Stochastic Sampling of the Hyperspherical von Mises–Fisher Distribution Without Rejection Methods

Gerhard Kurz and **Uwe D. Hanebeck**
Intelligent Sensor-Actuator-Systems Laboratory (ISAS)
Institute for Anthropomatics and Robotics
Karlsruhe Institute of Technology (KIT), Germany
gerhard.kurz@kit.edu, uwe.hanebeck@ieee.org

***Abstract*—We propose a novel sampling algorithm for the von Mises–Fisher distribution on the unit hypersphere. Unlike previous works, we show a solution for an arbitrary number of dimensions without requiring rejection sampling. As a result, the proposed algorithm has a deterministic runtime. The key idea consists in applying the inversion method to a one-dimensional subproblem and analytically calculating the integral occurring in the distribution function. The proposed method is most efficient for odd numbers of dimensions. We compare the algorithm to a state-of-the-art rejection sampling method in simulations.**

I. Introduction

Hyperspherical probability distributions are distributions defined on the unit hypersphere $S^{d-1} = \{\underline{x} \in \mathbb{R}^d : \|\underline{x}\| = 1\}$. They are of importance in a variety of applications related to estimation and sensor data fusion, which we will discuss below. One of the most popular hyperspherical distributions is the von Mises–Fisher distribution [1], sometimes also referred to as the Langevin distribution.

In recent years, a number of papers has been published where the von Mises–Fisher distribution is applied in a variety of contexts. Chiuso and Picci used the distribution in the context of camera-based tracking on the unit sphere [2]. Later, Markovic et al. considered an application involving an omnidirectional camera [3], [4]. Traa and Smaragdis proposed the use of the von Mises–Fisher distribution for multiple speaker tracking [5] based on microphone arrays. Other applications include high angular resolution diffusion MRI [6] as well as electron backscatter microscopy in the context of estimation of crystal orientations for crystallography [7], [8].

The probability density of a von Mises–Fisher distribution on the unit hypersphere S^{d-1} is given by

$$f(\underline{x}; \underline{\mu}, \kappa) = c_d(\kappa) \cdot \exp(\kappa \cdot \underline{\mu}^T \underline{x}) \ , \qquad (1)$$

where we have $\underline{x} \in S^{d-1}$, the location parameter $\underline{\mu} \in S^{d-1}$, and the concentration parameter $\kappa \geq 0$. The normalization constant can be obtained according to

$$c_d(\kappa) = \frac{\kappa^{d/2-1}}{(2\pi)^{d/2} I_{d/2-1}(\kappa)} \ ,$$

where $I_v(\kappa)$ is the modified Bessel function of the first kind.

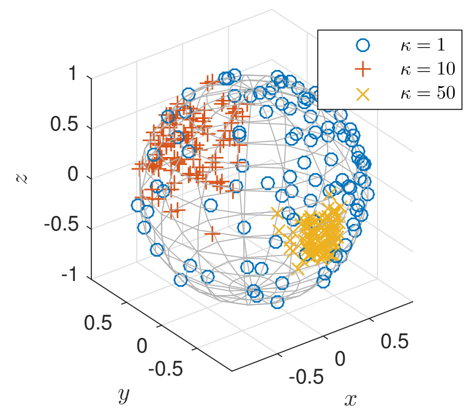

Figure 1. Samples from different von Mises–Fisher distributions on S^2.

For many applications it is important to obtain random samples from the von Mises–Fisher distribution, for example in the context of Monte Carlo methods and for evaluation or visualization purposes. An example of samples obtained from three different von Mises–Fisher distributions on S^2 ($d = 3$) is given in Fig. 1. The problem of obtaining random samples from a given probability distribution has been investigated for a long time [9], but common books on this topic typically do not include the von Mises–Fisher distribution [10].

There has been some discussion of sampling algorithms for the von Mises–Fisher distribution in literature. Ulrich [11] proposed a fairly general framework for sampling hyperspherical distributions using rejection sampling. A slightly modified version of this method that fixes certain issues was proposed by Wood [12] for the von Mises–Fisher as well as the Bingham distribution. Both of these methods sample from a related proposal density instead of the true density and accept the resulting samples only with a certain probability. This leads to a non-deterministic runtime because it is not known a priori how many samples are drawn before one is accepted. Even though the average runtime may be quite small and it might be possible to prove a stochastic bound, the worst-case runtime cannot be bounded. Also, the variable number of iterations required makes vectorization or parallelization of the algorithm more difficult.

Some authors have also proposed methods for the special case of sampling the von Mises–Fisher distribution on S^2 ($d = 3$), for example Fisher et al. [13, p. 59], Jung [14], and Jakob [15]. In this case, there exist simple solutions and no rejection sampling is necessary. However, these approaches cannot be generalized to a higher number of dimensions.

The case of S^1 ($d = 2$) has also been investigated. There, the von Mises–Fisher distribution coincides with the von Mises distribution. Algorithms for sampling the von Mises distribution based on rejection sampling have been published in [10, Sec. 7.3] and [16].

The contribution of this paper can be summarized as follows. We propose a novel method for stochastic sampling of the von Mises–Fisher distribution on the unit hypersphere S^{d-1}. For an odd number of dimensions d, the proposed method is analytical except for a bisection algorithm, which is, however, guaranteed to converge up to machine precision with a reasonable predefined number of steps. For an even number of dimensions d, the same method can be applied, but one of the intermediate results is given by a series representation, which cannot be evaluated analytically. Therefore, the proposed method is of particular interest for an odd number dimensions.

II. Proposed Algorithm

Without loss of generality, we only consider the case of $\underline{\mu} = [1, 0, \dots, 0]^T$ at first. Once samples for a von Mises–Fisher distribution with this $\underline{\mu}$ have been obtained, they can be rotated appropriately to a different location.

Based on the results of Wood [12, eq. (1)] and Ulrich [11], we can obtain samples for this distribution according to

$$\underline{s} = [w, \sqrt{(1-w^2)}\underline{v}^T]^T \ ,$$

where $\underline{v}$ is sampled uniformly[1] on S^{d-2} and $w \in [-1, 1]$ is distributed according to the probability density function

$$f^w(w) = c \cdot (1-w^2)^{(d-3)/2} \exp(\kappa w) \ , \tag{2}$$

where

$$c^{-1} = \int_{-1}^{1} (1-w^2)^{(d-3)/2} \exp(\kappa w) dw \ .$$

Just as in (1), κ refers to the concentration parameter of the von Mises–Fisher distribution under consideration. This density can be derived by using a substitution into spherical coordinates and only considering the angle between $\underline{x}$ and $\underline{\mu}$, similar to [17, Appendix C]. A plot of (2) is depicted in Fig. 2. As a result, the problem of sampling the von Mises–Fisher distribution can be reduced to the problem of sampling w from $f^w(w)$.

[1] Uniform sampling on a sphere can be easily achieved by sampling from a multivariate Gaussian with mean zero and identity covariance, and subsequent normalization to unit length.

A. Solution for $d = 3$

In order to introduce the basic idea, we first look at an easy special case that has previously been solved. For $d = 3$, $f^w(w)$ simplifies to $f^w(w) = c \cdot \exp(\kappa w)$. Thus, we obtain the distribution function in closed form according to

$$F^w(w) = \int_{-1}^{w} f^w(x) dx = \frac{c}{\kappa}(\exp(\kappa w) - \exp(-\kappa)) \ .$$

The normalization constant is given by

$$c^{-1} = \int_{-1}^{1} \exp(\kappa w) dw = \frac{1}{\kappa}(\exp(\kappa) - \exp(-\kappa)) = \frac{2}{\kappa}\sinh(\kappa) \ .$$

Then, we can apply the inversion method [9, Sec. 4.1]. For this purpose, we sample a random variable u uniformly on $[0, 1]$ and obtain

$$w = (F^w)^{-1}(u) = \frac{1}{\kappa} \log\left(2u \sinh(\kappa) + \exp(-\kappa)\right) \ .$$

This is equal to the result given by Jung [14]. According to [15, Sec. 3.1], this result can be reformulated as

$$w = 1 + \frac{1}{\kappa} \log\left(u + (1-u)\exp(-2\kappa)\right)$$

to increase numerical stability for large values of κ.

B. Generalization to an Arbitrary Number of Dimensions

To generalize this method to an arbitrary number of dimensions, we calculate the distribution function using the substitution $x = \cos(\phi)$, which yields

$$\begin{aligned} F^w(w) &= \int_{-1}^{w} f^w(x) dx \\ &= c \cdot \int_{-1}^{w} (1-x^2)^{(d-3)/2} \exp(\kappa x) dx \\ &= c \cdot \int_{0}^{\arccos(w)+\pi/2} \sin(\phi)^{d-2} \exp(\kappa \cos(\phi)) d\phi \ . \end{aligned}$$

The normalization constant c is thus given by

$$c^{-1} = \int_{0}^{\pi} \sin(\phi)^{d-2} \exp(\kappa \cos(\phi)) d\phi \ .$$

1) Odd Number of Dimensions: If the number of dimensions d is odd, we can obtain a closed-form solution as follows. For odd $d \geq 3$, we define $n = (d-3)/2$, and the indefinite integral (antiderivative) is given by

$$\begin{aligned} &\int \sin(\phi)^{d-2} \exp(\kappa \cos(\phi)) d\phi \qquad (3) \\ &= \int (1-\cos(\phi)^2)^n \sin(\phi) \exp(\kappa \cos(\phi)) d\phi \\ &= \int \left(\sum_{j=0}^{n} \binom{n}{j} (-\cos(\phi)^2)^j \right) \sin(\phi) \exp(\kappa \cos(\phi)) d\phi \\ &= \sum_{j=0}^{n} \binom{n}{j} (-1)^j \int \cos(\phi)^{2j} \sin(\phi) \exp(\kappa \cos(\phi)) d\phi \ . \end{aligned}$$

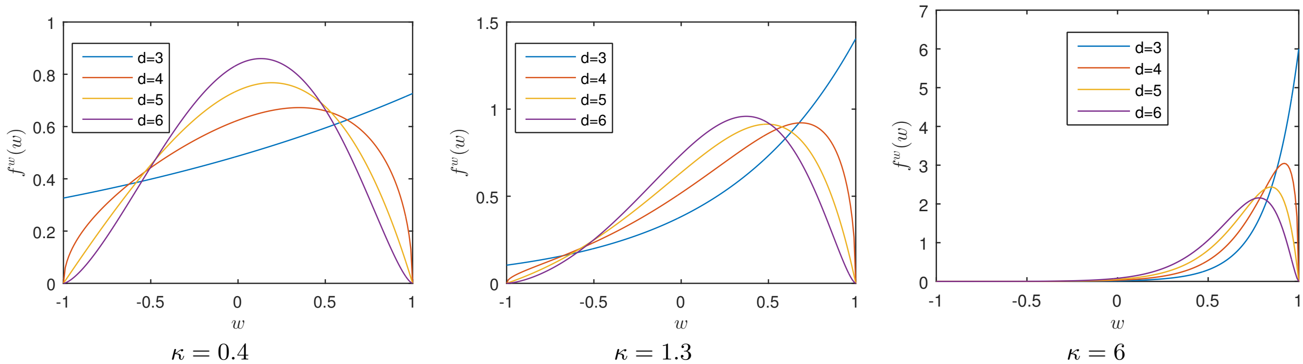

Figure 2. The probability density function $f^w(w)$ for different dimensions d and different values of the concentration parameter κ.

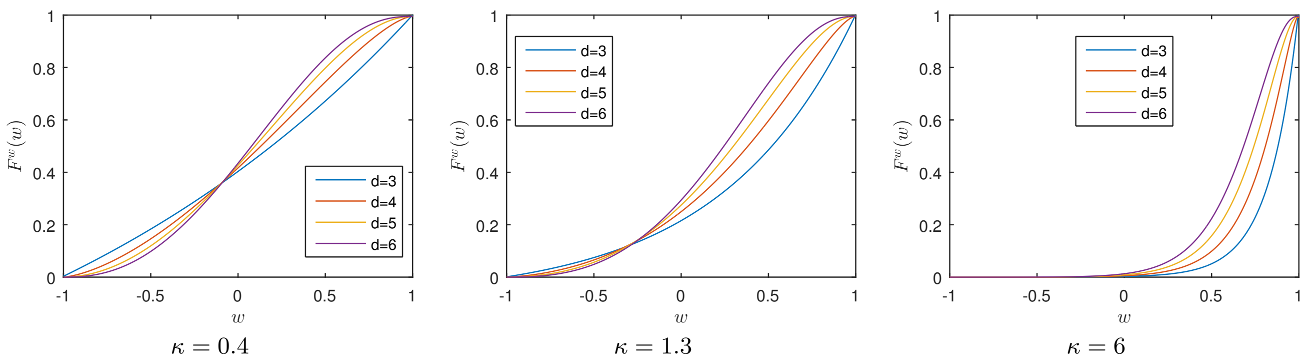

Figure 3. The distribution function $F^w(w)$ for different dimensions d and different values of the concentration parameter κ.

Furthermore, we can calculate the remaining integral using the following identities

$$\begin{aligned}
&\int \cos(\phi)^{2j} \sin(\phi) \exp(\kappa \cos(\phi)) d\phi \\
&= \int \sin(\phi) \frac{\partial^{2j}}{(\partial\kappa)^{2j}} \exp(\kappa \cos(\phi)) d\phi \\
&= \frac{\partial^{2j}}{(\partial\kappa)^{2j}} \int \sin(\phi) \exp(\kappa \cos(\phi)) d\phi \\
&= -\frac{\partial^{2j}}{(\partial\kappa)^{2j}} \frac{\exp(\kappa \cos(\phi))}{\kappa} + \text{const} .
\end{aligned}$$

The derivatives in the last step can easily be obtained for arbitrary $j \geq 0$, but the terms for large values of j can get quite complicated.

As a result, the distribution function $F^w(w)$ can be calculated in closed form for an arbitrary odd number of dimensions $d \geq 3$. Solutions of (3) for a few values of d are given in the Appendix.

2) Even Number of Dimensions: An even number of dimensions is more tricky and leads to a series representation rather than a closed-form solution. For an even number of dimensions $d \geq 2$, we can obtain the indefinite integral according to

$$\begin{aligned}
&\int \sin(\phi)^{d-2} \exp(\kappa \cos(\phi)) d\phi \\
&= \int (1 - \cos(\phi)^2)^n \exp(\kappa \cos(\phi)) d\phi \\
&= \int \left(\sum_{j=0}^{n} \binom{n}{j} (-\cos(\phi)^2)^j \right) \exp(\kappa \cos(\phi)) d\phi \\
&= \sum_{j=0}^{n} \binom{n}{j} (-1)^j \int \cos(\phi)^{2j} \exp(\kappa \cos(\phi)) d\phi ,
\end{aligned}$$

where $n = d/2 - 1$. Now, we can rewrite the remaining integral using the following identities

$$\begin{aligned}
&\int \cos(\phi)^{2j} \exp(\kappa \cos(\phi)) d\phi \\
&= \int \frac{\partial^{2j}}{(\partial\kappa)^{2j}} \exp(\kappa \cos(\phi)) d\phi \\
&= \frac{\partial^{2j}}{(\partial\kappa)^{2j}} \int \exp(\kappa \cos(\phi)) d\phi \\
&= -\frac{\partial^{2j}}{(\partial\kappa)^{2j}} \int \sum_{m=-\infty}^{\infty} I_m(\kappa) \cos(m\phi) d\phi \\
&= -\sum_{m=-\infty}^{\infty} \frac{\partial^{2j}}{(\partial\kappa)^{2j}} I_m(\kappa) \int \cos(m\phi) d\phi \\
&= -\sum_{m=-\infty}^{\infty} \frac{\sin(m\phi)}{m} \frac{\partial^{2j}}{(\partial\kappa)^{2j}} I_m(\kappa) + \text{const} \\
&= -\sum_{m=-\infty}^{\infty} \frac{\sin(m\phi)}{m} \frac{1}{2^{2j}} \sum_{n=0}^{2j} \binom{2j}{n} I_{m-2j+2n}(\kappa) + \text{const}
\end{aligned}$$

$$= -\frac{1}{2^{2j}} \sum_{n=0}^{2j} \binom{2j}{n} \sum_{m=-\infty}^{\infty} \frac{\sin(m\phi)}{m} I_{m-2j+2n}(\kappa) + \text{const} ,$$

where we use [18, eq. (9.6.34)] and [18, eq. (9.6.29)]. The infinite series contained in the result converges fairly quickly, so that considering just a few summands is usually sufficient.

3) Inversion of the Distribution Function: In order to sample from f^w, we can now sample u uniformly on $[0, 1]$ and calculate $w = (F^w)^{-1}(u)$. Even though $F^w(w)$ can be obtained in analytical form, calculation of its inverse is difficult for $d > 3$. For this reason, we propose the use of a numerical method to obtain w as the root of $F^w(w) - u$. It is easy to see that $F^w(w)$ is strictly increasing on $[-1, 1]$ and that there is always a unique solution. Thus, the bisection method [19, Sec. 9.1] is guaranteed to yield the result up to numerical precision in all cases after a reasonable number of steps[2].

In order to achieve even faster convergence, it is also possible to apply the one-dimensional Newton–Raphson method [19, Sec. 9.4], which converges quadratically under certain conditions. We use the initial value $w_0 = 0$ and the iteration

$$w_{k+1} = w_k - \frac{F^w(w_k) - u}{f^w(w_k)} .$$

Note that it is trivial in this case to calculate the derivative of $F^w(w) - u$, which is given by $f^w(w)$. However, convergence is not guaranteed, as $F^w(w) - u$ is not necessarily convex (see Fig. 3). We achieved convergence in most cases in our experiments, usually after just a few steps. To ensure guaranteed convergence a more sophisticated method may need to be used, e.g., Ridders' method [19, Sec. 9.2.1], Brent's method [19, Sec. 9.3], or a hybrid method combining the Newton–Raphson method and the bisection method.

C. Generalization to Arbitrary $\underline{\mu}$

As mentioned before, the proposed algorithm can easily be generalized to von Mises–Fisher distributions with arbitrary $\underline{\mu} \in S^{d-1}$. This is achieved by subsequently rotating all samples with a suitable rotation matrix $\mathbf{Q}(\underline{\mu})$. Because von Mises–Fisher distributions are radially symmetric around the axis of $\underline{\mu}$, an arbitrary rotation matrix satisfying the equality $\underline{\mu} = \mathbf{Q}(\underline{\mu}) \cdot [1, 0, \dots, 0]^T$ can be chosen. Thus, the first column of $\mathbf{Q}(\underline{\mu})$ is given by $\underline{\mu}$ and the other columns need to be chosen in such a way that $\mathbf{Q}(\underline{\mu})$ is an orthogonal matrix with determinant 1. This can be achieved by initializing $\mathbf{Q}(\underline{\mu})$ using a matrix with $\underline{\mu}$ in the first column, and then performing a Gram–Schmidt orthogonalization procedure or a different method of computing the QR decomposition [21, Sec. 5.2].

Pseudo code for the entire procedure is given in Algorithm 1.

Algorithm 1: von Mises–Fisher Sampling

Input: parameters $\underline{\mu}$, κ, number of iterations n with default $n = 50$
Output: sample $\underline{s}$
$d \leftarrow$dimension($\underline{\mu}$);
$u \leftarrow$uniform($[0, 1]$);
/* Bisection algorithm to sample w */
$w_{\text{min}} \leftarrow -1$;
$w_{\text{max}} \leftarrow 1$;
for $j \leftarrow 1$ **to** n **do**
  $w \leftarrow \frac{w_{\text{min}} + w_{\text{max}}}{2}$;
  if $F^w(w_k) > u(i)$ **then**
    $w_{\text{max}} \leftarrow w$;
  else
    $w_{\text{min}} \leftarrow w$;
  end
end
/* Generate sample on unit sphere */
$\underline{v} \leftarrow$uniform(S^{d-2});
$\underline{s} \leftarrow [w, \sqrt{1 - w^2} \cdot \underline{v}^T]^T$;
/* Rotate depending on $\underline{\mu}$ */
$\mathbf{M} \leftarrow [\underline{\mu}, \underline{0}, \dots \underline{0}]$;
$[\mathbf{Q}, \mathbf{R}] \leftarrow$QrDecomposition($\mathbf{M}$);
if $\mathbf{R}(1, 1) < 0$ **then**
  /* Make sure that the first column of $\mathbf{Q}$ is $\underline{\mu}$ and not $-\underline{\mu}$ */
  $\mathbf{Q} \leftarrow -\mathbf{Q}$;
end
$\underline{s} \leftarrow \mathbf{Q}\underline{s}$;
return $\underline{s}$;

III. Evaluation

In order to show the validity of the proposed method, we performed multiple simulations, where we compare the novel algorithm to the method proposed by Wood [12]. For this purpose, we consider a von Mises–Fisher probability density in d dimensions with $\underline{\mu} = [1, 0, \dots, 0]^T$ and several concentration parameters $\kappa \in \{0.1, 2, 150\}$. We draw different numbers of samples from this distribution using the two sampling methods and fit a new von Mises-Fisher distribution to these samples using the algorithm proposed by Sra [22]. Now, we consider the angular error $\angle(\underline{\mu}, \underline{\mu}^{\text{fitted}})$ between the location parameter of the original and the fitted distribution as well as the absolute error $|\kappa - \kappa^{\text{fitted}}|$ in the concentration parameter. For a sound sampling method, one would expect both errors to converge to zero for an increasing number of samples.

In Fig. 4–6, we show the results of 1000 Monte Carlo runs for different values of κ and $d \in \{5, 7, 9\}$. It can be seen that the error converges to zero for both approaches and it does so at an almost identical speed. This indicates that the proposed method performs similarly to Wood's method even though it does not require rejection sampling.

We also compared the computation time of Wood's approach and our method. A non-optimized MATLAB implementation of the proposed method is about two to four times slower than Wood's method, depending on the number of dimensions. However, a vectorized version of the proposed

[2] As each step halves the interval in which the solution can be located, we gain one bit of accuracy in each step. Thus, we need approximately 50 steps to reach machine precision for a IEEE 754 double variable [20].

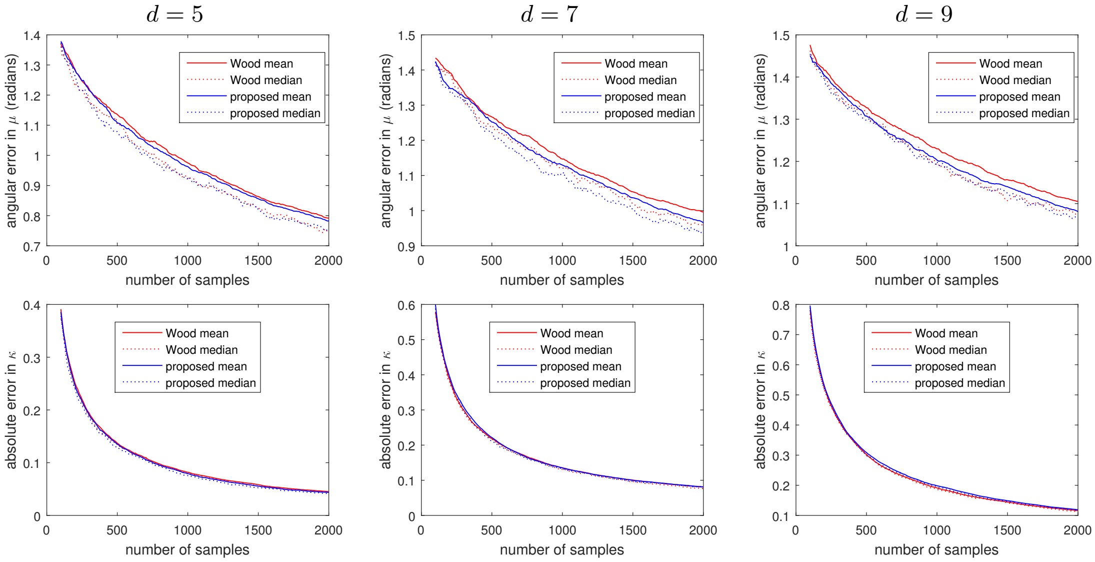

Figure 4. Evaluation results for $\kappa = 0.1$.

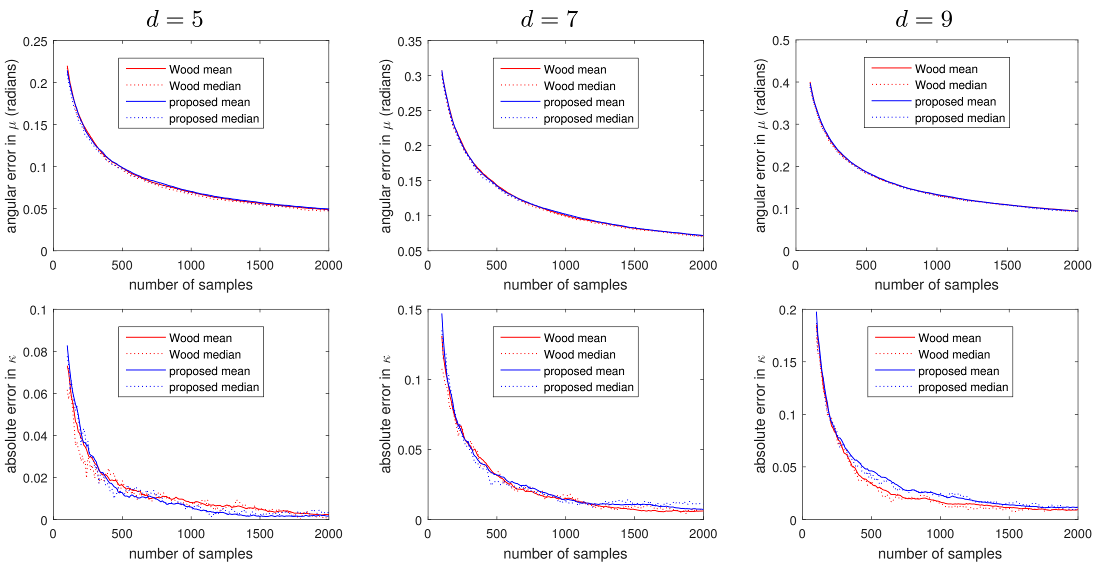

Figure 5. Evaluation results for $\kappa = 2$.

algorithm is faster up to a factor of two than Wood's method if many samples are requested in one function call.

Our implementations of both Wood's method and the proposed method will be made available as part of `libDirectional` [23], a MATLAB library for directional statistics and directional estimation.

IV. Conclusion

We have presented a novel method for sampling from a von Mises–Fisher distribution on the hypersphere without the need to resort to rejection sampling. As a result, the proposed algorithm has a deterministic runtime unlike rejection approaches.

The method can be efficiently implemented for an odd number of dimensions, but can in principle be applied to an even number of dimensions as well. Because samples are drawn from the unit sphere, it is unfortunately not easily possible to reduce sampling in an even dimension to a higher odd dimension, because omitting one of the dimensions of the samples destroys the unit-norm property.

Our evaluation indicates comparable results to the state-of-the-art method by Wood, which necessitates rejection sampling. Thus, the proposed method provides a superior alternative when the computation time needs to be bounded.

Appendix

Here, we give the solution of the indefinite integral (3) for low odd numbers of dimensions:

$d = 3$:

$$-\frac{e^{\kappa \cos(\phi)}}{\kappa}$$

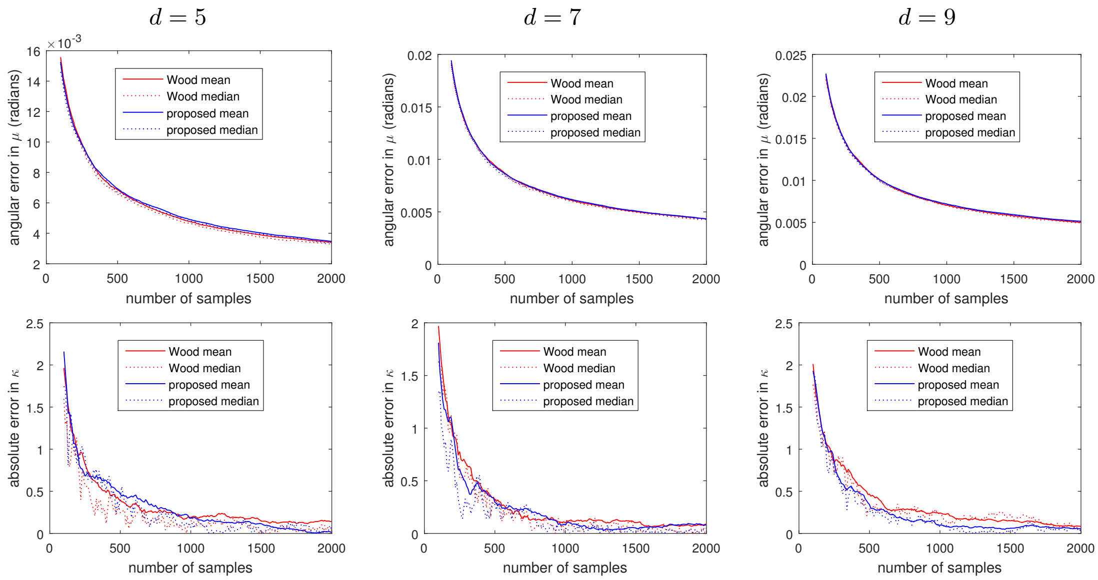

Figure 6. Evaluation results for $\kappa = 150$.

$d = 5$:

$$\frac{e^{\kappa\cos(\phi)}\left(\kappa^2\cos(2\phi) - \kappa^2 - 4\kappa\cos(\phi) + 4\right)}{2\kappa^3}$$

$d = 7$:

$$\begin{aligned}&-\left(e^{\kappa\cos(\phi)}(\kappa^4\cos(4\phi) + 3\kappa^4 - 8\kappa^3\cos(3\phi)\right.\\&-4\left(\kappa^2-12\right)\kappa^2\cos(2\phi) + 8\left(\kappa^2 - 24\right)\kappa\cos(\phi)\\&\left.+16\kappa^2 + 192)\right)/(8\kappa^5)\end{aligned}$$

$d = 9$:

$$\begin{aligned}&\left(e^{\kappa\cos(\phi)}(-6\kappa^6\cos(4\phi) + \kappa^6\cos(6\phi) - 10\kappa^6\right.\\&+36\kappa^5\cos(3\phi) - 12\kappa^5\cos(5\phi) + 120\kappa^4\cos(4\phi)\\&-24\kappa^4 - 960\kappa^3\cos(3\phi) + 3456\kappa^2\\&+3\left(5\kappa^4 - 32\kappa^2 + 1920\right)\kappa^2\cos(2\phi)\\&\left.-24\left(\kappa^4 + 24\kappa^2 + 960\right)\kappa\cos(\phi) + 23040)\right)/(32\kappa^7)\ .\end{aligned}$$

References

[1] R. Fisher, "Dispersion on a Sphere," *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 217, no. 1130, pp. 295–305, 1953.

[2] A. Chiuso and G. Picci, "Visual Tracking of Points as Estimation on the Unit Sphere," in *The Confluence of Vision and Control*. Springer, 1998, pp. 90–105.

[3] I. Markovic, M. Bukal, J. Cesic, and I. Petrovic, "Direction-only Tracking of Moving Objects on the Unit Sphere via Probabilistic Data Association," in *Proceedings of the 17th International Conference on Information Fusion (Fusion 2014)*, Salamanca, Spain, Jul. 2014.

[4] I. Markovic, F. Chaumette, and I. Petrovic, "Moving Object Detection, Tracking and Following Using an Omnidirectional Camera on a Mobile Robot," in *Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA 2014)*, Hong-Kong, Jun. 2014.

[5] J. Traa and P. Smaragdis, "Multiple Speaker Tracking With the Factorial von Mises–Fisher Filter," in *IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, Sep. 2014.

[6] A. Bhalerao and C.-F. Westin, "Hyperspherical von Mises–Fisher Mixture (HvMF) Modelling of High Angular Resolution Diffusion MRI," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2007)*. Springer, 2007, pp. 236–243.

[7] Y.-H. Chen, D. Wei, G. Newstadt, M. DeGraef, J. Simmons, and A. Hero, "Parameter Estimation in Spherical Symmetry Groups," *IEEE Signal Processing Letters*, vol. 22, no. 8, pp. 1152–1155, 2015.

[8] ——, "Statistical Estimation and Clustering of Group-invariant Orientation Parameters," in *Proceedings of the 18th International Conference on Information Fusion (Fusion 2015)*, Washington D. C., USA, Jul. 2015.

[9] B. D. Ripley, "Computer Generation of Random Variables: A Tutorial," *International Statistical Review*, vol. 51, no. 3, pp. pp. 301–319, 1983.

[10] L. Devroye, *Non-Uniform Random Variate Generation*, 1st ed. New York: Springer, 1986.

[11] G. Ulrich, "Computer Generation of Distributions on the m-Sphere," *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 33, no. 2, pp. 158–163, 1984.

[12] A. T. A. Wood, "Simulation of the von Mises–Fisher distribution," *Communications in Statistics—Simulation and Computation*, vol. 23, no. 1, pp. 157–164, 1994.

[13] N. I. Fisher, T. Lewis, and B. J. Embleton, *Statistical Analysis of Spherical Data*. Cambridge University Press, 1987.

[14] S. Jung, "Generating von Mises–Fisher Distribution on the Unit Sphere (S2)," Tech. Rep., 2009.

[15] W. Jakob, "Numerically Stable Sampling of the von Mises Fisher Distribution on S2 (and other Tricks)," Tech. Rep., 2015.

[16] D. J. Best and N. I. Fisher, "Efficient Simulation of the von Mises Distribution," *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 28, no. 2, pp. 152–157, 1979.

[17] G. Kurz, I. Gilitschenski, S. Julier, and U. D. Hanebeck, "Recursive Bingham Filter for Directional Estimation Involving 180 Degree Symmetry," *Journal of Advances in Information Fusion*, vol. 9, no. 2, pp. 90 – 105, Dec. 2014. [Online]. Available: http://isif.org/journal

[18] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 10th ed. New York: Dover, 1972.

[19] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes—The Art of Scientific Computing*, 3rd ed. Cambridge University Press, 2007.

[20] D. Goldberg, "What Every Computer Scientist Should Know About Floating-point Arithmetic," *ACM Computing Surveys*, vol. 23, no. 1, pp. 5–48, Mar. 1991.

[21] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. Baltimore, Maryland, USA: The Johns Hopkins University Press.

[22] S. Sra, "A Short Note on Parameter Approximation for von Mises–Fisher Distributions: And a Fast Implementation of Is(x)," *Computational Statistics*, vol. 27, no. 1, pp. 177–190, 2012.

[23] G. Kurz, I. Gilitschenski, F. Pfaff, and L. Drude, "libDirectional," 2015. [Online]. Available: https://github.com/libDirectional