

# The Conjugate Unscented Transform - An Approach to Evaluate Multi-Dimensional Expectation Integrals

Nagavenkat Adurthi

Graduate Student

nagavenk@buffalo.edu

Puneet Singla

Assistant Professor

psingla@buffalo.edu

Tarunraj Singh

Professor

tsingh@buffalo.edu

University at Buffalo, The State University of New York, Amherst, NY 14260-4400

**Abstract**—This paper presents an extension to the unscented transformation to evaluate expectation integrals in general  $N$ -dimensional space by satisfying higher order moment equations. New sets of sigma points are defined to satisfy moment equations up to order 8. The proposed methodology can be used as an efficient alternative to Gaussian quadrature rule with significantly reduced number of function evaluations but without any loss in accuracy. The proposed methodology is compared to the Gauss Hermite product rule and many other computational algorithms by considering three benchmark problems. Numerical simulation results illustrate the effectiveness of the proposed methodology in computing high dimension expectation integrals with significantly reduced number of function evaluations.

## I. INTRODUCTION

Numerous fields of science and engineering present the problem of uncertainty characterization and propagation through nonlinear dynamic systems with stochastic excitation and uncertain initial conditions. One may be interested in predicting the probability of collision of asteroid with Earth, diffusion of toxic materials through atmosphere, control of movement and planning of actions of autonomous systems, the optimization of financial investment policies, or simply the computation of the prediction step in the design of a Bayes filter. All these applications require the study of the relevant stochastic system and involve computing multi-dimensional expected value integrals involving appropriate probability density function (pdf). Analytical expressions for these multi-dimension integrals exist only for linear systems and only for few moments. For example, the well celebrated Kalman filter provides the analytical expressions for mean and covariance of linear system subject to Gaussian white noise and Gaussian initial condition errors. However, one often do not have direct analytical solution for these integrals and have to approximate integral values by making use of computational methods.

Several computation techniques exist in the literature to approximate the expectation integral values, the most popular being Monte Carlo (MC) methods, Gaussian Quadrature Rule, Unscented Transformation (UT) and Cubature methods. All these methods involve an approximation of the expectation integral as a weighted sum of integrand values at specified points within the domain of integration. These methods basically differ from each other in the generation of these specific points. For example, MC methods involves random samples from the specified pdf while Gaussian quadrature scheme involves deterministic points carefully

chosen to reproduce exactly the integrals involving polynomial functions. Both deterministic Gaussian quadrature and MC methods are very popular but both methods require extensive computational resources and effort, and becomes increasingly infeasible for computing expectation integrals in high-dimension space [1].

For one-dimensional integrals, one needs  $m$  quadrature points according to the Gaussian quadrature scheme to reproduce the expectation integrals involving  $2m - 1$  degree polynomial functions. However, for a generic  $N$ -dimensional system, one needs to take the tensor product of 1-dimensional  $m$  quadrature points and hence one would require a total of  $m^N$  quadrature points (also known as cubature points for  $N \geq 2$ ) [1]. Even for a moderate dimension system involving 6 random variables, the number of points required to evaluate the expectation integral with only 5 points along each direction is  $5^6 = 15,625$ . This is a huge number of points that might be computationally expensive to use especially when the evaluation of function at each cubature point itself can be an expensive procedure. But fortunately the Gaussian quadrature rule is *not minimal* for  $N \geq 2$  and there exists cubature rules with reduced number of points [1]. This forms the basis of our motivation to the work presented in this paper.

An extensive amount of work has been done in this field to develop cubature/quadrature rules which can exactly reproduce integrals involving polynomials of degree less than or equal to ' $d = 2m - 1$ ' with fewer points, i.e., less than  $m^N$ . A good description of non-product Gaussian cubature rules, particularly for symmetric density functions, that are second, third and fifth degree exact in any dimension is provided in [2]. In general to develop a cubature rule of degree ' $d$ ' that is applicable to any dimension is still an open problem. A cubature rule of degree 2 with  $N + 1$  points and a cubature method of degree 3 with  $2N$  points for a general centrally symmetric weight function (such as the normal and uniform pdf) in  $N$ -dimensions is developed in [3]. A fully symmetric integration rules with minimal points for 2-dimension are developed that are exact for degree 9 – 15 in [4]. A 19 point cubature rule is developed, for symmetric regions in 2-dimensions that are exact for degree 9 or less in [5]. For a 2-dimentional integral, with symmetric weight function 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 2-dimensional systems. The highly acclaimed Unscented Transform (UT) with only  $2N+1$  points is a degree 3 cubature method, where these cubature points are now called *sigma points* [7]. Similar to the UT method, a more recent development is the Cubature Kalman filter (CKF) which is again exact to degree 3 but uses only  $2N$  points [8]. It is to be noticed that the sigma/cubature points of  $2N+1$  UT and  $2N$  CKF, though they have different elegant approaches, are infact *special cases of a more general work* previously done in [3] for any symmetrical probability distribution function.

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 a similar manner and tries to overcome some of these difficulties.

The primary objective of this paper can be stated as "To find a fully symmetric sigma/cubature point set with all positive weights and reduced number of points that is *equivalent* to the set of cubature points of Gaussian quadrature product rule of *same order*". *Equivalent to same order* implies that for a polynomial of order  $2m-1$  in generic  $N$ -dimensions, both the new reduced sigma point set from the proposed method known as Conjugate Unscented Transform method ('CUT') and the  $m^N$  quadrature points from the Gaussian quadrature product rule result in same order of relative percentage error. The organization of paper is as follows: firstly, we review few popular cubature rules, particularly the Unscented Transform. A intuitive understanding of how these methods tend to pick the sigma/cubature points provides a valuable insight upon which our development of the framework/methodology for the Conjugate Unscented Transform is based. Secondly, we describe this framework/methodology to evaluate the sigma point sets for the three methods we intend to propose along with numerical results to show the efficacy of each method compared to Gauss Hermite Product rule.

## II. EXPECTATION INTEGRAL AND CUBATURE METHODS

Let us consider the problem of computing expected value of a function  $f(x)$  with respect to a Gaussian density function, Without loss of generality consider the mean and covariance of the the Gaussian function to be zero and unity, respectively.

$$E[f(x)] = \int f(x)N(x,0|1)dx \quad (1)$$

Taking the Taylor series expansion of  $f(x)$  about the mean and substituting in the aforementioned expression leads to

$$E[f(x)] = f(0) + \nabla f(0)E[x] + \frac{1}{2!}\nabla^2 f(0)E[x^2] + \frac{1}{3!}\nabla^3 f(0)E[x^3] + \frac{1}{4!}\nabla^4 f(0)E[x^4] \dots \quad (2)$$

Notice that the problem of evaluating the expected value of nonlinear function  $f(x)$  has reduced to computing higher order moments of  $x$ . Thus by increasing the number of

TABLE I  
MOMENT CONSTRAINT EQUATIONS FOR 2D TILL ORDER 4

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

terms in the Taylor series expansion, one can obtain more accurate value of the expectation integral of (1). Furthermore, consider the discrete approximation of (1) as a weighted average of  $f(x)$  evaluated at each quadrature/sigma point set  $(x_1, x_2, \dots, x_n)$  with corresponding weights  $(w_1, w_2, \dots, w_n)$ :

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

Now, substitution of the Taylor series expansion about the mean for  $f(x_i)$  in (3) leads to:

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

Comparing the coefficients of  $f(0)$  and derivatives of  $f(x)$  evaluated at 0 from (4) and (2) leads to the following set of equations known as the *moment constraint equations*:

$$\sum_{i=1}^n w_i x_i^m = E[x^m], \quad m = 0, 1, 2, \dots \quad (5)$$

Notice that these constraint equations simply conveys that the sigma points  $x_i$  should satisfy the infinite moment equations in the the domain of the pdf, i.e.,  $x$ -space. Thus, if  $f(x)$  is a polynomial function of degree  $d$  then sigma points should be chosen to reproduce first  $d$  moments of the pdf exactly to guarantee the exact evaluation of the expectation integral of (1). For non-polynomial functions, one have to judiciously choose the number of moment equations which should be satisfied. These one-dimensional moment constraint equations can easily be generalized for generic  $N$ -dimensional system as illustrated in [7]. For example, Table I shows all the moment equations till  $4^{th}$  order moment in two dimensions.

As discussed in the last section, one intends to seek a set of sigma points that satisfy all the moment constraint equations till the prescribed order  $d$ . Many computational methods have been discussed in literature to choose a set of points which satisfy the moment equations up to a desired order with Gaussian quadrature methods being most popular. Although Gaussian quadrature methods can be generalized to satisfy these moment equations to a prescribed order  $d$ , these methods suffer from *curse of dimensionality* with number of points increasing exponentially with increase in dimension. As an alternative to Gaussian quadrature scheme, many non-product methods [3], [4], [5], [6], [7], [8], [9] and [10] have been discussed in literature to seek minimal number of sigma points to satisfy the desired number of moment equations. However, the development of a cubature rule of

degree ‘ $d$ ’ that is applicable to any dimension is still an open problem. The main purpose of this work is to extend the effort of few previous non-product rules, particularly the conventional unscented transformation to compute the expectation integrals involving Gaussian pdf while satisfying the higher order moment equations.

### III. UNSCENTED TRANSFORMATION

In this section, we briefly discuss the main idea of the conventional UT. The conventional UT algorithm involves the selection of  $2N + 1$  sigma points such that first two moment equations are satisfied [7]. For a zero mean and unity covariance, these sigma points are chosen such that one of the point lies at the origin and the remaining  $2N$  points are constrained to lie symmetrically about the mean on  $N$ -orthogonal axes, known as principal axes, at a distance  $r_1$ . The variable  $r_1$  is chosen to be  $\sqrt{N + \kappa}$  such that the second order moment equations are satisfied and an analytical expression for sigma point weights can be obtained.

$$X_0 = \mu \quad W_0 = \kappa / (N + \kappa) \quad (6)$$

$$X_i = \mu + (\sqrt{(N + \kappa)P})_i \quad W_i = 1 / [2(N + \kappa)] \quad (7)$$

$$X_{i+N} = \mu - (\sqrt{(N + \kappa)P})_i \quad W_{i+N} = 1 / [2(N + \kappa)] \quad (8)$$

where  $i$  varies from 1 to  $N$ . It should be mentioned that the principal axes for the generation of sigma points correspond to each column of the identity covariance matrix or the square root of the covariance matrix in general. To find the sigma points of an arbitrary Normal pdf, one can always find the sigma points with respect to a normal pdf with zero mean and unity covariance and then apply *affine transformations* to each point. A description of such affine transformations can be found in [2], [8]. Furthermore, the tuning parameter  $\kappa$  is generally chosen such that one of the  $4^{th}$  moment equation error can be minimized which leads to the constraint of  $N + \kappa = 3$ . This methodology works well for systems till dimension 3 after which the weight corresponding to the central point becomes negative, which can lead to indefinite covariance matrix.

Tanne et al. [11] have developed higher order unscented transformation method for a general pdf in one dimensional system which can satisfy the moment equations up to  $8^{th}$  order. Lévesque generalizes the method presented in [11] to develop unscented transformation for generic  $N$ -dimension system [12]. However, this method fails to capture cross moments. It is clear that obtaining  $4^{th}$  order accurate sigma points for higher dimensions is still a challenge. One has to note that most of the non-product rules tend to judiciously select a specific structure of the points. Conventional UT and CKF methods are based upon the fact that constraining the points to lie on principal axis leads to *the minimal number of points* required to satisfy moment equations up to order 2.

In summary, the set of nonlinear moment constraint equations are very tedious to solve for a general  $N$ -dimensional system. The UT provides an insight that “the points can be constrained to some *carefully* selected axis” and one needs to solve for only the distances of these points from the origin

and their corresponding weights. One has to be judicious in the selection of appropriate axis. The set of moment constraint equations along with these additional constraints makes the system easier to solve. However, choosing an appropriate axis is not a *trivial* problem.

Let us consider the problem of selecting sigma points such that first four moment equations are satisfied. Assuming i.i.d random variables  $(X_1, X_2, \dots, X_N)$  with zero mean and identity covariance lead to following higher order moments till order 4:

$$\begin{aligned} E[X_i^2] &= 1, & E[X_i X_j] &= 0, & E[X_i^4] &= 3 \\ E[X_i^3 X_j] &= 0, & E[X_i^2 X_j^2] &= 1, & E[X_i^2 X_j X_k] &= 0 \\ E[X_i X_j X_k X_l] &= 0, \end{aligned} \quad (9)$$

Where  $\{i, j, k, l\} \subset \{1, 2, 3, \dots, N\}$ . Like in the conventional UT scheme, 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  $w_1$  leads to following constraint equations:

$$E[X_i^2] \equiv 2r_1^2 w_1 = 1 \quad (10)$$

$$E[X_i^4] \equiv 2r_1^4 w_1 = 3 \quad (11)$$

$$E[X_i^2 X_j^2] \equiv 0 \neq 1 \quad (12)$$

$$1 = w_0 + 2Nw_1 \quad (13)$$

It can be seen that the elegant choice of  $r_1 = \sqrt{N + \kappa}$  in (10) leads to the same weights in (7), (8) and eventually the central weight (13) leads to (6). It should be noticed that odd power moments of (9) are already satisfied due to the symmetry of  $2N + 1$  sigma points and thus, one has only three distinct moment equations given by (10)-(12). Furthermore, the  $4^{th}$  order cross moment  $E[X_i^2 X_j^2]$  in (12) cannot be satisfied by this particular selection of symmetric sigma points that are constrained to lie only on the principal axis. In fact, no cross moment of any order can be satisfied by the selection of symmetric points constrained to lie only on the principal axis. Thus the UT and CKF have error in the 4th order moments. In summary, either one needs to break the symmetry of the sigma points lying on the principal axis or one needs to look for alternative axis (other than principal axis) to define new sigma points for the satisfaction of higher order moment equation. Breaking the symmetry of sigma point is not desirable since that will require us to include odd order moment equations in formulation also. In this work, the second option is pursued to define additional axis called as *conjugate axis*.

### IV. CONJUGATE UNSCENTED TRANSFORMATION

In this section, an extension of conventional unscented transformation is described to guarantee the satisfaction of moment equations of generic order  $d$ . Based on the advantages and disadvantages of conventional UT, the main steps of this new approach can be enumerated as follows:

*Step 1:* Making use of the fact that the Gaussian PDF is symmetric, the new sigma points are constrained to lie on appropriately defined symmetric axis known as *Conjugate*

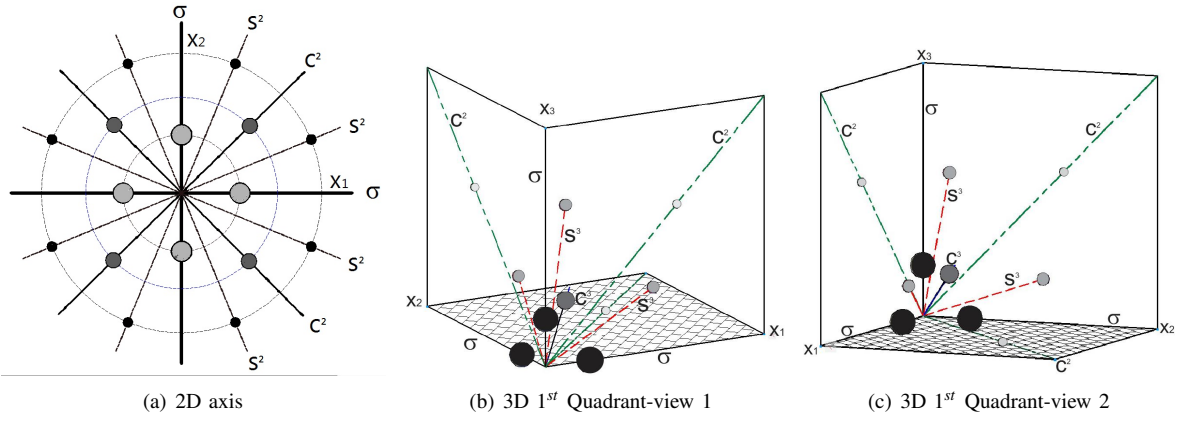


Fig. 1. Symmetric set of points and axis 2D and 3D space

*Axis.* For a generic  $N$ -dimensional system,  $M^{th}$ -conjugate axis with  $M \leq N$ , is defined as 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 will have  $2^N$  distinct points or  $2^{N-1}$  axis and the  $2^{nd}$ -Conjugate set of axis will have  $2N(N-1)$  distinct points or  $N(N-1)$  axis. The set of  $M^{th}$  conjugate axis labelled as  $c^M$ , where the points are listed as  $c_i^M$  with  $i = 1, 2, 3 \dots 2^M \binom{N}{M}$ .

$$c_i^M \in \{(\pm\sigma_{n_1} \pm \sigma_{n_2} \dots \pm \sigma_{n_M}) | \{n_1, n_2, \dots, n_M\} \subset D\} \quad (14)$$

where, the set  $D = \{1, 2, 3, \dots, N\}$  and  $\sigma_k \in \{\pm(I)_k | k \in D\}$ ,  $k = 1, 2, 3, \dots, 2N$  denotes the  $2N$  distinct points on  $N$  principal axes. Each point on the principle axis is at a unit distance from the origin and each point on the  $M^{th}$ -conjugate axis is at a distance of  $|c_i^M| = \sqrt{M}$  from the origin. The principle axes are labeled as  $\sigma$  and the conjugate axes are labeled as  $c^M$  with corresponding points in Fig. 1(a) for 2-dimensional case while Figs. 1(b) and 1(c) show two different perspective views of one of the quadrant for a 3-dimensional case. It should be mentioned that all the 8 quadrants in the 3-dimensional case are symmetrical.

*Step2:* Sigma points on the same set of symmetric axis are assumed to be equidistant from the mean and have equal weight. For the  $i^{th}$  set of axis, the distance scaling variables are labeled as  $r_i$  and weight variables are labeled as  $w_i$ . Notice that this helps us in automatic satisfaction of any odd order moments like in the conventional UT algorithm.

*Step 3:* The moment constraint equations of desired order are derived in terms of unknown variables  $r_i$  and  $w_i$  by making use of the higher order moments formed from Isserlis theorem [13]. For example the moment constraint equations till order 4 for a 2-dimensional system are shown in Table I. As the points  $(x_i, y_i)$  are constrained to lie on a line, they can be replaced in terms of a single variable namely the distance  $r_i$  from the origin.

*Step 4:* These nonlinear set of equations are solved for  $r_i$  and  $w_i$ , that result in the required sigma point set.

In the following subsections, these new sigma point sets are explicitly derived by making use of *conjugate axis* to

satisfy the moment constraint equations of order 4, 6 and 8 respectively.

#### A. 4<sup>th</sup> Conjugate Unscented Transformation (CUT4)

Sigma points that are 4<sup>th</sup> moment equivalent can completely satisfy the moment constraint equations up to 4<sup>th</sup> order. Julier and Uhlmann have described an approach in the Appendix of [14] to calculate the sigma points that can capture all the fourth order moments. But this method suffers from the presence of a negative/zero weight for dimensions higher than or equal to 4. To gain more insight these equations are re-derived and solved. Let us consider that one sigma point of weight  $w_0$  lies on the origin,  $2N$  points of weight  $w_1$  lie symmetrically on each principal axis at a distance of  $r_1$  and  $2N(N-1)$  points of weight  $w_2$  lie symmetrically on each 2<sup>nd</sup>-conjugate axis at a distance  $r_2|c_i^2|$ . Thus, the moment constraint equations up to order 4 are given as:

$$2r_1^2w_1 + 4(N-1)r_2^2w_2 = 1 \quad (15)$$

$$2r_1^4w_1 + 4(N-1)r_2^4w_2 = 3 \quad (16)$$

$$4r_2^4w_2 = 1 \quad (17)$$

$$1 - 2Nw_1 - 2N(N-1)w_2 = w_0 \quad (18)$$

Notice that variables  $r_1, w_1, r_2, w_2$  can be solved from (15)-(17) and then  $w_0$  can be found from (18). Thus, there are 4 main variables and 3 equations leading to an underdetermined system of equations indicating that there needs to be some kind of optimization procedure. In [14], authors chooses to minimize the error in one of the 6<sup>th</sup> order moment to solve this underdetermined system of equations. Before the optimization is framed, one can simplify the constraint equations as follows:

$$w_2 = \frac{1}{4r_2^4}, w_1 = \frac{4-N}{2r_1^4}, r_1^2r_2^2 = r_1^2(N-1) + r_2^2(4-N) \quad (19)$$

As a consequence of this simplification, there is only one constraint in (19) to solve in terms of the variables  $r_1$  and  $r_2$ . Even though it can be solved by minimizing the 6<sup>th</sup> moment constraint violation, the solution still always result in a negative/zero weight above dimension 3. We address this

TABLE II  
SIGMA POINTS FOR CUT4

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

problem of negative weight by choosing a different set of axis called conjugate axis, and hence the label ‘CUT4’. We seek to solve for sigma points that are 4<sup>th</sup> order equivalent, fully symmetric and have positive weights. A similar approach from which we draw our inspiration, can be found in formula (IV) of [9], where a degree 4 cubature rule is developed for the uniform distribution.

The new sigma points are selected such that  $2N$  points of weight  $w_1$  lie symmetrically on each principal axis at a distance of  $r_1$  and  $2^N$  points of weight  $w_2$  lie on the  $N^{\text{th}}$ -conjugate axis at a distance of  $r_2 \sqrt{N}$ . Substitution of this particular selection of sigma points as summarized in Table (II) leads to the following 4th order moment constraint equations:

$$2r_1^2 w_1 + 2^N r_2^2 w_2 = 1 \quad (20)$$

$$2r_1^4 w_1 + 2^N r_2^4 w_2 = 3 \quad (21)$$

$$2^N r_2^4 w_2 = 1 \quad (22)$$

It should be mention that this aforementioned general set of equations were derived by first forming them separately for dimension 3,4 and 5 and then finding a general set of equations. The weight  $w_0$  corresponding to the central point at the mean is given as  $w_0 = 1 - 2Nw_1 - 2^N w_2$ . A unique value for central weight  $w_0$  can be obtained by minimizing the 6<sup>th</sup> moment constraint violation error  $(2r_1^6 w_1 + 2^N r_2^6 w_2 - 15)^2$  subjected to the constraint of (20)-(22). The optimized solution for dimension one and two is shown in Table III while the analytical solution for dimension greater than 2 is given as:

$$r_1 = \sqrt{\frac{N+2}{2}}, \quad r_2 = \sqrt{\frac{N+2}{N-2}} \quad (23)$$

$$w_1 = \frac{1}{r_1^4} = \frac{4}{(N+2)^2}, \quad w_2 = \frac{1}{2^N r_2^4} = \frac{(N-2)^2}{2^N (N+2)^2} \quad (24)$$

According to CUT4, the expected integral as in (1) would need 14 and 1044 sigma points to exactly evaluate the expectation of a polynomial function of degree 4 in 3- and 10-dimensional space where as the GH Product rule would require a minimal of 27 and 59,049 points, respectively. This striking difference is a motivation to develop a 6<sup>th</sup> moment equivalent method CUT6 in the next section.

#### B. 6<sup>th</sup> order Conjugate Unscented Transform: CUT6

This section describes a procedure to solve all the moment constraint equations up to 6<sup>th</sup> order for dimensions  $N \leq 9$ . In this case, once again each sigma point on the principal axis have been assigned a weight of  $w_1$  and are constrained to lie symmetrically at a distance  $r_1$  from the origin. Furthermore,

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

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

points on the  $N^{\text{th}}$ -conjugate axis have been chosen with a weight value of  $w_2$  and are constrained to lie symmetrically at a distance of  $r_2 \sqrt{N}$ . Finally, the third set of points are selected with weight being  $w_3$  and constrained to lie symmetrically at a distance of  $r_3 \sqrt{2}$  along the  $2^{\text{nd}}$ -conjugate axis. These points are enumerated in Table IV and the moments till order 6 are:

$$\begin{aligned} E[X_i^2] &= 1, & E[X_i^4] &= 3, & E[X_i^2 X_j^2] &= 1 \\ E[X_i^6] &= 15 & E[X_i^4 X_j^2] &= 3 & E[X_i^2 X_j^2 X_k^2] &= 1 \end{aligned} \quad (25)$$

As the points chosen are fully symmetric only the moment with all even powers needs to be satisfied. The general set of moment constraint equations for  $N < 7$  are given as:

$$2r_1^2 w_1 + 2^N r_2^2 w_2 + 4(N-1)r_3^2 w_3 = 1 \quad (26)$$

$$2r_1^4 w_1 + 2^N r_2^4 w_2 + 4(N-1)r_3^4 w_3 = 3 \quad (27)$$

$$2^N r_2^4 w_2 + 4r_3^4 w_3 = 1 \quad (28)$$

$$2r_1^6 w_1 + 2^N r_2^6 w_2 + 4(N-1)r_3^6 w_3 = 15 \quad (29)$$

$$2^N r_2^6 w_2 + 4r_3^6 w_3 = 3 \quad (30)$$

$$2^N r_2^6 w_2 = 1 \quad (31)$$

$$1 - 2Nw_1 - 2^N w_2 - 2N(N-1)w_3 = w_0 \quad (32)$$

Solving equations (29)-(31) for unknown weights leads to following analytical expressions:

$$w_1 = \frac{8-n}{r_1^6}, \quad w_2 = \frac{1}{2^n r_2^6}, \quad w_3 = \frac{1}{2r_3^6} \quad (33)$$

The substitution of aforementioned analytical expression for weight variables in (26)-(28) leads to 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 given by (26)-(32). In this case, the central weight becomes negative for  $N > 6$ . Above dimension 6 the central weight becomes negative and above 8 one weight starts to become negative, hence this procedure is valid for  $N \leq 6$ . A very similar method having the same drawback is presented in [10], for which the author has derived the 7<sup>th</sup> order equivalent set of sigma points valid only for specific dimensions  $N = 3, 4, 6, 7$ . Furthermore, the problem of negative weight can

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

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

be avoided by choosing  $3^{rd}$ -conjugate axis instead of  $2^{nd}$ -conjugate axis which makes the proposed approach valid for  $7 \leq N \leq 9$ . The set of  $6^{th}$  order moment constraint equations in terms of these new sigma points, as enumerated in Table V, are given as:

$$2r_1^2 w_1 + 2^N r_2^2 w_2 + 4(N-1)(N-2)r_3^2 w_3 = 1 \quad (34)$$

$$2r_1^4 w_1 + 2^N r_2^4 w_2 + 4(N-1)(N-2)r_3^4 w_3 = 3 \quad (35)$$

$$2^N r_2^4 w_2 + 8(n-2)r_3^4 w_3 = 1 \quad (36)$$

$$2r_1^6 w_1 + 2^N r_2^6 w_2 + 4(N-1)r_3^6 w_3 = 15 \quad (37)$$

$$2^N r_2^6 w_2 + 8(n-2)r_3^6 w_3 = 3 \quad (38)$$

$$2^N r_2^6 w_2 + 8r_3^6 w_3 = 1 \quad (39)$$

$$1 - 2Nw_1 - 2^N w_2 - (4n(n-1)(n-2)/3)w_3 = w_0 \quad (40)$$

Once again exploiting the linearity in the weights, one can solve for weights analytically followed by the solution of  $r_1$ ,  $r_2$  and  $r_3$  from the remaining three equations. Notice that this aforementioned set of equations (34)-(39) are preferred till dimension 9 as the central weight becomes negative for  $N > 9$ . But if one allows the presence of negative weight the equations can be solved until dimension 13 after which some roots become complex. The results show a contrasting difference in the number of points. For example CUT6 would need only 49 and 1203 function evaluations while the GH product rule would require a minimal of 256 and 262,144 function evaluations for the exact computation of the expectation integral of a polynomial function of degree 6 in 4- and 9- dimensional space, respectively.

#### C. $8^{th}$ order Conjugate Unscented Transform: CUT8

This section describes an *attempt* to solve all the moment constraint equations up to order 8 by choosing appropriate axes. The analysis presented in this section has been described as an attempt since the solution presented is valid for  $N \leq 6$  with one of the weight being negative for  $N > 6$ . A procedure is provided that needs to be performed separately for each dimension. The procedure presented might apparently look laborious or in a way de-motivating but the *striking* results do justify our *endeavour*. A new set of axis is defined for a generic  $N$ -dimensional system known as *scaled conjugate axis*:

*$N^{th}$ -Scaled Conjugate axis*: For a  $N$ -dimensional system, the scaled conjugate axis 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. The set of  $N^{th}$ -Scaled conjugate axis are labeled as  $s^N(h)$ , and the points are listed as  $s_i^N(h)$  where  $i = 1, 2, 3, \dots, N2^N$

$$s_i^N(h) \in \{(\pm h \sigma_{n_1} \pm \sigma_{n_2} \pm \dots \pm \sigma_{n_N}) | \{n_1, n_2, \dots, n_N\} \subset D\} \quad (41)$$

Fig. 1(a) shows the scaled conjugate axis labeled as  $s^N(h)$  for a two dimensional case with the scaling parameter  $h$  being 2.414. Furthermore, Figs. 1(b) and 1(c) shows the conjugate axis in two different perspective views with the scaling parameter  $h$  being 2.74. The generic procedure valid for  $N \leq 6$  is described as follows:

- 1) The first set of points are selected 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 constrained to lie 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 assumed to lie 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 assumed to lie 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 constrained to lie 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 lie on the  $N^{th}$ -Scaled Conjugate axis such that  $X_i^{(6)} = r_6 s_i^N(h)$  and  $W_i^{(6)} = w_6$ . The scaling parameter  $h$  has to be appropriately selected.
- 7) Enumerate the points and form the moment constraint equations up to order 8 using the moments in (42).

Let us denote the collection of all sigma points by  $X$ , i.e.,  $X = \{X^{(1)}, X^{(2)}, X^{(3)}, X^{(4)}, X^{(5)}, X^{(6)}, X_0\}$ . Now, the even order moment constraint equations up to order 8 are:

$$\begin{aligned} E[X_i^2] &= 1, & E[X_i^4] &= 3, & E[X_i^2 X_j^2] &= 1, \\ E[X_i^6] &= 15, & E[X_i^4 X_j^2] &= 3, & E[X_i^8] &= 105, \\ E[X_i^6 X_j^2] &= 15, & E[X_i^4 X_j^4] &= 9, & E[X_i^4 X_j^2 X_k^2] &= 3, \\ E[X_i^2 X_j^2 X_k^2] &= 1, & E[X_i^2 X_j^2 X_k^2 X_l^2] &= 1, \end{aligned} \quad (42)$$

To further illustrate the procedure, let us consider a particular case of 5-dimensional system. The aforementioned selection of sigma points leads to the 11 moment constraint equations:

$$\begin{aligned} 2r_1^2 w_1 + 32r_2^2 w_2 + 16r_3^2 w_3 + 32r_4^2 w_4 + 48r_5^2 w_5 + 128r_6^2 w_6 + 32h^2 r_6^2 w_6 &= 1 \\ 2r_1^4 w_1 + 32r_2^4 w_2 + 16r_3^4 w_3 + 32r_4^4 w_4 + 48r_5^4 w_5 + 128r_6^4 w_6 + 32h^4 r_6^4 w_6 &= 3 \\ 32r_2^2 w_2 + 4r_3^4 w_3 + 32r_4^4 w_4 + 24r_5^4 w_5 + 96r_6^4 w_6 + 64h^2 r_6^4 w_6 &= 1 \\ 2r_1^6 w_1 + 32r_2^6 w_2 + 16r_3^6 w_3 + 32r_4^6 w_4 + 48r_5^6 w_5 + 128r_6^6 w_6 + 32h^6 r_6^6 w_6 &= 15 \\ 32r_2^6 w_2 + 4r_3^6 w_3 + 32r_4^6 w_4 + 24r_5^6 w_5 + 96r_6^6 w_6 + 32h^2 r_6^6 w_6 + 32h^4 r_6^6 w_6 &= 3 \\ 32r_2^6 w_2 + 32r_4^6 w_4 + 8r_5^6 w_5 + 64r_6^6 w_6 + 96h^2 r_6^6 w_6 &= 1 \\ 2r_1^8 w_1 + 32r_2^8 w_2 + 16r_3^8 w_3 + 32r_4^8 w_4 + 48r_5^8 w_5 + 128r_6^8 w_6 + 32h^8 r_6^8 w_6 &= 105 \\ 32r_2^8 w_2 + 4r_3^8 w_3 + 32r_4^8 w_4 + 24r_5^8 w_5 + 96r_6^8 w_6 + 32h^2 r_6^8 w_6 + 32h^4 r_6^8 w_6 &= 15 \\ 32r_2^8 w_2 + 4r_3^8 w_3 + 32r_4^8 w_4 + 24r_5^8 w_5 + 96r_6^8 w_6 + 64h^2 r_6^8 w_6 + 64h^4 r_6^8 w_6 &= 9 \\ 32r_2^8 w_2 + 32r_4^8 w_4 + 8r_5^8 w_5 + 64r_6^8 w_6 + 64h^2 r_6^8 w_6 + 32h^4 r_6^8 w_6 &= 3 \\ 32r_2^8 w_2 + 32r_4^8 w_4 + 32r_6^8 w_6 + 128h^2 r_6^8 w_6 &= 1 \end{aligned} \quad (43)$$

TABLE VI  
SOLUTIONS FOR  $2 \leq N \leq 6$ ,  $8^{th}$  MOMENT CONSTRAINT EQUATIONS, CUT8

Variable	2D	3D	4D	5D	6D
$r_1$	2.0681360611	2.2551372655	2.2017090714	2.3143708172	2.4494897427
$r_2$	0.8491938499	0.7174531274	0.7941993714	0.8390942773	0.8938246941221211
$r_3$	1.1386549808	1.8430194370	1.8725743605	1.8307521253	1.7320508075
$r_4$	1.8616199350	1.5584810327	1.3291164300	1.3970397430	1.531963037906212
$r_5$	—	—	2	2	2
$r_6$	—	1.3055615004	1.1258655812	1.1134786327	1.0954451150
$w_1$	0.0438226426	0.0246319934	0.0181100873	0.0105290342	0.0061728395
$w_2$	0.1405096621	0.081510094	0.0320632733	0.0151440196	0.0069134430
$w_3$	0.0009215768	0.00976723555	0.006614353	0.0052828996	0.0041152263
$w_4$	0.0124095396	0.0057724893	0.0034899065	0.0010671298	0.0002183265
$w_5$	—	—	0.0006510416	0.0006510416	0.00065104166
$w_6$	—	0.0002794729	0.0002521833	0.00013776017	0.00007849171
$h$	3	2.74	3	3	3

These set of 11 equations are found to work very well up to  $N \leq 6$  after which some weights become negative. Though this procedure needs to be repeated for different  $N$ , it is a one-time procedure and the results for each dimension can be stored. These 11 moment constraint equations can be solved in a similar manner as mentioned in the previous sections. One can analytically solve for six weight unknowns from the last six equations. The substitution of analytical expressions for weight variables into the first 5 equations leads to 5 equations of reduced order and 6 variables  $r_1, r_2, r_3, r_4, r_5$  and  $r_6$  to be solved for. One can minimize the error in the  $10^{th}$  moment equation violation error to find a unique solution for all six variables. Table VI lists the numerical solution for sigma points for  $N \leq 6$ . It should be noticed that there are only 8 moment constraint equations for 2-dimensional case and hence the sigma points  $(X^{(5)}, X^{(6)})$  are unnecessary, in effect there are only 8 variables  $r_1, r_2, r_3, r_4$  and  $w_1, w_2, w_3, w_4$  needs to be solved. Similarly, there are 10 moment constraint equations for 3-dimensional system and hence the sigma points  $(X^{(5)})$  are dropped and the 10 variables to solve for are  $r_1, r_2, r_3, r_4, r_6$  and  $w_1, w_2, w_3, w_4, w_6$ . Using the Table VI, CUT8 method would need only 355 and 745 function evaluations to compute the expectation integral for a polynomial function of degree 8 in the 5- and 6-dimensional space, respectively whereas the GH product rule would need 3,125 and 15,625 function evaluations for the same 5- and 6-dimensional space respectively. It is clear that the number of function evaluations required by the CUT methods are significantly less than that required by the GH quadrature rule while having the same order of accuracy. This difference in number of points becomes extremely contrasting for computing integrals in higher dimension spaces.

## V. NUMERICAL RESULTS

In the previous sections, the CUT methods were compared to the Gauss Hermite Product rule using polynomial functions only. In this section, two benchmark non-polynomial functions are considered to further illustrate the effectiveness of the proposed *conjugate unscented transformation* method in numerical evaluation of expectation integrals involving Gaussian pdf. The notations such as GH3, GH4, ... imply Gauss-Hermite product rule with 3, 4, ... quadrature points in a single direction and hence  $3^N, 4^N, \dots$  in total.

### A. Non-polynomial functions

1) *Polar to Cartesian coordinates* ( $x = r \cos \theta, y = r \sin \theta$ ): This benchmark problem was introduced in [14] to illustrate the effectiveness of the conventional UT. The example is re-simulated in Fig. 2(a) while making use of same parameters as described in [14]. The mean and  $1-\sigma$  contours are plotted for each method in Fig. 2(a). All the methods, except CKF, seem to do equally well compared to the MC simulations using  $3 \times 10^6$  samples in 2-dimensional space. CKF predicts the mean correctly but covariance is slightly under estimated due to error in all the  $4^{th}$  moments. For the same example when the parameters are changed as  $E[r] = 50$ ,  $E[\theta] = 0$ ,  $\sigma_r = 0.02^2$ ,  $\sigma_\theta = 30^2$ , the inclusion of higher moment constraints tend to do better than the conventional UT as seen in Fig. 2(b). The CKF slightly under estimates the covariance while the UT and GH3 slightly over estimates the covariance. CUT4 in 2D can additionally capture one of the  $6^{th}$  order moment and hence does better than GH3 and UT.

2)  $f(\mathbf{X}) = (\sqrt{1 + \mathbf{X}^T \mathbf{X}})^\alpha$ ,  $\mathbf{X} \in \mathbb{R}^N$ : This benchmark problem was introduced in [8]. It has been discussed that computing the expectation of  $f(\mathbf{X})$  for negative values for  $\alpha$  is a challenging task since  $\alpha < 0$  leads to a delta-sequence functions. For simulation purposes,  $\alpha = -3$  and the covariance of the zero mean Gaussian pdf is assumed to be 0.1 times the identity matrix. The covariance is intentionally scaled down by 0.1 to make the integral value converge with reasonable number of points. Fig. 3(a) shows the relative error with respect to the true value of the expectation integral as a function of GH quadrature points along each dimension while varying the dimension of input space from 2 to 6. Fig. 3(b) shows the relative error for dimensions 2 to 6 using the CUT methods. Furthermore, Fig. 3(c) shows the minimal number of points required for each method to achieve at least 0.5% relative error. It is clear that the proposed CUT8 method yields the accuracy of 0.5% as the GH4 quadrature rule but with a smaller fraction of the number of points.

## VI. CONCLUSIONS

New sets of sigma points are developed to compute multi-dimension expectation integrals involving Gaussian pdf by constraining the points to lie on principal axis and newly define conjugate axis. These new sets of sigma points are

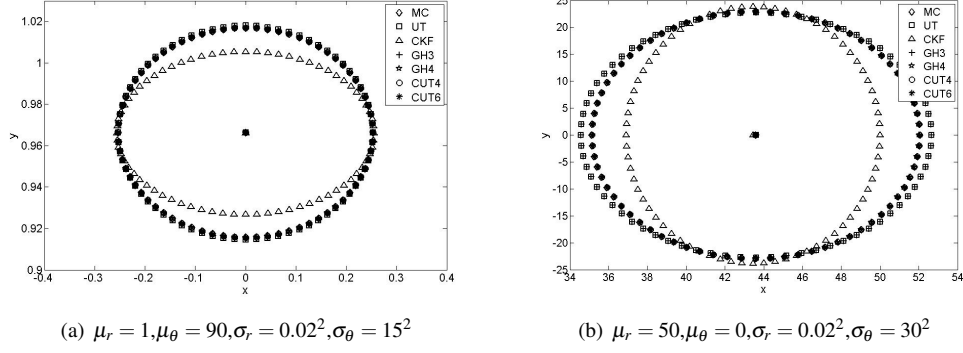


Fig. 2. Simulation Results for Polar to Cartesian Transformation. The notations GH2, GH3, GH4... stand for the Gauss Hermite Product rule with 2, 3, 4... points along each dimension.

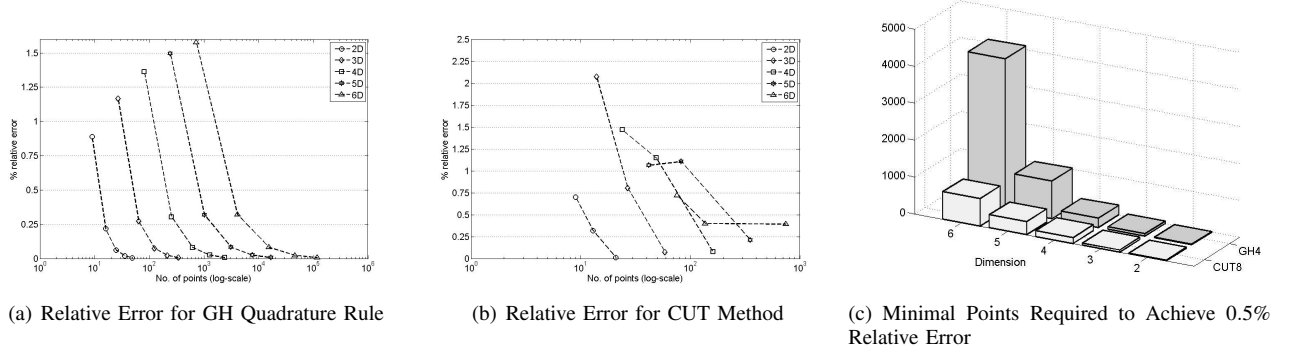


Fig. 3. GH vs. CUT: Simulation Results corresponding to  $f(\mathbf{X}) = (\sqrt{1 + \mathbf{X}^T \mathbf{X}})^{-3}$

guaranteed to exactly reproduce expectation integrals involving polynomial functions up to order 8 with significantly less number of points as compared to the conventional Gaussian quadrature rule. Numerical experiments involving non-polynomial functions are considered to show the efficacy of the developed methodology and numerical results obtained provides a basis for optimism.

## VII. ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Award No. CMMI- 0908403 and CMMI-1054759.

## REFERENCES

- [1] A. H. Stroud and D. Secrest, *Gaussian Quadrature Formulas*. Englewood Cliffs, NJ: Prentice Hall, 1966.
- [2] A. H. Stroud, *Approximate Calculation of Multiple Integrals*. Englewood Cliffs, NJ: Prentice Hall, 1971.
- [3] —, “Numerical integration formulas of degree two,” *Mathematics of Computation, American Mathematical Society*, vol. 14, no. 69, pp. 21–26, January 1960.
- [4] P. Rabinowitz and N. Richter, “Perfectly symmetric two-dimensional integration formulas with minimal numbers of points,” *Mathematics of Computation, American Mathematical Society*, vol. 23, no. 108, pp. 765–779, October 1969.
- [5] R. Piessens and A. Haegemans, “Cubature formulas of degree nine for symmetric planar regions,” *Mathematics of Computation, American Mathematical Society*, vol. 29, no. 131, pp. 810–815, July 1975.
- [6] R. Franke, “Minimal point cubatures of precision seven for symmetric planar regions,” *SIAM Journal on Numerical Analysis*, vol. 10, no. 5, pp. 849–862, October 1973.
- [7] S. Julier, J. Uhlmann, and H. Durrant-Whyte, “A new method for the nonlinear transformation of means and covariances in filters and estimators,” *IEEE Transactions on Automatic Control*, vol. AC-45, no. 3, pp. 477–482, March 2000.
- [8] I. Arasaratnam and S. Haykin, “A general isserlis theorem for mixed gaussian random variables,” *IEEE transactions on Automatic Control*, vol. 54, no. 6, pp. 1254 – 1269, June 2009.
- [9] A. Stroud, “Some fifth degree integration formulas for symmetric regions,” *Mathematics of Computation, American Mathematical Society*, vol. 20, no. 93, pp. 90–97, January 1966.
- [10] —, “Some seventh degree integration formulas for symmetric regions,” *SIAM Journal on Numerical Analysis*, vol. 4, no. 1, pp. 37–44, March 1967.
- [11] D. Tenne and T. Singh, “The higher order unscented filter,” vol. 3. Proc. Amer. Control Conf, 2003, p. 24412446.
- [12] J. F. Lvesque, “Unscented filtering and nonlinear estimation,” vol. 2. Proceedings of the Guidance, Navigation, and Control Conference (AIAA, March 2006, pp. 819830..
- [13] J. Michalowicz, J. Nichols, F. Bucholtz, and C. Olson, “A general isserlis theorem for mixed gaussian random variables,” *Statistics and Probability Letters*, vol. 81, no. 8, pp. 1233–1240, August 2011.
- [14] S. Julier and J. Uhlmann, “Unscented filtering and nonlinear estimation,” vol. 92. Proc. IEEE, March 2004, pp. 401–422.