is non-Gaussian and its odd-order moments are non-zero. A Monte Carlo (MC) simulation with 25 million samples has been performed and the resulting moments are shown in Table 1 along with the estimates of the Higher Order Unscented Filter and the standard Unscented Filter (UF). The parameters for the HOUF are derived from equation 22 and listed in Table 2. We arrive at two solutions, which in turn are identical since the indices are swapped, i.e. σ_1 becomes σ_2 and vice versa, etc. Observing Table 2,

Table 2: Solution of the Higher Order Unscented Filter

σ_1	σ_2	w_0	w_1	w_2	κ
1.6495	0.7827	0.5333	0.0113	0.2221	1
0.7827	1.6495	0.5333	0.2221	0.0113	1

the even order moments of the transformed random variable Z remain the same as the input X, whereas the odd order moments change significantly. The UF matches the MC's even order moments up to the fourth order, while the odd order moments differ. The HOUF closely approximates the even and odd order moments of the MC simulation up to the eights order moment.

For greater clarity, we expand the above example to a problem with known analytical solution. Let the random variable Y describe the distance of a particle above a barrier [12]. The particle at the initial position x_0 is traveling with a velocity of v_0, then the distance from the barrier is given as:

$$y = |x_0 + v_0 t|\ . \tag{24}$$

Now, let the initial position x_0 be a Gaussian distributed random variable with unit variance P_x and zero mean such that

$$Y = g(X,t) = |X|\ , \tag{25}$$

where $X = N(v_0 t, 1)$ is a Gaussian random variable with a time varying mean. The distance of the particle exemplifies a function with varying level of nonlinearity since for a larger distance of the particle from the barrier the transformation becomes almost linear, i.e. $Y = X$ and only the tails of the Gaussian distributed X are effected by the nonlinear transformation. The analytical solutions of the mean and variance can be derived with standard procedures of probability theory, which are:

$$\bar{y} = E\{Y\} = v_0 t\, \mathrm{erf}\left(\frac{v_0 t}{\sqrt{2P_x}}\right) + \sqrt{\frac{2P_x}{\pi}} e^{-\frac{1}{2}\frac{(v_0 t)^2}{P_x}}\ , \tag{26}$$

Table 1: Moments of the transformation $Z = |X|$

$P_x = 1$	**Moments**							
	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8
input X								
	0	1	0	3	0	15	0	105
transformed $Z = \|X\|$								
MC	0.7980	1.0005	1.5971	3.0035	6.3926	15.0283	38.2514	104.8600
UF	0.5774	1.0000	1.7321	3.0000	5.1962	9.0000	15.5885	27.0000
HOUF	0.6667	1.0007	1.6335	3.0056	6.3351	15.0467	38.8505	105.3871

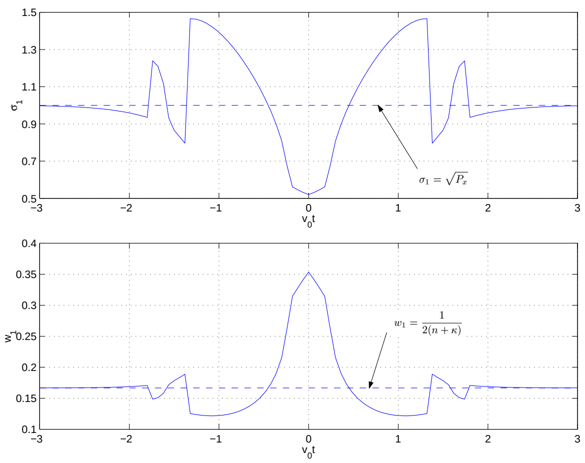


Figure 1: Optimal solution minimizing the mean and variance error

$$P_y = E\{(Y - \bar{y})^2\} = P_x + (v_0 t)^2 - \bar{y}^2 \ . \qquad (27)$$

The goal is to estimate the mean $\bar{y}$ and the variance P_y as accurately as possible as a function of the σ-set and the weights w_i. The optimal solution is obtained by constructing a cost function consisting of the square error of the mean and variance estimate.

$$J(\sigma_i, w_i) = (\bar{y} - \bar{\eta})^2 + (P_y - P_\eta)^2 \qquad (28)$$

The cost function $J(\sigma_i, w_i)$ can be minimized for the unscented filter (equations 12, 13) with respect to the design variables σ_1 and w_1, where $w_0 = \frac{\kappa}{n+\kappa}$ and $n + \kappa = 3$ has been assumed. The numerically determined optimal σ_1 and w_1 are shown in Figure 1 for a time varying mean $v_0 t$ within the interval $[-3 \ 3]$. The "jumps", for example at $v_0 t = 1.3$, result from the selected nonlinear function, the absolute value, which is part of the weighted sequence in equations 12 and 13. Furthermore, the suggested σ-set and weights by Uhlman and Julier [2] are illustrated by the dashed lines. The optimal solution approaches Uhlman's σ-set as the input mean increases and the influence of the nonlinearity vanishes.

Figure 2 shows the analytical mean and deviation, i.e. the square root of the variance, for a time varying mean of the input X. The optimal solution (Figure 1) closely matches the analytical solution. Both unscented filters reasonably approximate the mean of the particles distance, whereas the HOUF better describes its deviation, which is the result of matching the higher order moments.

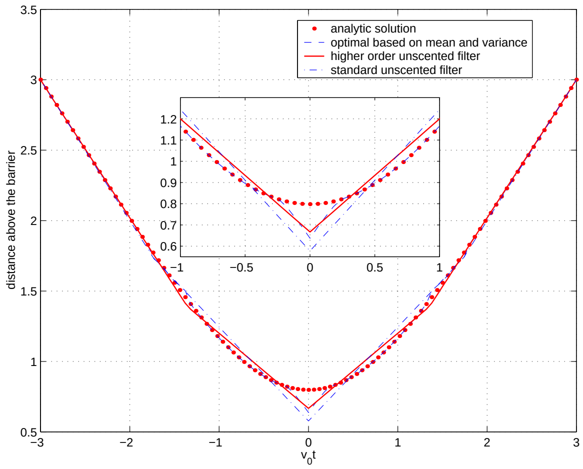
3.2 3-σ HOUF

The performance of the HOUF depends on the number of auxiliary σ's, where a larger σ-set should better approximate the true solution. The expansion to 3 σ's enables the HOUF to match up to the twelfth order moment. Solving an expanded set of equations similar to equations 20, yields the parameters, which are shown in Appendix B. Figure 3 compares the true solution with the HOUF for 2 and 3 σ's for the particle above a barrier example. It can be seen that 3-σ HOUF approaches the true solution.

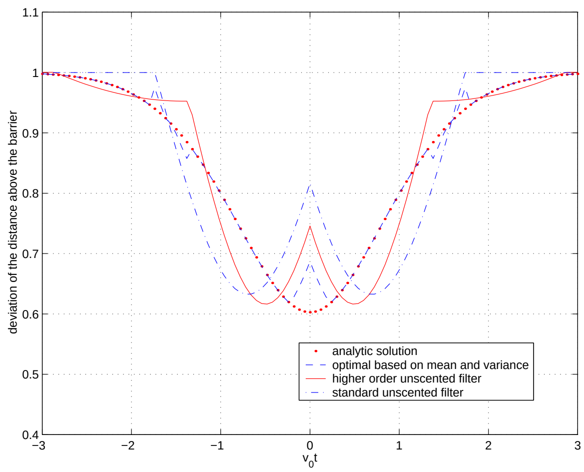
4 Conclusion

The unscented filter by Uhlman and Julier [2] matches up to the fifth order moment of the input variable (assuming a Gaussian distribution), whereas the transformed covariance introduces errors at the fourth and higher orders of a Taylor series expansion. The higher order unscented filter is capable of matching any desirable input moment by selecting a modified σ-set. This results in a higher order approximation of the transformed covariance. This has been shown on an example with variable degree of nonlinearity, i.e. the distance of a particle above a barrier.

The improved approximation of the transformed mean and covariance can be utilized in dynamic estimations such as Kalman filtering. With the assumption of Gaussian distribution of the state throughout its propagation the HOUF can approximate the mean and covariance with adjustable accuracy. On the contrary, a larger σ-set increases the computational load, which leads to a trade-off depending on the nonlinear transformation. As we have seen from the aforementioned example, after one transformation the Gaussian assumption does not hold and odd order central moments are generated. Based on the characteristics of



(a) Mean distance above the barrier



(b) Variance of the distance above the barrier

Figure 2: Deviation of the distance of a particle above a barrier

the nonlinear function it might be desirable to track higher order moments, which can be approximated by the HOUF after slight modification of the matching equations.

The HOUF derived in this paper currently applies to one dimensional systems, but its algorithm is amenable to address multi-dimensional extensions.

Acknowledgments

This research has been supported by the Center for Multisource Information Fusion (CMIF) and the Calspan/UB Research Center (CUBRC).

References

[1] Magnus Nørgaard, Niels K. Poulsen, and Ole Ravn. New developments in state estimation for nonlin-

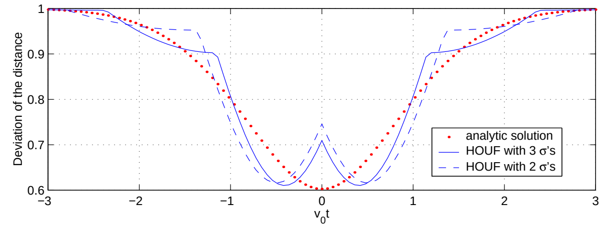


Figure 3: Comparison of 2- and 3-σ HOUF

ear systems. *Automatica*, 36(11):1627–1638, Nov. 2000.

[2] Simon J. Julier, Jeffrey K. Uhlmann, and Hugh F. Durrant-Whyte. New approach for filtering nonlinear systems. In *ACC*, volume 3, pages 1628–1632, 1995.

[3] N.J. Gordon, D.J. Salmond, and A.F.M. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEE Proc. on Radar & Signal Processing*, 140(2):107–113, Apr. 1993.

[4] Yaakov Bar-Shalom and Thomas E. Fortmann. *Tracking Data and Association*. Academic Press, Inc., 1988.

[5] Mohinder S. Grewal and Angus P. Andrews. *Kalman Filtering–Theory and Practice*. Prentice Hall Information and System Sciences Series, 1993.

[6] Kazufumi Ito and Kaiqi Xiong. Gaussian filters for nonlinear filtering problems. *IEEE Trans. on Automatic Control*, 45(5):910–927, May 2000.

[7] A. Doucet, N. de Freitas, and N. Gordon. *Sequential Monte Carlo Methods in Practice*. New York: Springer-Verlag, Jan. 2001.

[8] Rudolph van der Merwe, Arnaud Doucet, Nando de Freitas, and Eric Wan. The unscented particle filter. Technical Report CUED/F-INFENG/TR 380, Cambridge University Engineering Department, Aug. 2000.

[9] Sanjeev Arulampalam, Simon Maskell, Neil Gordon, and Tim Clapp. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Trans. on Signal Processing*, 50(2):174–188, Feb. 2002.

[10] S. Challa and Y. Bar-Shalom. Nonlinear filter design using fokker-planck-kolmogorov probability density evolutions. *IEEE Trans. on Aerospace and Electr. Systems*, 36(1):309–315, Jan. 2000.

[11] Andrew Wiles. Modular elliptic curves and fermat's last theorem. *The Annals of Mathematics*, 141(3):443–551, 1995.

[12] James R. van Zandt. A more robust unscented transform. Technical report, MITRE Corporation, http://www.mitre.org, July 2001.

A The Taylor Series Expansion

Describing the nonlinear transformation of equation 1 via a Taylor series about the mean $\bar{X}$, we have:

$$\begin{aligned} g(\bar{X}+\Delta X) = g(\bar{X}) \quad & + \quad \left.\frac{\mathrm{d}g}{\mathrm{d}x}\right|_{x=\bar{X}} \Delta X \\ & + \quad \frac{1}{2!}\left.\frac{\mathrm{d}^2 g}{\mathrm{d}x^2}\right|_{x=\bar{X}} \Delta X^2 \\ & + \quad \frac{1}{3!}\left.\frac{\mathrm{d}^3 g}{\mathrm{d}x^3}\right|_{x=\bar{X}} \Delta X^3 \\ & + \quad \ldots \end{aligned} \tag{29}$$

The mean and variance of the transformed random variable can be determined by taking expectations of the Taylor series

$$\begin{aligned} \bar{Y} &= E\{Y\} \\ &= g(\bar{X}) + \frac{1}{2!}\frac{\mathrm{d}^2 g}{\mathrm{d}x^2} P_x + \frac{1}{3!}\frac{\mathrm{d}^3 g}{\mathrm{d}x^3} E\{\Delta X^3\} + \ldots \end{aligned} \tag{30}$$

$$\begin{aligned} P_y &= E\left\{\left(Y-\bar{Y}\right)^2\right\} \\ &= \left[\left(\frac{\mathrm{d}g}{\mathrm{d}x}\right)^2 - \left(\frac{1}{2!}\frac{\mathrm{d}^2 g}{\mathrm{d}x^2}\right)^2 P_x\right] P_x + \\ & \left[\frac{\mathrm{d}g}{\mathrm{d}x}\frac{\mathrm{d}^2 g}{\mathrm{d}x^2} - \left(\frac{1}{3!}\frac{\mathrm{d}^3 g}{\mathrm{d}x^3}\right)^2 E\{\Delta X^3\} \right. \\ & \left. - \frac{1}{3!}\frac{\mathrm{d}^2 g}{\mathrm{d}x^2}\frac{\mathrm{d}^3 g}{\mathrm{d}x^3} P_x\right] E\{\Delta X^3\} + \ldots \end{aligned} \tag{31}$$

It is evident that the Taylor approximation of the mean and variance depends on the higher order moments of the input X.

B 3-σ HOUF optimal Parameters

Employing an analytical solver to the expanded set of equations 20, a solution for the optimal σ set can be obtained by sequentially solving the following three equations.

$$9\sigma_3^6 - 63P_x\sigma_3^4 + 105P_x^2\sigma_3^2 - 35P_x^3 = 0 \tag{32}$$

$$3\sigma_1^4 + (3\sigma_3^2 - 21P_x)\sigma_1^2 - 21P_x\sigma_3^2 + 3\sigma_3^4 + 35P_x = 0 \tag{33}$$

$$\sigma_2^2 + \sigma_1^2 + \sigma_3^2 - 7P_x = 0 \tag{34}$$

The closed form solution of the weights is:

$$w_1 = \frac{(126P_x^3 + 18\sigma_3^6 - 72\sigma_3^4P_x - 54\sigma_3^2P_x^2)\sigma_1^2 - 567P_x^4 + 98P_x^3\sigma_3^2 + 666P_x^2\sigma_3^4 + 27\sigma_3^8 - 270\sigma_3^6P_x}{54\sigma_1^2\sigma_2^2(\sigma_1^2-\sigma_2^2)(\sigma_3^2-\sigma_2^2)(\sigma_3^2-\sigma_1^2)} \tag{35}$$

$$w_2 = \frac{5P_x^2 - 3P_x\sigma_1^2 - 3P_x\sigma_3^2 + 3\sigma_1^2\sigma_2^2}{18\sigma_2^2(\sigma_1^2-\sigma_2^2)(\sigma_3^2-\sigma_2^2)} \tag{36}$$

$$w_3 = \frac{(\sigma_3^2 - 3P_x)^2 - \frac{8}{3}P_x^2}{\sigma_3^2(\sigma_3^2-\sigma_1^2)(\sigma_3^2-\sigma_2^2)} \tag{37}$$