

THE RADON SIGNED CUMULATIVE DISTRIBUTION TRANSFORM AND ITS APPLICATIONS IN CLASSIFICATION OF SIGNED IMAGES

LE GONG, SHIYING LI, NAQIB SAD PATHAN, MOHAMMAD SHIFAT-E-RABBI,
 GUSTAVO K. ROHDE[†], ABU HASNAT MOHAMMAD RUBAIYAT,
 AND SUMATI THAREJA

Date of Receiving : 28. 07. 2023
 Date of Revision : 21. 12. 2023
 Date of Acceptance : 21. 12. 2023

Abstract. Here we describe a new image representation technique based on the mathematics of transport and optimal transport. The method relies on the combination of the well-known Radon transform for images and a recent signal representation method called the Signed Cumulative Distribution Transform. The newly proposed method generalizes previous transport-related image representation methods to arbitrary functions (images), and thus can be used in more applications. We describe the new transform, and some of its mathematical properties and demonstrate its ability to partition image classes with real and simulated data. In comparison to existing transport transform methods, as well as deep learning-based classification methods, the new transform more accurately represents the information content of signed images, and thus can be used to obtain higher classification accuracies. The implementation of the proposed method in Python language is integrated as a part of the software package PyTransKit [12]

1. Introduction

Finding useful mathematical formulas for representing signal and image data can be critical for solving important engineering and scientific problems. Fourier representation methods, for example, can dramatically simplify the solution of shift invariant linear systems of equations (e.g. convolutions), and thus are extensively used to filter sound and other types of signals, in optics and image processing (e.g. deconvolution) and other important problems. Localized sparse representations (e.g. short time Fourier transforms, Wavelets) have been extensively utilized in signal compression, denoising,

2010 *Mathematics Subject Classification.* 65R10, 68U10, 68T45.

Key words and phrases. Radon Signed Cumulative Distribution Transform, Image Representation, Image Classification.

Communicated by: Nikhil Khanna

[†] *Corresponding author*

filtering and other applications given their ability to summarize important signal features using few parameters.

In the past few years, new signal and image representation methods (i.e. transforms) based on the mathematics of optimal mass transport have emerged. Unlike the aforementioned methods (Fourier and Wavelet representations), which represent functions as linear combinations of basis functions or frame vectors, the emerging transport-based techniques are nonlinear representation methods. They have been extensively used to render problems that are nonlinear, non convex, and difficult to solve in signal domain into linear, convex problems that have closed form solutions in transport transform domain. Examples include reconstructing images from few measurements [14, 6], estimating signal parameters such as time delay and frequency [25], as well as classifying data where they have been shown to be especially effective in allowing for the formulation of closed form, efficient, and accurate solution of a wide variety of signal and image classification problems [26, 27, 28, 34].

To date, several types of transforms based on optimal transport have appeared in the literature, nearly all can be interpreted as a variation of the linear optimal transport (LOT) technique [17, 33], which is naturally applicable to data that can be interpreted as probability distributions. These include 1D signal transforms for non-negative distributional data [21] and signed distributions [2], as well as transforms for representing non-negative and normalized images in two and higher dimensions [15]. However, transport-based representation techniques for signed images in two or higher dimensions are yet to be described, and thus the success of transport-based classification methods in applications that involve signed images [3, 9, 13] has yet to be replicated.

Signed signals in two or higher dimensions are frequently encountered in science and engineering applications. One example being magnetic resonance imaging (MRI) where voltage coils acquire a series of 1D signals which are then reconstructed into images using the Fourier transform method, yielding complicated functions that have positive and negative values [5]. Other examples include classification problems that employ scalograms derived from various types of Wavelet transforms [13]. Background subtraction is also a commonly performed pre-processing operation for microscopy and other types of optical imaging, typically yielding signed data [9]. Filtering techniques are also often employed to pre-process data prior to processing and classification. When signed filters are used (e.g. edge-type detection filters, convolutional neural networks), the resulting pre-processed images become signed [3].

Our goal in this manuscript is to describe a transport-based method for representing 2D signed images and demonstrate its ability to solve classification problems involving this type of data. The new image representation is denoted the Radon Signed Cumulative Distribution Transform (RSCDT) and consists of combining the recently developed signed cumulative distribution transform (SCDT) [2] together with the well-known Radon transform [24]. We describe ensuing mathematical properties of the representation, as well as a new Wasserstein-type metric for signed images which can be useful in the solution of pattern recognition problems. We then adapt the transport-based classification problem statement appearing in recent publications involving unsigned images [27] to signed images and describe its solution. Computational results showing the performance of the classification method based on the newly developed RSCDT in comparison to existing methods, including deep learning, are then shown,

TABLE 1. Description of symbols

Symbols	Description
$s(x) / s(\mathbf{x})$	Signal / image
Ω_s	Domain of s
$\tilde{s}(t, \theta)$	Radon transform of s
$\hat{s}(x) / \hat{s}(t, \theta)$	CDT / R-CDT transform of s
$\mathcal{R}(\cdot) / \mathcal{R}^{-1}(\cdot)$	Forward / inverse Radon transform operation
$g(x)$	Strictly increasing and differentiable function
$g^\theta(t)$	Strictly increasing and differentiable function, indexed by an angle θ
$s \circ g$	$s(g(x))$: composition of $s(x)$ with $g(x)$
$\tilde{s} \circ g^\theta$	$\tilde{s}(g^\theta(t), \theta)$: composition of $\tilde{s}(t, \theta)$ with $g^\theta(t)$ along the t dimension of $\tilde{s}(t, \theta)$
$\mathcal{G}$	Set of increasing diffeomorphisms $g : \mathbb{R} \rightarrow \mathbb{R}$
$\mathcal{G}_\mathcal{R}$	Set of increasing diffeomorphisms $g^\theta : \mathbb{R} \rightarrow \mathbb{R}$ parameterized by θ with $\theta \in [0, \pi]$
$\mathcal{T}$	Set of all possible increasing diffeomorphisms from $\mathbb{R}$ to $\mathbb{R}$

followed by concluding remarks. For readability, proofs for the given theorems and statements are provided in the appendix.

2. Preliminaries

2.1. Notation. Throughout the manuscript, we deal with signals s assuming these to be square integrable in their respective domains. That is, we assume that $\int_{\Omega_s} |s(x)|^2 dx < \infty$, where $\Omega_s \subseteq \mathbb{R}$ is the domain over which s is defined. In addition, we at times make use of the common notation: $\|s\|^2 = \langle s, s \rangle = \int_{\Omega_s} s(x)^* s(x) dx = \int_{\Omega_s} |s(x)|^2 dx$, where $\langle \cdot, \cdot \rangle$ is the inner product. Signals are assumed to be real, so the complex conjugate $*$ does not play a role. We will apply the same notation for functions whose input argument is two dimensional, i.e. images. Let $\mathbf{x} \in \Omega_s \subseteq \mathbb{R}^2$. A 2D continuous function representing the continuous image is denoted $s(\mathbf{x})$, $\mathbf{x} \in \Omega_s$. Signals or images are denoted $s^{(k)}$ when the class information is available, where the superscript (k) represents the class label.

Below we will also make use of one dimensional (1D) increasing diffeomorphisms (invertible and differentiable functions), which are denoted as $g(x)$ for signals and $g^\theta(t)$ when they need to be parameterized by an angle θ . The set of all possible increasing diffeomorphisms from $\mathbb{R}$ to $\mathbb{R}$ will be denoted as $\mathcal{T}$. Finally, at times we also utilize the ‘ $\circ$ ’ operator to denote composition. A summary of the symbols and notation used can be found in Table 1.

2.2. The Cumulative Distribution Transform (CDT). The CDT [21] is an invertible nonlinear 1D signal transform from the space of smooth probability densities to the space of diffeomorphisms. The CDT morphs a given input signal, defined as a probability density function (PDF), into another PDF in such a way that the Wasserstein distance between them is minimized. More formally, let $s(x)$, $x \in \Omega_s$ and $r(y)$, $y \in \Omega_r$,

which we denote as a chosen ‘reference’ signal, define a given signal and a reference signal, respectively, which we consider to be appropriately normalized such that $s > 0, r > 0$, and $\int_{\Omega_s} s(x)dx = \int_{\Omega_r} r(y)dy = 1$. The forward CDT transform¹ of $s(x)$ with respect to $r(y)$ is given by the strictly increasing function $\hat{s}(y)$ that satisfies

$$\int_{-\infty}^{\hat{s}(y)} s(u)du = \int_{-\infty}^y r(v)dv \quad (2.1)$$

As described in detail in [21], the CDT is a nonlinear and invertible operation, with the inverse being

$$s(x) = \frac{d\hat{s}^{-1}(x)}{dx} r(\hat{s}^{-1}(x)), \text{ and } \hat{s}^{-1}(\hat{s}(x)) = x.$$

Therefore, $\hat{s}$ can be seen as a representation of the original function s . Moreover, like the Fourier transform [4] for example, the CDT has a number of properties which will help us render signal and image classification problems easier to solve. Finally, taking advantage that many of the properties of the CDT are independent of the arbitrarily chosen reference, throughout the manuscript we use the uniform distribution for the reference function r , i.e.

$$r(x) = \begin{cases} 1, & x \in \Omega_r \\ 0, & \text{otherwise.} \end{cases} \quad (2.2)$$

Property II-B.1 (Composition): Let $s(x)$ denote a normalized signal and let $\hat{s}(x)$ be the CDT of $s(x)$. The CDT of $s_g = g'(s \circ g)$ is given by

$$\hat{s}_g = g^{-1} \circ \hat{s} \quad (2.3)$$

Here, $g \in \mathcal{T}$ is an invertible and differentiable function (diffeomorphism), $g' = dg(x)/dx$, and ‘ $\circ$ ’ denotes the composition operator with $(s \circ g)(x) = s(g(x))$. For a proof, see Appendix A in supplementary materials.

The CDT composition property implies that variations in a signal $s(x)$ caused by applying $g(x)$ to the independent variable (i.e. $x \rightarrow g(x)$) in the signal space so that $s(x) \rightarrow s(g(x))$, will cause an amplitude modulation in i.e. $\hat{s}(x) \rightarrow g^{-1} \circ \hat{s}(x)$ in the CDT space (see Figure 1).

We note that in the case when the reference function r is the uniform reference stated in (2.2) the integral defining the CDT (2.1) becomes simply $\int_{-\infty}^{\hat{s}(y)} s(u)du = y$, $y \in [0, 1]$. Using the cumulative distribution notation $S(y) = \int_{-\infty}^y s(u)du$, the CDT of signal s is simply: $\hat{s}(y) = S^{-1}(y)$, where S^{-1} refers to the inverse function of S (which in this case is monotonic). Noting the definition of the Wasserstein metric between two 1D distributions s_1 and s_2 is related to their inverse cumulative distributions ($W_2^2 = \|S_1^{-1} - S_2^{-1}\|^2$) [32], it is thus not hard to see that the usual Wasserstein distance between 1D distributions is related to the CDT, as can be seen in the following property,

Property II-B.2 (Embedding): The CDT induces an isometric embedding between the space of 1D signals with the 2-Wasserstein metric and the space of their CDT

¹ We are using a slightly different definition of the CDT than in [21]. The properties of the CDT outlined here hold in both definitions.

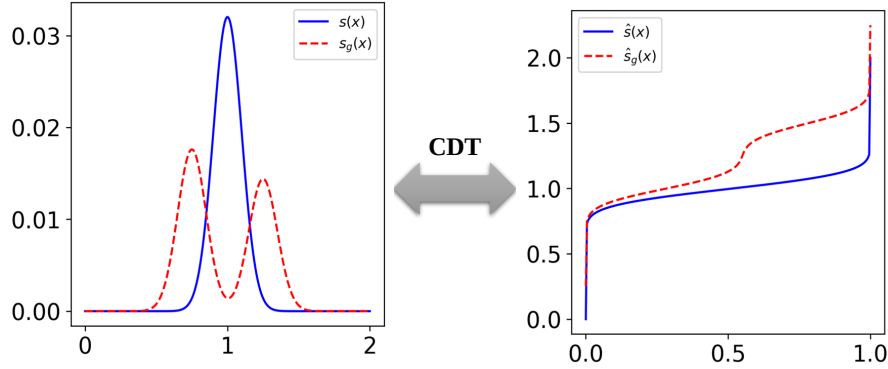


FIGURE 1. The cumulative distribution transform (CDT) of a signal (probability density function). Note that the CDT of an altered (transported) signal $s_g(x)$ (see text for definition) is related to the transform of s . In short, the CDT renders displacements into amplitude modulations in transform space.

transforms with a weighted-Euclidean metric [21], i.e.,

$$W_2^2(s_1, s_2) = \|(\hat{s}_1 - \hat{s}_2) \sqrt{r}\|_{L^2(\Omega_r)}^2, \quad (2.4)$$

for all signals s_1, s_2 . That is to say, if we wish to use the Wasserstein distance as a measure of similarity between s_1, s_2 , we can compute it as simply a weighted Euclidean norm in CDT space.

The property above naturally links the CDT and Wasserstein distances for PDFs. Wasserstein [32] distances are linked to optimal transport and have been used in a variety of applications in signal and image processing and machine learning (see [16] for a recent review).

2.3. The Radon transform. The d -dimensional Radon transform $\mathcal{R}$ maps a function $s \in L^1(\mathbb{R}^d)$ where $L^1(\mathbb{R}^d) := \{s : \mathbb{R}^d \rightarrow \mathbb{R} \mid \int_{\mathbb{R}^d} |s(\mathbf{x})| d\mathbf{x} \leq \infty\}$ into the set of its integrals over the hyperplanes of $\mathbb{R}^n$ and is defined as,

$$\tilde{s}(t, \theta) = \mathcal{R}s(t, \theta) := \int_{\mathbf{x} \cdot \theta = t} I(\mathbf{x}) d\sigma(\mathbf{x}), \quad \forall t \in \mathbb{R}, \quad \forall \theta \in \mathbb{S}^{d-1}, \quad (2.5)$$

where $d\sigma$ denotes the natural surface measure on the $(d-1)$ -dimensional hyperplane $\mathbf{x} \cdot \theta = t$ orthogonal to θ , and $\mathbb{S}^{d-1}$ is the unit sphere in $\mathbb{R}^d$. Note that $\mathcal{R} : L^1(\mathbb{R}^d) \rightarrow L^1(\mathbb{R} \times \mathbb{S}^{d-1})$. The Radon transform also possesses an inverse, which we illustrate in the two-dimensional case. Using the Fourier Slice Theorem [20, 23], the inverse Radon transform $s = \mathcal{R}^{-1}(\tilde{s})$ is defined as

$$s(\mathbf{x}) = \int_0^\pi \int_{-\infty}^\infty \tilde{s}(\mathbf{x} \cdot \xi_\theta - \tau, \theta) w(\tau) d\tau d\theta, \quad (2.6)$$

where, in $\mathbb{R}^2$, w is the ramp filter (i.e., $(\mathcal{F}w)(\xi) = |\xi|, \forall \xi$) and $\mathcal{F}$ is the one dimensional Fourier transform operator.

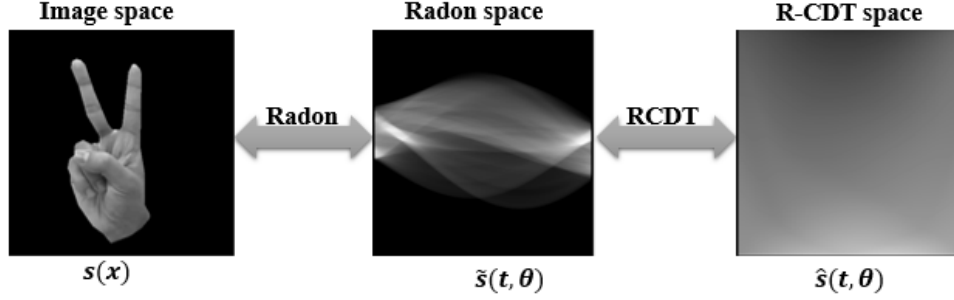


FIGURE 2. The process of calculating the Radon cumulative distribution transform (R-CDT) of an image $s(\mathbf{x})$ (defined as a 2-dimensional probability density function). The first step is to apply the Radon transform on $s(\mathbf{x})$ to obtain $\tilde{s}(t, \theta)$. The R-CDT $\hat{s}(t, \theta)$ is then obtained by applying the CDT over the t dimension of $\tilde{s}(t, \theta)$, $\forall \theta$.

Property II-C.1 (*Intensity equality*): Note that

$$\int_{\Omega_s} s(\mathbf{x}) d\mathbf{x} = \int_{-\infty}^{\infty} \tilde{s}(t, \theta) dt, \quad \forall \theta \in [0, \pi] \quad (2.7)$$

which implies that $\int_{-\infty}^{\infty} \tilde{s}(t, \theta_i) dt = \int_{-\infty}^{\infty} \tilde{s}(t, \theta_j) dt$ for any two choices $\theta_i, \theta_j \in [0, \pi]$.

2.4. Radon Cumulative Distribution Transform (R-CDT). The CDT framework was extended for 2D patterns (images as normalized density functions) through the sliced-Wasserstein distance in [15], and was denoted as R-CDT. The main idea behind the R-CDT is to first obtain a family of one dimensional representations of a two dimensional probability measure (e.g., an image) through the Radon transform and then apply the CDT over the t dimension in Radon transform space. More formally, let $s(\mathbf{x})$ and $r(\mathbf{x})$ define a given image and a reference image, respectively, which we consider to be appropriately normalized. The forward R-CDT of $s(\mathbf{x})$ with respect to $r(\mathbf{x})$ is given by the measure preserving function $\hat{s}(t, \theta)$ that satisfies

$$\int_{-\infty}^{\hat{s}(t, \theta)} \tilde{s}(u, \theta) du = \int_{-\infty}^t \tilde{r}(u, \theta) du, \quad \forall \theta \in [0, \pi] \quad (2.8)$$

As in the case of the CDT, a transformed signal in R-CDT space can be recovered via the following inverse formula [15],

$$s(\mathbf{x}) = \mathcal{R}^{-1} \left(\frac{\partial \hat{s}^{-1}(t, \theta)}{\partial t} \tilde{r}(\hat{s}^{-1}(t, \theta), \theta) \right) \quad (2.9)$$

The process of calculating the R-CDT transform is shown in Figure 2. As with the CDT, the R-CDT has a couple of properties outlined below which will be of interest when classifying images. Let us give some important definitions first. Let $\mathcal{G}_{\mathcal{R}}$ be the set defined as follows,

$$\mathcal{G}_{\mathcal{R}} = \{g = (g^\theta)_{\theta \in [0, \pi]} : g^\theta : \mathbb{R} \rightarrow \mathbb{R} \text{ is a strictly increasing bijection } \forall \theta \in [0, \pi]\} \quad (2.10)$$

Definition 2.1. Let $\star : \mathcal{G}_{\mathcal{R}} \times \mathcal{G}_{\mathcal{R}} \rightarrow \mathcal{G}_{\mathcal{R}}$ be an operator defined by

$$(g \star h)(\cdot, \theta) := (g^\theta \circ h^\theta)(\cdot), \quad \forall \theta \in [0, \pi].$$

It is not hard to see that $(\mathcal{G}_{\mathcal{R}}, \star)$ is a group (see A.1).

Definition 2.2. Let $H_{\mathcal{R}} \subseteq \mathcal{G}_{\mathcal{R}}$. We call $H_{\mathcal{R}}$ a convex subgroup of $(\mathcal{G}_{\mathcal{R}}, \star)$ if $H_{\mathcal{R}}$ is a convex set and $(H_{\mathcal{R}}, \star)$ is a group.

Property II-D.1 (Composition): Let $s(\mathbf{x})$ denote an appropriately normalized image and let $\tilde{s}(t, \theta)$ and $\hat{s}(t, \theta)$ be the Radon transform and the R-CDT transform of $s(\mathbf{x})$, respectively. The R-CDT transform of $s_{g^\theta} = \mathcal{R}^{-1} \left((g^\theta)' \tilde{s} \circ g^\theta \right)$ is given by

$$\hat{s}_{g^\theta} = (g^\theta)^{-1} \circ \hat{s}, \quad (2.11)$$

where $(g^\theta)' = dg^\theta(t)/dt$, $\tilde{s} \circ g^\theta := \tilde{s}(g^\theta(t), \theta)$, and $(g^\theta)^{-1} \circ \hat{s} = (g^\theta)^{-1}(\hat{s}(t, \theta))$. Here for each fixed θ , g^θ can be thought of an increasing and differentiable function with respect to t . The above equation hence follows from the composition property for 1D CDT. For a proof, see [15].

The R-CDT composition property implies that, variations along both independent and dependent axis directions in an image, caused by applying $g^\theta(t)$ to the independent t variable of its Radon transform, become changes solely along the dependent variable in R-CDT space.

Definition 2.3. Given R-CDT $\hat{s}(t, \theta)$ of some image s , computed w.r.t. reference image r , the R-CDT distance (also known as sliced Wasserstein distance [16]) between s and r is given by:

$$SW_2^2(s, r) = \int_0^\pi \int_{-\infty}^\infty (\hat{s}(t, \theta) - t)^2 \tilde{r}(t, \theta) dt d\theta. \quad (2.12)$$

Property II-D.2 (Embedding): R-CDT induces an isometric embedding between the space of images with sliced-Wasserstein metric and the space of their R-CDT transforms with a weighted-Euclidean metric, i.e.,

$$SW_2^2(s_1, s_2) = \left\| (\hat{s}_1 - \hat{s}_2) \sqrt{\tilde{r}} \right\|_{L^2(\Omega_{\tilde{r}})}^2 \quad (2.13)$$

for all images s_1 and s_2 . For a proof, see [15].

As the case with the 1D CDT shown above, the property above naturally links the R-CDT and sliced Wasserstein distances for PDFs and affords us a simple means of computing similarity among images [15]. We remark that throughout this manuscript we use the notation $\hat{s}$ for both CDT or R-CDT transforms of a signal or image s with respect to a fixed reference signal or image r , if a reference is not specified.

2.5. Signed Cumulative Distribution Transform. This section is a brief description of the SCDT (for details, see [2]). Fix a reference signal r that is a positive probability density. Let s be a non-negative signal with $\|s\|_1 = 1$. The cumulative distribution function F_s of s is defined as

$$F_s(x) := \int_{-\infty}^x s(t) dt. \quad (2.14)$$

The Cumulative Distribution Transform $\mathcal{C}(s)$ of s with respect to r , as defined in eq. (2.1) (which implies $F_s(\widehat{s}(t)) = F_r(t)$) [21], can also be written as (see [2])

$$\widehat{s} := F_s^\dagger \circ F_r, \quad (2.15)$$

where $F_s^\dagger$ is the generalized inverse of F_s defined by

$$F^\dagger(y) := \inf\{x : F(x) > y\}.$$

Note that the generalized inverse comes in play here, because the CDF F_s of s need not be invertible. Now, for a signed signal s , the transform is defined by utilizing the Jordan decomposition of s , namely

$$s^+(t) = \max\{0, s(t)\}, \quad s^-(t) = \max\{0, -s(t)\} \quad (2.16)$$

Then, the SCDT of s with respect to r is defined as

$$\widehat{s} = \mathbb{T}_r(s) := \left((s^+)^*, \|s^+\|_1, (s^-)^*, \|s^-\|_1 \right), \quad (2.17)$$

where

$$(s^\pm)^* := \begin{cases} \widehat{\frac{s^\pm}{\|s^\pm\|_1}} & \text{if } s^\pm \text{ is non-trivial} \\ 0 & \text{if } s^\pm = 0. \end{cases} \quad (2.18)$$

Let $\mathcal{N} := \{f : \mathbb{R} \rightarrow \mathbb{R} \text{ non-decreasing}\} \times \mathbb{R}^+$. Then, for a fixed reference r ,

$$\mathbb{T}_r : L^1(\mathbb{R}) \rightarrow \mathcal{N} \times \mathcal{N}$$

is a bijection, where for $(f, u, g, v) \in \mathcal{N} \times \mathcal{N}$ the inverse transform is given as,

$$(f, u, g, v) \longrightarrow \frac{u}{\|r\|_1} r(f^\dagger(\cdot))(f^\dagger)'(\cdot) - \frac{v}{\|r\|_1} r(g^\dagger(\cdot))(g^\dagger)'(\cdot) \quad (2.19)$$

provided $f^\dagger$ and $g^\dagger$ are differentiable.

Note that, the SCDT expands on the concept of the CDT by accommodating functions with both positive and negative values. By splitting a signed signal into its positive and negative components, the SCDT allows for the application of the CDT independently to each part. This enables the transformation of signed signals while preserving important properties and characteristics. The introduction of the SCDT also unveils the concept of the generalized inverse. The generalized inverse behaves similarly to the regular inverse for invertible functions, however, it extends to situations where functions may not be invertible. This becomes necessary in the case of the SCDT, given that the Jordan decomposition used above (see (2.16)) yields two non-negative signals s^+ and s^- which might contain zeroes, and hence their cumulation (an analogue of CDF, for functions that are not probability distributions (see [2])) stated in equation (2.14) might not be strictly increasing and thus, might not admit an inverse. Therefore, the generalized inverse, is valuable in situations where a direct inverse may not exist, still allowing for the recovery and analysis of signals even when invertibility is not guaranteed.

By incorporating the SCDT and the notion of the generalized inverse, the framework becomes more versatile and applicable to a broader range of signals. Let us now see some of the important properties of SCDT, that are a more generalized version of the properties of the CDT.

Lemma 2.4. (*Composition Property*) [2] Let $s \in L^1(\mathbb{R})$, and let $g : \mathbb{R} \rightarrow \mathbb{R}$ be a strictly increasing surjection. Consider s_g given by $s_g(t) = g'(t)(s(g(t)))$. Then, the SCDT of s_g satisfies

$$\widehat{s}_g = \left(g^{-1} \circ (s^+)^*, \|s^+\|_1, g^{-1} \circ (s^-)^*, \|s^-\|_1 \right).$$

In order to state the second lemma, we require a metric on the native space. This metric is a generalization of Wasserstein distance [31] defined for probability densities, to non-normalized signed signals.

Definition 2.5. (Signed Wasserstein Metric) Let $r, s \in L^1(\mathbb{R})$ such that $\int r(t)|t|^2 dt < \infty$ and $\int s(t)|t|^2 dt < \infty$, then

$$D_S^2(r, s) := D_{W^2}^2(r^+, s^+) + D_{W^2}^2(r^-, s^-) \quad (2.20)$$

where $D_{W^2}^2(p, q) = W_2^2\left(\frac{p}{\|p\|_1}, \frac{q}{\|q\|_1}\right) + (\|p\|_1 - \|q\|_1)^2$ and $W_2^2(\cdot, \cdot)$ is the usual Wasserstein distance [31].

Utilizing the SCDT and the metric D_S , the following result is a generalization of a well-known isometry.

Lemma 2.6. (*Isometry*) [2] Let $s_1, s_2 \in L^1(\mathbb{R})$ such that $\int s_1(t)|t|^2 dt < \infty$ and $\int s_2(t)|t|^2 dt < \infty$. Then,

$$\begin{aligned} D_S^2(s_1, s_2) &= \|\widehat{s}_1 - \widehat{s}_2\|_{(L^2(r) \times \mathbb{R})^2}^2 \\ &:= \|(s_1^+)^* - (s_2^+)^*\|_{L^2(r)}^2 + (\|s_1^+\|_1 - \|s_2^+\|_1)^2 \\ &\quad + \|(s_1^-)^* - (s_2^-)^*\|_{L^2(r)}^2 + (\|s_1^-\|_1 - \|s_2^-\|_1)^2, \end{aligned}$$

where $\|\cdot\|_{L^2(r)}$ is the norm defined by

$$\|f\|_{L^2(r)}^2 := \int |f(x)|^2 r(x) dx.$$

Now that we have all the preliminaries in place, we define the Radon Signed Cumulative Distribution Transform and see some of its interesting properties that will later help us in classification of signed signals in 2D.

3. Radon Signed Cumulative Distribution Transform

Often many image pre-processing steps like background subtraction leave images having both positive and negative values. An example of such image is shown in Figure 3. Here we expand the definition of the R-CDT provided above (see section 2.4) to allow us to deal with signed images. The idea can be simply summarized as the application of the SCDT to the Radon transform of a signed function. We describe below the mathematical properties of this new representation and show how they can be used to obtain classifiers which are at the same time accurate, closed form, and cheap to compute.

Definition 3.1. (RSCDT) Let $r : \mathbb{R}^2 \rightarrow \mathbb{R}$ be the positive reference image and $s : \mathbb{R}^2 \rightarrow \mathbb{R}$ be the signed image, then Radon Signed CDT of s with respect to r is a composition of the Radon transform and SCDT for each projection angle. Formally, we first take the Radon transform of both the reference and the signal as,

$$\tilde{r}(t, \theta) = \mathcal{R}(r(\mathbf{x})) \quad \text{and} \quad \tilde{s}(t, \theta) = \mathcal{R}(s(\mathbf{x})).$$

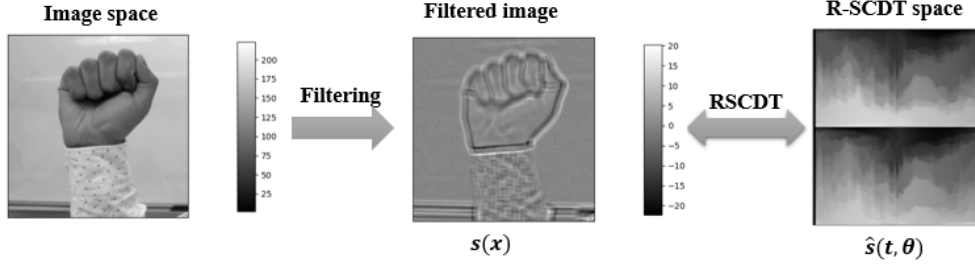


FIGURE 3. The process of obtaining the Radon Signed Cumulative Distribution Transform (R-SCDT) of a grayscale image with background is depicted. In the first step, a filtering technique is applied to suppress the background as well as to enhance the features of interest in the image. The filtered image contains negative pixel values as shown in the figure. To calculate the R-SCDT ($\hat{s}(t, \theta)$) of the filtered image, a two-step process is followed. First, the Radon transform is applied to the filtered image which computes the line integrals at different angles. Next, the signed cumulative distribution transform (SCDT) is applied to each projection obtained from the Radon transform.

Finally, the RSCDT of s with respect to the reference r , for each $\theta \in [0, \pi]$ is given as,

$$\hat{s}(\cdot, \theta) = \mathbb{T}_{\tilde{r}(\cdot, \theta)}(\tilde{s}(\cdot, \theta)) = \left((\tilde{s}^+)^*(\cdot, \theta), \|\tilde{s}^+(\cdot, \theta)\|_1, (\tilde{s}^-)^*(\cdot, \theta), \|\tilde{s}^-(\cdot, \theta)\|_1 \right) \quad (3.1)$$

where $(\cdot)^*$ and $\mathbb{T}_{\tilde{r}(\cdot, \theta)}$ are as defined in (2.18) and (2.19), respectively.

As in the case of its family transforms (CDT, R-CDT and SCDT), a transformed image in the RSCDT space can be recovered via following inverse formula

$$s = \mathcal{R}^{-1}([\mathbb{T}_{\tilde{r}(\cdot, \theta)}^{-1}(\hat{s}(\cdot, \theta))]_{\theta \in [0, \pi]})$$

where $\mathcal{R}^{-1}$ is the inverse of the Radon transform (see (2.6)) and $\mathbb{T}^{-1}$ is the inverse of the SCDT (see (2.19)).

3.1. Properties of RSCDT. Like its family transforms, RSCDT has a few properties outlined below which will be of interest when classifying images.

Proposition 3.2. (Composition Property) Let $s(\mathbf{x})$ denote a signed image and let $\tilde{s}(t, \theta)$ and $\hat{s}(t, \theta)$ be the Radon transform and RSCDT of $s(\mathbf{x})$ respectively. For $g \in \mathcal{G}_{\mathcal{R}}$ (see 2.10) and for each $\theta \in [0, \pi]$, define

$$s_g = \mathcal{R}^{-1}((g^\theta)' \tilde{s} \circ g^\theta) \quad (3.2)$$

where $(g^\theta)'(t) = dg^\theta(t)/dt$, $\tilde{s} \circ g^\theta(\cdot) := \tilde{s}(g^\theta(\cdot), \theta)$. Then the RSCDT of s_g is given by

$$\hat{s}_g(t, \theta) = \left((g^\theta)^{-1} \circ (\tilde{s}^+)^*(t, \theta), \|\tilde{s}^+(\cdot, \theta)\|_1, (g^\theta)^{-1} \circ (\tilde{s}^-)^*(t, \theta), \|\tilde{s}^-(\cdot, \theta)\|_1 \right) \quad (3.3)$$

Like in the previous versions of the transform, the composition property of the RSCDT implies that variations along both the independent and dependent axis directions in an image, caused by $g^\theta(t)$ to the independent variable in the Radon space becomes solely along the dependant variable in the RSCDT space. The following

corollary, gives an explicit example of the composition property for an easy to understand variation of translation in the independent variables.

Corollary 3.3. Let $s(\mathbf{x})$ denote a signed image and let $\tilde{s}(t, \theta)$ and $\hat{s}(t, \theta)$ be the Radon transform and RSCDT of $s(\mathbf{x})$ respectively. Let

$$s_1(x_1, x_2) = s(x_1 - x_0, y_1 - y_0),$$

Then the RSCDT $\hat{s}_1(\cdot, \theta)$ of s_1 for every $\theta \in [0, \pi]$ is given by

$$\hat{s}_1(t, \theta) = \left((g^\theta)^{-1} \circ (\tilde{s}^+)^*(t, \theta), \|\tilde{s}^+(\cdot, \theta)\|_1, (g^\theta)^{-1} \circ (\tilde{s}^-)^*(t, \theta), \|\tilde{s}^-(\cdot, \theta)\|_1 \right) \quad (3.4)$$

where, $g^\theta(t) = t - x_0 \cos \theta - y_0 \sin \theta$.

In certain applications data sets can be rendered convex by family transforms of RSCDT in the respective transform domains. For example, we saw in case of 1D in [21] sets generated by translations of a template signal can have a complex geometry in the signal domain. However, they have a very simple convex structure in the transform domain. This convexification property is useful in classification problems since two disjoint convex data sets can be separated by a linear classifier. The following convexity property is a generalization of the convexity property that was proved in [2] for signed signals in 1D.

Proposition 3.4. (Convexification Property) Let $\phi(\mathbf{x})$ denote a signed image and let $\tilde{\phi}(t, \theta)$ and $\hat{\phi}(t, \theta)$ be the Radon transform and RSCDT of $\phi(\mathbf{x})$ respectively. Let $H_{\mathcal{R}} \subseteq \mathcal{G}_{\mathcal{R}}$. Consider the generative model,

$$\mathbb{S}_{\phi, H_{\mathcal{R}}} := \{\phi_g : \phi_g = \mathcal{R}^{-1}((g^\theta)' \tilde{\phi} \circ g^\theta), g \in H_{\mathcal{R}}\} \quad (3.5)$$

If $H_{\mathcal{R}}^{-1}$ is a convex set then $\hat{\mathbb{S}}_{\phi, H_{\mathcal{R}}} := \{\hat{\phi}_g : \phi_g \in \mathbb{S}_{\phi, H_{\mathcal{R}}}\}$ is convex. Here, $H_{\mathcal{R}}^{-1} := \{g^{-1} = (g^\theta)^{-1}_{\theta \in [0, \pi]}, g \in H_{\mathcal{R}}\}$.

Next is a corollary to the convexification property above that is based on the fact that if G is a group then $G = G^{-1}$.

Corollary 3.5. Let $\phi(\mathbf{x})$ be as in the above proposition. Let $H_{\mathcal{R}}$ be a convex sub-group of $\mathcal{G}_{\mathcal{R}}$ then, $\hat{\mathbb{S}}_{\phi, H_{\mathcal{R}}} := \{\hat{\phi}_g : \phi_g \in \mathbb{S}_{\phi, H_{\mathcal{R}}}\}$ (see (3.5)) is convex.

3.2. Metric Structure. The Wasserstein distance in the space $\mathcal{P}(\mathbb{R})$ (probability space on the real line) is intimately related to the L^2 distance in the transport transform domain as shown in [1, 30, 32]. We then saw a Wasserstein like metric defined in [2] for signed signals and its relationship with L^2 distance in the transform domain for 1D signed signals. This relation has proven to be useful in some applications since it renders certain optimization problems involving the Wasserstein like distances into standard least squares optimization problems. In this section, we generalize the Wasserstein-like distance defined in [2] for signed measures and the sliced Wasserstein distance [15] to define a metric space structure on the space of signed signals (images) in 2D. We then see an analogue of the isometry [2] between the metric on signed images and an analogue of the L^2 distance in the transform domain.

Definition 3.6. (Signed Sliced-Wasserstein Distance) For $s_1, s_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ the distance is defined as

$$\begin{aligned} D(s_1, s_2) &:= \left(\int_0^\pi D_S^2(\tilde{s}_1(\cdot, \theta), \tilde{s}_2(\cdot, \theta)) d\theta \right)^{\frac{1}{2}} \\ &= \left(\int_0^\pi D_{W^2}^2(\tilde{s}_1^+(\cdot, \theta), \tilde{s}_2^+(\cdot, \theta)) d\theta + \int_0^\pi D_{W^2}^2(\tilde{s}_1^-(\cdot, \theta), \tilde{s}_2^-(\cdot, \theta)) d\theta \right)^{\frac{1}{2}}, \end{aligned}$$

where D_S and D_{W^2} are as defined in (2.20).

Proposition 3.7. (Isometry/Embedding) RSCDT induces an isometric embedding between the space of images with Signed Sliced-Wasserstein metric defined above and the space of their RSCDT transforms with a Euclidean-type metric,

$$D(s_1, s_2) = \left(\int_{\theta=0}^\pi \|\hat{s}_1(\cdot, \theta) - \hat{s}_2(\cdot, \theta)\|_{(L^2(\tilde{r}(\cdot, \theta)) \times \mathbb{R})^2}^2 d\theta \right)^{\frac{1}{2}}, \quad (3.6)$$

A compact notation for the above equation would be,

$$D(s_1, s_2) = \|\hat{s}_1 - \hat{s}_2\|_{(L^2(\tilde{r}) \times L^2[0, \pi])^2} \quad (3.7)$$

where

$$\begin{aligned} \|\hat{s}\|_{(L^2(\tilde{r}) \times L^2[0, \pi])^2}^2 &:= \int_0^\pi \|\hat{s}(\cdot, \theta)\|_{(L^2(\tilde{r}(\cdot, \theta)) \times \mathbb{R})^2}^2 d\theta \\ &= \int_0^\pi \left(\int_{\mathbb{R}} |\tilde{s}^+(t, \theta)|^2 \tilde{r}(t, \theta) dt + \|\tilde{s}^+(\cdot, \theta)\|_1^2 \right. \\ &\quad \left. + \int_{\mathbb{R}} |\tilde{s}^-(t, \theta)|^2 \tilde{r}(t, \theta) dt + \|\tilde{s}^-(\cdot, \theta)\|_1^2 \right) d\theta \end{aligned}$$

4. Image Classification using RSCDT

Here we discuss a mathematical model-based problem statement for the type of classification problems we discuss in this paper. Furthermore, we illustrate how the composition and convexification properties of the RSCDT play a crucial role in facilitating the classification of images.

4.1. Signal Class Model and Problem Statement. In numerous applications, we focus on classifying images that are instances of a certain prototype (or template) observed under some often unknown deformation patterns. Let's consider the problem of classifying handwritten digits, such as the MNIST dataset [18]. In such datasets, it is reasonable to assume that each observed digit image can be considered as an instance of a template (or templates) subjected to unknown deformations or variations. For instance, a suitable model for a particular class in the dataset, like the digit 1, would involve a fixed pattern for the digit 1 but with different translations, meaning the digit can appear randomly positioned within the image's field of view. Alternatively, the digit could also exhibit variations in size or slight nonrigid deformations. We denote the set of transformations (translations, nonrigid transformations, etc.) that make the image classification non trivial and hard to solve as *confounds*. The following mathematical model for image (2D signals) classes formalize these concepts.

Definition 4.1 (2D signal class model). Let $\mathcal{G}_{\mathcal{R}} \subset \mathcal{T}$ be the set of confounds. The 2D mass (image intensity) preserving class model for the k^{th} class is defined to be the set

$$\mathbb{S}^{(k)} = \left\{ s_j^{(k)} | s_j^{(k)} = \mathcal{R}^{-1} \left((g_j^\theta)' \tilde{\varphi}^{(k)} \circ g_j^\theta \right), \forall g_j \in \mathcal{G}_{\mathcal{R}} \right\}, \quad (4.1)$$

where $\varphi^{(k)}$ is the template pattern of class k and $s_j^{(k)}$ is the j -th image from that class. Given the signal class model, the mathematical description of the classification problem is defined as:

Definition 4.2 (Classification problem). Let $\mathcal{G}_{\mathcal{R}} \subset \mathcal{T}$ denote the set of confounds and $\mathbb{S}^{(k)}$ be defined as in equation (4.1). Given training samples $\{s_1^{(1)}, s_2^{(1)}, \dots\}$ (class 1), $\{s_1^{(2)}, s_2^{(2)}, \dots\}$ (class 2), $\dots$ as training data, determine the class (k) of an unknown image s .

4.2. Proposed Solution. The signal class model specified in equation (4.1) typically produces signal classes that are nonconvex. However, the RSCDT can simplify the data geometry and thereby simplify the classification problem. Hence, following the methodology proposed in [27], we employ the composition property of the RSCDT on $s_j^{(k)}$ in eq. (4.1), which yields the signal class model in the transform domain as:

$$\widehat{\mathbb{S}}^{(k)} = \left\{ \widehat{s}_j^{(k)} | \widehat{s}_j^{(k)} = (g_j^\theta)^{-1} \circ \tilde{\varphi}^{(k)}, \forall g_j \in \mathcal{G}_{\mathcal{R}} \right\}. \quad (4.2)$$

By applying the convexification property of the RSCDT, as described in Proposition 3.4, it is evident that the class model given in eq. (4.2) forms a convex set if $\mathcal{G}_{\mathcal{R}}$ is a convex group [27]. Furthermore, since the RSCDT is a one-to-one mapping, it follows that if $\mathbb{S}^{(k)} \cap \mathbb{S}^{(p)} = \emptyset$, then $\widehat{\mathbb{S}}^{(k)} \cap \widehat{\mathbb{S}}^{(p)} = \emptyset$. Consequently, we define a subspace generated by the convex set $\widehat{\mathbb{S}}^{(k)}$ as follows:

$$\widehat{\mathbb{V}}^{(k)} = \text{span} \left(\widehat{\mathbb{S}}^{(k)} \right) = \left\{ \sum_{j \in \mathbf{J}} \alpha_j \widehat{s}_j^{(k)} | \alpha_j \in \mathbb{R}, \mathbf{J} \text{ is finite} \right\}. \quad (4.3)$$

It can be demonstrated, under specific assumptions, that the convex space associated with a particular class in the transformed domain does not overlap with the subspace corresponding to a different class [27], i.e. $\widehat{\mathbb{S}}^{(k)} \cap \widehat{\mathbb{V}}^{(p)} = \emptyset, k \neq p$. It follows from the analysis that for a test sample s generated according to the mathematical model for an unknown class k , we have $d^2(\widehat{s}, \widehat{\mathbb{V}}^{(k)}) \sim 0$, and $d^2(\widehat{s}, \widehat{\mathbb{V}}^{(p)}) > 0$ for $p \neq k$. In this context, $d^2(\cdot, \cdot)$ represents the Euclidean distance between $\widehat{s}$ and the nearest point in $\widehat{\mathbb{S}}^{(k)}$ or $\widehat{\mathbb{V}}^{(k)}$. Therefore, the unknown class of the test image can be predicted by employing a nearest subspace search method in the RSCDT space:

$$\underset{k}{\operatorname{argmin}} d^2(\widehat{s}, \widehat{\mathbb{V}}^{(k)}). \quad (4.4)$$

We refer to this approach as the RSCDT-NS method.

4.2.1. Training Algorithm. Based on the aforementioned analysis, we propose a non-iterative training algorithm following the approach outlined in [27]. The algorithm proceeds as follows: First, the transforms of the training samples are computed to obtain $\widehat{\mathbb{S}}^{(k)}$ for all classes. Next, we approximate $\widehat{\mathbb{V}}^{(k)}$ by taking the span of $\widehat{\mathbb{S}}^{(k)}$ resulting in, $\widehat{\mathbb{V}}^{(k)} = \text{span} \left\{ \widehat{s}_1^{(k)}, \widehat{s}_2^{(k)}, \dots \right\}$. The set $\left\{ \widehat{s}_1^{(k)}, \widehat{s}_2^{(k)}, \dots \right\}$ is orthogonalized to obtain the set of

basis vectors $\{b_1^{(k)}, b_2^{(k)}, \dots\}$ that spans the space $\widehat{\mathcal{V}}^{(k)}$ for class k . Finally, the matrix $B^{(k)}$ is formed with the computed basis vectors as its columns, i.e., $B^{(k)} = [b_1^{(k)}, b_2^{(k)}, \dots]$. This matrix is later used for predicting the class label of unknown samples. It is important to note that while this algorithm is designed for classifying signal patterns that can be modeled as instances of a particular template (as defined in equation 4.1), the proposed RSCDT-NS method does not require the estimation of the template $\varphi^{(k)}$ for each class. Instead, one can directly use the training samples for a specific class k to generate the matrix $B^{(k)}$ using the aforementioned algorithm. In addition, it is important to note that the convexity results, and thus our ability to model signed image classes as vector spaces, is independent of the reference chosen to define the RSCDT.

4.2.2. Testing Algorithm. Let's consider the problem of predicting the class of a test image s using the proposed RSCDT-NS method. Firstly, we obtain the RSCDT $\widehat{s}$ of the test sample s . Next, we estimate the distance between $\widehat{s}$ and the subspace model for each class using $d^2(\widehat{s}, \widehat{\mathcal{V}}^{(k)}) \sim \|\widehat{s} - B^{(k)}B^{(k)T}\widehat{s}\|_{L_2}^2$. The class of $\widehat{s}$ is then estimated by solving the optimization problem:

$$\underset{k}{\operatorname{argmin}} \quad \|\widehat{s} - A^{(k)}\widehat{s}\|_{L_2}^2, \quad (4.5)$$

where $A^{(k)} = B^{(k)}B^{(k)T}$ represents an orthogonal projection matrix onto the subspace spanned by the columns of $B^{(k)}$.

5. Computational Results

We evaluated our proposed image classification method on three datasets, showcasing its effectiveness. In the first dataset, a simulated dataset comprising two classes, we demonstrate a scenario where the existing transport-based classification method (e.g., RCDT-NS) fails to differentiate between the two classes. However, our proposed method (RSCDT-NS) achieves 100% accuracy in classifying the two classes. Next, we present another scenario in which negative pixel values can arise during image classification, highlighting the need for a signed transport-based method to handle such situations effectively. In the third case, we applied our proposed method to a real dataset and observed that the signed transport-based method, RSCDT-NS, outperforms both the unsigned methods and state-of-the-art convolutional neural network-based algorithms.

5.1. Result on simulated data. This study employs simulated data consisting of two categories of image data. The first category comprises three circular shapes with randomized sizes and positions. These circular shapes are characterized by two positive pixel values and one negative pixel value. The second category of image data features three circular shapes of arbitrary sizes and random positions. However, in contrast to the first category, two of these circular shapes exhibit negative pixel values, while the remaining one displays positive pixel values (refer to Figure 4). The simulation results, presented in Table 2, compare the performance of two transport-based techniques. The first method utilizes the RSCDT transform on images followed by the nearest subspace (NS) classification technique. The second method applies the RCDT transform on the image after taking the absolute value, as RCDT can only be used on positive distributions. From the results in Table 2, it is evident that the RSCDT-NS method

achieves 100% accuracy in distinguishing the two classes. However, the RCDT-NS method fails to correctly classify the two groups and only achieves chance accuracy.

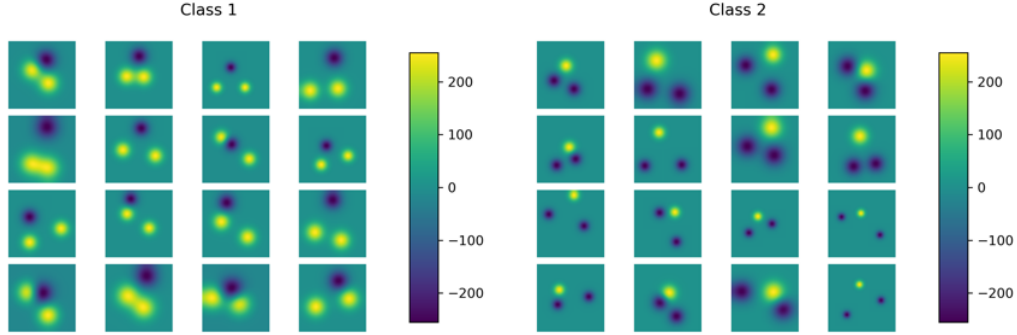


FIGURE 4. (a) Simulated images from class 01, showcasing three circular shapes with randomized sizes and positions. These circular shapes are distinguished by two positive pixel values and one negative pixel value. (b) Simulated images from class 02, featuring three circular shapes with randomized sizes and positions. Unlike class 01, two of these circular shapes exhibit negative pixel values, while the remaining one displays positive pixel values.

TABLE 2. **Comparative analysis of RSCDT-NS and RCDT-NS techniques on a simulated dataset**

Method	Accuracy (%)
RSCDT-NS	100
RCDT-NS	49.00

5.2. Result on 2D Geometric Shape Dataset. The proposed method has also been applied to a 2D geometric shape dataset [7]. This dataset consists of nine geometric shapes. Each shape within a particular class can be considered as a scaled, translated, and/or rotated version of a common template, while utilizing different background and foreground colors (refer to Figure 5). To recognize a shape, the images have been converted to grayscale and then difference of Gaussian (DoG) [8] has been calculated in order to enhance the edge feature. The calculation of DoG introduces negative pixel values. As the RCDT-based transport method [27] cannot directly be implemented on images with negative pixel values, we have employed the proposed RSCDT-based nearest subspace method (RSCDT-NS) to classify the different shapes. The achieved results using RSCDT-NS have been compared with other methods, as presented in

Table 3. It is observed that RSCDT based nearest subspace (RSCDT-NS) method outperforms all other techniques and provides an accuracy of 94.88%. Other transport based method RCDT-NS has also been employed directly on raw images as well as on filtered images after taking the absolute value in order to compare with the proposed method. The performance of both the methods are found low. The effect of adding an offset value with filtered image, to make the image of interest have positive values, has also been assessed for comparative purposes. However, adding a random constant to the images does not follow the generative method delineated in equation (3.5), (4.1), and (4.2) and hence this approach does not yield satisfactory results. Additionally, the performances of several state-of-the-art convolutional neural network (CNN) based frameworks, say Densenet121[11], ResNet-18[10], ShuffleNet[19] and VGG-16[29] have been also evaluated. In terms of classification accuracy, the best performance among the CNN-based networks has been obtained through VGG-16 using 200 training images per class which is 79.11%. On the other hand, the transport-based method proposed for the signed distribution RSCDT-NS can achieve test accuracy 83.88% using only 20 images per class in training phase and can provide 94.88% test accuracy using only 200 images per class in training phase.

TABLE 3. Comparative analysis of different transport based transformations and classification techniques on 2D geometric shape dataset.

Method	No of training samples per class				
	10	20	50	100	200
RSCDT-NS	65.33	83.88	93.11	92.0	94.88
RCDT-NS (Filtered image)	51.44	56.77	78.77	80.22	82.55
RCDT-NS (Filtered image with offset)	11.33	10.88	12.88	14.22	12.66
RCDT-NS (Raw image)	12.88	15.11	15.00	16.44	19.66
DenseNet121	61.22	63.44	63.33	63.56	61.78
ResNet-18	58.83	59.87	64.00	65.11	58.33
ShuffleNet	41.56	44.22	42.67	44.42	44.22
VGG16	60.44	63.94	69.67	76.33	79.11

5.3. Result on sign language dataset. Finally, we meticulously evaluate the performance of our proposed RSCDT-NS method on a real-world sign language dataset [22], which encompasses images of 24 distinct classes of static signs, each representing a specific sign in sign language. Our primary objective in this study is to thoroughly assess the effectiveness of the RSCDT-NS method under various training scenarios.

To gain a deeper understanding of the dataset, we present a visualization of sample images from each class in Figure 6(a). Upon closer inspection, we observe that each image in the dataset contains background pixel information, which can potentially hinder the classification task when employing available transport-based methods such

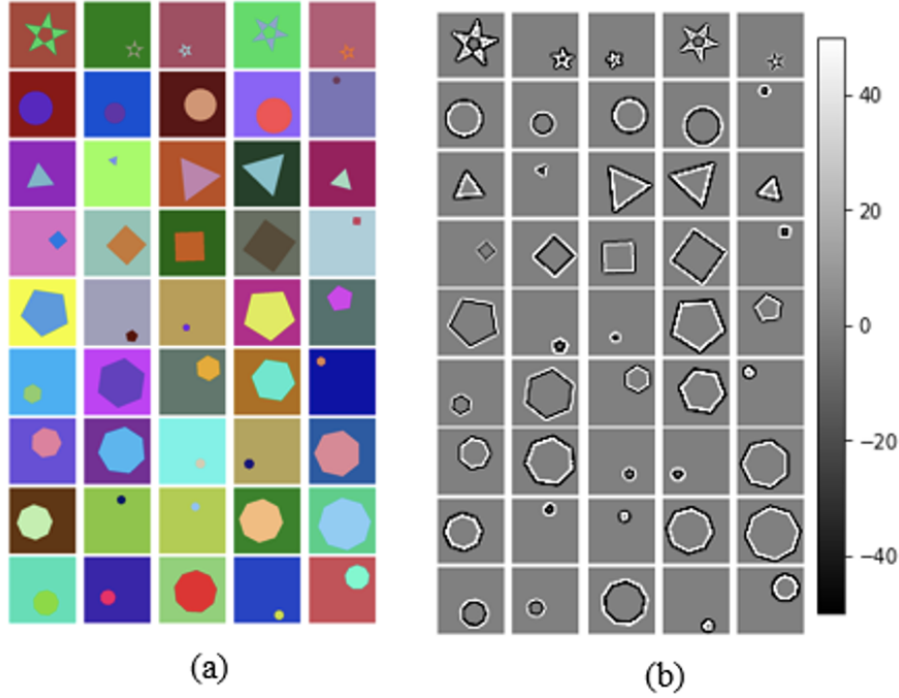


FIGURE 5. (a) Sample images from a 2D geometric pattern dataset. Each row in the figure represents images from a specific class. (b) Images are converted to grayscale and then Difference of Gaussian has been obtained to enhance the edge features in different shapes. This process results in the presence of negative pixel values.

as RCDT-NS. In order to mitigate the influence of the background and enhance the discriminative power of the images, we first convert the images to grayscale and subsequently calculate the difference of Gaussian (DoG) in order to enhance the edge features [8]. This process introduces negative pixel values in the images as illustrated in Figure 6(b). We then apply our proposed transport-based method, RSCDT-NS, specifically developed to handle images with negative pixel values. This preprocessing step significantly improves the suitability of the images for subsequent analysis.

Results in table 4 show that the RSCDT-NS method outperforms the RCDT-NS method, regardless of whether the RCDT is directly applied to the raw images or the RCDT-NS is applied to the absolute filtered images. In this table the performance of the RCDT-NS method applied to filtered images with a constant added so as to render the filtered images positive valued is also evaluated. This robust performance demonstrates the superior capability of the newly introduced RSCDT-NS method in effectively capturing the discriminative features inherent in the sign language dataset.

Furthermore, our comprehensive evaluation illustrates that the performance of the RSCDT-NS method surpasses that of state-of-the-art convolutional neural network (CNN) based methods. Despite CNN based methods being a popular choice for image

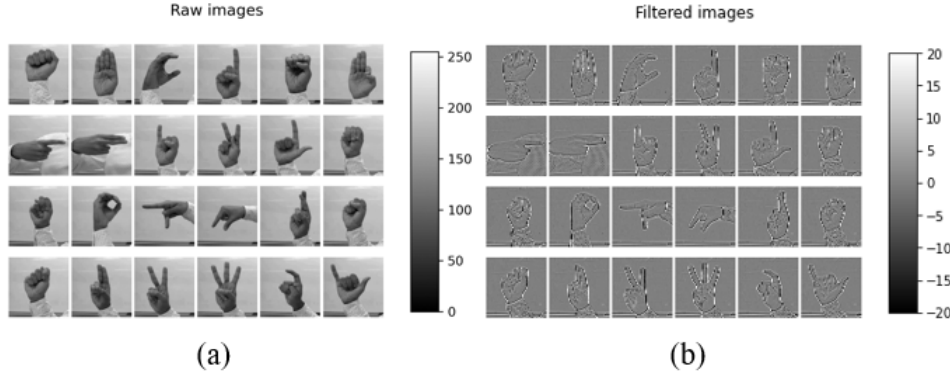


FIGURE 6. (a) Images from a sign language dataset, comprising 24 types of static signs.(b) Corresponding images after undergoing difference of Gaussian (DoG) process. This operation leads to the emergence of negative pixel values.

classification tasks, it exhibits subpar results when trained with a limited number of images per class. The best classification accuracy using CNN based technique has been achieved through ResNet-18 (69.79%). In stark contrast, the RSCDT-NS method showcases its robustness and efficacy by achieving an impressive accuracy of 91.46% using the same number of training images, firmly establishing its superiority in the field of sign language recognition.

These results provide compelling evidence that the proposed RSCDT-NS method holds significant potential for advancing signed image recognition systems, surpassing both available transport-based methods and cutting-edge deep learning approaches. By leveraging the distinctive characteristics of the sign language dataset, the RSCDT-NS method demonstrates its prowess in extracting and utilizing discriminative features, ultimately leading to highly accurate and reliable sign language recognition.

5.4. Limitations. The RSCDT method detailed in this paper shows high effectiveness for image classification, achieving notable accuracy with limited training data and eliminating the need for extensive hyper-parameter tuning. However, this efficiency is contingent on the data adhering to the signal class model delineated in equations (3.5), (4.1), and (4.2). The method's effectiveness declines when dealing with test samples that diverge from this specified model. For instance, this paper employs a signal class model derived from a single template under specific confounds, in line with the model in equation (3.5). When a test sample is produced using either multiple templates or a single template accompanied by confounds that do not align with the model, the method's ability to accurately classify the sample may be compromised.

This limitation is demonstrated in Figure 7, where three different scenarios highlight possible classification failures under these conditions. The first scenario involves an object partially obscured by another object of similar appearance. The second scenario

TABLE 4. **Performance Evaluation of Different Transport-Based Transformations and Classification Techniques on a real sign language dataset**

Method	No of training samples per class				
	4	8	12	16	20
RSCDT-NS	43.75	65.42	77.71	83.75	91.46
RCDT-NS(Filtered image)	40.41	55.66	66.67	76.88	83.96
RCDT-NS (Filtered image with offset)	14.38	19.16	17.71	20.63	22.50
RCDT-NS (Raw image)	22.91	30.83	30.62	48.33	45.83
DenseNet121	48.75	51.25	49.58	52.92	54.79
ResNet-18	43.33	62.29	65.62	69.38	69.79
ShuffleNet	12.29	30.00	27.92	47.52	56.71
VGG16	45.83	53.12	57.08	61.04	60.83

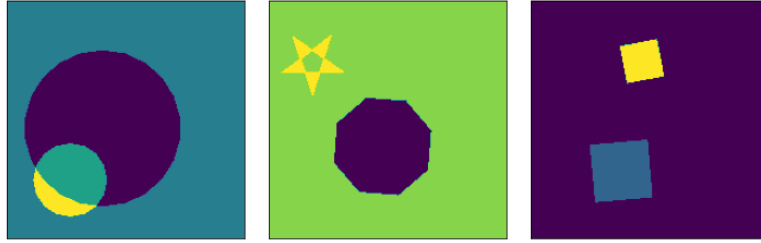


FIGURE 7. Illustrative examples to show when the model might struggle with classifying images. The first image on the left has an object blocking part of another similar object. The second image displays two objects of different shapes. The last image on the right has several objects of the same type. In each of these scenarios, the model might not perform well.

presents several objects originating from different signal class models. In the third scenario, the image features repetitive similar objects. These images may be viewed as a composite of various templates, each with its own distinct confounds, as opposed to being based on a single template. This deviation from the suggested signal class model could lead to the trained model misidentifying the shapes within the test images.

6. Summary, Discussion, and Conclusion

Image representation methods are important components of modern automated image analysis and classification methods. Classification problems, for example, can be difficult to solve in some representations, while dramatically simpler to solve in others. While Fourier and Wavelet transforms have been instrumental in allowing for the simple and effective solutions of numerous filtering, compression, and estimation problems,

when it comes to classification, methods based on ad hoc feature extraction and deep learning have taken center stage given their simplicity and effectiveness. The lack of theoretical foundations for these methods, however, has helped slow the progress of such methodology in recent years particular in terms of performance given limited training sets, interpretability, and assurance guarantees.

Recently methods for representing images using optimal transport have emerged [1, 15, 17], and allowed for the simple non iterative (closed form) solution of a certain category of image classification problems [27, 28]. Up until now, however, the methodology was limited to images that can be interpreted as probability density functions: they must be positive functions and integrate to 1. Here we have extended the previously introduced R-CDT method [15] by combining the R-CDT [15] and SCDT [2] methods and introduced the Radon Signed Cumulative Distribution Transform (RSCDT). Unlike the previous RCDT method, the RSCDT now permits one to use the technique for general (signed) functions, as opposed to only positive functions that integrate to one. We described many properties of the RSCDT method (including composition and convexity of image classes, and a new mathematical distance), and described how to use these in designing simple (closed form), and accurate image classification methods using the nearest subspace technique [27]. Comparisons with state of the art deep learning methods for image classification revealed that the newly proposed RSCDT representation method, when combined with nearest subspace methods [27], can yield superior accuracy. We further note that additional accuracy gains can be obtained by enhancing the nearest subspace method with additional subspace dimensions to encode affine transformations [28] as well as utilizing multiple “local” subspaces to enhance the accuracy of each individual class [26]. The combination of these enhanced classification methods with the newly introduced RSCDT method will be the subject of future work.

In summary, we believe that the RSCDT method introduced here will facilitate the analysis of signed functions (images) in a number of applications. By allowing for invertible transformation of signed images, the RSCDT representation method does not involve information loss. When applied to signed images that can be modeled according to the class model proposed in equation (4.1), the transform can render image classes convex, and thus easy to partition.

7. Acknowledgements

The authors dedicate this paper to Prof. Akram Aldroubi, who has been a teacher, mentor and friend to each one of us. His expertise, wisdom, and willingness to share his knowledge have greatly contributed to our development in the field. We are truly fortunate to have had the opportunity to learn from him and to have him as a guiding force in our academic endeavors.

Authors also acknowledge funding from the NIH (GM130825) and ONR (N000142212505) for supporting this work.

References

- [1] A. Aldroubi, S. Li, and G. K. Rohde, Partitioning signal classes using transport transforms for data analysis and machine learning, *Sampl. Theory Signal Process. Data Anal.*, 19 (2021), no. 1, Paper No. 6, 25 pp.

- [2] A. Aldroubi, R. D. Martin, I. Medri, G. K. Rohde, and S. Thareja, The signed cumulative distribution transform for 1-D signal analysis and classification, *Found. Data Sci.*, 4 (2022), no. 1, 137–163.
- [3] M. Basu, Gaussian-based edge-detection methods-a survey, *IEEE Trans. Syst., Man, Cybern., Part C*, 32 (2002), 252–260.
- [4] R. N. Bracewell, *The Fourier transform and its applications*, Third edition, McGraw-Hill Ser. Electr. Engrg. Circuits Systems, McGraw-Hill Book Co., New York, 1986.
- [5] R. W. Brown, Y.-C. Norman Cheng, E. M. Haacke, M. R. Thompson, and Ramesh Venkatesan, *Magnetic resonance imaging: physical principles and sequence design*, John Wiley & Sons, 2014.
- [6] L. Cattell, C. H. Meyer, F. H. Epstein, and G. K. Rohde, Reconstructing high-resolution cardiac mr movies from under-sampled frames, In: *Proceedings of the 2017 51st Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, USA: IEEE, 2017, pp. 1255–1259.
- [7] A. E. Korchi and Y. Ghanou, 2D geometric shapes dataset - for machine learning and pattern recognition, *Data in Brief*, 32 (2020), Art. ID 106090, 5 pp.
- [8] R. C. Gonzalez and R. E. Woods, *Digital image processing*, vol. 21, 2011.
- [9] N. Hamilton, Quantification and its applications in fluorescent microscopy imaging, *Traffic*, 10 (2009), no. 8, 951–961.
- [10] K. He, X. Zhang, S. Ren, and J. Sun, Deep residual learning for image recognition, In: *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 27–30 June 2016, Las Vegas, NV, USA, pp. 770–778.
- [11] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, Densely connected convolutional networks, In: *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 21–26 July 2017, Honolulu, HI, USA, pp. 4700–4708.
- [12] Imaging and Data Science Lab, *PyTransKit*. <https://github.com/rohdelab/PyTransKit>.
- [13] U. Jung and B.-H. Koh, Wavelet energy-based visualization and classification of high-dimensional signal for bearing fault detection, *Knowl. Inf. Syst.*, 44 (2015), no. 1, 197–215.
- [14] S. Kolouri and G. K. Rohde, Transport-based single frame super resolution of very low resolution face images, In: *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 7–12 June 2015, Boston, MA, USA, pp. 4876–4884.
- [15] S. Kolouri, S. R. Park, and G. K. Rohde, The Radon cumulative distribution transform and its application to image classification, *IEEE Trans. Image Process.*, 25 (2016), no. 2, 920–934.
- [16] S. Kolouri, S. R. Park, M. Thorpe, D. Slepcev, and G. K. Rohde, Optimal mass transport: Signal processing and machine-learning applications, *IEEE Signal Process. Mag.*, 34 (2017), no. 4, 43–59.
- [17] S. Kolouri, A. B. Tosun, J. A. Ozolek, and G. K. Rohde, A continuous linear optimal transport approach for pattern analysis in image datasets, *Pattern Recognit.*, 51 (2016), 453–462.
- [18] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, Gradient-based learning applied to document recognition, *Proc. IEEE*, 86 (1998), no. 11, 2278–2324.
- [19] N. Ma, X. Zhang, H.-T. Zheng, and J. Sun, ShuffleNet V2: Practical guidelines for efficient CNN architecture design, In: *Proceedings of the European Conference on Computer Vision (ECCV)*, 8–14 September 2018, Munich, Germany, pp. 116–131.
- [20] F. Natterer, *The mathematics of computerized tomography*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001.
- [21] S. R. Park, S. Kolouri, S. Kundu, and G. K. Rohde, The cumulative distribution transform and linear pattern classification, *Appl. Comput. Harmon. Anal.*, 45 (2018), no. 3, 616–641.

- [22] R. F. Pinto, C. D. B. Borges, A. M. A. Almeida, and I. C. Paula, Static hand gesture recognition based on convolutional neural networks, *J. Electr. Comput. Eng.*, (2019), Art. ID 4167890, 12 pp.
- [23] E. T. Quinto, An introduction to x-ray tomography and radon transforms, An introduction to X-ray tomography and Radon transforms, In: American Mathematical Society Short Course, 3–4 January 2005, Atlanta, GA, *Proc. Symp. Appl. Math.* vol. 63, 2006, AMS, Providence RI, pp. 1–23.
- [24] J. Radon, On the determination of functions from their integrals along certain manifolds, *Math. Phys. Klass.*, 69 (1917), 262–277.
- [25] A. H. M. Rubaiyat, K. M. Hallam, J. M. Nichols, M. N. Hutchinson, S. Li, and G. K. Rohde, Parametric signal estimation using the cumulative distribution transform, *IEEE Trans. Signal Process.*, 68 (2020), 3312–3324.
- [26] A. H. M. Rubaiyat, S. Li, X. Yin, M. Shifat-E-Rabbi, Y. Zhuang, and G. K. Rohde, End-to-end signal classification in signed cumulative distribution transform space, eprint arXiv:2205.00348 [eess.SP], Apr 2022.
- [27] M. Shifat-E-Rabbi, X. Yin, A. H. M. Rubaiyat, S. Li, S. Kolouri, A. Aldroubi, J. M. Nichols, and G. K. Rohde, Radon cumulative distribution transform subspace modeling for image classification, *J. Math. Imaging Vision*, 63 (2021), no. 9, 1185–1203.
- [28] M. Shifat-E-Rabbi, Y. Zhuang, S. Li, A. H. M. Rubaiyat, X. Yin, and G. K. Rohde, Invariance encoding in sliced-Wasserstein space for image classification with limited training data, *Pattern Recognit.*, 137 (2023), Art. ID 109268, 12 pp.
- [29] K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition, eprint arXiv:1409.1556 [cs.CV], Sep 2014.
- [30] M. Thorpe, Introduction to optimal transport. <https://www.math.cmu.edu/~mthorpe/OTNotes>.
- [31] C. Villani, *Topics in optimal transportation*, Grad. Stud. Math., 58, American Mathematical Society, Providence, RI, 2003.
- [32] C. Villani, *Optimal transport: old and new*, Grundlehren Math. Wiss., 338 [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 2009.
- [33] W. Wang, D. Slepčev, S. Basu, J. A. Ozolek, and G. K. Rohde, A linear optimal transportation framework for quantifying and visualizing variations in sets of images, *Int. J. Comput. Vis.*, 101 (2013), no. 2, 254–269.
- [34] Y. Zhuang, S. Li, M. Shifat-E-Rabbi, X. Yin, A. H. M. Rubaiyat, and G. K. Rohde, Local sliced-wasserstein feature sets for illumination-invariant face recognition, eprint arXiv:2202.10642 [cs.CV], Feb 2022.

Appendix A.

A.1. Proof that $\mathcal{G}_{\mathcal{R}}$ is a group.

Proof. Here we show that $(\mathcal{G}_{\mathcal{R}}, \star)$ is a group. Recall, from equation (2.10),

$$\mathcal{G}_{\mathcal{R}} = \{g = (g^\theta)_{\theta \in [0, \pi]} : g^\theta : \mathbb{R} \rightarrow \mathbb{R} \text{ is a strictly increasing bijection } \forall \theta \in [0, \pi]\}$$

and the operation $\star$ is defined as, $\star : \mathcal{G}_{\mathcal{R}} \times \mathcal{G}_{\mathcal{R}} \rightarrow \mathcal{G}_{\mathcal{R}}$, $(g \star h)(\cdot, \theta) := (g^\theta \circ h^\theta)(\cdot)$ for all $\theta \in [0, \pi]$ and $g, h \in \mathcal{G}_{\mathcal{R}}$. Now, let's look at the individual group properties,

- Closure : for any $g, h \in \mathcal{G}_{\mathcal{R}}$, $g \star h$ is component wise composition of strictly increasing functions for each $\theta \in [0, \pi]$, hence $g \star h$ is also a strictly increasing surjection. Thus, $g \star h \in \mathcal{G}_{\mathcal{R}}$.
- Associativity : Composition of functions is always an associative operation.
- Identity : $I = (I^\theta)_{\theta \in [0, \pi]}$ such that $I^\theta(t) = t$, for all $t \in \mathbb{R}$ and $\theta \in [0, \pi]$.
- Inverse : Note that if α is a strictly increasing bijection then so is α^{-1} . So, for any $g = (g^\theta)_{\theta \in [0, \pi]} \in \mathcal{G}_{\mathcal{R}}$, define $g^{-1} = ((g^\theta)^{-1})_{\theta \in [0, \pi]}$ to be the inverse of g . Clearly, $g \star g^{-1} = (g^\theta \circ (g^\theta)^{-1})_{\theta \in [0, \pi]} = (I^\theta)_{\theta \in [0, \pi]} = I$.

□

A.2. Proof of Composition Property (see Proposition 3.2).

Proof. Let $s_g = \mathcal{R}^{-1}((g^\theta)' \tilde{s} \circ g^\theta)$ where $g \in \mathcal{G}_{\mathcal{R}}$ and the composition operation is in terms of the first variable of $\tilde{s}(\cdot, \theta)$. Then by the definition of $\mathcal{R}$, we have $\tilde{s}_g = (g^\theta)' \tilde{s} \circ g^\theta$, which implies the following in terms of the corresponding CDFs: $F_{\tilde{s}_g(\cdot, \theta)} = F_{\tilde{s}(\cdot, \theta)} \circ g^\theta, \forall \theta$. Now the goal is to evaluate SCDT for each $\tilde{s}_g(\cdot, \theta)$, which we denote $\tilde{s}_{g^\theta}(\cdot)$ for simplicity. Similarly we denote $(\tilde{s}_g)^\pm(\cdot, \theta)$ as $\tilde{s}_{g^\theta}^\pm(\cdot)$. First note that

$$\|\tilde{s}_{g^\theta}^\pm\|_1 = \|\tilde{s}^\pm(\cdot, \theta)\|_1, \forall \theta. \quad (\text{A.1})$$

Indeed, since g^θ is strictly increasing, we have

$$\tilde{s}_{g^\theta}^\pm = (g^\theta)' \tilde{s}^\pm \circ g^\theta, \quad (\text{A.2})$$

and by the change of variables formula

$$\|\tilde{s}_{g^\theta}^\pm\|_1 = \int_{\mathbb{R}} \tilde{s}_{g^\theta}^\pm(t) dt = \int_{\mathbb{R}} (g^\theta(t))' \tilde{s}^\pm(g^\theta(t), \theta) dt = \int_{\mathbb{R}} \tilde{s}^\pm(u, \theta) du = \|\tilde{s}^\pm(\cdot, \theta)\|_1.$$

Similarly one can show that

$$F_{\tilde{s}_{g^\theta}^\pm} = F_{\tilde{s}^\pm(\cdot, \theta)} \circ g^\theta, \forall \theta. \quad (\text{A.3})$$

Then by definition of the SCDT in (2.17), the composition property follows from equations (A.1) and (A.3) proved above. Indeed,

$$\begin{aligned}
\widehat{s}_g &= \left((\widetilde{s}_{g^\theta}^+)^*, \|\widetilde{s}_{g^\theta}^+\|_1, (\widetilde{s}_{g^\theta}^-)^*, \|\widetilde{s}_{g^\theta}^-\|_1 \right) \\
&= \left(F^{\dagger}_{\frac{\widetilde{s}_g^+}{\|\widetilde{s}_g^+\|_1}} \circ F_{\frac{\widetilde{r}}{\|\widetilde{r}\|_1}}, \|\widetilde{s}_{g^\theta}^+\|_1, F^{\dagger}_{\frac{\widetilde{s}_g^-}{\|\widetilde{s}_g^-\|_1}} \circ F_{\frac{\widetilde{r}}{\|\widetilde{r}\|_1}}, \|\widetilde{s}_{g^\theta}^-\|_1 \right) \\
&= \left(\left(F_{\frac{\widetilde{s}^+}{\|\widetilde{s}^+\|_1}} \circ g^\theta \right)^\dagger \circ F_{\frac{\widetilde{r}}{\|\widetilde{r}\|_1}}, \|\widetilde{s}^+\|_1, \left(F_{\frac{\widetilde{s}^-}{\|\widetilde{s}^-\|_1}} \circ g^\theta \right)^\dagger \circ F_{\frac{\widetilde{r}}{\|\widetilde{r}\|_1}}, \|\widetilde{s}^-\|_1 \right) \\
&= \left((g^\theta)^{-1} \circ F^{\dagger}_{\frac{\widetilde{s}^+}{\|\widetilde{s}^+\|_1}} \circ F_{\frac{\widetilde{r}}{\|\widetilde{r}\|_1}}, \|\widetilde{s}^+\|_1, (g^\theta)^{-1} \circ F^{\dagger}_{\frac{\widetilde{s}^-}{\|\widetilde{s}^-\|_1}} \circ F_{\frac{\widetilde{r}}{\|\widetilde{r}\|_1}}, \|\widetilde{s}^-\|_1 \right) \\
&= \left((g^\theta)^{-1} \circ (\widetilde{s}^+)^*, \|\widetilde{s}^+\|_1, (g^\theta)^{-1} \circ (\widetilde{s}^-)^*, \|\widetilde{s}^-\|_1 \right),
\end{aligned}$$

where the second to the last equality follows from the fact that g^θ is strictly increasing. $\square$

A.3. Proof of the Convexification Property (see Proposition 3.4).

Proof. Assume $H_{\mathcal{R}}^{-1}$ is a convex set. We show that $\widehat{S}_{\phi, H_{\mathcal{R}}}$ is convex. To that end, let $\widehat{\phi}_g, \widehat{\phi}_h \in \widehat{S}_{\phi, H_{\mathcal{R}}}$.

Remark A.1. For the set of equations below, note that $g, h \in H_{\mathcal{R}}$, i.e. $g := (g^\theta)_{\theta \in [0, \pi]}$ and $h := (h^\theta)_{\theta \in [0, \pi]}$. The equality below holds true for all $\theta \in [0, \pi]$ because the operations are done component wise. To simplify our calculations here, we avoid the cumbersome θ notation.

For $\lambda \in [0, 1]$, consider the convex combination,

$$\begin{aligned}
\lambda \widehat{\phi}_g + (1 - \lambda) \widehat{\phi}_h &= \lambda \left((\widetilde{\phi}_g^+)^*, \|\widetilde{\phi}_g^+\|_1, (\widetilde{\phi}_g^-)^*, \|\widetilde{\phi}_g^-\|_1 \right) + (1 - \lambda) \left((\widetilde{\phi}_h^+)^*, \|\widetilde{\phi}_h^+\|_1, (\widetilde{\phi}_h^-)^*, \|\widetilde{\phi}_h^-\|_1 \right) \\
&= \left(\lambda (\widetilde{\phi}_g^+)^* + (1 - \lambda) (\widetilde{\phi}_h^+)^*, \lambda \|\widetilde{\phi}_g^+\| + (1 - \lambda) \|\widetilde{\phi}_h^+\|, \right. \\
&\quad \left. \lambda (\widetilde{\phi}_g^-)^* + (1 - \lambda) (\widetilde{\phi}_h^-)^*, \lambda \|\widetilde{\phi}_g^-\| + (1 - \lambda) \|\widetilde{\phi}_h^-\| \right) \\
&= \left(\lambda (\widetilde{\phi}_g^+)^* + (1 - \lambda) (\widetilde{\phi}_h^+)^*, \|\widetilde{\phi}^+\|, \lambda (\widetilde{\phi}_g^-)^* + (1 - \lambda) (\widetilde{\phi}_h^-)^*, \|\widetilde{\phi}^-\| \right) \\
&= \left(\lambda (g^{-1} \circ (\widetilde{\phi}^+)^*) + (1 - \lambda) (h^{-1} \circ (\widetilde{\phi}^+)^*), \|\widetilde{\phi}^+\|, \right. \\
&\quad \left. \lambda (g^{-1} \circ (\widetilde{\phi}^-)^*) + (1 - \lambda) (h^{-1} \circ (\widetilde{\phi}^-)^*), \|\widetilde{\phi}^-\| \right) \\
&= \left((\lambda g^{-1} + (1 - \lambda) h^{-1}) \circ (\widetilde{\phi}^+)^*, \|\widetilde{\phi}^+\|, (\lambda g^{-1} + (1 - \lambda) h^{-1}) \circ (\widetilde{\phi}^-)^*, \|\widetilde{\phi}^-\| \right) \\
&= \left(p^{-1} \circ (\widetilde{\phi}^+)^*, \|\widetilde{\phi}^+\|, p^{-1} \circ (\widetilde{\phi}^-)^*, \|\widetilde{\phi}^-\| \right)
\end{aligned}$$

where $p^{-1} = \lambda g^{-1} + (1 - \lambda) h^{-1} \in H_{\mathcal{R}}$ and the norms $\|(\widetilde{\phi})_g^\pm\|_1 = \|(\widetilde{\phi})_h^\pm\|_1 = \|(\widetilde{\phi})^\pm\|_1$, as shown in the proof of Composition property A.2 above. Thus, $\lambda \widehat{\phi}_g + (1 - \lambda) \widehat{\phi}_h \in \widehat{S}_{\phi, H_{\mathcal{R}}}$. $\square$

A.4. Proof that $D(\cdot, \cdot)$ is a metric (see def. 3.6) and D induces an isometry (see prop. 3.7).

Proof. Here we show that D is a metric. Recall by definition 3.6,

$$\begin{aligned} D(s_1, s_2) &:= \left(\int_0^\pi D_S^2(\tilde{s}_1(\cdot, \theta), \tilde{s}_2(\cdot, \theta)) d\theta \right)^{\frac{1}{2}} \\ &= \left(\int_0^\pi D_{W^2}^2(\tilde{s}_1^+(\cdot, \theta), \tilde{s}_2^+(\cdot, \theta)) d\theta + \int_0^\pi D_{W^2}^2(\tilde{s}_1^-(\cdot, \theta), \tilde{s}_2^-(\cdot, \theta)) d\theta \right)^{\frac{1}{2}}, \end{aligned}$$

where D_S and D_{W^2} are as defined in (2.20).

- Non - Negativity ($D(\cdot, \cdot) \geq 0$) : Note that $D_S^2(\cdot, \cdot) \geq 0$, therefore, $D(\cdot, \cdot) \geq 0$.
- $D(s_1, s_2) = 0 \implies s_1 = s_2$: Assume, $D(s_1, s_2) = 0$. This implies, $\int_0^\pi D_S^2(\tilde{s}_1(\cdot, \theta), \tilde{s}_2(\cdot, \theta)) d\theta = 0$. Since, $D_S^2(\cdot, \cdot) \geq 0$, therefore, $D_S(\tilde{s}_1(\cdot, \theta), \tilde{s}_2(\cdot, \theta)) = 0 \quad \forall \theta \in [0, \pi]$. Hence, $\tilde{s}_1(\cdot, \theta) = \tilde{s}_2(\cdot, \theta) \quad \forall \theta \in [0, \pi]$ which further implies, $s_1 = s_2$.
- Symmetry ($D(s_1, s_2) = D(s_2, s_1)$) : This follows from the fact that $D_S(\cdot, \cdot)$ is a metric (see [2]) and hence symmetric. Therefore, $D_S^2(\tilde{s}_1(\cdot, \theta), \tilde{s}_2(\cdot, \theta)) = D_S^2(\tilde{s}_2(\cdot, \theta), \tilde{s}_1(\cdot, \theta)) \quad \forall \theta \in [0, \pi]$. Integrating both sides with respect to $\theta \in [0, \pi]$, we have $D(s_1, s_2) = D(s_2, s_1)$.
- Triangle inequality ($D(s_1, s_2) \leq D(s_1, r) + D(r, s_2)$) : By definition,

$$\begin{aligned} D^2(s_1, r) &= \int_0^\pi D_S^2(\tilde{s}_1(\cdot, \theta), \tilde{r}(\cdot, \theta)) d\theta \\ &= \int_0^\pi \|\hat{s}_1(\cdot, \theta) - \hat{r}(\cdot, \theta)\|_{L^2(\tilde{r}(\cdot, \theta))}^2 d\theta \\ &= \int_0^\pi \int_{\mathbb{R}} (\hat{s}_1(t, \theta) - t)^2 \tilde{r}(t, \theta) dt d\theta \end{aligned}$$

where the second equality follows from the isometry of $D_S(\cdot, \cdot)$ [2]. Similarly,

$$D^2(s_2, r) = \int_0^\pi \int_{\mathbb{R}} (\hat{s}_2(t, \theta) - t)^2 \tilde{r}(t, \theta) dt d\theta \quad (\text{A.4})$$

Let $h(t, \theta)$ be the RSCDT of s_2 with respect to s_1 , i.e.

$$D^2(s_2, s_1) = \int_0^\pi \int_{\mathbb{R}} (h(t, \theta) - t)^2 \tilde{s}_1(t, \theta) dt d\theta \quad (\text{A.5})$$

Following the composition property, we have $h(\cdot, \theta) \circ \hat{s}_1(\cdot, \theta) = \hat{s}_2(\cdot, \theta)$ and $\tilde{r}(\cdot, \theta) = \hat{s}_1'(\cdot, \theta) \tilde{s}_1(\hat{s}_1(\cdot, \theta), \theta)$ for every $\theta \in [0, \pi]$. Using change of variables

$$\begin{aligned}
 t &= \widehat{s}(u, \theta) \text{ in (A.5),} \\
 D^2(s_2, s_1) &= \int_0^\pi \int_{\mathbb{R}} (h(\widehat{s}_1(u, \theta), \theta) - \widehat{s}_1(u, \theta))^2 \widehat{s}_1(\widehat{s}_1(u, \theta), \theta) \frac{d\widehat{s}_1(u, \theta)}{du} du d\theta \\
 &= \int_0^\pi \int_{\mathbb{R}} (\widehat{s}_2(u, \theta) - \widehat{s}_1(u, \theta))^2 \widetilde{r}(u, \theta) du d\theta \\
 &= \|\widehat{s}_2 - \widehat{s}_1\|_{(L^2(\widetilde{r}) \times L^2[0, \pi])^2}^2
 \end{aligned}$$

This establishes the isometry i.e. $D(s_2, s_1) = \|\widehat{s}_2 - \widehat{s}_1\|_{(L^2(\widetilde{r}) \times L^2[0, \pi])^2}$. Using this relation, we finally have the required triangle inequality for the metric,

$$\begin{aligned}
 D(s_2, s_1) &= \|\widehat{s}_2 - \widehat{s}_1\|_{(L^2(\widetilde{r}) \times L^2[0, \pi])^2} \\
 &\leq \|\widehat{s}_2 - I\|_{(L^2(\widetilde{r}) \times L^2[0, \pi])^2} + \|I - \widehat{s}_1\|_{(L^2(\widetilde{r}) \times L^2[0, \pi])^2} \\
 &= D(s_2, r) + D(r, s_1)
 \end{aligned}$$

□

(Le Gong) DEPARTMENT OF MATHEMATICS, VANDERBILT UNIVERSITY, NASHVILLE, TN 37212, USA

E-mail address: le.gong@vanderbilt.edu

(Shiying Li) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL, CB #3250, CHAPEL HILL, NORTH CAROLINA 27599-3250

E-mail address: shiy1@unc.edu

(Naqib Sad Pathan) DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING, UNIVERSITY OF VIRGINIA, CHARLOTTESVILLE, VA 22904, USA

E-mail address: qpb3vt@virginia.edu

(Mohammad Shifat-E-Rabbi) DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING, NORTH SOUTH UNIVERSITY, BASHUNDHARA, DHAKA, 1229 BANGLADESH

E-mail address: rabbi.mohammad@northsouth.edu

(Gustavo K. Rohde) DEPARTMENT OF BOMEDICAL ENGINEERING, DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING, UNIVERSITY OF VIRGINIA, CHARLOTTESVILLE, VA 22908, USA

E-mail address: gustavo.rohde@gmail.com

(Abu Hasnat Mohammad Rubaiyat) U.S. NAVAL RESEARCH LABORATORY, WASHINGTON, DC 20375, USA

E-mail address: ar3fx@virginia.edu

(Sumati Thareja) DEPARTMENT OF MATHEMATICS, VANDERBILT UNIVERSITY, NASHVILLE, TN 37212, USA

E-mail address: sumatithareja5@gmail.com