The Conjugate Unscented Transform - An Approach to Evaluate Stochastic Integrals Involving the Normal Weight Function

Nagavenkat Adurthi, Graduate Student, nagavenk@buffalo.edu
Puneet Singla, Assistant Professor, psingla@buffalo.edu
Tarunraj Singh, Professor, tsingh@buffalo.edu
University at Buffalo, State University of New York, Amherst, NY 14260-4400

***Abstract*— This paper presents few methods to evaluate fully symmetric sigma points with positive weights for the Unscented Transform that can capture higher order moments of the normal probability density function. This work can be considered as a potential extension to the $2n+1$ Unscented Transform rule for nonlinear filtering or it can be used as an efficient Gaussian Cubature rule with reduced number of points compared to the Gauss Hermite product rule. Mainly three sets of sigma points are explored, firstly the set of sigma points that are 4th degree exact for any dimension. Secondly, the set of sigma points of that are 6th degree exact for dimensions $2 \leq N \leq 9$ and thirdly the set of sigma points that are 8th degree exact for dimensions $2 \leq N \leq 6$. Each method is compared to the Gauss Hermite product rule which is known to be exact for polynomial functions. The results are very encouraging for the fact that the same order of accuracy can be achieved with the a small fraction of the number of points used by the Gauss Hermite product rule. Finally results for a few non-polynomial type functions are shown that would motivate one to develop even higher order set of sigma points with reduced number of points.**

I. INTRODUCTION

Integrals involving the normal probability density function such as in (1) frequently arise in statistics and filtering and often do not have an analytical answer. One would have to resort to numerical computation methods such as the Gauss Hermite quadrature rule, where the quadrature points are the roots of Hermite polynomial of required degree. For 1 dimensional integrals the Gauss Hermite quadrature rules are simple to compute or are well documented, are accurate and are by far the minimal number of points required. To evaluate higher dimensional integrals, the quadrature points(now called cubature for $N \geq 2$) can be constructed from the 1 dimensional quadrature points, which gives rise to the 'Gauss Hermite Product rule'. A complete description on product rules can be found in chapter 3 of [1]. In 1 dimensional integrals, 'm' quadrature points are required for a polynomial of degree $2m-1$, the proof of which is given in [1]. Hence for a N-dimensional system one would require m^N cubature points. This number grows exponentially with the dimension. Even for a lower dimension such as 6, the number of points required to evaluate the integral when $f(x)$ is a polynomial of degree 9 is $5^6 = 15625$. This is a huge number of points that might be computationally expensive to use especially when the evaluation of $f(x)$ at each cubature point is in itself expensive. But fortunately the Gauss Hermite product rule is *not minimal*,with reference to [1], and there exists cubature rules with reduced number of points. This forms the basis of our motivation to develop efficient cubature methods.

An extensive amount of work has been done in this field to develop cubature rules with fewer points. Particularly [2] where the author has developed non-product Gaussian Cubature rules for second, third and fifth degree polynomials in any dimension. In general to develop a cubature rule that is applicable to any dimension is a difficult procedure. Thus the main focus has been to construct cubature rules only for few lower dimension integrals but to a higher degree with minimal number of points. A cubature rule of degree 2 with $N+1$ points and a cubature method of degree 3 with $2N$ points is developed in [3] for a general centrally symmetric weight function such as the normal and uniform PDF and it is claimed that this is the minimum number of points possible. In [4] a fully symmetric integration rules with minimal points for 2-Dimension are developed that are exact for degree 9-15. A 19 point cubature rule is developed in [5] for symmetric regions in 2-dimensions that are exact for degree 9. For a 2D integral, with symmetric weight function such as Normal and Uniform density functions, a 12 point cubature rule for degree 7 is developed in [6]. It is even claimed that this is the minimal number of points required and there exist many such 12 point cubature rule for degree 7 in 2D system.

Most of the non-product cubature rule possess certain similarities, they exploit the symmetry of the weight function, assume a structure for the cubature points, solve a system of nonlinear equations. Hence these methods also suffer some similar drawbacks such as inconsistency of the set of nonlinear equations, the presence of negative weights/complex roots and the inability to provide a generalized solution that works for any dimension. The present work heads in the similar manner and tries to overcome some of these difficulties. Thus the primary objective of this paper can be stated as "'To find a sigma point set with all positive weights and reduced number of points that is *equivalent* to the set of quadrature points of Gauss-Hermite Product rule of *same order*. By *equivalent to same order* we mean that for a polynomial of order $2m-1$ in N-Dimensions, both the new reduced sigma point set from the proposed Conjugate Unscented Transform method ('CUT') and the m^N quadrature points from Gauss-Hemrite product rule result in same order of relative error compared to the 'true value'.

A. Importance of Higher-Order Moments

Consider the integral in (1) for evaluating the true expected value of a function and also the approximated expected value in (2) using Taylor series expansion of the function about the mean of the Normal density function.

$$E[f(x)] = \int f(x)N(x,\mu|P)dx \tag{1}$$

$$E[f(x)] = f(\mu) + \nabla f(\mu)E[\delta x] + \frac{1}{2!}\nabla^2 f(\mu)E[\delta x^2] + \frac{1}{3!}\nabla^3 f(\mu)E[\delta x^3] + \frac{1}{4!}\nabla^4 f(\mu)E[\delta x^4]... \tag{2}$$

where $\delta x = (x-\mu)$. To elaborate, more the higher order moments are evaluated more is the integral value closer to the 'true value'. By 'true' value we mean the analytical answer if it is available or the *converged* numerical value evaluated by the Gauss-Hermite product rule. The word *converged* is stressed because for a non-polynomial function one usually does not know the required number of points for convergence.

B. Sigma Point Set to capture moments of Gaussian PDF

With out loss of generality consider a Gaussian PDF with 0 mean and identity covariance of appropriate dimension. Under this assumption all the odd order central and raw moments of the Gaussian PDF are 0. For a sigma point set $(x_1, x_2, ..., x_n)$ with corresponding weights $(w_1, w_2, ...w_3)$, the expected value or weighted integral value is approximated by a weighted sum as in (3). Consider the Taylor series approximation of the function value at each point about the mean as in (4) which are then substituted in (3). The resultant equation (5) is in terms of the sigma points and associated weights. Upon comparing the coefficients of (5) to (2), the resulting set of equations are called the moment constraint eqautions (6).

$$E[f(x)] = \sum_{i=1}^{n} w_i f(x_i) \tag{3}$$

$$f(x_i) = f(0) + \frac{1}{2!}\nabla^2 f(0)x_i^2 + \frac{1}{4!}\nabla^4 f(0)x_i^4 + \frac{1}{6!}\nabla^6 f(0)x_i^6 ... \tag{4}$$

$$E[f(x)] = f(0)(\sum_{i=1}^{n} w_i) + \frac{1}{2!}\nabla^2 f(0)(\sum_{i=1}^{n} w_i x_i^2) + \frac{1}{4!}\nabla^4 f(0)(\sum_{i=1}^{n} w_i x_i^4) + \frac{1}{6!}\nabla^6 f(0)(\sum_{i=1}^{n} w_i x_i^6) \tag{5}$$

$$\sum_{i=1}^{n} w_i x_i^m = E[\delta x^m], \quad m = 0, 1, 2, \cdots \tag{6}$$

When more sigma points are evaluated such that they can satisfy the higher order moment constraint equations (6), the integral value evaluated by (3) tends to converge to the true value of the integral. Though only a 1-Dimensional case is shown the moment constraint equations can also be extended to N-Dimensions.

TABLE I

MOMENT CONSTRAINT EQUATIONS FOR 2D TILL ORDER 4

Continuous	Discrete	Continuous	Dicrete
$E[x]$	$\sum_{i=1}^{n} w_i x_i$	$E[y]$	$\sum_{i=1}^{n} w_i y_i$
$E[x^2]$	$\sum_{i=1}^{n} w_i x_i^2$	$E[xy]$	$\sum_{i=1}^{n} w_i x_i y_i$
$E[y^2]$	$\sum_{i=1}^{n} w_i y_i^2$	$E[x^4]$	$\sum_{i=1}^{n} w_i x_i^2$
$E[y^4]$	$\sum_{i=1}^{n} w_i y_i^2$	$E[x^3 y]$	$\sum_{i=1}^{n} w_i x_i^3 y_i$
$E[y^3 x]$	$\sum_{i=1}^{n} w_i y_i^3 x_i$	$E[x^2 y^2]$	$\sum_{i=1}^{n} w_i x_i^2 y_i^2$

II. METHODOLOGY TO SOLVE THE MOMENT CONSTRAINT EQUATIONS

A common methodology that is followed throughout the paper can be summarized as

1) As we know that the Gaussian PDF is symmetric, we exploit this by constraining the sigma points to various symmetric axis. We later define some of these axis as and when they are required.
2) Sigma points on the same set of symmetric axis are equidistant from the mean and have equal weight. For the i^{th} set of axis , the distance scaling variables are labeled r_i and weight variables are labeled w_i.
3) We find all the moments of the continuous multidimensional normal PDF. The moments of a normal PDF are given by the Isserlis theorem which is formally restated in [7]. By which, given any covariance matrix 'P' we can find all higher order moments just from the entries of P.
4) We enumerate the list of points and then form the moment constraint equations till required order in terms of the variable r_i and w_i. For example the moment constraint equations till order 4 for a 2D system are shown in Table I. As the points (x_i, y_i) are constrained to a line, they can be replaced in terms of a single variable namely the distance r_i from the origin.
5) These nonlinear set of equations are solved for r_i and w_i, that give the sigma point set.

A. Sigma points that are 2^{nd} moment equivalent

Sigma points that are 2^{nd} moment equivalent can completely satisfy the moment constraint equations till 2^{nd} order. To facilitate with the analysis in this section the following definition is introduced with respect to a zero mean and identity covariance matrix Normal PDF

Principal axis: The orthogonal axis in cartesian space intersecting at the origin. These are the axis corresponding to each column of the identity covariance matrix. Thus there are N principal axis or $2N$ distinct points on the principal axis for N-Dimensional system. Let the set $D = \{1, 2, 3, ..., N\}$ We list the points on the principal axis as

$$\sigma_i \in \{\pm(I)_j | j \in D\} \quad i = 1, 2, 3, ..., 2N. \tag{7}$$

Each point on the principle axis is at a distance $|\sigma_i|$ or in this case at 1 unit distance from the origin, where $|x| = \sqrt{x_1^2 + x_2^2 \cdots x_N^2}$

1) Sigma points constrained to the principal axis: For an i.i.d random variables $(X_1, X_2, ..., X_N)$ the N-Dimensional normal PDF has Identity covariance and zero mean. The first four moments are

$$\begin{aligned} E[X_i^2] &= 1, \quad E[X_iX_j] = 0, \quad E[X_i^4] = 3 \\ E[X_i^3X_j] &= 0, \quad E[X_i^2X_j^2] = 1, \quad E[X_i^2X_jX_k] = 0 \\ E[X_iX_jX_kX_l] &= 0, \end{aligned} \tag{8}$$

Where $\{i, j, k, l\} \subset \{1, 2, 3, ..., N\}$. Consider a fully symmetric set of sigma points that lie on the principal axis such that the distance of each point from the origin r_1 and each has weight of w_1. The moments that have any odd powers for the random variable in the set of moments in (8) are already satisfied due to symmetry of sigma points. The non-zero/even moment constraint equations till order 4, interms of the variables r_1 and w_1 are

$$E[X_i^2] \equiv 2r_1^2 w_1 = 1 \tag{9}$$

$$E[X_i^4] \equiv 2r_1^4 w_1 = 3 \tag{10}$$

$$E[X_i^2X_j^2] \equiv 0 \neq 1 \tag{11}$$

Particularly the 4^{th} order cross moment $E[X_i^2X_j^2]$ in (11) cannot be satisfied by the sigma points that are constrained to be only on the principal axis. Infact no cross moment of any order can be satisfied by points just on the principal axis.The central weight is calculated as

$$w_0 = 1 - 2Nw_1 \tag{12}$$

In the following sub sections, methods that tend to find points on the principal axis to satisfy the 2^{nd} order moment constraint equations are illustrated with respect to the framework presented in this section.

2) $2N + 1$ sigma points of the Unscented Transform-UKF: The $2N + 1$ sigma points for the Unscented Transform in [8] are chosen such that 1 point is the origin and $2N$ points of each weight w_1 are constrained to lie on the principal axis at distance scaled by r_1. The distance of each point on the principal axis is chosen such that the second order moments are satisfied. A tuning parameter κ is introduced such that one of the 4th moment can be tuned. This works well for systems till dimension 3 after which the central weight becomes negative. The present framework is shown to be inline with the $2n + 1$ Unscented sigma points by working our way backwards. The suggested points are

$$X_0 = \mu \qquad W_0 = \kappa/(N + \kappa) \tag{13}$$

$$X_i = \mu + (\sqrt{(N + \kappa)P})_i \qquad W_i = 1/[2(N + \kappa)] \tag{14}$$

$$X_{i+N} = \mu - (\sqrt{(N + \kappa)P})_i \qquad W_{i+N} = 1/[2(N + \kappa)] \tag{15}$$

With $\mu = \vec{0}_{(N\times1)}$ and $P = I_{(N\times N)}$, by an elegant selection of $r_1 = \sqrt{N + \kappa}$ one could solve for w_1 from (9), w_0 from (12) and κ from (10) as

$$w_1 = 1/2(N + \kappa) \tag{16}$$

$$w_0 = 1 - 2Nw_1 = \kappa/(N + \kappa) \tag{17}$$

$$N + \kappa = 3 \tag{18}$$

Thus under a normal distribution assumption the tuning parameter $\kappa = 3 - N$. When the dimension of system is greater than 3, $\kappa < 0$ and the central weight becomes negative leading to some numerical issues as outlined in [1]. Thus the fully symmetric $2N + 1$ Unscneted Transform sigma points are second moment equivalent and can integrate polynomials of degree 3 or less accurately.

3) $2n$ Cubature points-CKF: The authors of [9] provide a mathematically rigourous and elegant way of of finding the second moment equivalent $2N$ cubature points that can again accurately integrate polynomials of degree 3 or less. Spherical-radial coordinates are employed to find these $2N$ symmetric cubature points on the principal axis. Contrary to the $2N + 1$ Unscented transform sigma points, they only have $2N$ cubature points on the principal axis and no point on the mean. This is equivalent to setting $\kappa = 0$ in the Unscented transform sigma points thus making the central weight $w_0 = 0$. An alternate derivation, using the present framework, to the one given in [9] is provided which is less rigourous but more intuitive. As the central weight w_0 is 0, we get w_1 from (12), r_1 from (9) as

$$w_1 = 1/(2N) \quad \text{and} \quad r_1 = \sqrt{N} \tag{19}$$

The absolute error in 4^{th} moment equation (10) is

$$|2N^2 \frac{1}{2N} - 3| \equiv |N - 3| \tag{20}$$

For systems till dimension 3, the Unscented Transform points can better match one of the 4^{th} order moment than the CKF-Cubature points. After dimension 3 both the method have error in the 4th order moments as they were not designed to capture all the 4th order moments.

B. Sigma points that are 4th moment equivalent

Sigma points that are 4^{th} moment equivalent can completely satisfy the moment constraint equations till 4^{th} order. The following definition is introduced

M^{th}-Conjugate axis: For a N-Dimensional system where $M \leq N$,the axis that are constructed from all the combinations, including the sign permutations, of the set of principal axis taken M at a time. For example, the N^{th}-Conjugate set of axis for N-Dimensional system have 2^N distinct points or 2^{N-1} axis and the 2^{nd}-Conjugate set of axis for N-Dimensional system has $2N(N-1)$ distinct points or $N(N-1)$ axis. We label the set of M^{th} conjugate axis as c^M, where the points are listed as c_i^M, where $i = 1, 2, ..., 2^M \binom{N}{M}$

$$c_i^M \in \{(\pm\sigma_{n_1} \pm \sigma_{n_2} \pm ... \pm \sigma_{n_M}) | \{n_1, n_2, ..., n_M\} \subset D\} \tag{21}$$

For zero mean and identity covariance, each point on the M^{th}-Conjugate axis is at a distance of $|c_i^M| = \sqrt{M}$ from the origin.

1) Sigma points constrained to principal axis and 2^{nd}-Conjugate axis: In appendex IV of [10], the authors provides a method to calculate the cubature points that can capture all the fourth order moments. But this method also suffers from the presence of a negative/zero weight for dimensions higher than or equal to 4.The reason can be shown analytically. Let the points be chosen such that 1 point of weight w_0 lies on the origin, $2N$ points of weight w_1 each lie on the principal axis at a distance of r_1 and $2N(N-1)$ points of weight w_2 each lie on the 2^{nd}- Conjugate axis at a distance $r_2|c_i^2|$. Following the similar framework, the moment constraint equations till order 4 are

$$2r_1^2w_1 + 4(N-1)r_2^2w_2 = 1 \tag{22}$$
$$2r_1^4w_1 + 4(N-1)r_2^4w_2 = 3 \tag{23}$$
$$4r_2^4w_2 = 1 \tag{24}$$
$$1 - 2Nw_1 - 2N(N-1)w_2 = w_0 \tag{25}$$

The main variables to solve for are r_1, w_1, r_2, w_2 from (22)-(24) and then w_0 is found from (25). Thus there are 4 main variables and 3 equations indicating that there needs to be some kind of optimization procedure. The authors of [10] chooses to minimize the error in one of the 6th order moment. Before the optimization is framed, one can simplify the constraint equations as

$$w_2 = \frac{1}{4r_2^4}, w_1 = \frac{4-N}{2r_1^4}, \quad r_1^2r_2^2 = r_1^2(N-1) + r_2^2(4-N) \tag{26}$$

There is only one constraint in (26) to solve interms of the variables r_1 and r_2. Even though it can be solved by minimizing the 6th moment error one would still result in a negative/zero weight above dimension 3.

2) Sigma points constrained to the principal axis and the N^{th}-Conjugate axis: In this section we solve for sigma points that are 4th order equivalent, fully symmetric and has all the weights to be positive. The method is abbreviated as 'CUT4'. A similar work can be found in formula (IV) of [11], where the cubature rule is developed for uniform distribution. The points are selected such that $2N$ points of weight w_1 each lie on the principal axis at a distance of r_1 and 2^N points of weight w_2 lie on the N^{th}- Conjugate axis at a distance of $r_2\sqrt{N}$. Once the points are enumerated as in Table (III) and substituted into the moment constraint equations till order 4, the set of equations (27)-(29) are to be solved. This general set of equations were formed by first forming them separately for dimension 3,4 and 5 and then finding a general set of equations. This set of generalized equations were numerically tested to be exact till dimension 10. Though it is also shown to be analytically correct for any dimension.

$$2r_1^2w_1 + 2^Nr_2^2w_2 = 1 \tag{27}$$
$$2r_1^4w_1 + 2^Nr_2^4w_2 = 3 \tag{28}$$
$$2^Nr_2^4w_2 = 1 \tag{29}$$

And the central weight at the mean is given by

$$1 - 2Nw_1 - 2^Nw_2 = w_0 \tag{30}$$

TABLE II

CUT4: OPTIMIZED SOLUTION FOR $N = 1$ AND $N = 2$

Variable	$N = 1$	$N = 2$
r_1	1.4861736616297834	2.6060099476935847
r_2	3.2530871022700643	1.190556300661233
w_0	0.5811010092660772	0.41553535186548973
w_1	0.20498484723245053	0.021681819434216532
w_2	0.00446464813451093	0.12443434259941118
No. of points	5	9

TABLE III

SIGMA POINTS FOR CUT4

	Position	Weights
$1 \le i \le 2N$	$X_i = r_1\sigma_i$	$W_i = w_1$
$1 \le i \le 2^N$	$X_{i+2N} = r_2c_i^N$	$W_{i+2N} = w_2$
Central weight	$X_0 = \mathbf{0}$	$W_0 = w_0$
$n = 2N + 2^N\,(+1)$		

The central weight can be assigned a value such that the square of the 6th moment constraint error $(2r_1^6w_1 + 2^Nr_2^6w_2 - 15)^2$ is minimized subjected to the constraint of (27)-(29) or even to be 0 thus reducing one point. The solution of the former scheme is shown in Table (II) for dimensions 1 and 2 while the latter scheme is shown for dimensions greater than 2 .

In summary for $N > 2$

$$r1 = \sqrt{\frac{N+2}{2}}, \qquad r2 = \sqrt{\frac{N+2}{N-2}} \tag{31}$$

$$w_1 = \frac{1}{r_1^4} = \frac{4}{(N+2)^2}, \quad w_2 = \frac{1}{2^Nr_2^4} = \frac{(N-2)^2}{2^N(N+2)^2} \tag{32}$$

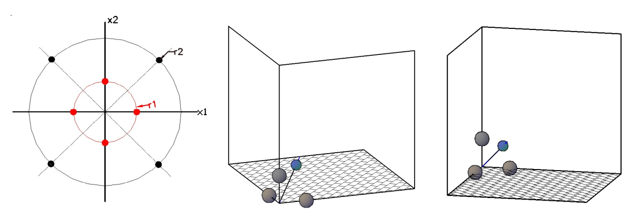

Fig. 1. 2D and 3D Cubature points to satisfy 4th order moments

C. Sigma points that are 6^{th} moment equivalent

This section describes a procedure to solve all the moment constraint equations till 6^{th} order for dimensions $N \le 9$. Firstly the moments till 6^{th} order for a Normal PDF are evaluated, then the moment constraint equations are formed by selecting the appropriate axis. The points on the principal axis have each a weight of w_1 and are at a distance r_1 from the origin. The points on the N^{th}-Conjugate axis have each a weight of w_2 and are located at a distance of $r_2\sqrt{N}$. The third set of points have weight w_3 and are located at a distance of $r_3\sqrt{2}$ along the 2^{nd}-Conjugate axis. The points are shown in

TABLE IV
SIGMA POINTS FOR CUT6, ($N \le 6$)

	Position	Weights
$1 \le i \le 2N$	$X_i = r_1\sigma_i$	$W_i = w_1$
$1 \le i \le 2^N$	$X_{i+2N} = r_2 c_i^N$	$W_{i+2N} = w_2$
$1 \le i \le 2N(N-1)$	$X_{i+2N+2^N} = r_3 c_i^2$	$W_{i+2N+2^N} = w_3$
Central weight	$X_0 = \mathbf{0}$	$W_0 = w_0$
$n = 2N^2 + 2^N + 1$		

TABLE V
SIGMA POINTS FOR CUT6, ($7 \le N \le 9$)

	Position	Weights
$1 \le i \le 2N$	$X_i = r_1\sigma_i$	$W_i = w_1$
$1 \le i \le 2^N$	$X_{i+2N} = r_2 c_i^N$	$W_{i+2N} = w_2$
$1 \le i \le 4n(n-1)(n-2)/3$	$X_{i+2N+2^N} = r_3 c_i^3$	$W_{i+2N+2^N} = w_3$
Central weight	$X_0 = \mathbf{0}$	$W_0 = w_0$
$n = 2N + 2^N + 4n(n-1)(n-2)/3 + 1$		

Fig(2) and enumerated in Table (IV).The moments for the continuous normal PDF are

$$\begin{aligned} E[X_i^2] &= 1, & E[X_i^4] &= 3, & E[X_i^2X_j^2] &= 1 \\ E[X_i^6] &= 15 & E[X_i^4X_j^2] &= 3 & E[X_i^2X_j^2X_k^2] &= 1 \end{aligned} \quad (33)$$

As the points chosen are fully symmetric only the moment with all even powers are to be satisfied. The general set of moment constraint equations is
For $N < 7$

$$2r_1^2w_1 + 2^N r_2^2 w_2 + 4(N-1)r_3^2 w_3 = 1 \quad (34)$$
$$2r_1^4w_1 + 2^N r_2^4 w_2 + 4(N-1)r_3^4 w_3 = 3 \quad (35)$$
$$2^N r_2^4 w_2 + 4r_3^4 w_3 = 1 \quad (36)$$
$$2r_1^6w_1 + 2^N r_2^6 w_2 + 4(N-1)r_3^6 w_3 = 15 \quad (37)$$
$$2^N r_2^6 w_2 + 4r_3^6 w_3 = 3 \quad (38)$$
$$2^N r_2^6 w_2 = 1 \quad (39)$$
$$1 - 2Nw_1 - 2^N w_2 - 2N(N-1)w_3 = w_0 \quad (40)$$

Solving the 3 equations (37)-(39) for the weights analytically ,

$$w_1 = \frac{8-n}{r_1^6}, \quad w_2 = \frac{1}{2^n r_2^6}, \quad w_3 = \frac{1}{2r_3^6} \quad (41)$$

Once the weights are symbolically solved, they can be substituted into (34)-(36), then there are only 3 polynomial equations in terms of r_1,r_2 and r_3. This reduced system of equations is much easier to solve than the original system of equations. Above dimension 6 the central weight becomes negative and above 8 one weight starts to become negative, hence this procedure is valid for $N \le 6$. A very similar method having the same drawback is presented in [12], for which the author derives the cubature rule for degree 7 and for specific dimensions $N = 3, 4, 6, 7$.

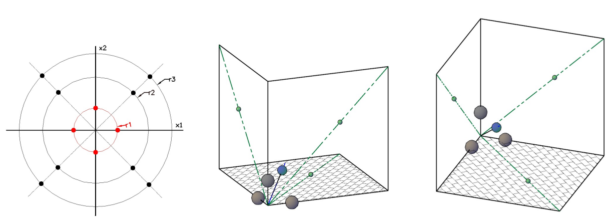

Fig. 2. 2D and 3D Sigma points to satisfy 6th order moments

The problem of negative weight can be solved by choosing another set of axis as, For $7 \le N \le 9$, choose 3^{rd}-Conjugate axis instead of 2^{nd}-conjugate axis. The set of 6th moment constraint equations in terms of the points, as enumerated in Table (V), constrained to principal axis (r_1,w_1), N^{th}-conjugate axis (r_2,w_2) and 3^{rd}- conjugate axis (r_3,w_3) are For $7 \le N \le 9$

$$2r_1^2w_1 + 2^N r_2^2 w_2 + 4(N-1)(N-2)r_3^2 w_3 = 1 \quad (42)$$
$$2r_1^4w_1 + 2^N r_2^4 w_2 + 4(N-1)(N-2)r_3^4 w_3 = 3 \quad (43)$$
$$2^N r_2^4 w_2 + 8(n-2)r_3^4 w_3 = 1 \quad (44)$$
$$2r_1^6w_1 + 2^N r_2^6 w_2 + 4(N-1)r_3^6 w_3 = 15 \quad (45)$$
$$2^N r_2^6 w_2 + 8(n-2)r_3^6 w_3 = 3 \quad (46)$$
$$2^N r_2^6 w_2 + 8r_3^6 w_3 = 1 \quad (47)$$
$$1 - 2Nw_1 - 2^N w_2 - (4n(n-1)(n-2)/3)w_3 = w_0 \quad (48)$$

Exploiting the linearity in the weights, one can again reduce the overall order and number of variables in the system. This set of equations (42)-(47) are preferred till dimension 9 only as above dimension 9 the central weight becomes negative. But if one allows the presence of negative weight the equations can be solved untill dimension 13 after which some roots become complex. In summary

D. *Sigma points that are 8^{th} order equivalent*

This section describes an *attempt* to solve all the moment constraint equations till order 8 by choosing appropriate axis. We describe the analysis in this section to be an attempt as we were able to 'exactly' solve the 8th order moment constraint equations till only dimension 6, above which the weight becomes negative. We provide an analysis that is done separately for each dimension. The procedure presented might apparently look laborious or in a way de-motivating but the *striking* results do infact justify our *endeavour*. A new set of axis is defined for a N-D system with zero mean and Identity covariance as follows
N^{th}-Scaled Conjugate axis: For a N-Dimensional system ,the axis that are constructed from all the combinations,including sign permutations, of the set of principal axis such that in every combination exactly one principal axis is scaled by a scaling parameter 'h' .The N^{th}-Scaled Conjugate set of axis for N-Dimensional system has $N2^N$ distinct points or $N2^{N-1}$ axis. We label the set of N^{th}-Scaled conjugate axis as $s^N(h)$,

and the points are listed as $s_i^N(h)$ where $i = 1,2,3,...,N2^N$

$$s_i^N(h) \in \{(\pm h\sigma_{n_1} \pm \sigma_{n_2} \pm \ldots \pm \sigma_{n_N}) | \{n_1, n_2, ..., n_N\} \subset D\} \tag{49}$$

The procedure followed is described for a general dimension of $N \leq 6$.

1) The first set of points are on the principal axis such that $X_i^{(1)} = r_1\sigma_i$ and $W_i^{(1)} = w_1$.
2) The second set of points are on the N^{th}-Conjugate axis such that $X_i^{(2)} = r_2 c_i^N$ and $W_i^{(2)} = w_2$.
3) The third set of points are on the 2^{nd}-Conjugate axis such that $X_i^{(3)} = r_3 c_i^2$ and $W_i^{(3)} = w_3$.
4) The fourth set of points are again on the N^{th}-Conjugate axis such that $X_i^{(4)} = r_4 c_i^N$ and $W_i^{(4)} = w_4$.
5) The fifth set of points are on the 3^{rd}-Conjugate axis such that $X_i^{(5)} = r_5 c_i^3$ and $W_i^{(5)} = w_5$.
6) The sixth set of points are on the N^{th}-Scaled Conjugate axis such that $X_i^{(6)} = r_6 s_i^N(h)$ and $W_i^{(6)} = w_6$.
7) The scaling parameter 'h' has to be appropriate chosen
8) Enumerate the points and form the moment constraint equations till order 8 using the moments in (50).

The sigma point set is $\{X^{(1)}, X^{(2)}, X^{(3)}, X^{(4)}, X^{(5)}, X^{(6)}, X_0\}$ and $\{W^{(1)}, W^{(2)}, W^{(3)}, W^{(4)}, W^{(5)}, W^{(6)}, W_0\}$. The even/non-zero moments till order 8 are

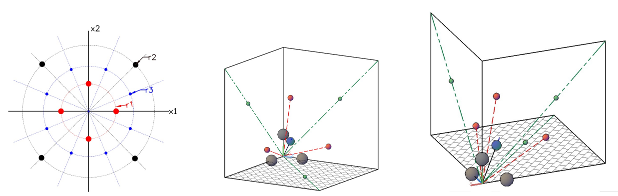

Fig. 3. 2D and 3D Sigma points to satisfy 8th order moments

$$\begin{aligned}
E[X_i^2] &= 1, & E[X_i^4] &= 3, & E[X_i^2X_j^2] &= 1,\\
E[X_i^6] &= 15, & E[X_i^4X_j^2] &= 3, & E[X_i^8] &= 105,\\
E[X_i^6X_j^2] &= 15, & E[X_i^4X_j^4] &= 9, & E[X_i^4X_j^2X_k^2] &= 3,\\
E[X_i^2X_j^2X_k^2] &= 1, & E[X_i^2X_j^2X_k^2X_l^2] &= 1, &&
\end{aligned} \tag{50}$$

A particular case of 5D system is developed as an example. The set of moment equations for 5D system till order 8 are derived and solved for the corresponding sigma points. The points are enumerated in the procedure described and once substituted into the appropriate moment constraint equations result in exactly 11 moment constraint equations as in (51).

$$\begin{aligned}
2r_1^2w_1 + 32r_2^2w_2 + 16r_3^2w_3 + 32r_4^2w_4 + 48r_5^2w_5 + 128r_6^2w_6 + 32h^2r_6^2w_6 &= 1\\
2r_1^4w_1 + 32r_2^4w_2 + 16r_3^4w_3 + 32r_4^4w_4 + 48r_5^4w_5 + 128r_6^4w_6 + 32h^4r_6^4w_6 &= 3\\
32r_2^4w_2 + 4r_3^4w_3 + 32r_4^4w_4 + 24r_5^4w_5 + 96r_6^4w_6 + 64h^2r_6^4w_6 &= 1\\
2r_1^6w_1 + 32r_2^6w_2 + 16r_3^6w_3 + 32r_4^6w_4 + 48r_5^6w_5 + 128r_6^6w_6 + 32h^6r_6^6w_6 &= 15\\
32r_2^6w_2 + 4r_3^6w_3 + 32r_4^6w_4 + 24r_5^6w_5 + 96r_6^6w_6 + 32h^2r_6^6w_6 + 32h^4r_6^6w_6 &= 3\\
32r_2^6w_2 + 32r_4^6w_4 + 8r_5^6w_5 + 64r_6^6w_6 + 96h^2r_6^6w_6 &= 1\\
2r_1^8w_1 + 32r_2^8w_2 + 16r_3^8w_3 + 32r_4^8w_4 + 48r_5^8w_5 + 128r_6^8w_6 + 32h^8r_6^8w_6 &= 105\\
32r_2^8w_2 + 4r_3^8w_3 + 32r_4^8w_4 + 24r_5^8w_5 + 96r_6^8w_6 + 32h^2r_6^8w_6 + 32h^6r_6^8w_6 &= 15\\
32r_2^8w_2 + 4r_3^8w_3 + 32r_4^8w_4 + 24r_5^8w_5 + 96r_6^8w_6 + 64h^4r_6^8w_6 &= 9\\
32r_2^8w_2 + 32r_4^8w_4 + 8r_5^8w_5 + 64r_6^8w_6 + 64h^2r_6^8w_6 + 32h^4r_6^8w_6 &= 3\\
32r_2^8w_2 + 32r_4^8w_4 + 32r_6^8w_6 + 128h^2r_6^8w_6 &= 1
\end{aligned} \tag{51}$$

These set of 11 equations that are formed from points as described in the procedure tend to work very well till dimension 6 after which some weights become negative. Though this is done separately for each dimension it is a one-time procedure, as the results for each dimension till 6 can be evaluated and the points can be stored. As the sigma points are evaluated for an i.i.d set of normal random variables, each sigma point can undergo an affine transformation based on the mean and covariance matrix of the arbitrary Normal PDF. A detailed description of such transformation is given in chapter one of [2] and proposition 4.2 of [9].These 11 moment constraint equations can be solved in a similar manner to the previous sections. Firstly by analytically solving for the 6 weights from the last 6 equations, which is easily achieved with any symbolic computation environment and then substituting them into the first 5 equations leaves one with only 5 equations of reduced order and 6 variables r_1, r_2, r_3, r_4, r_5 and r_6. One could minimize the error in 10th moment or as our objective is to solve these set of equations one could assume a value for a variable and solve all the other 5 variables. Table (VI) is computed by assuming a value for r_5 and solving the 5 equations in 5 variables. In 2D case, there are only 8 moment constraint equations and hence only 8 variables r_1, r_2, r_3, r_4 and w_1, w_2, w_3, w_4 to solve for. For the 3D case, there are 10 moments constraint equations and the 10 variables to solve for are r_1, r_2, r_3, r_4, r_6 and w_1, w_2, w_3, w_4, w_6. The solutions for all dimensions $2 \leq N \leq 6$ are shown in Table (VI). The solutions have been verified till dimension 6 and can be directly used to generate the sigma points as mentioned in the procedure. One has to be careful as while rounding off the solution to even the 4th decimal place could result in significant error.

III. Results

This section illustrates few examples using the methods mentioned in this paper.

A. Polynomial functions

All the methods are compared against the Gauss Hermite Product rule for multidimensional integral as it is known to exactly integrate polynomials.In general for a polynomial of degree $2m-1$ the Gauss Hermite Product rule for N^{th}-Dimensional integral requires m^N cubature points. In every case the integral that is being evaluated in the domian $(-\infty, \infty)$ is of the form

$$\mu = E[f(\mathbf{X})] = \int f(\mathbf{X}) N(\mathbf{X}, 0|P) d\mathbf{X} \tag{52}$$

where $\mathbf{X} = [X_1, X_2, ..., X_N]^T$ and $f(X)$ is a polynomial function such that $f : R^N \rightarrow R$ i.e. f is real valued. The gaussian weighting function has mean $\mathbf{0}$ and covariance matrix denoted as P, appropriately defined in each case. A sample

TABLE VI

SOLUTIONS FOR $2 \leq N \leq 6$, 8^{th} MOMENT CONSTRAINT EQUATIONS

Variable	2D	3D	4D	5D	6D
r_1	2.068136061121187	2.255137265545780	2.201709071472343	2.314370817280745	2.449489742783178
r_2	0.8491938499087475	0.7174531274600530	0.7941993714175681	0.8390942773980102	0.8938246941221211
r_3	1.138654980847415	1.843019437068797	1.872574360506295	1.830752125326649	1.732050807568877
r_4	1.861619935018895	1.558481032725744	1.329116430064565	1.397039743064496	1.531963037906212
r_5	–	–	2	2	2
r_6	–	1.305561500466050	1.125865581272049	1.113478632736702	1.095445115010332
w_1	0.04382264267013926	0.024631993437193266	0.01811008737283111	0.010529034221546607	0.006172839506172839
w_2	0.1405096621714662	0.08151009408908164	0.032063273384586845	0.015144019639537572	0.006913443044833937
w_3	0.0009215768861610588	0.009767235524166815	0.006614353755080834	0.0052828996967816825	0.004115226337448559
w_4	0.01240953967762697	0.00577248937435553	0.003489906522946932	0.0010671298950159158	0.0002183265828666806
w_5	–	–	0.0006510416666666666	0.0006510416666666666	0.0006510416666666666
w_6	–	0.000279472936899139	0.00025218336987488566	0.00013776017592074394	0.00007849171328446504
h	3	2.74	3	3	3

TABLE VII

GH VS CUT4: % REL. ERROR AND NO. OF POINTS

method	n_1	μ_1 % error	n_2	μ_2 % error
GH2	8	52.36	1024	16.39
GH3	27	$4.89e-014$	59049	$7.23e-012$
CUT4	14	0	1044	$6.72e-012$

TABLE VIII

GH VS CUT6: % REL. ERROR AND NO. OF POINTS

method	n_3	μ_3 % error	n_4	μ_4 % error
GH3	81	12.45	19683	4.18
GH4	256	$2.31e-013$	262144	$1.37e-009$
CUT6	49	$6.49e-013$	1203	$6.26e-009$

TABLE IX

GH VS CUT8: % REL. ERROR AND NO. OF POINTS

method	n_5	μ_5 % error	n_6	μ_6 % error
GH4	1024	3.45	4096	2.49
GH5	3125	$4.73e-012$	15625	$9.29e-012$
CUT8	355	$7.52e-012$	745	$6.63e-012$

function is taken from [9] as

$$f(X) = (\sqrt{1+X^TX})^p \tag{53}$$

where the value of p is varied in each case to get various degree polynomials. The notations GH2,GH3,GH4$\cdots$ stand for the Gauss Hermite Product rule with 2,3,4$\cdots$ points in 1 dimension, thereby $2^N, 3^N, 4^N \cdots$ total points. In each case the relative error in the integral value, the corresponding covariance matrix of the weight function used and the total number of points are represented by μ_i, P_i and n_i respectively.

1) Polynomials of degree 5 or less: Taking $p = 4$ gives a polynomial of degree 4. Table (VII) compares the integral value evaluate by 4th moment equivalent sigma points (CUT4) and the Gauss Hermite product rule. The Gauss Hermite product rule would need atleast 3^N for a N-Dimensional system. μ_1 and μ_2 are calculated using a covariance matrices P_1 (randomly generated) and P_2 (identity matrix) respectively. The relative error is calculated using GH5. The number of points required by the CUT4 method is significantly less than that required Gauss Hermite rule while having the same order of accuracy. This difference in number of points becomes extremely contrasting for higher dimensions such as $N = 10$ thus highlighting the advantage of CUT4.

$$P_1 = \begin{bmatrix} 114.2595 & 90.1397 & 8.9751 \\ 90.1397 & 92.2504 & 29.1237 \\ 8.9751 & 29.1237 & 84.0908 \end{bmatrix}, P_2 = 100I_{(10\times10)}$$

2) Polynomials of degree 7 or less: Taking $p = 6$ gives a polynomial of degree 6. Table (VIII) compares the integral value evaluate by 6th moment equivalent sigma points (CUT6) and the Gauss Hermite product rule. The Gauss Hermite product rule would need atleast 4^N for a N-Dimensional system. μ_3 and μ_4 are calculated using a covariance matrices P_3 and P_4 respectively. The relative error is calculated using GH7. Hence with the sheer reduction in number of points used by CUT6 one could capture the integral with same order of accuracy as that of Gauss Hermite . Where $P_3 = 100I_{(4\times4)}$ and $P_4 = 100I_{(9\times9)}$

3) Polynomials of degree 9 or less: Taking $p = 8$ gives a polynomial of order 8. Table (IX) compares the integral value evaluate by 8th moment equivalent sigma points (CUT8) and the Gauss Hermite product rule. The Gauss hermite product rule would need atleast 5^N. μ_5 and μ_6 are calculated using a covariance matrices P_5 and P_6 respectively. The relative error is calculated using GH9. CUT8 is able to match GH5 to same order of accuracy with fewer number of points. Where $P_5 = 100I_{(5\times5)}$, $P_6 = 100I_{(6\times6)}$

B. Non-polynomial nonlinearity

1) Polar to Cartesian coordinates: The example on polar to cartesian coordinates transformation from [8] is one way to graphically illustrulate the advantage in capturing higher order moments. The example is re-simulated in Fig(4), using the parameters decribed in [8] as $\mu_r = 1$, $\mu_\theta = 90$, $\sigma_r = 0.02^2$, $\sigma_\theta = 15^2$. The mean and 1-σ contours are plotted for each method. All the methods,except CKF, seem to do equally well compared to the Monte Carlo simulations using 3×10^6 samples in 2D. CKF predicts the mean correctly but covariance is slightly under estimated due to error in all the 4th moments. The Unscented Transform can do better as having $\kappa = 1$ captures one 4th order moment. For the same example when the parameters are changed as $\mu_r = 50$, $\mu_\theta = 0$, $\sigma_r = 0.02^2$, $\sigma_\theta = 30^2$, though not a realistic choice, the higher moment methods tend to do better as seen in Fig(5). The CKF slightly under estimates the covariance while the UT and GH3 slightly over estimates the covariance. CUT4 in 2D can also capture one of the 6th order moment and hence does better than GH3.

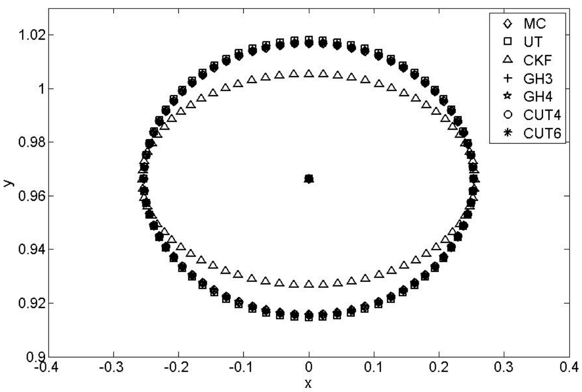

Fig. 4. Polar to Cartesian, $\mu_r = 1, \mu_\theta = 90, \sigma_r = 0.02^2, \sigma_\theta = 15^2$

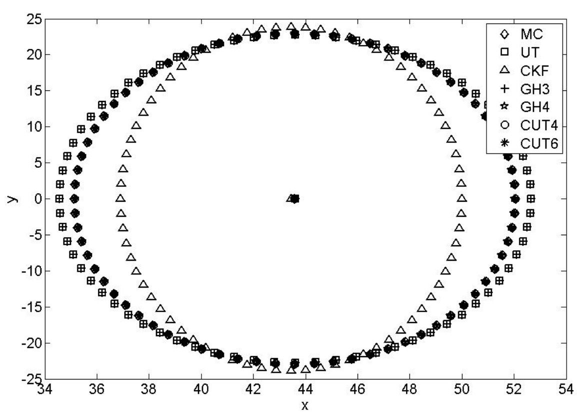

Fig. 5. Polar to Cartesian, $\mu_r = 50, \mu_\theta = 0, \sigma_r = 0.02^2, \sigma_\theta = 30^2$

2) $p = -3$: Having p to be negative makes the function to behave as a delta function and hence difficult to integrate using quadrature points. This may be because most of the quadrature points fall in the region where the function value is near to zero and hence does not help in the integral value. The convergence of the integral is highly effected by increasing the covariance of the gaussian PDF. This is characterized by the oscillatory type behaviour of the integral. Even taking a large number of points does not always guarantee the convergence. But by decreasing the magnitude of covariance, most of the quadrature points tend to fall within the region where the function value is significant. Under this condition one can consider the integral value to converge. This example has been shown just to motivate the fact that 'as one can capture more moments the integral value tends to converge, this statement is in agreement with equation (2). The covariance is taken as $P = 0.1I$ where I is the identity matrix of appropriate dimension. The covariance is intentionally scaled down by 0.1 to make the integral converge. In Fig(6), the Gauss Hermite product rule is used to compare the convergence of the integral in 2D,3D,4D,5D and 6D(belonging to each curve in Fig(6)) as the number of points are increased. The points are increased from 3 to 7 in single dimension resulting in GH3,GH4,GH5,GH6 and GH7 for each curve. In Fig(7), the CUT4,CUT6 and CUT8 is used in 2D,3D,4D,5D and 6D to compare the performance of the methods in evaluating the same integral. From Fig(7), it can be seen that CUT8 can achieve a relative error just below 0.5% in all dimensions considered. To achieve the same accuracy of 0.5%, the Gauss Hermite Product rule would need atleast 4 points(GH4) in a single dimension, thus 4^N in total. In Fig(8), the number of points required by GH4 and CUT8 are compared for each dimension.

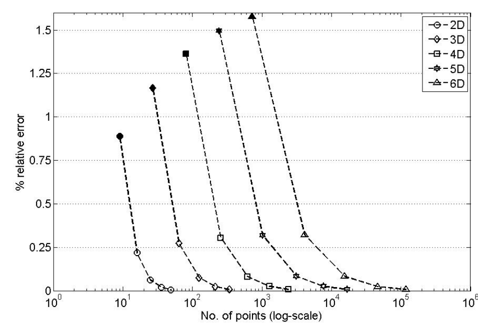

Fig. 6. % rel. error for GH3 to GH7 in each Dim.

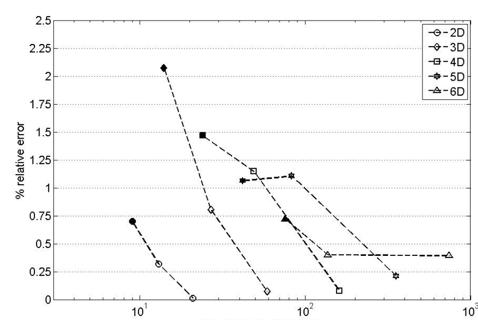

Fig. 7. % rel. error for CUT4,CUT6 and CUT8 in each Dim.

IV. CONCLUSIONS

Integrals involving normal weight function are extensively used in areas of statistics and nonlinear filtering.......

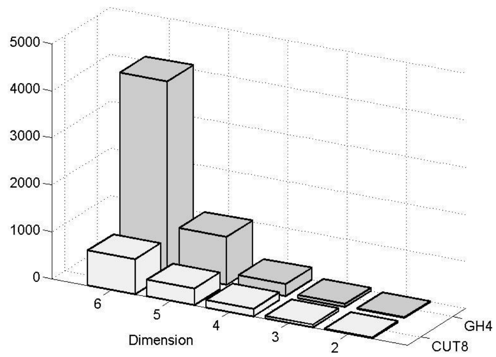

Fig. 8. GH4 vs CUT8: Points required to achieve 0.5% rel error

It has been customary to use the Guass-Hermite product rule to evaluate these integrals but this not only involves a lot of computational cost but even computational time. Especially for applications in online or real-time filtering, one would prefer a cubature rule with as minimal points as possible without the compromise in accuracy. This has been a motivating factor to develop higher order sigma points. Thus we have been able to provide a sigma point set that is 4th order equivalent, a sigma point set that is 6th order equivalent till dimension 9 and finally a sigma point set that is 8th order equivalent till dimension 6. When integrating polynomial functions, each sigma point set can be considered a direct replacement for the equivalent Gauss Hermite product rule of same order.

V. ACKNOWLEDGMENTS

REFERENCES

[1] Stroud A. H., Secrest D, "‘Gaussian Quadrature Formulas"’,Prentice hall, 1966

[2] Stroud A. H., "‘Approximate Calculation of Multiple Integrals"’, Prentice Hall, 1971

[3] Stroud A. H.,"’Numerical Integration Formulas of Degree Two"’,Mathematics of Computation, Vol. 14, No. 69 (Jan., 1960), pp. 21-26

[4] Philip R., Nira R.,"’ Perfectly Symmetric Two-Dimensional Integration Formulas with Minimal Numbers of Points"’,Mathematics of Computation, Vol. 23, No. 108 (Oct., 1969), pp. 765-779.

[5] Robert Piessens and Ann Haegemans,"’Cubature Formulas of Degree Nine for Symmetric Planar Regions"’,Mathematics of Computation, Vol. 29, No. 131 (Jul., 1975), pp. 810-815

[6] Richard Franke,"’Minimal Point Cubatures of Precision Seven for Symmetric Planar Regions"’,SIAM Journal on Numerical Analysis, Vol. 10, No. 5 (Oct., 1973), pp. 849-862

[7] Michalowicz J. V., Nicholas J M,...,"’A general Isserlis theorem for mixed Gaussian random variables "’, Statistics and probability letters...

[8] S. J. Julier, J. K. Ulhmann and Whyte,"’A new method for nonlinear transformation of means and covariances in filters and estimators"’, IEEE Trans Automat. Control vol.45, no. 3,pp 472-482 Mar. 2000

[9] I Arasaratnam and Simon Haykin,"’Cubature Kalman"’,IEEE transactions on Automat. Control, vol 54, no. 6, June 2009

[10] SIMON J. JULIER, AND JEFFREY K. UHLMANN,"’Unscented Filtering and Nonlinear Estimation"’,PROCEEDINGS OF THE IEEE, VOL. 92, NO. 3, MARCH 2004

[11] A. H. Stroud,"’Some Fifth Degree Integration Formulas for Symmetric Regions"’,Mathematics of Computation, Vol. 20, No. 93 (Jan., 1966), pp. 90-97

[12] A. H. Stroud,"’Some Seventh Degree Integration Formulas for Symmetric Regions"’,SIAM Journal on Numerical Analysis, Vol. 4, No. 1 (Mar., 1967), pp. 37-44