Letter

Model reduction of traveling-wave problems via Radon cumulative distribution transform

Jie Ren [1], William R. Wolf [2], and Xuerui Mao [1,*]

[1] *Department of Mechanical Engineering, Faculty of Engineering, University of Nottingham, Nottingham NG7 2RD, United Kingdom*

[2] *School of Mechanical Engineering, Universidade Estadual de Campinas, Campinas, São Paulo 13083-860, Brazil*

(Received 1 February 2021; accepted 3 August 2021; published 17 August 2021)

Traveling-wave problems, due to their sizable Kolmogorov *n*-width, have brought critical challenges to conventional model reduction techniques. This study aims to provide insights into this problem by exploiting the Radon cumulative distribution transform (R-CDT) [Kolouri, Park, and Rohde, IEEE Trans. Image Process. **25**, 920 (2016)], which has emerged in the sector of computer vision science. The core lies in the unique property of the nonlinear invertible R-CDT that renders both traveling and scaling components into amplitude modulations. In contrast to the physical space, a substantial model reduction is achieved in the R-CDT space while sustaining high accuracy. The method is parameter-free and data-driven and lends itself to problems regardless of the dimensions or boundary conditions. Examples start with a one-dimensional Burgers' equation subject to nonperiodic boundary conditions, where both traveling and diffusion dominate the physics. In higher-dimensional problems, we show the model reduction of traveling Gaussian solitons. In addition to foreseeable motions, the proposed method is capable of handling random traveling with a nondifferentiable trajectory.

DOI: 10.1103/PhysRevFluids.6.L082501

I. INTRODUCTION

Model reduction of physical data, either simulated or measured, plays a significant role in understanding the concealed mechanisms while granting plentiful downstream applications, e.g., data compression, model-based control, and predictions [1]. The reduction has been backed by the fact that the inherent physics is usually sparse in the degree of freedom [2] despite the bulky size of the generated data. However, with standard model reduction techniques, e.g., proper orthogonal decomposition (POD) and dynamic mode decomposition (DMD), the accuracy and efficiency immensely rely on the narrowness of the data's Kolmogorov *n*-width, which traveling-wave problems fail to minimize.

There has been a growing interest in advancing model reduction techniques for traveling-wave problems. In Table I we compare some recently developed methods. Notably, one line of thought dates back to the shifted-based template-fitting method [3], in which the data can be properly shifted in space such that a traveling wave appears to be stationary in the transformed coordinates. To achieve adequate model reduction and diminish the loss of accuracy, this type of method will require one to determine the appropriate shifting matrix and introduce algorithms to minimize the error.

The optimal shift matrix recently became addressed in an optimization framework developed by Mendible *et al.* [4]. This method, unsupervised traveling-wave identification with shifting and truncation (UnTWIST), seeks the analytical representations of multiple traveling velocities by com-

*maoxuerui@sina.com

TABLE I. Comparison of recently developed dimensionality-reduction methods for traveling-wave problems.

Method	Spatial dimension	Boundary condition
UnTWIST [4]	1D	Periodic
Shifted-POD [5]	1D, 2D	Periodic
cDMD [6]	1D, 2D	Periodic
Transport reversal [7]	1D but extendable	Any
Data calibration [8]	1D	Periodic
TSMOR [9]	1D, 2D	Any
AMD [10]	1D, 2D	Any
AADEIM [11]	1D, 2D, 3D	Any
Radon-CDT (present method)	1D, 2D, 3D	Any

bining ridge detection, spectral clustering, and sparse relaxed regularized regression. A physically precise shift is thus made possible for each traveling wave carried by the system. Besides algorithms recognizing the transport velocities, shifted-POD [5] provides a pioneering approach accounting for the postshift stage, in which the data shifting introduces errors when multiple different waves are present. Shifted-POD iteratively minimizes the residual between the data snapshot and its POD approximation with a least-squares optimization, capable of achieving machine precision. In line with the idea of shift-based methods, characteristic DMD (cDMD) [6] rotates the coordinate system in both space and time. The rotating angle is determined by the group velocity of the traveling waves, such that the physics in the new frame develops along the characteristic line, supporting a significant singular value drop.

However, shift-based methods confront data loss unless the boundary locks up the traveling waves with periodic or reflecting conditions. Transport reversal [7] generalized the template fitting approach [3] by introducing a set of enhancement algorithms (e.g., greedy iteration, cutoff vectors). The method thus manages more general problems (with varying shapes, nonperiodic boundary conditions).

Another prospect to model traveling-wave problems is through a preconditioning operation. Data calibration [8] makes use of an invertible mapping to calibrate the data, where the Kolmogorov n-width becomes reduced such that the required modes are much less. The method has been developed in the 1D framework with periodic boundary conditions. The proposed transform still requires *a priori* expertise on the behavior of the physical data.

Apart from works described above, various new or improved model reduction methods have been proposed. For example, transported snapshot model order reduction (TSMOR) [9] is specially designed to handle problems with moving shocks (discontinuities). The transports are approximated as a low-order polynomial expansion. Advection mode decomposition (AMD) [10] defines an advection-mode hierarchy based on Wasserstein distance to separate the advection feature of the flow while the residuals are decomposed with standard POD. Peherstorfer [11] proposed the adaptive bases and adaptive sampling discrete empirical interpolation method (AADEIM), exploiting the temporal and spatial locality of convective structures. The local basis gets updated online following an adaptive sampling scheme querying the full model at a few selected spatial coordinates.

While the above-reviewed research handle traveling-wave problems in various aspects, assessment of crucial properties (e.g., problem dimension, boundary condition as tabulated in Table I) shows that the model reduction towards practical applications is far from mature. In particular, a majority of the methods are not ready for 3D problems. Some methods are target-oriented and require essential knowledge of the physical problem to be known in advance, while shift-based methods put restrictions on boundary conditions. In this study, we aim to boost the effort with ideas generated in the sector of computer vision science. We propose to build reduced-order models in the Radon cumulative distribution transformed (R-CDT) space, such that the Kolmogorov n-width becomes significantly reduced. The method solves multidimensional problems and possesses

flexibility with boundary conditions. It is also parameter-free and explicit; thus training as generally required by machine-learning [12] is not necessary. The rest of the Letter begins by introducing the methodology followed by examples with increasing dimensions.

II. METHODOLOGY

The heart of the proposed R-CDT space is the cumulative distribution transform (CDT) [13], while the Radon transform is applied to break higher-dimensional problems into a series of 1D signals. CDT interprets 1D signals as probability density functions. To meet this requirement, a given input signal $f(x)$ and a reference signal $r(x)$, both defined on $[x_1, x_2]$, are first normalized such that

$$\int_{x_1}^{x_2} f(x)\,dx = \int_{x_1}^{x_2} r(x)\,dx = 1, \text{ with } f(x) > 0,\ r(x) > 0. \tag{1}$$

The forward CDT of $f(x)$ with respect to $r(x)$ is defined by the strictly increasing function $\hat{f}(x)$ that satisfies

$$\int_{x_1}^{\hat{f}(x)} f(x')\,dx' = \int_{x_1}^{x} r(x')\,dx'. \tag{2}$$

The inverse of CDT (iCDT) is obtained by differentiating (2) with respect to x:

$$f(x) = r(\hat{f}^{-1}(x))\frac{d}{dx}\hat{f}^{-1}(x), \text{ where } \hat{f}^{-1}(\hat{f}(x)) = x. \tag{3}$$

A denormalization step is applied to bring the signal to its original range. CDT establishes a nonlinear invertible one-to-one mapping between the signal and its transformation and, therefore, retains all the information contained in the normalized signal. The advantage of CDT over the other linear transforms, e.g., Fourier transform or wavelet transform, lies in its exclusive composition properties that render traveling and scaling components into amplitude modulations. For example, the CDT of a traveling signal $f(\alpha x - ct)$ with phase velocity c and scaling factor α is given by $[\hat{f}(x) + ct]/\alpha$. A proof is given in the Appendix.

Figure 1 demonstrates this property. The traveling wave [colored by the time traveled in Fig. 1(a)] converts to a purely stationary signal with growing amplitude in the CDT space shown in Fig. 1(b). In Figs. 1(c) and 1(d), we examine the correlation matrix $\boldsymbol{R}_{ij} = \langle f(x, t_i), f(x, t_j)\rangle$. Generally speaking, a fast decay of eigenvalues of $\boldsymbol{R}$ warrants a narrow Kolmogorov n-width and an effective model reduction. This decay is determined by the nonorthogonality of $\boldsymbol{R}$, which can be viewed as a deviation of $\boldsymbol{R}$ from the identity matrix (the identity matrix gives rise to nondecaying eigenvalues). As shown in Figs. 1(c) and 1(d), the diagonally dominant correlation matrix in the physical space becomes a globally distributed matrix in the CDT space [Fig. 1(d)], retaining a reduced Kolmogorov n-width. In the present study, we define the range of function $\hat{f}(x)$ to be the same as its domain, namely, $\hat{f}(x_1) = x_1$, $\hat{f}(x_2) = x_2$. This confinement will help to ensure the robustness and reversibility of CDT provided the signal is differentiable and properly normalized. From (2), $r(x)$ determines the shape of the transformed profile $\hat{f}(x)$. Nonetheless, it has been tested that different choices of $r(x)$, for example, an arbitrary snapshot of the input data at a chosen time step t_r, $r(x) = f(x, t_r)$ or standard analytic functions $r(x) = \sin(x) + 2$, lead to similar performance of model reduction. In the examples showed below, results are obtained with $r(x) = f(x, t_r)$. CDT, in nature, has the same mathematical foundations as optimal transportation (OT) [14]. Here "optimal" lies in the minimal Wasserstein distance between the signal and template, defined as

$$\|L\|_2 = \left\{\int_{x_1}^{x_2} [\hat{f}(x) - x]^2 r(x)\,dx\right\}^{\frac{1}{2}}, \tag{4}$$

reflecting the geometric connotation of CDT.

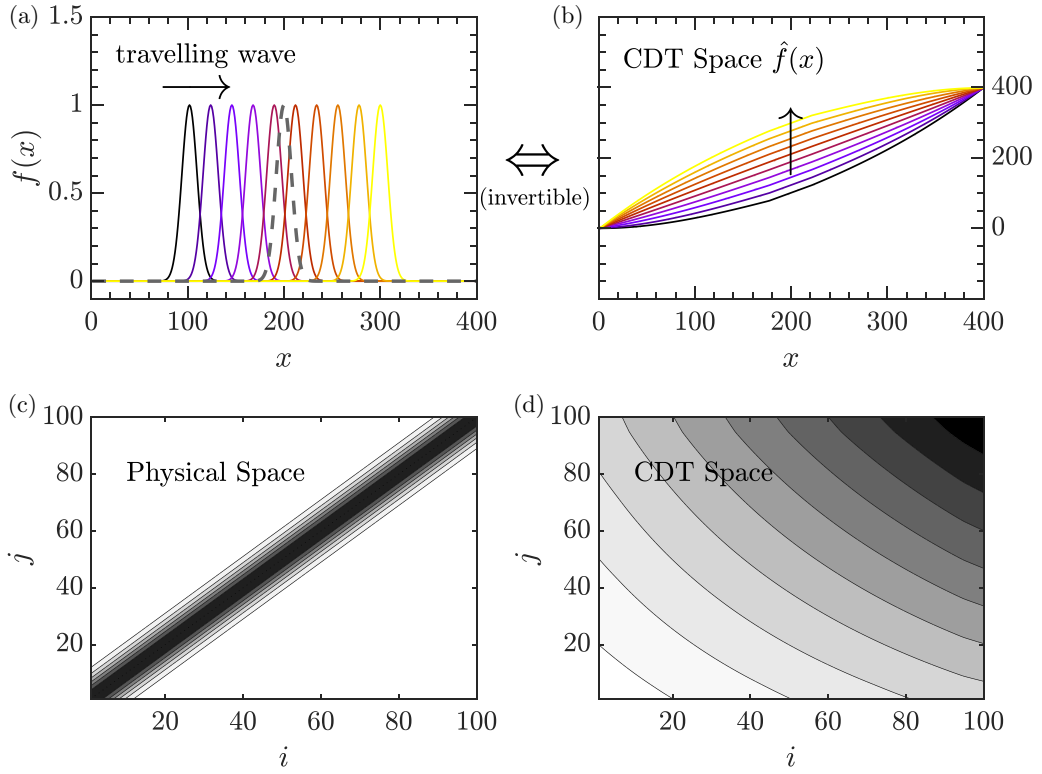

FIG. 1. Traveling Gaussian soliton $f(x,t)$ (a) and its CDT transform $\hat{f}(x,t)$ (b). Solid lines are colored by time traveled, while the dashed line indicates the template. The pseudocolor plot of the correlation matrix $\boldsymbol{R}_{ij} = \langle f(x,t_i), f(x,t_j)\rangle$ in physical (c) and CDT (d) space.

For higher-dimensional problems, we apply the Radon transform to the signal before CDT. The Radon transform (see [15] for a comprehensive introduction) finds its roots in multidisciplinary sciences, including medical imaging, geophysics, material science, optics, etc. The inverse transform (iRadon) represents the mathematical framework of reconstruction from probe-measured (e.g., by x rays or microwaves) distributions. The Radon transform integrates an image or volume f along different angles ($0 \leqslant \theta, \phi < \pi$) leading to a number of projected profiles $\tilde{f}$, defined as

$$\begin{cases} \text{2D:} & \tilde{f}(s,\theta) = \iint_{-\infty}^{+\infty} f(x,y)\delta(s - x\cos\theta - y\sin\theta)\,dx\,dy, \\ \text{3D:} & \tilde{f}(s,\theta,\phi) = \iiint_{-\infty}^{+\infty} f(x,y,z)\delta(s - x\sin\theta\cos\phi - y\sin\theta\sin\phi - z\cos\theta)\,dx\,dy\,dz, \end{cases} \tag{5}$$

and δ is the Dirac delta function. The Radon-CDT (R-CDT) transform $\hat{f}$ [16,17] is then obtained by applying CDT along each projection angle:

$$\int_{s_1}^{\hat{f}(s,\theta)} \tilde{f}(s',\theta)\,ds' = \int_{s_1}^{s} \tilde{r}(s',\theta)\,ds', \quad \int_{s_1}^{\hat{f}(s,\theta,\phi)} \tilde{f}(s',\theta,\phi)\,ds' = \int_{s_1}^{s} \tilde{r}(s',\theta,\phi)\,ds', \tag{6}$$

where $\tilde{r}$ is the Radon transform of the template $r(x,y;z)$. Below we show the properties of R-CDT for 2D problems, while a similar relation holds for the 3D framework. Following the definition (5), the Radon transform of a traveling signal $g(x,y) = f(\alpha x - b, \alpha y - c)$ is

$$\tilde{g}(s,\theta) = \tilde{f}(\alpha s - \alpha b\cos\theta - \alpha c\sin\theta, \theta). \tag{7}$$

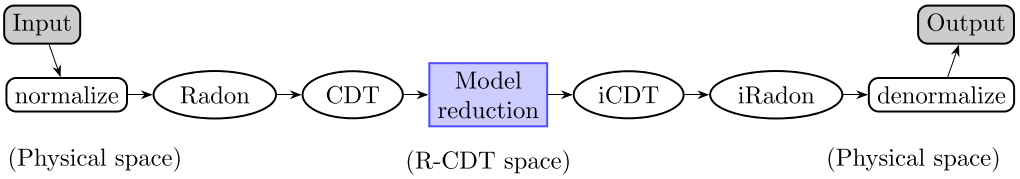

FIG. 2. Flow chart of the proposed method.

This implies that the 2D traveling $\boldsymbol{v} = (b, c)$ and scaling α in physical space becomes a 1D traveling of $\alpha b \cos\theta + \alpha c \sin\theta$ and scaling of α in the Radon space. From the composition properties of CDT, the following relation thus holds in the R-CDT space:

$$\hat{g}(s,\theta) = \frac{1}{\alpha^2}\hat{f}(s,\theta) + \frac{b}{\alpha}\cos\theta + \frac{c}{\alpha}\sin\theta. \tag{8}$$

In contrast to the appearance in physical space, $g(x, y) = f(\alpha x - b, \alpha y - c)$, the traveling and scaling component turns into amplitude modulations in the R-CDT space. Moreover, the application of 1D OT combined with Radon transform has been successful in solving multidimensional optimal transport problems with applications in image processing [18] and model reduction of BGK equations [19] to name a few.

Figure 2 provides a work-flow summary of the proposed method. In examples presented in this article, we build the reduced-order model (ROM) in the R-CDT space by applying a standard POD. As both CDT and Radon transforms are invertible, the inverse transform (iRadon and iCDT) and denormalization are then performed to reconstruct the physical space. When the data type is of a vectorial nature, the reduced order models are built in the R-CDT space of each component respectively.

III. EXAMPLES

A. The 1-D problem: Viscous Burgers' equation

The viscous Burgers' equation is given by

$$u_t + uu_x = \nu u_{xx}. \tag{9}$$

As a model equation mimicking the formation of shock waves, the solution takes the form of diffusive waves that travel with speeds proportional to the wave height. Similar to the case described in Mendible *et al.* [4], we consider $\nu = 0.1$, a Gaussian soliton, $u(x, t = 0) = \exp[-(x + 2)^2]$ is prescribed as an initial condition, and we limit the domain to $-8 \leqslant x \leqslant 8$ with travel-in and -out boundary conditions. The solution is obtained for $t \in [0, 40]$ with a size of 268 grid points $\times$ 201 time steps.

Figure 3(a) presents the evolution of the profile. The initial Gaussian soliton travels downstream with decaying amplitude. By the end of the time limit, part of the wave traveled out of the domain. Upon comparing the data [Fig. 3(b)] and its two-mode ROM [Fig. 3(c)] in the x-t frame, it is evident that the ROM considerably distorted the data. On the other hand, the data in the CDT space [Fig. 3(d)] can be methodically reduced to two modes [Fig. 3(e)]. The iCDT finally recovers the reduced data in the physical space [Fig. 3(f)]. There is hardly a discernible difference between the reconstructed data and its origin.

To quantify the performance of the ROM in CDT space, Fig. 4 plots the decay of singular values σ (obtained from POD and normalized by the first value σ_1) and the error ϵ as functions of the number of modes. The singular value drop is much faster in the CDT space, implying reduced Kolmogorov n-width. The error has been defined as the root-mean-square level between

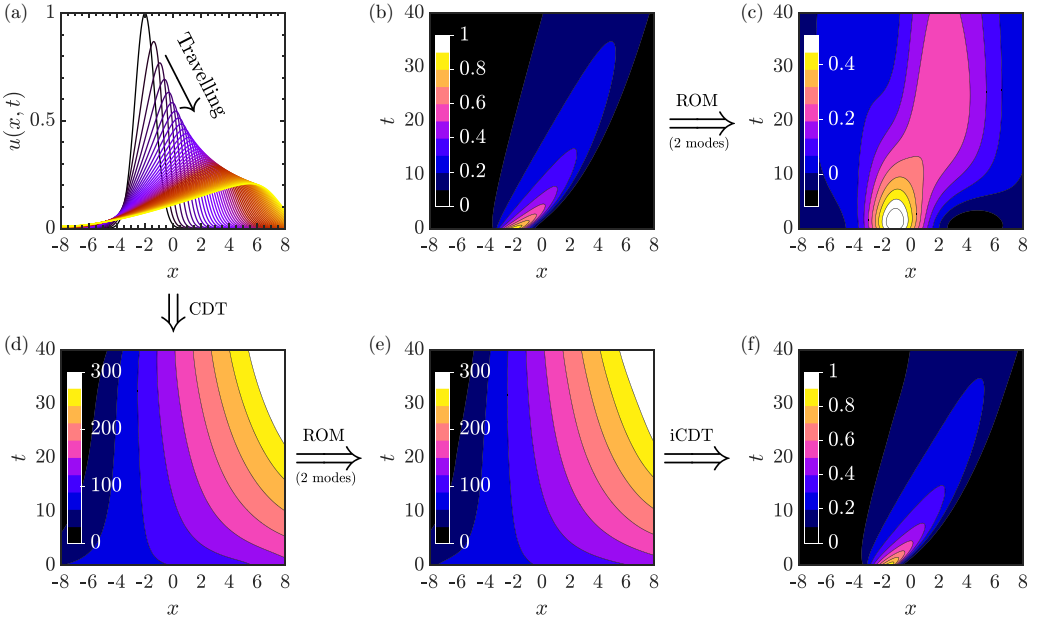

FIG. 3. Model reduction of the viscous Burgers' equation. The evolution of the wave profiles is visualized in panel (a), where the lines are colored by the time traveled. We show the original signal and its two-mode ROM in the x-t frame in panels (b) and (c). With the proposed method, the physical signal is transformed into the CDT space (d), where the two-mode ROM is built (e), and this is followed by an iCDT to recover the data in physical space (f).

the reconstructed data u and ground-truth value u_0:

$$\epsilon = \sqrt{\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}|u(x_i,t_j)-u_0(x_i,t_j)|^2}. \tag{10}$$

Here N and M are the numbers of grid points and time steps. Figure 4(b) indicates that the error is essentially reduced (up to 88% less), in particular in the regimes where fewer modes are kept.

B. Higher-dimensional problems: Traveling Gaussian solitons

Here we show the application of R-CDT space for higher-dimensional problems. The tested problems are the 2D and 3D traveling Gaussian solitons, given by

$$u(x,y,z,t) = \exp\{-a_x[x-\xi_x(t)]^2 - a_y[y-\xi_y(t)]^2 - a_z[z-\xi_z(t)]^2\}. \tag{11}$$

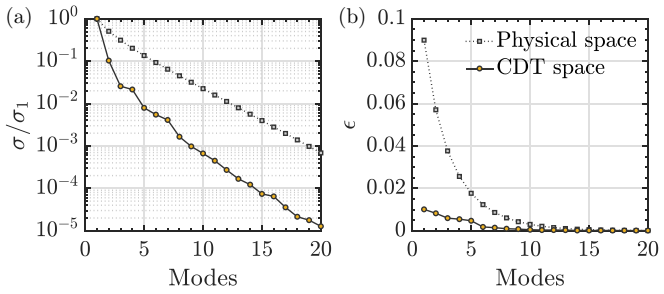

FIG. 4. Comparison of the singular value decay (a) and error of reduced order models (b) in the physical and CDT space.

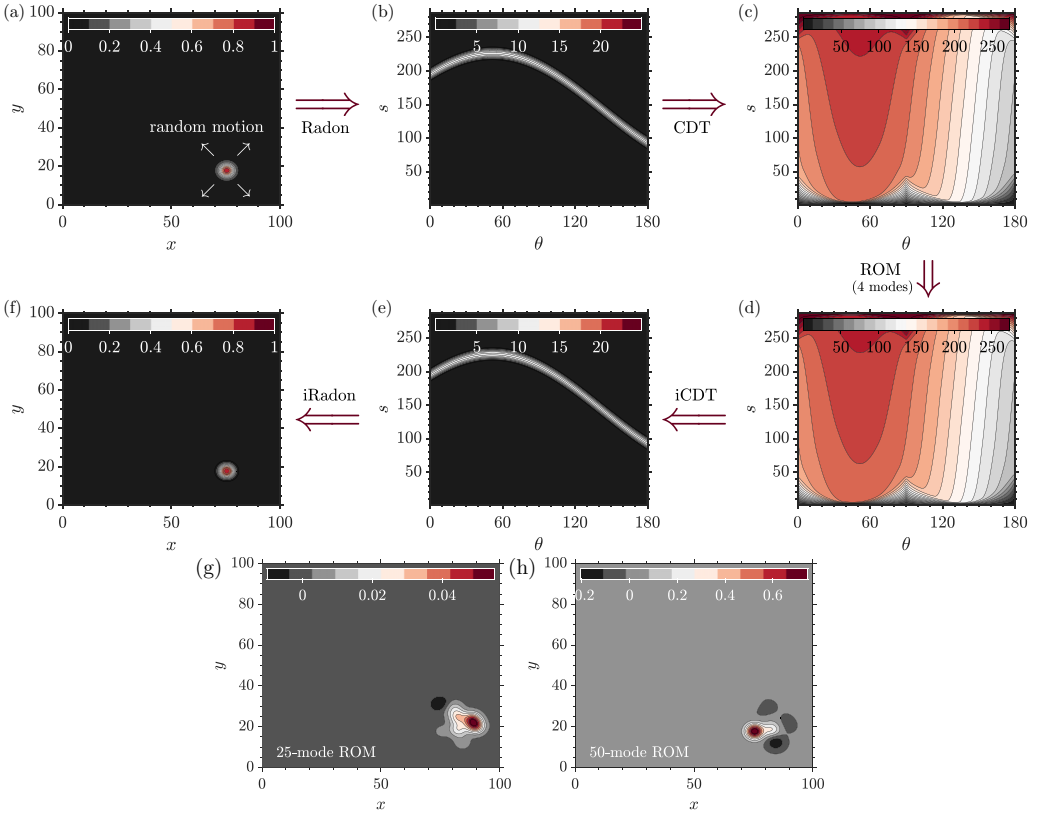

FIG. 5. The procedure to build a reduced-order model via R-CDT. The original signal (a) is transformed into the R-CDT space by sequentially applying the Radon transform (b) and CDT (c). The four-mode ROM is built in the R-CDT space (d); this is followed by iCDT (e) and iRadon to recover the physical space (f). For comparison, the 25-mode and 50-mode ROMs built in the physical space are shown in panels (g) and (h). This figure shows a snapshot corresponding to $t = 10$; a video covering the whole time horizon is available in the Supplemental Material [20].

The parameters a_x, a_y, a_z determine the size of the soliton along each coordinate (x, y, z). The time functions $(\xi_x(t), \xi_y(t), \xi_z(t))$ control the center of the soliton at time t. In the examples below, we have specified $a_x = a_y = -0.08$ for 2D and $a_x = a_y = a_z = -0.05$ for 3D problems.

In the 2D example, we demonstrate the method for random traveling with a nondifferentiable trajectory. The domain is $(x, y) = [0, 100] \times [0, 100]$. The time functions $\xi_x(t), \xi_y(t), \xi_z(t)$ are randomly generated with values between 10 and 90. We show the instantaneous signal in Fig. 5(a); the R-CDT space is obtained by sequentially performing the Radon [Fig. 5(b)] and CDT [Fig. 5(c)] transformation. We show the four-mode ROM in Fig. 5(d), iCDT in Fig. 5(e), and iRadon transform in Fig. 5(f) recovering the physical space. To compare, we show the 25-mode [Fig. 5(g)] and 50-mode ROMs [Fig. 5(h)] built in the physical space. As can be seen, the random motions are hardly modeled in the physical space even with 50 modes. We show the singular value decay and the error in Fig. 6. For random motion problems, as can be seen here, the singular value drops very slowly in the physical space. The error reduces almost linearly by increasing the number of modes. Also, the proposed R-CDT space essentially outperforms the physical space for 3D problems. Figure 7(a) visualizes a Gaussian soliton traveling uniformly in a 3D field. The singular values in R-CDT space reach a much faster drop [Fig. 7(b)], leading to smaller errors of ROM.

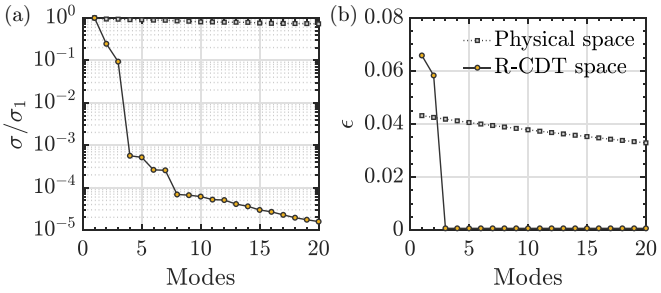

FIG. 6. Comparison of the singular value decay (a) and error of reduced order models (b) in the physical and R-CDT space.

The reconstructed data from the standard four-mode ROM [Fig. 7(c)] are compared with their counterpart in R-CDT space [Fig. 7(d)].

IV. DISCUSSION AND CONCLUDING REMARKS

In this Letter, we have proposed the Radon cumulative distribution transform (R-CDT) space for the model reduction of traveling waves up to three-dimensional problems. As opposed to physical space, the Kolmogorov n-width becomes sharply reduced in the R-CDT space. Accordingly, a more precise but smaller reduced-order model becomes feasible. The strength of the method lies in its parameter-free and data-driven nature. Neither recognition of the physical process (e.g.,

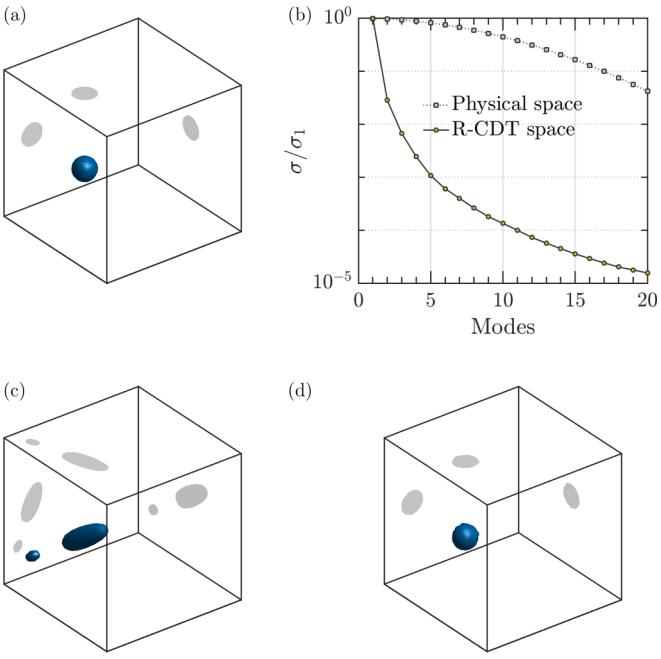

FIG. 7. The 3D traveling Gaussian soliton visualized with an isosurface of $f(x, y, z, t) = 0.05$ (a). Comparison of the singular value drop in the physical and R-CDT space (b). The reconstructed data from four-mode ROM built in the physical (c) and R-CDT (d) spaces, respectively. Shadows in panels (a), (c), and (d) indicate the projection of isosurfaces onto three coordinate planes. A movie showing the temporal evolution is available in the Supplemental Material [20].

detect traveling waves and obtain their velocities) nor data training as usually required by machine learning is necessary. There is no limitation on the boundary condition since a spatial/temporal shift is not performed. The method is particularly outstanding for traveling-dominated signals even with random trajectories. The weakness may emanate from the accuracy of the inverse transforms (iRadon and iCDT). When noise or more complex signals dominate, the error produced by a reduced-order model (ROM) in the R-CDT space may get amplified during the iCDT and iRadon process. This can be overcome by improving the algorithms of inverse transforms or combining more advanced techniques building the ROM. For a more complex data set, e.g., coupled with turbulence, future extension of the method can focus on separating the traveling component and building reduced-order models, respectively.

The R-CDT space shares a broad junction to the state-of-art model reduction techniques. In this work, we have shown that even standard POD provides very efficient models in the R-CDT space. We hope this work creates further grounds for the model reduction of traveling-wave problems.

ACKNOWLEDGMENTS

This investigation is funded by the European Unions Horizon 2020 future and emerging technologies program with agreement No. 828799. Calculations were performed on HPC-Midlands funded by the Engineering and Physical Sciences Research Council (grant number EP/K000063/1). J.R. acknowledges support from the Alexander von Humboldt Foundation. The authors would like to thank Prof. Gustavo K. Rohde and Dr. Hamidreza Eivazi for useful discussions.

APPENDIX: COMPOSITION PROPERTIES OF CDT

Let $g(x)$ be an invertible, differentiable function, the CDT of $f(g(x))g'(x)$ with respect to $r(x)$ is $\hat{f}_g(x) = g^{-1}(\hat{f}(x))$. The proof of this property follows the definition of CDT. From (2),

$$\int_{x_1}^{\hat{f}(x)} f(u)\,du = \int_{x_1}^{\hat{f}_g(x)} f(g(u))g'(u)\,du = \int_{x_1}^{x} r(u)\,du. \tag{A1}$$

Let $g(u) = v$, then $g'(u)du = dv$, and (A1) becomes

$$\int_{x_1}^{\hat{f}(x)} f(u)\,du = \int_{x_1}^{g(\hat{f}_g(x))} f(v)\,dv = \int_{x_1}^{x} r(u)\,du. \tag{A2}$$

Therefore,

$$\hat{f}(x) = g(\hat{f}_g(x)) \Leftrightarrow \hat{f}_g(x) = g^{-1}(\hat{f}(x)). \tag{A3}$$

From the composition property, let $g(x) = \alpha x - ct$, representing a traveling wave with phase velocity c and scaling factor α. Then the CDT of $f(\alpha x - ct)$ is given by $[\hat{f}(x) + ct]/\alpha$.

[1] C. W. Rowley and S. T. Dawson, Model reduction for flow analysis and control, Annu. Rev. Fluid Mech. **49**, 387 (2017).

[2] S. L. Brunton, J. L. Proctor, and J. N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, Proc. Natl. Acad. Sci. USA **113**, 3932 (2016).

[3] C. W. Rowley and J. E. Marsden, Reconstruction equations and the Karhunen–Loève expansion for systems with symmetry, Physica D **142**, 1 (2000).

[4] A. Mendible, S. L. Brunton, A. Y. Aravkin, W. Lowrie, and J. N. Kutz, Dimensionality reduction and reduced-order modeling for traveling wave physics, Theor. Comput. Fluid Dyn. **34**, 385 (2020).

[5] J. Reiss, P. Schulze, J. Sesterhenn, and V. Mehrmann, The shifted proper orthogonal decomposition: A mode decomposition for multiple transport phenomena, SIAM J. Sci. Comput. **40**, A1322 (2018).

[6] J. Sesterhenn and A. Shahirpour, A characteristic dynamic mode decomposition, Theor. Comput. Fluid Dyn. **33**, 281 (2019).
[7] D. Rim, S. Moe, and R. J. LeVeque, Transport reversal for model reduction of hyperbolic partial differential equations, SIAM/ASA J. Uncertain. Quant. **6**, 118 (2018).
[8] N. Cagniart, Y. Maday, and B. Stamm, Model order reduction for problems with large convection effects, in *Contributions to Partial Differential Equations and Applications*, edited by B. N. Chetverushkin, W. Fitzgibbon, Y. A. Kuznetsov, P. Neittaanmäki, J. Periaux, O. Pironneau (Springer, 2019), pp. 131–150.
[9] N. J. Nair and M. Balajewicz, Transported snapshot model order reduction approach for parametric, steady-state fluid flows containing parameter-dependent shocks, Int. J. Numer. Methods Eng. **117**, 1234 (2019).
[10] A. Iollo and D. Lombardi, Advection modes by optimal mass transfer, Phys. Rev. E **89**, 022923 (2014).
[11] B. Peherstorfer, Model reduction for transport-dominated problems via online adaptive bases and adaptive sampling, SIAM J. Sci. Comput. **42**, A2803 (2020).
[12] K. Lee and K. T. Carlberg, Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders, J. Comput. Phys. **404**, 108973 (2020).
[13] S. R. Park, S. Kolouri, S. Kundu, and G. K. Rohde, The cumulative distribution transform and linear pattern classification, Appl. Comput. Harmonic Anal. **45**, 616 (2018).
[14] G. Monge, Mémoire sur la théorie des déblais et des remblais, Histoire de l'Académie Royale des Sciences de Paris (1781).
[15] S. R. Deans, *The Radon Transform and Some of Its Applications* (Dover Publications, Mineola, NY, 2007).
[16] 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**, 920 (2016).
[17] 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, arXiv:2004.03669 (2020).
[18] N. Bonneel, J. Rabin, G. Peyré, and H. Pfister, Sliced and Radon Wasserstein barycenters of measures, J. Math. Imaging Vision **51**, 22 (2015).
[19] F. Bernard, A. Iollo, and S. Riffaud, Reduced-order model for the BGK equation based on POD and optimal transport, J. Comput. Phys. **373**, 545 (2018).
[20] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevFluids.6.L082501 for the movies of temporal evolution shown in fig. 5 and fig. 7.