A Discrete Radon Transform Based on the Area of Cube-Plane Intersection

Robert Beinert, Jonas Bresch, Michael Quellmalz

***Abstract*—The Radon transform is a fundamental tool for analyzing data in tomographic imaging, optimal transport, crystallography, and geometric analysis. Numerical computations require an accurate discretization. To deal with voxelized images and objects, we derive a closed-form, piecewise polynomial expression for the Radon transform of an axis-aligned cube in arbitrary dimension d. Building on this formula, we propose a discrete Radon transform in $\mathbb{R}^d$ that is both analytically exact for voxelized data and computationally efficient. For improved numerical stability, we introduce a regularized variant replacing the Radon transform of a cube, i.e. the $(d-1)$-dimensional area of the intersection between that cube and a hyperplane, by the d-dimensional volume of the intersection between the cube and a thin slab around the hyperplane. Numerical experiments demonstrate the effectiveness of the proposed approach in several applications including 3D shape matching, classification, and sliced Wasserstein barycenters. The computational efficiency in higher dimensions is verified by a comparison with Monte Carlo integration.**

***Index Terms*—Radon transform, Fourier transform, polynomial splines, discretization, shape matching, classification, sliced Wasserstein distance and barycenter.**

I. Introduction

In imaging, inverse problems, and geometric analysis the Radon transform holds a pivotal role by linking functions to their integrals over hyperplanes. Thereby, it encodes geometric information about a function that can be probed analytically and reconstructed algorithmically. Understanding its behavior for simple geometric bodies is thus both practically and theoretically important. In particular, the problem of computing the $(d-1)$-dimensional volume of the intersection between a hyperplane and an axis-aligned hypercube has a long history: classical closed-form results for special directions date back to Sommerfeld [68] and Pólya [59]. The latter article "Berechnung eines bestimmten Integrals" ("Computation of a certain integral") seems not to be well-known and apparently unknown to [8]. More recently, sharp and asymptotic bounds and extremal results were obtained in [7], [30], [40], [41], [50].

From the geometric-analysis side, a detailed understanding of cube–hyperplane intersections informs local extrema of section volumes [44], [53], [58] and combinatorial bounds on intersection vertices [15], [27], [49]. While prior works give sharp bounds and treat a number of special configurations (notably the main diagonal and central slices), explicit expressions for general directions and offsets, which are amenable to numerical evaluation, have been missing in a single, unified form, cf. Remark 7.

From the computational imaging side, there is an extensive literature on discrete Radon transforms, including Fourier-based approaches and numerical approximations for tomographic imaging [3]–[6], [12]. Applications of 3D Radon representations include rendering, shape matching and object retrieval [22], [23], [35], [45], [73], crystallography, medical/seismic imaging, biometric authentication [42], [46], [48], [66], and kernel-based methods [62], [63]. Practical algorithms related to voxel-friendly approximations, spline convolutions and nonuniform Fourier methods are discussed in [32], [37], [47], [52], [70]. However, many existing discrete schemes either reduce the 3D problem to repeated 2D line integrals, rely on Fourier interpolation that introduces artifacts for voxelized data, or employ Monte Carlo approximations [1], [20], [24] that are costly for high resolution. Further studies on the discretization are done via the operator norm [17] and adjoint mismatch [18].

Motivated by these limitations, we derive an explicit formula for the d-dimensional Radon transform of a hypercube. Building on this, we propose a voxel-aware discrete Radon transform that uses the exact cube–plane intersection area for summing voxel contributions. This discretization is both more accurate than simple center-projection binning, cf. [33] for $d = 2$, and substantially faster than high-quality Monte Carlo sampling.

Main contributions:

(i) An explicit closed form expression for the Radon transform of a hypercube in $\mathbb{R}^d$ (Theorem 1) in terms of a piecewise polynomial of degree $d-1$ that is $d-2$ times differentiable.
(ii) An explicit formula for a regularized Radon transform computing the d-dimensional volume of the intersection of the cube with a thin slab, which is the ε-neighborhood of a hyperplane (Corollary 6).
(iii) A voxel-aware discrete Radon transform utilizing the exact formulas from (i) and (ii).
(iv) Numerical simulations demonstrating the approach on 3D shape matching and classification tasks, sliced Wasserstein minimization and barycenters, and comparisons with Monte Carlo integration.

Organization: Section II provides explicit formulas for the Radon transform of a hypercube in $\mathbb{R}^d$, special cases $d \in \{2, 3\}$, and the cube–slab intersection volume. Two application areas are considered: Firstly, Section III introduces a voxel-aware discrete Radon transform, and performs shape matching via variants of the trace transform [22] and the normalized

R. Beinert, J. Bresch, and M. Quellmalz are with Institute of Mathematics, Technische Universität Berlin, Berlin, Germany. beinert,bresch,quellmalz@math.tu-berlin.de (all authors contributed equally).

Radon cumulative distribution transform [10]. Secondly, Section IV provides proof-of-concept experiments for the approximation of empirical measures and Wasserstein barycenters, both based on a free-support Radon model. Finally, Section V compares Monte Carlo sampling for approximating the 4D Radon transform with our explicit formula.

II. Multi-Dimensional Radon transform

A. Basic definitions

For an absolutely integrable function $f \in \mathrm{L}^1(\mathbb{R}^d)$ and a fixed direction $\boldsymbol{\theta} \in \mathbb{S}^{d-1} \coloneqq \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\| = 1\}$, the *Radon transform* is defined as

$$R_{\boldsymbol{\theta}}[f](t) \coloneqq \int_{H_{\boldsymbol{\theta}}(t)} f(\boldsymbol{x}) \,\mathrm{d}\sigma, \qquad t \in \mathbb{R}, \tag{1}$$

where σ denotes the surface measure on the *hyperplane*

$$H_{\boldsymbol{\theta}}(t) \coloneqq \{\boldsymbol{x} \in \mathbb{R}^d : \langle \boldsymbol{x}, \boldsymbol{\theta} \rangle = t\}. \tag{2}$$

The Radon transform is well-defined for almost all $t \in \mathbb{R}$ and obeys the symmetry $R_{-\boldsymbol{\theta}}[f](-t) = R_{\boldsymbol{\theta}}[f](t)$.

Furthermore, the *Fourier transform* is defined by

$$\mathcal{F}_d[f](\boldsymbol{v}) \coloneqq (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(\boldsymbol{x}) \,\mathrm{e}^{\mathrm{i}\langle \boldsymbol{x}, \boldsymbol{v} \rangle} \,\mathrm{d}\boldsymbol{x}, \qquad \boldsymbol{v} \in \mathbb{R}^d.$$

The Fourier and Radon transform are related via the Fourier slice theorem [54], which for all $s \in \mathbb{R}$ and $\boldsymbol{\theta} \in \mathbb{S}^{d-1}$ states

$$\mathcal{F}_1[R_{\boldsymbol{\theta}}[f]](s) = (2\pi)^{\frac{d-1}{2}} \mathcal{F}_d[f](s\boldsymbol{\theta}). \tag{3}$$

For $a > 0$, we denote the *indicator function* of the interval $(-a, a]$ by

$$\mathbf{1}_a(x) \coloneqq \begin{cases} 1, & \text{if } x \in (-a, a], \\ 0, & \text{otherwise,} \end{cases}$$

and the indicator function of the *hypercube*

$$(-\boldsymbol{a}, \boldsymbol{a}] \coloneqq \bigtimes_{i=1}^{d} (-a_i, a_i], \qquad \boldsymbol{a} \in \mathbb{R}^d_{>0}, \tag{4}$$

by

$$\mathbf{1}_{\boldsymbol{a}}(\boldsymbol{x}) \coloneqq \mathbf{1}_{a_1}(x_1) \cdots \mathbf{1}_{a_d}(x_d), \qquad \boldsymbol{x} \in \mathbb{R}^d.$$

Furthermore, we define the *positive part* of $x \in \mathbb{R}$ by

$$(x)_+ \coloneqq \begin{cases} x, & \text{if } x > 0, \\ 0, & \text{if } x \leqslant 0. \end{cases}$$

We frequently use its ℓ-th power $(x)_+^\ell$ for $\ell \in \mathbb{N}_0$, where we set $0^0 \coloneqq 0$. Moreover, we define the vector

$$\boldsymbol{x}^\circ \coloneqq \boldsymbol{x}|_{\operatorname{supp}(\boldsymbol{x})} \in \mathbb{R}^\ell, \qquad \boldsymbol{x} \in \mathbb{R}^d,$$

whose j-th entry is the j-th non-zero entry of $\boldsymbol{x}$ and where

$$\ell = \|\boldsymbol{x}\|_0 \coloneqq |\operatorname{supp}(\boldsymbol{x})|$$

denotes the number of non-zero entries of $\boldsymbol{x}$. The product of all entries of $\boldsymbol{x} \in \mathbb{R}^d$ is denoted by

$$P(\boldsymbol{x}) \coloneqq \prod_{j=1}^{d} x_j,$$

and the componentwise product of two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ by $\boldsymbol{x} \odot \boldsymbol{y} \in \mathbb{R}^d$.

The Fourier transform of $\mathbf{1}_{\boldsymbol{a}}$ with $\boldsymbol{a} \in \mathbb{R}^d_{>0}$ can be expressed as

$$\mathcal{F}_d[\mathbf{1}_{\boldsymbol{a}}](\boldsymbol{v}) = \left(\frac{2}{\pi}\right)^{d/2} \prod_{j=1}^{d} a_j \operatorname{sinc}(a_j v_j), \tag{5}$$

see [56, Ex. 2.3], with the *cardinal sine* function

$$\operatorname{sinc}(x) \coloneqq \begin{cases} \frac{\sin(x)}{x}, & x \in \mathbb{R}\backslash\{0\}, \\ 1, & x = 0. \end{cases}$$

B. The Radon transform of a hypercube

The next theorem provides a closed form for the Radon transform of the indicator function $\mathbf{1}_{\boldsymbol{a}}$ of the hypercube $(-\boldsymbol{a}, \boldsymbol{a}]$. This coincides with the surface area of $(-\boldsymbol{a}, \boldsymbol{a}] \cap H_{\boldsymbol{\theta}}(t)$.

Theorem 1. *Let $\boldsymbol{\theta} \in \mathbb{S}^{d-1}$, $t \in \mathbb{R}$, and $\boldsymbol{a} \in \mathbb{R}^d_{>0}$. Furthermore, let $\ell \coloneqq \|\boldsymbol{\theta}\|_0$. Then, for the Radon transformation of the hypercube, $A^{\boldsymbol{a}}_{\boldsymbol{\theta}}(t) \coloneqq \mathcal{R}_{\boldsymbol{\theta}}[\mathbf{1}_{\boldsymbol{a}}](t)$, holds*

$$A^{\boldsymbol{a}}_{\boldsymbol{\theta}}(t) = \frac{2^{d-\ell} P(\boldsymbol{a})}{P((\boldsymbol{a} \odot \boldsymbol{\theta})^\circ)\,(\ell-1)!} \times \sum_{\boldsymbol{k} \in \{-1,1\}^\ell} P(\boldsymbol{k}) \left(t + \langle \boldsymbol{k}, (\boldsymbol{a} \odot \boldsymbol{\theta})^\circ \rangle\right)_+^{\ell-1} \tag{6}$$

The proof requires the following lemma on convolutions. Related results about splines can be found in [36], [19, § 1]. The *convolution* of two functions $g, h \in \mathrm{L}^1(\mathbb{R})$ is defined as

$$(g * h)(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(y) h(y - x) \,\mathrm{d}y, \qquad x \in \mathbb{R}.$$

Iteratively, we denote the *k-fold convolution* of k functions $g_j \in \mathrm{L}^1(\mathbb{R})$ by

$$\overset{k}{\underset{j=1}{*}} g_j \coloneqq g_1 * \ldots * g_k. \tag{7}$$

The Fourier convolution theorem [56, Thm. 2.15] relates the Fourier transform and the convolution of two function by

$$\mathcal{F}_1[g * h] = \mathcal{F}_1[g] \cdot \mathcal{F}_1[h], \tag{8}$$

which directly extends to the k-fold convolution from (7).

Lemma 2. *Let $k \in \mathbb{N}$ and $\boldsymbol{b} \in \mathbb{R}^k_{>0}$. Then the k-fold convolution of indicator functions is*

$$\overset{k}{\underset{j=1}{*}} \mathbf{1}_{b_j}(t) = \frac{(2\pi)^{(1-k)/2}}{(k-1)!} \sum_{\boldsymbol{k} \in \{-1,1\}^k} P(\boldsymbol{k}) (t + \langle \boldsymbol{k}, \boldsymbol{b} \rangle)_+^{k-1}, \quad t \in \mathbb{R}. \tag{9}$$

Proof. For $k = 1$, we have

$$\mathbf{1}_{b_1}(t) = (t + b_1)_+^0 - (t - b_1)_+^0, \qquad t \in \mathbb{R}.$$

Now consider $\boldsymbol{b} \in \mathbb{R}^k_{\geqslant 0}$ and define for $\boldsymbol{c} \in \mathbb{R}^d$ the by the last component reduced vector $\boldsymbol{c}' \coloneqq (c_1, \dots, c_{k-1})^\top \in \mathbb{R}^{d-1}$. Then, we have for any $t \in \mathbb{R}$ that

$$\begin{aligned}
&\left(\mathbf{1}_{b_k} * \tfrac{(2\pi)^{(2-k)/2}}{(k-2)!} \sum_{\boldsymbol{k}' \in \{-1,1\}^{k-1}} P(\boldsymbol{k}')(t + \langle \boldsymbol{k}', \boldsymbol{b}' \rangle)_+^{k-2}\right)(t) \\
&= \tfrac{(2\pi)^{(1-k)/2}}{(k-2)!} \sum_{\boldsymbol{k}' \in \{-1,1\}^{k-1}} \int_{t-b_k}^{t+b_k} P(\boldsymbol{k}')(s + \langle \boldsymbol{k}', \boldsymbol{b}' \rangle)_+^{k-2} \,\mathrm{d}s \\
&= \tfrac{(2\pi)^{(1-k)/2}}{(k-1)!} \Bigg[\sum_{\boldsymbol{k}' \in \{-1,1\}^{k-1}} P(\boldsymbol{k}')(t + b_k + \langle \boldsymbol{k}', \boldsymbol{b}' \rangle)_+^{k-1} \\
&\qquad\qquad - P(\boldsymbol{k}')(t - b_k + \langle \boldsymbol{k}', \boldsymbol{b}' \rangle)_+^{k-1} \Bigg] \\
&= \tfrac{(2\pi)^{(1-k)/2}}{(k-1)!} \sum_{\boldsymbol{k} \in \{-1,1\}^k} P(\boldsymbol{k})(t + \langle \boldsymbol{k}, \boldsymbol{b} \rangle)_+^{k-1}.
\end{aligned}$$

□

Proof of Theorem 1. By (5) and the Fourier slice theorem (3), we obtain for almost all $s \in \mathbb{R}$ that

$$\begin{aligned}
\mathcal{F}_1[\mathcal{R}_{\boldsymbol{\theta}}[\mathbf{1}_{\boldsymbol{a}}]](s) &= (2\pi)^{\frac{d-1}{2}} \mathcal{F}_d[\mathbf{1}_{\boldsymbol{a}}](s\boldsymbol{\theta}) \\
&= \frac{2^d}{\sqrt{2\pi}} \prod_{j=1}^d a_j \operatorname{sinc}(a_j s \theta_j) \\
&= \frac{2^d}{\sqrt{2\pi} P(\boldsymbol{\theta}^\circ)} \prod_{\theta_j \neq 0} a_j \theta_j \operatorname{sinc}(a_j s \theta_j) \prod_{\theta_j = 0} a_j \\
&= \frac{2^{d-\ell} P(\boldsymbol{a})}{P((\boldsymbol{a} \odot \boldsymbol{\theta})^\circ)} (2\pi)^{\frac{\ell-1}{2}} \prod_{\theta_j \neq 0} \mathcal{F}_1[\mathbf{1}_{a_j \theta_j}](s) \\
&= \frac{2^{d-\ell} P(\boldsymbol{a})}{P((\boldsymbol{a} \odot \boldsymbol{\theta})^\circ)} (2\pi)^{\frac{\ell-1}{2}} \mathcal{F}_1\Big[\underset{j=1}{\overset{\ell}{\circledast}} \mathbf{1}_{(\boldsymbol{a} \odot \boldsymbol{\theta})_j^\circ} \Big](s)
\end{aligned}$$

by the Fourier convolution theorem (8). Lemma 2 yields a representative of the Radon transform of $\mathbf{1}_{\boldsymbol{a}}$ in the space of continuous functions by the right-hand-side of (9). This finishes the proof. □

The summation index $\boldsymbol{k} \in \{-1,1\}^\ell$ for $\ell = d$ in (6) corresponds to the 2^d vertices of the hypercube $(-\boldsymbol{a}, \boldsymbol{a}]$, which has the corners $\boldsymbol{a} \odot \boldsymbol{k}$, where $\boldsymbol{k} \in \{-1,1\}^d$.

The intersection of $(-\boldsymbol{a}, \boldsymbol{a}]$ with the hyperplane $H_{\boldsymbol{\theta}}(t)$ is a convex polytope with at most $\lfloor \frac{d+1}{2} \rfloor \binom{d}{\lfloor d/2 \rfloor}$ vertices, and this bound is attained [53]. These can be used for computing the area by triangulation, but their number grows asymptotically like $2^{d+1/2}$, whereas our summation has at most 2^d summands.

C. Explicit formulas from Theorem 1

First, we note the degenerate case $\ell = 1$, i.e. the Radon transform of the hypercube in direction of the axes $\boldsymbol{\theta} = \boldsymbol{e}_i$, which denotes the ith unit vector in $\mathbb{R}^d$:

$$A^{\boldsymbol{a}}_{\boldsymbol{e}_i}(t) = 2^{d-1} \mathbf{1}_{a_i}(t) \tfrac{P(\boldsymbol{a})}{a_i}, \quad t \in \mathbb{R}, \quad i \in [\![d]\!], \tag{10}$$

where $[\![n]\!] \coloneqq \{1, \dots, n\}$ for $n \in \mathbb{N}$.

Second, we state explicit versions of Theorem 1 for the important cases of dimension $d \in \{2, 3\}$. For a 2D rectangle, we have the following.

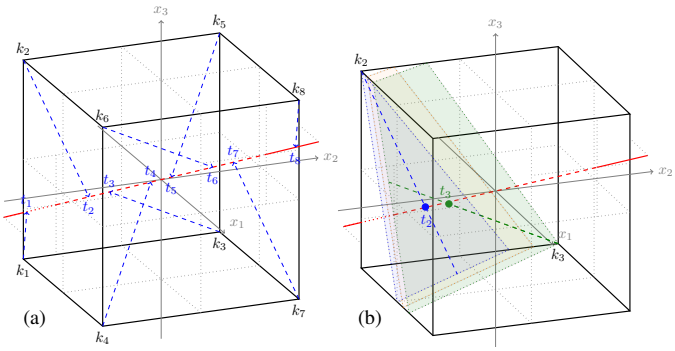

Fig. 1. (a): Projections t_i of the corners k_i of the cube to the red line in direction θ. (b): Planes $H_{\boldsymbol{\theta}}(t)$ for $t = t_2$ in blue, $t = t_3$ in green, and $t \in (t_2, t_3)$ in yellow, with three or five corners.

Corollary 3 (2D Radon transform)**.** *Let $\boldsymbol{\theta} \in \mathbb{S}^1$ with $\ell = \|\boldsymbol{\theta}\|_0$ and $\boldsymbol{a} \in \mathbb{R}^2_{>0}$. If $\ell = 1$, the area $A^{\boldsymbol{a}}_{\boldsymbol{\theta}}$ is given in* (10)*. If $\ell = 2$, we set $(t_i)_{i=1}^4$ as a nondecreasing rearrangement of $(\langle \boldsymbol{\theta}, \boldsymbol{a} \odot \boldsymbol{k} \rangle)_{\boldsymbol{k} \in \{-1,1\}^2}$, and it holds*

$$A^{\boldsymbol{a}}_{\boldsymbol{\theta}}(t) = \frac{1}{|\theta_1 \theta_2|} \begin{cases} 0, & t \leqslant t_1, \\ t - t_1, & t \in (t_1, t_2], \\ t_2 - t_1, & t \in (t_2, t_3], \\ t_4 - t, & t \in (t_3, t_4], \\ 0, & t \geqslant t_4. \end{cases}$$

In the 3D case, the intersection of a plane with a cube is a polygon with 3, 4, 5, or 6 edges (or in the degenerate case a line segment, singleton, or empty), see Fig. 1 (b).

Corollary 4 (3D Radon transform)**.** *Let $\boldsymbol{\theta} \in \mathbb{S}^2$ with $\ell = \|\boldsymbol{\theta}\|_0$ and $\boldsymbol{a} \in \mathbb{R}^3_{>0}$. If $\ell = 1$, the area $A^{\boldsymbol{a}}_{\boldsymbol{\theta}}(t)$ is given in* (10)*. If $\ell = 2$, we have $A^{\boldsymbol{a}}_{\boldsymbol{\theta}}(t) = 2a_j A^{\boldsymbol{a}^\circ}_{\boldsymbol{\theta}^\circ}(t)$ given in Corollary 3, where j is the index with $\theta_j = 0$. If $\ell = 3$, we set $(t_i)_{i=1}^8$ as a nondecreasing rearrangement of $(\langle \boldsymbol{\theta}, \boldsymbol{a} \odot \boldsymbol{k} \rangle)_{\boldsymbol{k} \in \{-1,1\}^3}$. Then it holds*

$$A^{\boldsymbol{a}}_{\boldsymbol{\theta}}(t) = \begin{cases} 0 & t \leqslant t_1, \\ \frac{(t-t_1)^2}{2|\theta_1\theta_2\theta_3|} & t \in (t_1, t_2], \\ \frac{(t_2-t_1)(2t-t_2-t_1)}{2|\theta_1\theta_2\theta_3|} & t \in (t_2, t_3], \\ \frac{t(t+2(-t_3+t_2-t_1))+t_3^2-t_2^2+t_1^2}{2|\theta_1\theta_2\theta_3|} & t \in (t_3, t_4], \\ \frac{2t(t_4-t_3+t_2-t_1)-t_4^2+t_3^2-t_2^2+t_1^2}{2|\theta_1\theta_2\theta_3|}, & t \in (t_4, 0]. \end{cases}$$

For $t > 0$, $A^{\boldsymbol{a}}_{\boldsymbol{\theta}}$ can be obtained from $A^{\boldsymbol{a}}_{\boldsymbol{\theta}}(t) = A^{\boldsymbol{a}}_{\boldsymbol{\theta}}(-t)$, and the symmetry $t_i = -t_{9-i} \leqslant 0$ for all $i \in [\![4]\!]$.

The points t_i are the projections of the vertices $\boldsymbol{a} \odot \boldsymbol{k}$ to the line in direction $\boldsymbol{\theta}$, see Fig. 1 (a). If some of the t_i coincide, the respective intervals in the last corollary are empty.

In general, the function $A^{\boldsymbol{a}}_{\boldsymbol{\theta}}$ is a piecewise polynomial in t of degree $\ell - 1$ and it is $\ell - 2$ times continuously differentiable if $\ell \geqslant 2$. Fig. 2 indicates a continuous behavior in 3D between the non-degenerate and the degenerate case, the latter coincides with the non-degenerate case from the 2D Radon transform, cf. Corollary 3.

Remark 5 (Optimal directions)**.** Fig. 3 visualizes the spherical function

$$\mathbb{S}^2 \ni \boldsymbol{\theta} \mapsto A^{\boldsymbol{a}}_{\boldsymbol{\theta}}(t), \quad \text{for some} \quad t \in \mathbb{R},$$

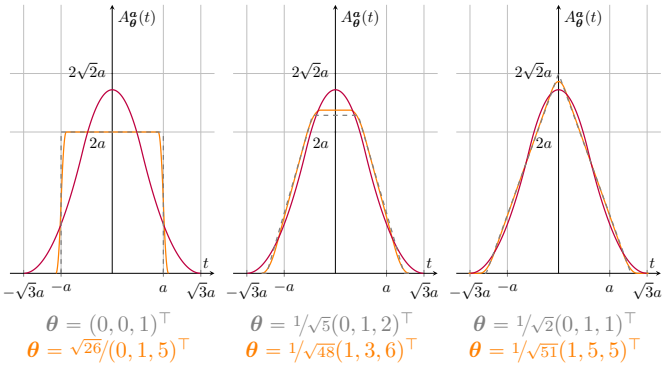

Fig. 2. The function $A^a_\theta(t)$ for $a = (a,a,a)^\top$, $a > 0$, is the 3D Radon transform of the cube. Functions depending on t for $\theta = 1/\sqrt{3}(1,1,1)^\top \in \mathbb{S}^2$ (purple) as reference, and varying $\theta \in \mathbb{S}^2$ in the degenerate (dashed gray) and general case (orange).

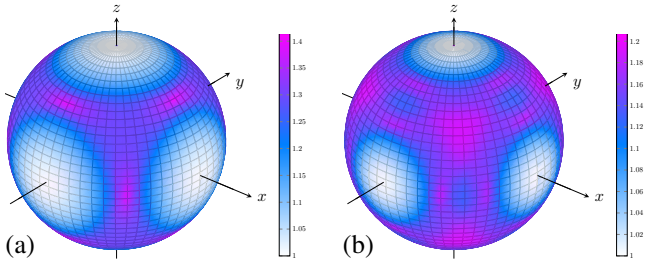

Fig. 3. The function $\mathbb{S}^2 \ni \theta \mapsto A^a_\theta(t)$ for $a = (1/2, 1/2, 1/2)^\top$, which is the 3D Radon transform of the cube. (a): $t = 0$. (b): $t = 1/7$.

for $a = 1/2(1,1,1)^\top \in \mathbb{R}^3$. In [40] it was shown that it has 12 global maxima $1/\sqrt{3}(\pm 1, \pm 1, \pm 1)^\top$ for all $1/2\sqrt{3} < t \leqslant \sqrt{3}/2$, which remain at least local maxima if $t < 1/2\sqrt{3}$, see Fig. 3 (b). Moreover, $\theta \mapsto A^a_\theta(0)$ has 12 global maxima [50], namely $\theta \in 1/\sqrt{2}\{(\pm 1, \pm 1, 0)^\top, (\pm 1, 0, \pm 1)^\top, (0, \pm 1, \pm 1)^\top\}$, and 6 global minima $\theta \in \{\pm \mathrm{e}_i : i \in [\![3]\!]\}$, see Fig. 3 (a).

D. Volume of a slab

Closely related to Theorem 1, we replace the hyperplane by a d-dimensional slab. Denote by $V^a_\theta(t_1, t_2)$ the volume of the intersection of the hypercube $(-a, a]$ from (4) with the *slab*

$$S_\theta(t_1, t_2) := \{x \in \mathbb{R}^d : \langle x, \theta \rangle \in [t_1, t_2]\} \tag{11}$$

for $\theta \in \mathbb{S}^{d-1}$ and $t_1, t_2 \in \mathbb{R}$.

Corollary 6. *For $a \in \mathbb{R}^d_{>0}$, $\theta \in \mathbb{S}^{d-1}$, and $t_1, t_2 \in \mathbb{R}$ with $t_1 < t_2$, it holds with the notation from Theorem 1 that*

$$V^a_\theta(t_1, t_2) = \frac{2^{d-\ell} P(a)}{P((a \odot \theta)^\circ)\, \ell!} \sum_{k \in \{-1,1\}^\ell} P(k) \times \Big(\big(t_2 + \langle k, (a \odot \theta)^\circ \rangle\big)_+^\ell - \big(t_1 + \langle k, (a \odot \theta)^\circ \rangle\big)_+^\ell \Big). \tag{12}$$

Proof. The slab $S_\theta(t_1, t_2)$ from (11) is the union of hyperplanes (2), namely

$$S_\theta(t_1, t_2) = \bigcup_{s \in [t_1, t_2]} H_\theta(s).$$

Since $\|\theta\| = 1$, we have

$$V^a_\theta(t_1, t_2) = \int_{(-a,a] \cap S_\theta(t_1,t_2)} 1 \,\mathrm{d}x = \int_{t_1}^{t_2} A^a_\theta(s) \,\mathrm{d}s.$$

Integrating (6) for t, we obtain (12). □

The volume converges to the area A^a_θ if the width of the slab approaches zero, in particular it holds

$$\lim_{\varepsilon \searrow 0} \tfrac{1}{2\varepsilon} V^a_\theta(t - \varepsilon, t + \varepsilon) = A^a_\theta(t), \qquad \forall t \in \mathbb{R}. \tag{13}$$

Remark 7 (Related results)**.** Explicit volume formulas have been stated in different forms in [8], [57], [59]. However, this literature considered the special cases of (12), where either $\theta = e/\sqrt{d} \in \mathbb{S}^{d-1}$ and $a \in \mathbb{R}^d_{>0}$ is arbitrary [59], or $\theta \in \mathbb{S}^{d-1}$ is arbitrary and $a = e$ is fixed [8], [41], [57], where $e := (1, \ldots, 1)^\top \in \mathbb{R}^d$. These two cases are related via a coordinate transform: Let $a \in \mathbb{R}^d_{>0}$ and note that (12) remains valid for $\theta \in \mathbb{R}^d$. For $x \in \mathbb{R}^d$, we perform the coordinate transform $x = a \odot y$ with $\mathrm{d}x = P(a)\,\mathrm{d}a$. Then $x \in (-a, a]$ if and only if $y \in (-e, e]$ and

$$x \in H_e(t) \Leftrightarrow \langle e, x \rangle = t \Leftrightarrow \langle a, y \rangle = t \Leftrightarrow y \in H_a(t),$$

for any $t \in \mathbb{R}$. Hence, we have

$$V^a_e(t_1, t_2) = \int_{(-a,a] \cap S_e(t_1,t_2)} 1 \,\mathrm{d}x = \int_{(-e,e] \cap S_a(t_1,t_2)} 1 \,\mathrm{d}y = P(a) \cdot V^e_a(t_1, t_2).$$

and therewith, by (13), it follows

$$A^a_e(t) = P(a) A^e_a(t), \qquad t \in \mathbb{R}.$$

III. Discrete Radon Transform and Shape Matching

In this section, we substantiate the theoretical findings from Theorem 1 with numerical applications. Therefore, we consider the task of object recognition and classification in 3D utilizing the discrete Radon transform of voxelized images outlined in Section III-B. First, in Section III-C, we perform feature extraction utilizing the Radon shape matching [22] based on the *trace transform* [34]. Second, in Section III-D, we study the recently proposed *Normalized Radon Cumulative Distribution Transform* (NR-CDT) [9]–[11]. Finally, we consider the classification of affinely transformed 3D objects in Section III-E.

A. Datasets

The ModelNet10 dataset[1] [72] contains ten classes of 3D objects: bathtubs, beds, chairs, desks, dressers, monitors, nightstands, sofas, tables, and toilets, see Fig. 4 and 5.

We consider a subset from ModelNet10 comprising the first ten samples from each of the classes. Initially, we do not rely on affine object classes for which a theoretical (linear) separability result holds. Therefore, the considered experiments may validate robustness of the transforms. Afterwards, we consider classes of affinely transformed objects in Section III-E.

[1]Online available: https://modelnet.cs.princeton.edu

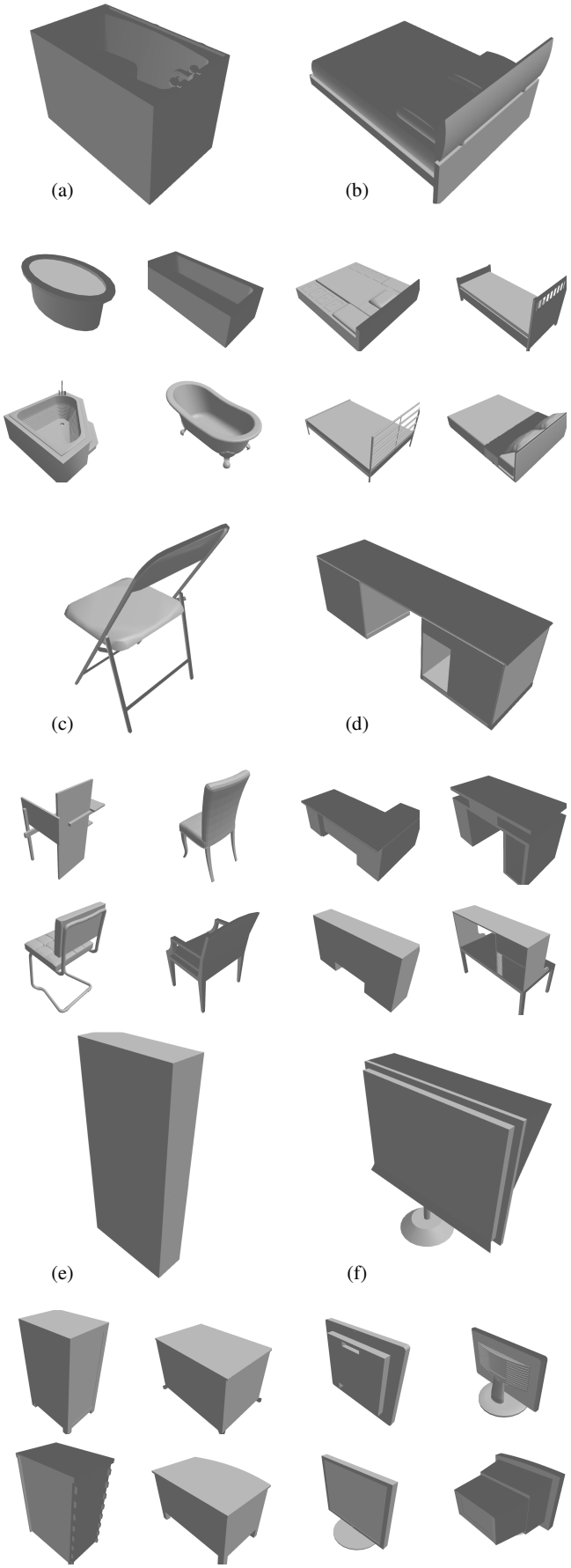

Fig. 4. Selection of the first to fives samples of the ModelNet10 dataset for (a) bathtubs, (b) beds, (c) chairs, (d) desks, (e) dresser, and (f) monitors.

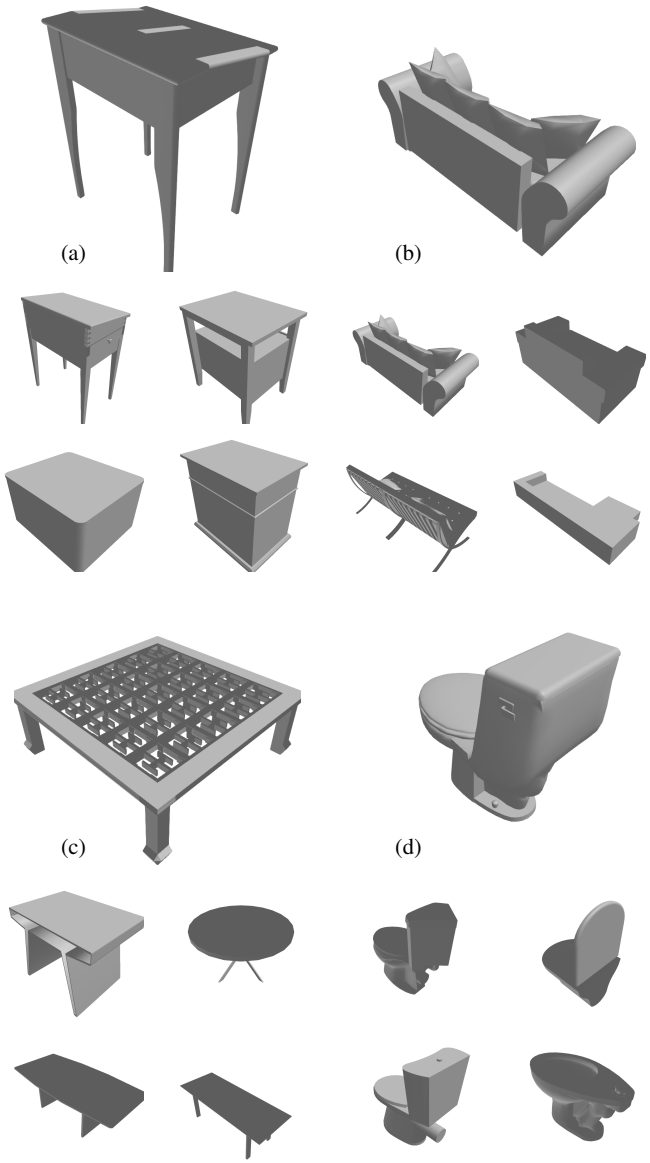

Fig. 5. Selection of the first to fives samples of the ModelNet10 dataset for (a) nightstands, (b) sofas, (c) tables, and (d) toilets.

B. Discretization of the Radon transform

For a numerical realization of the Radon transform on the voxelized objects, we need a discrete model. For $N \in \mathbb{N}$, we define the symmetric index set

$$\mathcal{I}_N := \{-N+1/2, -N+3/2, \dots, N-1/2\}$$

and the regular grid $\mathcal{I}_N^d$ as d-fold Cartesian product. We consider a d-dimensional image $F \in \mathbb{R}^{\mathcal{I}_N^d}$ with N voxels in each dimension. Each voxel $\boldsymbol{n} = (n_1, \dots, n_d) \in \mathcal{I}_N^d$ is imagined as cube with side length (i.e. voxel size) $s > 0$ centered at $s \cdot \boldsymbol{n} \in \mathbb{R}^d$, i.e., the cube $(s(\boldsymbol{n} - 1/2), s(\boldsymbol{n} + 1/2)]$. Then F corresponds to the piecewice constant function

$$f_F(\boldsymbol{x}) := \sum_{\boldsymbol{n} \in \mathcal{I}_N^d} F(\boldsymbol{n})\, \mathbf{1}_{(s(\boldsymbol{n}-1/2), s(\boldsymbol{n}+1/2)]}(\boldsymbol{x}), \qquad \boldsymbol{x} \in \mathbb{R}^d.$$

The cubes touch but do not overlap.

Let $\boldsymbol{\theta} \in \mathbb{S}^{d-1}$ and $t \in \mathbb{R}$. By linearity, we have

$$\mathcal{R}_{\boldsymbol{\theta}}[f_F](t) = \sum_{\boldsymbol{n} \in \mathcal{I}_N^d} F(\boldsymbol{n})\, \mathcal{R}_{\boldsymbol{\theta}}\big[\mathbf{1}_{(s(\boldsymbol{n}-1/2), s(\boldsymbol{n}+1/2))}\big](t).$$

The shift identity for the Radon transform,

$$\mathcal{R}_{\boldsymbol{\theta}}[f(\cdot - \boldsymbol{y})](t) = \mathcal{R}_{\boldsymbol{\theta}}[f](t - \langle \boldsymbol{\theta}, \boldsymbol{y} \rangle), \qquad \boldsymbol{y} \in \mathbb{R}^d,$$

follows directly from the definition (1). Together with the observation that $\mathbf{1}_{(s(\boldsymbol{n}-\frac{1}{2}), s(\boldsymbol{n}+\frac{1}{2}))}(\boldsymbol{x}) = \mathbf{1}_{s/2\boldsymbol{e}}(\boldsymbol{x} - s \cdot \boldsymbol{n})$, we obtain that

$$\mathcal{R}_{\boldsymbol{\theta}}[f_F](t) = \sum_{\boldsymbol{n} \in \mathcal{I}_N^d} F(\boldsymbol{n})\, \mathcal{R}_{\boldsymbol{\theta}}\big[\mathbf{1}_{s/2\boldsymbol{e}}\big](t - s\langle \boldsymbol{n}, \boldsymbol{\theta} \rangle).$$

Then we have

$$\mathcal{R}_{\boldsymbol{\theta}}[f_F](t) = \sum_{\boldsymbol{n} \in \mathcal{I}_N^d} F(\boldsymbol{n})\, A_{\boldsymbol{\theta}}^{s/2\boldsymbol{e}}(t - s\langle \boldsymbol{n}, \boldsymbol{\theta} \rangle) \tag{14}$$

with $A_{\boldsymbol{\theta}}^{\boldsymbol{a}}$ given in (6). Since the box $s/2(-\boldsymbol{e}, \boldsymbol{e}]$ has a diagonal of $s\sqrt{d}/2$, the summands in (14) vanish if $|t - s\langle n, \boldsymbol{\theta} \rangle| > s\sqrt{d}/2$. For improved stability, we replace $A_{\boldsymbol{\theta}}^{\boldsymbol{a}}(t)$ by its regularized variant $(2\varepsilon)^{-1} V_{\boldsymbol{\theta}}^{\boldsymbol{a}}(t - \varepsilon, t + \varepsilon)$, see (12).

Remark 8 (Existing approaches)**.** Besides our approach, there are other ways to discretize the Radon transform. Most work covers only the case $d = 2$. A simple, well-known approach [14] projects the center of a voxel to the line $\boldsymbol{\theta}^\perp$, and then uses binning in t. For general d, we set b as bin size in t, and approximate

$$\mathcal{R}_{\boldsymbol{\theta}}[f_F](t) \approx \sum_{\substack{\boldsymbol{n} \in \mathcal{I}_N^d \\ |\langle \boldsymbol{n}, \boldsymbol{\theta} \rangle - t| < b}} F(\boldsymbol{n}).$$

This approach is usually implemented by first computing the projections $\langle \boldsymbol{n}, \boldsymbol{\theta} \rangle$ for all voxels $\boldsymbol{n}$ and then binning them to a grid in t. Therefore, this approach is generally faster than ours, but less accurate. It can be improved by splitting each voxel into subvoxels. The convergence of different approaches was analyzed in [16], [33].

Another approach, known as slant stacking [69] in 2D discretizes the line integral by a quadrature on the line $\boldsymbol{\theta}^\perp$, where the function values of f are obtained via interpolation of F; taking f_F here corresponds to a the nearest neighbor interpolation. Discrete convolutions in combination with data interpolation was utilized for the 2D Radon transform in [2]. Monte Carlo approximations were discussed in [1], [20], [24].

Unfortunately, there are fewer readily available implementations of the 3D Radon transform. Although some packages like `RadonKA`[2] [70] provide "3D Radon transforms", these are vectorized versions of the 2D Radon transform (i.e. integrating along all lines perpendicular to the third coordinate) rather than integrating along planes. In [69], closed forms for few simple objects, e.g. Gaußian bells and ellipsoids in 3D are provided.

C. Radon shape matching via feature extraction [22]

We first revise the construction of the feature extractors, and then provide the numerical results for the shape matching.

[2]Julia package `RadonKA` https://github.com/roflmaostc/RadonKA.jl

TABLE I
SUMMARY OF THE COMBINATION FOR THE EXTRACTORS FROM [22].

$\pi(t, \vartheta_1, \vartheta_2)$	$(t, \vartheta_1, \vartheta_2)$	$(t, \vartheta_2, \vartheta_1)$	$(\vartheta_2, t, \vartheta_1)$	$(\vartheta_1, t, \vartheta_2)$
$F_{i_1}, F_{i_2}, F_{i_3}$	F_1, F_1, F_1 F_1, F_1, F_2 F_1, F_1, F_4 F_1, F_2, F_1 F_1, F_3, F_1 F_1, F_3, F_4 F_1, F_4, F_1 F_1, F_4, F_2 F_2, F_1, F_1 F_2, F_1, F_4 F_2, F_3, F_4 F_2, F_4, F_1 F_2, F_4, F_2 F_3, F_1, F_2 F_3, F_1, F_4 F_3, F_2, F_1 F_3, F_4, F_1	F_1, F_1, F_2 F_1, F_1, F_3 F_1, F_1, F_4 F_1, F_2, F_1 F_1, F_2, F_4 F_1, F_4, F_1 F_1, F_4, F_3 F_2, F_1, F_1 F_2, F_1, F_3 F_2, F_1, F_4 F_2, F_2, F_2 F_2, F_4, F_1 F_2, F_4, F_3 F_3, F_1, F_1 F_3, F_2, F_4 F_3, F_4, F_4	F_1, F_1, F_4 F_1, F_2, F_1 F_1, F_2, F_4 F_2, F_1, F_1 F_2, F_2, F_1 F_3, F_1, F_1 F_3, F_1, F_2 F_3, F_1, F_4 F_3, F_2, F_4 F_3, F_3, F_4	F_1, F_1, F_4 F_1, F_2, F_1 F_1, F_3, F_1 F_1, F_3, F_2 F_2, F_1, F_1 F_2, F_2, F_1 F_3, F_1, F_1 F_3, F_2, F_1

1) Preprocessing: According to [22], the following preprocessing steps on the reduced ModelNet10 dataset from Section III-A are applied: (i) centering the objects by the means of mass points, i.e. centroid of the voxels, (ii) *Principal Component Analysis* (PCA) alignment using the covariance/inertia matrix for rotating towards the principal axes, (iii) scaling into the unit box, and (iv) normalizing the mass after revoxelizing the objects.

2) Construction of the features: Following the definition of the feature extractors from [22, Eq. (10–13)], the sphere is discretized by a uniform grid on the spherical coordinates given by

$$\boldsymbol{\theta}(\boldsymbol{\vartheta}) \coloneqq \begin{bmatrix} \sin\vartheta_1 \sin\vartheta_2 \\ \cos\vartheta_1 \sin\vartheta_2 \\ \cos\vartheta_2 \end{bmatrix}, \quad \begin{cases} \boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2), \\ \vartheta_1 \in [0, 2\pi) \\ \vartheta_2 \in [0, \pi]. \end{cases} \tag{15}$$

The extractors $F_i \colon \mathbb{R}^d \to \mathbb{R}$ given by

$$F_1(g) \coloneqq \max_{i=1\ldots,D} g_i, \qquad F_2(g) \coloneqq \sum_{i=1}^{D-1} \tfrac{1}{2}|g_{i+1} - g_i|,$$

$$F_3(g) \coloneqq \sum_{i=1}^{D} g_i, \qquad F_4(g) \coloneqq F_1(g) + F_1(-g),$$

are applied recursively on

$$\mathcal{R}[f_F] : \mathbb{R}^{D_1} \times [0, 2\pi)^{D_2} \times [0, \pi]^{D_3} \to \mathbb{R}, \\ (t, \boldsymbol{\theta}(\vartheta_1, \vartheta_2)) \mapsto \mathcal{R}_{\boldsymbol{\theta}(\vartheta_1, \vartheta_2)}[f](t).$$

The ordering $(F_{i_1}, F_{i_2}, F_{i_3})$ given by the permutation $\pi = (\pi_1, \pi_2, \pi_3) \in S_3(\{t, \vartheta_1, \vartheta_2\})$ of the 51 selected extractors is summarized in Table I [22, Tab. 1]. We obtain the extractors

$$D_{i_1, i_2, i_3} \coloneqq F_{i_3}(F_{i_2}(F_{i_1}\mathcal{R}[f_F]\pi(\cdot_1, \cdot_2, \cdot_3))\pi(\cdot_2, \cdot_3))(\cdot_3)$$

3) Numerical results: After preprocessing the (reduced) ModelNet10 dataset with $64 \times 64 \times 64$ voxels, our discrete Radon transform $\mathcal{R}_{\boldsymbol{\theta}(\boldsymbol{\vartheta})}[f_F](t)$ is applied with different resolutions: equispaced radii $t \in [-\sqrt{3}/2, \sqrt{3}/2]$ and an equispaced grid in spherical coordinates $\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2)$, see (15).

Instead of the similarity measure from [22], we perform a 1-*nearest neighbor* (1-NN) classification on the extracted object features with different sizes of reference (sub)datasets: for N

TABLE II
ACCURACY RESULTS FOR THE RADON SHAPE MATCHING [22] WITH 1-NN CLASSIFICATION ON THE REDUCED MODELNET10 DATASET ($N = 3$): BATHTUBS, CHAIRS, AND DRESSERS.

radii	angles	R	$\|\cdot\|_1$	$\|\cdot\|_2$	$\|\cdot\|_\infty$
512	$(20, 16)$	1	0.451 ± 0.156	0.425 ± 0.136	0.412 ± 0.163
		3	0.480 ± 0.127	0.487 ± 0.121	0.457 ± 0.115
		5	0.633 ± 0.100	0.588 ± 0.104	0.544 ± 0.097
	$(30, 21)$	1	0.491 ± 0.167	0.484 ± 0.133	0.514 ± 0.176
		3	0.531 ± 0.153	0.507 ± 0.126	0.553 ± 0.155
		5	0.690 ± 0.107	0.635 ± 0.102	0.571 ± 0.098
2048	$(20, 16)$	1	0.502 ± 0.120	0.432 ± 0.145	0.333 ± 0.156
		3	0.555 ± 0.095	0.461 ± 0.111	0.361 ± 0.111
		5	0.687 ± 0.085	0.556 ± 0.067	0.574 ± 0.101
	$(30, 21)$	1	0.513 ± 0.111	0.415 ± 0.164	0.425 ± 0.178
		3	0.585 ± 0.089	0.451 ± 0.137	0.466 ± 0.153
		5	0.712 ± 0.069	0.680 ± 0.116	0.666 ± 0.124

TABLE III
ACCURACY RESULTS FOR THE RADON SHAPE MATCHING [22] WITH 1-NN CLASSIFICATION ON THE ENTIRE MODELNET10 DATASET ($N = 10$).

radii	angles	R	$\|\cdot\|_1$	$\|\cdot\|_2$	$\|\cdot\|_\infty$
2048	$(30, 21)$	1	0.240 ± 0.044	0.222 ± 0.045	0.215 ± 0.054
		3	0.312 ± 0.048	0.312 ± 0.054	0.277 ± 0.038
		5	0.363 ± 0.044	0.351 ± 0.060	0.333 ± 0.051

classes, the $N \times 10$ dataset is randomly divided into $N \times R$ train and $N \times (10-R)$ test samples, and the data extractors are afterwards used to classify the test data. Therefore each test sample is assigned to the reference sample with the smallest distance of the extractors. Here, we consider the Manhattan norm $\|\cdot\|_1$, the Euclidean norm $\|\cdot\|_2$, and the Chebyshev norm $\|\cdot\|_\infty$. The calculations are repeated 20 times with different random splits. The mean and standard deviation are reported in Table II for the reduced ($N = 3$) and in Table III for the entire ModelNet10 dataset ($N = 10$).

We observe the best classification result of 71%, cf. Table II, which is significantly better than guessing (33%) on the 3-class problem with bathtubs, chairs, and dressers. For the 10-class problem in Table III, we observe the best classification accuray of 36% with the Manhattan norm. The respective distance map is provided in Fig. 6 (left), from which the class structure is just slightly visible. The confusion map in Fig. 6 (right) indicates uncertainties between several classes; the maximal rate of correct assignments is 64%.

The similarity measure used in [22] provides even worse results. We were not able to reproduce the exact numbers from [22] due to unavailable code, the lack of documentation of the preprocessing, and the ambiguity how the Radon transform is calculated. However, compared to their dataset, our choice is much more difficult with more varieties in the shapes within each class.

D. Classification via NR-CDT [10]

The Radon transform can be utilized for the classification of images using the so-called R-CDT approach [39]. Its extension, the NR-CDT [9], [11], is invariant to affine transformations, and was applied to pattern recognition [10]. Differently, we apply the theory on voxels and not on meshes (i.e. point clouds), which makes the computation of the 3D Radon transform crucial. We first revise the construction of

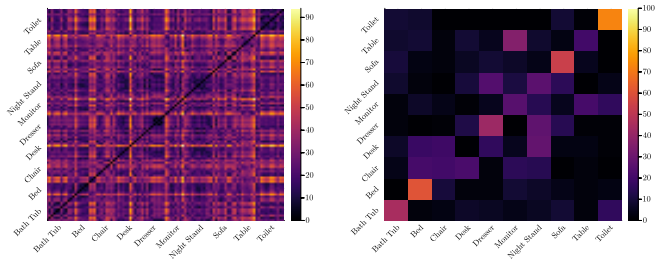

Fig. 6. Distance map (left) and confusion map (in %, right) for the Radon shape matching of the entire ModelNet10 dataset with respect to the Manhattan norm $\|\cdot\|_1$, i.e. the best performing distance in Table III.

the NR-CDT for functions, and then provide the numerical results on the ModelNet10 dataset for comparison with the feature extractors from Section III-C.

1) Normalized Radon cumulative distribution transform: Let $f \in \mathrm{L}^1(\mathbb{R}^d)$ be nonnegative and $\int_{\mathbb{R}^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 1$. Following [10], we define the *Radon cumulative distribution transform* (Radon-CDT) of f by

$$\widehat{\mathcal{R}_{\boldsymbol{\theta}}[f]}\colon \mathbb{R} \to \mathbb{R}, \quad \xi \mapsto \inf\Big\{ s \in \mathbb{R} \;\Big|\; \int_{-\infty}^{s} \mathcal{R}_{\boldsymbol{\theta}}[f](t)\,\mathrm{d}t > \xi \Big\}.$$

This is the quantile function, i.e., the generalized inverse of the cumulative distribution function of $\mathcal{R}_{\boldsymbol{\theta}}[f]$, see [9], [10]. We define the *mean* and *standard derivation* of a function $g \in \mathrm{L}^\infty([0,1])$ by

$$\mathrm{mean}(g) \coloneqq \int_0^1 g(s)\,\mathrm{d}s, \quad \mathrm{std}(g)^2 \coloneqq \int_0^1 (g(s) - \mathrm{mean}(g))^2\,\mathrm{d}s.$$

Following [10], the *normalized Radon-CDT* (NR-CDT)

$$\mathcal{N}_{\boldsymbol{\theta}}[f]\colon \mathbb{R} \to \mathbb{R},\ t \mapsto \frac{\widehat{\mathcal{R}_{\boldsymbol{\theta}}[f]}(t) - \mathrm{mean}(\widehat{\mathcal{R}_{\boldsymbol{\theta}}[f]})}{\mathrm{std}(\widehat{\mathcal{R}_{\boldsymbol{\theta}}[f]})},$$

is well defined if $\mathrm{supp}(f) \subset \mathbb{R}^d$ is compact and $\dim(\mathrm{supp}(f)) > d - 1$, cf. [10, Prop. 4, Lem. 1].

Finally, we define the *max-normalized* R-CDT ($_\mathrm{m}$NR-CDT)

$$\mathcal{N}_\mathrm{m}[f]\colon \mathbb{R} \to \mathbb{R}, \quad \xi \mapsto \sup_{\boldsymbol{\theta} \in \mathbb{S}^{d-1}} \mathcal{N}_{\boldsymbol{\theta}}[f](\xi).$$

The $_\mathrm{m}$NR-CDT is invariant under affine transformations, see [9, Thm. 1]. More specifically, consider for a function f its affine transformation

$$f_{\boldsymbol{A},\boldsymbol{y}} : \mathbb{R}^d \to \mathbb{R},\ \boldsymbol{x} \mapsto f(\boldsymbol{A}^{-1}(\boldsymbol{x} - \boldsymbol{y})), \tag{16}$$

for $\boldsymbol{A} \in \mathrm{GL}(d)$ and $\boldsymbol{y} \in \mathbb{R}^d$, then it holds

$$\mathcal{N}_\mathrm{m}[f_{\boldsymbol{A},\boldsymbol{y}}] = \mathcal{N}_\mathrm{m}[f]. \tag{17}$$

2) Numerical results: We apply the $_\mathrm{m}$NR-CDT on both datasets, i.e. the reduced (bathtubs, chairs, and dressers) and entire ModelNet10 dataset, where each class consists of ten samples. The discretization uses three parameters: i) the number of evaluation points for the CDT; ii) the equispaced radii as in Section III-C3; iii) the angles $\boldsymbol{\theta}$, which are not restricted to grids unlike in Section III-C. There are many notions of "almost equispaced" points $\boldsymbol{\theta} \in \mathbb{S}^2$ such as spherical designs [28], [71] or quasi-Monte Carlo designs [31], [64]. We chose

TABLE IV
ACCURACY RESULTS FOR THE $_{\mathrm{M}}$NR-CDT [10] WITH THE 1-NN CLASSIFICATION ON THE REDUCED MODELNET10 DATASET ($N = 3$): BATHTUBS, CHAIRS, AND DRESSERS.

radii	angles	R	$\|\cdot\|_1$	$\|\cdot\|_2$	$\|\cdot\|_\infty$
64	64	1	0.853 ± 0.089	0.826 ± 0.080	0.703 ± 0.085
		3	0.890 ± 0.061	0.852 ± 0.075	0.776 ± 0.087
		5	0.922 ± 0.067	0.900 ± 0.063	0.783 ± 0.110
128	128	1	0.901 ± 0.040	0.855 ± 0.073	0.757 ± 0.078
		3	0.904 ± 0.048	0.866 ± 0.059	0.759 ± 0.094
		5	0.900 ± 0.066	0.869 ± 0.100	0.820 ± 0.094
256	256	1	0.892 ± 0.031	0.835 ± 0.100	0.696 ± 0.083
		3	0.914 ± 0.093	0.869 ± 0.068	0.785 ± 0.079
		5	0.933 ± 0.094	0.896 ± 0.070	0.836 ± 0.076

TABLE V
ACCURACY RESULTS FOR THE $_{\mathrm{M}}$NR-CDT [10] WITH THE 1-NN CLASSIFICATION ON THE ENTIRE MODELNET10 DATASET ($N = 10$).

radii	angles	R	$\|\cdot\|_1$	$\|\cdot\|_2$	$\|\cdot\|_\infty$
256	256	1	0.307 ± 0.040	0.312 ± 0.051	0.245 ± 0.037
		3	0.375 ± 0.041	0.401 ± 0.052	0.312 ± 0.046
		5	0.419 ± 0.047	0.472 ± 0.042	0.388 ± 0.060

the Fibonacci points [26], [29] given in spherical coordinates (15) by

$$\boldsymbol{\vartheta} = \begin{bmatrix} \vartheta_1 \\ \vartheta_2 \end{bmatrix} \in \left\{ \begin{bmatrix} i\varphi \\ \arccos\left(1 - \frac{2(i-1)+1}{n}\right) \end{bmatrix} \,\middle|\, i \in [\![n]\!] \right\}, \quad (18)$$

with the golden angle $\varphi := \pi(3 - \sqrt{5})$. The resulting $_{\mathrm{m}}$NR-CDTs are compared with different norms, cf. Section III-C3.

For the shape matching on the ModelNet10 dataset, we perform a 1-NN classification with splitting sizes $R \in \{1, 3, 5\}$ as in Section III-C3. The accuracies (mean±std.) for different discretizations of the sphere $\mathbb{S}^2$ and the radii based on 20 runs of the 1-NN are reported in Table IV for the reduced dataset ($N = 3$) and in Table V for the entire dataset ($N = 10$).

We observe accuracies up to 93% for the shape matching, see Table IV, in the reduced classes (bathtubs, chairs, and dressers), which is a significant improvement to the Radon shape matching from Section III-C, cf. Table II. Especially, in the $_{\mathrm{m}}$NR-CDT space, the chairs are perfectly distinguished from both other classes.

For the entire dataset, we obtain an accuracy up to 47%, which is significantly better than random guessing (10%), as well as the accuracies from Section III-C, cf. Table III. The separability of the classes in the $_{\mathrm{m}}$NR-CDT-space is clearly visible from the distance map in Fig. 7 (left). Moreover, in the confusion map, see Fig. 7 (right), there are few classes (nightstands, tables, sofas, and bathtubs) that cannot be assigned well. However, the diagonal is more dominant than in Fig. 6 (right).

E. Classification of affinely transformed objects

Contrary to the previous subsections, we consider affienly transformed objects in 3D. Here, we take the first sample from each class of the ModelNet10 dataset, cf. Fig. 4 and 5, as a template, and generate ten objects by applying random affine transformations (anisotropic scaling, shearing, shifting, rotation) to the template, see (16). The numerical studies (1-NN) from Sections III-C3 and III-D2 are repeated for the densest discretization. We obtain an accuracy for the Radon shape matching [22] of up to 42%, which is outperformed by the $_{\mathrm{m}}$NR-CDT embedding [10] with a nearly perfect classification result of 99%, as theoretically expected, cf. (17).

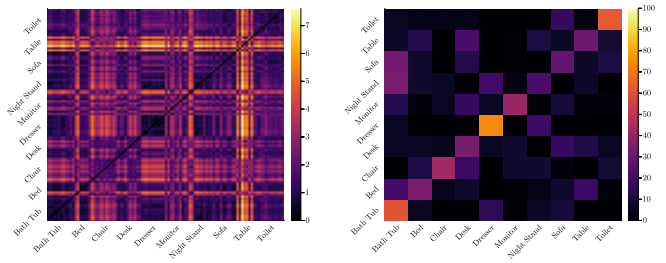

Fig. 7. Distance map (left) and confusion map (in %, right) of the entire ModelNet10 dataset in the $_{\mathrm{m}}$NR-CDT-space with respect to the Euclidean norm $\|\cdot\|_2$, i.e. the distance with the best outcome of the accuracies from Table V.

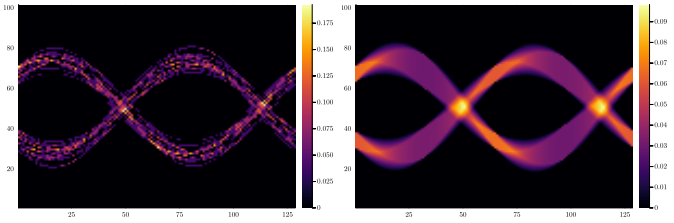

Fig. 8. 2D Radon transforms of the empirical (target) measure Δ_n (left) and the uniform mixture Υ_k solving (23), corresponding to Fig. 9 (top).

IV. APPROXIMATION AND MEANS OF EMPIRICAL DATA

In contrast to Section III, which relies on a voxel-based data structure, here we are interested in a free-support setting for measures. Firstly, we consider clustering in 2D and 3D, more precisely, we aim to fit a given empirical measure by a mixture of measures on cubes via the sliced Wasserstein distance, see Section IV-A. Secondly, we compute sliced Wasserstein barycenters, see Section IV-B.

A. Optimal histogram via sliced Wasserstein distance

Approximating an empirical measure $\Delta_n := {}^1\!/_n \sum_{i=1}^n \delta_{\boldsymbol{x}_i}$, where $\delta_{\boldsymbol{x}}$ denotes the Dirac measure at a point $\boldsymbol{x} \in \mathbb{X} \subset \mathbb{R}^d$, by another distribution is a important task [13], [38]. We are interested in linear combinations

$$\Upsilon_k := \sum_{j=1}^{k} \gamma_j u_j, \quad (19)$$

where u_j is the uniform measure on a cube

$$(\boldsymbol{c}_j - \boldsymbol{w}_j, \boldsymbol{c}_j + \boldsymbol{w}_j]$$

with center $\boldsymbol{c}_j \in \mathbb{X}$ and width $\boldsymbol{w}_j \in \mathbb{X}$, and $\gamma_j := {}^{u_j(\mathbb{X})}\!/_{\sum_{\ell=1}^k u_\ell(\mathbb{X})}$. Note that the ansatz of minimizing the Kullback–Leibler divergence or any other Csiszár divergence to the given measure is not defined in this case [21].

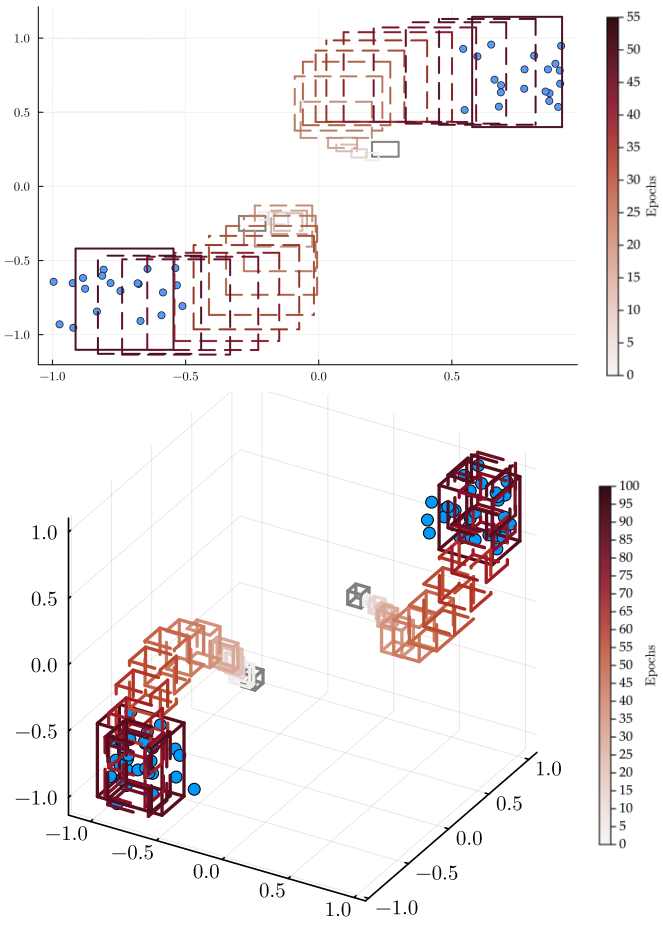

Fig. 9. Visualization of the iterates in $\mathbb{R}^d$ with $d = 2$ (top) and $d = 3$ (bottom), corresponding to Fig. 8. The uniform mixture Υ_2 (red cubes) approximates the optimal histogram for the empirical measure Δ_n (blue dots).

An alternative approach is to consider the sliced Wasserstein distance. To this end, the *Wasserstein distance* [65] of measures $\chi, \eta \in \mathcal{P}(\mathbb{R})$ is defined by

$$\mathrm{W}_2^2(\chi,\eta) := \inf_{\gamma\in\Pi(\chi,\eta)} \int_{\mathbb{R}\times\mathbb{R}} |x-y|^2 \,\mathrm{d}\gamma(x,y),$$

where Π is the set of measures $\gamma \in \mathcal{P}(\mathbb{R}\times\mathbb{R})$ with marginals χ and η. We have

$$\mathrm{W}_2^2(\chi,\eta) = \int_0^1 |F_\chi^{[-1]}(s) - F_\eta^{[-1]}(s)|^2 \,\mathrm{d}\rho(s) \qquad (20)$$

for some reference measure $\rho \in \mathcal{P}(\mathbb{R})$ with $\chi \ll \rho$, with the cumulative distribution function $F_\chi(s) := \chi((-\infty, s])$ and its generalized inverse $g^{[-1]}(t) := \inf\{s \mid g(s) > t\}$. The *sliced Wasserstein distance* of $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ is defined by

$$\mathrm{SW}_2^2(\mu,\nu) := \int_{\mathbb{S}^{d-1}} \mathrm{W}_2^2(\mathcal{R}_{\boldsymbol{\theta}}[\mu], \mathcal{R}_{\boldsymbol{\theta}}[\nu]) \,\mathrm{d}u_{\mathbb{S}^{d-1}}(\boldsymbol{\theta}), \qquad (21)$$

where $u_{\mathbb{S}^{d-1}}$ is the uniform measure on $\mathbb{S}^{d-1}$ and the Radon transform of measures, corresponding to (1) for the density, is defined by

$$\mathcal{R}_{\boldsymbol{\theta}}[\mu] := \langle\cdot,\boldsymbol{\theta}\rangle_{\#}\mu = \mu\circ(\langle\cdot,\boldsymbol{\theta}\rangle)^{-1} \in \mathcal{P}(\mathbb{R}), \quad \boldsymbol{\theta}\in\mathbb{S}^{d-1}. \qquad (22)$$

Compared to the Wasserstein distance, its sliced version combines similar properties with faster computation, cf. [51], [55], [60], [61].

Then, we propose to approximate the empirical measure Δ_n by solving

$$\operatorname*{arg\,min}_{\boldsymbol{c},\boldsymbol{w}\in\mathbb{X}^k} \mathrm{SW}_2^2(\Upsilon_k, \Delta_n), \qquad (23)$$

where Υ_k is parameterized by $\boldsymbol{c}$ and $\boldsymbol{w}$. For the numerical computation, the Wasserstein distance W_2^2 is computed by (20), the Radon transform of Υ_k is computed by a linear combination of (6), and the Radon transform (22) of the empirical measure is given by

$$\mathcal{R}_{\boldsymbol{\theta}}[\Delta_n] = \sum_{i=1}^n \delta_{\langle \boldsymbol{x}_i, \boldsymbol{\theta}\rangle},$$

where we apply nearest neighbor interpolation.

We provide two experiments for $d \in \{2,3\}$ and consider Δ_n sampled from the uniform measure on $[-1,-1/2]^d \cup [1/2,1]^d$ with $n = 40$ samples for $d = 2$ and $n = 60$ for $d = 3$. We use 101 radii equispaced in $[-\sqrt{d},\sqrt{d}]$, and 128 angles equispaced on $\mathbb{S}^1$, and Fibonacci points (18) on $\mathbb{S}^2$. We solve (23), with the ADAM optimizer [25] using 100 epochs, learning rate 0.05, smoothing parameters $(0.9, 0.99)$, and initialization $\boldsymbol{c}_{1,2} = \pm 1/4\boldsymbol{e} \in \mathbb{R}^d$ and $\boldsymbol{w}_{1,2} = 1/10\boldsymbol{e} \in \mathbb{R}^d$. The Radon transforms of the target Δ_n and optimal measure Υ_2 are visualized in Fig. 8. The iterates are visualized in Fig. 9. We observe that the computed measures Υ_n perfectly cluster in both settings. The corresponding Radon transforms reproduce the main character of the target.

B. Sliced Wasserstein barycenter

The *sliced Wasserstein barycenter* [13] of two probability measures $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R}^d)$ with finite second moment, such that $\mathcal{R}[\mu_i] \in \mathcal{P}_2(\mathbb{S}^{d-1}\times\mathbb{R})$ for $i \in [\![2]\!]$, is defined for some $\lambda \in (0,1)$ by

$$\operatorname*{arg\,min}_{\mu\in\mathcal{P}(\mathbb{R}^d)} \lambda \mathrm{SW}_2^2(\mu_1,\mu) + (1-\lambda)\mathrm{SW}_2^2(\mu_2,\mu), \qquad (24)$$

where the sliced Wasserstein distance is defined in (21).

The computation of the Radon transform is performed in the manner of Section IV-A with 42 angles $\boldsymbol{\theta}$ and 31 radii t. The reference measures μ_j are either two hemispheres rotated by 90° around the z-axis, or a sphere and a hemisphere, see Fig. 10. The minimization uses the same setting as in Section IV-A and relies on a free support discretization of μ in the form Υ_k, see (19), with $k = 200$ centers in $\mathbb{R}^3$, The barycenters are visualized in Fig. 10, and the corresponding Radon transform in Fig. 11.

V. Comparison with Monte Carlo integration

Another implementation of the Radon transform, cf. Rem. 8, is based on a Monte Carlo integration, see [1], [20], [24]. Similar to (13), we approximate the intersection area of $(-1,1]^d$ and the hyperplane $H_{\boldsymbol{\theta}}(t)$ by

$$A_{\boldsymbol{\theta}}^{\boldsymbol{e}}(t) \approx \tfrac{1}{2\varepsilon} V_{\boldsymbol{\theta}}^{\boldsymbol{e}}(t-\varepsilon, t+\varepsilon) = \frac{\mathbb{E}_{\boldsymbol{x}\sim\mathcal{U}((-1,1]^d)}[|\langle\boldsymbol{\theta},\boldsymbol{x}\rangle - t| \leqslant \varepsilon]}{2^{1-d}\varepsilon} \qquad (25)$$

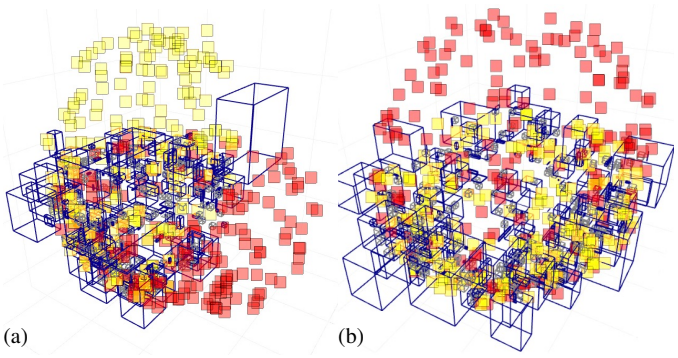

Fig. 10. Visualization of the calculated sliced Wasserstein barycenters (blue boxes) between the voxelized target measures for (a) two rotated hemispheres (red, yellow) and (b) sphere (red) and hemisphere (yellow).

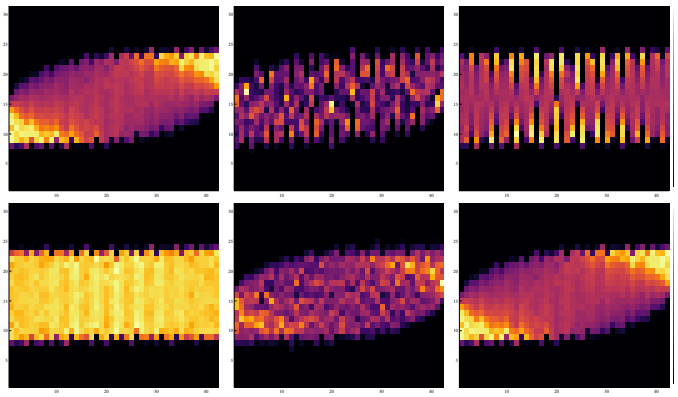

Fig. 11. 3D Radon transform of the two reference measures μ_1 (left) and μ_2 (right), and the computed barycenter ($\lambda = 1/2$) from (24) (middle). The measures are visualized in Fig. 10.

with N random samples of $\boldsymbol{x}$ from the uniform distribution $\mathcal{U}((-1,1]^d)$. We use $d = 4$, $N = 2^q$ samples with $q \in \{16, 20, 24\}$, and $\varepsilon = 10^{-3}$ for the Monte Carlo sampling. We generate quasi-uniform points on the sphere $\mathbb{S}^3$ by applying the inverse Gaußian cumulative distribution function to Sobol' points [67] in $[0,1]^4$ and normalizing to the sphere, following the construction in [43, § 5.2]. Notably, the evaluation of our formula (6) for 128 angles $\boldsymbol{\theta}$ and 128 radii t takes less than 1 second, whereas the Monte Carlo sampling for just 1 angle and 128 radii takes about 5 seconds with $N = 2^{24}$ samples for an appropriate result with less than 1% error. The error depending on N is plotted in Fig. 12.

Details of the Implementation

We implemented the multidimensional discrete Radon transform and the experiments in Julia.[3] They are performed on a MacBookPro 2020 with an Intel Core i5 CPU (4 cores, 1.4 GHz) and 8 GB RAM.

Acknowledgments

The authors thank M. Piening for drawing the attention on the experiment in Section IV-A.

[3]Code available at https://github.com/JJEWBresch/XdRadonTransform.

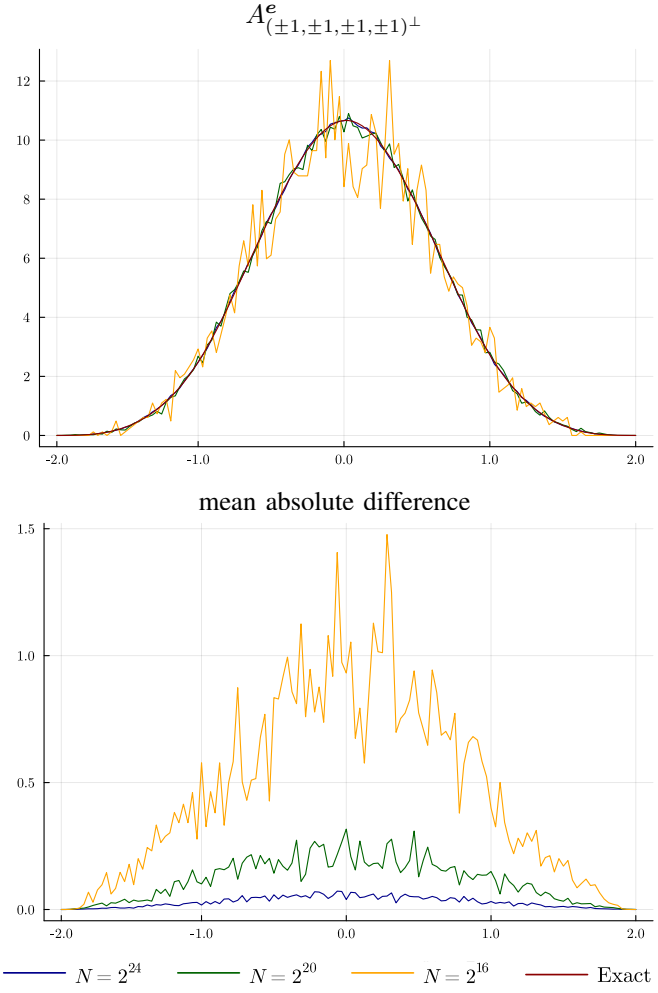

Fig. 12. Comparison of the exact formula (6) and the Monte Carlo approximation (25) of the Radon transform $A^{\boldsymbol{e}}_{\boldsymbol{\theta}}(t)$ in $d = 4$ (top), and the mean absolute difference for 20 repetitions (bottom), both depending on t.

Funding

Funding by the DFG under the SFB 10.55776/F68 "Tomography Across the Scales" (STE 571/19-1, project number: 495365311) is gratefully acknowledged. For the purpose of open access, the authors have applied a CC BY public copyright license to any authors-accepted manuscript version arising from this submission.

References

[1] C. Agarwal, A. Mhatre, S. Patra, S. Chaudhury, and A. Goswami. Algebraic reconstruction technique combined with Monte Carlo method for weight matrix calculation in gamma ray transmission tomography. *Discov. Appl. Sci.*, 1:1157, 2019. doi:10.1007/s42452-019-1201-1.

[2] F. Andersson, M. Carlsson, and V. V. Nikitin. Fast algorithms and efficient GPU implementations for the Radon transform and the backprojection operator represented as convolution operators. *SIAM J. Imaging Sci.*, 9(2), 2016. doi:10.1137/15M1023762.

[3] A. Averbuch, R. R. Coifman, D. L. Donoho, M. Israeli, and Y. Shkolnisky. A framework for discrete integral transformations I – The pseudopolar Fourier transform. *SIAM J. Sci. Comput.*, 30(2):764–784, 2008. doi:10.1137/060650283.

[4] A. Averbuch, R. R. Coifman, D. L. Donoho, M. Israeli, Y. Shkolnisky, and I. Sedelnikov. A framework for discrete integral transformations II – The 2D discrete Radon transform. *SIAM J. Sci. Comput.*, 30(2):785–803, 2008. doi:10.1137/060650301.

[5] A. Averbuch, G. Shabat, and Y. Shkolnisky. Direct inversion of the three-dimensional pseudo-polar Fourier transform. *SIAM J. Sci. Comput.*, 38(2):A1100–A1120, 2016. doi:10.1137/15M1031916.

[6] A. Averbuch and Y. Shkolnisky. 3D Fourier based discrete Radon transform. *Appl. Comput. Harmon. Anal.*, 15(1):33–69, 2003. doi:10.1016/S1063-5203(03)00030-7.

[7] K. Ball. Cube slicing in $\mathbb{R}^n$. *Proc. Am. Math. Soc.*, 97(3):465–473, 1986. doi:10.1090/S0002-9939-1986-0840631-0.

[8] D. L. Barrow and P. W. Smith. Spline notation applied to a volume problem. *Am. Math. Mon.*, 86(1):50–51, 1979. doi:10.1080/00029890.1979.1199473.

[9] M. Beckmann, R. Beinert, and J. Bresch. Max-normalized Radon cumulative distribution transform for limited data classification. In *International Conference on Scale Space and Variational Methods in Computer Vision (SSVM)*, pages 241–254, 2024. doi:10.1007/978-3-031-92366-1_19.

[10] M. Beckmann, R. Beinert, and J. Bresch. Generalizations of the normalized Radon cumulative distribution transform for limited data recognition, 2025. arXiv:2512.08099. doi:10.48550/arXiv.2512.08099.

[11] M. Beckmann, R. Beinert, and J. Bresch. Normalized Radon cumulative distribution transforms for invariance and robustness in optimal transport based image classification, 2025. arXiv:2506.08761. doi:10.48550/arXiv.2506.08761.

[12] G. Beylkin. Discrete Radon transform. *IEEE Trans. Acoust., Speech, Signal Process.*, 35(2):162–172, 1987. doi:10.1109/TASSP.1987.1165108.

[13] N. Bonneel, J. Rabin, G. Peyré, and H. Pfister. Sliced and Radon Wasserstein barycenters of measures. *J. Math. Imaging Vis.*, 51(1):22–45, 2015. doi:10.1007/s10851-014-0506-3.

[14] R. N. Bracewell. *Two-Dimensional Imaging*. Prentice Hall, Englewood Cliffs, NJ, USA, 1 edition, 1995.

[15] M.-C. Brandenburg, J. A. D. Loera, and C. Meroni. The best ways to slice a polytope. *Math. Comput.*, 94(352):1003–1042, 2025. doi:10.1090/mcom/4006.

[16] K. Bredies and R. Huber. Convergence analysis of pixel-driven Radon and fanbeam transforms. *SIAM J. Numer. Anal.*, 59(3):1399–1432, 2021. doi:10.1137/20M1326635.

[17] J. Bresch, D. A. Lorenz, F. Schneppe, and M. Winkler. Matrix-free stochastic calculation of operator norms without using adjoints, 2024. arXiv:2410.08297. doi:10.48550/arXiv.2410.08297.

[18] J. Bresch, D. A. Lorenz, F. Schneppe, and M. Winkler. Computing adjoint mismatch of linear maps, 2025. arXiv:2503.21361. doi:10.48550/arXiv.2503.21361.

[19] S. R. Carl Boor, Klaus Höllig. *Box Splines*. Applied Mathematical Sciences. Springer, New York, NY, USA, 1993. doi:10.1007/978-1-4757-2244-4.

[20] B. Chyba, M. Mantler, and M. Reiter. Monte Carlo simulation of projections in computed tomography. *Powder Diffr.*, 23(2):150–153, 2008. doi:10.1154/1.2919045.

[21] I. Csiszár. A class of measures of informativity of observation channels. *Period. Math. Hung.*, 2:191–213, 1972. doi:10.1007/BF02018661.

[22] P. Daras, D. Zarpalas, D. Tzovaras, and M. Strintzis. Shape matching using the 3D Radon transform. In *2nd International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT)*, pages 953–960. IEEE, 2004. doi:10.1109/TDPVT.2004.1335419.

[23] P. Daras, D. Zarpalas, D. Tzovaras, and M. Strintzis. Efficient 3D model search and retrieval using generalized 3D Radon transforms. *IEEE Trans. Multimed.*, 8(1):101–114, 2006. doi:10.1109/TMM.2005.861287.

[24] M. V. de Hoop and C. Spencer. Quasi-Monte Carlo integration over $\mathbb{S}^2 \times \mathbb{S}^2$ for migration inversion. *Inverse Probl.*, 12(3):219, 1996. doi:10.1088/0266-5611/12/3/004.

[25] J. B. Diederik P. Kingma. Adam: A method for stochastic optimization, 2015. arXiv:1412.6980. doi:10.48550/arXiv.1412.6980.

[26] A. González. Measurement of areas on a sphere using Fibonacci and latitude–longitude lattices. *Math. Geosci.*, 42(1):49–64, 2010. doi:10.1007/s11004-009-9257-x.

[27] C. Groenland and T. Johnston. Intersection sizes of linear subspaces with the hypercube. *J. Comb. Theory Ser. A*, 170:105142, 2020. doi:10.1016/j.jcta.2019.105142.

[28] M. Gräf and D. Potts. On the computation of spherical designs by a new optimization approach based on fast spherical Fourier transforms. *Numer. Math.*, 119:699–724, 2011. doi:10.1007/s00211-011-0399-7.

[29] J. H. Hannay and J. F. Nye. Fibonacci numerical integration on a sphere. *J. Phys. A: Math. Gen.*, 37(48):11591, 2004. doi:10.1088/0305-4470/37/48/005.

[30] D. Hensley. Slicing the cube in $\mathbb{R}^n$ and probability (bounds for the measure of a central cube slice in $\mathbb{R}^n$ by probability methods). *Proc. Am. Math. Soc.*, 73(1):95–100, 1979. doi:10.2307/2042889.

[31] J. Hertrich, T. Jahn, and M. Quellmalz. Fast summation of radial kernels via QMC slicing. In *International Conference on Learning Representations (ICLR)*, 2025. URL: https://openreview.net/forum?id=iNmVX9lx9l.

[32] S. Horbelt, M. Liebling, and M. Unser. Discretization of the Radon transform and of its inverse by spline convolutions. *IEEE Trans. Med. Imaging*, 21:363–376, 05 2002. doi:10.1109/TMI.2002.1000260.

[33] R. Huber. Convergence of ray- and pixel-driven discretization frameworks in the strong operator topology. *Inverse Probl. Imag.*, 2025. doi:10.3934/ipi.2026007.

[34] A. Kadyrov and M. Petrou. The trace transform and its applications. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(8):811–828, 2001. doi:10.1109/34.946986.

[35] K. Katayama, Y. Yamaji, S. Toyoizumi, and T. Hirashima. Content-based retrieval of 3D CAD subassemblies using 3D Radon transform. *J. Inf. Process.*, 31:734–742, 2023. doi:10.2197/ipsjjip.31.734.

[36] F. Killmann and E. von Collani. A note on the convolution of the uniform and related distributions and their use in quality control. *Econ. Qual. Contr.*, 16:17–41, 2001. doi:10.1515/EQC.2001.17.

[37] T. Knopp, S. Kunis, and D. Potts. A note on the iterative MRI reconstruction from nonuniform k-space data. *Int. J. Biomed. Imaging*, 2007:24727, 2007. doi:10.1155/2007/24727.

[38] S. Kolouri, S. Park, M. Thorpe, D. Slepcev, and G. Rohde. Generalized sliced-Wasserstein distances. In *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)*, pages 261–272, Red Hook, NY, USA, 2019. Curran Associates Inc. URL: https://dl.acm.org/doi/10.5555/3454287.3454311.

[39] S. Kolouri, S. R. Park, and G. K. Rohde. The Radon cumulative distribution transform and its application to image classification. *IEEE Trans. Image Process.*, 25(2):920–934, 2016. doi:10.1109/TIP.2015.2509419.

[40] H. König. Non-central sections of the simplex, the cross-polytope and the cube. *Adv. Math.*, 376:107458, 2021. doi:10.1016/j.aim.2020.107458.

[41] H. König and A. Koldobsky. Volumes of low-dimensional slabs and sections in the cube. *Adv. Appl. Math.*, 47(4):894–907, 2011. doi:10.1016/j.aam.2011.05.001.

[42] S. Lanzavecchia and P. Luigi Bellon. Fast computation of 3D Radon transform via a direct Fourier method. *Bioinform.*, 14(2):212–216, 1998. doi:10.1093/bioinformatics/14.2.212.

[43] C. Lemieux. *Monte Carlo and Quasi-Monte Carlo Sampling*. Springer Series in Statistics. Springer, New York, NY, USA, 1st edition, 2009. doi:10.1007/978-0-387-78165-5.

[44] J. A. D. Loera, G. Lopez-Campos, and A. J. Torres. On the number of vertices in a hyperplane section of a polytope, 2025. arXiv:2412.12419. doi:10.48550/arXiv.2412.12419.

[45] J. Ma and Z. Ma. 3D Radon transform for shape retrieval using bag-of-visual-features. *Int. Arab J. Inf. Technol.*, 17(4):471–479, 2020. doi:10.34028/iajit/17/4/5.

[46] W. A. Mahmoud and M. Shaker. 3D ear print authentication using 3D Radon transform. In *2nd International Conference on Information & Communication Technologies*, volume 1, pages 1052–1056. IEEE, 2006. doi:10.1109/ICTTA.2006.1684519.

[47] J. G. Marichal-Hernandez, Óscar Gómez-Cárdenes, F. L. R. González, D. H. Kim, and J. M. Rodríguez-Ramos. Three-dimensional multiscale discrete Radon and John transforms. *Opt. Eng.*, 59(9):093104, 2020. doi:10.1117/1.OE.59.9.093104.

[48] M. Meduňa, F. Isa, F. Bressan, and H. von Känel. The Radon transform as a tool for 3D reciprocal-space mapping of epitaxial microcrystals. *J. Appl. Crystallogr.*, 55(4):823–836, 2022. doi:10.1107/S1600576722004885.

[49] N. Melo and A. Winter. Intersection patterns of linear subspaces with the hypercube. *J. Comb. Theory Ser. A.*, 164:60–71, 2019. doi:10.1016/j.jcta.2018.12.006.

[50] J. Moody, C. Stone, D. Zach, and A. Zvavitch. A remark on extremal non-central sections of the unit cube. *Asympt. Geometr. Anal., Fields Inst. Comm.*, 68:211–228, 2013. doi:10.1007/978-1-4614-6406-8_9.

[51] K. Nguyen. An introduction to sliced optimal transport: Foundations, advances, extensions, and applications. *Found. Trends Comput. Graph. Vis.*, 17(3-4):171–406, 2025. doi:10.1561/0600000119.

[52] D. Nikolaev, E. Ershov, A. Kroshnin, E. Limonova, A. Mukovozov, and I. Faradzhev. On a fast Hough/Radon transform as a compact summation scheme over digital straight line segments. *Math.*, 11(15):3336, 2023. doi:10.3390/math11153336.

[53] P. E. O'Neil. Hyperplane cuts of an n-cube. *Discrete Math.*, 1(2):193–195, 1971. doi:10.1016/0012-365X(71)90025-2.

[54] V. P. Palamodov. *Reconstruction from Integral Data*. Chapman & Hall, London, UK, 1 edition, 2016. doi:10.1007/978-3-319-34445-4.

[55] G. Peyré and M. Cuturi. Computational optimal transport. *Found. Trends Mach. Learn.*, 11(5-6):355–607, 2019. doi:10.1561/2200000073.

[56] G. Plonka, D. Potts, G. Steidl, and M. Tasche. *Numerical Fourier Analysis*. Applied and Numerical Harmonic Analysis. Birkhäuser, Basel, Switzerland, 2nd edition, 2023. doi:10.1007/978-3-031-35005-4.

[57] L. Pournin. Shallow sections of the hypercube. *Isr. J. Math.*, 255:685–704, 2023. doi:10.1007/s11856-022-2400-9.

[58] L. Pournin. Local extrema for hypercube sections. *J. Anal. Math.*, 152:557–594, 2024. doi:10.1007/s11854-023-0304-1.

[59] G. Pólya. Berechnung eines bestimmten Integrals. *Math. Ann.*, 74:204–212, 1913. doi:10.1007/BF01456040.

[60] M. Quellmalz, R. Beinert, and G. Steidl. Sliced optimal transport on the sphere. *Inverse Probl.*, 39(10):105005, 2023. doi:10.1088/1361-6420/acf156.

[61] M. Quellmalz, L. Buecher, and G. Steidl. Parallelly sliced optimal transport on spheres and on the rotation group. *J. Math. Imaging Vis.*, 66:951–976, 2024. doi:10.1007/s10851-024-01206-w.

[62] N. Rux, M. Quellmalz, and G. Steidl. Slicing of radial functions: a dimension walk in the Fourier space. *Sampl. Theory Signal Process. Data Anal.*, 23, 2025. doi:10.1007/s43670-025-00100-9.

[63] N. Rux, M. Quellmalz, and G. Steidl. Smoothed distance kernels for MMDs and applications in Wasserstein gradient flows. *Adv. Comput. Math.*, 2026. doi:10.1007/s10444-026-10289-5.

[64] E. Saff, I. Sloan, and R. Womersley. QMC designs: Optimal order quasi Monte Carlo integration schemes on the sphere. *Math. Comput.*, 83(290):2821–2851, 2014. URL: https://www.jstor.org/stable/24488679.

[65] F. Santambrogio. *Optimal Transport for Applied Mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and Their Applications*. Birkhäuser, Cham, Switzerland, 2015. doi:10.1007/978-3-319-20828-2.

[66] W. Shi, W. Wang, Y. Shi, S. Chen, Z. Li, and N. Wang. 3D high-resolution Radon transform based on a strong sparse L_{p-1} norm and its applications. *J. Geophys. Eng.*, 21(3):1008–1026, 2024. doi:10.1093/jge/gxae052.

[67] I. Sobol'. On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Comput. Math. & Math. Phys.*, 7(4):86–112, 1967. doi:10.1016/0041-5553(67)90144-9.

[68] A. Sommerfeld. Eine besondere anschauliche Ableitung des Gaußischen Fehlergesetzes. In S. Meyer, editor, *Festschrift Ludwig Boltzmann gewidmet zum sechzigsten Geburtstage 20. Februar 1904*, pages 848–859, Leipzig, 1904. Barth. Boltzmann–Festschrift.

[69] P. A. Toft. *The Radon Transform – Theory and Implementation*. Phd thesis, Technical University of Denmark, 1996. URL: https://orbit.dtu.dk/en/publications/the-radon-transform-theory-and-implementation.

[70] F. Wechsler, C. Gigli, J. Madrid-Wolff, and C. Moser. Wave optical model for tomographic volumetric additive manufacturing. *Opt. Express*, 32:14705–14712, 2024. doi:10.1364/OE.521322.

[71] R. S. Womersley. Efficient spherical designs with good geometric properties. In J. Dick, F. Kuo, and H. Wozniakowski, editors, *Contemporary Computational Mathematics - A Celebration of the 80th Birthday of Ian Sloan*, pages 1243–1285. Springer, 2018. doi:10.1007/978-3-319-72456-0_57.

[72] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, and J. Xiao. 3D ShapeNets: A deep representation for volumetric shapes. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1912–1920, 2015. doi:10.1109/CVPR.2015.7298801.

[73] D. Zarpalas, P. Daras, A. Axenopoulos, D. Tzovaras, and M. G. Strintzis. 3D model search and retrieval using the spherical trace transform. *EURASIP J. Adv. Signal Process.*, pages 953–960, 2006. doi:10.1155/2007/23912.