Deterministic Sampling of the Watson Distribution

Vlad M. Korsakov, **Daniel Frisch**, and **Uwe D. Hanebeck**
Intelligent Sensor-Actuator-Systems Laboratory (ISAS)
Institute for Anthropomatics and Robotics
Karlsruhe Institute of Technology (KIT), Germany
vlad.korsakov@student.kit.edu, daniel.frisch@kit.edu, uwe.hanebeck@kit.edu

***Abstract*—We propose a rejection-free method to obtain homogeneous, deterministic samples from the Watson distribution in $\mathbb{S}^1$, $\mathbb{S}^2$, and $\mathbb{S}^3$. The Watson distribution is a special case of the Bingham distribution in that it is rotationally symmetric around its mean-direction parameter. We propose a solution that generates the samples efficiently: Low-dispersion samples are obtained by transforming low-discrepancy, uniform samples via an orthogonal inverse transform. We compare several strategies for computing this transform efficiently. In computational experiments estimating certain expected values, we demonstrate a much better convergence rate of the proposed deterministic samples compared to random samples.**

I. Introduction

The Watson distribution, also known as the Dimroth-Scheidegger-Watson Distribution [1], is a density on the unit sphere $\mathbb{S}^{p-1} \subset \mathbb{R}^p$. It is defined by

$$f(\underline{x} \mid \underline{\mu}, \kappa) = M\left(\frac{1}{2}, \frac{p}{2}, \kappa\right)^{-1} \cdot e^{\kappa \cdot (\underline{x}^\top \underline{\mu})^2} \quad ,$$

with mean direction vector $\underline{\mu}$ and concentration parameter κ, where

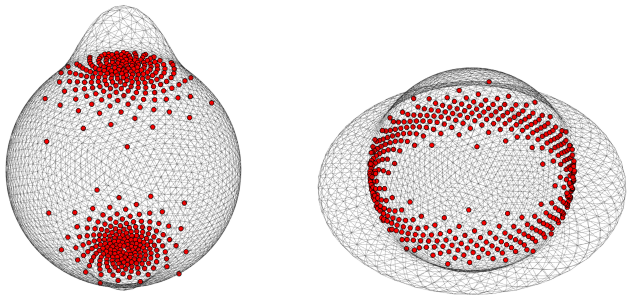
$$M(a, b, z) = \sum_{n=0}^{\infty} \frac{(a)_n}{(b)_n} \cdot \frac{z^n}{n!}$$

is the confluent hypergeometric function of the first kind [1], [2]. The meshes in Figs. 1 and 2 visualize various Watson distributions in $\mathbb{S}^2$.

A. Application Scenarios

Many measurements naturally live on a unit sphere and have an antipodal symmetry, meaning that the directions $\underline{x}$ and $-\underline{x}$ describe the same physical state. Gaussians in $\mathbb{R}^p$ ignore the spherical geometry, and even directional models such as the von Mises–Fisher distribution are not antipodally symmetric. The Watson distribution is one of the simplest models that has this symmetry, which is why efficient sampling from it is useful in practice.

A typical example is rigid-body orientation estimation. Orientations are often parameterized by unit quaternions, because they do not suffer from singularities or the "gimbal lock" of Euler-angle parameterizations [3], remain well-defined where Rodrigues vectors become singular at 180° rotations [4, p. 30], and provide a more compact representation than rotation matrices [4, p. 48]. A unit quaternion and its antipode represent the same rotation, so the distribution on the quaternion sphere $\mathbb{S}^3$ has to be antipodally symmetric. The Watson distribution matches this symmetry of $SO(3)$ [3] and is used for orientation estimation [5], [3] and pose estimation [6]. In sample-based filters [7], new samples have to be drawn from a Watson distribution at every time step, so the cost per sample and the quality of the approximation matter for both runtime and accuracy.

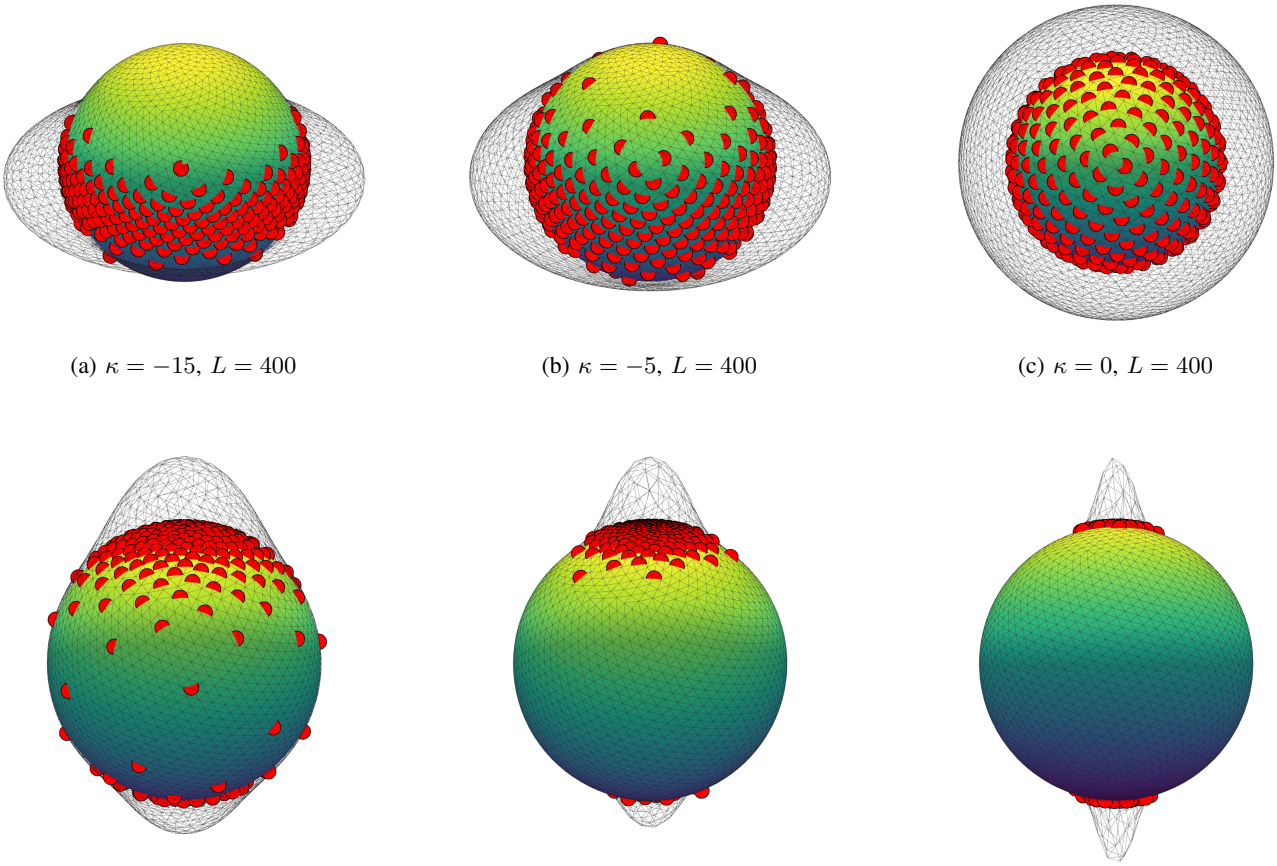


(a) $\kappa = 10$ (b) $\kappa = -20$

Fig. 1: Sets of 400 Samples (red dots) on $\mathbb{S}^2$ obtained with our proposed method for different parameters κ. The mesh represents the respective Watson density function on the sphere. (The sphere itself is not shown here, but in Fig. 2.)

Directional embeddings also occur in several machine learning and data-mining contexts [8], [9], where unit-norm feature vectors carry no preferred sign. With thinly spread data, the Watson distribution can model the data better than the simpler von Mises–Fisher distribution [9]. Sampling is needed for generating synthetic test data and as a building block in EM-style learning of Watson mixture models. Other application scenarios include clustering [10], RGB-D image segmentation [11], and medical imaging, e.g., the diffusion MRI model NODDI, which uses the Watson distribution to describe fiber-orientation dispersion in brain tissue [12], [13].

In each of these settings, expectations under the Watson distribution have to be approximated numerically, e.g., for filter updates, likelihoods, or downstream cost functions. The existing options have known drawbacks: random sampling converges only at the $L^{-1/2}$ Monte-Carlo rate [14], while grid- and subdivision-based schemes [15], [16] pay full computational cost for points that contribute very little to the result. This motivates the deterministic, low-discrepancy sampling approach proposed here.



(a) $\kappa = -15$, $L = 400$ (b) $\kappa = -5$, $L = 400$ (c) $\kappa = 0$, $L = 400$

(d) $\kappa = 5$, $L = 400$ (e) $\kappa = 15$, $L = 400$ (f) $\kappa = 50$, $L = 400$

Fig. 2: Sets of 400 Samples on $\mathbb{S}^2$ with our proposed method and different parameters κ.

B. State of the Art: Quaternion Filtering

Due to the fact that the antipode of a unit quaternion represents the same rotation [15], quaternion filters often employ distributions with antipodal symmetry such as the Watson distribution or its generalization, the Bingham distribution [15]. Existing quaternion filters often rely on random sampling schemes [6], [17].

Techniques to draw random samples from the Watson distribution have been explored in [18] and [10]. However, for numerical integration, random samples have the disadvantage of slow convergence, where the standard deviation of the result decreases with the square root of the number of samples [14].

As alternatives to random sampling, in [15], grid-based quaternion filtering techniques were explored, but this method has the disadvantage that grid points with low weights do not contribute much to the solution, but equally to computational complexity. Another technique explored in [16] is tesseract subdivision, but it also has the same disadvantage as grid-based techniques. Geometry-driven deterministic samples for the related Bingham distribution have been proposed in [19], [20], but no analogous construction is available for the Watson distribution.

C. State of the Art: Transformable Deterministic Sampling

Several related transformation-based deterministic sampling techniques have already been explored. In [21], deterministic uniform sampling on spheres and $SO(3)$ is presented. Furthermore, deterministic samples have been proposed for the Gaussian [22], [23], [24] and von Mises–Fisher [25] distributions, for the circle [26], on the 2-sphere via Dirac mixture reduction [27], as well as for arbitrary densities via rejection [28] or equal-volume sphere packing [29]. Deterministic sampling for all densities that are separable in spherical coordinates was proposed in [30]. Additionally, [31] explored replacing the random sampling step in particle filtering with a deterministic alternative.

D. Contribution

We propose and evaluate multiple techniques for computing deterministic samples of the Watson distribution in $\mathbb{S}^1$, $\mathbb{S}^2$, and $\mathbb{S}^3$. In the case of $\mathbb{S}^2$, we additionally introduce a closed-form solution for the orthogonal inverse transform (Section IV). We present several numerical computation techniques, including inverse interpolation and an inverse ordinary differential equation (ODE), as well as a substitution to avoid numerical

problems with the inverse ODE, and an ODE formulation with sample events (Section VI). These numerical methods can not only be used on $\mathbb{S}^2$, but also in all other dimensions. We provide an empirical comparison of runtime performance of the different methods (Fig. 3). In addition, we provide an empirical comparison of the performance of several sample types by using them for numerical integration (Section VII).

E. Structure

The paper is structured as follows. In Section II, we list a few methods producing *uniform* low-discrepancy points with the required dimensions and periodicity. These are then transformed via the quantile functions derived in Section III for $\mathbb{S}^1$, Section IV for $\mathbb{S}^2$, and Section V for $\mathbb{S}^3$. Some of these quantile functions cannot be represented in closed form but have to be evaluated numerically with one of the methods described in Section VI, where there are always very fast choices, see Fig. 3. Finally, we perform an evaluation in Section VII.

II. Uniform Reference Point Sets

A. Equidistant Point Sets

For the one-dimensional / univariate (1D) case, the equidistant point set $\{x_i\}_{i=1}^{L}$, where L denotes the total number of points,

$$x_i = i \cdot \frac{1}{L} \ , \qquad i \in \{0, 1, \dots, L-1\} \ , \tag{1}$$

or its centered variant

$$x_i = \frac{2i-1}{2L} \ , \qquad i \in \{1, 2, \dots, L\} \ , \tag{2}$$

can be used [25].

B. Fibonacci–Rank-1 Lattice

For the two-dimensional (2D) case, the Fibonacci–Rank-1 lattice can be used. It is the best possible low-discrepancy point set in the periodic unit square or torus $[0,1)^2$ [25], [32]. It is defined as

$$\underline{x}_i = i \cdot \begin{bmatrix} \frac{1}{\mathrm{Fib}_{k+1}} \\ \frac{\mathrm{Fib}_k}{\mathrm{Fib}_{k+1}} \end{bmatrix} \bmod 1 \ , \tag{3}$$

$$i \in \{0, 1, \dots, \mathrm{Fib}_{k+1} - 1\} \ , \tag{4}$$

or its centered variant

$$\underline{x}_i = i \cdot \begin{bmatrix} \frac{1}{\mathrm{Fib}_{k+1}} \\ \frac{\mathrm{Fib}_k}{\mathrm{Fib}_{k+1}} \end{bmatrix} + \frac{1}{2 \cdot \mathrm{Fib}_{k+1}} \bmod 1 \ , \tag{5}$$

$$i \in \{0, 1, \dots, \mathrm{Fib}_{k+1} - 1\} \ , \tag{6}$$

where Fib_k is the k-th Fibonacci number [25]. It is periodic in both directions. One disadvantage of this point set is that the total number of points L has to be a Fibonacci number.

C. Fibonacci–Kronecker Lattice

Unlike the Fibonacci–Rank-1 lattice, the Fibonacci–Kronecker lattice can be constructed for an arbitrary number of points. It is defined by

$$\underline{x}_i = i \cdot \begin{bmatrix} \frac{1}{L} \\ \frac{1}{\Phi} \end{bmatrix} \bmod 1 \ , \qquad i \in \{0, 1, \dots, L-1\} \ , \tag{7}$$

and its centered variant

$$\underline{x}_i = \begin{bmatrix} \frac{2i-1}{2L} \\ \frac{i}{\Phi} \end{bmatrix} \bmod 1 \ , \qquad i \in \{1, 2, \dots, L\} \ , \tag{8}$$

where Φ is the inverse golden ratio

$$\frac{1}{\Phi} = \frac{\sqrt{5}-1}{2} = 0.618033\dots$$

[25], [30].

The Fibonacci–Kronecker lattice consists of equidistant points along the first coordinate and the Fibonacci–Kronecker (also called Golden-Fibonacci) sequence along the second. It is periodic along the latter coordinate and can therefore be projected onto a sphere, with that axis mapped to the periodic spherical coordinate [30].

D. Rank-1 Lattice

A generalization of the Fibonacci–Rank-1 lattice is the class of rank-1 lattices. They can be constructed via

$$\mathcal{P}_L(\underline{g}) = \{i \cdot \underline{g} \bmod 1\}_{i=1}^{L} \ , \tag{9}$$

where $\underline{g}$ is a generator vector. Possible generators were explored in [33], [34, 5.3.2.B] and [34, p. 3.1.2].

E. Sobol

Another well-known alternative is the Sobol sequence [30], [35], [36]. It forms a low-discrepancy point set with periodic structure in every coordinate direction and is implemented in many major programming languages [30]. Note that the Sobol sequence is designed to provide $n = 2^m$ low-discrepancy points in $[0,1)^d$. Choosing a sample size n that is not a power of two can significantly degrade its balance properties [37].

III. Watson Distribution on $\mathbb{S}^1$

On $\mathbb{S}^1$, we parametrize the unit circle using the angle $\varphi \in [0, 2\pi)$ via

$$[\varphi] \to \begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix} \ .$$

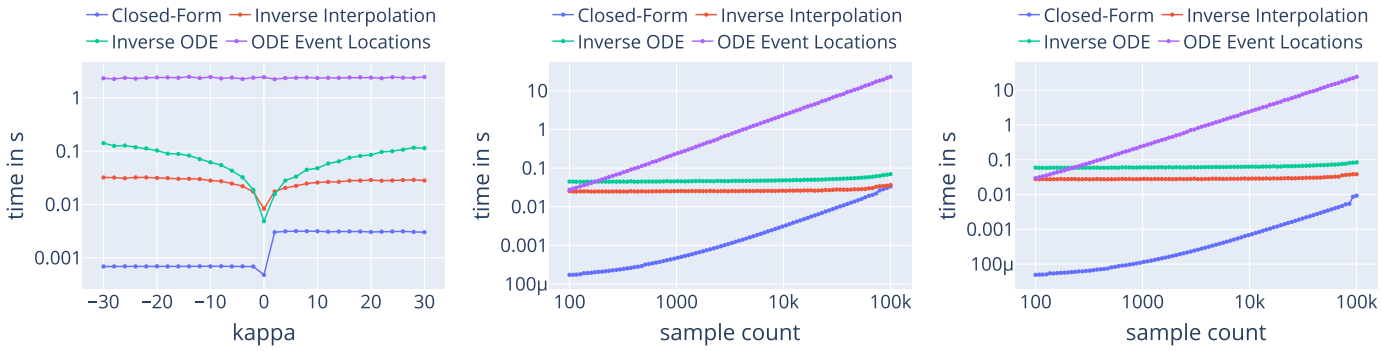
For $\underline{\mu} = \begin{bmatrix} 1 & 0 \end{bmatrix}^\top$, this yields the density function

$$f(\varphi \mid \kappa) \propto \exp(\kappa \cdot (\cos(\varphi))^2) \ .$$

The associated cumulative density function (CDF) is given by

$$F(\varphi \mid \kappa) \propto \int_0^{\varphi} e^{\kappa \cdot (\cos(\varphi))^2} \, \mathrm{d}\varphi \ . \tag{10}$$

Unlike the $\mathbb{S}^2$ case, this integral does not simplify to a closed form, so one of the other methods described in Section VI must be used for computation.



(a) 10k samples, variable κ (b) $\kappa = 10$, variable sample count (c) $\kappa = -10$, variable sample count

Fig. 3: Computation time for generating identical samples on $\mathbb{S}^2$ using our closed-form approach versus the numerical method in Section VI, shown for different κ and L.

IV. Watson Distribution on $\mathbb{S}^2$

On $\mathbb{S}^2$, the following transformation maps to spherical coordinates

$$\begin{bmatrix}\theta\\ \varphi\end{bmatrix} \to \begin{bmatrix}\sin(\theta)\cdot\cos(\varphi)\\ \sin(\theta)\cdot\sin(\varphi)\\ \cos(\theta)\end{bmatrix} \quad . \tag{11}$$

A. Generic Procedure

When transformed to the spherical coordinate system and choosing $\underline{\mu} = \begin{bmatrix}0 & 0 & 1\end{bmatrix}^\top$, the density function of the Watson distribution can be observed to be constant with respect to φ, and we get

$$f(\theta, \varphi \mid \kappa) \propto \exp(\kappa \cdot (\cos(\theta))^2) \quad . \tag{12}$$

This does not depend on φ and is therefore separable into two factors $f(\theta, \varphi) = f_1(\theta) \cdot f_2(\varphi)$, where f_1 is constant. From this, the samples could already be computed as described in Section VI-B.

B. Closed-Form Solution

As an alternative, we will now pursue a closed-form solution (that will turn out to be faster, see Fig. 3).

1) Cumulative Function: The CDF

$$F(\theta, \varphi \mid \kappa) \tag{13}$$

$$\propto \int_0^{\varphi}\int_0^{\theta} e^{\kappa\cdot(\cos(\theta))^2}\sin(\theta)\,\mathrm{d}\theta\,\mathrm{d}\varphi \tag{14}$$

$$\propto \begin{cases}\varphi\cdot(\mathrm{Erfi}(\sqrt{\kappa}) - \mathrm{Erfi}(\sqrt{\kappa}\cdot\cos(\theta))) & , \quad \kappa > 0\\ \varphi\cdot(\mathrm{Erf}(\sqrt{-\kappa}) - \mathrm{Erf}(\sqrt{-\kappa}\cdot\cos(\theta))) & , \quad \kappa < 0\end{cases} , \tag{15}$$

where Erf is the error function and Erfi is the imaginary error function $\mathrm{Erfi}(z) = -i\cdot\mathrm{Erf}(i\cdot z)$, is separable into two factors $F(\theta, \varphi) = F_1(\theta)\cdot F_2(\varphi)$

$$F_1(\theta) \propto \begin{cases}\mathrm{Erfi}(\sqrt{\kappa}) - \mathrm{Erfi}(\sqrt{\kappa}\cdot\cos\theta) & , \quad \kappa > 0\\ \mathrm{Erf}(\sqrt{-\kappa}) - \mathrm{Erf}(\sqrt{-\kappa}\cdot\cos\theta) & , \quad \kappa < 0\end{cases} , \tag{16}$$

$$F_2(\varphi) = \frac{\varphi}{2\pi} \quad . \tag{17}$$

Algorithm 1 Deterministic sampling on $\mathbb{S}^2$

Require: concentration κ, sample count L
1: generate uniform points $\{(p_i, q_i)\}_{i=1}^{L} \subset [0,1)^2$ from a 2D low-discrepancy sequence (Section II)
2: **for** $i = 1, \ldots, L$ **do**
3: $\quad\theta_i \leftarrow Q_\theta(p_i; \kappa)$ via (18)
4: $\quad\varphi_i \leftarrow Q_\varphi(q_i) = 2\pi q_i$
5: $\quad\underline{x}_i \leftarrow [\sin\theta_i\cos\varphi_i,\ \sin\theta_i\sin\varphi_i,\ \cos\theta_i]^\top$
6: **end for**
7: **return** $\{\underline{x}_i\}_{i=1}^{L}$

2) Quantile Function: Inverting and simplifying yields the quantile functions, where $p, q \in [0, 1]$ are uniform inputs,

$$Q_\theta(p) = \tag{18}$$

$$\begin{cases}\cos^{-1}\left(\frac{1}{\sqrt{\kappa}}\cdot\mathrm{Erfi}^{-1}((1-2p)\cdot\mathrm{Erfi}(\sqrt{\kappa}))\right) & , \quad \kappa > 0\\ \cos^{-1}\left(\frac{1}{\sqrt{-\kappa}}\cdot\mathrm{Erf}^{-1}((1-2p)\cdot\mathrm{Erf}(\sqrt{-\kappa}))\right) & , \quad \kappa < 0\end{cases} ,$$

$$Q_\varphi(q) = 2\pi q \quad . \tag{19}$$

These quantile functions can be used to transform uniform reference point sets. In particular, insert the uniform coordinates from a two-dimensional low-discrepancy sequence into Q_θ and Q_φ, respectively, and transform the resulting spherical coordinates to Cartesian with (11). The full procedure is summarized in Algorithm 1.

For example, choosing the centered variant of the Fibonacci–Kronecker lattice as described in (8) yields

$$\underline{x}_i^f = \begin{bmatrix}\sqrt{1-w^2}\cdot\cos\left(\frac{2\pi\cdot i}{\Phi}\right)\\ \sqrt{1-w^2}\cdot\sin\left(\frac{2\pi\cdot i}{\Phi}\right)\\ w\end{bmatrix} ,$$

where

$$w = \begin{cases}\frac{1}{\sqrt{\kappa}}\cdot\mathrm{Erfi}^{-1}\left(\frac{1-2i+L}{L}\cdot\mathrm{Erfi}(\sqrt{\kappa})\right) & , \quad \kappa > 0\\ \frac{1}{\sqrt{-\kappa}}\cdot\mathrm{Erf}^{-1}\left(\frac{1-2i+L}{L}\cdot\mathrm{Erf}(\sqrt{-\kappa})\right) & , \quad \kappa < 0\\ \frac{1-2i+L}{L} & , \quad \kappa = 0\end{cases} .$$

Note that in most scientific programming languages, the inverse imaginary error function $\mathrm{Erfi}^{-1}(\cdot)$ is currently not implemented. We found that a fast way of computing it is solving $\mathrm{Erfi}(x) = y$ with Newton's method, where Erfi itself can be evaluated, e.g., with the Faddeeva package [38].

Figures 1 and 2 show samples produced by this procedure for several concentration parameters. For positive κ, the samples concentrate around the two poles defined by $\pm\underline{\mu}$ (bipolar regime), with concentration growing as κ increases. For negative κ, the samples concentrate on the equator orthogonal to $\underline{\mu}$ (girdle regime). For $\kappa = 0$, the Watson distribution reduces to the uniform distribution on the sphere, and the samples cover $\mathbb{S}^2$ uniformly.

V. Watson Distribution on $\mathbb{S}^3$

Similarly, on $\mathbb{S}^3$, the coordinate transformation

$$\begin{bmatrix}\psi\\ \theta\\ \varphi\end{bmatrix} \to \begin{bmatrix}\cos\psi\\ \sin\psi\cdot\cos\theta\\ \sin\psi\cdot\sin\theta\cdot\cos\varphi\\ \sin\psi\cdot\sin\theta\cdot\sin\varphi\end{bmatrix}$$

can be used. For $\underline{\mu} = \begin{bmatrix}1 & 0 & 0 & 0\end{bmatrix}^\top$, this yields the density function

$$f(\psi,\theta,\varphi \mid \kappa) \propto \exp(\kappa\cdot(\cos(\psi))^2) \ .$$

The associated CDF

$$\begin{aligned} &F(\psi,\theta,\varphi)\\ &\propto \int_0^{\varphi}\int_0^{\theta}\int_0^{\psi} e^{\kappa\cdot(\cos(\psi))^2}\cdot\sin(\theta)\cdot(\sin(\psi))^2\,\mathrm{d}\psi\,\mathrm{d}\theta\,\mathrm{d}\varphi \qquad (20)\\ &\propto \int_0^{\psi} e^{\kappa\cdot(\cos(\psi))^2}\cdot\varphi\cdot(1-\cos(\theta))\cdot(\sin(\psi))^2\,\mathrm{d}\psi \qquad (21)\end{aligned}$$

is separable into three components, two of which have closed form quantile functions, with $p \in [0,1]$,

$$\begin{aligned} Q_\theta(p) &= \cos^{-1}(1-2p) \qquad (22)\\ Q_\varphi(p) &= 2\pi p \ . \qquad (23)\end{aligned}$$

$Q_\psi(p)$ does not have a closed form, so one of the methods described in Section VI has to be used. The inverse interpolation method in Section VI-A is advisable here.

VI. Computation of Quantile Functions from Densities

This section will describe and evaluate various numerical methods to compute quantile functions. Some of them have already been proposed in [30], but we will reiterate them here for reference, tailor them to the specific case of Watson distributions, including a novel substitution for the inverse ODE method, and compare their computational complexity.

A. Numerical Integration and Inverse Interpolation

The CDF of, e.g., the Watson distribution can also be computed numerically. To do this, we formulate an ODE

$$F'(x) = f_x(x) \ , \qquad F(0) = 0 \ , \qquad (24)$$

where f_x is the 1D marginal density of the corresponding variable. Solving the ODE gives us a numerical representation (input-output-tuples) of the CDF. Swapping the tuple entries and then interpolating yields an approximation of the desired quantile function Q. This method is universally applicable.

1) Inverse Interpolation for $\mathbb{S}^2$: On $\mathbb{S}^2$, the 1D marginal of the polar angle is

$$\int_0^{2\pi} f(\theta,\varphi)\cdot\sin(\theta)\,\mathrm{d}\varphi = 2\pi f(\theta,0)\cdot\sin(\theta) \ ,$$

where f is the normalized Watson density.

Using notation commonly used in ODE solver documentation [30], we integrate the CDF of the polar angle θ on $[0,\pi]$,

$$u'(\theta) = 2\pi\cdot f(\theta,0)\cdot\sin(\theta) \ , \qquad u(0) = 0 \ , \qquad (25)$$

where $u(\theta) = F_1(\theta)$. We can then use a typical numerical ODE solver to compute (θ,u)-tuples approximating F_1. Exchanging the coordinates and interpolating them then yields an approximation of the quantile function Q [30].

2) Inverse Interpolation for $\mathbb{S}^3$: Similarly, for $\mathbb{S}^3$, we have the 1D marginal of ψ

$$\begin{aligned} &\int_0^{2\pi}\int_0^{\pi} f(\psi,\theta,\varphi)\cdot(\sin(\psi))^2\cdot\sin(\theta)\,\mathrm{d}\theta\,\mathrm{d}\varphi \qquad (26)\\ &\quad = 4\pi\cdot f(\psi,0,0)\cdot(\sin(\psi))^2 \ , \qquad (27)\end{aligned}$$

where f is the normalized Watson density. We then solve the ODE

$$u'(\psi) = 4\pi\cdot f(\psi,0,0)\cdot(\sin(\psi))^2 \ , \quad u(0) = 0 \ , \qquad (28)$$

and compute the quantile function Q in the same manner as in Section VI-A1.

B. Inverse ODE

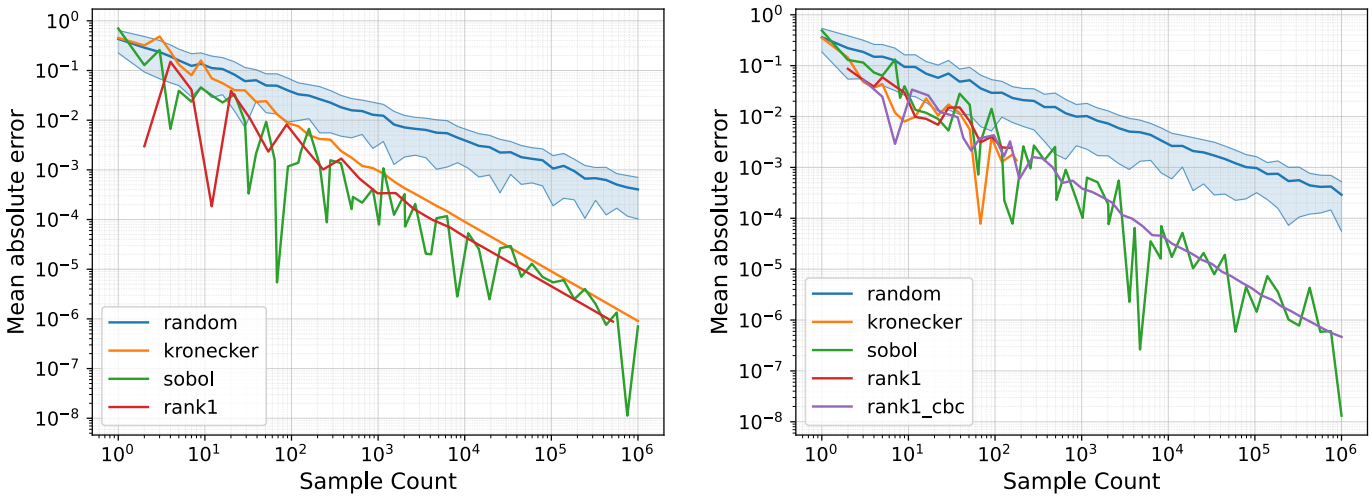
Alternatively, we can directly write Q_i as an ODE in terms of f_x [30]

$$u'(p) = \frac{1}{f_x(u(p))} \ , \qquad u(0) = 0 \ . \qquad (29)$$

1) Inverse ODE for $\mathbb{S}^2$: In particular, for θ on $\mathbb{S}^2$, this yields

$$u'(p) = \frac{1}{2\pi\cdot f(u(p),0)\cdot\sin(u(p))} \ , \quad u(0) = 0 \ . \qquad (30)$$

However, the denominator, due to the sine function factor, becomes zero at the poles ($\theta = 0,\pi$), causing numerical issues due to division by zero.



(a) $\underline{x}_0 = \begin{bmatrix}4 & 5 & 6\end{bmatrix}^\top$ and $\kappa = 10$ on $\mathbb{S}^2$

(b) $\underline{x}_0 = \begin{bmatrix}4 & 5 & 6 & 7\end{bmatrix}^\top$ and $\kappa = 10$ on $\mathbb{S}^3$

Fig. 4: Mean absolute integration error for various sampling methods (colored lines) and numbers of samples (abscissa).

2) Inverse ODE for $\mathbb{S}^2$ with Substitution: This can be circumvented via the following substitution. From (14) it looks advantageous to substitute $-\cos(\theta) \to z$, $z \in [-1, 1]$. This eliminates the $\sin(u(p))$ factor and leads to the ODE

$$u'(p) = \frac{1}{2\pi \cdot f(\cos^{-1}(u(p)), 0)} \ , \qquad u(0) = -1 \tag{31}$$

which can now be solved without numerical problems. Note that $z = -\cos(\theta)$ can be identified with the Euclidean z coordinate due to the spherical coordinate system (11).

C. ODE with Event Locations

Using the original ODE from Section VI-A, ODE solvers can also take a set of events as input and then find the locations where those events occur [34, Sec. 7.4.5]. This can be used to directly evaluate $Q_i(p_{\text{target}_j})$ by defining L event functions

$$e_j(\theta, u) = u - p_{\text{target}_j} \ , \qquad j \in \{1, \dots, L\} \ , \tag{32}$$

and letting the ODE solver find the locations where these events occur. These event locations can then be directly interpreted as the transformed samples [34, Sec. 7.4.5]. This method is universally applicable, however quite slow, see Fig. 3.

D. Performance Comparison

We evaluated the performance of computing samples for various κ and sample counts L on $\mathbb{S}^2$ using the closed form proposed in Section IV and the numerical methods described in Section VI, i.e., numeric integration with inverse interpolation, inverse ODE, and inverse ODE with event locations. All approaches implement the same orthogonal inverse transform and produce the same samples, except for negligible numerical deviations.

As expected, the closed-form solution consistently performed the best. Numerical integration with inverse interpolation and the inverse ODE performed similarly and slightly worse than the closed form. For the inverse ODE with event locations, we observed a large performance drop at higher sample counts because each event location adds significantly more overhead than in the other methods. See Fig. 3 for details.

VII. Evaluation

This section evaluates different kinds of samples created with the proposed algorithm and compares them against existing methods. We numerically integrate

$$\int_{\mathbb{S}^n} f(\underline{x}) \cdot g(\underline{x}) \, \mathrm{d}\underline{x} \approx \frac{1}{L} \sum_i g(\underline{x}_i) \ , \qquad \underline{x}_i \sim f(\underline{x}) \ , \tag{33}$$

with $\underline{x} \in \mathbb{S}^n$ and where f is the Watson probability density function (PDF) and

$$g(\underline{x}) = ||\underline{x} - \underline{x}_0|| \tag{34}$$

is the exemplary test function, here the distance of $\underline{x}$ from the point $\underline{x}_0$.

Using $\kappa = 10$ and $\underline{x}_0 = \begin{bmatrix}4 & 5 & 6\end{bmatrix}^\top$ for $\mathbb{S}^2$ as well as $\underline{x}_0 = \begin{bmatrix}4 & 5 & 6 & 7\end{bmatrix}^\top$ for $\mathbb{S}^3$, we compared the absolute integration error for varying sample counts from 1 to 10^6. The reference integration result was computed with `scipy.integrate.nquad` from the SciPy Python library.

$\mathbb{S}^2$ samples were computed using the closed-form formula described in Section IV. Fig. 4a shows the absolute errors for random samples, samples using the Fibonacci–Kronecker lattice described in Section II-C, Sobol samples and samples based on the Fibonacci–Rank-1 lattice described in Section II-B.

Furthermore, $\mathbb{S}^3$ samples were computed numerically using the method described in Section VI-A. Fig. 4b shows the absolute errors for random samples, Sobol samples and samples based on different kinds of rank-1 lattices using various generators. The samples labeled `kronecker` and `rank1` were constructed with generators detailed in [34, 5.3.2.B] and [34, p. 3.1.2], respectively. Note that these generators are only available for $L \leq 177$ and $L \leq 155$.

The samples labeled `rank1_cbc` were constructed using the component-by-component construction method for rank-1 lattice rules with a prime number of points and product weights described in [33].

In both $\mathbb{S}^2$ and $\mathbb{S}^3$ (Fig. 4), the deterministic sample sets clearly outperform random samples, exhibiting an empirical convergence of roughly L^{-1} rather than the $L^{-1/2}$ rate of random samples. Among the deterministic sets, the differences between Sobol, Fibonacci–Kronecker, Fibonacci–Rank-1, and the rank-1 lattices are comparatively small, with all of them lying close to the same L^{-1} trend.

The Python source code of our implementation for generating the deterministic samples as well as reproducing the evaluation is available at [39], [40]. In addition, it is available on CodeOcean and linked on the IEEE Xplore page.

VIII. Conclusion

A. Advantages

We proposed a rejection-free approach for deterministic sampling of the Watson distribution on $\mathbb{S}^1$, $\mathbb{S}^2$, and $\mathbb{S}^3$ by transforming low-discrepancy uniform samples with an orthogonal inverse transform. We presented and compared different methods of computing quantile functions, including a closed form for $\mathbb{S}^2$. With the proposed deterministic samples, we observed an empirical convergence of roughly L^{-1} — much faster than the $L^{-1/2}$ for random samples. For example, we can reach the error 0.01 in Fig. 4b with about 10 deterministic instead of about 100 random samples.

B. Limitations

In the $\mathbb{S}^1$ and $\mathbb{S}^3$ case, closed-form mappings are not available for all coordinates, so numerical methods have to be used (some of which are still very fast).

Furthermore, the Watson distribution is spherically symmetric and therefore, unlike the Bingham distribution, not able to represent correlations between different angles.

C. Outlook

Future work includes extending the approach to higher dimensions. In addition, deterministic sampling on the more general Bingham distribution has to be investigated. This would enable efficient low-discrepancy sampling for more general antipodally symmetric models.

References

[1] Kanti V Mardia and Peter E Jupp. *Directional Statistics*. John Wiley & Sons, 2009.

[2] Lukas Sablica, Kurt Hornik, and Josef Leydold. "Random Sampling from the Watson Distribution". English. In: Research Report Series / Department of Statistics and Mathematics 134 (2022). DOI: 10.57938/6e5ffebc-7ad2-4225-9909-d8833529e808.

[3] Igor Gilitschenski, Gerhard Kurz, Simon J. Julier, and Uwe D. Hanebeck. "Unscented Orientation Estimation Based on the Bingham Distribution". In: *IEEE Transactions on Automatic Control* 61.1 (Jan. 2016), pp. 172–177. DOI: 10.1109/TAC.2015.2423831.

[4] Hashim A. Hashim. *Special Orthogonal Group SO(3), Euler Angles, Angle-axis, Rodriguez Vector and Unit-Quaternion: Overview, Mapping and Challenges*. 2024. URL: https://arxiv.org/abs/1909.06669.

[5] Yu-Hui Chen, Dennis Wei, Gregory Newstadt, Marc DeGraef, Jeffrey Simmons, and Alfred Hero. "Statistical estimation and clustering of group-invariant orientation parameters". In: *2015 18th International Conference on Information Fusion (Fusion)*. 2015, pp. 719–726.

[6] Hugh Durrant-Whyte, Nicholas Roy, and Pieter Abbeel. "Monte Carlo Pose Estimation with Quaternion Kernels and the Bingham Distribution". In: *Robotics: Science and Systems VII*. 2012, pp. 97–104.

[7] Daniel Frisch and Uwe D. Hanebeck. "Progressive Bayesian Filtering with Coupled Gaussian and Dirac Mixtures". In: *Proceedings of the 23rd International Conference on Information Fusion (Fusion 2020)*. Virtual, July 2020. DOI: 10.23919/FUSION45008.2020.9190540.

[8] Inderjit S. Dhillon and Suvrit Sra. "Matrix nearness problems in data mining". In: 2007. URL: https://api.semanticscholar.org/CorpusID:120493554.

[9] Avleen S. Bijral, Markus Breitenbach, and Greg Grudic. "Mixture of Watson Distributions: A Generative Model for Hyperspherical Embeddings". In: *Proceedings of the Eleventh International Conference on Artificial Intelligence and Statistics*. Ed. by Marina Meila and Xiaotong Shen. Vol. 2. Proceedings of Machine Learning Research. San Juan, Puerto Rico: PMLR, Mar. 2007, pp. 35–42. URL: https://proceedings.mlr.press/v2/bijral07a.html.

[10] Stephen J. Maybank, Liu Liu, and Dacheng Tao. "Generalised Watson Distribution on the Hypersphere with Applications to Clustering". In: *Journal of Mathematical Imaging and Vision* 65.2 (Apr. 1, 2023), pp. 302–322. DOI: 10.1007/s10851-022-01118-7.

[11] Md Hasnat, Olivier Alata, and Alain Tremeau. "Unsupervised RGB-D image segmentation using joint clustering and region merging". In: Sept. 2014. DOI: 10.5244/C.28.17.

[12] Hui Zhang, Torben Schneider, Claudia A. Wheeler-Kingshott, and Daniel C. Alexander. "NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain." eng. In: *NeuroImage* 61.4 (July 2012), pp. 1000–1016. DOI: 10.1016/j.neuroimage.2012.03.072.

[13] Kurt G. Schilling, Vaibhav Janve, Yurui Gao, Iwona Stepniewska, Bennett A. Landman, and Adam W. Anderson. "Histological validation of diffusion MRI fiber orientation distributions and dispersion". In: *NeuroImage* 165 (2018), pp. 200–221. DOI: https://doi.org/10.1016/j.neuroimage.2017.10.046.

[14] Russel E. Caflisch. "Monte Carlo and Quasi-Monte Carlo Methods". en. In: *Acta Numerica* 7 (Jan. 1998), pp. 1–49. DOI: 10.1017/S0962492900002804. (Visited on 11/05/2021).

[15] Kailai Li, Florian Pfaff, and Uwe D. Hanebeck. "Grid-Based Quaternion Filter for SO(3) Estimation". In: *Proceedings of the 2020 European Control Conference (ECC 2020)*. Virtual, May 2020. DOI: 10.23919/ECC51009.2020.9143723.

[16] Gerhard Kurz, Florian Pfaff, and Uwe D. Hanebeck. "Discretization of SO(3) Using Recursive Tesseract Subdivision". In: *Proceedings of the 2017 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2017)*. Daegu, Republic of Korea, Nov. 2017. DOI: 10.1109/MFI.2017.8170406.

[17] Kailai Li, Gerhard Kurz, Lukas Bernreiter, and Uwe D. Hanebeck. "Simultaneous Localization and Mapping Using a Novel Dual Quaternion Particle Filter". In: *Proceedings of the 21st International Conference on Information Fusion (Fusion 2018)*. Cambridge, United Kingdom, July 2018. DOI: 10.23919/ICIF.2018.8455347.

[18] Lukas Sablica, Kurt Hornik, and Josef Leydold. "Efficient Sampling From the Watson Distribution in Arbitrary Dimensions". In: *Journal of Computational and Graphical Statistics* 34.3 (2025), pp. 923–933. DOI: 10.1080/10618600.2024.2416521.

[19] Kailai Li, Daniel Frisch, Benjamin Noack, and Uwe D. Hanebeck. "Geometry-Driven Deterministic Sampling for Nonlinear Bingham Filtering". In: *Proceedings of the 2019 European Control Conference (ECC 2019)*. Naples, Italy, June 2019. DOI: 10.23919/ECC.2019.8796102.

[20] Kailai Li, Daniel Frisch, Susanne Radtke, Benjamin Noack, and Uwe D. Hanebeck. "Wavefront Orientation Estimation Based on Progressive Bingham Filtering". In: *Proceedings of the IEEE ISIF Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF 2018)*. Oct. 2018. DOI: 10.1109/SDF.2018.8547094.

[21] A. Yershova and S.M. LaValle. "Deterministic sampling methods for spheres and SO(3)". In: *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*. Vol. 4. 2004, 3974–3980 Vol.4. DOI: 10.1109/ROBOT.2004.1308891.

[22] Daniel Frisch and Uwe D. Hanebeck. "The Generalized Fibonacci Grid as Low-Discrepancy Point Set for Optimal Deterministic Gaussian Sampling". In: *Journal of Advances in Information Fusion* 18.1 (June 2023), pp. 16–34. URL: https://isif.org/media/generalized-fibonacci-grid-low-discrepancy-point-set-optimal-deterministic-gaussian-sampling.

[23] Daniel Frisch and Uwe D. Hanebeck. "Deterministic Gaussian Sampling With Generalized Fibonacci Grids". In: *Proceedings of the 24th International Conference on Information Fusion (Fusion 2021)*. Sun City, South Africa, Nov. 2021. DOI: 10.23919/FUSION49465.2021.9626975.

[24] Daniel Frisch and Uwe D. Hanebeck. "Efficient Deterministic Conditional Sampling of Multivariate Gaussian Densities". In: *Proceedings of the 2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2020)*. Virtual, Sept. 2020. DOI: 10.1109/MFI49285.2020.9235212.

[25] Daniel Frisch and Uwe D. Hanebeck. "Deterministic Von Mises-Fisher Sampling on the Sphere Using Fibonacci Lattices". In: *Proceedings of the combined IEEE 2023 Symposium Sensor Data Fusion and International Conference on Multisensor Fusion and Integration (SDF-MFI 2023)*. Bonn, Germany, Nov. 2023. DOI: 10.1109/SDF-MFI59545.2023.10361396.

[26] Daniel Frisch and Uwe D. Hanebeck. "Deterministic Sampling on the Circle Using Projected Cumulative Distributions". In: *Proceedings of the 25th International Conference on Information Fusion (Fusion 2022)*. Linköping, Sweden, July 2022. DOI: 10.23919/FUSION49751.2022.9841299.

[27] Daniel Frisch, Kailai Li, and Uwe D. Hanebeck. "Optimal Reduction of Dirac Mixture Densities on the 2-Sphere". In: *Proceedings of the 1st Virtual IFAC World Congress (IFAC-V 2020)*. July 2020. DOI: 10.1016/j.ifacol.2020.12.1856.

[28] Daniel Frisch and Uwe D. Hanebeck. "Rejection Sampling from Arbitrary Multivariate Distributions Using Generalized Fibonacci Lattices". In: *Proceedings of the 25th International Conference on Information Fusion (Fusion 2022)*. Linköping, Sweden, July 2022. DOI: 10.23919/FUSION49751.2022.9841322.

[29] Daniel Frisch and Uwe D. Hanebeck. "Deterministic Sampling of Arbitrary Densities Using Equal Sphere Packing of Volume under the Density (PoVuD)". In: *Proceedings of the 26th International Conference on Information Fusion (Fusion 2023)*. Charleston, USA, June 2023. DOI: 10.23919/FUSION52260.2023.10224093.

[30] Daniel Frisch and Uwe D. Hanebeck. "Deterministic Sampling with Separation of Variables in Spherical Coordinates". In: *Proceedings of the 28th International Conference on Information Fusion (FUSION 2025)*. Rio de Janeiro, Brazil, July 2025. DOI: 10.23919/FUSION65864.2025.11124009.

[31] Andrey A Popov and Renato Zanetti. "Deterministic Optimal Transport-based Gaussian Mixture Particle Filtering for Verifiable Applications". In: *arXiv preprint arXiv:2501.17302* (2025).

[32] Aicke Hinrichs and Jens Oettershagen. "Optimal Point Sets for Quasi-Monte Carlo Integration of Bivariate Periodic Functions with Bounded Mixed Derivatives". en. In: *Monte Carlo and Quasi-Monte Carlo Methods*. Ed. by Ronald Cools and Dirk Nuyens. Springer Proceedings in Mathematics & Statistics. Cham: Springer International Publishing, 2016, pp. 385–405. DOI: 10.1007/978-3-319-33507-0_19.

[33] Dirk Nuyens and Ronald Cools. "Fast Component-by-Component Construction, a Reprise for Different Kernels". In: *Monte Carlo and Quasi-Monte Carlo Methods 2004*. Ed. by Harald Niederreiter and Denis Talay. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 373–387.

[34] Daniel Frisch. "Transformable Deterministic Sampling". PhD thesis. Karlsruher Institut für Technologie (KIT), 2025. 301 pp. DOI: 10.5445/IR/1000179985.

[35] Stephen Joe and Frances Y. Kuo. "Remark on Algorithm 659: Implementing Sobol's Quasirandom Sequence Generator". In: *ACM Trans. Math. Softw.* 29.1 (Mar. 2003), pp. 49–57. DOI: 10.1145/641876.641879.

[36] I.M. Sobol. "Uniformly Distributed Sequences with an Additional Uniform Property". In: *USSR Computational Mathematics and Mathematical Physics* 16.5 (1976), pp. 236–242. DOI: 10.1016/0041-5553(76)90154-3.

[37] Art B. Owen. *On dropping the first Sobol' point*. en. Aug. 2020. URL: https://arxiv.org/abs/2008.08051v4 (visited on 01/25/2026).

[38] USA Massachusetts Institute of Technology Boston. *Faddeeva Package*. 2015. URL: http://ab-initio.mit.edu/wiki/index.php/Faddeeva_Package.

[39] Vlad Korsakov. *Code to Evaluate and Compare Deterministic Watson Samples.* 2026. URL: https://github.com/Vlad-Kor/FUSION26_Watson.

[40] Vlad Korsakov. *Code to Evaluate and Compare Deterministic Watson Samples.* 2026. URL: https://github.com/KIT-ISAS/deterministic_sampling_Watson_FUSION26_Korsakov.