Beckmann, M., Beinert, R., and Bresch, J. (2025) – Max-normalized Radon CDT for limited data classification, **Lect. Notes Comput. Sci.**

Beckmann, M., Beinert, R., and Bresch, J. (2025) – Normalized Radon CDT for invariance and robustness in optimal transport based image classification, **SIAM J. Imaging Sci.**

Beckmann, M., Beinert, R., and Bresch, J. (2025) – Generalization of the normalized Radon CDT for limited data recognition, **arXiv.2512.08099**

# Generalization of Normalized Radon Cumulative Distribution Transforms in the Light as Feature Extractors for Image Classification

Jonas Bresch | Technical University Berlin | joint work with
Robert Beinert and Matthias Beckmann

Funded by:

Berlin Mathematical School BMS

DAAD

## Motivation

– Automatic classification of images, for instance watermarks in historical manuscripts[1]" (see below).

[1] Hauser, D. et al. (2024) – "On image processing and pattern recognition for thermograms of watermarks in manuscripts – A first proof-of-concept,
[2] Wang, W. et al. (2013) – A linear optimal transportation framework for quantifying and visualizing variations in sets of images, **Int. J. Comput. Vis.**
[3] Moosmüller, C. and Cloninger, A. (2023) – Linear optimal transport embedding: Provable Wasserstein classification for certain rigid transformations and perturbations, **IMA J. Appl. Math.**

# Motivation

- Automatic classification of images, for instance watermarks in historical manuscripts[1]" (see below).
- **Challenges:** hundreds of classes, few samples per class.


[1] Hauser, D. et al. (2024) – "On image processing and pattern recognition for thermograms of watermarks in manuscripts – A first proof-of-concept,
[2] Wang, W. et al. (2013) – A linear optimal transportation framework for quantifying and visualizing variations in sets of images, **Int. J. Comput. Vis.**
[3] Moosmüller, C. and Cloninger, A. (2023) – Linear optimal transport embedding: Provable Wasserstein classification for certain rigid transformations and perturbations, **IMA J. Appl. Math.**

## Motivation

- Automatic classification of images, for instance watermarks in historical manuscripts[1]" (see below).
- **Challenges:** hundreds of classes, few samples per class.
- Reasonable assumption: **"affine"** class structure.

[1] Hauser, D. et al. (2024) – "On image processing and pattern recognition for thermograms of watermarks in manuscripts – A first proof-of-concept,
[2] Wang, W. et al. (2013) – A linear optimal transportation framework for quantifying and visualizing variations in sets of images, **Int. J. Comput. Vis.**
[3] Moosmüller, C. and Cloninger, A. (2023) – Linear optimal transport embedding: Provable Wasserstein classification for certain rigid transformations and perturbations, **IMA J. Appl. Math.**

## Motivation

- Automatic classification of images, for instance watermarks in historical manuscripts[1]" (see below).
- **Challenges:** hundreds of classes, few samples per class.
- Reasonable assumption: **"affine"** class structure.
- Need **invariance** to translations/dilations and, generally, stability under (non-)affine deformations[2][3].

[1] Hauser, D. et al. (2024) – "On image processing and pattern recognition for thermograms of watermarks in manuscripts – A first proof-of-concept,
[2] Wang, W. et al. (2013) – A linear optimal transportation framework for quantifying and visualizing variations in sets of images, **Int. J. Comput. Vis.**
[3] Moosmüller, C. and Cloninger, A. (2023) – Linear optimal transport embedding: Provable Wasserstein classification for certain rigid transformations and perturbations, **IMA J. Appl. Math.**

# Contributions

– Feature extractor via Radon cumulative distribution transfrom (R-CDT) [4]

[4] Kolouri, S., Park, S. R., and Rohde, G. K. (2016) – The Radon cumulative distribution transform and its application to image classification, **IEEE Trans. Image Process.**
[5] Natterer, F. (2001) – The Mathematics of Computerized Tomography,
[6] Deng, L. (2012) – The MNIST database of handwritten digit images for machine learning research, **IEEE Signal Process. Mag.**
[7] Liu, C.-L. et al. (2011) – "CASIA online and offline Chinese handwriting databases,
[8] Koch, S. et al. (2019) – "ABC: A big CAD model dataset for geometric deep learning,
[9] Bogo, F. et al. (2014) – "FAUST: Dataset and evaluation for 3D mesh registration,
[10] Flusser, J. and Suk, T. (1993) – Pattern recognition by affine moment invariants, **Pattern Recognit.**

## Contributions

– Feature extractor via Radon cumulative distribution transfrom (R-CDT) [4]
→ based on one-dimensional optimal transport maps (CDTs),


[4] Kolouri, S., Park, S. R., and Rohde, G. K. (2016) – The Radon cumulative distribution transform and its application to image classification, **IEEE Trans. Image Process.**
[5] Natterer, F. (2001) – The Mathematics of Computerized Tomography,
[6] Deng, L. (2012) – The MNIST database of handwritten digit images for machine learning research, **IEEE Signal Process. Mag.**
[7] Liu, C.-L. et al. (2011) – "CASIA online and offline Chinese handwriting databases,
[8] Koch, S. et al. (2019) – "ABC: A big CAD model dataset for geometric deep learning,
[9] Bogo, F. et al. (2014) – "FAUST: Dataset and evaluation for 3D mesh registration,
[10] Flusser, J. and Suk, T. (1993) – Pattern recognition by affine moment invariants, **Pattern Recognit.**

## Contributions

– Feature extractor via Radon cumulative distribution transfrom (R-CDT) [4]

→ based on one-dimensional optimal transport maps (CDTs), and generalized to

[4] Kolouri, S., Park, S. R., and Rohde, G. K. (2016) – The Radon cumulative distribution transform and its application to image classification, **IEEE Trans. Image Process.**
[5] Natterer, F. (2001) – The Mathematics of Computerized Tomography,
[6] Deng, L. (2012) – The MNIST database of handwritten digit images for machine learning research, **IEEE Signal Process. Mag.**
[7] Liu, C.-L. et al. (2011) – "CASIA online and offline Chinese handwriting databases,
[8] Koch, S. et al. (2019) – "ABC: A big CAD model dataset for geometric deep learning,
[9] Bogo, F. et al. (2014) – "FAUST: Dataset and evaluation for 3D mesh registration,
[10] Flusser, J. and Suk, T. (1993) – Pattern recognition by affine moment invariants, **Pattern Recognit.**

# Contributions

– Feature extractor via Radon cumulative distribution transfrom (R-CDT) [4]

→ based on one-dimensional optimal transport maps (CDTs), and generalized to

- $\mathbb{R}^2$, via the 2-**dimensional** Radon transform for R-CDT for $\mathcal{P}(\mathbb{R}^2)$,


[4] Kolouri, S., Park, S. R., and Rohde, G. K. (2016) – The Radon cumulative distribution transform and its application to image classification, **IEEE Trans. Image Process.**
[5] Natterer, F. (2001) – The Mathematics of Computerized Tomography,
[6] Deng, L. (2012) – The MNIST database of handwritten digit images for machine learning research, **IEEE Signal Process. Mag.**
[7] Liu, C.-L. et al. (2011) – “CASIA online and offline Chinese handwriting databases,
[8] Koch, S. et al. (2019) – “ABC: A big CAD model dataset for geometric deep learning,
[9] Bogo, F. et al. (2014) – “FAUST: Dataset and evaluation for 3D mesh registration,
[10] Flusser, J. and Suk, T. (1993) – Pattern recognition by affine moment invariants, **Pattern Recognit.**

# Contributions

- Feature extractor via Radon cumulative distribution transfrom (R-CDT) [4]

→ based on one-dimensional optimal transport maps (CDTs), and generalized to

  - $\mathbb{R}^2$, via the $2$-**dimensional** Radon transform for R-CDT for $\mathcal{P}(\mathbb{R}^2)$,
  - $\mathbb{R}^d$, via the $d$-**dimensional Radon** transform for R-CDT for $\mathcal{P}(\mathbb{R}^d)$,


[4] Kolouri, S., Park, S. R., and Rohde, G. K. (2016) – The Radon cumulative distribution transform and its application to image classification, **IEEE Trans. Image Process.**
[5] Natterer, F. (2001) – The Mathematics of Computerized Tomography,
[6] Deng, L. (2012) – The MNIST database of handwritten digit images for machine learning research, **IEEE Signal Process. Mag.**
[7] Liu, C.-L. et al. (2011) – “CASIA online and offline Chinese handwriting databases,
[8] Koch, S. et al. (2019) – “ABC: A big CAD model dataset for geometric deep learning,
[9] Bogo, F. et al. (2014) – “FAUST: Dataset and evaluation for 3D mesh registration,
[10] Flusser, J. and Suk, T. (1993) – Pattern recognition by affine moment invariants, **Pattern Recognit.**

# Contributions

– Feature extractor via Radon cumulative distribution transfrom (R-CDT) [4]

→ based on one-dimensional optimal transport maps (CDTs), and generalized to

- $\mathbb{R}^2$, via the $2$-**dimensional** Radon transform for R-CDT for $\mathcal{P}(\mathbb{R}^2)$,
- $\mathbb{R}^d$, via the $d$-**dimensional Radon** transform for R-CDT for $\mathcal{P}(\mathbb{R}^d)$,
- "more" general data on $\mathbb{X}$, via the **generalized, circular** and **rotation Radon transform**[5] for R-CDT for $\mathcal{P}(\mathbb{X})$.

[4] Kolouri, S., Park, S. R., and Rohde, G. K. (2016) – The Radon cumulative distribution transform and its application to image classification, **IEEE Trans. Image Process.**
[5] Natterer, F. (2001) – The Mathematics of Computerized Tomography,
[6] Deng, L. (2012) – The MNIST database of handwritten digit images for machine learning research, **IEEE Signal Process. Mag.**
[7] Liu, C.-L. et al. (2011) – "CASIA online and offline Chinese handwriting databases,
[8] Koch, S. et al. (2019) – "ABC: A big CAD model dataset for geometric deep learning,
[9] Bogo, F. et al. (2014) – "FAUST: Dataset and evaluation for 3D mesh registration,
[10] Flusser, J. and Suk, T. (1993) – Pattern recognition by affine moment invariants, **Pattern Recognit.**

## Contributions

- Feature extractor via Radon cumulative distribution transfrom (R-CDT) [4]

→ based on one-dimensional optimal transport maps (CDTs), and generalized to
  - $\mathbb{R}^2$, via the $2$-**dimensional** Radon transform for R-CDT for $\mathcal{P}(\mathbb{R}^2)$,
  - $\mathbb{R}^d$, via the $d$-**dimensional Radon** transform for R-CDT for $\mathcal{P}(\mathbb{R}^d)$,
  - "more" general data on $\mathbb{X}$, via the **generalized, circular** and **rotation Radon transform**[5] for R-CDT for $\mathcal{P}(\mathbb{X})$.
- Introduce $h$-**normalized Radon-CDT** ($_{\mathrm{h}}$NR-CDT) via two-step normalization.


[4] Kolouri, S., Park, S. R., and Rohde, G. K. (2016) – The Radon cumulative distribution transform and its application to image classification, **IEEE Trans. Image Process.**
[5] Natterer, F. (2001) – The Mathematics of Computerized Tomography,
[6] Deng, L. (2012) – The MNIST database of handwritten digit images for machine learning research, **IEEE Signal Process. Mag.**
[7] Liu, C.-L. et al. (2011) – "CASIA online and offline Chinese handwriting databases,
[8] Koch, S. et al. (2019) – "ABC: A big CAD model dataset for geometric deep learning,
[9] Bogo, F. et al. (2014) – "FAUST: Dataset and evaluation for 3D mesh registration,
[10] Flusser, J. and Suk, T. (1993) – Pattern recognition by affine moment invariants, **Pattern Recognit.**

## Contributions

- Feature extractor via Radon cumulative distribution transfrom (R-CDT) [4]

→ based on one-dimensional optimal transport maps (CDTs), and generalized to
  - $\mathbb{R}^2$, via the 2-**dimensional** Radon transform for R-CDT for $\mathcal{P}(\mathbb{R}^2)$,
  - $\mathbb{R}^d$, via the $d$-**dimensional Radon** transform for R-CDT for $\mathcal{P}(\mathbb{R}^d)$,
  - "more" general data on $\mathbb{X}$, via the **generalized, circular** and **rotation Radon transform**[5] for R-CDT for $\mathcal{P}(\mathbb{X})$.
- Introduce $h$-**normalized Radon-CDT** ($_h$NR-CDT) via two-step normalization.
- **Linear separability** for


[4] Kolouri, S., Park, S. R., and Rohde, G. K. (2016) – The Radon cumulative distribution transform and its application to image classification, **IEEE Trans. Image Process.**
[5] Natterer, F. (2001) – The Mathematics of Computerized Tomography,
[6] Deng, L. (2012) – The MNIST database of handwritten digit images for machine learning research, **IEEE Signal Process. Mag.**
[7] Liu, C.-L. et al. (2011) – "CASIA online and offline Chinese handwriting databases,
[8] Koch, S. et al. (2019) – "ABC: A big CAD model dataset for geometric deep learning,
[9] Bogo, F. et al. (2014) – "FAUST: Dataset and evaluation for 3D mesh registration,
[10] Flusser, J. and Suk, T. (1993) – Pattern recognition by affine moment invariants, **Pattern Recognit.**

## Contributions

- Feature extractor via Radon cumulative distribution transfrom (R-CDT) [4]

→ based on one-dimensional optimal transport maps (CDTs), and generalized to
  - $\mathbb{R}^2$, via the 2-**dimensional** Radon transform for R-CDT for $\mathcal{P}(\mathbb{R}^2)$,
  - $\mathbb{R}^d$, via the $d$-**dimensional Radon** transform for R-CDT for $\mathcal{P}(\mathbb{R}^d)$,
  - "more" general data on $\mathbb{X}$, via the **generalized, circular** and **rotation Radon transform**[5] for R-CDT for $\mathcal{P}(\mathbb{X})$.
- Introduce $h$-**normalized Radon-CDT** (${}_{\mathrm{h}}$NR-CDT) via two-step normalization.
- **Linear separability** for
  - affinely transformed measure classes → single-point sets in ${}_{\mathrm{h}}$NR-CDT-space.

[4] Kolouri, S., Park, S. R., and Rohde, G. K. (2016) – The Radon cumulative distribution transform and its application to image classification, **IEEE Trans. Image Process.**
[5] Natterer, F. (2001) – The Mathematics of Computerized Tomography,
[6] Deng, L. (2012) – The MNIST database of handwritten digit images for machine learning research, **IEEE Signal Process. Mag.**
[7] Liu, C.-L. et al. (2011) – "CASIA online and offline Chinese handwriting databases,
[8] Koch, S. et al. (2019) – "ABC: A big CAD model dataset for geometric deep learning,
[9] Bogo, F. et al. (2014) – "FAUST: Dataset and evaluation for 3D mesh registration,
[10] Flusser, J. and Suk, T. (1993) – Pattern recognition by affine moment invariants, **Pattern Recognit.**

# Contributions

- Feature extractor via Radon cumulative distribution transfrom (R-CDT) [4]

→ based on one-dimensional optimal transport maps (CDTs), and generalized to
  - $\mathbb{R}^2$, via the $2$-**dimensional** Radon transform for R-CDT for $\mathcal{P}(\mathbb{R}^2)$,
  - $\mathbb{R}^d$, via the $d$-**dimensional Radon** transform for R-CDT for $\mathcal{P}(\mathbb{R}^d)$,
  - "more" general data on $\mathbb{X}$, via the **generalized, circular** and **rotation Radon transform**[5] for R-CDT for $\mathcal{P}(\mathbb{X})$.
- Introduce $h$-**normalized Radon-CDT** (${}_{\mathrm{h}}$NR-CDT) via two-step normalization.
- **Linear separability** for
  - affinely transformed measure classes → single-point sets in ${}_{\mathrm{h}}$NR-CDT-space.
  - distorted measure classes → balls in ${}_{\mathrm{h}}$NR-CDT-space.

[4] Kolouri, S., Park, S. R., and Rohde, G. K. (2016) – The Radon cumulative distribution transform and its application to image classification, **IEEE Trans. Image Process.**
[5] Natterer, F. (2001) – The Mathematics of Computerized Tomography,
[6] Deng, L. (2012) – The MNIST database of handwritten digit images for machine learning research, **IEEE Signal Process. Mag.**
[7] Liu, C.-L. et al. (2011) – "CASIA online and offline Chinese handwriting databases,
[8] Koch, S. et al. (2019) – "ABC: A big CAD model dataset for geometric deep learning,
[9] Bogo, F. et al. (2014) – "FAUST: Dataset and evaluation for 3D mesh registration,
[10] Flusser, J. and Suk, T. (1993) – Pattern recognition by affine moment invariants, **Pattern Recognit.**

## Contributions

– Feature extractor via Radon cumulative distribution transfrom (R-CDT) [4]
→ based on one-dimensional optimal transport maps (CDTs), and generalized to
  - $\mathbb{R}^2$, via the $2$-**dimensional** Radon transform for R-CDT for $\mathcal{P}(\mathbb{R}^2)$,
  - $\mathbb{R}^d$, via the $d$-**dimensional Radon** transform for R-CDT for $\mathcal{P}(\mathbb{R}^d)$,
  - "more" general data on $\mathbb{X}$, via the **generalized, circular** and **rotation Radon transform**[5] for R-CDT for $\mathcal{P}(\mathbb{X})$.

– Introduce $h$-**normalized Radon-CDT** ($_{\mathrm{h}}$NR-CDT) via two-step normalization.
– **Linear separability** for
  - affinely transformed measure classes → single-point sets in $_{\mathrm{h}}$NR-CDT-space.
  - distorted measure classes → balls in $_{\mathrm{h}}$NR-CDT-space.

– Classifier: nearest template and neighbor, (linear) SVM.


[4] Kolouri, S., Park, S. R., and Rohde, G. K. (2016) – The Radon cumulative distribution transform and its application to image classification, **IEEE Trans. Image Process.**
[5] Natterer, F. (2001) – The Mathematics of Computerized Tomography,
[6] Deng, L. (2012) – The MNIST database of handwritten digit images for machine learning research, **IEEE Signal Process. Mag.**
[7] Liu, C.-L. et al. (2011) – "CASIA online and offline Chinese handwriting databases,
[8] Koch, S. et al. (2019) – "ABC: A big CAD model dataset for geometric deep learning,
[9] Bogo, F. et al. (2014) – "FAUST: Dataset and evaluation for 3D mesh registration,
[10] Flusser, J. and Suk, T. (1993) – Pattern recognition by affine moment invariants, **Pattern Recognit.**

## Contributions

– Feature extractor via Radon cumulative distribution transfrom (R-CDT) [4]

→ based on one-dimensional optimal transport maps (CDTs), and generalized to

- $\mathbb{R}^2$, via the $2$-**dimensional** Radon transform for R-CDT for $\mathcal{P}(\mathbb{R}^2)$,
- $\mathbb{R}^d$, via the $d$-**dimensional Radon** transform for R-CDT for $\mathcal{P}(\mathbb{R}^d)$,
- "more" general data on $\mathbb{X}$, via the **generalized, circular** and **rotation Radon transform**[5] for R-CDT for $\mathcal{P}(\mathbb{X})$.

– Introduce $h$-**normalized Radon-CDT** ($_{\mathrm{h}}$NR-CDT) via two-step normalization.

– **Linear separability** for

- affinely transformed measure classes → single-point sets in $_{\mathrm{h}}$NR-CDT-space.
- distorted measure classes → balls in $_{\mathrm{h}}$NR-CDT-space.

– Classifier: nearest template and neighbor, (linear) SVM.

– Numeric: on synthetic (Image and Rotation), LinMNIST[6], Chinese MNIST[7]," ABC[8]," and FAUST[9]" dataset.

[4] Kolouri, S., Park, S. R., and Rohde, G. K. (2016) – The Radon cumulative distribution transform and its application to image classification, **IEEE Trans. Image Process.**
[5] Natterer, F. (2001) – The Mathematics of Computerized Tomography,
[6] Deng, L. (2012) – The MNIST database of handwritten digit images for machine learning research, **IEEE Signal Process. Mag.**
[7] Liu, C.-L. et al. (2011) – "CASIA online and offline Chinese handwriting databases,
[8] Koch, S. et al. (2019) – "ABC: A big CAD model dataset for geometric deep learning,
[9] Bogo, F. et al. (2014) – "FAUST: Dataset and evaluation for 3D mesh registration,
[10] Flusser, J. and Suk, T. (1993) – Pattern recognition by affine moment invariants, **Pattern Recognit.**

## Contributions

– Feature extractor via Radon cumulative distribution transfrom (R-CDT) [4]

→ based on one-dimensional optimal transport maps (CDTs), and generalized to

- $\mathbb{R}^2$, via the $2$-**dimensional** Radon transform for R-CDT for $\mathcal{P}(\mathbb{R}^2)$,
- $\mathbb{R}^d$, via the $d$-**dimensional Radon** transform for R-CDT for $\mathcal{P}(\mathbb{R}^d)$,
- "more" general data on $\mathbb{X}$, via the **generalized, circular** and **rotation Radon transform**[5] for R-CDT for $\mathcal{P}(\mathbb{X})$.

– Introduce $h$-**normalized Radon-CDT** ($_{\mathrm{h}}$NR-CDT) via two-step normalization.

– **Linear separability** for

- affinely transformed measure classes → single-point sets in $_{\mathrm{h}}$NR-CDT-space.
- distorted measure classes → balls in $_{\mathrm{h}}$NR-CDT-space.

– Classifier: nearest template and neighbor, (linear) SVM.

– Numeric: on synthetic (Image and Rotation), LinMNIST[6], Chinese MNIST[7]," ABC[8]," and FAUST[9]" dataset.

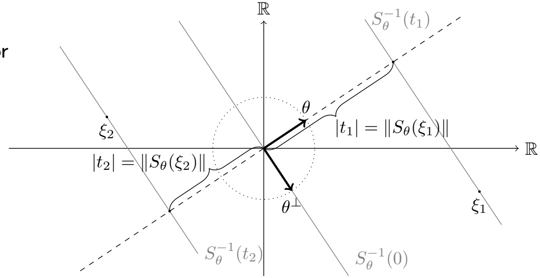
– Comparison: with LOT[2][3], R-CDT[4], momentum-based (MB) methods[10], and neuronal networks (NN).

[4] Kolouri, S., Park, S. R., and Rohde, G. K. (2016) – The Radon cumulative distribution transform and its application to image classification, **IEEE Trans. Image Process.**
[5] Natterer, F. (2001) – The Mathematics of Computerized Tomography,
[6] Deng, L. (2012) – The MNIST database of handwritten digit images for machine learning research, **IEEE Signal Process. Mag.**
[7] Liu, C.-L. et al. (2011) – "CASIA online and offline Chinese handwriting databases,
[8] Koch, S. et al. (2019) – "ABC: A big CAD model dataset for geometric deep learning,
[9] Bogo, F. et al. (2014) – "FAUST: Dataset and evaluation for 3D mesh registration,
[10] Flusser, J. and Suk, T. (1993) – Pattern recognition by affine moment invariants, **Pattern Recognit.**

## 2D Radon Transfrom on Functions vs. on Measures

- For $\boldsymbol{\theta} \in \mathbb{S}_1 := \{\mathbf{x} \in \mathbb{R}^2 \mid \|\mathbf{x}\| = 1\}$, define the slicing operator

$$S_{\boldsymbol{\theta}} \colon \mathbb{R}^2 \to \mathbb{R}, \quad \mathbf{x} \mapsto \langle \mathbf{x}, \boldsymbol{\theta} \rangle, \quad \mathbf{x} \in \mathbb{R}^2.$$

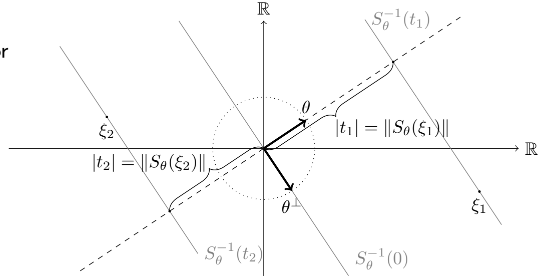


## 2D Radon Transfrom on Functions vs. on Measures

- For $\boldsymbol{\theta} \in \mathbb{S}_1 := \{\mathbf{x} \in \mathbb{R}^2 \mid \|\mathbf{x}\| = 1\}$, define the slicing operator

$$S_{\boldsymbol{\theta}} \colon \mathbb{R}^2 \to \mathbb{R}, \quad \mathbf{x} \mapsto \langle \mathbf{x}, \boldsymbol{\theta} \rangle, \quad \mathbf{x} \in \mathbb{R}^2.$$

- with inverse $S_{\boldsymbol{\theta}}^{-1}(t) = \{t\boldsymbol{\theta} + \tau\boldsymbol{\theta}^\perp \mid \tau \in \mathbb{R}\} \subset \mathbb{R}^2$.



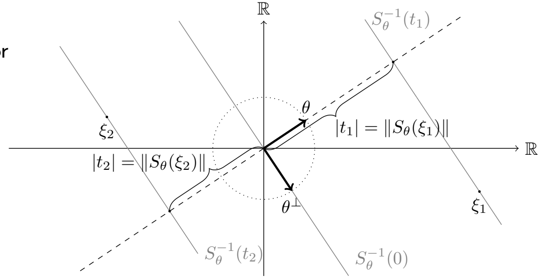
## 2D Radon Transfrom on Functions vs. on Measures

- For $\boldsymbol{\theta} \in \mathbb{S}_1 := \{\mathbf{x} \in \mathbb{R}^2 \mid \|\mathbf{x}\| = 1\}$, define the slicing operator

$$S_{\boldsymbol{\theta}} \colon \mathbb{R}^2 \to \mathbb{R}, \quad \mathbf{x} \mapsto \langle \mathbf{x}, \boldsymbol{\theta} \rangle, \quad \mathbf{x} \in \mathbb{R}^2.$$

- with inverse $S_{\boldsymbol{\theta}}^{-1}(t) = \{t\boldsymbol{\theta} + \tau\boldsymbol{\theta}^{\perp} \mid \tau \in \mathbb{R}\} \subset \mathbb{R}^2$.
- Let $f \in \mathrm{L}^1(\mathbb{R}^2)$, its **Radon transform** $\mathcal{R}[f] \colon \mathbb{R} \times \mathbb{S}_1 \to \mathbb{R}$ is the line integral, here $\mathrm{d}s$ denotes the arc length of $S_{\boldsymbol{\theta}}^{-1}(t)$

$$\mathcal{R}[f](t, \boldsymbol{\theta}) := \int_{S_{\boldsymbol{\theta}}^{-1}(t)} f(s) \, \mathrm{d}s, \quad (t, \boldsymbol{\theta}) \in \mathbb{R} \times \mathbb{S}_1,$$



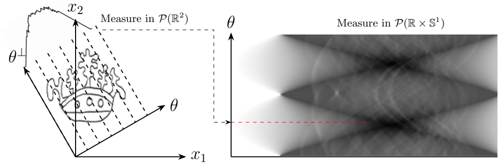
## 2D Radon Transfrom on Functions vs. on Measures

- For $\boldsymbol{\theta} \in \mathbb{S}_1 := \{\mathbf{x} \in \mathbb{R}^2 \mid \|\mathbf{x}\| = 1\}$, define the slicing operator

$$S_{\boldsymbol{\theta}} \colon \mathbb{R}^2 \to \mathbb{R}, \quad \mathbf{x} \mapsto \langle \mathbf{x}, \boldsymbol{\theta} \rangle, \quad \mathbf{x} \in \mathbb{R}^2.$$

- with inverse $S_{\boldsymbol{\theta}}^{-1}(t) = \{t\boldsymbol{\theta} + \tau\boldsymbol{\theta}^{\perp} \mid \tau \in \mathbb{R}\} \subset \mathbb{R}^2$.
- Let $f \in \mathrm{L}^1(\mathbb{R}^2)$, its **Radon transform** $\mathcal{R}[f] \colon \mathbb{R} \times \mathbb{S}_1 \to \mathbb{R}$ is the line integral, where $\mathrm{d}s$ denotes the arc length of $S_{\boldsymbol{\theta}}^{-1}(t)$

$$\mathcal{R}[f](t, \boldsymbol{\theta}) := \int_{S_{\boldsymbol{\theta}}^{-1}(t)} f(s)\, \mathrm{d}s, \quad (t, \boldsymbol{\theta}) \in \mathbb{R} \times \mathbb{S}_1,$$




- For direction $\boldsymbol{\theta} \in \mathbb{S}_1$, the **restricted Radon transform** $\mathcal{R}_{\boldsymbol{\theta}}$ of signed, regular, finite measures $\mu \in \mathcal{M}(\mathbb{R}^2)$ is

$$\mathcal{R}_{\boldsymbol{\theta}} \colon \mathcal{M}(\mathbb{R}^2) \to \mathcal{M}(\mathbb{R}), \quad \mu \mapsto (S_{\boldsymbol{\theta}})_{\#}\mu = \mu \circ S_{\boldsymbol{\theta}}^{-1},$$

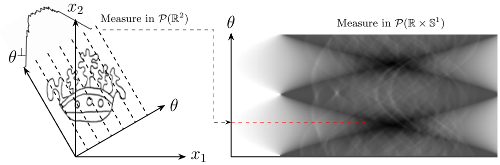
## 2D Radon Transfrom on Functions vs. on Measures

- For $\boldsymbol{\theta} \in \mathbb{S}_1 \coloneqq \{\mathbf{x} \in \mathbb{R}^2 \mid \|\mathbf{x}\| = 1\}$, define the slicing operator

$$S_{\boldsymbol{\theta}} \colon \mathbb{R}^2 \to \mathbb{R}, \quad \mathbf{x} \mapsto \langle \mathbf{x}, \boldsymbol{\theta} \rangle, \quad \mathbf{x} \in \mathbb{R}^2.$$

- with inverse $S_{\boldsymbol{\theta}}^{-1}(t) = \{t\boldsymbol{\theta} + \tau\boldsymbol{\theta}^\perp \mid \tau \in \mathbb{R}\} \subset \mathbb{R}^2$.
- Let $f \in \mathrm{L}^1(\mathbb{R}^2)$, its **Radon transform** $\mathcal{R}[f] \colon \mathbb{R} \times \mathbb{S}_1 \to \mathbb{R}$ is the line integral, where $\mathrm{d}s$ denotes the arc length of $S_{\boldsymbol{\theta}}^{-1}(t)$

$$\mathcal{R}[f](t, \boldsymbol{\theta}) \coloneqq \int_{S_{\boldsymbol{\theta}}^{-1}(t)} f(s)\,\mathrm{d}s, \quad (t, \boldsymbol{\theta}) \in \mathbb{R} \times \mathbb{S}_1,$$




- For direction $\boldsymbol{\theta} \in \mathbb{S}_1$, the **restricted Radon transform** $\mathcal{R}_{\boldsymbol{\theta}}$ of signed, regular, finite measures $\mu \in \mathcal{M}(\mathbb{R}^2)$ is

$$\mathcal{R}_{\boldsymbol{\theta}} \colon \mathcal{M}(\mathbb{R}^2) \to \mathcal{M}(\mathbb{R}), \quad \mu \mapsto (S_{\boldsymbol{\theta}})_{\#}\mu = \mu \circ S_{\boldsymbol{\theta}}^{-1},$$

- the *Radon transform* $\mathcal{R} \colon \mathcal{M}(\mathbb{R}^2) \to \mathcal{M}(\mathbb{R} \times \mathbb{S}_1)$ via $\mathcal{R}[\mu] \coloneqq \mathcal{I}_{\#}[\mu \times u_{\mathbb{S}_1}]$ with

$$\mathcal{I}(\mathbf{x}, \boldsymbol{\theta}) \coloneqq (S_{\boldsymbol{\theta}}(\mathbf{x}), \boldsymbol{\theta}) \qquad (\mathbf{x}, \boldsymbol{\theta}) \in \mathbb{R}^2 \times \mathbb{S}_1.$$

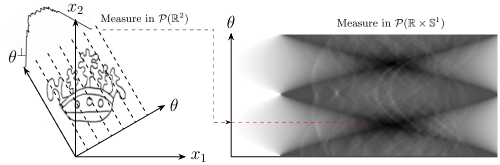
## 2D Radon Transfrom on Functions vs. on Measures

- For $\boldsymbol{\theta} \in \mathbb{S}_1 := \{\mathbf{x} \in \mathbb{R}^2 \mid \|\mathbf{x}\| = 1\}$, define the slicing operator

$$S_{\boldsymbol{\theta}} \colon \mathbb{R}^2 \to \mathbb{R}, \quad \mathbf{x} \mapsto \langle \mathbf{x}, \boldsymbol{\theta} \rangle, \quad \mathbf{x} \in \mathbb{R}^2.$$

- with inverse $S_{\boldsymbol{\theta}}^{-1}(t) = \{t\boldsymbol{\theta} + \tau\boldsymbol{\theta}^{\perp} \mid \tau \in \mathbb{R}\} \subset \mathbb{R}^2$.
- Let $f \in \mathrm{L}^1(\mathbb{R}^2)$, its **Radon transform** $\mathcal{R}[f] \colon \mathbb{R} \times \mathbb{S}_1 \to \mathbb{R}$ is the line integral, where $\mathrm{d}s$ denotes the arc length of $S_{\boldsymbol{\theta}}^{-1}(t)$

$$\mathcal{R}[f](t, \boldsymbol{\theta}) := \int_{S_{\boldsymbol{\theta}}^{-1}(t)} f(s) \, \mathrm{d}s, \quad (t, \boldsymbol{\theta}) \in \mathbb{R} \times \mathbb{S}_1,$$




- For direction $\boldsymbol{\theta} \in \mathbb{S}_1$, the **restricted Radon transform** $\mathcal{R}_{\boldsymbol{\theta}}$ of signed, regular, finite measures $\mu \in \mathcal{M}(\mathbb{R}^2)$ is

$$\mathcal{R}_{\boldsymbol{\theta}} \colon \mathcal{M}(\mathbb{R}^2) \to \mathcal{M}(\mathbb{R}), \quad \mu \mapsto (S_{\boldsymbol{\theta}})_{\#}\mu = \mu \circ S_{\boldsymbol{\theta}}^{-1},$$

- the *Radon transform* $\mathcal{R} \colon \mathcal{M}(\mathbb{R}^2) \to \mathcal{M}(\mathbb{R} \times \mathbb{S}_1)$ via $\mathcal{R}[\mu] := \mathcal{I}_{\#}[\mu \times u_{\mathbb{S}_1}]$ with

$$\mathcal{I}(\mathbf{x}, \boldsymbol{\theta}) := (S_{\boldsymbol{\theta}}(\mathbf{x}), \boldsymbol{\theta}) \qquad (\mathbf{x}, \boldsymbol{\theta}) \in \mathbb{R}^2 \times \mathbb{S}_1.$$

- Notice: $\mathcal{R}[f\lambda_{\mathbb{R}^2}] = \mathcal{R}[f]\sigma_{\mathbb{R} \times \mathbb{S}_1}$ for all $f \in \mathrm{L}^1(\mathbb{R}^2)$ w.r.t. the surface measure on $\mathbb{R} \times \mathbb{S}_1$.

## Properties of the 2D Radon Transform

**Proposition (Disintegration, Connection)**

## Properties of the 2D Radon Transform

**Proposition (Disintegration, Connection)**

*Let* $\mu \in \mathcal{M}(\mathbb{R}^2)$. *For* $g \in C_0(\mathbb{R} \times \mathbb{S}_1), h \in \mathrm{L}^\infty(\mathbb{R} \times \mathbb{S}_1)$

## Properties of the 2D Radon Transform

### Proposition (Disintegration, Connection)

*Let* $\mu \in \mathcal{M}(\mathbb{R}^2)$. *For* $g \in C_0(\mathbb{R} \times \mathbb{S}_1), h \in \mathrm{L}^\infty(\mathbb{R} \times \mathbb{S}_1)$ *holds*

$$\langle \mathcal{R}[\mu], g \rangle = \int_{\mathbb{S}_1} \langle \mathcal{R}_{\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \, \mathrm{d}u_{\mathbb{S}_1}(\boldsymbol{\theta})$$

$$\langle \mathcal{R}[\mu], h \rangle = \langle \mu, \mathcal{R}^*[h] \rangle, \quad \mathcal{R}^*[h] := \int_{\mathbb{S}_1} h(S(\cdot, \boldsymbol{\theta}), \boldsymbol{\theta}) \, \mathrm{d}u_{\mathbb{S}_1}(\boldsymbol{\theta})$$

## Properties of the 2D Radon Transform

### Proposition (Disintegration, Connection)

*Let* $\mu \in \mathcal{M}(\mathbb{R}^2)$. *For* $g \in C_0(\mathbb{R} \times \mathbb{S}_1), h \in \mathrm{L}^\infty(\mathbb{R} \times \mathbb{S}_1)$ *holds*

$$\langle \mathcal{R}[\mu], g \rangle = \int_{\mathbb{S}_1} \langle \mathcal{R}_{\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \, \mathrm{d}u_{\mathbb{S}_1}(\boldsymbol{\theta})$$

$$\langle \mathcal{R}[\mu], h \rangle = \langle \mu, \mathcal{R}^*[h] \rangle, \quad \mathcal{R}^*[h] := \int_{\mathbb{S}_1} h(S(\cdot, \boldsymbol{\theta}), \boldsymbol{\theta}) \, \mathrm{d}u_{\mathbb{S}_1}(\boldsymbol{\theta})$$

### Proposition (Affinely Transformed Measures)

## Properties of the 2D Radon Transform

### Proposition (Disintegration, Connection)

*Let* $\mu \in \mathcal{M}(\mathbb{R}^2)$. *For* $g \in C_0(\mathbb{R} \times \mathbb{S}_1), h \in \mathrm{L}^\infty(\mathbb{R} \times \mathbb{S}_1)$ *holds*

$$\langle \mathcal{R}[\mu], g \rangle = \int_{\mathbb{S}_1} \langle \mathcal{R}_{\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \, \mathrm{d}u_{\mathbb{S}_1}(\boldsymbol{\theta})$$

$$\langle \mathcal{R}[\mu], h \rangle = \langle \mu, \mathcal{R}^*[h] \rangle, \quad \mathcal{R}^*[h] := \int_{\mathbb{S}_1} h(S(\cdot, \boldsymbol{\theta}), \boldsymbol{\theta}) \, \mathrm{d}u_{\mathbb{S}_1}(\boldsymbol{\theta})$$

### Proposition (Affinely Transformed Measures)

*For any* $\boldsymbol{\theta} \in \mathbb{S}_1$, $\mathbf{A} \in \mathrm{GL}(2)$ *and* $\mathbf{y} \in \mathbb{R}^2$,

## Properties of the 2D Radon Transform

### Proposition (Disintegration, Connection)

*Let* $\mu \in \mathcal{M}(\mathbb{R}^2)$. *For* $g \in C_0(\mathbb{R} \times \mathbb{S}_1), h \in \mathrm{L}^\infty(\mathbb{R} \times \mathbb{S}_1)$ *holds*

$$\langle \mathcal{R}[\mu], g \rangle = \int_{\mathbb{S}_1} \langle \mathcal{R}_{\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \, \mathrm{d}u_{\mathbb{S}_1}(\boldsymbol{\theta})$$

$$\langle \mathcal{R}[\mu], h \rangle = \langle \mu, \mathcal{R}^*[h] \rangle, \quad \mathcal{R}^*[h] := \int_{\mathbb{S}_1} h(S(\cdot, \boldsymbol{\theta}), \boldsymbol{\theta}) \, \mathrm{d}u_{\mathbb{S}_1}(\boldsymbol{\theta})$$

### Proposition (Affinely Transformed Measures)

*For any* $\boldsymbol{\theta} \in \mathbb{S}_1, \mathbf{A} \in \mathrm{GL}(2)$ *and* $\mathbf{y} \in \mathbb{R}^2$, *and*

$$\mu_{\mathbf{A},\mathbf{y}} := (\mathbf{A} \cdot + \mathbf{y})_{\#}\mu = \mu \circ (\mathbf{A}^{-1}(\cdot - \mathbf{y}))$$

## Properties of the 2D Radon Transform

### Proposition (Disintegration, Connection)

*Let* $\mu \in \mathcal{M}(\mathbb{R}^2)$. *For* $g \in C_0(\mathbb{R} \times \mathbb{S}_1), h \in \mathrm{L}^\infty(\mathbb{R} \times \mathbb{S}_1)$ *holds*

$$\langle \mathcal{R}[\mu], g \rangle = \int_{\mathbb{S}_1} \langle \mathcal{R}_{\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \, \mathrm{d}u_{\mathbb{S}_1}(\boldsymbol{\theta})$$

$$\langle \mathcal{R}[\mu], h \rangle = \langle \mu, \mathcal{R}^*[h] \rangle, \quad \mathcal{R}^*[h] := \int_{\mathbb{S}_1} h(S(\cdot, \boldsymbol{\theta}), \boldsymbol{\theta}) \, \mathrm{d}u_{\mathbb{S}_1}(\boldsymbol{\theta})$$

### Proposition (Affinely Transformed Measures)

*For any* $\boldsymbol{\theta} \in \mathbb{S}_1, \mathbf{A} \in \mathrm{GL}(2)$ *and* $\mathbf{y} \in \mathbb{R}^2$, *and* $\mu_{\mathbf{A},\mathbf{y}} := (\mathbf{A} \cdot + \mathbf{y})_{\#}\mu = \mu \circ (\mathbf{A}^{-1}(\cdot - \mathbf{y}))$ *holds*

$$\begin{aligned}\mathcal{R}_{\boldsymbol{\theta}}[\mu_{\mathbf{A},\mathbf{y}}] &= (\|\mathbf{A}^\top \boldsymbol{\theta}\| \cdot + \langle \mathbf{y}, \boldsymbol{\theta} \rangle)_{\#} \mathcal{R}_{\mathbf{A}^\top \boldsymbol{\theta}/\|\mathbf{A}^\top \boldsymbol{\theta}\|}[\mu] \\ &= \mathcal{R}_{\mathbf{A}^\top \boldsymbol{\theta}/\|\mathbf{A}^\top \boldsymbol{\theta}\|}[\mu] \circ \left( \frac{\cdot - \langle \mathbf{y}, \boldsymbol{\theta} \rangle}{\|\mathbf{A}^\top \boldsymbol{\theta}\|} \right).\end{aligned}$$

# Properties of the 2D Radon Transform

## Proposition (Disintegration, Connection)

*Let* $\mu \in \mathcal{M}(\mathbb{R}^2)$. *For* $g \in C_0(\mathbb{R} \times \mathbb{S}_1), h \in \mathrm{L}^\infty(\mathbb{R} \times \mathbb{S}_1)$ *holds*

$$\langle \mathcal{R}[\mu], g \rangle = \int_{\mathbb{S}_1} \langle \mathcal{R}_{\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \, \mathrm{d}u_{\mathbb{S}_1}(\boldsymbol{\theta})$$

$$\langle \mathcal{R}[\mu], h \rangle = \langle \mu, \mathcal{R}^*[h] \rangle, \quad \mathcal{R}^*[h] := \int_{\mathbb{S}_1} h(S(\cdot, \boldsymbol{\theta}), \boldsymbol{\theta}) \, \mathrm{d}u_{\mathbb{S}_1}(\boldsymbol{\theta})$$

## Proposition (Affinely Transformed Measures)

*For any* $\boldsymbol{\theta} \in \mathbb{S}_1, \mathbf{A} \in \mathrm{GL}(2)$ *and* $\mathbf{y} \in \mathbb{R}^2$, *and* $\mu_{\mathbf{A},\mathbf{y}} := (\mathbf{A} \cdot + \mathbf{y})_\# \mu = \mu \circ (\mathbf{A}^{-1}(\cdot - \mathbf{y}))$ *holds*

$$\mathcal{R}_{\boldsymbol{\theta}}[\mu_{\mathbf{A},\mathbf{y}}] = (\|\mathbf{A}^\top \boldsymbol{\theta}\| \cdot + \langle \mathbf{y}, \boldsymbol{\theta} \rangle)_\# \mathcal{R}_{\mathbf{A}^\top \boldsymbol{\theta}/\|\mathbf{A}^\top \boldsymbol{\theta}\|}[\mu]$$
$$= \mathcal{R}_{\mathbf{A}^\top \boldsymbol{\theta}/\|\mathbf{A}^\top \boldsymbol{\theta}\|}[\mu] \circ \left( \frac{\cdot - \langle \mathbf{y}, \boldsymbol{\theta} \rangle}{\|\mathbf{A}^\top \boldsymbol{\theta}\|} \right).$$

| transformation | $\mathbf{A}$ | $\mathbf{y}$ | $\mathcal{R}_{\boldsymbol{\theta}(\vartheta)}[\mu_{\mathbf{A},\mathbf{y}}]$, $\vartheta \in (-\frac{\pi}{2}, \frac{\pi}{2})$ with $\boldsymbol{\theta}(\vartheta) := (\cos(\vartheta), \sin(\vartheta))^\top$ |
|---|---|---|---|
| translation | $\mathbf{I}$ | $\mathbb{R}^2$ | $\mathcal{R}_{\boldsymbol{\theta}(\vartheta)}[\mu] \circ (\cdot - \langle \mathbf{y}, \boldsymbol{\theta}(\vartheta) \rangle)$ |
| rotation | $\begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}$ | $\mathbf{0}$ | $\mathcal{R}_{\boldsymbol{\theta}(\vartheta - \varphi)}[\mu]$ |
| reflection | $\begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ \sin(\varphi) & -\cos(\varphi) \end{pmatrix}$ | $\mathbf{0}$ | $\mathcal{R}_{\boldsymbol{\theta}(\varphi - \vartheta)}[\mu]$ |
| anisotropic scaling | $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ | $\mathbf{0}$ | $\mathcal{R}_{\boldsymbol{\theta}(\arctan(\frac{b}{a}\tan(\vartheta)))}[\mu] \circ ([a^2 \cos^2(\vartheta) + b^2 \sin^2(\vartheta)]^{-1/2} \cdot)$ |
| vertical shear | $\begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}$ | $\mathbf{0}$ | $\mathcal{R}_{\boldsymbol{\theta}(\arctan(c + \tan(\vartheta)))}[\mu] \circ ([1 + c^2 \cos^2(\vartheta) + c \sin(2\vartheta)]^{-1/2} \cdot)$ |

## Generalized Radon Transform on Measures

- Let $\mathbb{X}$ be a **Polish space**,

## Generalized Radon Transform on Measures

– Let $\mathbb{X}$ be a **Polish space**, i.e. a separable completely metrizable topological space,

## Generalized Radon Transform on Measures

- Let $\mathbb{X}$ be a **Polish space**, i.e. a separable completely metrizable topological space, and let $\mathcal{M}(\mathbb{X})$ denotes the set of finite, signed, and regular (Borel) measures

## Generalized Radon Transform on Measures

– Let $\mathbb{X}$ be a **Polish space**, i.e. a separable completely metrizable topological space, and let $\mathcal{M}(\mathbb{X})$ denotes the set of finite, signed, and regular (Borel) measures and $\mathcal{P}(\mathbb{X})$ denotes the set of probability measures.

## Generalized Radon Transform on Measures

- Let $\mathbb{X}$ be a **Polish space**, i.e. a separable completely metrizable topological space, and let $\mathcal{M}(\mathbb{X})$ denotes the set of finite, signed, and regular (Borel) measures and $\mathcal{P}(\mathbb{X})$ denotes the set of probability measures.
- The **generalized slicing operator**

## Generalized Radon Transform on Measures

- Let $\mathbb{X}$ be a **Polish space**, i.e. a separable completely metrizable topological space, and let $\mathcal{M}(\mathbb{X})$ denotes the set of finite, signed, and regular (Borel) measures and $\mathcal{P}(\mathbb{X})$ denotes the set of probability measures.
- The **generalized slicing operator** with the **direction set** $\Theta$ and measurable **defining function** $\phi \colon \mathbb{X} \times \Theta \to \mathbb{R}$

## Generalized Radon Transform on Measures

- Let $\mathbb{X}$ be a **Polish space**, i.e. a separable completely metrizable topological space, and let $\mathcal{M}(\mathbb{X})$ denotes the set of finite, signed, and regular (Borel) measures and $\mathcal{P}(\mathbb{X})$ denotes the set of probability measures.
- The **generalized slicing operator** with the **direction set** $\Theta$ and measurable **defining function** $\phi\colon \mathbb{X} \times \Theta \to \mathbb{R}$ is defined by

$$S_{\phi,\boldsymbol{\theta}}\colon \mathbb{X} \to \mathbb{R}, \qquad \mathbf{x} \mapsto \phi(\mathbf{x}, \boldsymbol{\theta}), \qquad \boldsymbol{\theta} \in \Theta,$$

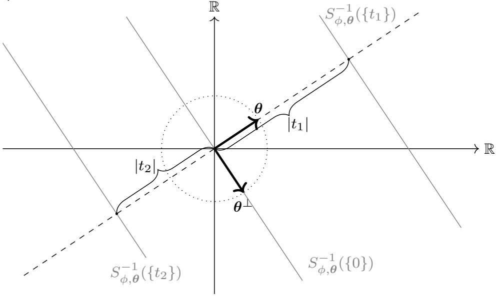
## Generalized Radon Transform on Measures

- Let $\mathbb{X}$ be a **Polish space**, i.e. a separable completely metrizable topological space, and let $\mathcal{M}(\mathbb{X})$ denotes the set of finite, signed, and regular (Borel) measures and $\mathcal{P}(\mathbb{X})$ denotes the set of probability measures.
- The **generalized slicing operator** with the **direction set** $\Theta$ and measurable **defining function** $\phi\colon \mathbb{X}\times\Theta\to\mathbb{R}$ is defined by

$$S_{\phi,\boldsymbol{\theta}}\colon \mathbb{X}\to\mathbb{R}, \qquad \mathbf{x}\mapsto\phi(\mathbf{x},\boldsymbol{\theta}), \qquad \boldsymbol{\theta}\in\Theta,$$

- The **generalized restricted Radon transform** as

$$\mathcal{R}_{\phi,\boldsymbol{\theta}}\colon \mathcal{M}(\mathbb{X})\to\mathcal{M}(\mathbb{R}), \qquad \mu\mapsto(S_{\phi,\boldsymbol{\theta}})_{\#}\mu \coloneqq \mu\circ S^{-1}_{\phi,\boldsymbol{\theta}}.$$



# Generalized Radon Transform on Measures

- Let $\mathbb{X}$ be a **Polish space**, i.e. a separable completely metrizable topological space, and let $\mathcal{M}(\mathbb{X})$ denotes the set of finite, signed, and regular (Borel) measures and $\mathcal{P}(\mathbb{X})$ denotes the set of probability measures.
- The **generalized slicing operator** with the **direction set** $\Theta$ and measurable **defining function** $\phi \colon \mathbb{X} \times \Theta \to \mathbb{R}$ is defined by

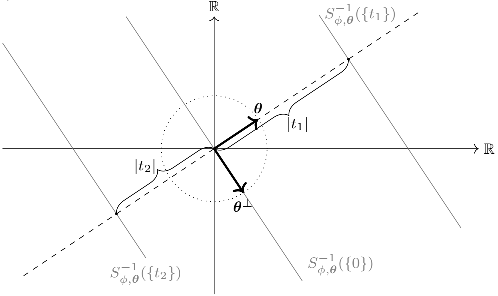
$$S_{\phi,\boldsymbol{\theta}} \colon \mathbb{X} \to \mathbb{R}, \qquad \mathbf{x} \mapsto \phi(\mathbf{x}, \boldsymbol{\theta}), \qquad \boldsymbol{\theta} \in \Theta,$$

- The **generalized restricted Radon transform** as

$$\mathcal{R}_{\phi,\boldsymbol{\theta}} \colon \mathcal{M}(\mathbb{X}) \to \mathcal{M}(\mathbb{R}), \qquad \mu \mapsto (S_{\phi,\boldsymbol{\theta}})_{\#}\mu := \mu \circ S^{-1}_{\phi,\boldsymbol{\theta}}.$$

- The **gluing measure** $\gamma \in \mathcal{P}(\Theta)$ and **gluing map**

$$\mathcal{I}_{\phi} \colon \mathbb{X} \times \Theta \to \mathbb{R} \times \Theta, \quad (\mathbf{x}, \boldsymbol{\theta}) \mapsto (S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \boldsymbol{\theta}).$$



# Generalized Radon Transform on Measures

- Let $\mathbb{X}$ be a **Polish space**, i.e. a separable completely metrizable topological space, and let $\mathcal{M}(\mathbb{X})$ denotes the set of finite, signed, and regular (Borel) measures and $\mathcal{P}(\mathbb{X})$ denotes the set of probability measures.
- The **generalized slicing operator** with the **direction set** $\Theta$ and measurable **defining function** $\phi \colon \mathbb{X} \times \Theta \to \mathbb{R}$ is defined by

$$S_{\phi,\boldsymbol{\theta}} \colon \mathbb{X} \to \mathbb{R}, \qquad \mathbf{x} \mapsto \phi(\mathbf{x}, \boldsymbol{\theta}), \qquad \boldsymbol{\theta} \in \Theta,$$

- The **generalized restricted Radon transform** as

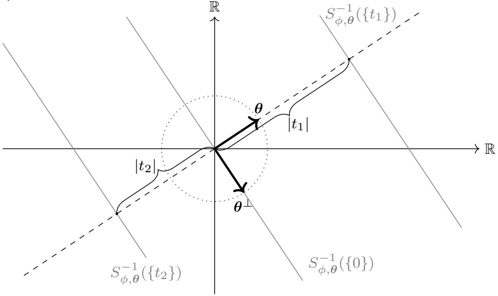
$$\mathcal{R}_{\phi,\boldsymbol{\theta}} \colon \mathcal{M}(\mathbb{X}) \to \mathcal{M}(\mathbb{R}), \qquad \mu \mapsto (S_{\phi,\boldsymbol{\theta}})_{\#}\mu \coloneqq \mu \circ S_{\phi,\boldsymbol{\theta}}^{-1}.$$

- The **gluing measure** $\gamma \in \mathcal{P}(\Theta)$ and **gluing map**

$$\mathcal{I}_{\phi} \colon \mathbb{X} \times \Theta \to \mathbb{R} \times \Theta, \quad (\mathbf{x}, \boldsymbol{\theta}) \mapsto (S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \boldsymbol{\theta}).$$

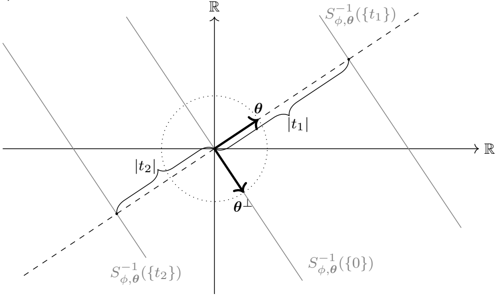
- The **generalized Radon transform** if $\Theta$ is a Polish space, too, by

$$\mathcal{R}_{\phi,\gamma} \colon \mathcal{M}(\mathbb{X}) \to \mathcal{M}(\mathbb{R} \times \Theta), \qquad \mu \mapsto (\mathcal{I}_{\phi})_{\#}[\mu \times \gamma].$$



## Generalized Radon Transform on Measures

- Let $\mathbb{X}$ be a **Polish space**, i.e. a separable completely metrizable topological space, and let $\mathcal{M}(\mathbb{X})$ denotes the set of finite, signed, and regular (Borel) measures and $\mathcal{P}(\mathbb{X})$ denotes the set of probability measures.
- The **generalized slicing operator** with the **direction set** $\Theta$ and measurable **defining function** $\phi\colon \mathbb{X} \times \Theta \to \mathbb{R}$ is defined by
$$S_{\phi,\boldsymbol{\theta}}\colon \mathbb{X} \to \mathbb{R}, \qquad \mathbf{x} \mapsto \phi(\mathbf{x}, \boldsymbol{\theta}), \qquad \boldsymbol{\theta} \in \Theta,$$
- The **generalized restricted Radon transform** as
$$\mathcal{R}_{\phi,\boldsymbol{\theta}}\colon \mathcal{M}(\mathbb{X}) \to \mathcal{M}(\mathbb{R}), \qquad \mu \mapsto (S_{\phi,\boldsymbol{\theta}})_{\#}\mu := \mu \circ S^{-1}_{\phi,\boldsymbol{\theta}}.$$
- The **gluing measure** $\gamma \in \mathcal{P}(\Theta)$ and **gluing map**
$$\mathcal{I}_\phi\colon \mathbb{X} \times \Theta \to \mathbb{R} \times \Theta, \quad (\mathbf{x}, \boldsymbol{\theta}) \mapsto (S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \boldsymbol{\theta}).$$
- The **generalized Radon transform** if $\Theta$ is a Polish space, too, by
$$\mathcal{R}_{\phi,\gamma}\colon \mathcal{M}(\mathbb{X}) \to \mathcal{M}(\mathbb{R} \times \Theta), \qquad \mu \mapsto (\mathcal{I}_\phi)_{\#}[\mu \times \gamma].$$
- Restricting the domain of the generalized (restricted) Radon transform to probability measures on $\mathbb{X}$ yields a map towards probability measures on $\mathbb{R} \times \Theta$ (respectively $\mathbb{R}$).



# Properties of the Generalized Radon Transform

## Proposition (Disintegration, Connection)

## Properties of the Generalized Radon Transform

**Proposition (Disintegration, Connection)**

*Let* $\mu \in \mathcal{M}(\mathbb{X})$. *Then for all continuous functions* $g \in C_0(\mathbb{R} \times \Theta)$ *vanishing at infinity*

## Properties of the Generalized Radon Transform

**Proposition (Disintegration, Connection)**

*Let* $\mu \in \mathcal{M}(\mathbb{X})$. *Then for all continuous functions* $g \in C_0(\mathbb{R} \times \Theta)$ *vanishing at infinity holds*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], g \rangle = \int_{\Theta} \langle \mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \, \mathrm{d}\gamma(\boldsymbol{\theta}).$$

## Properties of the Generalized Radon Transform

**Proposition (Disintegration, Connection)**

*Let* $\mu \in \mathcal{M}(\mathbb{X})$. *Then for all continuous functions* $g \in C_0(\mathbb{R} \times \Theta)$ *vanishing at infinity holds*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], g \rangle = \int_{\Theta} \langle \mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \,\mathrm{d}\gamma(\boldsymbol{\theta}).$$

**Proposition (Backprojection, Adjoint)**

## Properties of the Generalized Radon Transform

### Proposition (Disintegration, Connection)

*Let* $\mu \in \mathcal{M}(\mathbb{X})$*. Then for all continuous functions* $g \in C_0(\mathbb{R} \times \Theta)$ *vanishing at infinity holds*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], g \rangle = \int_\Theta \langle \mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \, \mathrm{d}\gamma(\boldsymbol{\theta}).$$

### Proposition (Backprojection, Adjoint)

$$\mathcal{R}^*_{\phi,\gamma}[g](\mathbf{x}) := \int_\Theta h(S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \theta) \, \mathrm{d}\gamma(\boldsymbol{\theta}), \quad \mathbf{x} \in \mathbb{X}.$$

# Properties of the Generalized Radon Transform

## Proposition (Disintegration, Connection)

*Let* $\mu \in \mathcal{M}(\mathbb{X})$. *Then for all continuous functions* $g \in C_0(\mathbb{R} \times \Theta)$ *vanishing at infinity holds*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], g \rangle = \int_\Theta \langle \mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \, \mathrm{d}\gamma(\boldsymbol{\theta}).$$

## Proposition (Backprojection, Adjoint)

$$\mathcal{R}^*_{\phi,\gamma}[g](\mathbf{x}) := \int_\Theta h(S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \theta) \, \mathrm{d}\gamma(\boldsymbol{\theta}), \quad \mathbf{x} \in \mathbb{X}.$$

*For essentially bounded measurable function* $h$ *i.e.*

$$\exists C > 0 \text{ s.t. } |h(t, \boldsymbol{\theta})| \leq C \text{ a.e. } (t, \boldsymbol{\theta}) \in \mathbb{R} \times \Theta,$$

## Properties of the Generalized Radon Transform

### Proposition (Disintegration, Connection)

*Let $\mu \in \mathcal{M}(\mathbb{X})$. Then for all continuous functions $g \in C_0(\mathbb{R} \times \Theta)$ vanishing at infinity holds*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], g \rangle = \int_{\Theta} \langle \mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \,\mathrm{d}\gamma(\boldsymbol{\theta}).$$

### Proposition (Backprojection, Adjoint)

$$\mathcal{R}^*_{\phi,\gamma}[g](\mathbf{x}) := \int_{\Theta} h(S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \theta) \,\mathrm{d}\gamma(\boldsymbol{\theta}), \quad \mathbf{x} \in \mathbb{X}.$$

*For essentially bounded measurable function $h$ i.e.*

$$\exists C > 0 \text{ s.t. } |h(t, \boldsymbol{\theta})| \leq C \text{ a.e. } (t, \boldsymbol{\theta}) \in \mathbb{R} \times \Theta,$$

*then*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], h \rangle = \langle \mu, \mathcal{R}^*_{\phi,\gamma}[h] \rangle \;\forall\, h \in \mathrm{L}^{\infty}_{\lambda_{\mathbb{R}} \times \gamma}(\mathbb{R} \times \Theta).$$

## Properties of the Generalized Radon Transform

### Proposition (Disintegration, Connection)

*Let $\mu \in \mathcal{M}(\mathbb{X})$. Then for all continuous functions $g \in C_0(\mathbb{R} \times \Theta)$ vanishing at infinity holds*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], g \rangle = \int_\Theta \langle \mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \, \mathrm{d}\gamma(\boldsymbol{\theta}).$$

### Proposition (Backprojection, Adjoint)

$$\mathcal{R}^*_{\phi,\gamma}[g](\mathbf{x}) := \int_\Theta h(S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \theta) \, \mathrm{d}\gamma(\boldsymbol{\theta}), \quad \mathbf{x} \in \mathbb{X}.$$

*For essentially bounded measurable function $h$ i.e.*

$$\exists C > 0 \text{ s.t. } |h(t, \boldsymbol{\theta})| \leq C \text{ a.e. } (t, \boldsymbol{\theta}) \in \mathbb{R} \times \Theta,$$

*then*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], h \rangle = \langle \mu, \mathcal{R}^*_{\phi,\gamma}[h] \rangle \; \forall \, h \in \mathrm{L}^\infty_{\lambda_{\mathbb{R}} \times \gamma}(\mathbb{R} \times \Theta).$$

### Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$)

# Properties of the Generalized Radon Transform

## Proposition (Disintegration, Connection)

*Let* $\mu \in \mathcal{M}(\mathbb{X})$. *Then for all continuous functions* $g \in C_0(\mathbb{R} \times \Theta)$ *vanishing at infinity holds*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], g \rangle = \int_\Theta \langle \mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \, \mathrm{d}\gamma(\boldsymbol{\theta}).$$

## Proposition (Backprojection, Adjoint)

$$\mathcal{R}^*_{\phi,\gamma}[g](\mathbf{x}) := \int_\Theta h(S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \theta) \, \mathrm{d}\gamma(\boldsymbol{\theta}), \quad \mathbf{x} \in \mathbb{X}.$$

*For essentially bounded measurable function* $h$ *i.e.*

$$\exists C > 0 \text{ s.t. } |h(t, \boldsymbol{\theta})| \leq C \text{ a.e. } (t, \boldsymbol{\theta}) \in \mathbb{R} \times \Theta,$$

*then*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], h \rangle = \langle \mu, \mathcal{R}^*_{\phi,\gamma}[h] \rangle \; \forall \, h \in \mathrm{L}^\infty_{\lambda_\mathbb{R} \times \gamma}(\mathbb{R} \times \Theta).$$

## Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$)

*For any* $\boldsymbol{\theta} \in \mathbb{S}_{d-1}, \mathbf{A} \in \mathrm{GL}(d), \mathbf{y} \in \mathbb{R}^d$,

## Properties of the Generalized Radon Transform

### Proposition (Disintegration, Connection)

*Let $\mu \in \mathcal{M}(\mathbb{X})$. Then for all continuous functions $g \in C_0(\mathbb{R} \times \Theta)$ vanishing at infinity holds*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], g \rangle = \int_\Theta \langle \mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \, \mathrm{d}\gamma(\boldsymbol{\theta}).$$

### Proposition (Backprojection, Adjoint)

$$\mathcal{R}^*_{\phi,\gamma}[g](\mathbf{x}) := \int_\Theta h(S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \theta) \, \mathrm{d}\gamma(\boldsymbol{\theta}), \quad \mathbf{x} \in \mathbb{X}.$$

*For essentially bounded measurable function $h$ i.e.*

$$\exists C > 0 \text{ s.t. } |h(t, \boldsymbol{\theta})| \le C \text{ a.e. } (t, \boldsymbol{\theta}) \in \mathbb{R} \times \Theta,$$

*then*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], h \rangle = \langle \mu, \mathcal{R}^*_{\phi,\gamma}[h] \rangle \ \forall\, h \in \mathrm{L}^\infty_{\lambda_\mathbb{R} \times \gamma}(\mathbb{R} \times \Theta).$$

### Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$)

*For any $\boldsymbol{\theta} \in \mathbb{S}_{d-1}, \mathbf{A} \in \mathrm{GL}(d), \mathbf{y} \in \mathbb{R}^d$, using $\phi(\mathbf{x}, \boldsymbol{\theta}) := \langle \mathbf{x}, \boldsymbol{\theta} \rangle$*

## Properties of the Generalized Radon Transform

**Proposition (Disintegration, Connection)**

*Let $\mu \in \mathcal{M}(\mathbb{X})$. Then for all continuous functions $g \in C_0(\mathbb{R} \times \Theta)$ vanishing at infinity holds*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], g \rangle = \int_\Theta \langle \mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu], g(\cdot,\boldsymbol{\theta}) \rangle \,\mathrm{d}\gamma(\boldsymbol{\theta}).$$

**Proposition (Backprojection, Adjoint)**

$$\mathcal{R}^*_{\phi,\gamma}[g](\mathbf{x}) := \int_\Theta h(S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \theta) \,\mathrm{d}\gamma(\boldsymbol{\theta}), \quad \mathbf{x} \in \mathbb{X}.$$

*For essentially bounded measurable function $h$ i.e.*

$$\exists C > 0 \text{ s.t. } |h(t,\boldsymbol{\theta})| \le C \text{ a.e. } (t,\boldsymbol{\theta}) \in \mathbb{R} \times \Theta,$$

*then*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], h \rangle = \langle \mu, \mathcal{R}^*_{\phi,\gamma}[h] \rangle \;\forall\, h \in \mathrm{L}^\infty_{\lambda_\mathbb{R} \times \gamma}(\mathbb{R} \times \Theta).$$

**Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$)**

*For any $\boldsymbol{\theta} \in \mathbb{S}_{d-1}, \mathbf{A} \in \mathrm{GL}(d), \mathbf{y} \in \mathbb{R}^d$, using $\phi(\mathbf{x},\boldsymbol{\theta}) := \langle \mathbf{x}, \boldsymbol{\theta} \rangle$ and $\mu_{\mathbf{A},\mathbf{y}} := (\mathbf{A} \cdot + \mathbf{y})_\# \mu$*

# Properties of the Generalized Radon Transform

**Proposition (Disintegration, Connection)**

*Let $\mu \in \mathcal{M}(\mathbb{X})$. Then for all continuous functions $g \in C_0(\mathbb{R} \times \Theta)$ vanishing at infinity holds*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], g\rangle = \int_\Theta \langle \mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu], g(\cdot,\boldsymbol{\theta})\rangle \, \mathrm{d}\gamma(\boldsymbol{\theta}).$$

**Proposition (Backprojection, Adjoint)**

$$\mathcal{R}^*_{\phi,\gamma}[g](\mathbf{x}) := \int_\Theta h(S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \theta)\, \mathrm{d}\gamma(\boldsymbol{\theta}), \quad \mathbf{x} \in \mathbb{X}.$$

*For essentially bounded measurable function $h$ i.e.*

$$\exists C > 0 \text{ s.t. } |h(t,\boldsymbol{\theta})| \le C \text{ a.e. } (t,\boldsymbol{\theta}) \in \mathbb{R} \times \Theta,$$

*then*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], h\rangle = \langle \mu, \mathcal{R}^*_{\phi,\gamma}[h]\rangle \; \forall\, h \in \mathrm{L}^\infty_{\lambda_\mathbb{R}\times\gamma}(\mathbb{R}\times\Theta).$$

**Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$)**

*For any $\boldsymbol{\theta} \in \mathbb{S}_{d-1}, \mathbf{A} \in \mathrm{GL}(d), \mathbf{y} \in \mathbb{R}^d$, using $\phi(\mathbf{x},\boldsymbol{\theta}) := \langle \mathbf{x}, \boldsymbol{\theta}\rangle$ and $\mu_{\mathbf{A},\mathbf{y}} := (\mathbf{A}\cdot + \mathbf{y})_\# \mu$ holds*

$$\mathcal{R}_{\boldsymbol{\theta}}[\mu_{\mathbf{A},\mathbf{y}}] = (\|\mathbf{A}^\top\boldsymbol{\theta}\| \cdot + \langle \mathbf{y}, \boldsymbol{\theta}\rangle)_\# \mathcal{R}_{\mathbf{A}^\top\boldsymbol{\theta}/\|\mathbf{A}^\top\boldsymbol{\theta}\|}[\mu]$$

# Properties of the Generalized Radon Transform

## Proposition (Disintegration, Connection)

*Let $\mu \in \mathcal{M}(\mathbb{X})$. Then for all continuous functions $g \in C_0(\mathbb{R} \times \Theta)$ vanishing at infinity holds*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], g \rangle = \int_\Theta \langle \mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \,\mathrm{d}\gamma(\boldsymbol{\theta}).$$

## Proposition (Backprojection, Adjoint)

$$\mathcal{R}^*_{\phi,\gamma}[g](\mathbf{x}) := \int_\Theta h(S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \theta) \,\mathrm{d}\gamma(\boldsymbol{\theta}), \quad \mathbf{x} \in \mathbb{X}.$$

*For essentially bounded measurable function $h$ i.e.*

$$\exists C > 0 \text{ s.t. } |h(t, \boldsymbol{\theta})| \leq C \text{ a.e. } (t, \boldsymbol{\theta}) \in \mathbb{R} \times \Theta,$$

*then*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], h \rangle = \langle \mu, \mathcal{R}^*_{\phi,\gamma}[h] \rangle \;\forall\, h \in \mathrm{L}^\infty_{\lambda_\mathbb{R} \times \gamma}(\mathbb{R} \times \Theta).$$

## Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$)

*For any $\boldsymbol{\theta} \in \mathbb{S}_{d-1}, \mathbf{A} \in \mathrm{GL}(d), \mathbf{y} \in \mathbb{R}^d$, using $\phi(\mathbf{x}, \boldsymbol{\theta}) := \langle \mathbf{x}, \boldsymbol{\theta} \rangle$ and $\mu_{\mathbf{A},\mathbf{y}} := (\mathbf{A} \cdot + \mathbf{y})_\# \mu$ holds*

$$\mathcal{R}_{\boldsymbol{\theta}}[\mu_{\mathbf{A},\mathbf{y}}] = (\|\mathbf{A}^\top \boldsymbol{\theta}\| \cdot + \langle \mathbf{y}, \boldsymbol{\theta} \rangle)_\# \mathcal{R}_{\mathbf{A}^\top \boldsymbol{\theta} / \|\mathbf{A}^\top \boldsymbol{\theta}\|}[\mu]$$

## Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{R}^d$)

*For any $\boldsymbol{\theta} \in \mathbb{R}^d, \mathbf{Q} \in \mathrm{O}(d), \mathbf{y} \in \mathbb{R}^d$,*

# Properties of the Generalized Radon Transform

### Proposition (Disintegration, Connection)

*Let* $\mu \in \mathcal{M}(\mathbb{X})$. *Then for all continuous functions* $g \in C_0(\mathbb{R} \times \Theta)$ *vanishing at infinity holds*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], g \rangle = \int_\Theta \langle \mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \, \mathrm{d}\gamma(\boldsymbol{\theta}).$$

### Proposition (Backprojection, Adjoint)

$$\mathcal{R}^*_{\phi,\gamma}[g](\mathbf{x}) := \int_\Theta h(S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \theta) \, \mathrm{d}\gamma(\boldsymbol{\theta}), \quad \mathbf{x} \in \mathbb{X}.$$

*For essentially bounded measurable function* $h$ *i.e.*

$$\exists C > 0 \text{ s.t. } |h(t, \boldsymbol{\theta})| \leq C \text{ a.e. } (t, \boldsymbol{\theta}) \in \mathbb{R} \times \Theta,$$

*then*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], h \rangle = \langle \mu, \mathcal{R}^*_{\phi,\gamma}[h] \rangle \; \forall \, h \in \mathrm{L}^\infty_{\lambda_\mathbb{R} \times \gamma}(\mathbb{R} \times \Theta).$$

### Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$)

*For any* $\boldsymbol{\theta} \in \mathbb{S}_{d-1}, \mathbf{A} \in \mathrm{GL}(d), \mathbf{y} \in \mathbb{R}^d$, *using* $\phi(\mathbf{x}, \boldsymbol{\theta}) := \langle \mathbf{x}, \boldsymbol{\theta} \rangle$ *and* $\mu_{\mathbf{A},\mathbf{y}} := (\mathbf{A} \cdot + \mathbf{y})_\# \mu$ *holds*

$$\mathcal{R}_{\boldsymbol{\theta}}[\mu_{\mathbf{A},\mathbf{y}}] = (\|\mathbf{A}^\top \boldsymbol{\theta}\| \cdot + \langle \mathbf{y}, \boldsymbol{\theta} \rangle)_\# \mathcal{R}_{\mathbf{A}^\top \boldsymbol{\theta} / \|\mathbf{A}^\top \boldsymbol{\theta}\|}[\mu]$$

### Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{R}^d$)

*For any* $\boldsymbol{\theta} \in \mathbb{R}^d, \mathbf{Q} \in \mathrm{O}(d), \mathbf{y} \in \mathbb{R}^d$, *using* $\phi(\mathbf{x}, \boldsymbol{\theta}) := \|\mathbf{x} - \boldsymbol{\theta}\|$,

## Properties of the Generalized Radon Transform

### Proposition (Disintegration, Connection)

*Let* $\mu \in \mathcal{M}(\mathbb{X})$. *Then for all continuous functions* $g \in C_0(\mathbb{R} \times \Theta)$ *vanishing at infinity holds*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], g \rangle = \int_\Theta \langle \mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \, \mathrm{d}\gamma(\boldsymbol{\theta}).$$

### Proposition (Backprojection, Adjoint)

$$\mathcal{R}^*_{\phi,\gamma}[g](\mathbf{x}) := \int_\Theta h(S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \theta) \, \mathrm{d}\gamma(\boldsymbol{\theta}), \quad \mathbf{x} \in \mathbb{X}.$$

*For essentially bounded measurable function* $h$ *i.e.*

$$\exists C > 0 \text{ s.t. } |h(t, \boldsymbol{\theta})| \le C \text{ a.e. } (t, \boldsymbol{\theta}) \in \mathbb{R} \times \Theta,$$

*then*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], h \rangle = \langle \mu, \mathcal{R}^*_{\phi,\gamma}[h] \rangle \; \forall \, h \in \mathrm{L}^\infty_{\lambda_\mathbb{R} \times \gamma}(\mathbb{R} \times \Theta).$$

### Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$)

*For any* $\boldsymbol{\theta} \in \mathbb{S}_{d-1}, \mathbf{A} \in \mathrm{GL}(d), \mathbf{y} \in \mathbb{R}^d$, *using* $\phi(\mathbf{x}, \boldsymbol{\theta}) := \langle \mathbf{x}, \boldsymbol{\theta} \rangle$ *and* $\mu_{\mathbf{A},\mathbf{y}} := (\mathbf{A} \cdot + \mathbf{y})_\# \mu$ *holds*

$$\mathcal{R}_{\boldsymbol{\theta}}[\mu_{\mathbf{A},\mathbf{y}}] = (\|\mathbf{A}^\top \boldsymbol{\theta}\| \cdot + \langle \mathbf{y}, \boldsymbol{\theta} \rangle)_\# \mathcal{R}_{\mathbf{A}^\top \boldsymbol{\theta} / \|\mathbf{A}^\top \boldsymbol{\theta}\|}[\mu]$$

### Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{R}^d$)

*For any* $\boldsymbol{\theta} \in \mathbb{R}^d, \mathbf{Q} \in \mathrm{O}(d), \mathbf{y} \in \mathbb{R}^d$, *using* $\phi(\mathbf{x}, \boldsymbol{\theta}) := \|\mathbf{x} - \boldsymbol{\theta}\|$, *and* $\mu_{s\mathbf{Q},\mathbf{y}} := (s\mathbf{Q} \cdot + \mathbf{y})_\# \mu$

## Properties of the Generalized Radon Transform

### Proposition (Disintegration, Connection)

*Let $\mu \in \mathcal{M}(\mathbb{X})$. Then for all continuous functions $g \in C_0(\mathbb{R} \times \Theta)$ vanishing at infinity holds*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], g \rangle = \int_\Theta \langle \mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \, \mathrm{d}\gamma(\boldsymbol{\theta}).$$

### Proposition (Backprojection, Adjoint)

$$\mathcal{R}^*_{\phi,\gamma}[g](\mathbf{x}) := \int_\Theta h(S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \theta) \, \mathrm{d}\gamma(\boldsymbol{\theta}), \quad \mathbf{x} \in \mathbb{X}.$$

*For essentially bounded measurable function $h$ i.e.*

$$\exists C > 0 \text{ s.t. } |h(t, \boldsymbol{\theta})| \leq C \text{ a.e. } (t, \boldsymbol{\theta}) \in \mathbb{R} \times \Theta,$$

*then*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], h \rangle = \langle \mu, \mathcal{R}^*_{\phi,\gamma}[h] \rangle \; \forall\, h \in \mathrm{L}^\infty_{\lambda_\mathbb{R} \times \gamma}(\mathbb{R} \times \Theta).$$

### Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$)

*For any $\boldsymbol{\theta} \in \mathbb{S}_{d-1}, \mathbf{A} \in \mathrm{GL}(d), \mathbf{y} \in \mathbb{R}^d$, using $\phi(\mathbf{x}, \boldsymbol{\theta}) := \langle \mathbf{x}, \boldsymbol{\theta} \rangle$ and $\mu_{\mathbf{A},\mathbf{y}} := (\mathbf{A} \cdot + \mathbf{y})_\# \mu$ holds*

$$\mathcal{R}_{\boldsymbol{\theta}}[\mu_{\mathbf{A},\mathbf{y}}] = (\|\mathbf{A}^\top \boldsymbol{\theta}\| \cdot + \langle \mathbf{y}, \boldsymbol{\theta} \rangle)_\# \mathcal{R}_{\mathbf{A}^\top \boldsymbol{\theta} / \|\mathbf{A}^\top \boldsymbol{\theta}\|}[\mu]$$

### Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{R}^d$)

*For any $\boldsymbol{\theta} \in \mathbb{R}^d, \mathbf{Q} \in \mathrm{O}(d), \mathbf{y} \in \mathbb{R}^d$, using $\phi(\mathbf{x}, \boldsymbol{\theta}) := \|\mathbf{x} - \boldsymbol{\theta}\|$, and $\mu_{s\mathbf{Q},\mathbf{y}} := (s\mathbf{Q} \cdot + \mathbf{y})_\# \mu$ holds*

$$\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu_{s\mathbf{Q},\mathbf{y}}] = (s\, \cdot)_\# \mathcal{R}_{\phi, \mathbf{Q}^\top(\boldsymbol{\theta} - \mathbf{y})}[\mu].$$

# Properties of the Generalized Radon Transform

## Proposition (Disintegration, Connection)

*Let $\mu \in \mathcal{M}(\mathbb{X})$. Then for all continuous functions $g \in C_0(\mathbb{R} \times \Theta)$ vanishing at infinity holds*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], g \rangle = \int_\Theta \langle \mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \, \mathrm{d}\gamma(\boldsymbol{\theta}).$$

## Proposition (Backprojection, Adjoint)

$$\mathcal{R}^*_{\phi,\gamma}[g](\mathbf{x}) := \int_\Theta h(S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \theta) \, \mathrm{d}\gamma(\boldsymbol{\theta}), \quad \mathbf{x} \in \mathbb{X}.$$

*For essentially bounded measurable function $h$ i.e.*

$$\exists C > 0 \text{ s.t. } |h(t, \boldsymbol{\theta})| \leq C \text{ a.e. } (t, \boldsymbol{\theta}) \in \mathbb{R} \times \Theta,$$

*then*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], h \rangle = \langle \mu, \mathcal{R}^*_{\phi,\gamma}[h] \rangle \; \forall\, h \in \mathrm{L}^\infty_{\lambda_\mathbb{R} \times \gamma}(\mathbb{R} \times \Theta).$$

## Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$)

*For any $\boldsymbol{\theta} \in \mathbb{S}_{d-1}, \mathbf{A} \in \mathrm{GL}(d), \mathbf{y} \in \mathbb{R}^d$, using $\phi(\mathbf{x}, \boldsymbol{\theta}) := \langle \mathbf{x}, \boldsymbol{\theta} \rangle$ and $\mu_{\mathbf{A},\mathbf{y}} := (\mathbf{A} \cdot + \mathbf{y})_\# \mu$ holds*

$$\mathcal{R}_{\boldsymbol{\theta}}[\mu_{\mathbf{A},\mathbf{y}}] = (\|\mathbf{A}^\top \boldsymbol{\theta}\| \cdot + \langle \mathbf{y}, \boldsymbol{\theta} \rangle)_\# \mathcal{R}_{\mathbf{A}^\top \boldsymbol{\theta} / \|\mathbf{A}^\top \boldsymbol{\theta}\|}[\mu]$$

## Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{R}^d$)

*For any $\boldsymbol{\theta} \in \mathbb{R}^d, \mathbf{Q} \in \mathrm{O}(d), \mathbf{y} \in \mathbb{R}^d$, using $\phi(\mathbf{x}, \boldsymbol{\theta}) := \|\mathbf{x} - \boldsymbol{\theta}\|$, and $\mu_{s\mathbf{Q},\mathbf{y}} := (s\mathbf{Q} \cdot + \mathbf{y})_\# \mu$ holds*

$$\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu_{s\mathbf{Q},\mathbf{y}}] = (s\, \cdot)_\# \mathcal{R}_{\phi, \mathbf{Q}^\top(\boldsymbol{\theta} - \mathbf{y})}[\mu].$$

## Proposition (Affine Transformation $\mathbb{X} = \Theta = \mathrm{SO}(3)$)

## Properties of the Generalized Radon Transform

### Proposition (Disintegration, Connection)

*Let $\mu \in \mathcal{M}(\mathbb{X})$. Then for all continuous functions $g \in C_0(\mathbb{R} \times \Theta)$ vanishing at infinity holds*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], g \rangle = \int_\Theta \langle \mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \, \mathrm{d}\gamma(\boldsymbol{\theta}).$$

### Proposition (Backprojection, Adjoint)

$$\mathcal{R}^*_{\phi,\gamma}[g](\mathbf{x}) := \int_\Theta h(S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \theta) \, \mathrm{d}\gamma(\boldsymbol{\theta}), \quad \mathbf{x} \in \mathbb{X}.$$

*For essentially bounded measurable function $h$ i.e.*

$$\exists C > 0 \text{ s.t. } |h(t, \boldsymbol{\theta})| \leq C \text{ a.e. } (t, \boldsymbol{\theta}) \in \mathbb{R} \times \Theta,$$

*then*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], h \rangle = \langle \mu, \mathcal{R}^*_{\phi,\gamma}[h] \rangle \; \forall \, h \in \mathrm{L}^\infty_{\lambda_\mathbb{R} \times \gamma}(\mathbb{R} \times \Theta).$$

### Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$)

*For any $\boldsymbol{\theta} \in \mathbb{S}_{d-1}, \mathbf{A} \in \mathrm{GL}(d), \mathbf{y} \in \mathbb{R}^d$, using $\phi(\mathbf{x}, \boldsymbol{\theta}) := \langle \mathbf{x}, \boldsymbol{\theta} \rangle$ and $\mu_{\mathbf{A},\mathbf{y}} := (\mathbf{A} \cdot + \mathbf{y})_\# \mu$ holds*

$$\mathcal{R}_{\boldsymbol{\theta}}[\mu_{\mathbf{A},\mathbf{y}}] = (\|\mathbf{A}^\top \boldsymbol{\theta}\| \cdot + \langle \mathbf{y}, \boldsymbol{\theta} \rangle)_\# \mathcal{R}_{\mathbf{A}^\top \boldsymbol{\theta} / \|\mathbf{A}^\top \boldsymbol{\theta}\|}[\mu]$$

### Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{R}^d$)

*For any $\boldsymbol{\theta} \in \mathbb{R}^d, \mathbf{Q} \in \mathrm{O}(d), \mathbf{y} \in \mathbb{R}^d$, using $\phi(\mathbf{x}, \boldsymbol{\theta}) := \|\mathbf{x} - \boldsymbol{\theta}\|$, and $\mu_{s\mathbf{Q},\mathbf{y}} := (s\mathbf{Q} \cdot + \mathbf{y})_\# \mu$ holds*

$$\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu_{s\mathbf{Q},\mathbf{y}}] = (s \, \cdot)_\# \mathcal{R}_{\phi,\mathbf{Q}^\top(\boldsymbol{\theta} - \mathbf{y})}[\mu].$$

### Proposition (Affine Transformation $\mathbb{X} = \Theta = \mathrm{SO}(3)$)

*For any $\boldsymbol{\theta}, \mathbf{R} \in \mathrm{SO}(3)$,*

## Properties of the Generalized Radon Transform

### Proposition (Disintegration, Connection)

*Let* $\mu \in \mathcal{M}(\mathbb{X})$. *Then for all continuous functions* $g \in C_0(\mathbb{R} \times \Theta)$ *vanishing at infinity holds*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], g \rangle = \int_\Theta \langle \mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \, \mathrm{d}\gamma(\boldsymbol{\theta}).$$

### Proposition (Backprojection, Adjoint)

$$\mathcal{R}^*_{\phi,\gamma}[g](\mathbf{x}) := \int_\Theta h(S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \theta) \, \mathrm{d}\gamma(\boldsymbol{\theta}), \quad \mathbf{x} \in \mathbb{X}.$$

*For essentially bounded measurable function* $h$ *i.e.*

$$\exists C > 0 \text{ s.t. } |h(t, \boldsymbol{\theta})| \le C \text{ a.e. } (t, \boldsymbol{\theta}) \in \mathbb{R} \times \Theta,$$

*then*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], h \rangle = \langle \mu, \mathcal{R}^*_{\phi,\gamma}[h] \rangle \; \forall\, h \in \mathrm{L}^\infty_{\lambda_\mathbb{R} \times \gamma}(\mathbb{R} \times \Theta).$$

### Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$)

*For any* $\boldsymbol{\theta} \in \mathbb{S}_{d-1}, \mathbf{A} \in \mathrm{GL}(d), \mathbf{y} \in \mathbb{R}^d$, *using* $\phi(\mathbf{x}, \boldsymbol{\theta}) := \langle \mathbf{x}, \boldsymbol{\theta} \rangle$ *and* $\mu_{\mathbf{A},\mathbf{y}} := (\mathbf{A} \cdot + \mathbf{y})_\# \mu$ *holds*

$$\mathcal{R}_{\boldsymbol{\theta}}[\mu_{\mathbf{A},\mathbf{y}}] = (\|\mathbf{A}^\top \boldsymbol{\theta}\| \cdot + \langle \mathbf{y}, \boldsymbol{\theta} \rangle)_\# \mathcal{R}_{\mathbf{A}^\top \boldsymbol{\theta} / \|\mathbf{A}^\top \boldsymbol{\theta}\|}[\mu]$$

### Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{R}^d$)

*For any* $\boldsymbol{\theta} \in \mathbb{R}^d, \mathbf{Q} \in \mathrm{O}(d), \mathbf{y} \in \mathbb{R}^d$, *using* $\phi(\mathbf{x}, \boldsymbol{\theta}) := \|\mathbf{x} - \boldsymbol{\theta}\|$, *and* $\mu_{s\mathbf{Q},\mathbf{y}} := (s\mathbf{Q} \cdot + \mathbf{y})_\# \mu$ *holds*

$$\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu_{s\mathbf{Q},\mathbf{y}}] = (s \cdot)_\# \mathcal{R}_{\phi, \mathbf{Q}^\top(\boldsymbol{\theta} - \mathbf{y})}[\mu].$$

### Proposition (Affine Transformation $\mathbb{X} = \Theta = \mathrm{SO}(3)$)

*For any* $\boldsymbol{\theta}, \mathbf{R} \in \mathrm{SO}(3)$, *using* $\phi(\mathbf{x}, \boldsymbol{\theta}) := \arccos\left(\frac{\mathrm{tr}(\boldsymbol{\theta}^\top \mathbf{x}) - 1}{2}\right)$,

## Properties of the Generalized Radon Transform

### Proposition (Disintegration, Connection)

*Let* $\mu \in \mathcal{M}(\mathbb{X})$. *Then for all continuous functions* $g \in C_0(\mathbb{R} \times \Theta)$ *vanishing at infinity holds*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], g \rangle = \int_\Theta \langle \mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu], g(\cdot,\boldsymbol{\theta}) \rangle \, \mathrm{d}\gamma(\boldsymbol{\theta}).$$

### Proposition (Backprojection, Adjoint)

$$\mathcal{R}^*_{\phi,\gamma}[g](\mathbf{x}) := \int_\Theta h(S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \theta) \, \mathrm{d}\gamma(\boldsymbol{\theta}), \quad \mathbf{x} \in \mathbb{X}.$$

*For essentially bounded measurable function* $h$ *i.e.*

$$\exists C > 0 \text{ s.t. } |h(t,\boldsymbol{\theta})| \le C \text{ a.e. } (t,\boldsymbol{\theta}) \in \mathbb{R} \times \Theta,$$

*then*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], h \rangle = \langle \mu, \mathcal{R}^*_{\phi,\gamma}[h] \rangle \; \forall \, h \in \mathrm{L}^\infty_{\lambda_\mathbb{R} \times \gamma}(\mathbb{R} \times \Theta).$$

### Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$)

*For any* $\boldsymbol{\theta} \in \mathbb{S}_{d-1}, \mathbf{A} \in \mathrm{GL}(d), \mathbf{y} \in \mathbb{R}^d$, *using* $\phi(\mathbf{x},\boldsymbol{\theta}) := \langle \mathbf{x}, \boldsymbol{\theta} \rangle$ *and* $\mu_{\mathbf{A},\mathbf{y}} := (\mathbf{A} \cdot + \mathbf{y})_\# \mu$ *holds*

$$\mathcal{R}_{\boldsymbol{\theta}}[\mu_{\mathbf{A},\mathbf{y}}] = (\|\mathbf{A}^\top \boldsymbol{\theta}\| \cdot + \langle \mathbf{y}, \boldsymbol{\theta} \rangle)_\# \mathcal{R}_{\mathbf{A}^\top \boldsymbol{\theta} / \|\mathbf{A}^\top \boldsymbol{\theta}\|}[\mu]$$

### Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{R}^d$)

*For any* $\boldsymbol{\theta} \in \mathbb{R}^d, \mathbf{Q} \in \mathrm{O}(d), \mathbf{y} \in \mathbb{R}^d$, *using* $\phi(\mathbf{x},\boldsymbol{\theta}) := \|\mathbf{x} - \boldsymbol{\theta}\|$, *and* $\mu_{s\mathbf{Q},\mathbf{y}} := (s\mathbf{Q} \cdot + \mathbf{y})_\# \mu$ *holds*

$$\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu_{s\mathbf{Q},\mathbf{y}}] = (s \, \cdot)_\# \mathcal{R}_{\phi,\mathbf{Q}^\top(\boldsymbol{\theta} - \mathbf{y})}[\mu].$$

### Proposition (Affine Transformation $\mathbb{X} = \Theta = \mathrm{SO}(3)$)

*For any* $\boldsymbol{\theta}, \mathbf{R} \in \mathrm{SO}(3)$, *using* $\phi(\mathbf{x},\boldsymbol{\theta}) := \arccos\left(\frac{\mathrm{tr}(\boldsymbol{\theta}^\top \mathbf{x}) - 1}{2}\right)$, *and* $\mu_\mathbf{R} := (\mathbf{R} \, \cdot)_\# \mu$

# Properties of the Generalized Radon Transform

### Proposition (Disintegration, Connection)

*Let* $\mu \in \mathcal{M}(\mathbb{X})$. *Then for all continuous functions* $g \in C_0(\mathbb{R} \times \Theta)$ *vanishing at infinity holds*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], g \rangle = \int_\Theta \langle \mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu], g(\cdot, \boldsymbol{\theta}) \rangle \, \mathrm{d}\gamma(\boldsymbol{\theta}).$$

### Proposition (Backprojection, Adjoint)

$$\mathcal{R}^*_{\phi,\gamma}[g](\mathbf{x}) := \int_\Theta h(S_{\phi,\boldsymbol{\theta}}(\mathbf{x}), \theta) \, \mathrm{d}\gamma(\boldsymbol{\theta}), \quad \mathbf{x} \in \mathbb{X}.$$

*For essentially bounded measurable function* $h$ *i.e.*

$$\exists C > 0 \text{ s.t. } |h(t, \boldsymbol{\theta})| \le C \text{ a.e. } (t, \boldsymbol{\theta}) \in \mathbb{R} \times \Theta,$$

*then*

$$\langle \mathcal{R}_{\phi,\gamma}[\mu], h \rangle = \langle \mu, \mathcal{R}^*_{\phi,\gamma}[h] \rangle \; \forall \, h \in \mathrm{L}^\infty_{\lambda_\mathbb{R} \times \gamma}(\mathbb{R} \times \Theta).$$

### Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$)

*For any* $\boldsymbol{\theta} \in \mathbb{S}_{d-1}, \mathbf{A} \in \mathrm{GL}(d), \mathbf{y} \in \mathbb{R}^d$, *using* $\phi(\mathbf{x}, \boldsymbol{\theta}) := \langle \mathbf{x}, \boldsymbol{\theta} \rangle$ *and* $\mu_{\mathbf{A},\mathbf{y}} := (\mathbf{A} \cdot + \mathbf{y})_\# \mu$ *holds*

$$\mathcal{R}_{\boldsymbol{\theta}}[\mu_{\mathbf{A},\mathbf{y}}] = (\|\mathbf{A}^\top \boldsymbol{\theta}\| \cdot + \langle \mathbf{y}, \boldsymbol{\theta} \rangle)_\# \mathcal{R}_{\mathbf{A}^\top \boldsymbol{\theta} / \|\mathbf{A}^\top \boldsymbol{\theta}\|}[\mu]$$

### Proposition (Affine Transformation $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{R}^d$)

*For any* $\boldsymbol{\theta} \in \mathbb{R}^d, \mathbf{Q} \in \mathrm{O}(d), \mathbf{y} \in \mathbb{R}^d$, *using* $\phi(\mathbf{x}, \boldsymbol{\theta}) := \|\mathbf{x} - \boldsymbol{\theta}\|$, *and* $\mu_{s\mathbf{Q},\mathbf{y}} := (s\mathbf{Q} \cdot + \mathbf{y})_\# \mu$ *holds*

$$\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu_{s\mathbf{Q},\mathbf{y}}] = (s \, \cdot)_\# \mathcal{R}_{\phi, \mathbf{Q}^\top(\boldsymbol{\theta} - \mathbf{y})}[\mu].$$

### Proposition (Affine Transformation $\mathbb{X} = \Theta = \mathrm{SO}(3)$)

*For any* $\boldsymbol{\theta}, \mathbf{R} \in \mathrm{SO}(3)$, *using* $\phi(\mathbf{x}, \boldsymbol{\theta}) := \arccos\left(\frac{\mathrm{tr}(\boldsymbol{\theta}^\top \mathbf{x}) - 1}{2}\right)$, *and* $\mu_\mathbf{R} := (\mathbf{R} \, \cdot)_\# \mu$ *holds*

$$\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu_\mathbf{R}] = \mathcal{R}_{\phi, \mathbf{R}^\top \boldsymbol{\theta}}[\mu].$$

# Optimal Transport on the Reel Line – Domain of the Restricted, Generalized Radon Projection

- **Aim:** Generalize the R-CDT[11], for probability measures $\mathcal{P}(\mathbb{R}) \subset \mathcal{M}(\mathbb{R})$, similar to [12].

[12] Diaz Martin, R., Medri, I. V., and Rohde, G. K. (2024) – Data representation with optimal transport, **arXiv:2406.15503**
[13] Villani, C. (2003) – Topics in Optimal Transportation,

## Optimal Transport on the Reel Line – Domain of the Restricted, Generalized Radon Projection

- **Aim:** Generalize the R-CDT[11], for probability measures $\mathcal{P}(\mathbb{R}) \subset \mathcal{M}(\mathbb{R})$, similar to [12].
- The **cumulative distribution function** $F_\mu \colon \mathbb{R} \to [0,1]$ is given by $F_\mu(t) := \mu((-\infty, t])$,

[12] Diaz Martin, R., Medri, I. V., and Rohde, G. K. (2024) – Data representation with optimal transport, **arXiv:2406.15503**
[13] Villani, C. (2003) – Topics in Optimal Transportation,

# Optimal Transport on the Reel Line – Domain of the Restricted, Generalized Radon Projection

- **Aim:** Generalize the R-CDT[11], for probability measures $\mathcal{P}(\mathbb{R}) \subset \mathcal{M}(\mathbb{R})$, similar to [12].
- The **cumulative distribution function** $F_\mu \colon \mathbb{R} \to [0,1]$ is given by $F_\mu(t) := \mu((-\infty, t])$,
- its generalized inverse (**quantile function**) is given by

$$F_\mu^{[-1]}(t) := \inf\{s \in \mathbb{R} \mid F_\mu(s) > t\}, \quad t \in \mathbb{R}.$$

[12] Diaz Martin, R., Medri, I. V., and Rohde, G. K. (2024) – Data representation with optimal transport, **arXiv:2406.15503**
[13] Villani, C. (2003) – Topics in Optimal Transportation,

## Optimal Transport on the Reel Line – Domain of the Restricted, Generalized Radon Projection

- **Aim:** Generalize the R-CDT[11], for probability measures $\mathcal{P}(\mathbb{R}) \subset \mathcal{M}(\mathbb{R})$, similar to [12].
- The **cumulative distribution function** $F_\mu\colon \mathbb{R} \to [0,1]$ is given by $F_\mu(t) \coloneqq \mu((-\infty, t])$,
- its generalized inverse (**quantile function**) is given by

$$F_\mu^{[-1]}(t) \coloneqq \inf\{s \in \mathbb{R} \mid F_\mu(s) > t\}, \quad t \in \mathbb{R}.$$

- Based on $\rho \in \mathcal{P}(\mathbb{R})$ (not give mass to atoms), define the **cumulative distribution transform (CDT)**

$$\widehat{\mu}\colon \mathbb{R} \to \mathbb{R}, \qquad \widehat{\mu} \coloneqq F_\mu^{[-1]} \circ F_\rho.$$

[12] Diaz Martin, R., Medri, I. V., and Rohde, G. K. (2024) – Data representation with optimal transport, **arXiv:2406.15503**
[13] Villani, C. (2003) – Topics in Optimal Transportation,

# Optimal Transport on the Reel Line – Domain of the Restricted, Generalized Radon Projection

- **Aim:** Generalize the R-CDT[11], for probability measures $\mathcal{P}(\mathbb{R}) \subset \mathcal{M}(\mathbb{R})$, similar to [12].
- The **cumulative distribution function** $F_\mu : \mathbb{R} \to [0,1]$ is given by $F_\mu(t) := \mu((-\infty, t])$,
- its generalized inverse (**quantile function**) is given by

$$F_\mu^{[-1]}(t) := \inf\{s \in \mathbb{R} \mid F_\mu(s) > t\}, \quad t \in \mathbb{R}.$$

- Based on $\rho \in \mathcal{P}(\mathbb{R})$ (not give mass to atoms), define the **cumulative distribution transform (CDT)**

$$\widehat{\mu} \colon \mathbb{R} \to \mathbb{R}, \qquad \widehat{\mu} := F_\mu^{[-1]} \circ F_\rho.$$

→ For convex cost $c \colon \mathbb{R} \to [0, \infty)$, the CDT solves the Monge–Kantorovich transportation problem [13]

$$\widehat{\mu} = \operatorname*{arg\,min}_{T_\# \rho = \mu} \int_{\mathbb{R}} c(s - T(s)) \,\mathrm{d}\rho(t), \qquad \text{where } T \text{ is measurable.}$$

[12] Diaz Martin, R., Medri, I. V., and Rohde, G. K. (2024) – Data representation with optimal transport, **arXiv:2406.15503**
[13] Villani, C. (2003) – Topics in Optimal Transportation,

# Optimal Transport on the Reel Line – Domain of the Restricted, Generalized Radon Projection

- **Aim:** Generalize the R-CDT[11], for probability measures $\mathcal{P}(\mathbb{R}) \subset \mathcal{M}(\mathbb{R})$, similar to [12].
- The **cumulative distribution function** $F_\mu : \mathbb{R} \to [0,1]$ is given by $F_\mu(t) := \mu((-\infty, t])$,
- its generalized inverse (**quantile function**) is given by

$$F_\mu^{[-1]}(t) := \inf\{s \in \mathbb{R} \mid F_\mu(s) > t\}, \quad t \in \mathbb{R}.$$

- Based on $\rho \in \mathcal{P}(\mathbb{R})$ (not give mass to atoms), define the **cumulative distribution transform (CDT)**

$$\widehat{\mu} \colon \mathbb{R} \to \mathbb{R}, \qquad \widehat{\mu} := F_\mu^{[-1]} \circ F_\rho.$$

→ For convex cost $c \colon \mathbb{R} \to [0, \infty)$, the CDT solves the Monge–Kantorovich transportation problem [13]

$$\widehat{\mu} = \operatorname*{arg\,min}_{T_\# \rho = \mu} \int_{\mathbb{R}} c(s - T(s))\, \mathrm{d}\rho(t), \qquad \text{where } T \text{ is measurable.}$$

→ $\mu \in \mathcal{P}_2(\mathbb{R}) = \{\mu \in \mathcal{P}(\mathbb{R}) \mid \int_{\mathbb{R}} t \,\mathrm{d}\mu(t) < \infty\}$, then $\widehat{\mu} \in \mathrm{L}^2_\rho(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \mid \text{measurable}, \int_{\mathbb{R}} |f(t)|^2 \,\mathrm{d}\rho(t) < \infty\}$.

---


[12] Diaz Martin, R., Medri, I. V., and Rohde, G. K. (2024) – Data representation with optimal transport, **arXiv:2406.15503**
[13] Villani, C. (2003) – Topics in Optimal Transportation,

## Optimal Transport on the Reel Line – Domain of the Restricted, Generalized Radon Projection

– **Aim:** Generalize the R-CDT[11], for probability measures $\mathcal{P}(\mathbb{R}) \subset \mathcal{M}(\mathbb{R})$, similar to [12].

– The **cumulative distribution function** $F_\mu \colon \mathbb{R} \to [0,1]$ is given by $F_\mu(t) := \mu((-\infty, t])$,

– its generalized inverse (**quantile function**) is given by

$$F_\mu^{[-1]}(t) := \inf\{s \in \mathbb{R} \mid F_\mu(s) > t\}, \quad t \in \mathbb{R}.$$

– Based on $\rho \in \mathcal{P}(\mathbb{R})$ (not give mass to atoms), define the **cumulative distribution transform (CDT)**

$$\widehat{\mu} \colon \mathbb{R} \to \mathbb{R}, \qquad \widehat{\mu} := F_\mu^{[-1]} \circ F_\rho.$$

→ For convex cost $c \colon \mathbb{R} \to [0, \infty)$, the CDT solves the Monge–Kantorovich transportation problem [13]

$$\widehat{\mu} = \operatorname*{arg\,min}_{T_\# \rho = \mu} \int_{\mathbb{R}} c(s - T(s)) \, \mathrm{d}\rho(t), \qquad \text{where } T \text{ is measurable.}$$

→ $\mu \in \mathcal{P}_2(\mathbb{R}) = \{\mu \in \mathcal{P}(\mathbb{R}) \mid \int_{\mathbb{R}} t \, \mathrm{d}\mu(t) < \infty\}$, then $\widehat{\mu} \in \mathrm{L}^2_\rho(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \mid \text{measurable}, \int_{\mathbb{R}} |f(t)|^2 \, \mathrm{d}\rho(t) < \infty\}$.

→ For $\mu, \nu \in \mathcal{P}_2(\mathbb{R})$ the Wasserstein-2 distance[4] fulfills

$$W_2^2(\mu, \nu) := \inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{R}^2} c(x-y)^2 \, \mathrm{d}\pi(x,y) = \|\widehat{\mu} - \widehat{\nu}\|_\rho^2 := \int_{\mathbb{R}} |\widehat{\mu}(t) - \widehat{\nu}(t)|^2 \, \mathrm{d}\rho(t).$$

[12] Diaz Martin, R., Medri, I. V., and Rohde, G. K. (2024) – Data representation with optimal transport, **arXiv:2406.15503**
[13] Villani, C. (2003) – Topics in Optimal Transportation,

# Generalized Normalized Radon CDTs – Feature Extractors – Normalization & Sorting

- We define the **generalized R-CDT** by $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t).$

## Generalized Normalized Radon CDTs – Feature Extractors – Normalization & Sorting

– We define the **generalized R-CDT** by $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t)$.

→ For $\mu \in \mathcal{P}_2(\mathbb{X})$, it is $\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] \in \mathcal{P}_2(\mathbb{R})$ and consequently $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} \in \mathrm{L}^2_{\rho\times\gamma}(\mathbb{R} \times \Theta)$

# Generalized Normalized Radon CDTs – Feature Extractors – Normalization & Sorting

- We define the **generalized R-CDT** by $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t)$.

→ For $\mu \in \mathcal{P}_2(\mathbb{X})$, it is $\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] \in \mathcal{P}_2(\mathbb{R})$ and consequently $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} \in \mathrm{L}^2_{\rho\times\gamma}(\mathbb{R}\times\Theta)$

- with the sliced Wasserstein distance of $\mu, \nu \in \mathcal{P}_2(\mathbb{X})$ given by

$$\|\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} - \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu]}\|^2_{\rho\times\gamma} := \int_\Theta \int_\mathbb{R} |\widehat{\mathcal{R}_{\phi,\gamma}[\mu]}(t,\boldsymbol{\theta}) - \widehat{\mathcal{R}_{\phi,\gamma}[\nu]}(t,\boldsymbol{\theta})|^2 \,\mathrm{d}\rho(t)\,\mathrm{d}\gamma(\boldsymbol{\theta}) = \mathrm{SW}_2^2(\mu,\nu) = \int_\Theta \mathrm{W}_2^2(\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] - \mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu])\,\mathrm{d}\gamma(\boldsymbol{\theta}).$$

# Generalized Normalized Radon CDTs – Feature Extractors – Normalization & Sorting

- We define the **generalized R-CDT** by $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t)$.

→ For $\mu \in \mathcal{P}_2(\mathbb{X})$, it is $\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] \in \mathcal{P}_2(\mathbb{R})$ and consequently $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} \in \mathrm{L}^2_{\rho\times\gamma}(\mathbb{R}\times\Theta)$

- with the sliced Wasserstein distance of $\mu, \nu \in \mathcal{P}_2(\mathbb{X})$ given by

$$\|\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} - \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu]}\|^2_{\rho\times\gamma} \coloneqq \int_\Theta \int_\mathbb{R} |\widehat{\mathcal{R}_{\phi,\gamma}[\mu]}(t,\boldsymbol{\theta}) - \widehat{\mathcal{R}_{\phi,\gamma}[\nu]}(t,\boldsymbol{\theta})|^2 \,\mathrm{d}\rho(t)\,\mathrm{d}\gamma(\boldsymbol{\theta}) = \mathrm{SW}_2^2(\mu,\nu) = \int_\Theta \mathrm{W}_2^2(\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] - \mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu])\,\mathrm{d}\gamma(\boldsymbol{\theta}).$$

---

1) **Normalization:** The **generalized NR-CDT** $\mathcal{N}_\phi[\mu](t,\boldsymbol{\theta}) \coloneqq \mathcal{N}_{\phi,\boldsymbol{\theta}}[\mu](t)$, for $t \in \mathbb{R}, \boldsymbol{\theta} \in \Theta$, is defined by

$$\mathcal{N}_{\phi,\boldsymbol{\theta}}[\mu](t) \coloneqq \frac{\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t) - \mathrm{mean}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]})}{\mathrm{std}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]})}, \qquad t \in \mathbb{R}.$$

---

## Generalized Normalized Radon CDTs – Feature Extractors – Normalization & Sorting

- We define the **generalized R-CDT** by $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t)$.

→ For $\mu \in \mathcal{P}_2(\mathbb{X})$, it is $\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] \in \mathcal{P}_2(\mathbb{R})$ and consequently $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} \in \mathrm{L}^2_{\rho\times\gamma}(\mathbb{R}\times\Theta)$

- with the sliced Wasserstein distance of $\mu, \nu \in \mathcal{P}_2(\mathbb{X})$ given by

$$\|\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} - \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu]}\|^2_{\rho\times\gamma} \coloneqq \int_\Theta \int_\mathbb{R} |\widehat{\mathcal{R}_{\phi,\gamma}[\mu]}(t,\boldsymbol{\theta}) - \widehat{\mathcal{R}_{\phi,\gamma}[\nu]}(t,\boldsymbol{\theta})|^2 \,\mathrm{d}\rho(t)\,\mathrm{d}\gamma(\boldsymbol{\theta}) = \mathrm{SW}_2^2(\mu,\nu) = \int_\Theta \mathrm{W}_2^2(\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] - \mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu])\,\mathrm{d}\gamma(\boldsymbol{\theta}).$$

---

1) **Normalization:** The **generalized NR-CDT** $\mathcal{N}_\phi[\mu](t,\boldsymbol{\theta}) \coloneqq \mathcal{N}_{\phi,\boldsymbol{\theta}}[\mu](t)$, for $t \in \mathbb{R}, \boldsymbol{\theta} \in \Theta$, is defined by

$$\mathcal{N}_{\phi,\boldsymbol{\theta}}[\mu](t) \coloneqq \frac{\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t) - \mathrm{mean}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]})}{\mathrm{std}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]})}, \qquad t \in \mathbb{R}.$$

---

- For **well-definiteness**: restrict to **uniformly bounded** $\phi$:

## Generalized Normalized Radon CDTs – Feature Extractors – Normalization & Sorting

- We define the **generalized R-CDT** by $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t)$.

→ For $\mu \in \mathcal{P}_2(\mathbb{X})$, it is $\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] \in \mathcal{P}_2(\mathbb{R})$ and consequently $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} \in \mathrm{L}^2_{\rho\times\gamma}(\mathbb{R} \times \Theta)$

- with the sliced Wasserstein distance of $\mu, \nu \in \mathcal{P}_2(\mathbb{X})$ given by

$$\|\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} - \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu]}\|^2_{\rho\times\gamma} \coloneqq \int_\Theta \int_\mathbb{R} |\widehat{\mathcal{R}_{\phi,\gamma}[\mu]}(t,\boldsymbol{\theta}) - \widehat{\mathcal{R}_{\phi,\gamma}[\nu]}(t,\boldsymbol{\theta})|^2 \,\mathrm{d}\rho(t)\,\mathrm{d}\gamma(\boldsymbol{\theta}) = \mathrm{SW}_2^2(\mu,\nu) = \int_\Theta \mathrm{W}_2^2(\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] - \mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu])\,\mathrm{d}\gamma(\boldsymbol{\theta}).$$

---

1) **Normalization:** The **generalized NR-CDT** $\mathcal{N}_\phi[\mu](t,\boldsymbol{\theta}) \coloneqq \mathcal{N}_{\phi,\boldsymbol{\theta}}[\mu](t)$, for $t \in \mathbb{R}, \boldsymbol{\theta} \in \Theta$, is defined by

$$\mathcal{N}_{\phi,\boldsymbol{\theta}}[\mu](t) \coloneqq \frac{\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t) - \mathrm{mean}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]})}{\mathrm{std}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]})}, \qquad t \in \mathbb{R}.$$

---

- For **well-definiteness**: restrict to **uniformly bounded** $\phi$: $|\phi(\mathbf{x},\boldsymbol{\theta}) - \phi(\mathbf{y},\boldsymbol{\theta})| \le M_L \;\; \forall\, d_\mathbb{X}(\mathbf{x},\mathbf{y}) \le L, \forall\, \boldsymbol{\theta} \in \Theta$,

# Generalized Normalized Radon CDTs – Feature Extractors – Normalization & Sorting

- We define the **generalized R-CDT** by $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t)$.

→ For $\mu \in \mathcal{P}_2(\mathbb{X})$, it is $\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] \in \mathcal{P}_2(\mathbb{R})$ and consequently $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} \in \mathrm{L}^2_{\rho\times\gamma}(\mathbb{R}\times\Theta)$

- with the sliced Wasserstein distance of $\mu, \nu \in \mathcal{P}_2(\mathbb{X})$ given by

$$\|\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} - \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu]}\|^2_{\rho\times\gamma} := \int_\Theta \int_\mathbb{R} |\widehat{\mathcal{R}_{\phi,\gamma}[\mu]}(t,\boldsymbol{\theta}) - \widehat{\mathcal{R}_{\phi,\gamma}[\nu]}(t,\boldsymbol{\theta})|^2 \,\mathrm{d}\rho(t)\,\mathrm{d}\gamma(\boldsymbol{\theta}) = \mathrm{SW}_2^2(\mu,\nu) = \int_\Theta \mathrm{W}_2^2(\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] - \mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu])\,\mathrm{d}\gamma(\boldsymbol{\theta}).$$

---

1) **Normalization:** The **generalized NR-CDT** $\mathcal{N}_\phi[\mu](t,\boldsymbol{\theta}) := \mathcal{N}_{\phi,\boldsymbol{\theta}}[\mu](t)$, for $t \in \mathbb{R}, \boldsymbol{\theta} \in \Theta$, is defined by

$$\mathcal{N}_{\phi,\boldsymbol{\theta}}[\mu](t) := \frac{\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t) - \mathrm{mean}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]})}{\mathrm{std}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]})}, \qquad t \in \mathbb{R}.$$

---

- For **well-definiteness**: restrict to **uniformly bounded** $\phi$: $|\phi(\mathbf{x},\boldsymbol{\theta}) - \phi(\mathbf{y},\boldsymbol{\theta})| \le M_L \;\; \forall\, d_\mathbb{X}(\mathbf{x},\mathbf{y}) \le L, \forall\, \boldsymbol{\theta} \in \Theta$,
- and consider **measures** $\mu$ in $\mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) := \{\mu \in \mathcal{P}(\mathbb{X}) \mid \mathrm{supp}(\mu) \subset\subset \mathbb{X} \;\; \wedge \;\; \exists\, c > 0 : \mathrm{std}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}) \ge c \;\forall\, \theta \in \Theta\}$.

## Generalized Normalized Radon CDTs – Feature Extractors – Normalization & Sorting

– We define the **generalized R-CDT** by $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t)$.

→ For $\mu \in \mathcal{P}_2(\mathbb{X})$, it is $\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] \in \mathcal{P}_2(\mathbb{R})$ and consequently $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} \in \mathrm{L}^2_{\rho\times\gamma}(\mathbb{R}\times\Theta)$

– with the sliced Wasserstein distance of $\mu, \nu \in \mathcal{P}_2(\mathbb{X})$ given by

$$\|\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} - \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu]}\|^2_{\rho\times\gamma} := \int_\Theta \int_\mathbb{R} |\widehat{\mathcal{R}_{\phi,\gamma}[\mu]}(t,\boldsymbol{\theta}) - \widehat{\mathcal{R}_{\phi,\gamma}[\nu]}(t,\boldsymbol{\theta})|^2 \,\mathrm{d}\rho(t)\,\mathrm{d}\gamma(\boldsymbol{\theta}) = \mathrm{SW}_2^2(\mu,\nu) = \int_\Theta \mathrm{W}_2^2(\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] - \mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu])\,\mathrm{d}\gamma(\boldsymbol{\theta}).$$

---

1) **Normalization:** The **generalized NR-CDT** $\mathcal{N}_\phi[\mu](t,\boldsymbol{\theta}) := \mathcal{N}_{\phi,\boldsymbol{\theta}}[\mu](t)$, for $t \in \mathbb{R}, \boldsymbol{\theta} \in \Theta$, is defined by

$$\mathcal{N}_{\phi,\boldsymbol{\theta}}[\mu](t) := \frac{\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t) - \mathrm{mean}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]})}{\mathrm{std}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]})}, \qquad t \in \mathbb{R}.$$

---

– For **well-definiteness**: restrict to **uniformly bounded** $\phi$: $|\phi(\mathbf{x},\boldsymbol{\theta}) - \phi(\mathbf{y},\boldsymbol{\theta})| \le M_L \;\; \forall\, d_\mathbb{X}(\mathbf{x},\mathbf{y}) \le L, \forall \boldsymbol{\theta} \in \Theta$,

– and consider **measures** $\mu$ in $\mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) := \{\mu \in \mathcal{P}(\mathbb{X}) \mid \mathrm{supp}(\mu) \subset\subset \mathbb{X} \;\; \wedge \;\; \exists\, c > 0 : \mathrm{std}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}) \ge c \;\forall\, \theta \in \Theta\}$.

→ Support of $\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]$ is bounded,

# Generalized Normalized Radon CDTs – Feature Extractors – Normalization & Sorting

- We define the **generalized R-CDT** by $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t)$.

→ For $\mu \in \mathcal{P}_2(\mathbb{X})$, it is $\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] \in \mathcal{P}_2(\mathbb{R})$ and consequently $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} \in \mathrm{L}^2_{\rho\times\gamma}(\mathbb{R}\times\Theta)$

- with the sliced Wasserstein distance of $\mu, \nu \in \mathcal{P}_2(\mathbb{X})$ given by

$$\|\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} - \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu]}\|^2_{\rho\times\gamma} := \int_\Theta \int_\mathbb{R} |\widehat{\mathcal{R}_{\phi,\gamma}[\mu]}(t,\boldsymbol{\theta}) - \widehat{\mathcal{R}_{\phi,\gamma}[\nu]}(t,\boldsymbol{\theta})|^2 \,\mathrm{d}\rho(t)\,\mathrm{d}\gamma(\boldsymbol{\theta}) = \mathrm{SW}_2^2(\mu,\nu) = \int_\Theta \mathrm{W}_2^2(\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] - \mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu])\,\mathrm{d}\gamma(\boldsymbol{\theta}).$$

---

1) **Normalization:** The **generalized NR-CDT** $\mathcal{N}_\phi[\mu](t,\boldsymbol{\theta}) := \mathcal{N}_{\phi,\boldsymbol{\theta}}[\mu](t)$, for $t \in \mathbb{R}, \boldsymbol{\theta} \in \Theta$, is defined by

$$\mathcal{N}_{\phi,\boldsymbol{\theta}}[\mu](t) := \frac{\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t) - \mathrm{mean}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]})}{\mathrm{std}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]})}, \qquad t \in \mathbb{R}.$$

---

- For **well-definiteness**: restrict to **uniformly bounded** $\phi$: $|\phi(\mathbf{x},\boldsymbol{\theta}) - \phi(\mathbf{y},\boldsymbol{\theta})| \le M_L \;\; \forall\, d_\mathbb{X}(\mathbf{x},\mathbf{y}) \le L, \forall\, \boldsymbol{\theta} \in \Theta$,
- and consider **measures** $\mu$ in $\mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) := \{\mu \in \mathcal{P}(\mathbb{X}) \mid \mathrm{supp}(\mu) \subset\subset \mathbb{X} \;\; \wedge \;\; \exists\, c > 0 : \mathrm{std}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}) \ge c \;\forall\, \theta \in \Theta\}$.

→ Support of $\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]$ is bounded, and by continuity means and standard derivations of $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}$ exist, i.e. bounded away from zero

## Generalized Normalized Radon CDTs – Feature Extractors – Normalization & Sorting

- We define the **generalized R-CDT** by $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t)$.

→ For $\mu \in \mathcal{P}_2(\mathbb{X})$, it is $\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] \in \mathcal{P}_2(\mathbb{R})$ and consequently $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} \in \mathrm{L}^2_{\rho\times\gamma}(\mathbb{R} \times \Theta)$

- with the sliced Wasserstein distance of $\mu, \nu \in \mathcal{P}_2(\mathbb{X})$ given by

$$\|\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} - \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu]}\|^2_{\rho\times\gamma} := \int_\Theta \int_\mathbb{R} |\widehat{\mathcal{R}_{\phi,\gamma}[\mu]}(t,\boldsymbol{\theta}) - \widehat{\mathcal{R}_{\phi,\gamma}[\nu]}(t,\boldsymbol{\theta})|^2 \,\mathrm{d}\rho(t)\,\mathrm{d}\gamma(\boldsymbol{\theta}) = \mathrm{SW}_2^2(\mu,\nu) = \int_\Theta \mathrm{W}_2^2(\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] - \mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu])\,\mathrm{d}\gamma(\boldsymbol{\theta}).$$

---

1) **Normalization:** The **generalized NR-CDT** $\mathcal{N}_\phi[\mu](t,\boldsymbol{\theta}) := \mathcal{N}_{\phi,\boldsymbol{\theta}}[\mu](t)$, for $t \in \mathbb{R}, \boldsymbol{\theta} \in \Theta$, is defined by

$$\mathcal{N}_{\phi,\boldsymbol{\theta}}[\mu](t) := \frac{\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t) - \mathrm{mean}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]})}{\mathrm{std}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]})}, \qquad t \in \mathbb{R}.$$

---

- For **well-definiteness**: restrict to **uniformly bounded** $\phi$: $|\phi(\mathbf{x},\boldsymbol{\theta}) - \phi(\mathbf{y},\boldsymbol{\theta})| \le M_L \ \ \forall\, d_\mathbb{X}(\mathbf{x},\mathbf{y}) \le L, \forall \boldsymbol{\theta} \in \Theta$,
- and consider **measures** $\mu$ in $\mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) := \{\mu \in \mathcal{P}(\mathbb{X}) \mid \mathrm{supp}(\mu) \subset\subset \mathbb{X} \ \wedge \ \exists\, c > 0 : \mathrm{std}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}) \ge c \ \forall\, \theta \in \Theta\}$.

→ Support of $\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]$ is bounded, and by continuity means and standard derivations of $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}$ exist, i.e. bounded away from zero → by construction, $\mathcal{N}_\phi[\mu]$ is bounded on $\mathrm{supp}(\rho) \times \Theta$.

# Generalized Normalized Radon CDTs – Feature Extractors – Normalization & Sorting

– We define the **generalized R-CDT** by $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t)$.

→ For $\mu \in \mathcal{P}_2(\mathbb{X})$, it is $\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] \in \mathcal{P}_2(\mathbb{R})$ and consequently $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} \in \mathrm{L}^2_{\rho\times\gamma}(\mathbb{R}\times\Theta)$

– with the sliced Wasserstein distance of $\mu, \nu \in \mathcal{P}_2(\mathbb{X})$ given by

$$\|\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} - \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu]}\|^2_{\rho\times\gamma} := \int_\Theta \int_\mathbb{R} |\widehat{\mathcal{R}_{\phi,\gamma}[\mu]}(t,\boldsymbol{\theta}) - \widehat{\mathcal{R}_{\phi,\gamma}[\nu]}(t,\boldsymbol{\theta})|^2 \,\mathrm{d}\rho(t)\,\mathrm{d}\gamma(\boldsymbol{\theta}) = \mathrm{SW}_2^2(\mu,\nu) = \int_\Theta \mathrm{W}_2^2(\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] - \mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu])\,\mathrm{d}\gamma(\boldsymbol{\theta}).$$

---

1) **Normalization:** The **generalized NR-CDT** $\mathcal{N}_\phi[\mu](t,\boldsymbol{\theta}) := \mathcal{N}_{\phi,\boldsymbol{\theta}}[\mu](t)$, for $t \in \mathbb{R}, \boldsymbol{\theta} \in \Theta$, is defined by

$$\mathcal{N}_{\phi,\boldsymbol{\theta}}[\mu](t) := \frac{\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t) - \mathrm{mean}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]})}{\mathrm{std}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]})}, \qquad t \in \mathbb{R}.$$

---

– For **well-definiteness**: restrict to **uniformly bounded** $\phi$: $|\phi(\mathbf{x},\boldsymbol{\theta}) - \phi(\mathbf{y},\boldsymbol{\theta})| \le M_L \;\; \forall\, d_\mathbb{X}(\mathbf{x},\mathbf{y}) \le L, \forall \boldsymbol{\theta} \in \Theta$,

– and consider **measures** $\mu$ in $\mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) := \{\mu \in \mathcal{P}(\mathbb{X}) \mid \mathrm{supp}(\mu) \subset\subset \mathbb{X} \;\;\wedge\;\; \exists c > 0 : \mathrm{std}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}) \ge c \;\forall\, \theta \in \Theta\}$.

→ Support of $\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]$ is bounded, and by continuity means and standard derivations of $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}$ exist, i.e. bounded away from zero → by construction, $\mathcal{N}_\phi[\mu]$ is bounded on $\mathrm{supp}(\rho) \times \Theta$.

---

2) **Sorting:** dimension reduction via bounded $h\colon \mathrm{L}^\infty(\Theta) \to \mathbb{R}$, i.e. boundedness-preserving (bounded to bounded sets):

$$h(g \circ \psi) = h(g) \qquad \text{for all} \qquad g \in \mathrm{L}^\infty(\Theta) \qquad \text{and any bijection} \qquad \psi\colon \Theta \to \Theta$$

---

# Generalized Normalized Radon CDTs – Feature Extractors – Normalization & Sorting

– We define the **generalized R-CDT** by $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t)$.

→ For $\mu \in \mathcal{P}_2(\mathbb{X})$, it is $\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] \in \mathcal{P}_2(\mathbb{R})$ and consequently $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} \in \mathrm{L}^2_{\rho\times\gamma}(\mathbb{R}\times\Theta)$

– with the sliced Wasserstein distance of $\mu, \nu \in \mathcal{P}_2(\mathbb{X})$ given by

$$\|\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]} - \widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu]}\|^2_{\rho\times\gamma} \coloneqq \int_\Theta \int_\mathbb{R} |\widehat{\mathcal{R}_{\phi,\gamma}[\mu]}(t,\boldsymbol{\theta}) - \widehat{\mathcal{R}_{\phi,\gamma}[\nu]}(t,\boldsymbol{\theta})|^2 \,\mathrm{d}\rho(t)\,\mathrm{d}\gamma(\boldsymbol{\theta}) = \mathrm{SW}_2^2(\mu,\nu) = \int_\Theta \mathrm{W}_2^2(\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu] - \mathcal{R}_{\phi,\boldsymbol{\theta}}[\nu])\,\mathrm{d}\gamma(\boldsymbol{\theta}).$$

---

1) **Normalization:** The **generalized NR-CDT** $\mathcal{N}_\phi[\mu](t,\boldsymbol{\theta}) \coloneqq \mathcal{N}_{\phi,\boldsymbol{\theta}}[\mu](t)$, for $t \in \mathbb{R}, \boldsymbol{\theta} \in \Theta$, is defined by

$$\mathcal{N}_{\phi,\boldsymbol{\theta}}[\mu](t) \coloneqq \frac{\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}(t) - \mathrm{mean}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]})}{\mathrm{std}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]})}, \qquad t \in \mathbb{R}.$$

---

– For **well-definiteness**: restrict to **uniformly bounded** $\phi$: $|\phi(\mathbf{x},\boldsymbol{\theta}) - \phi(\mathbf{y},\boldsymbol{\theta})| \le M_L \;\; \forall\, d_\mathbb{X}(\mathbf{x},\mathbf{y}) \le L, \forall \boldsymbol{\theta} \in \Theta$,

– and consider **measures** $\mu$ in $\mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) \coloneqq \{\mu \in \mathcal{P}(\mathbb{X}) \mid \mathrm{supp}(\mu) \subset\subset \mathbb{X} \;\; \wedge \;\; \exists\, c > 0 : \mathrm{std}(\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}) \ge c \;\forall\, \theta \in \Theta\}$.

→ Support of $\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]$ is bounded, and by continuity means and standard derivations of $\widehat{\mathcal{R}_{\phi,\boldsymbol{\theta}}[\mu]}$ exist, i.e. bounded away from zero → by construction, $\mathcal{N}_\phi[\mu]$ is bounded on $\mathrm{supp}(\rho)\times\Theta$.

---

2) **Sorting:** dimension reduction via bounded $h\colon \mathrm{L}^\infty(\Theta) \to \mathbb{R}$, i.e. boundedness-preserving (bounded to bounded sets):

$$h(g \circ \psi) = h(g) \qquad \text{for all} \qquad g \in \mathrm{L}^\infty(\Theta) \qquad \text{and any bijection} \qquad \psi\colon \Theta \to \Theta$$

---

→ The **generalized $_\mathbf{h}$NR-CDT** $\mathcal{N}_{h,\phi}[\mu]\colon \mathbb{R} \to \mathbb{R}$ is defined by $\mathcal{N}_{h,\phi}[\mu](t) \coloneqq h(\mathcal{N}_\phi[\mu](t,\cdot))$, for $t \in \mathbb{R}$ and $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X})$.

## Properties of the Generalized Normalized Radon CDT – $_h$NR-CDT

- It holds $\mathcal{N}_{h,\phi}[\mu] \in \mathrm{L}^\infty_\rho(\mathbb{R})$ for $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X})$

## Properties of the Generalized Normalized Radon CDT – $_h$NR-CDT

– It holds $\mathcal{N}_{h,\phi}[\mu] \in \mathrm{L}^{\infty}_{\rho}(\mathbb{R})$ for $\mu \in \mathcal{P}^{+}_{\phi,\mathrm{c}}(\mathbb{X}) \to$ **well-definiteness** of the measure transform, i.e. feature extractor.

## Properties of the Generalized Normalized Radon CDT – $_{\mathrm{h}}$NR-CDT

– It holds $\mathcal{N}_{h,\phi}[\mu] \in \mathrm{L}^{\infty}_{\rho}(\mathbb{R})$ for $\mu \in \mathcal{P}^{+}_{\phi,\mathrm{c}}(\mathbb{X}) \rightarrow$ **well-definiteness** of the measure transform, i.e. feature extractor.

**Theorem (Separability)**

## Properties of the Generalized Normalized Radon CDT – $_{\mathrm{h}}$NR-CDT

– It holds $\mathcal{N}_{h,\phi}[\mu] \in \mathrm{L}^\infty_\rho(\mathbb{R})$ for $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) \to$ **well-definiteness** of the measure transform, i.e. feature extractor.

**Theorem (Separability)**

*For the settings $\mathbb{X}$ and $\Theta$, and $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X})$ holds*

## Properties of the Generalized Normalized Radon CDT – $_{\mathrm{h}}$NR-CDT

– It holds $\mathcal{N}_{h,\phi}[\mu] \in \mathrm{L}^{\infty}_{\rho}(\mathbb{R})$ for $\mu \in \mathcal{P}^{+}_{\phi,\mathrm{c}}(\mathbb{X}) \to$ **well-definiteness** of the measure transform, i.e. feature extractor.

**Theorem (Separability)**

*For the settings $\mathbb{X}$ and $\Theta$, and $\mu \in \mathcal{P}^{+}_{\phi,\mathrm{c}}(\mathbb{X})$ holds*

i) $\mathcal{N}_{h,\phi}[T_{\#}\mu] = \mathcal{N}_{h,\phi}[\mu]$ *for the specific* $T \in \bullet_{\mathbb{X},\Theta}$

## Properties of the Generalized Normalized Radon CDT – ${}_h$NR-CDT

– It holds $\mathcal{N}_{h,\phi}[\mu] \in \mathrm{L}^\infty_\rho(\mathbb{R})$ for $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) \rightarrow$ **well-definiteness** of the measure transform, i.e. feature extractor.

**Theorem (Separability)**

*For the settings $\mathbb{X}$ and $\Theta$, and $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X})$ holds*

i) $\mathcal{N}_{h,\phi}[T_\# \mu] = \mathcal{N}_{h,\phi}[\mu]$ *for the specific* $T \in \bullet_{\mathbb{X},\Theta}$

ii) *each subset* $\hat{\mathbb{F}}_\mu$ *of* $\mathbb{F}_\mu := \{T_\# \mu \mid T \in \bullet_{\mathbb{X},\Theta}\}$ *reduces in* ${}_h$*NR-CDT-space to single points sets,*

## Properties of the Generalized Normalized Radon CDT – $_h$NR-CDT

– It holds $\mathcal{N}_{h,\phi}[\mu] \in \mathrm{L}^\infty_\rho(\mathbb{R})$ for $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) \to$ **well-definiteness** of the measure transform, i.e. feature extractor.

**Theorem (Separability)**

*For the settings $\mathbb{X}$ and $\Theta$, and $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X})$ holds*

i) $\mathcal{N}_{h,\phi}[T_\# \mu] = \mathcal{N}_{h,\phi}[\mu]$ *for the specific* $T \in \bullet_{\mathbb{X},\Theta}$

ii) *each subset* $\hat{\mathbb{F}}_\mu$ *of* $\mathbb{F}_\mu \coloneqq \{T_\# \mu \mid T \in \bullet_{\mathbb{X},\Theta}\}$ *reduces in* $_h$*NR-CDT-space to single points sets,*

iii) *separability of* $\mathcal{N}_{h,\phi}[\hat{\mathbb{F}}_{\mu_i}]$ *if* $\mathcal{N}_{h,\phi}[\mu_i] \neq \mathcal{N}_{h,\phi}[\mu_j]$, $i \neq j$.

## Properties of the Generalized Normalized Radon CDT – ${}_h$NR-CDT

– It holds $\mathcal{N}_{h,\phi}[\mu] \in \mathrm{L}^\infty_\rho(\mathbb{R})$ for $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) \to$ **well-definiteness** of the measure transform, i.e. feature extractor.

**Theorem (Separability)**

*For the settings $\mathbb{X}$ and $\Theta$, and $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X})$ holds*

i) $\mathcal{N}_{h,\phi}[T_\# \mu] = \mathcal{N}_{h,\phi}[\mu]$ *for the specific* $T \in \bullet_{\mathbb{X},\Theta}$

ii) *each subset* $\hat{\mathbb{F}}_\mu$ *of* $\mathbb{F}_\mu := \{T_\# \mu \mid T \in \bullet_{\mathbb{X},\Theta}\}$ *reduces in* ${}_h$*NR-CDT-space to single points sets,*

iii) *separability of* $\mathcal{N}_{h,\phi}[\hat{\mathbb{F}}_{\mu_i}]$ *if* $\mathcal{N}_{h,\phi}[\mu_i] \neq \mathcal{N}_{h,\phi}[\mu_j]$, $i \neq j$.

– Let $\eta\colon \mathbb{R} \to \mathbb{R}$ be boundedness-preserving and $H\colon \mathcal{B}(\mathbb{R}) \to \mathbb{R}$ bounded,

## Properties of the Generalized Normalized Radon CDT – ${}_h$NR-CDT

– It holds $\mathcal{N}_{h,\phi}[\mu] \in \mathrm{L}^\infty_\rho(\mathbb{R})$ for $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) \to$ **well-definiteness** of the measure transform, i.e. feature extractor.

**Theorem (Separability)**

*For the settings $\mathbb{X}$ and $\Theta$, and $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X})$ holds*

i) $\mathcal{N}_{h,\phi}[T_\# \mu] = \mathcal{N}_{h,\phi}[\mu]$ *for the specific* $T \in \bullet_{\mathbb{X},\Theta}$

ii) *each subset* $\hat{\mathbb{F}}_\mu$ *of* $\mathbb{F}_\mu := \{T_\# \mu \mid T \in \bullet_{\mathbb{X},\Theta}\}$ *reduces in* ${}_h$*NR-CDT-space to single points sets,*

iii) *separability of* $\mathcal{N}_{h,\phi}[\hat{\mathbb{F}}_{\mu_i}]$ *if* $\mathcal{N}_{h,\phi}[\mu_i] \neq \mathcal{N}_{h,\phi}[\mu_j]$, $i \neq j$.

– Let $\eta \colon \mathbb{R} \to \mathbb{R}$ be boundedness-preserving and $H \colon \mathcal{B}(\mathbb{R}) \to \mathbb{R}$ bounded,

– By setting $h(g) = H(\{\eta(g(\boldsymbol{\theta})) \mid \boldsymbol{\theta} \in \Theta\})$ for $g \in \mathrm{L}^\infty(\Theta)$, the required construction for $h$ is obtained,

# Properties of the Generalized Normalized Radon CDT – ${}_h$NR-CDT

– It holds $\mathcal{N}_{h,\phi}[\mu] \in \mathrm{L}^\infty_\rho(\mathbb{R})$ for $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) \to$ **well-definiteness** of the measure transform, i.e. feature extractor.

**Theorem (Separability)**

*For the settings $\mathbb{X}$ and $\Theta$, and $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X})$ holds*

i) $\mathcal{N}_{h,\phi}[T_\# \mu] = \mathcal{N}_{h,\phi}[\mu]$ *for the specific* $T \in \bullet_{\mathbb{X},\Theta}$

ii) *each subset* $\hat{\mathbb{F}}_\mu$ *of* $\mathbb{F}_\mu \coloneqq \{T_\# \mu \mid T \in \bullet_{\mathbb{X},\Theta}\}$ *reduces in* ${}_h$*NR-CDT-space to single points sets,*

iii) *separability of* $\mathcal{N}_{h,\phi}[\hat{\mathbb{F}}_{\mu_i}]$ *if* $\mathcal{N}_{h,\phi}[\mu_i] \neq \mathcal{N}_{h,\phi}[\mu_j]$, $i \neq j$.

– Let $\eta\colon \mathbb{R} \to \mathbb{R}$ be boundedness-preserving and $H\colon \mathcal{B}(\mathbb{R}) \to \mathbb{R}$ bounded,

– By setting $h(g) = H(\{\eta(g(\boldsymbol{\theta})) \mid \boldsymbol{\theta} \in \Theta\})$ for $g \in \mathrm{L}^\infty(\Theta)$, the required construction for $h$ is obtained, including

## Properties of the Generalized Normalized Radon CDT – ${}_h$NR-CDT

– It holds $\mathcal{N}_{h,\phi}[\mu] \in \mathrm{L}^\infty_\rho(\mathbb{R})$ for $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) \rightarrow$ **well-definiteness** of the measure transform, i.e. feature extractor.

**Theorem (Separability)**

*For the settings $\mathbb{X}$ and $\Theta$, and $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X})$ holds*

i) $\mathcal{N}_{h,\phi}[T_\# \mu] = \mathcal{N}_{h,\phi}[\mu]$ *for the specific* $T \in \bullet_{\mathbb{X},\Theta}$

ii) *each subset* $\hat{\mathbb{F}}_\mu$ *of* $\mathbb{F}_\mu \coloneqq \{T_\# \mu \mid T \in \bullet_{\mathbb{X},\Theta}\}$ *reduces in* ${}_h$*NR-CDT-space to single points sets,*

iii) *separability of* $\mathcal{N}_{h,\phi}[\hat{\mathbb{F}}_{\mu_i}]$ *if* $\mathcal{N}_{h,\phi}[\mu_i] \neq \mathcal{N}_{h,\phi}[\mu_j]$, $i \neq j$.

– Let $\eta \colon \mathbb{R} \to \mathbb{R}$ be boundedness-preserving and $H \colon \mathcal{B}(\mathbb{R}) \to \mathbb{R}$ bounded,

– By setting $h(g) = H(\{\eta(g(\boldsymbol{\theta})) \mid \boldsymbol{\theta} \in \Theta\})$ for $g \in \mathrm{L}^\infty(\Theta)$, the required construction for $h$ is obtained, including

a) $\eta = |\cdot|$ and $H = \sup$, i.e.,

$$\sup_{\boldsymbol{\theta} \in \Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|,$$

## Properties of the Generalized Normalized Radon CDT – ${}_h$NR-CDT

– It holds $\mathcal{N}_{h,\phi}[\mu] \in \mathrm{L}^\infty_\rho(\mathbb{R})$ for $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) \to$ **well-definiteness** of the measure transform, i.e. feature extractor.

**Theorem (Separability)**

*For the settings $\mathbb{X}$ and $\Theta$, and $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X})$ holds*

i) $\mathcal{N}_{h,\phi}[T_\# \mu] = \mathcal{N}_{h,\phi}[\mu]$ *for the specific* $T \in \bullet_{\mathbb{X},\Theta}$

ii) *each subset* $\hat{\mathbb{F}}_\mu$ *of* $\mathbb{F}_\mu := \{T_\# \mu \mid T \in \bullet_{\mathbb{X},\Theta}\}$ *reduces in* ${}_h$*NR-CDT-space to single points sets,*

iii) *separability of* $\mathcal{N}_{h,\phi}[\hat{\mathbb{F}}_{\mu_i}]$ *if* $\mathcal{N}_{h,\phi}[\mu_i] \neq \mathcal{N}_{h,\phi}[\mu_j]$, $i \neq j$.

– Let $\eta\colon \mathbb{R} \to \mathbb{R}$ be boundedness-preserving and $H\colon \mathcal{B}(\mathbb{R}) \to \mathbb{R}$ bounded,

– By setting $h(g) = H(\{\eta(g(\boldsymbol{\theta})) \mid \boldsymbol{\theta} \in \Theta\})$ for $g \in \mathrm{L}^\infty(\Theta)$, the required construction for $h$ is obtained, including

a) $\eta = |\cdot|$ and $H = \sup$, i.e.,
$$\sup_{\boldsymbol{\theta}\in\Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|,$$

b) $\eta = |\cdot|$ and $H = \inf$, i.e.,
$$\inf_{\boldsymbol{\theta}\in\Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|,$$

# Properties of the Generalized Normalized Radon CDT – $_h$NR-CDT

– It holds $\mathcal{N}_{h,\phi}[\mu] \in \mathrm{L}^\infty_\rho(\mathbb{R})$ for $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) \to$ **well-definiteness** of the measure transform, i.e. feature extractor.

**Theorem (Separability)**

*For the settings $\mathbb{X}$ and $\Theta$, and $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X})$ holds*

i) $\mathcal{N}_{h,\phi}[T_\# \mu] = \mathcal{N}_{h,\phi}[\mu]$ *for the specific* $T \in \bullet_{\mathbb{X},\Theta}$

ii) *each subset* $\hat{\mathbb{F}}_\mu$ *of* $\mathbb{F}_\mu := \{T_\# \mu \mid T \in \bullet_{\mathbb{X},\Theta}\}$ *reduces in* $_h$*NR-CDT-space to single points sets,*

iii) *separability of* $\mathcal{N}_{h,\phi}[\hat{\mathbb{F}}_{\mu_i}]$ *if* $\mathcal{N}_{h,\phi}[\mu_i] \neq \mathcal{N}_{h,\phi}[\mu_j]$, $i \neq j$.

– Let $\eta\colon \mathbb{R} \to \mathbb{R}$ be boundedness-preserving and $H\colon \mathcal{B}(\mathbb{R}) \to \mathbb{R}$ bounded,

– By setting $h(g) = H(\{\eta(g(\boldsymbol{\theta})) \mid \boldsymbol{\theta} \in \Theta\})$ for $g \in \mathrm{L}^\infty(\Theta)$, the required construction for $h$ is obtained, including

a) $\eta = |\cdot|$ and $H = \sup$, i.e.,
$$\sup_{\boldsymbol{\theta}\in\Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|,$$

b) $\eta = |\cdot|$ and $H = \inf$, i.e.,
$$\inf_{\boldsymbol{\theta}\in\Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|,$$

c) $\eta = |\cdot|$ and $H = \sup - \inf$, i.e.,
$$\sup_{\boldsymbol{\theta}\in\Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)| - \inf_{\boldsymbol{\theta}\in\Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|,$$

# Properties of the Generalized Normalized Radon CDT – $_h$NR-CDT

– It holds $\mathcal{N}_{h,\phi}[\mu] \in \mathrm{L}^\infty_\rho(\mathbb{R})$ for $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X})$ $\rightarrow$ **well-definiteness** of the measure transform, i.e. feature extractor.

**Theorem (Separability)**

*For the settings $\mathbb{X}$ and $\Theta$, and $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X})$ holds*

i) $\mathcal{N}_{h,\phi}[T_\# \mu] = \mathcal{N}_{h,\phi}[\mu]$ *for the specific* $T \in \bullet_{\mathbb{X},\Theta}$

ii) *each subset* $\hat{\mathbb{F}}_\mu$ *of* $\mathbb{F}_\mu := \{T_\# \mu \mid T \in \bullet_{\mathbb{X},\Theta}\}$ *reduces in $_h$NR-CDT-space to single points sets,*

iii) *separability of* $\mathcal{N}_{h,\phi}[\hat{\mathbb{F}}_{\mu_i}]$ *if* $\mathcal{N}_{h,\phi}[\mu_i] \neq \mathcal{N}_{h,\phi}[\mu_j]$, $i \neq j$.

– Let $\eta\colon \mathbb{R} \to \mathbb{R}$ be boundedness-preserving and $H\colon \mathcal{B}(\mathbb{R}) \to \mathbb{R}$ bounded,

– By setting $h(g) = H(\{\eta(g(\boldsymbol{\theta})) \mid \boldsymbol{\theta} \in \Theta\})$ for $g \in \mathrm{L}^\infty(\Theta)$, the required construction for $h$ is obtained, including

a) $\eta = |\cdot|$ and $H = \sup$, i.e.,
$$\sup_{\boldsymbol{\theta}\in\Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|,$$

b) $\eta = |\cdot|$ and $H = \inf$, i.e.,
$$\inf_{\boldsymbol{\theta}\in\Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|,$$

c) $\eta = |\cdot|$ and $H = \sup - \inf$, i.e.,
$$\sup_{\boldsymbol{\theta}\in\Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)| - \inf_{\boldsymbol{\theta}\in\Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|,$$

d) $\eta = \mathrm{Id}$ and $H = \sup - \inf$, i.e.,
$$\sup_{\boldsymbol{\theta}\in\Theta} \mathcal{N}_{\boldsymbol{\theta}}[\mu](t) - \inf_{\boldsymbol{\theta}\in\Theta} \mathcal{N}_{\boldsymbol{\theta}}[\mu](t).$$

## Properties of the Generalized Normalized Radon CDT – $_h$NR-CDT

– It holds $\mathcal{N}_{h,\phi}[\mu] \in \mathrm{L}^\infty_\rho(\mathbb{R})$ for $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) \to$ **well-definiteness** of the measure transform, i.e. feature extractor.

**Theorem (Separability)**

*For the settings $\mathbb{X}$ and $\Theta$, and $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X})$ holds*

i) $\mathcal{N}_{h,\phi}[T_\# \mu] = \mathcal{N}_{h,\phi}[\mu]$ *for the specific* $T \in \bullet_{\mathbb{X},\Theta}$

ii) *each subset* $\hat{\mathbb{F}}_\mu$ *of* $\mathbb{F}_\mu := \{T_\# \mu \mid T \in \bullet_{\mathbb{X},\Theta}\}$ *reduces in* $_h$*NR-CDT-space to single points sets,*

iii) *separability of* $\mathcal{N}_{h,\phi}[\hat{\mathbb{F}}_{\mu_i}]$ *if* $\mathcal{N}_{h,\phi}[\mu_i] \neq \mathcal{N}_{h,\phi}[\mu_j]$, $i \neq j$.

– Further transformations are:

– Let $\eta \colon \mathbb{R} \to \mathbb{R}$ be boundedness-preserving and $H \colon \mathcal{B}(\mathbb{R}) \to \mathbb{R}$ bounded,

– By setting $h(g) = H(\{\eta(g(\boldsymbol{\theta})) \mid \boldsymbol{\theta} \in \Theta\})$ for $g \in \mathrm{L}^\infty(\Theta)$, the required construction for $h$ is obtained, including

a) $\eta = |\cdot|$ and $H = \sup$, i.e.,
$$\sup_{\boldsymbol{\theta}\in\Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|,$$

b) $\eta = |\cdot|$ and $H = \inf$, i.e.,
$$\inf_{\boldsymbol{\theta}\in\Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|,$$

c) $\eta = |\cdot|$ and $H = \sup - \inf$, i.e.,
$$\sup_{\boldsymbol{\theta}\in\Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)| - \inf_{\boldsymbol{\theta}\in\Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|,$$

d) $\eta = \mathrm{Id}$ and $H = \sup - \inf$, i.e.,
$$\sup_{\boldsymbol{\theta}\in\Theta} \mathcal{N}_{\boldsymbol{\theta}}[\mu](t) - \inf_{\boldsymbol{\theta}\in\Theta} \mathcal{N}_{\boldsymbol{\theta}}[\mu](t).$$

# Properties of the Generalized Normalized Radon CDT – $_h$NR-CDT

– It holds $\mathcal{N}_{h,\phi}[\mu] \in \mathrm{L}^\infty_\rho(\mathbb{R})$ for $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) \rightarrow$ **well-definiteness** of the measure transform, i.e. feature extractor.

**Theorem (Separability)**

*For the settings $\mathbb{X}$ and $\Theta$, and $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X})$ holds*

i) $\mathcal{N}_{h,\phi}[T_\# \mu] = \mathcal{N}_{h,\phi}[\mu]$ *for the specific* $T \in \bullet_{\mathbb{X},\Theta}$

ii) *each subset* $\hat{\mathbb{F}}_\mu$ *of* $\mathbb{F}_\mu := \{T_\# \mu \mid T \in \bullet_{\mathbb{X},\Theta}\}$ *reduces in* $_h$*NR-CDT-space to single points sets,*

iii) *separability of* $\mathcal{N}_{h,\phi}[\hat{\mathbb{F}}_{\mu_i}]$ *if* $\mathcal{N}_{h,\phi}[\mu_i] \neq \mathcal{N}_{h,\phi}[\mu_j]$, $i \neq j$.

– Further transformations are:

e) TV-normalized: $h(g) = \sup\limits_{P \in \mathbb{P}} \sum\limits_{i=1}^{n_P} |g(\boldsymbol{\theta}_{i+1}) - g(\boldsymbol{\theta}_i)|$, $\Theta = \mathbb{S}_1$,
with $\mathbb{P} = \{P = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_{n_P+1}) \mid P$ is partition of $\mathbb{S}_1\}$,

– Let $\eta \colon \mathbb{R} \to \mathbb{R}$ be boundedness-preserving and $H \colon \mathcal{B}(\mathbb{R}) \to \mathbb{R}$ bounded,

– By setting $h(g) = H(\{\eta(g(\boldsymbol{\theta})) \mid \boldsymbol{\theta} \in \Theta\})$ for $g \in \mathrm{L}^\infty(\Theta)$, the required construction for $h$ is obtained, including

a) $\eta = |\cdot|$ and $H = \sup$, i.e.,
$\sup_{\boldsymbol{\theta} \in \Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|$,

b) $\eta = |\cdot|$ and $H = \inf$, i.e.,
$\inf_{\boldsymbol{\theta} \in \Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|$,

c) $\eta = |\cdot|$ and $H = \sup - \inf$, i.e.,
$\sup_{\boldsymbol{\theta} \in \Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)| - \inf_{\boldsymbol{\theta} \in \Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|$,

d) $\eta = \mathrm{Id}$ and $H = \sup - \inf$, i.e.,
$\sup_{\boldsymbol{\theta} \in \Theta} \mathcal{N}_{\boldsymbol{\theta}}[\mu](t) - \inf_{\boldsymbol{\theta} \in \Theta} \mathcal{N}_{\boldsymbol{\theta}}[\mu](t)$.

## Properties of the Generalized Normalized Radon CDT – $_h$NR-CDT

– It holds $\mathcal{N}_{h,\phi}[\mu] \in \mathrm{L}^\infty_\rho(\mathbb{R})$ for $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) \to$ **well-definiteness** of the measure transform, i.e. feature extractor.

**Theorem (Separability)**

*For the settings $\mathbb{X}$ and $\Theta$, and $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X})$ holds*

i) $\mathcal{N}_{h,\phi}[T_\# \mu] = \mathcal{N}_{h,\phi}[\mu]$ *for the specific* $T \in \bullet_{\mathbb{X},\Theta}$

ii) *each subset* $\hat{\mathbb{F}}_\mu$ *of* $\mathbb{F}_\mu := \{T_\# \mu \mid T \in \bullet_{\mathbb{X},\Theta}\}$ *reduces in* $_h$*NR-CDT-space to single points sets,*

iii) *separability of* $\mathcal{N}_{h,\phi}[\hat{\mathbb{F}}_{\mu_i}]$ *if* $\mathcal{N}_{h,\phi}[\mu_i] \neq \mathcal{N}_{h,\phi}[\mu_j]$, $i \neq j$.

– Further transformations are:

e) TV-normalized: $h(g) = \sup\limits_{P \in \mathbb{P}} \sum\limits_{i=1}^{n_P} |g(\boldsymbol{\theta}_{i+1}) - g(\boldsymbol{\theta}_i)|$, $\Theta = \mathbb{S}_1$, with $\mathbb{P} = \{P = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_{n_P+1}) \mid P$ is partition of $\mathbb{S}_1\}$,

f) average-normalized: $h(g) = \int_\Theta g(\boldsymbol{\theta}) \,\mathrm{d}\gamma(\boldsymbol{\theta})$ violating separability under arbitrary affine transformations,

– Let $\eta\colon \mathbb{R} \to \mathbb{R}$ be boundedness-preserving and $H\colon \mathcal{B}(\mathbb{R}) \to \mathbb{R}$ bounded,

– By setting $h(g) = H(\{\eta(g(\boldsymbol{\theta})) \mid \boldsymbol{\theta} \in \Theta\})$ for $g \in \mathrm{L}^\infty(\Theta)$, the required construction for $h$ is obtained, including

a) $\eta = |\cdot|$ and $H = \sup$, i.e.,
$$\sup_{\boldsymbol{\theta}\in\Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|,$$

b) $\eta = |\cdot|$ and $H = \inf$, i.e.,
$$\inf_{\boldsymbol{\theta}\in\Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|,$$

c) $\eta = |\cdot|$ and $H = \sup - \inf$, i.e.,
$$\sup_{\boldsymbol{\theta}\in\Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)| - \inf_{\boldsymbol{\theta}\in\Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|,$$

d) $\eta = \mathrm{Id}$ and $H = \sup - \inf$, i.e.,
$$\sup_{\boldsymbol{\theta}\in\Theta} \mathcal{N}_{\boldsymbol{\theta}}[\mu](t) - \inf_{\boldsymbol{\theta}\in\Theta} \mathcal{N}_{\boldsymbol{\theta}}[\mu](t).$$

## Properties of the Generalized Normalized Radon CDT – ${}_h$NR-CDT

– It holds $\mathcal{N}_{h,\phi}[\mu] \in \mathrm{L}^\infty_\rho(\mathbb{R})$ for $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) \to$ **well-definiteness** of the measure transform, i.e. feature extractor.

**Theorem (Separability)**

*For the settings $\mathbb{X}$ and $\Theta$, and $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X})$ holds*

i) $\mathcal{N}_{h,\phi}[T_\# \mu] = \mathcal{N}_{h,\phi}[\mu]$ *for the specific* $T \in \bullet_{\mathbb{X},\Theta}$

ii) *each subset* $\hat{\mathbb{F}}_\mu$ *of* $\mathbb{F}_\mu := \{T_\# \mu \mid T \in \bullet_{\mathbb{X},\Theta}\}$ *reduces in* ${}_h$*NR-CDT-space to single points sets,*

iii) *separability of* $\mathcal{N}_{h,\phi}[\hat{\mathbb{F}}_{\mu_i}]$ *if* $\mathcal{N}_{h,\phi}[\mu_i] \neq \mathcal{N}_{h,\phi}[\mu_j]$, $i \neq j$.

– Further transformations are:

e) TV-normalized: $h(g) = \sup_{P \in \mathbb{P}} \sum_{i=1}^{n_P} |g(\boldsymbol{\theta}_{i+1}) - g(\boldsymbol{\theta}_i)|$, $\Theta = \mathbb{S}_1$, with $\mathbb{P} = \{P = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_{n_P+1}) \mid P$ is partition of $\mathbb{S}_1\}$,

f) average-normalized: $h(g) = \int_\Theta g(\boldsymbol{\theta}) \, \mathrm{d}\gamma(\boldsymbol{\theta})$ violating separability under arbitrary affine transformations, i.e. only single point sets if $T \in \bullet_{\mathbb{X},\Theta}$ s.t. $\kappa(T^\top T) = 1$.

– Let $\eta \colon \mathbb{R} \to \mathbb{R}$ be boundedness-preserving and $H \colon \mathcal{B}(\mathbb{R}) \to \mathbb{R}$ bounded,

– By setting $h(g) = H(\{\eta(g(\boldsymbol{\theta})) \mid \boldsymbol{\theta} \in \Theta\})$ for $g \in \mathrm{L}^\infty(\Theta)$, the required construction for $h$ is obtained, including

a) $\eta = |\cdot|$ and $H = \sup$, i.e.,
$$\sup_{\boldsymbol{\theta} \in \Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|,$$

b) $\eta = |\cdot|$ and $H = \inf$, i.e.,
$$\inf_{\boldsymbol{\theta} \in \Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|,$$

c) $\eta = |\cdot|$ and $H = \sup - \inf$, i.e.,
$$\sup_{\boldsymbol{\theta} \in \Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)| - \inf_{\boldsymbol{\theta} \in \Theta} |\mathcal{N}_{\boldsymbol{\theta}}[\mu](t)|,$$

d) $\eta = \mathrm{Id}$ and $H = \sup - \inf$, i.e.,
$$\sup_{\boldsymbol{\theta} \in \Theta} \mathcal{N}_{\boldsymbol{\theta}}[\mu](t) - \inf_{\boldsymbol{\theta} \in \Theta} \mathcal{N}_{\boldsymbol{\theta}}[\mu](t).$$

## Numerics – Large Class Numbers & Few References

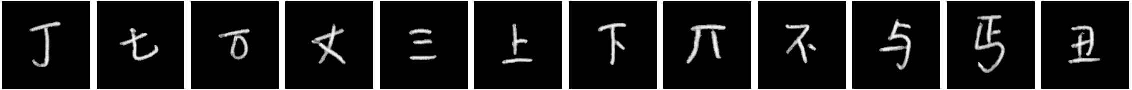
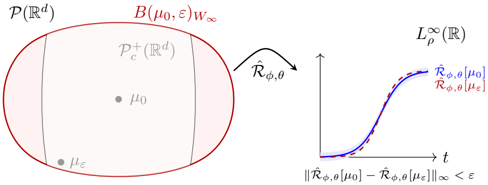


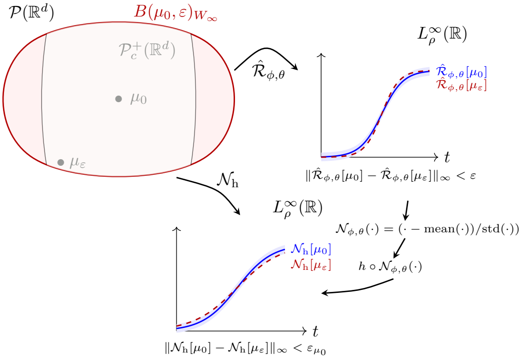
Table: NT and NN classification accuracies for the Chinese character dataset, where each class consists of 50 samples, generated by random affine transformations with scaling $[0.5, 1.0]$, shearing $[-25^\circ, 25^\circ]$, rotation angles $[0^\circ, 360^\circ]$ and pixel shifts $[-20, 20]$.

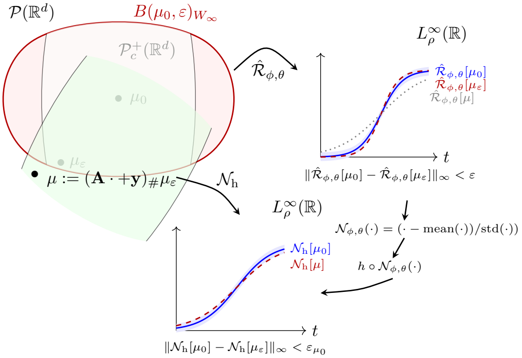
| # classes | method | NT # angles 2 | 4 | 8 | 16 | 32 | 64 | 128 | NN # training samples 5 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 100 | Euclidean | 0.0096 | | | | | | | $0.0202 \pm 0.0021$ | $0.0242 \pm 0.0022$ |
| | R-CDT | 0.0160 | 0.0180 | 0.0172 | 0.0164 | 0.0168 | 0.0170 | 0.0170 | $0.0188 \pm 0.0020$ | $0.0217 \pm 0.0021$ |
| | $_{\mathrm{m}}$NR-CDT | **0.0924** | **0.1038** | **0.3124** | **0.8422** | **0.9836** | **1.0000** | **1.0000** | $\mathbf{1.0000 \pm 0.0000}$ | $\mathbf{1.0000 \pm 0.0000}$ |
| | $_{\mathrm{a}}$NR-CDT | 0.0810 | 0.0172 | 0.1860 | 0.4950 | 0.6486 | 0.6814 | 0.6852 | $0.8345 \pm 0.0097$ | $0.9139 \pm 0.0062$ |
| 1000 | $_{\mathrm{m}}$NR-CDT | **0.0848** | **0.0620** | **0.2178** | **0.7368** | **0.9791** | **0.9975** | **0.9981** | $\mathbf{0.9987 \pm 0.0001}$ | $\mathbf{0.9987 \pm 0.0001}$ |
| | $_{\mathrm{a}}$NR-CDT | 0.0629 | 0.0362 | 0.0751 | 0.3475 | 0.5554 | 0.6030 | 0.6104 | $0.7651 \pm 0.0026$ | $0.8631 \pm 0.0018$ |

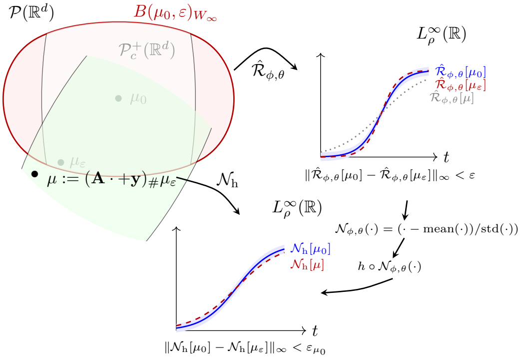
## Robustness of the Separability in Wasserstein Spaces for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$



# Robustness of the Separability in Wasserstein Spaces for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$



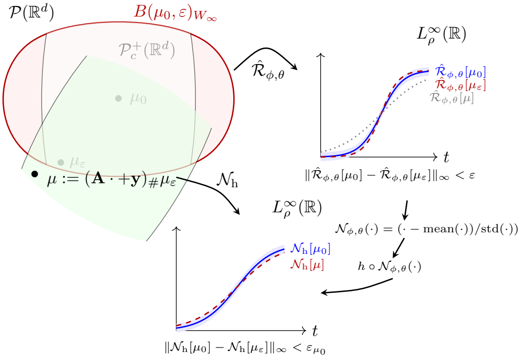

$\mathcal{P}(\mathbb{R}^d)$
$B(\mu_0, \varepsilon)_{W_\infty}$
$\mathcal{P}_c^+(\mathbb{R}^d)$
$\mu_0$
$\mu_\varepsilon$
$\mu := (\mathbf{A} \cdot + \mathbf{y})_\# \mu_\varepsilon$
$\hat{\mathcal{R}}_{\phi,\theta}$
$L_\rho^\infty(\mathbb{R})$
$\hat{\mathcal{R}}_{\phi,\theta}[\mu_0]$
$\hat{\mathcal{R}}_{\phi,\theta}[\mu_\varepsilon]$
$\hat{\mathcal{R}}_{\phi,\theta}[\mu]$
$t$
$\|\hat{\mathcal{R}}_{\phi,\theta}[\mu_0] - \hat{\mathcal{R}}_{\phi,\theta}[\mu_\varepsilon]\|_\infty < \varepsilon$
$\mathcal{N}_{\phi,\theta}(\cdot) = (\cdot - \mathrm{mean}(\cdot))/\mathrm{std}(\cdot))$
$h \circ \mathcal{N}_{\phi,\theta}(\cdot)$
$\mathcal{N}_{\mathrm{h}}$
$L_\rho^\infty(\mathbb{R})$
$\mathcal{N}_{\mathrm{h}}[\mu_0]$
$\mathcal{N}_{\mathrm{h}}[\mu]$
$t$
$\|\mathcal{N}_{\mathrm{h}}[\mu_0] - \mathcal{N}_{\mathrm{h}}[\mu_\varepsilon]\|_\infty < \varepsilon_{\mu_0}$




For $\mu \in \mathcal{P}_1(R^d)$ define the zero-mean quantile

$$\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu](t) := \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu](t) - \mathrm{mean}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu]) \quad \forall\, \boldsymbol{\theta} \in \mathbb{S}_1 \ \forall\, t \in \mathbb{R},$$

# Robustness of the Separability in Wasserstein Spaces for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$

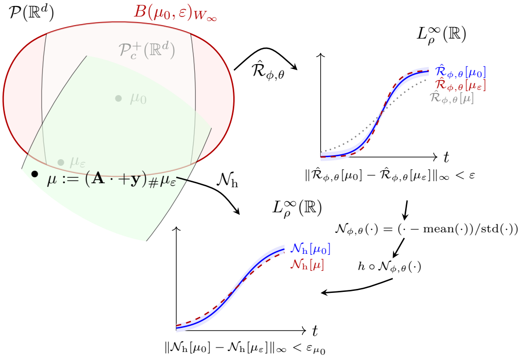



## Proposition

For $\mu \in \mathcal{P}_1(R^d)$ define the zero-mean quantile

$$\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu](t) := \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu](t) - \mathrm{mean}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu]) \quad \forall \boldsymbol{\theta} \in \mathbb{S}_1 \ \forall t \in \mathbb{R},$$

# Robustness of the Separability in Wasserstein Spaces for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$



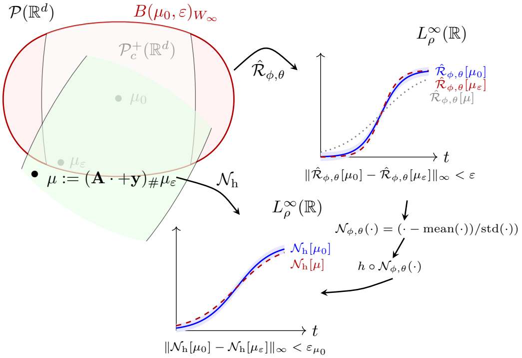

**Proposition**

*Let $\mu_0, \mu_\epsilon \in \mathcal{P}(\mathbb{R}^2)$ with $W_\infty(\mu_0, \mu_\epsilon) \leq \epsilon$.*

For $\mu \in \mathcal{P}_1(R^d)$ define the zero-mean quantile

$$\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu](t) := \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu](t) - \mathrm{mean}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu]) \quad \forall\, \boldsymbol{\theta} \in \mathbb{S}_1 \ \forall\, t \in \mathbb{R},$$

## Robustness of the Separability in Wasserstein Spaces for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$



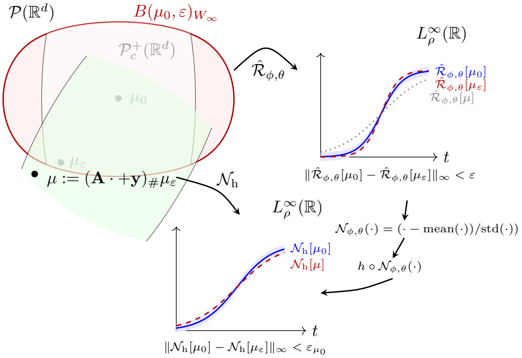

**Proposition**

*Let $\mu_0, \mu_\epsilon \in \mathcal{P}(\mathbb{R}^2)$ with $W_\infty(\mu_0, \mu_\epsilon) \leq \epsilon$. Then*

For $\mu \in \mathcal{P}_1(R^d)$ define the zero-mean quantile

$$\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu](t) := \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu](t) - \mathrm{mean}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu]) \quad \forall \boldsymbol{\theta} \in \mathbb{S}_1 \ \forall t \in \mathbb{R},$$

# Robustness of the Separability in Wasserstein Spaces for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$



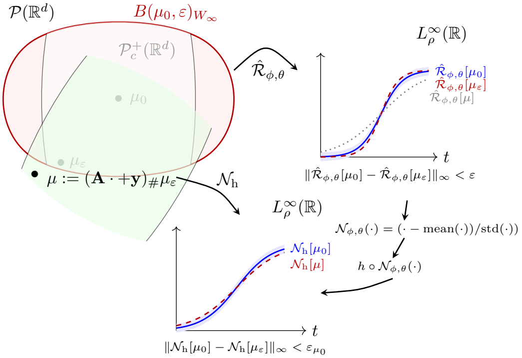

**Proposition**

*Let $\mu_0, \mu_\epsilon \in \mathcal{P}(\mathbb{R}^2)$ with $W_\infty(\mu_0, \mu_\epsilon) \leq \epsilon$. Then*
i) $\|\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0] - \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \leq \epsilon$,

For $\mu \in \mathcal{P}_1(R^d)$ define the zero-mean quantile

$$\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu](t) := \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu](t) - \mathrm{mean}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu]) \quad \forall\, \boldsymbol{\theta} \in \mathbb{S}_1 \;\forall\, t \in \mathbb{R},$$

# Robustness of the Separability in Wasserstein Spaces for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$



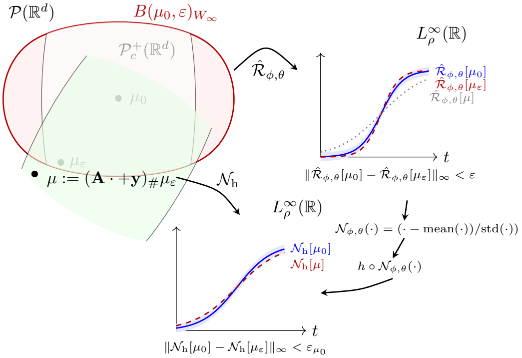

**Proposition**

*Let $\mu_0, \mu_\epsilon \in \mathcal{P}(\mathbb{R}^2)$ with $W_\infty(\mu_0, \mu_\epsilon) \le \epsilon$. Then*
i) $\|\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0] - \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \le \epsilon$, ii) $\|\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_0] - \widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \le 2\epsilon$.

For $\mu \in \mathcal{P}_1(R^d)$ define the zero-mean quantile

$$\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu](t) := \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu](t) - \mathrm{mean}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu]) \quad \forall\, \boldsymbol{\theta} \in \mathbb{S}_1 \;\forall\, t \in \mathbb{R},$$

# Robustness of the Separability in Wasserstein Spaces for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$




**Proposition**

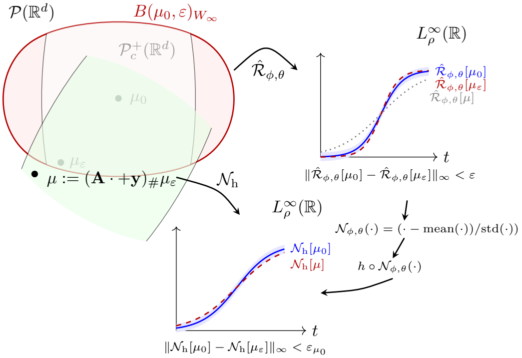
*Let $\mu_0, \mu_\epsilon \in \mathcal{P}(\mathbb{R}^2)$ with $W_\infty(\mu_0, \mu_\epsilon) \leq \epsilon$. Then*
i) $\|\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0] - \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \leq \epsilon$, ii) $\|\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_0] - \widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \leq 2\epsilon$.

**Proposition**

For $\mu \in \mathcal{P}_1(R^d)$ define the zero-mean quantile

$$\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu](t) := \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu](t) - \mathrm{mean}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu]) \quad \forall \boldsymbol{\theta} \in \mathbb{S}_1 \ \forall t \in \mathbb{R},$$

# Robustness of the Separability in Wasserstein Spaces for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$




## Proposition

*Let $\mu_0, \mu_\epsilon \in \mathcal{P}(\mathbb{R}^2)$ with $W_\infty(\mu_0, \mu_\epsilon) \le \epsilon$. Then*
i) $\|\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0] - \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \le \epsilon$, ii) $\|\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_0] - \widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \le 2\epsilon$.

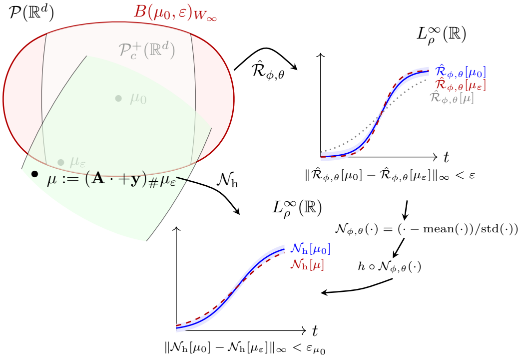
## Proposition

*Let $\mu_0 \in \mathcal{P}_c^*(\mathbb{R}^d)$ and $\mu_\epsilon \in \mathcal{P}(\mathbb{R}^d)$ with $W_\infty(\mu_0, \mu_\epsilon) \le \epsilon$,*

For $\mu \in \mathcal{P}_1(R^d)$ define the zero-mean quantile

$$\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu](t) := \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu](t) - \mathrm{mean}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu]) \quad \forall\, \boldsymbol{\theta} \in \mathbb{S}_1 \ \forall\, t \in \mathbb{R},$$

## Robustness of the Separability in Wasserstein Spaces for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$




**Proposition**

*Let $\mu_0, \mu_\epsilon \in \mathcal{P}(\mathbb{R}^2)$ with $W_\infty(\mu_0, \mu_\epsilon) \le \epsilon$. Then*
i) $\|\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0] - \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \le \epsilon$, ii) $\|\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_0] - \widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \le 2\epsilon$.

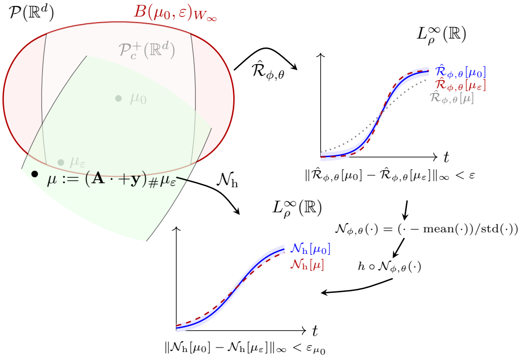
**Proposition**

*Let $\mu_0 \in \mathcal{P}_c^*(\mathbb{R}^d)$ and $\mu_\epsilon \in \mathcal{P}(\mathbb{R}^d)$ with $W_\infty(\mu_0, \mu_\epsilon) \le \epsilon$, and let $c_0 := \min_{\boldsymbol{\theta} \in \mathbb{S}_1} \mathrm{std}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0]) > 0$.*

For $\mu \in \mathcal{P}_1(R^d)$ define the zero-mean quantile

$$\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu](t) := \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu](t) - \mathrm{mean}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu]) \quad \forall \boldsymbol{\theta} \in \mathbb{S}_1 \ \forall t \in \mathbb{R},$$

# Robustness of the Separability in Wasserstein Spaces for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$




**Proposition**

*Let $\mu_0, \mu_\epsilon \in \mathcal{P}(\mathbb{R}^2)$ with $W_\infty(\mu_0, \mu_\epsilon) \le \epsilon$. Then*
i) $\|\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0] - \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \le \epsilon$, ii) $\|\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_0] - \widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \le 2\epsilon$.

**Proposition**

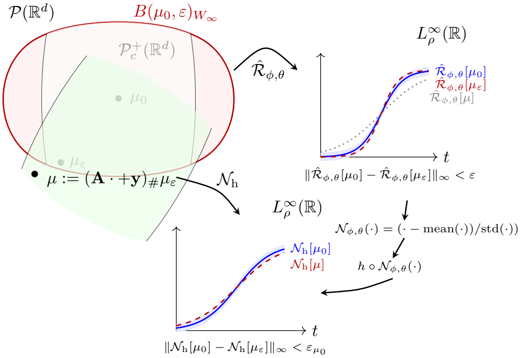
*Let $\mu_0 \in \mathcal{P}_c^*(\mathbb{R}^d)$ and $\mu_\epsilon \in \mathcal{P}(\mathbb{R}^d)$ with $W_\infty(\mu_0, \mu_\epsilon) \le \epsilon$, and let $c_0 := \min_{\boldsymbol{\theta}\in\mathbb{S}_1} \mathrm{std}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0]) > 0$. If $\epsilon < c_0/2$, then*

$$\|\mathcal{N}_{\mathrm{m}}[\mu_0] - \mathcal{N}_{\mathrm{m}}[\mu_\epsilon]\|_\infty \le \frac{4(\mathrm{diam}(\mu_0) + 2\epsilon)}{c_0(c_0 - 2\epsilon)}\epsilon.$$

For $\mu \in \mathcal{P}_1(R^d)$ define the zero-mean quantile

$$\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu](t) := \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu](t) - \mathrm{mean}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu]) \quad \forall \boldsymbol{\theta} \in \mathbb{S}_1 \ \forall t \in \mathbb{R},$$

# Robustness of the Separability in Wasserstein Spaces for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$




**Proposition**

*Let $\mu_0, \mu_\epsilon \in \mathcal{P}(\mathbb{R}^2)$ with $W_\infty(\mu_0, \mu_\epsilon) \leq \epsilon$. Then*
i) $\|\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0] - \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \leq \epsilon$, ii) $\|\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_0] - \widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \leq 2\epsilon$.

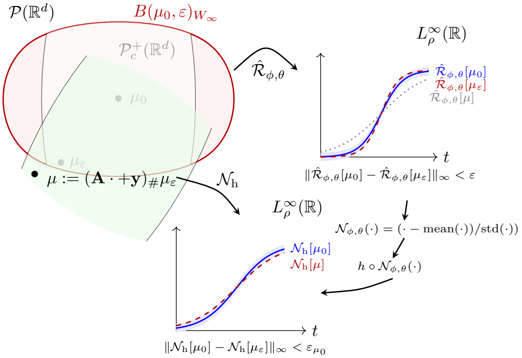
**Proposition**

*Let $\mu_0 \in \mathcal{P}_c^*(\mathbb{R}^d)$ and $\mu_\epsilon \in \mathcal{P}(\mathbb{R}^d)$ with $W_\infty(\mu_0, \mu_\epsilon) \leq \epsilon$, and let $c_0 := \min_{\boldsymbol{\theta} \in \mathbb{S}_1} \mathrm{std}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0]) > 0$. If $\epsilon < c_0/2$, then*
$\|\mathcal{N}_{\mathrm{m}}[(\mathbf{A} \cdot + \mathbf{y})_{\#}\mu_0] - \mathcal{N}_{\mathrm{m}}[(\mathbf{B} \cdot + \mathbf{z})_{\#}\mu_\epsilon]\|_\infty \leq \frac{4(\mathrm{diam}(\mu_0)+2\epsilon)}{c_0(c_0-2\epsilon)}\epsilon.$

For $\mu \in \mathcal{P}_1(R^d)$ define the zero-mean quantile

$$\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu](t) := \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu](t) - \mathrm{mean}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu]) \quad \forall \boldsymbol{\theta} \in \mathbb{S}_1 \ \forall t \in \mathbb{R},$$

# Robustness of the Separability in Wasserstein Spaces for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$




For $\mu \in \mathcal{P}_1(R^d)$ define the zero-mean quantile

$$\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu](t) := \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu](t) - \mathrm{mean}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu]) \quad \forall \boldsymbol{\theta} \in \mathbb{S}_1 \ \forall t \in \mathbb{R},$$

## Proposition

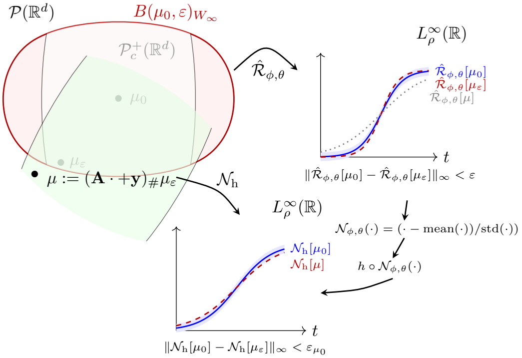
*Let $\mu_0, \mu_\epsilon \in \mathcal{P}(\mathbb{R}^2)$ with $W_\infty(\mu_0, \mu_\epsilon) \leq \epsilon$. Then*
i) $\|\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0] - \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \leq \epsilon$, ii) $\|\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_0] - \widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \leq 2\epsilon$.

## Proposition

*Let $\mu_0 \in \mathcal{P}_c^*(\mathbb{R}^d)$ and $\mu_\epsilon \in \mathcal{P}(\mathbb{R}^d)$ with $W_\infty(\mu_0, \mu_\epsilon) \leq \epsilon$, and let $c_0 := \min_{\boldsymbol{\theta} \in \mathbb{S}_1} \mathrm{std}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0]) > 0$. If $\epsilon < c_0/2$, then*
$$\|\mathcal{N}_{\mathrm{m}}[(\mathbf{A} \cdot + \mathbf{y})_{\#}\mu_0] - \mathcal{N}_{\mathrm{m}}[(\mathbf{B} \cdot + \mathbf{z})_{\#}\mu_\epsilon]\|_\infty \leq \frac{4(\mathrm{diam}(\mu_0) + 2\epsilon)}{c_0(c_0 - 2\epsilon)}\epsilon.$$

## Theorem (Separability with Robustness)

## Robustness of the Separability in Wasserstein Spaces for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$




For $\mu \in \mathcal{P}_1(R^d)$ define the zero-mean quantile

$$\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu](t) := \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu](t) - \mathrm{mean}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu]) \quad \forall \boldsymbol{\theta} \in \mathbb{S}_1 \ \forall t \in \mathbb{R},$$

### Proposition

*Let $\mu_0, \mu_\epsilon \in \mathcal{P}(\mathbb{R}^2)$ with $W_\infty(\mu_0, \mu_\epsilon) \le \epsilon$. Then*
i) $\|\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0] - \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \le \epsilon$, ii) $\|\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_0] - \widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \le 2\epsilon$.

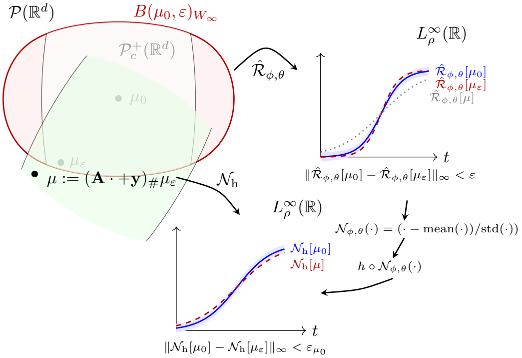
### Proposition

*Let $\mu_0 \in \mathcal{P}_c^*(\mathbb{R}^d)$ and $\mu_\epsilon \in \mathcal{P}(\mathbb{R}^d)$ with $W_\infty(\mu_0, \mu_\epsilon) \le \epsilon$, and let $c_0 := \min_{\boldsymbol{\theta} \in \mathbb{S}_1} \mathrm{std}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0]) > 0$. If $\epsilon < c_0/2$, then*
$\|\mathcal{N}_\mathrm{m}[(\mathbf{A}\cdot + \mathbf{y})_\# \mu_0] - \mathcal{N}_\mathrm{m}[(\mathbf{B}\cdot + \mathbf{z})_\# \mu_\epsilon]\|_\infty \le \frac{4(\mathrm{diam}(\mu_0) + 2\epsilon)}{c_0(c_0 - 2\epsilon)}\epsilon.$

### Theorem (Separability with Robustness)

*Let $\mu_0, \nu_0 \in \mathcal{P}_c^*(\mathbb{R}^d)$ with $\mathcal{N}_\mathrm{m}[\mu_0] \ne \mathcal{N}_\mathrm{m}[\nu_0]$,*

## Robustness of the Separability in Wasserstein Spaces for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$




For $\mu \in \mathcal{P}_1(R^d)$ define the zero-mean quantile

$$\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu](t) := \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu](t) - \mathrm{mean}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu]) \quad \forall \boldsymbol{\theta} \in \mathbb{S}_1 \ \forall t \in \mathbb{R},$$

### Proposition

*Let* $\mu_0, \mu_\epsilon \in \mathcal{P}(\mathbb{R}^2)$ *with* $W_\infty(\mu_0, \mu_\epsilon) \leq \epsilon$. *Then*
i) $\|\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0] - \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \leq \epsilon$, ii) $\|\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_0] - \widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \leq 2\epsilon$.

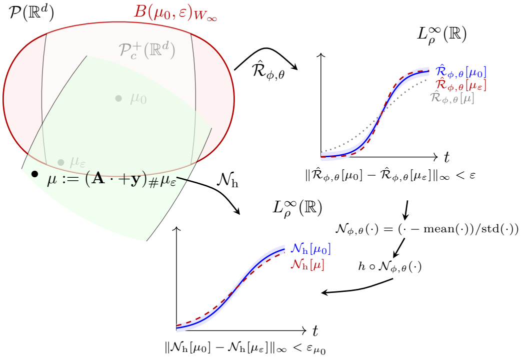
### Proposition

*Let* $\mu_0 \in \mathcal{P}_c^*(\mathbb{R}^d)$ *and* $\mu_\epsilon \in \mathcal{P}(\mathbb{R}^d)$ *with* $W_\infty(\mu_0, \mu_\epsilon) \leq \epsilon$, *and let* $c_0 := \min_{\boldsymbol{\theta} \in \mathbb{S}_1} \mathrm{std}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0]) > 0$. *If* $\epsilon < c_0/2$, *then*
$\|\mathcal{N}_{\mathrm{m}}[(\mathbf{A}\cdot+\mathbf{y})_{\#}\mu_0] - \mathcal{N}_{\mathrm{m}}[(\mathbf{B}\cdot+\mathbf{z})_{\#}\mu_\epsilon]\|_\infty \leq \frac{4(\mathrm{diam}(\mu_0)+2\epsilon)}{c_0(c_0-2\epsilon)}\epsilon.$

### Theorem (Separability with Robustness)

*Let* $\mu_0, \nu_0 \in \mathcal{P}_c^*(\mathbb{R}^d)$ *with* $\mathcal{N}_{\mathrm{m}}[\mu_0] \neq \mathcal{N}_{\mathrm{m}}[\nu_0]$, *for* $\bullet \in \{\mu, \nu\}$ *let* $c_\bullet := \min_{\boldsymbol{\theta} \in \mathbb{S}_1} \mathrm{std}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\bullet_0])$.

# Robustness of the Separability in Wasserstein Spaces for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$




For $\mu \in \mathcal{P}_1(R^d)$ define the zero-mean quantile

$$\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu](t) := \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu](t) - \mathrm{mean}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu]) \quad \forall \boldsymbol{\theta} \in \mathbb{S}_1 \ \forall t \in \mathbb{R},$$

## Proposition

*Let $\mu_0, \mu_\epsilon \in \mathcal{P}(\mathbb{R}^2)$ with $W_\infty(\mu_0, \mu_\epsilon) \le \epsilon$. Then*
i) $\|\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0] - \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \le \epsilon$, ii) $\|\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_0] - \widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \le 2\epsilon$.

## Proposition

*Let $\mu_0 \in \mathcal{P}_c^*(\mathbb{R}^d)$ and $\mu_\epsilon \in \mathcal{P}(\mathbb{R}^d)$ with $W_\infty(\mu_0, \mu_\epsilon) \le \epsilon$, and let $c_0 := \min_{\boldsymbol{\theta} \in \mathbb{S}_1} \mathrm{std}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0]) > 0$. If $\epsilon < c_0/2$, then*

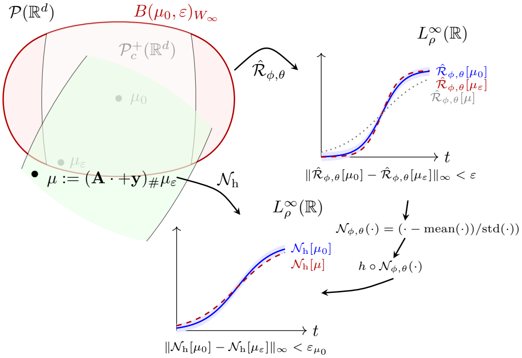
$$\|\mathcal{N}_{\mathrm{m}}[(\mathbf{A} \cdot + \mathbf{y})_{\#}\mu_0] - \mathcal{N}_{\mathrm{m}}[(\mathbf{B} \cdot + \mathbf{z})_{\#}\mu_\epsilon]\|_\infty \le \frac{4(\mathrm{diam}(\mu_0) + 2\epsilon)}{c_0(c_0 - 2\epsilon)}\epsilon.$$

## Theorem (Separability with Robustness)

*Let $\mu_0, \nu_0 \in \mathcal{P}_c^*(\mathbb{R}^d)$ with $\mathcal{N}_{\mathrm{m}}[\mu_0] \neq \mathcal{N}_{\mathrm{m}}[\nu_0]$, for $\bullet \in \{\mu, \nu\}$ let $c_\bullet := \min_{\boldsymbol{\theta} \in \mathbb{S}_1} \mathrm{std}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\bullet_0])$. Let $\epsilon < \frac{\min\{c_\mu, c_\nu\}}{2}$ satisfy*

$$4\epsilon\left(\frac{\mathrm{diam}(\mu_0) + 2\epsilon}{c_\mu(c_\mu - 2\epsilon)} + \frac{\mathrm{diam}(\nu_0) + 2\epsilon}{c_\nu(c_\nu - 2\epsilon)}\right) < \|\mathcal{N}_{\mathrm{m}}[\mu_0] - \mathcal{N}_{\mathrm{m}}[\nu_0]\|_\infty.$$

# Robustness of the Separability in Wasserstein Spaces for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$




For $\mu \in \mathcal{P}_1(R^d)$ define the zero-mean quantile

$$\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu](t) := \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu](t) - \mathrm{mean}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu]) \quad \forall \boldsymbol{\theta} \in \mathbb{S}_1 \; \forall t \in \mathbb{R},$$

## Proposition

*Let* $\mu_0, \mu_\epsilon \in \mathcal{P}(\mathbb{R}^2)$ *with* $W_\infty(\mu_0, \mu_\epsilon) \le \epsilon$. *Then*
i) $\|\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0] - \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \le \epsilon$, ii) $\|\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_0] - \widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \le 2\epsilon$.

## Proposition

*Let* $\mu_0 \in \mathcal{P}_c^*(\mathbb{R}^d)$ *and* $\mu_\epsilon \in \mathcal{P}(\mathbb{R}^d)$ *with* $W_\infty(\mu_0, \mu_\epsilon) \le \epsilon$, *and let* $c_0 := \min_{\boldsymbol{\theta} \in \mathbb{S}_1} \mathrm{std}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0]) > 0$. *If* $\epsilon < c_0/2$, *then*
$\|\mathcal{N}_{\mathrm{m}}[(\mathbf{A} \cdot + \mathbf{y})_\# \mu_0] - \mathcal{N}_{\mathrm{m}}[(\mathbf{B} \cdot + \mathbf{z})_\# \mu_\epsilon]\|_\infty \le \frac{4(\mathrm{diam}(\mu_0) + 2\epsilon)}{c_0(c_0 - 2\epsilon)} \epsilon$.

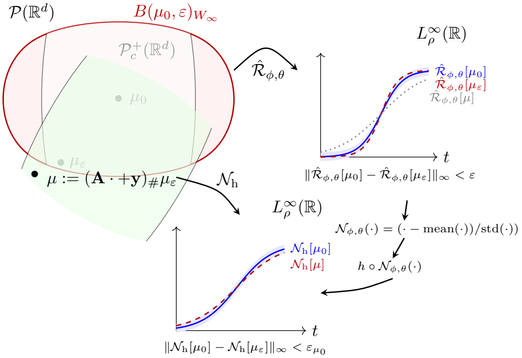
## Theorem (Separability with Robustness)

*Let* $\mu_0, \nu_0 \in \mathcal{P}_c^*(\mathbb{R}^d)$ *with* $\mathcal{N}_{\mathrm{m}}[\mu_0] \neq \mathcal{N}_{\mathrm{m}}[\nu_0]$, *for* $\bullet \in \{\mu, \nu\}$ *let* $c_\bullet := \min_{\boldsymbol{\theta} \in \mathbb{S}_1} \mathrm{std}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\bullet_0])$. *Let* $\epsilon < \frac{\min\{c_\mu, c_\nu\}}{2}$ *satisfy*

$$4\epsilon \left( \frac{\mathrm{diam}(\mu_0) + 2\epsilon}{c_\mu (c_\mu - 2\epsilon)} + \frac{\mathrm{diam}(\nu_0) + 2\epsilon}{c_\nu (c_\nu - 2\epsilon)} \right) < \|\mathcal{N}_{\mathrm{m}}[\mu_0] - \mathcal{N}_{\mathrm{m}}[\nu_0]\|_\infty .$$

*Let* $\mathbb{F}_\bullet := \left\{ (\mathbf{A} \cdot + \mathbf{y})_\# \bullet \mid \bullet \in \mathcal{P}(\mathbb{R}^d), W_\infty(\bullet, \bullet_0) \le \epsilon \right\}$.

# Robustness of the Separability in Wasserstein Spaces for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$




For $\mu \in \mathcal{P}_1(R^d)$ define the zero-mean quantile

$$\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu](t) := \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu](t) - \mathrm{mean}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu]) \quad \forall \boldsymbol{\theta} \in \mathbb{S}_1 \ \forall t \in \mathbb{R},$$

## Proposition

*Let $\mu_0, \mu_\epsilon \in \mathcal{P}(\mathbb{R}^2)$ with $W_\infty(\mu_0, \mu_\epsilon) \leq \epsilon$. Then*
i) $\|\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0] - \widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \leq \epsilon$, ii) $\|\widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_0] - \widetilde{\mathcal{N}}_{\boldsymbol{\theta}}[\mu_\epsilon]\|_\infty \leq 2\epsilon$.

## Proposition

*Let $\mu_0 \in \mathcal{P}_c^*(\mathbb{R}^d)$ and $\mu_\epsilon \in \mathcal{P}(\mathbb{R}^d)$ with $W_\infty(\mu_0, \mu_\epsilon) \leq \epsilon$, and let $c_0 := \min_{\boldsymbol{\theta} \in \mathbb{S}_1} \mathrm{std}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\mu_0]) > 0$. If $\epsilon < c_0/2$, then*
$\|\mathcal{N}_\mathrm{m}[(\mathbf{A} \cdot + \mathbf{y})_\# \mu_0] - \mathcal{N}_\mathrm{m}[(\mathbf{B} \cdot + \mathbf{z})_\# \mu_\epsilon]\|_\infty \leq \frac{4(\mathrm{diam}(\mu_0) + 2\epsilon)}{c_0(c_0 - 2\epsilon)} \epsilon.$

## Theorem (Separability with Robustness)

*Let $\mu_0, \nu_0 \in \mathcal{P}_c^*(\mathbb{R}^d)$ with $\mathcal{N}_\mathrm{m}[\mu_0] \neq \mathcal{N}_\mathrm{m}[\nu_0]$, for $\bullet \in \{\mu, \nu\}$ let $c_\bullet := \min_{\boldsymbol{\theta} \in \mathbb{S}_1} \mathrm{std}(\widehat{\mathcal{R}}_{\boldsymbol{\theta}}[\bullet_0])$. Let $\epsilon < \frac{\min\{c_\mu, c_\nu\}}{2}$ satisfy*

$$4\epsilon \left( \frac{\mathrm{diam}(\mu_0) + 2\epsilon}{c_\mu(c_\mu - 2\epsilon)} + \frac{\mathrm{diam}(\nu_0) + 2\epsilon}{c_\nu(c_\nu - 2\epsilon)} \right) < \|\mathcal{N}_\mathrm{m}[\mu_0] - \mathcal{N}_\mathrm{m}[\nu_0]\|_\infty.$$

*Let $\mathbb{F}_\bullet := \left\{ (\mathbf{A} \cdot + \mathbf{y})_\# \bullet \mid \bullet \in \mathcal{P}(\mathbb{R}^d),\ W_\infty(\bullet, \bullet_0) \leq \epsilon \right\}$.*
*Then, $\tilde{\mathbb{F}}_\mu \subset \mathbb{F}_\mu$ and $\tilde{\mathbb{F}}_\nu \subset \mathbb{F}_\nu$ are linearly separable.*

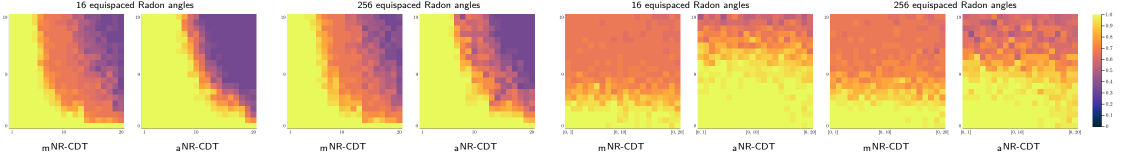



Figure: Phase transition of NT classification accuracies for 3-class academic datasets with 10 samples per class. **Left:** different salt noise (vertical: component numbers, horizontal: noise strengths); **Right:** non-affine deformations with frequencies in $[0, 2]$ and variable amplitude ranges (horizontal), as well as salt noise of strength 9 and variable location numbers (vertical).

# Numerics – Robustness Study by Salt Noise and Elastic Deformations for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$

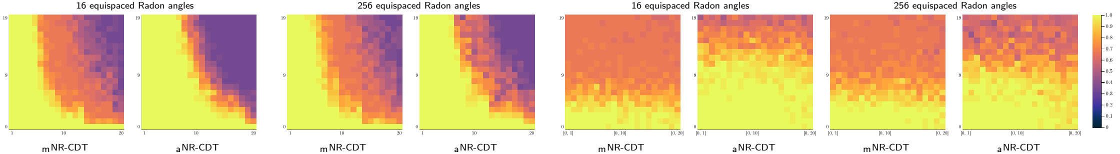
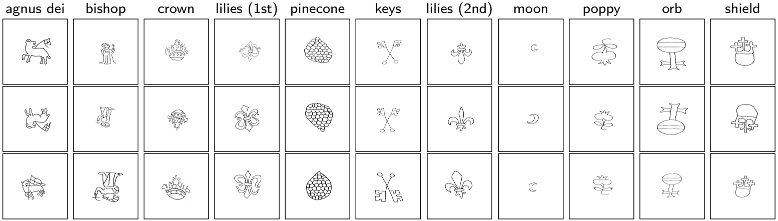



Figure: Phase transition of NT classification accuracies for 3-class academic datasets with 10 samples per class. **Left:** different salt noise (vertical: component numbers, horizontal: noise strengths); **Right:** non-affine deformations with frequencies in $[0, 2]$ and variable amplitude ranges (horizontal), as well as salt noise of strength 9 and variable location numbers (vertical).



# Numerics – Robustness Study by Salt Noise and Elastic Deformations for $\mathbb{X} = \mathbb{R}^d, \Theta = \mathbb{S}_{d-1}$

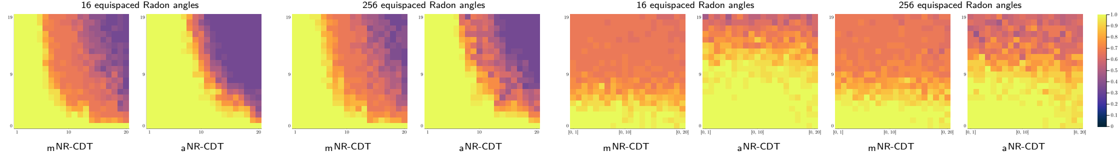



Figure: Phase transition of NT classification accuracies for 3-class academic datasets with 10 samples per class. **Left:** different salt noise (vertical: component numbers, horizontal: noise strengths); **Right:** non-affine deformations with frequencies in $[0, 2]$ and variable amplitude ranges (horizontal), as well as salt noise of strength 9 and variable location numbers (vertical).

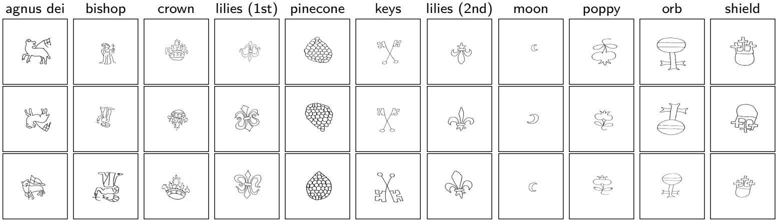
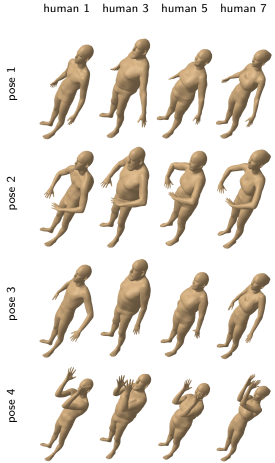


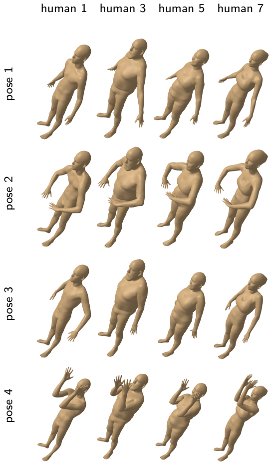

Table: NT classification accuracies for the historic watermark dataset.

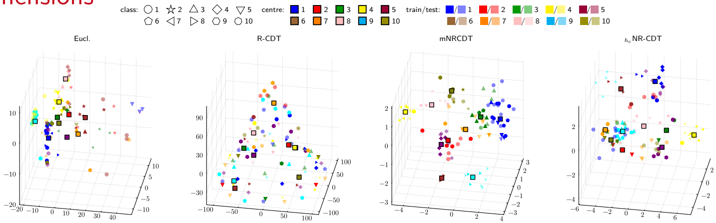
| # angles | Eucl. | MB | R-CDT | $_{\mathrm{m}}$NR-CDT | $_{h_{\mathrm{d}}}$NR-CDT | $_{\mathrm{tv}}$NR-CDT |
|---|---|---|---|---|---|---|
| 2 | | | 0.140 | 0.516 | 0.203 | 0.200 |
| 4 | | | 0.219 | **0.781** | 0.703 | 0.719 |
| 8 | | | 0.188 | **0.828** | 0.782 | 0.766 |
| 16 | | | 0.203 | **0.859** | 0.781 | 0.766 |
| 32 | 0.078 | 0.406 | 0.203 | **0.859** | 0.800 | 0.822 |
| 64 | | | 0.203 | 0.859 | 0.813 | **0.891** |
| 128 | | | 0.203 | 0.859 | 0.813 | **0.891** |
| 256 | | | 0.203 | 0.859 | 0.813 | **0.891** |
| 512 | | | 0.203 | 0.859 | 0.813 | **0.922** |
| 1024 | | | 0.203 | 0.859 | 0.796 | **0.938** |

# Numerics – Clustering in Higher Dimensions



# Numerics – Clustering in Higher Dimensions

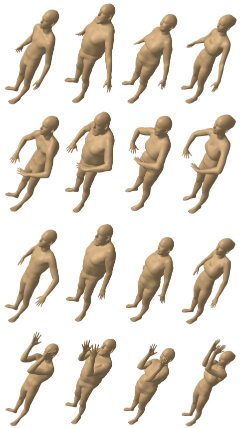


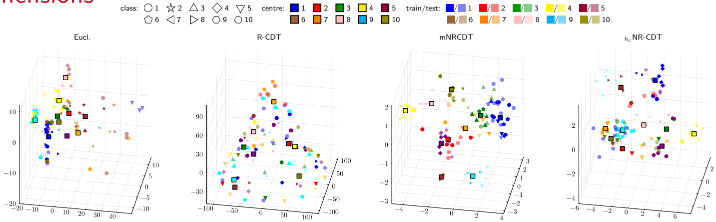



# Numerics – Clustering in Higher Dimensions

human 1 human 3 human 5 human 7

pose 1
pose 2
pose 3
pose 4

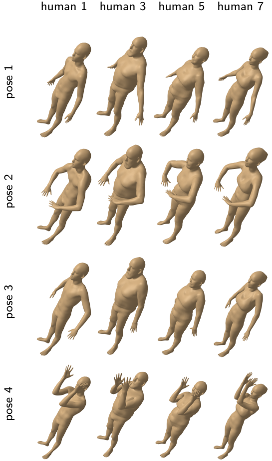


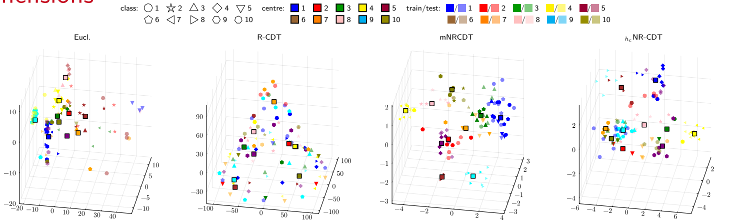



| train | Eucl. | R-CDT | $_{\mathrm{m}}$NR-CDT | $_{h_c}$NR-CDT |
|---|---|---|---|---|
| 1 | $0.658 \pm 0.054$ | $0.475 \pm 0.060$ | $\mathbf{0.706 \pm 0.044}$ | $0.672 \pm 0.030$ |
| 3 | $0.727 \pm 0.035$ | $0.575 \pm 0.041$ | $\mathbf{0.801 \pm 0.050}$ | $0.784 \pm 0.042$ |
| 5 | $0.757 \pm 0.037$ | $0.591 \pm 0.040$ | $\mathbf{0.838 \pm 0.044}$ | $0.828 \pm 0.038$ |

Table: 1-NN classification accuracies for the FAUST dataset with 10 samples per class and using 256 directions.

# Numerics – Clustering in Higher Dimensions




| train | Eucl. | R-CDT | $_{\rm m}$NR-CDT | $_{h_c}$NR-CDT |
|---|---|---|---|---|
| 1 | $0.658 \pm 0.054$ | $0.475 \pm 0.060$ | $\mathbf{0.706 \pm 0.044}$ | $0.672 \pm 0.030$ |
| 3 | $0.727 \pm 0.035$ | $0.575 \pm 0.041$ | $\mathbf{0.801 \pm 0.050}$ | $0.784 \pm 0.042$ |
| 5 | $0.757 \pm 0.037$ | $0.591 \pm 0.040$ | $\mathbf{0.838 \pm 0.044}$ | $0.828 \pm 0.038$ |

Table: 1-NN classification accuracies for the FAUST dataset with 10 samples per class and using 256 directions.

| | Eucl. | R-CDT | $_{\rm m}$NR-CDT | $_{h_c}$NR-CDT |
|---|---|---|---|---|
| $\mathrm{RI}_{\mathrm{train}}$ $(\uparrow)$ | 0.8261 | 0.8424 | 0.9208 | **0.9208** |
| $\mathrm{VI}_{\mathrm{train}}$ $(\downarrow)$ | 1.7299 | 2.7545 | **0.8316** | 0.9300 |
| $\mathrm{RI}_{\mathrm{test}}$ $(\uparrow)$ | 0.7910 | 0.8424 | 0.9346 | **0.9444** |
| $\mathrm{VI}_{\mathrm{test}}$ $(\downarrow)$ | 1.6587 | 2.7964 | **0.6975** | 0.7074 |

Table: Quality measures for 10-means clustering of the FAUST dataset with 10 samples per class, where 5 samples are used for training and the rest for testing.

## Included Publications


[1] Beckmann, M., Beinert, R., and Bresch, J. Max-normalized Radon CDT for limited data classification. In: **Lect. Notes Comput. Sci.** Lect. Notes Comput. Sci. 15667 (2025), pp. 241–254. DOI: 10.1007/978-3-031-92366-1_19.

[2] Beckmann, M., Beinert, R., and Bresch, J. Normalized Radon CDT for invariance and robustness in optimal transport based image classification. In: **SIAM J. Imaging Sci.** (2025). arXiv:2506.08761, accepted in SIAM J. Imaging Sci. (December, 2025). DOI: 10.48550/arXiv.2506.08761.

[3] Beckmann, M., Beinert, R., and Bresch, J. Generalization of the normalized Radon CDT for limited data recognition. In: **arXiv.2512.08099** (2025). arXiv:2512.08099, submitted to J. Math. Imaging Vis., September. DOI: 10.48550/arXiv.2512.08099.

# References I


[1] Hauser, D. et al. “On image processing and pattern recognition for thermograms of watermarks in manuscripts – A first proof-of-concept.” In: **International Conference on Document Analysis and Recognition (ICDAR)**. Vol. 14806. Lect. Notes Comput. Sci. Athens, Greece: Springer, 2024, pp. 91–107. DOI: 10.1007/978-3-031-70543-4_6.

[2] Wang, W. et al. A linear optimal transportation framework for quantifying and visualizing variations in sets of images. In: **Int. J. Comput. Vis.** 101.2 (2013), pp. 254–269. DOI: 10.1007/s11263-012-0566-z.

[3] Moosmüller, C. and Cloninger, A. Linear optimal transport embedding: Provable Wasserstein classification for certain rigid transformations and perturbations. In: **IMA J. Appl. Math.** 12.1 (2023), pp. 363–389. DOI: 10.1093/imaiai/iaac023.

[4] Kolouri, S., Park, S. R., and Rohde, G. K. The Radon cumulative distribution transform and its application to image classification. In: **IEEE Trans. Image Process.** 25.2 (2016), pp. 920–934. DOI: 10.1109/TIP.2015.2509419.

[5] Natterer, F. The Mathematics of Computerized Tomography. Classics in Applied Mathematics. Philadelphia, PA, USA: SIAM, 2001, pp. xiv + 226. DOI: 10.1137/1.9780898719284.

# References II


[6] Deng, L. The MNIST database of handwritten digit images for machine learning research. In: **IEEE Signal Process. Mag.** 29.6 (2012), pp. 141–142. DOI: 10.1109/MSP.2012.2211477.

[7] Liu, C.-L. et al. "CASIA online and offline Chinese handwriting databases." In: **International Conference on Document Analysis and Recognition (ICDAR)**. Beijing, China: IEEE, 2011, pp. 37–41. DOI: 10.1109/ICDAR.2011.17.

[8] Koch, S. et al. "ABC: A big CAD model dataset for geometric deep learning." In: **Conference on Computer Vision and Pattern Recognition (CVPR)**. Long Beach, CA, USA: IEEE, 2019, pp. 9593–9603. DOI: 10.1109/CVPR.2019.00983.

[9] Bogo, F. et al. "FAUST: Dataset and evaluation for 3D mesh registration." In: **Conference on Computer Vision and Pattern Recognition (CVPR)**. Columbus, OH, USA: IEEE, 2014, pp. 3794–3801. DOI: 10.1109/CVPR.2014.491.

[10] Flusser, J. and Suk, T. Pattern recognition by affine moment invariants. In: **Pattern Recognit.** 26.1 (1993), pp. 167–174. DOI: 10.1016/0031-3203(93)90098-H.

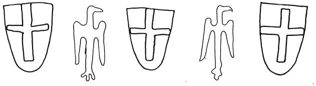
## References III


[11] Diaz Martin, R., Medri, I. V., and Rohde, G. K. Data representation with optimal transport. arXiv:2406.15503. 2024. DOI: 10.48550/arXiv.2406.15503.

[12] Villani, C. Topics in Optimal Transportation. 1st ed. Vol. 58. Graduate Studies in Mathematics. Providence, RI, USA: AMS, 2003, p. 370. DOI: 10.1090/gsm/058.

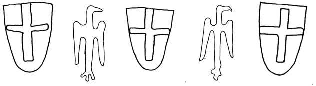
## Motivation: Objects as Prob. Measures

- Classification with **many classes** and **few samples**
- Assumption: (approx.) **affine** / structured class families
- Need: a) **invariance** to affine transformations
  b) **stability** to deformations
- Application: historic watermarks, automated pattern recognition



## Motivation: Objects as Prob. Measures

- Classification with **many classes** and **few samples**
- Assumption: (approx.) **affine** / structured class families
- Need: a) **invariance** to affine transformations
  b) **stability** to deformations
- Application: historic watermarks, automated pattern recognition



## Radon on Measures & Core Idea

- Restricted Radon transfrom via measurable $\phi : \mathbb{X} \times \Theta \to \mathbb{R}$

$$\mathcal{R}_{\phi,\theta} : \mathcal{M}(\mathbb{X}) \to \mathcal{M}(\mathbb{R})$$
$$\mu \mapsto (S_{\phi,\theta})_{\#}\mu, \ S_{\phi,\theta} \coloneqq \phi(\cdot, \boldsymbol{\theta})$$

- Slice per direction

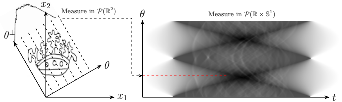
$$\mu \in \mathcal{P}(\mathbb{X}) \quad \Rightarrow \quad \mathcal{R}_{\phi,\theta}[\mu] \in \mathcal{P}(\mathbb{R})$$

- 1D embedding via CDT, wrt. ref $\rho$

$$\widehat{\mu} = \underset{T_{\#}\rho=\mu}{\arg\min} \int_{\mathbb{R}} c(s - T(s))\, \mathrm{d}\rho(s), \quad \text{atomless}\, \rho, \mu \in \mathcal{P}(\mathbb{R})$$

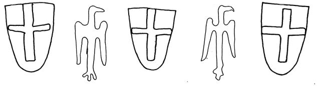
- R-CDT: collect sliced CDTs

$$\widehat{\mathcal{R}_{\phi,\theta}[\mu]} \in \mathrm{L}^2_\rho(\mathbb{R}) \text{ for } \mu \in \mathcal{P}_2(\mathbb{X})$$



## Motivation: Objects as Prob. Measures

- Classification with **many classes** and **few samples**
- Assumption: (approx.) **affine** / structured class families
- Need: a) **invariance** to affine transformations
  b) **stability** to deformations
- Application: historic watermarks, automated pattern recognition



## Radon on Measures & Core Idea

- Restricted Radon transfrom via measurable $\phi : \mathbb{X} \times \Theta \to \mathbb{R}$

$$\mathcal{R}_{\phi,\theta} : \mathcal{M}(\mathbb{X}) \to \mathcal{M}(\mathbb{R})$$
$$\mu \mapsto (S_{\phi,\theta})_{\#}\mu, \; S_{\phi,\theta} \coloneqq \phi(\cdot, \boldsymbol{\theta})$$

- Slice per direction

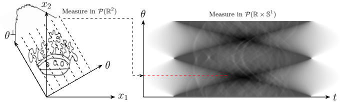
$$\mu \in \mathcal{P}(\mathbb{X}) \quad \Rightarrow \quad \mathcal{R}_{\phi,\theta}[\mu] \in \mathcal{P}(\mathbb{R})$$

- 1D embedding via CDT, wrt. ref $\rho$

$$\widehat{\mu} = \operatorname*{arg\,min}_{T_{\#}\rho=\mu} \int_{\mathbb{R}} c(s - T(s))\,\mathrm{d}\rho(s), \quad \text{atomless } \rho, \mu \in \mathcal{P}(\mathbb{R})$$

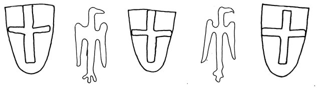
- R-CDT: collect sliced CDTs

$$\widehat{\mathcal{R}_{\phi,\theta}[\mu]} \in \mathrm{L}^2_\rho(\mathbb{R}) \text{ for } \mu \in \mathcal{P}_2(\mathbb{X})$$



## Motivation: Objects as Prob. Measures

- Classification with **many classes** and **few samples**
- Assumption: (approx.) **affine** / structured class families
- Need: a) **invariance** to affine transformations
  b) **stability** to deformations
- Application: historic watermarks, automated pattern recognition



## Normalization: NR-CDT and Sorting

- For $\mu \in \mathcal{P}^{+}_{\phi,\mathrm{c}}(\mathbb{X}) \subset \mathcal{P}_{\mathrm{c}}(\mathbb{X}) \subset \mathcal{P}_2(\mathbb{X})$ R-CDT is bounded with
  $$\mathrm{std}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]}) \geq c > 0 \quad \text{f.a.} \quad \theta \in \Theta.$$
- **Normalize** per slice:
  $$\mathcal{N}_{\phi,\theta}[\mu](t) = \frac{\widehat{\mathcal{R}_{\phi,\theta}[\mu]}(t) - \mathrm{mean}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]})}{\mathrm{std}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]})}$$
- **Sorting** the directions $\theta \in \Theta$ via:
  $$h : \mathrm{L}^{\infty}(\Theta) \to \mathbb{R}$$
- **Define** ${}_h$NR-CDT: $h\Big(\mathcal{N}_{\phi}[\mu](t,\cdot)\Big)$

## Radon on Measures & Core Idea

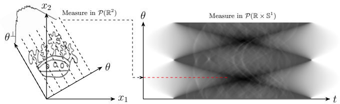
- Restricted Radon transfrom via measurable $\phi : \mathbb{X} \times \Theta \to \mathbb{R}$

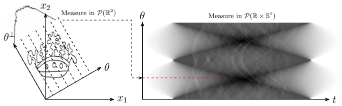
$$\mathcal{R}_{\phi,\theta} : \mathcal{M}(\mathbb{X}) \to \mathcal{M}(\mathbb{R})$$
$$\mu \mapsto (S_{\phi,\theta})_{\#}\mu, \; S_{\phi,\theta} \coloneqq \phi(\cdot,\boldsymbol{\theta})$$

- Slice per direction
  $$\mu \in \mathcal{P}(\mathbb{X}) \quad \Rightarrow \quad \mathcal{R}_{\phi,\theta}[\mu] \in \mathcal{P}(\mathbb{R})$$
- 1D embedding via CDT, wrt. ref $\rho$

$$\widehat{\mu} = \mathop{\arg\min}_{T_{\#}\rho=\mu} \int_{\mathbb{R}} c(s - T(s))\,\mathrm{d}\rho(s), \quad \text{atomless } \rho, \mu \in \mathcal{P}(\mathbb{R})$$

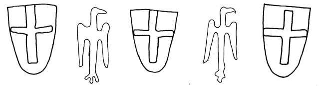
- R-CDT: collect sliced CDTs
  $$\widehat{\mathcal{R}_{\phi,\theta}[\mu]} \in \mathrm{L}^2_{\rho}(\mathbb{R}) \text{ for } \mu \in \mathcal{P}_2(\mathbb{X})$$






## Motivation: Objects as Prob. Measures

- Classification with **many classes** and **few samples**
- Assumption: (approx.) **affine** / structured class families
- Need: a) **invariance** to affine transformations
  b) **stability** to deformations
- Application: historic watermarks, automated pattern recognition



## Normalization: NR-CDT and Sorting

- For $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) \subset \mathcal{P}_\mathrm{c}(\mathbb{X}) \subset \mathcal{P}_2(\mathbb{X})$ R-CDT is bounded with
  $$\mathrm{std}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]}) \geq c > 0 \quad \text{f.a.} \quad \theta \in \Theta.$$
- **Normalize** per slice:
  $$\mathcal{N}_{\phi,\theta}[\mu](t) = \frac{\widehat{\mathcal{R}_{\phi,\theta}[\mu]}(t) - \mathrm{mean}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]})}{\mathrm{std}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]})}$$
- **Sorting** the directions $\theta \in \Theta$ via:
  $$h : \mathrm{L}^\infty(\Theta) \to \mathbb{R}$$
- **Define** $_\mathrm{h}$NR-CDT: $h\Big(\mathcal{N}_\phi[\mu](t,\cdot)\Big)$

## Radon on Measures & Core Idea

- Restricted Radon transfrom via measurable $\phi : \mathbb{X} \times \Theta \to \mathbb{R}$

$$\mathcal{R}_{\phi,\theta} : \mathcal{M}(\mathbb{X}) \to \mathcal{M}(\mathbb{R})$$
$$\mu \mapsto (S_{\phi,\theta})_\# \mu, \; S_{\phi,\theta} \coloneqq \phi(\cdot, \boldsymbol{\theta})$$

- Slice per direction
  $$\mu \in \mathcal{P}(\mathbb{X}) \quad \Rightarrow \quad \mathcal{R}_{\phi,\theta}[\mu] \in \mathcal{P}(\mathbb{R})$$
- 1D embedding via CDT, wrt. ref $\rho$

$$\widehat{\mu} = \operatorname*{arg\,min}_{T_\# \rho = \mu} \int_\mathbb{R} c(s - T(s))\,\mathrm{d}\rho(s), \quad \text{atomless } \rho, \mu \in \mathcal{P}(\mathbb{R})$$

- R-CDT: collect sliced CDTs
  $$\widehat{\mathcal{R}_{\phi,\theta}[\mu]} \in \mathrm{L}^2_\rho(\mathbb{R}) \text{ for } \mu \in \mathcal{P}_2(\mathbb{X})$$

## Theoretical Results

- **Translation:** for suitable $(\mathbb{X}, \Theta, \phi)$
  $$\widehat{\mathcal{R}_\phi[T_\# \mu]} = \hat{T} \circ \widehat{\mathcal{R}_\phi[\mu]}$$
  for specific $T$ correspond to $\hat{T}$

i) **Invariance:** $\mathcal{N}_{h,\phi}[T_\# \mu] = \mathcal{N}_{h,\phi}[\mu]$

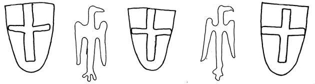
ii) **Singleton sets** in $_\mathrm{h}$NR-CDT-space
$$\mathbb{F}_\mu = \{T_\# \mu\} \xrightarrow{{}_\mathrm{h}\text{NR-CDT}} \mathcal{N}_{h,\phi}[\mathbb{F}_\mu] = \mathcal{N}_{h,\phi}[\mu]$$

iii) **Robustness** of $_\mathrm{m}$NR-CDT and $_\mathrm{a}$NR-CDT in Wasserstein-spaces
$$h \in \Big\{\sup_{\boldsymbol{\theta}\in\Theta}, \int_\Theta\Big\}.$$

## Motivation: Objects as Prob. Measures

- Classification with **many classes** and **few samples**
- Assumption: (approx.) **affine** / structured class families
- Need: a) **invariance** to affine transformations

  b) **stability** to deformations
- Application: historic watermarks, automated pattern recognition



## Normalization: NR-CDT and Sorting

- For $\mu \in \mathcal{P}^{+}_{\phi,\mathrm{c}}(\mathbb{X}) \subset \mathcal{P}_{\mathrm{c}}(\mathbb{X}) \subset \mathcal{P}_2(\mathbb{X})$ R-CDT is bounded with

$$\operatorname{std}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]}) \geq c > 0 \quad \text{f.a.} \quad \theta \in \Theta.$$

- **Normalize** per slice:

$$\mathcal{N}_{\phi,\theta}[\mu](t) = \frac{\widehat{\mathcal{R}_{\phi,\theta}[\mu]}(t) - \operatorname{mean}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]})}{\operatorname{std}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]})}$$

- **Sorting** the directions $\theta \in \Theta$ via:

$$h : \mathrm{L}^\infty(\Theta) \to \mathbb{R}$$

- **Define** ${}_{\mathrm{h}}$NR-CDT: $h\Big(\mathcal{N}_\phi[\mu](t,\cdot)\Big)$

## Radon on Measures & Core Idea

- Restricted Radon transfrom via measurable $\phi : \mathbb{X} \times \Theta \to \mathbb{R}$

$$\mathcal{R}_{\phi,\theta} : \mathcal{M}(\mathbb{X}) \to \mathcal{M}(\mathbb{R})$$
$$\mu \mapsto (S_{\phi,\theta})_{\#}\mu, \; S_{\phi,\theta} \coloneqq \phi(\cdot, \boldsymbol{\theta})$$

- Slice per direction

$$\mu \in \mathcal{P}(\mathbb{X}) \quad \Rightarrow \quad \mathcal{R}_{\phi,\theta}[\mu] \in \mathcal{P}(\mathbb{R})$$

- 1D embedding via CDT, wrt. ref $\rho$

$$\widehat{\mu} = \operatorname*{arg\,min}_{T_{\#}\rho=\mu} \int_{\mathbb{R}} c(s - T(s))\,\mathrm{d}\rho(s), \quad \text{atomless } \rho, \mu \in \mathcal{P}(\mathbb{R})$$

- R-CDT: collect sliced CDTs

$$\widehat{\mathcal{R}_{\phi,\theta}[\mu]} \in \mathrm{L}^2_\rho(\mathbb{R}) \text{ for } \mu \in \mathcal{P}_2(\mathbb{X})$$

## Theoretical Results

- **Translation:** for suitable $(\mathbb{X}, \Theta, \phi)$

$$\widehat{\mathcal{R}_\phi[T_{\#}\mu]} = \hat{T} \circ \widehat{\mathcal{R}_\phi[\mu]}$$

  for specific $T$ correspond to $\hat{T}$

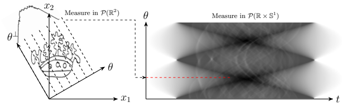
i) **Invariance:** $\mathcal{N}_{h,\phi}[T_{\#}\mu] = \mathcal{N}_{h,\phi}[\mu]$

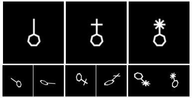
ii) **Singleton sets** in ${}_{\mathrm{h}}$NR-CDT-space

$$\mathbb{F}_\mu = \{T_{\#}\mu\} \xrightarrow{{}_{\mathrm{h}}\text{NR-CDT}} \mathcal{N}_{h,\phi}[\mathbb{F}_\mu] = \mathcal{N}_{h,\phi}[\mu]$$

iii) **Robustness** of ${}_{\mathrm{m}}$NR-CDT and ${}_{\mathrm{a}}$NR-CDT in Wasserstein-spaces

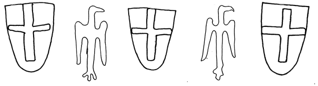
$$h \in \Big\{\sup_{\boldsymbol{\theta}\in\Theta}, \int_\Theta\Big\}.$$



## Motivation: Objects as Prob. Measures

- Classification with **many classes** and **few samples**
- Assumption: (approx.) **affine** / structured class families
- Need: a) **invariance** to affine transformations
  b) **stability** to deformations
- Application: historic watermarks, automated pattern recognition



## Normalization: NR-CDT and Sorting

- For $\mu \in \mathcal{P}^{+}_{\phi,\mathrm{c}}(\mathbb{X}) \subset \mathcal{P}_{\mathrm{c}}(\mathbb{X}) \subset \mathcal{P}_2(\mathbb{X})$ R-CDT is bounded with
  $$\operatorname{std}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]}) \ge c > 0 \quad \text{f.a.} \quad \theta \in \Theta.$$
- **Normalize** per slice:
  $$\mathcal{N}_{\phi,\theta}[\mu](t) = \frac{\widehat{\mathcal{R}_{\phi,\theta}[\mu]}(t) - \operatorname{mean}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]})}{\operatorname{std}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]})}$$
- **Sorting** the directions $\theta \in \Theta$ via:
  $$h : \mathrm{L}^\infty(\Theta) \to \mathbb{R}$$
- **Define** ${}_{\mathrm{h}}$NR-CDT: $h\big(\mathcal{N}_\phi[\mu](t,\cdot)\big)$

## Radon on Measures & Core Idea

- Restricted Radon transfrom via measurable $\phi : \mathbb{X} \times \Theta \to \mathbb{R}$

$$\mathcal{R}_{\phi,\theta} : \mathcal{M}(\mathbb{X}) \to \mathcal{M}(\mathbb{R})$$
$$\mu \mapsto (S_{\phi,\theta})_{\#}\mu, \; S_{\phi,\theta} \coloneqq \phi(\cdot, \boldsymbol{\theta})$$

- Slice per direction
  $$\mu \in \mathcal{P}(\mathbb{X}) \quad \Rightarrow \quad \mathcal{R}_{\phi,\theta}[\mu] \in \mathcal{P}(\mathbb{R})$$
- 1D embedding via CDT, wrt. ref $\rho$

$$\widehat{\mu} = \operatorname*{arg\,min}_{T_{\#}\rho = \mu} \int_{\mathbb{R}} c(s - T(s)) \, \mathrm{d}\rho(s), \quad \text{atomless } \rho, \mu \in \mathcal{P}(\mathbb{R})$$

- R-CDT: collect sliced CDTs
  $$\widehat{\mathcal{R}_{\phi,\theta}[\mu]} \in \mathrm{L}^2_\rho(\mathbb{R}) \text{ for } \mu \in \mathcal{P}_2(\mathbb{X})$$

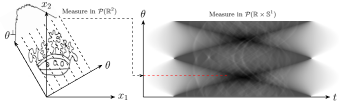
## Theoretical Results

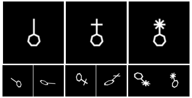
- **Translation:** for suitable $(\mathbb{X}, \Theta, \phi)$
  $$\widehat{\mathcal{R}_\phi[T_{\#}\mu]} = \hat{T} \circ \widehat{\mathcal{R}_\phi[\mu]}$$
  for specific $T$ correspond to $\hat{T}$

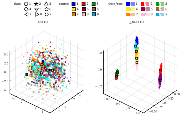
i) **Invariance:** $\mathcal{N}_{h,\phi}[T_{\#}\mu] = \mathcal{N}_{h,\phi}[\mu]$

ii) **Singleton sets** in ${}_{\mathrm{h}}$NR-CDT-space
$$\mathbb{F}_\mu = \{T_{\#}\mu\} \overset{{}_{\mathrm{h}}\text{NR-CDT}}{\longrightarrow} \mathcal{N}_{h,\phi}[\mathbb{F}_\mu] = \mathcal{N}_{h,\phi}[\mu]$$

iii) **Robustness** of ${}_{\mathrm{m}}$NR-CDT and ${}_{\mathrm{a}}$NR-CDT in Wasserstein-spaces
$$h \in \big\{\sup_{\boldsymbol{\theta} \in \Theta}, \textstyle\int_\Theta\big\}.$$

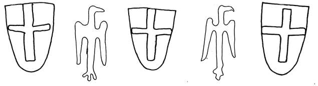




## Motivation: Objects as Prob. Measures

- Classification with **many classes** and **few samples**
- Assumption: (approx.) **affine** / structured class families
- Need: a) **invariance** to affine transformations
  b) **stability** to deformations
- Application: historic watermarks, automated pattern recognition



## Normalization: NR-CDT and Sorting

- For $\mu \in \mathcal{P}^{+}_{\phi,\mathrm{c}}(\mathbb{X}) \subset \mathcal{P}_{\mathrm{c}}(\mathbb{X}) \subset \mathcal{P}_2(\mathbb{X})$ R-CDT is bounded with
  $$\mathrm{std}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]}) \geq c > 0 \quad \text{f.a.} \quad \theta \in \Theta.$$
- **Normalize** per slice:
  $$\mathcal{N}_{\phi,\theta}[\mu](t) = \frac{\widehat{\mathcal{R}_{\phi,\theta}[\mu]}(t) - \mathrm{mean}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]})}{\mathrm{std}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]})}$$
- **Sorting** the directions $\theta \in \Theta$ via:
  $$h : \mathrm{L}^\infty(\Theta) \to \mathbb{R}$$
- **Define** ${}_{\mathrm{h}}$NR-CDT: $h\big(\mathcal{N}_\phi[\mu](t,\cdot)\big)$

## Radon on Measures & Core Idea

- Restricted Radon transfrom via measurable $\phi : \mathbb{X} \times \Theta \to \mathbb{R}$

$$\mathcal{R}_{\phi,\theta} : \mathcal{M}(\mathbb{X}) \to \mathcal{M}(\mathbb{R})$$
$$\mu \mapsto (S_{\phi,\theta})_{\#}\mu, \; S_{\phi,\theta} \coloneqq \phi(\cdot, \boldsymbol{\theta})$$

- Slice per direction
  $$\mu \in \mathcal{P}(\mathbb{X}) \quad \Rightarrow \quad \mathcal{R}_{\phi,\theta}[\mu] \in \mathcal{P}(\mathbb{R})$$
- 1D embedding via CDT, wrt. ref $\rho$

$$\widehat{\mu} = \underset{T_{\#}\rho=\mu}{\arg\min} \int_{\mathbb{R}} c(s - T(s))\,\mathrm{d}\rho(s), \quad \text{atomless } \rho, \mu \in \mathcal{P}(\mathbb{R})$$

- R-CDT: collect sliced CDTs
  $$\widehat{\mathcal{R}_{\phi,\theta}[\mu]} \in \mathrm{L}^2_\rho(\mathbb{R}) \text{ for } \mu \in \mathcal{P}_2(\mathbb{X})$$

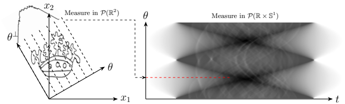
## Theoretical Results

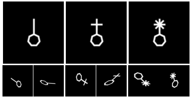
- **Translation:** for suitable $(\mathbb{X}, \Theta, \phi)$
  $$\widehat{\mathcal{R}_\phi[T_{\#}\mu]} = \hat{T} \circ \widehat{\mathcal{R}_\phi[\mu]}$$
  for specific $T$ correspond to $\hat{T}$

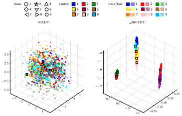
i) **Invariance:** $\mathcal{N}_{h,\phi}[T_{\#}\mu] = \mathcal{N}_{h,\phi}[\mu]$

ii) **Singleton sets** in ${}_{\mathrm{h}}$NR-CDT-space
$$\mathbb{F}_\mu = \{T_{\#}\mu\} \xrightarrow{{}_{\mathrm{h}}\text{NR-CDT}} \mathcal{N}_{h,\phi}[\mathbb{F}_\mu] = \mathcal{N}_{h,\phi}[\mu]$$

iii) **Robustness** of ${}_{\mathrm{m}}$NR-CDT and ${}_{\mathrm{a}}$NR-CDT in Wasserstein-spaces
$$h \in \left\{\sup_{\boldsymbol{\theta}\in\Theta}, \int_\Theta\right\}.$$

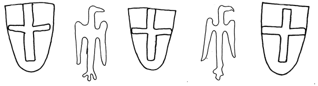




## Motivation: Objects as Prob. Measures

- Classification with **many classes** and **few samples**
- Assumption: (approx.) **affine** / structured class families
- Need: a) **invariance** to affine transformations
  b) **stability** to deformations
- Application: historic watermarks, automated pattern recognition



## Normalization: NR-CDT and Sorting

- For $\mu \in \mathcal{P}^+_{\phi,\mathrm{c}}(\mathbb{X}) \subset \mathcal{P}_\mathrm{c}(\mathbb{X}) \subset \mathcal{P}_2(\mathbb{X})$ R-CDT is bounded with
$$\operatorname{std}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]}) \geq c > 0 \quad \text{f.a.} \quad \theta \in \Theta.$$
- **Normalize** per slice:
$$\mathcal{N}_{\phi,\theta}[\mu](t) = \frac{\widehat{\mathcal{R}_{\phi,\theta}[\mu]}(t) - \operatorname{mean}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]})}{\operatorname{std}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]})}$$
- **Sorting** the directions $\theta \in \Theta$ via:
$$h : \mathrm{L}^\infty(\Theta) \to \mathbb{R}$$
- **Define** ${}_\mathrm{h}$NR-CDT: $h\big(\mathcal{N}_\phi[\mu](t,\cdot)\big)$

## Radon on Measures & Core Idea

- Restricted Radon transfrom via measurable $\phi : \mathbb{X} \times \Theta \to \mathbb{R}$
$$\mathcal{R}_{\phi,\theta} : \mathcal{M}(\mathbb{X}) \to \mathcal{M}(\mathbb{R})$$
$$\mu \mapsto (S_{\phi,\theta})_\# \mu, \; S_{\phi,\theta} \coloneqq \phi(\cdot, \boldsymbol{\theta})$$
- Slice per direction
$$\mu \in \mathcal{P}(\mathbb{X}) \quad \Rightarrow \quad \mathcal{R}_{\phi,\theta}[\mu] \in \mathcal{P}(\mathbb{R})$$
- 1D embedding via CDT, wrt. ref $\rho$
$$\widehat{\mu} = \operatorname*{arg\,min}_{T_\# \rho = \mu} \int_\mathbb{R} c(s - T(s)) \, \mathrm{d}\rho(s), \quad \text{atomless } \rho, \mu \in \mathcal{P}(\mathbb{R})$$
- R-CDT: collect sliced CDTs
$$\widehat{\mathcal{R}_{\phi,\theta}[\mu]} \in \mathrm{L}^2_\rho(\mathbb{R}) \text{ for } \mu \in \mathcal{P}_2(\mathbb{X})$$

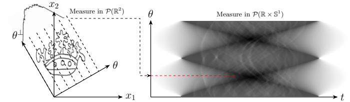
## Theoretical Results

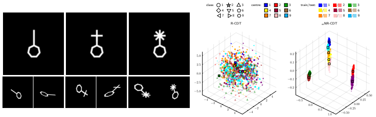
- **Translation:** for suitable $(\mathbb{X}, \Theta, \phi)$
$$\widehat{\mathcal{R}_\phi[T_\# \mu]} = \hat{T} \circ \widehat{\mathcal{R}_\phi[\mu]}$$
for specific $T$ correspond to $\hat{T}$

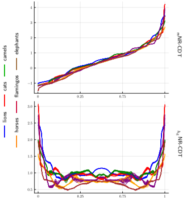
i) **Invariance:** $\mathcal{N}_{h,\phi}[T_\# \mu] = \mathcal{N}_{h,\phi}[\mu]$

ii) **Singleton sets** in ${}_\mathrm{h}$NR-CDT-space
$$\mathbb{F}_\mu = \{T_\# \mu\} \xrightarrow{{}_\mathrm{h}\text{NR-CDT}} \mathcal{N}_{h,\phi}[\mathbb{F}_\mu] = \mathcal{N}_{h,\phi}[\mu]$$

iii) **Robustness** of ${}_\mathrm{m}$NR-CDT and ${}_\mathrm{a}$NR-CDT in Wasserstein-spaces
$$h \in \big\{\sup_{\boldsymbol{\theta} \in \Theta}, \textstyle\int_\Theta\big\}.$$

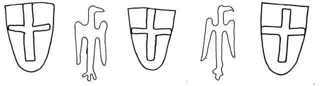


## Motivation: Objects as Prob. Measures

- Classification with **many classes** and **few samples**
- Assumption: (approx.) **affine** / structured class families
- Need: a) **invariance** to affine transformations
  b) **stability** to deformations
- Application: historic watermarks, automated pattern recognition



## Normalization: NR-CDT and Sorting

- For $\mu \in \mathcal{P}^{+}_{\phi,\mathrm{c}}(\mathbb{X}) \subset \mathcal{P}_{\mathrm{c}}(\mathbb{X}) \subset \mathcal{P}_2(\mathbb{X})$ R-CDT is bounded with
  $$\mathrm{std}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]}) \geq c > 0 \quad \text{f.a.} \quad \theta \in \Theta.$$
- **Normalize** per slice:
  $$\mathcal{N}_{\phi,\theta}[\mu](t) = \frac{\widehat{\mathcal{R}_{\phi,\theta}[\mu]}(t) - \mathrm{mean}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]})}{\mathrm{std}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]})}$$
- **Sorting** the directions $\theta \in \Theta$ via:
  $$h : \mathrm{L}^\infty(\Theta) \to \mathbb{R}$$
- **Define** ${}_{\mathrm{h}}$NR-CDT: $h\Big(\mathcal{N}_\phi[\mu](t,\cdot)\Big)$

## Radon on Measures & Core Idea

- Restricted Radon transfrom via measurable $\phi : \mathbb{X} \times \Theta \to \mathbb{R}$

$$\mathcal{R}_{\phi,\theta} : \mathcal{M}(\mathbb{X}) \to \mathcal{M}(\mathbb{R})$$
$$\mu \mapsto (S_{\phi,\theta})_{\#}\mu, \;\; S_{\phi,\theta} \coloneqq \phi(\cdot, \boldsymbol{\theta})$$

- Slice per direction
  $$\mu \in \mathcal{P}(\mathbb{X}) \quad \Rightarrow \quad \mathcal{R}_{\phi,\theta}[\mu] \in \mathcal{P}(\mathbb{R})$$
- 1D embedding via CDT, wrt. ref $\rho$

$$\widehat{\mu} = \operatorname*{arg\,min}_{T_{\#}\rho=\mu} \int_{\mathbb{R}} c(s - T(s))\,\mathrm{d}\rho(s), \quad \text{atomless } \rho, \mu \in \mathcal{P}(\mathbb{R})$$

- R-CDT: collect sliced CDTs
  $$\widehat{\mathcal{R}_{\phi,\theta}[\mu]} \in \mathrm{L}^2_\rho(\mathbb{R}) \text{ for } \mu \in \mathcal{P}_2(\mathbb{X})$$

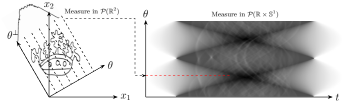
## Theoretical Results

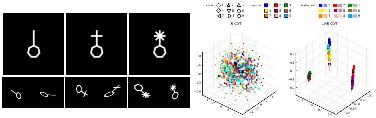
- **Translation:** for suitable $(\mathbb{X}, \Theta, \phi)$
  $$\widehat{\mathcal{R}_{\phi}[T_{\#}\mu]} = \hat{T} \circ \widehat{\mathcal{R}_{\phi}[\mu]}$$
  for specific $T$ correspond to $\hat{T}$

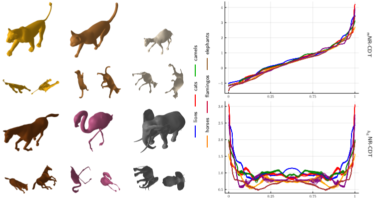
i) **Invariance:** $\mathcal{N}_{h,\phi}[T_{\#}\mu] = \mathcal{N}_{h,\phi}[\mu]$

ii) **Singleton sets** in ${}_{\mathrm{h}}$NR-CDT-space
$$\mathbb{F}_\mu = \{T_{\#}\mu\} \xrightarrow{{}_{\mathrm{h}}\text{NR-CDT}} \mathcal{N}_{h,\phi}[\mathbb{F}_\mu] = \mathcal{N}_{h,\phi}[\mu]$$

iii) **Robustness** of ${}_{\mathrm{m}}$NR-CDT and ${}_{\mathrm{a}}$NR-CDT in Wasserstein-spaces
$$h \in \left\{\sup_{\boldsymbol{\theta}\in\Theta}, \int_{\Theta}\right\}.$$



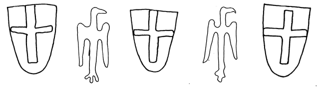




| # classes | method | NT # angles 8 | 16 | 32 | 64 | 128 | 1-NN # training samples 5 | 10 |
|---|---|---|---|---|---|---|---|---|
| 100 | Euclidean | 0.0096 | | | | | $0.0202 \pm 0.0021$ | $0.0242 \pm 0.0022$ |
| | R-CDT | 0.0192 | 0.0188 | 0.0190 | 0.0190 | 0.0192 | $0.0178 \pm 0.0022$ | $0.0210 \pm 0.0024$ |
| | ${}_{\mathrm{m}}$NR-CDT | **0.4952** | **0.9032** | **0.9922** | **0.9992** | **1.0000** | $\mathbf{1.0000 \pm 0.0000}$ | $\mathbf{1.0000 \pm 0.0000}$ |
| 1000 | ${}_{\mathrm{m}}$NR-CDT | **0.3433** | **0.8256** | **0.9842** | **0.9977** | **1.0000** | $\mathbf{0.9987 \pm 0.0001}$ | $\mathbf{0.9997 \pm 0.0001}$ |

## Motivation: Objects as Prob. Measures

- Classification with **many classes** and **few samples**
- Assumption: (approx.) **affine** / structured class families
- Need: a) **invariance** to affine transformations
  b) **stability** to deformations
- Application: historic watermarks, automated pattern recognition



## Normalization: NR-CDT and Sorting

- For $\mu \in \mathcal{P}^+_{\phi,c}(\mathbb{X}) \subset \mathcal{P}_c(\mathbb{X}) \subset \mathcal{P}_2(\mathbb{X})$ R-CDT is bounded with
$$\operatorname{std}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]}) \geq c > 0 \quad \text{f.a.} \quad \theta \in \Theta.$$
- **Normalize** per slice:
$$\mathcal{N}_{\phi,\theta}[\mu](t) = \frac{\widehat{\mathcal{R}_{\phi,\theta}[\mu]}(t) - \operatorname{mean}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]})}{\operatorname{std}(\widehat{\mathcal{R}_{\phi,\theta}[\mu]})}$$
- **Sorting** the directions $\theta \in \Theta$ via:
$$h : \mathrm{L}^\infty(\Theta) \to \mathbb{R}$$
- **Define** $_{\mathrm{h}}$NR-CDT: $h\big(\mathcal{N}_\phi[\mu](t, \cdot)\big)$

## Radon on Measures & Core Idea

- Restricted Radon transfrom via measurable $\phi : \mathbb{X} \times \Theta \to \mathbb{R}$

$$\mathcal{R}_{\phi,\theta} : \mathcal{M}(\mathbb{X}) \to \mathcal{M}(\mathbb{R})$$
$$\mu \mapsto (S_{\phi,\theta})_\# \mu, \; S_{\phi,\theta} \coloneqq \phi(\cdot, \boldsymbol{\theta})$$

- Slice per direction
$$\mu \in \mathcal{P}(\mathbb{X}) \quad \Rightarrow \quad \mathcal{R}_{\phi,\theta}[\mu] \in \mathcal{P}(\mathbb{R})$$
- 1D embedding via CDT, wrt. ref $\rho$

$$\widehat{\mu} = \operatorname*{arg\,min}_{T_\# \rho = \mu} \int_{\mathbb{R}} c(s - T(s)) \, \mathrm{d}\rho(s), \quad \text{atomless } \rho, \mu \in \mathcal{P}(\mathbb{R})$$

- R-CDT: collect sliced CDTs
$$\widehat{\mathcal{R}_{\phi,\theta}[\mu]} \in \mathrm{L}^2_\rho(\mathbb{R}) \text{ for } \mu \in \mathcal{P}_2(\mathbb{X})$$

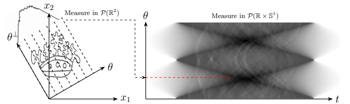
## Theoretical Results

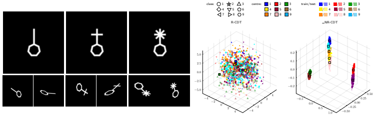
- **Translation:** for suitable $(\mathbb{X}, \Theta, \phi)$
$$\widehat{\mathcal{R}_\phi[T_\# \mu]} = \hat{T} \circ \widehat{\mathcal{R}_\phi[\mu]}$$
for specific $T$ correspond to $\hat{T}$

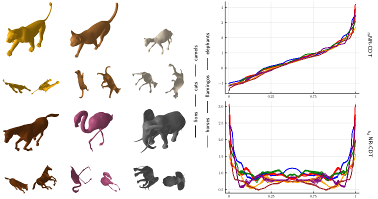
i) **Invariance:** $\mathcal{N}_{h,\phi}[T_\# \mu] = \mathcal{N}_{h,\phi}[\mu]$

ii) **Singleton sets** in $_{\mathrm{h}}$NR-CDT-space
$$\mathbb{F}_\mu = \{T_\# \mu\} \overset{_{\mathrm{h}}\text{NR-CDT}}{\longrightarrow} \mathcal{N}_{h,\phi}[\mathbb{F}_\mu] = \mathcal{N}_{h,\phi}[\mu]$$

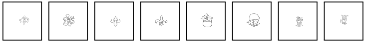
iii) **Robustness** of $_{\mathrm{m}}$NR-CDT and $_{\mathrm{a}}$NR-CDT in Wasserstein-spaces
$$h \in \left\{\sup_{\boldsymbol{\theta} \in \Theta}, \int_\Theta\right\}.$$




| # classes | method | NT # angles 8 | 16 | 32 | 64 | 128 | 1-NN # training samples 5 | 10 |
|---|---|---|---|---|---|---|---|---|
| 100 | Euclidean | 0.0096 | | | | | $0.0202 \pm 0.0021$ | $0.0242 \pm 0.0022$ |
| | R-CDT | 0.0192 | 0.0188 | 0.0190 | 0.0190 | 0.0192 | $0.0178 \pm 0.0022$ | $0.0210 \pm 0.0024$ |
| | $_{\mathrm{m}}$NR-CDT | **0.4952** | **0.9032** | **0.9922** | **0.9992** | **1.0000** | $\mathbf{1.0000 \pm 0.0000}$ | $\mathbf{1.0000 \pm 0.0000}$ |
| 1000 | $_{\mathrm{m}}$NR-CDT | **0.3433** | **0.8256** | **0.9842** | **0.9977** | **1.0000** | $\mathbf{0.9987 \pm 0.0001}$ | $\mathbf{0.9997 \pm 0.0001}$ |



| # angles | 8 | | 16 | | 32 | | 64 | | 128 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\Vert\cdot\Vert_\infty$ | $\Vert\cdot\Vert_2$ | $\Vert\cdot\Vert_\infty$ | $\Vert\cdot\Vert_2$ | $\Vert\cdot\Vert_\infty$ | $\Vert\cdot\Vert_2$ | $\Vert\cdot\Vert_\infty$ | $\Vert\cdot\Vert_2$ | $\Vert\cdot\Vert_\infty$ | $\Vert\cdot\Vert_2$ |
| R-CDT | 0.4285 | 0.1785 | 0.3928 | 0.1785 | 0.4642 | 0.1785 | 0.4285 | 0.1785 | 0.4285 | 0.1785 |
| $_{\mathrm{m}}$NR-CDT | 0.8571 | 0.9642 | 0.9285 | **1.0000** | 0.9285 | **1.0000** | 0.9295 | 0.9642 | 0.9285 | 0.9642 |
| $_{\mathrm{a}}$NR-CDT | 0.9642 | 0.8571 | **1.0000** | 0.9642 | **1.0000** | 0.9642 | **1.0000** | 0.9642 | **1.0000** | 0.9642 |
| Eucl. | 0.2857 | 0.1785 | | | | | | | | |