Linear Regression Kalman Filtering Based on Hyperspherical Deterministic Sampling

Gerhard Kurz and Uwe D. Hanebeck

***Abstract*— Nonlinear filtering based on Gaussian densities is commonly performed using so-called Linear Regression Kalman Filters (LRKFs). These filters rely on sample-based approximations of Gaussian densities. We propose a novel sampling scheme that is based on decomposing the problem of sampling a multivariate Gaussian into sampling a Chi distribution and sampling uniformly on the surface of a hypersphere. The proposed sampling scheme has significant advantages compared to existing methods because it produces a user-selectable number of samples with uniform, nonnegative weights and it does not require any numerical optimization. We evaluate the novel method in simulations and provide comparisons to multiple state-of-the-art approaches.**

I. Introduction

Nonlinear estimation has been of interest for a long time and work in this area dates back to the 1960s, when the Extended Kalman Filter (EKF) was proposed [1]. Since then, numerous approaches have been investigated. An important class of methods are the so-called Linear Regression Kalman Filters (LRKFs) [2], which use sample-based statistical linearization in conjunction with the standard Kalman filter [3]. LRKFs, in particular the Unscented Kalman Filter (UKF) [4], have gained a lot of popularity due to their simplicity and the ability to provide good results in many scenarios at a reasonable computational cost. The key distinction between different LRKFs is the way they choose the set of samples used for the statistical linearization.

In this paper, we present a novel sampling scheme that can be used within the LRKF framework. The key idea is splitting the approximation of a multivariate standard Gaussian in $\mathbb{R}^d$ into approximation of the direction as well as the length of the vectors. The direction can be sampled according to a uniform distribution on the hypersphere and the length can be sampled according to a Chi distribution[1]. For this purpose, we use an equal area partitioning approach on the hypersphere in conjunction with a one-dimensional Chi distribution sampling based on minimization of the squared L^2 distance of the cumulative distribution functions. This idea is illustrated in Fig. 1. Our method creates a layered structure, where the samples are placed on nested hyperspheres.

The high-degree Cubature Kalman Filter (CKF) [5] also relies on similar ideas. It uses numerical integration rules that are based on splitting the integral into a spherical and a radial part. For the third-degree CKF [6], all samples are located on a sphere, and for the fifth-degree CKF, all samples—except for a sample at the origin—are located on a sphere.

The authors are with the Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology (KIT), Germany. E-mail: gerhard.kurz@kit.edu, uwe.hanebeck@ieee.org

[1]The original version of this paper mistakenly stated that the length follows a Gaussian distribution. This is the revised version of the paper, where we have corrected this issue and rerun all simulations using the correct distribution.

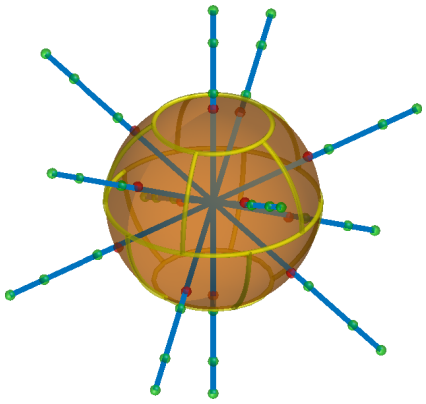


Fig. 1: Idea of the proposed deterministic sampling scheme. A uniform sampling on the sphere (red balls) is computed using equal area partitioning. Samples of one-dimensional Chi distribution (green balls) are placed on the lines emanating from the center of the sphere through the spherical samples.

The downside compared to our proposed method is that the number of samples is fixed by the dimension and the cubature rule. Also, for the fifth-degree rule, negative weights cannot be avoided for dimensions $d \geq 5$, which can cause problems that we will discuss later.

Furthermore, the proposed approach is somewhat similar to the filter by Huber et al. [7], which samples a one-dimensional Gaussian distribution and replicates it on each axis of the coordinate system to obtain samples for a multivariate Gaussian distribution. One of the key advantages of the proposed approach is that the sample locations are not limited to the axes of the coordinate system, which leads to much better results.

Another related approach is the randomized UKF by Straka et al. [8]. It takes multiple UKF sample sets and performs a random rotation as well as a random scaling on each. While this approach also relies on the separation of the direction from the length of the sample vectors, it is inherently nondeterministic, which makes it difficult to generate reproducible results. The quality of the approximation can vary strongly depending on the chosen random values. Another problem is the presence of negative weights.

Other LRKFs include the Gaussian Hermite Kalman Filter [9], [12], a filter based on a quadrature scheme that requires exponentially many samples with respect to the dimension. Furthermore, there is the Smart Sampling Kalman Filter (S^2KF) [10], which is based on placing the samples according to an optimality criterion that minimizes a distance function between the Gaussian distribution and the samples. A similar approach with additional moment constraints is taken in [13]. A downside of these methods is that a computationally expensive numerical optimization has to be performed to obtain the sample set. Thus, the optimization is usually

Sampling Scheme	Number of Samples	Weights	Deterministic	Numerical Optimization
UKF [4]	$2d+1$	neg. weights optional[a]	yes	no
RUKF [8]	multiple of $2d+1$	neg. weights	no	no
3rd degree CKF [6]	$2d$	equal	yes	no
5th degree CKF [5]	$2d^2+1$	neg. weights[b]	yes	no
GHKF [9]	m^d	nonequal, positive	yes	no
S^2KF [10]	arbitrary $\geq 2d$	equal	yes[c]	yes[d]
proposed	arbitrary $\geq 2d$	equal	yes	no

[a] uniform weights are always possible
[b] for dimensions ≥ 5
[c] depending on initialization
[d] an implementation of the numerical optimization procedure is available online [11]

TABLE I: Overview of LRKFs.

performed offline for a standard normal distribution and the samples are transformed to the current distribution using a Mahalanobis transformation [14, Sec. 3.3]. This makes it difficult to change the number of samples or dimensions at runtime. An overview of some of the most popular approaches is given in Table I.

In general, a sampling scheme for an LRKF has certain desirable properties. First of all, we want to maintain the first and second moment of the Gaussian distribution to ensure that the LRKF is equivalent to the optimal Kalman filter for linear systems. Matching the first two moments means that a Gaussian can be converted to samples and vice versa without losing any information, which is why all LRKFs considered here fulfill this property. It may be desirable to maintain higher moments of the Gaussian distribution, because Gaussian quadrature [15, Sec. 2.7] rules guarantee optimality of certain moments for polynomial systems up to a predetermined degree. In practice, however, it may be more important to have a fairly even distribution of the samples such that the space is well covered and the shape of the Gaussian distribution is well matched. It is also beneficial if the number of samples can be adjusted to perform a trade-off between computational effort and accuracy. Unfortunately, many approaches only allow a fixed number of samples or a choice that is limited to very coarse steps. Another issue is the computational effort needed to obtain the samples. Some approaches use very costly numerical optimization methods that have to be performed offline, in particular for many samples or dimensions. Moreover, we would like the sample set to be deterministic, so results are easily reproducible and the accuracy of the estimate does not depend on the choice of random numbers.

Finally, it is advantageous if the weights of the samples are as uniform as possible because in this case, all samples contribute equally to the result. In particular, samples with negative weights should be avoided as they can result in covariance matrices that are not positive definite, as we illustrate in the following example.

Example 1 (Negative Weights)**.** *Consider a standard normal distribution $\mathcal{N}(0,1)$, which we want to propagate through the simple yet nonlinear function $f(x) = (x-0.2)^2$.*

For this purpose, we use an LRKF with negative weights, say, the UKF with negative weight $\gamma_1 = -1$ ($W_0 = -1$ according to the notation used in [4, Sec. IV-A]). This yields the sample locations

$$[0, -\sqrt{2}/2, \sqrt{2}/2] \approx [0, -0.7071, 0.7071]$$

and weights $[\gamma_1, \gamma_2, \gamma_3] = [-1, 1, 1]$. The propagated samples are at locations

$$[\underline{r}_1, \underline{r}_2, \underline{r}_3] = [0.0400, 0.8228, 0.2572]$$

and their weights stay the same. We obtain the sample mean

$$\underline{\mu} = \textstyle\sum_{i=1}^{N} \gamma_i \cdot \underline{r}_i = 1.04 \ ,$$

which is identical to the analytic solution. However, the covariance is

$$\begin{aligned}\mathbf{C} &= \textstyle\sum_{i=1}^{N} \gamma_i \cdot (\underline{r}_i - \underline{\mu})(\underline{r}_i - \underline{\mu})^T \\ &= -1 \cdot 1 + 1 \cdot 0.0472 + 1 \cdot 0.6128 = -0.34 < 0 \ ,\end{aligned}$$

i.e., not positive definite. As a result, future filtering steps based on this covariance are impossible. A common practical solution is to omit steps that lead to non-positive-definite covariance matrices, which is clearly suboptimal.

II. Key Idea

In the following, we will show how to combine samples form a Chi distribution and uniform hyperspherical samples to obtain a sampling scheme for multivariate Gaussian densities. Consider a multivariate standard Gaussian distribution in d dimensions, which can be reformulated as

$$\begin{aligned}&\mathcal{N}(\underline{x}; \underline{0}, \mathbf{I}) = \frac{1}{(2\pi)^{d/2}} \exp\left(-\underline{x}^T \underline{x}/2\right) \\ &= \frac{1}{(2\pi)^{d/2}} \exp\left(-\|\underline{x}\|^2/2\right) = \frac{1}{(2\pi)^{(d-1)/2}} \mathcal{N}(\|\underline{x}\|; 0, 1) \ .\end{aligned}$$

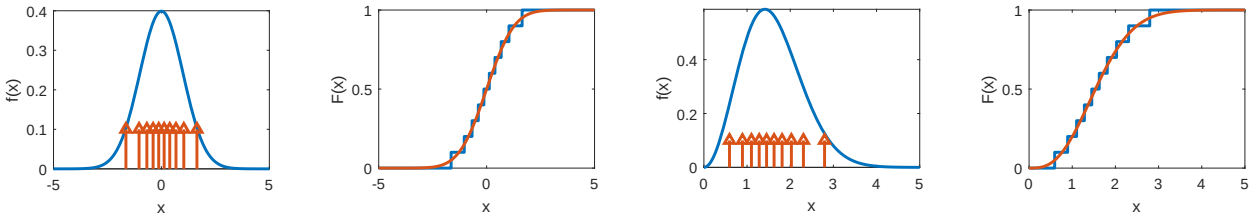
It can be seen that this distribution is uniform distribution with respect to the direction $\frac{\underline{x}}{\|\underline{x}\|}$. Note that the direction does not appear in the equation due to uniformity.

The distribution of the norm $r := \|x\| \geq 0$ can be derived according to

$$\begin{aligned}f(r) &= \int_{\{\underline{x} \in \mathbb{R}^d : \|\underline{x}\| = r\}} \mathcal{N}(\underline{x}; \underline{0}, \mathbf{I}) \, \mathrm{d}\underline{x} \\ &= \frac{1}{(2\pi)^{d/2}} \exp\left(-r^2/2\right) \int_{\{\underline{x} \in \mathbb{R}^d : \|\underline{x}\| = r\}} \mathrm{d}\underline{x} \\ &= \frac{1}{(2\pi)^{d/2}} \exp\left(-r^2/2\right) \frac{2\pi^{d/2}}{\Gamma(d/2)} r^{d-1} \\ &= \frac{r^{d-1} \exp\left(-r^2/2\right)}{2^{d/2-1}\Gamma(d/2)} \ ,\end{aligned}$$

where we use the fact that the surface of a $(d-1)$-dimensional sphere in d dimensional space is $\frac{2\pi^{d/2}}{\Gamma(d/2)} r^{d-1}$ with Gamma function $\Gamma(\cdot)$. As can be seen, the resulting density coincides with a Chi distribution with d degrees of freedom.

As a result, we can sample the Chi distribution on $\|\underline{x}\|$ and the uniform distribution on $\frac{\underline{x}}{\|\underline{x}\|}$ separately (using the



(a) Gaussian pdf. (b) Gaussian cdf. (c) Chi pdf. (d) Chi cdf.

Fig. 2: Example for a 1D deterministic approximation of a Gaussian distribution and a Chi distribution for $d = 3$ degrees of freedom.

techniques presented in Sec. IV and Sec. III) and then obtain samples for the multivariate Gaussian by considering their Cartesian product.

To obtain our sample set, we scale each hyperspherical sample $\underline{s}_i$ with each of the Chi distribution samples β_j according to

$$\{\underline{r}_1, \dots, \underline{r}_N\} = \{\underline{s}_i \cdot \beta_j | 1 \leq i \leq M, 1 \leq j \leq L\} \ ,$$

where M is the number of spherical samples and L is the number of Gaussian samples or layers. Hence, each Gaussian sample creates one layer of samples. The weights of the combined samples are obtained by multiplying the weights of the individual samples. In our case, we obtain uniform weights $\gamma_1 = \dots = \gamma_N = \frac{1}{M \cdot L}$ for all samples.

Samples for a non-standard Gaussian distribution $\mathcal{N}(\underline{\mu}, \mathbf{C})$ can be obtained by performing a Mahalanobis transformation, i.e, $\underline{r}_j^{\text{transformed}} = \underline{\mu} + \sqrt{\mathbf{C}} \cdot \underline{r}_j$ for $1 \leq j \leq N$, where $\sqrt{\mathbf{C}}$ can be computed using the Cholesky decomposition.

III. Deterministic Hyperspherical Sampling

In this section, we focus on the problem of computing a set of samples uniformly distributed on the surface of the unit hypersphere $S^{d-1} = \{\underline{x} \in \mathbb{R}^d : \|\underline{x}\| = 1\}$, i.e., the set of vectors with Euclidean norm 1. We seek to obtain a subset $\{\underline{s}_1, \dots, \underline{s}_M\} \subset S^{d-1}$ of $M \in \mathbb{N}$ vectors that covers the surface of the unit hypersphere evenly. While uniform random sampling on the surface of the unit hypersphere is very easy, deterministic sampling can be quite tricky. This problem has been studied quite extensively, especially for the case $d = 3$, i.e., the sphere in $\mathbb{R}^3$ [16]. It arises in many applications, sometimes in slight variations as far as the measure of uniformity is concerned. For instance, in physics, the Thompson problem [17], [18, Sec. 18.7] considers the question of distributing electrons on the surface of a sphere such that the Coulomb energy is minimized. Another related example is the Tammes problem in botany, which considers the distribution of pores on pollen grains [18, Sec. 18.9]. Optimal quantization, though usually on real vector spaces rather than the sphere, is also a closely related issue [19], [20]. Many approaches for the spherical problem can be found in literature, e.g., [21], [16], and also some for the hyperspherical case [22], [23].

We use the equal area partitioning approach proposed by Leopardi [24], [25], which partitions the surface of the sphere into regions of equal area. The regions are chosen such that their diameter, i.e., the largest distance between any two points in the region, is small. Leopardi proves an upper bound for the diameter that converges to zero if the number of regions goes to infinity. We use the center of each region as a hyperspherical sample. The algorithm is based on recursive partitioning of the hypersphere in d dimensions by reducing its partition to that of the hypersphere in $d-1$ dimensions. A MATLAB implementation is available online [26].

Leopardi's algorithm has a number of beneficial properties. As it is not based on numerical optimization of an optimality criterion, it is extremely fast, does not get stuck in local optima, and is independent of initialization. Also, it can be applied to an arbitrary number of dimensions $d \geq 2$ and an arbitrary numbers of samples $M \in \mathbb{N}$.

Due to the recursive construction, the resulting samples are not spread perfectly evenly on the hypersphere, but it can be verified empirically that the results are very close to a uniform distribution. Furthermore, the bound on the diameter of the regions and the symmetry of the construction provide theoretical guarantees. Examples for the resulting partitions and samples are shown in Fig. 4.

IV. Deterministic Chi Distribution Sampling

In this section, we consider the problem of approximating a Chi distribution with a set of samples. To be specific, we are looking for the parameters $w_1, \dots, w_L > 0$ and $\beta_1, \dots, \beta_L$ of a Dirac mixture

$$f(x; \beta_1, \dots, \beta_L, w_1, \dots, w_L) = \sum_{i=1}^{L} w_i \delta(x - \beta_i) \quad (1)$$

with $L \in \mathbb{N}$ components, where $\sum_{i=1}^{L} w_i = 1$ holds. In the following, we restrict ourselves to equally weighted samples, i.e., $w_1 = \dots = w_L = \frac{1}{L}$.

In [27], Schrempf et al. proposed a sampling scheme for one-dimensional densities based on minimization of the squared L^2 distance of cumulative distribution function. The distance measure for densities $f_1(\cdot)$ and $f_2(\cdot)$ is given by

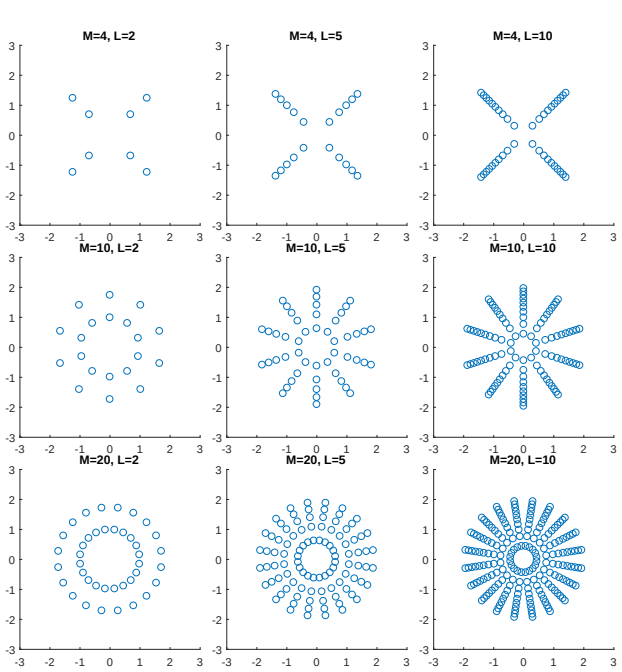
$$D = \int_{-\infty}^{\infty} (F_1(x) - F_2(x))^2 \, \mathrm{d}x \ ,$$

where $F_i(x) = \int_{-\infty}^{x} f_i(t) \, \mathrm{d}t$ is the cumulative distribution function. If $f_2(\cdot)$ is a Dirac mixture as defined in (1) and is assumed to have equal weights, then it can be shown that the optimal approximation of $f_1(\cdot)$ given by

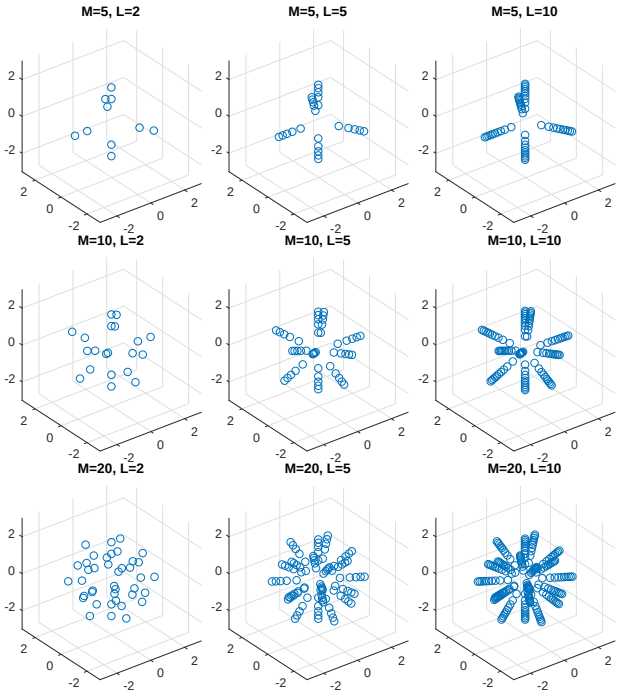
$$\beta_i = F_1^{-1}\big((2i-1)/(2L)\big) \quad (2)$$

according to [27, Theorem III.1].

However, this approach does not, in general, preserve any moments of the density $f_1(\cdot)$. For symmetric densities (such as the Gaussian), the first moment is maintained, but higher moments are not. In [7], Huber et al. introduced a constraint for the second moment to resolve this issue,



(a) Examples in 2D.



(b) Examples in 3D.

Fig. 3: Examples of the sample sets produced by the novel sampling scheme.

which requires numerical optimization. The main problem with this approach is that using the one-dimensional samples in multiple dimensions will once again violate the second moment constraint. For this reason, our approach does not use any moment constraints, but instead enforces the second moment retroactively in n dimensions by using the Mahalanobis transformation (see Sec. V).

As discussed in Sec. III, we obtain samples from the hypersphere, which correspond to rays starting at origin to a certain direction. Thus, we need to approximate the a Chi distribution for each ray by computing its inverse cdf as given by (2). For practical implementation, the square root of the Chi squared distribution inverse cdf can be used because implementations of this function are more readily available (e.g., `chi2inv` in MATLAB).

An example of the results can be seen in Fig. 2.

V. Moment Correction

As the samples obtained this way do not necessarily have exactly the mean and covariance of a standard normal distribution, we subtract the actual mean

$$\underline{\mu} := \mathbb{E}(\underline{x}) = \sum_{i=1}^{N} \gamma_i \underline{r}_i$$

from each sample. Then, we apply the Mahalanobis transformation to correct the covariance as is done in the S^{2}KF [10]. For this purpose, we compute the sample covariance

$$\mathbf{C} := \mathbb{E}(\underline{x}\underline{x}^T) = \sum_{i=1}^{N} \gamma_i \underline{r}_i \underline{r}_i^T$$

and obtain its Cholesky decomposition, i.e., we obtain an upper triangular matrix $\mathbf{R}$ such that $\mathbf{R}^T \cdot \mathbf{R} = \mathbf{C}$. Then, we multiply each sample from the left with $\mathbf{R}^{-1}$, which yields samples with covariance $\mathbf{I}$. Examples of the resulting sample sets in 2D and 3D for different values of L and M are shown in Fig. 3.

VI. Filter

We can use the samples derived above in the standard LRKF framework [2], [10, Sec. 2].

A. Prediction Step, Time Update

The system model is given by

$$\underline{x}_{k+1} = \underline{a}_k(\underline{x}_k, \underline{w}_k) \ ,$$

where $\underline{w}_k \sim \mathcal{N}(\underline{0}, \mathbf{C}_k^w)$ is non-additive zero-mean Gaussian process noise. To handle the non-additive noise, we sample from joint density of state and noise

$$[\underline{x}_k, \underline{w}_k]^T \sim \mathcal{N}([\hat{\underline{x}}_k^e, \underline{0}]^T, \mathrm{diag}(\mathbf{C}_k^e, \mathbf{C}_k^w)) \ .$$

Based on the samples $[\underline{x}_{k,1}, \underline{w}_{k,1}]^T, \ldots, [\underline{x}_{k,N}, \underline{w}_{k,N}]^T$ with weights $\gamma_1, \ldots, \gamma_N$, we can obtain mean and covariance of the predicted density according to

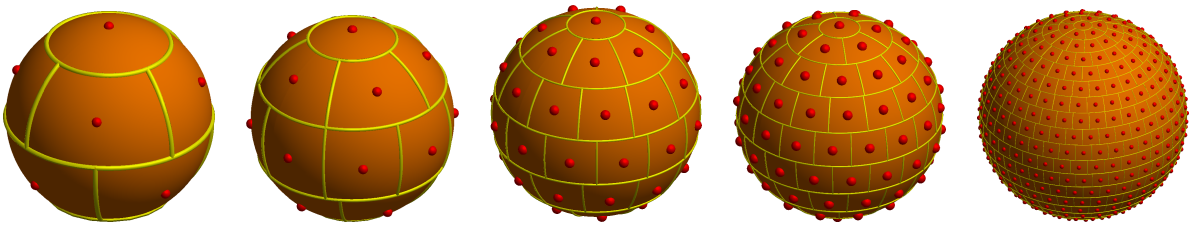
$$\begin{aligned} \hat{\underline{x}}_{k+1}^p &\approx \textstyle\sum_{i=1}^{N} \gamma_i \underline{a}_k(\underline{x}_{k,i}, \underline{w}_{k,i}) \ , \\ \mathbf{C}_{k+1}^p &\approx \textstyle\sum_{i=1}^{N} \gamma_i (\underline{a}_k(\underline{x}_{k,i}, \underline{w}_{k,i}) - \hat{\underline{x}}_{k+1}^p) \\ &\quad \cdot (\underline{a}_k(\underline{x}_{k,i}, \underline{w}_{k,i}) - \hat{\underline{x}}_{k+1}^p)^T \ . \end{aligned}$$

For additive noise, we only need to sample from the state density and the equations can be simplified [28, Sec. 2.4.4].

B. Correction Step, Measurement Update

We assume a measurement model

$$\underline{y}_k = \underline{h}_k(\underline{x}_k, \underline{v}_k) \ ,$$



$N = 10$ $N = 20$ $N = 50$ $N = 100$ $N = 500$

Fig. 4: Equal area partitioning on the sphere S^2.

where $\underline{v}_k \sim \mathcal{N}(\underline{0}, \mathbf{C}_k^v)$ is non-additive zero-mean Gaussian measurement noise. Once again, we sample from the joint density of state and noise

$$[\underline{x}_k, \underline{v}_k]^T \sim \mathcal{N}([\hat{\underline{x}}_k^p, \underline{0}]^T, \mathrm{diag}(\mathbf{C}_k^p, \mathbf{C}_k^v)) \ .$$

Then, we compute the expected measurement

$$\hat{\underline{y}}_k \approx \sum_{i=1}^N \gamma_i \underline{h}_k(\underline{x}_{k,i}, \underline{v}_{k,i})$$

and its covariance

$$\mathbf{C}_k^y \approx \sum_{i=1}^N \gamma_i (\underline{h}_k(\underline{x}_{k,i}, \underline{v}_{k,i}) - \hat{\underline{y}}_k) \cdot (\underline{h}_k(\underline{x}_{k,i}, \underline{v}_{k,i}) - \hat{\underline{y}}_k)^T$$

as well as the cross-covariance of state and measurement

$$\mathbf{C}_k^{x,y} \approx \sum_{i=1}^N \gamma_i (\underline{x}_{k,i} - \hat{\underline{x}}_k^p) \cdot (\underline{h}_k(\underline{x}_{k,i}, \underline{v}_{k,i}) - \hat{\underline{y}}_k)^T \ .$$

Using the actual measurement $\tilde{\underline{y}}_k$, we perform the Kalman update to obtain the mean and covariance of the estimated state according to

$$\hat{\underline{x}}_k^e = \hat{\underline{x}}_k^p + \mathbf{C}_k^{x,y} \cdot (\mathbf{C}_k^y)^{-1} \cdot (\tilde{\underline{y}}_k - \hat{\underline{y}}_k) \ ,$$
$$\mathbf{C}_k^e = \mathbf{C}_k^p - \mathbf{C}_k^{x,y} \cdot (\mathbf{C}_k^y)^{-1} \cdot (\mathbf{C}_k^{x,y})^T \ .$$

If the noise is additive, it is sufficient to sample from the state density and we can simplify the equations [28, Sec. 2.4.4].

VII. Evaluation

We evaluate the proposed approach in comparison with LRKFs found in literature. For this purpose, we rely on the implementations available in the Nonlinear Estimation Toolbox [11] for MATLAB.

A. Sample Analysis

Before we evaluate the novel approach in a filtering scenario, we take a close look at the sample placement and perform a comparison to other sampling schemes. Examples of samples produced by different methods for a two-dimensional scenario can be seen in Fig. 5. The S^{2}KF produces the most homogeneous samples in this case.

To visualize the behavior of the sampling schemes in higher dimensions, we consider the distance of each sample to the origin. Although this does not capture all aspects of the sampling schemes, it helps to understand the radial component. The results for the proposed filter, the S^{2}KF and the RUKF are shown in Fig. 6 and Fig. 7. We do not depict the UKF and the CKF as their samples are exactly located on a hypersphere (except for a single sample at the origin). It can be seen that the samples of the proposed filter are not exactly located on hyperspheres due to the covariance correction (see Sec. V). In the 5D case, the S^{2}KF also creates a structure resembling multiple hyperspheres in some cases, whereas the RUKF is quite random. While the different behaviors of the considered sampling schemes are highly interesting, it is not obvious which choice of samples yields the best results in a filtering application, so we will look at this aspect in Sec. VII-B.

Furthermore, we investigate the expected squared distance of a Gaussian-distributed random vector from the closest sample. Intuitively, the closer this vector is located to a sample point, the better the nonlinear mapping will be approximated. For $\underline{x} \sim \mathcal{N}(\underline{0}, \mathbf{I})$, we consider

$$\mathbb{E}\left(\min_{1 \leq i \leq N} \|\underline{x} - \underline{r}_i\|_2^2\right) \ .$$

This expectation value can be computed for a given sample set using numerical integration. It corresponds to the squared distortion measure used in [20, Sec. 2.1]. The results for the 2D case are shown in Fig. 8. It can be seen that the S^{2}KF performs very well because it spreads the samples very evenly. The proposed approach can achieve similar results for a sufficient number of layers L. It can also be observed that increasing the number of layers only pays off when the total number of samples N is sufficiently high.

B. Filter Comparison

We evaluate the novel filter in scenarios with different numbers of dimensions. As the measurement update is typically the more challenging part, we assume a linear random walk system model

$$\underline{x}_{k+1} = \underline{x}_k + \underline{w}_k \ ,$$

where $\underline{w}_k \sim \mathcal{N}(\underline{0}, 10^{-2} \cdot \mathbf{I})$ is additive zero-mean system noise. The measurement model is given by

$$\underline{h}(\underline{x}_k) = \underline{x}_k + \begin{bmatrix} \sin(c \cdot x_{k,2}) \\ \vdots \\ \sin(c \cdot x_{k,d}) \\ \sin(c \cdot x_{k,1}) \end{bmatrix} + \underline{v}_k \ ,$$

where $\underline{v}_k \sim \mathcal{N}(\underline{0}, 10^{-4} \cdot \mathbf{I})$ is zero-mean measurement noise. We choose the constant parameter $c = 10$. This model is interesting because the behavior strongly depends on the uncertainty at the current time step. Also, it can be used in an arbitrary number of dimensions. The true initial state is given by $\underline{x}_0 = \underline{0}$ and the initial estimate is given by $\underline{x}_0^e \sim \mathcal{N}(\underline{0}, 10^{-2} \cdot \mathbf{I})$.

In our evaluation, we compare the proposed approach to the EKF, the UKF, the 5th degree CKF, the S^{2}KF, the randomized UKF, and the GHKF (with $m = 3$ points in each dimension). For the proposed approach, the S^{2}KF, and the RUKF, we used $N = 2d^2 + 1$ samples. Furthermore, we use $L = 5$

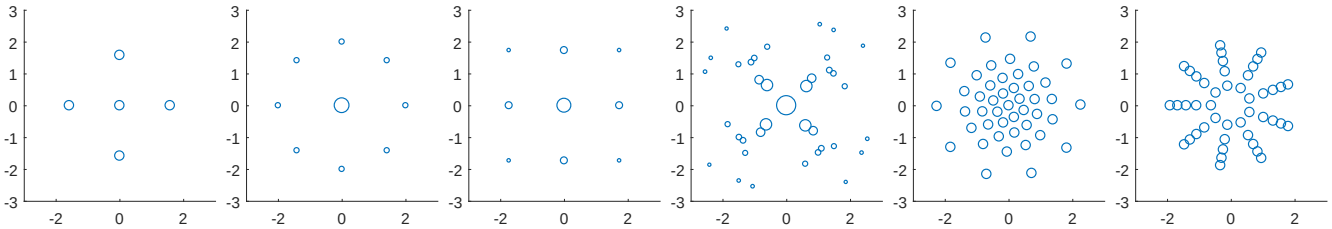


Fig. 5: Comparison of sampling schemes in 2D. The weights are illustrated using the size of the samples and negative weights are shown in red.

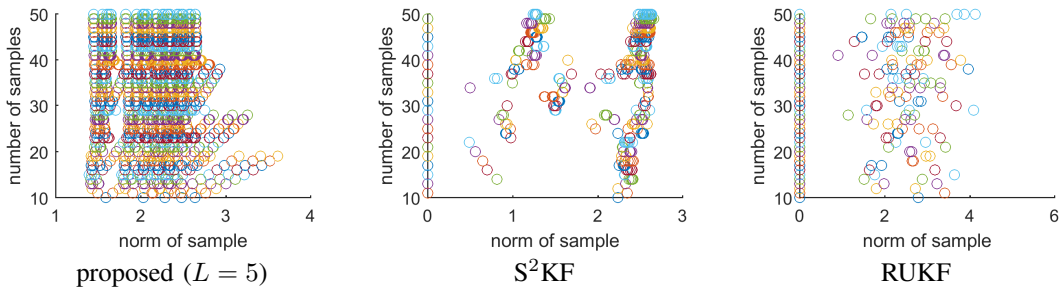


Fig. 6: Norm of all samples in 5D for different number of samples N.

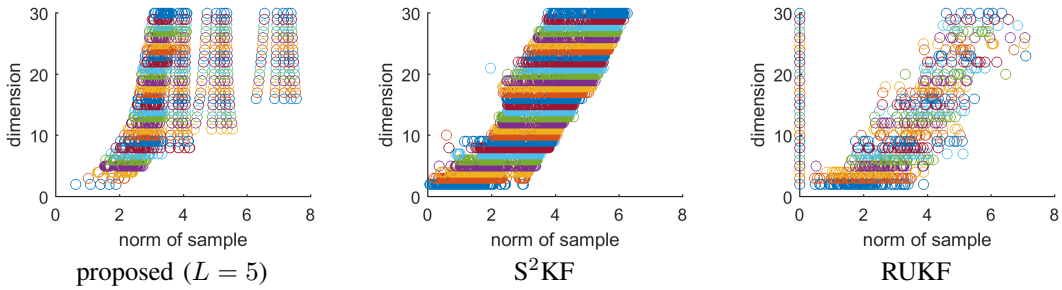


Fig. 7: Norm of all samples for different number of dimensions and $N = 500$.

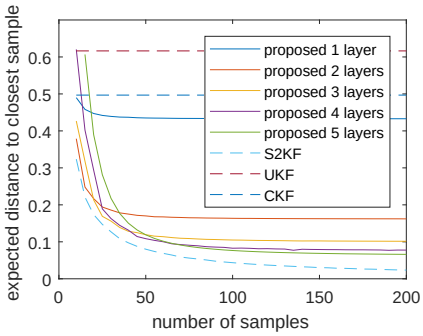


Fig. 8: Expected distance to the closest sample in 2D. Note that the number of samples for the UKF and CKF are fixed.

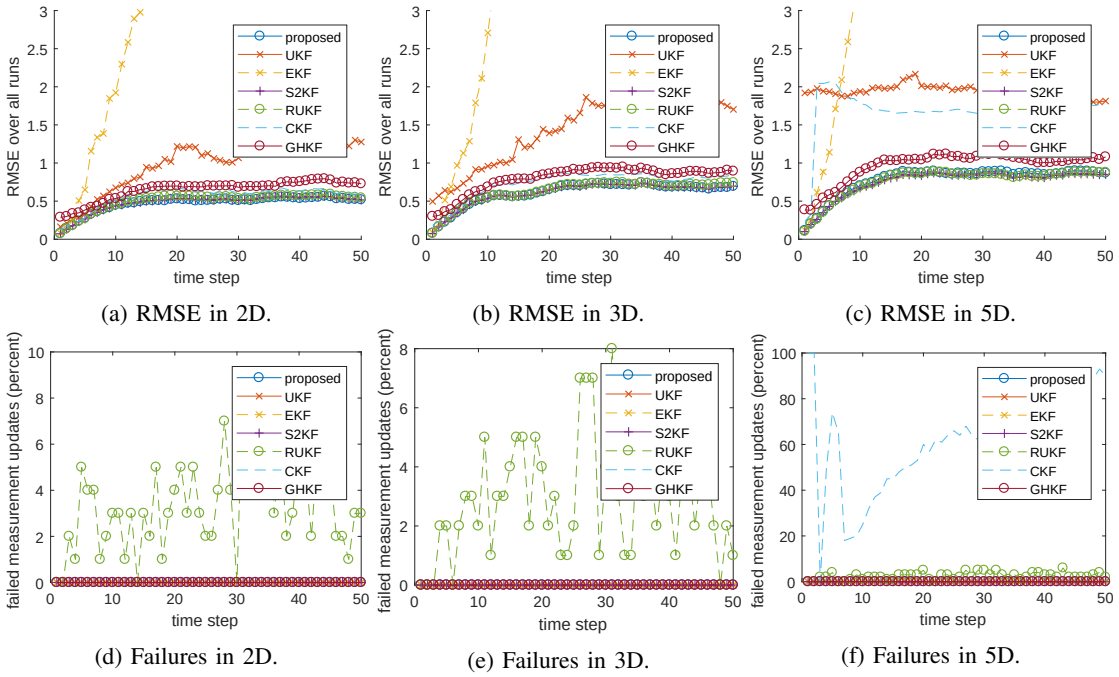
layers in our method, which has empirically been found to work quite well as long as N is large enough.

We performed the evaluation in 2D, 3D, and 5D for 50 time steps each. The results of 100 Monte Carlo runs are shown in Fig. 9. We use the RMSE averaged across all runs as the evaluation criterion. The novel method achieves a similar estimation accuracy as the S^{2}KF and RUKF in the considered scenario, while being deterministic and avoiding slow precomputation. These three methods clearly outperform the EKF and UKF. The CKF works quite well in 2D and 3D, but is a lot worse in 5D because the negative weights causes frequent failures of the update step. The RUKF also exhibits occasional failures, but in the considered scenarios, the effect on its performance seems limited.

VIII. Conclusion

We have presented a novel sampling scheme based on a combination of Chi distribution sampling and uniform hyperspherical sampling. This approach has significant advantages compared to state-of-the-art methods, namely, no negative weights, a flexible number of components, no randomness, and no need for costly numerical optimization.

Based on this sampling scheme, we have proposed a new filter based on the LRKF-principle. We have evaluated the novel filter in multiple simulations and shown that its performance is clearly superior to the EKF, UKF, CKF, and GHKF. It is comparable to the RUFK and the S^{2}KF, while avoiding their disadvantages, namely the nondeterminism and negative weights of the RUKF and the expensive precomputation of the S^{2}KF.



(a) RMSE in 2D. (b) RMSE in 3D. (c) RMSE in 5D.

(d) Failures in 2D. (e) Failures in 3D. (f) Failures in 5D.

Fig. 9: Evaluation results.

ACKNOWLEDGMENT

The authors thank Jannik Steinbring for his support. This work was supported by the German Research Foundation (DFG) under grant HA 3789/13-1.

REFERENCES

[1] G. L. Smith, S. F. Schmidt, and L. A. McGee, *Application of Statistical Filter Theory to the Optimal Estimation of Position and Velocity on Board a Circumlunar Vehicle*. National Aeronautics and Space Administration, 1962.

[2] T. Lefebvre, H. Bruyninckx, and J. D. Schuller, "Comment on "A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators" [with Authors' Reply]," *IEEE Transactions on Automatic Control*, vol. 47, no. 8, pp. 1406–1409, Aug. 2002.

[3] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME Journal of Basic Engineering*, vol. 82, pp. 35–45, 1960.

[4] S. J. Julier and J. K. Uhlmann, "Unscented Filtering and Nonlinear Estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.

[5] B. Jia, M. Xin, and Y. Cheng, "High-degree Cubature Kalman Filter," *Automatica*, vol. 49, no. 2, pp. 510–518, 2013.

[6] I. Arasaratnam and S. Haykin, "Cubature Kalman Filters," *IEEE Transactions on Automatic Control*, vol. 54, no. 6, pp. 1254–1269, 2009.

[7] M. F. Huber and U. D. Hanebeck, "Gaussian Filter based on Deterministic Sampling for High Quality Nonlinear Estimation," in *Proceedings of the 17th IFAC World Congress (IFAC 2008)*, vol. 17, no. 2, Seoul, Republic of Korea, Jul. 2008.

[8] O. Straka, J. Dunik, and M. Simandl, "Randomized Unscented Kalman Filter in Target Tracking," in *15th International Conference on Information Fusion (Fusion 2012)*, 2012, pp. 503–510.

[9] K. Ito and K. Xiong, "Gaussian Filters for Nonlinear Filtering Problems," *IEEE Transactions on Automatic Control*, vol. 45, no. 5, pp. 910–927, May 2000.

[10] J. Steinbring, M. Pander, and U. D. Hanebeck, "The Smart Sampling Kalman Filter with Symmetric Samples," *Journal of Advances in Information Fusion*, vol. 11, no. 1, pp. 71–90, Jun. 2016.

[11] J. Steinbring, "Nonlinear Estimation Toolbox," 2015. [Online]. Available: https://bitbucket.org/nonlinearestimation/toolbox

[13] U. D. Hanebeck, "Sample Set Design for Nonlinear Kalman Filters viewed as a Moment Problem," in *Proceedings of the 17th International Conference on Information Fusion (Fusion 2014)*, Salamanca, Spain, Jul. 2014.

[12] Z. Wang and Y. Li, "Cross Approximation-based Quadrature Filter," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, Las Vegas, NV, USA, 2016.

[14] W. Härdle and L. Simar, *Applied Multivariate Statistical Analysis*, 2nd ed. Berlin, Germany: Springer, 2007.

[15] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration*, 2nd ed., ser. Compute Science and Applied Mathematics. San Diego, California, USA: Academic Press, Inc., 1984.

[16] E. B. Saff and A. B. Kuijlaars, "Distributing Many Points on a Sphere," *The Mathematical Intelligencer*, vol. 19, no. 1, pp. 5–11, 1997.

[17] J. J. Thomson, "On the Structure of the Atom: An Investigation of the Stability and Periods of Oscillation of a Number of Corpuscles Arranged at Equal Intervals around the Circumference of a Circle; with Application of the Results to the Theory of Atomic Structure," *Philosophical Magazine Series 6*, vol. 7, no. 39, pp. 237–265, 1904.

[18] T. Aste and D. Weaire, *The Pursuit of Perfect Packing*, 2nd ed. Taylor & Francis, 2008.

[19] Q. Du, V. Faber, and M. Gunzburger, "Centroidal Voronoi Tessellations: Applications and Algorithms," *SIAM review*, vol. 41, no. 4, pp. 637–676, 1999.

[20] G. Pagès and J. Printems, "Optimal Quadratic Quantization for Numerics: The Gaussian Case," *Monte Carlo Methods and Applications*, vol. 9, no. 2, pp. 135–165, 2003.

[21] A. Katanforoush and M. Shahshahani, "Distributing Points on the Sphere, I," *Experimental Mathematics*, vol. 12, no. 2, pp. 199–209, 2003.

[22] H. Peng and Y. Yu, "Project Report for CS59000 OPT: Optimization on the Surface of the (Hyper)-Sphere," Tech. Rep., May 2012.

[23] F.-F. Henrich and K. Obermayer, "Active Learning by Spherical Subdivision," *Journal of Machine Learning Research*, vol. 9, no. Jan, pp. 105–130, 2008.

[24] P. Leopardi, "A Partition of the Unit Sphere into Regions of Equal Area and Small Diameter," *Electronic Transactions on Numerical Analysis*, vol. 25, no. 12, pp. 309–327, 2006.

[25] ——, "Distributing Points on the Sphere: Partitions, Separation, Quadrature and Energy," Ph.D. dissertation, University of New South Wales, Apr. 2007.

[26] ——, "Recursive Zonal Equal Area Sphere Partitioning Toolbox 1.10," http://eqsp.sourceforge.net/, Jun. 2005.

[27] O. C. Schrempf, D. Brunn, and U. D. Hanebeck, "Density Approximation Based on Dirac Mixtures with Regard to Nonlinear Estimation and Filtering," in *Proceedings of the 2006 IEEE Conference on Decision and Control (CDC 2006)*, San Diego, California, USA, Dec. 2006.

[28] I. Gilitschenski, "Deterministic Sampling for Nonlinear Dynamic State Estimation," Ph.D. dissertation, Karlsruhe Institute of Technology, Intelligent Sensor-Actuator-Systems Laboratory, 2015.