Speckle Denoising of Dynamic Contrast-Enhanced Ultrasound Using Low-Rank Tensor Decomposition

Metin Calis, Massimo Mischi, *Senior Member, IEEE*, Alle-Jan van der Veen, *Fellow, IEEE*, and Borbala Hunyadi, *Senior Member, IEEE*

***Abstract*— Dynamic contrast-enhanced ultrasound (DCEUS) is an imaging modality for assessing microvascular perfusion and dispersion kinetics. However, the presence of speckle noise may hamper the quantitative analysis of the contrast kinetics. Common speckle denoising techniques based on low-rank approximations typically model the speckle noise as white Gaussian noise (WGN) after the log transformation and apply matrix-based algorithms. We address the high dimensionality of the 4D DCEUS data and apply low-rank tensor decomposition techniques to denoise speckles. Although there are many tensor decompositions that can describe low rankness, we limit our research to multilinear rank and tubal rank. We introduce a gradient-based extension of the multilinear singular value decomposition to model low multilinear rankness, assuming that the log-transformed speckle noise follows a Fisher-tippet distribution. In addition, we apply an algorithm based on tensor singular value decomposition to model low tubal rankness, assuming that the log-transformed speckle noise is WGN with sparse outliers. The effectiveness of the methods is evaluated through simulations and phantom studies. Additionally, the tensor-based algorithms' real-world performance is assessed using DCEUS prostate recordings. Comparative analyses with existing DCEUS denoising literature are conducted, and the algorithms' capabilities are showcased in the context of prostate cancer classification. The addition of Fisher-tippet distribution did not improve the results of tr-MLSVD in the in vivo case. However, most cancer markers are better distinguishable when using a tensor denoising technique than state-of-the-art approaches.**

***Index Terms*— Dynamic contrast-enhanced ultrasound, low-rank tensor decomposition, multilinear singular value decomposition, prostate cancer, speckle denoising.**

Received 18 October 2024; revised 19 February 2025; accepted 11 March 2025. Date of publication 14 March 2025; date of current version 2 July 2025. This work was supported in part by Angiogenesis Analytics and in part by Holland High Tech with a Publieke-Privaat Samenwerking [Public-Private Partnership (PPS)] supplement for Research and Development in the Topsector High Tech Systemen en Materialen [High Tech Systems and Material (HTSM)] under Project 20.091. *(Corresponding author: Metin Calis.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by Medical Research Ethics Committee (MEC AMC).

Metin Calis, Alle-Jan van der Veen, and Borbala Hunyadi are with the Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: m.calis@tudelft.nl; a.j.vanderveen@tudelft.nl; b.hunyadi@tudelft.nl).

Massimo Mischi is with the Department of Electrical Engineering, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands (e-mail: m.mischi@tue.nl).

Digital Object Identifier 10.1109/TMI.2025.3551660

I. Introduction

DYNAMIC contrast-enhanced ultrasound (DCEUS) is an ultrasound imaging modality that enables the characterization of the blood perfusion patterns through the microvasculature [1]. DCEUS is investigated for many applications, such as the localization of prostate cancer. With the help of intravenously injected microbubbles, 4D DCEUS enables the visualization and analysis of the entire prostate gland. The newly formed angiogenic vessels associated with the tumor growth create a distinctive and irregular microvascular architecture that is used as an indicator of prostate cancer [2]. Through analysis of the DCEUS loops by convective dispersion modeling, a number of quantitative dispersion parameters that reflect angiogenic changes in the underlying microvascular architecture have been proposed. The time evolution of the microbubbles inside the vasculature, also called the time-intensity curves (TICs), are fit with the local density random walk model, and distinctive perfusion and dispersion dynamics have been observed in malignant and benign regions [3]. Indeed, a significant amount of work has been published that proves the benefit of using contrast-ultrasound dispersion techniques (CUDI) to identify angiogenesis [4], [5], [6], [7].

The general framework for the classification of angiogenesis consists of several steps. Firstly, the microbubbles are administrated intravenously, and the Digital Imaging and Communications in Medicine (DICOM) data is recorded. Commercial ultrasound scanners utilize various algorithms such as harmonic imaging, phase inversion, or amplitude modulation to enhance the non-linear echoes from the microbubbles while suppressing the approximately linear echoes from the tissue, especially at the employed low ultrasound pressure [8]. The log-compressed envelope of the resulting radio-frequency data is exported for further preprocessing, feature extraction, and classification. Even after the aforementioned contrast enhancement techniques, DCEUS recordings suffer from speckle noise. This noise results from the coherent imaging of microbubbles in a resolution cell. Assuming that a large number of randomly distributed scatterers exist in this cell and there are no strong reflectors, the real and the imaginary parts of the complex echo can be modeled by a zero-mean Gaussian density [9], [10], [11]. The magnitude of this echo is Rayleigh distributed, where the mean is proportional to the standard deviation of the complex echo. This creates multiplicative noise, which becomes additive in the log domain after the log compression performed by the scanner. At lower bubble concentrations, noise is more dominated by the pharmacokinetic statistics of the bubbles, i.e., their probability of being in or outside the cell [12].

Many existing features are initially developed and tested in 2D data [13], [14]. Although most of the CUDI features are extended for the three-dimensional space [5], [6] [15], the denoising algorithms do not take the multidimensional structure of the recording into consideration. The two most widely used denoising techniques for DCEUS sequences include a wavelet [11], and an SVD-based [16] algorithm. The authors in [11] model the log-transformed noise as WGN with outliers. They apply the robust smoother-cleaner [17] to the finest level of decomposition using a median filter with a window of five and biorthogonal wavelets with three and nine vanishing moments. The resulting signal is denoised using a regular wavelet shrinkage algorithm with soft thresholding. In [18], a low-pass filter is applied to TICs where the cut-off frequency is set to be 0.5 Hz. In [16], the authors investigated various matrix decomposition techniques and found that the truncated singular value decomposition (SVD) gives the best results for the localization of prostate cancer. The assumption is that distinct columns of the resulting factor matrices will capture the signal and noise subspace. Denoising is achieved by reconstructing the data after setting the singular values of the noise subspace to zero. The DCEUS recording is flattened into a spatiotemporal matrix where the columns correspond to the time, and the rows correspond to the space to apply SVD. However, this flattening removes spatial information and destroys the 3D structure. Therefore, previously in [19], we applied the multilinear singular value decomposition to the DCEUS recording and truncated the factor matrices at each unfolding by ranks estimated by a robust information-theoretic method [20]. When the dispersion and perfusion features were used, a slightly better separation between the malignant and benign regions was observed compared to [16]. We further showed in [21] an improved classification performance in a multi-parametric setting. The multilinear singular value decomposition aims to decompose the tensor into its orthonormal basis through least squares. This type of decomposition is suitable for WGN noise [22].

We implement and compare various low-rank tensor-based denoising techniques that incorporate the Fisher-tippet assumption on the speckle noise. First, we propose a tensor-based denoising algorithm that uses the prior noise distribution to estimate a low-rank tensor via gradient descent by utilizing the general estimation framework (GTE) [23]. The difference between the proposed algorithm and [23] is twofold. The derivatives are calculated more efficiently, enabling their application for large recordings, and the noise is distributed by Fisher-tippet. In [19], we investigated several low-rank tensor denoising techniques for denoising Fisher-tippet noise through simulation and found out that orientation invariant tensor nuclear norm [24] performs well, especially for low SNR scenarios. Hence, we apply the orientation invariant tensor nuclear norm (OITNN) algorithm to model the log-transformed DCEUS recordings as a low-rank tensor with sparse and WGN, which has not been done before. The performance of the aforementioned tensor techniques will be compared with the denoising techniques proposed in the literature, and the performance of prostate cancer localization will be reported. We incorporate a linear classifier for each feature and report the area under the ROC curve.

In the literature, the low-rank despeckling techniques are commonly applied to the nonlocal patches extracted from the ultrasound recording [25], [26], [27]. The DCEUS recordings are different than the fundamental mode ultrasound images. Rather than imaging a static morphology, in DCEUS, moving bubbles inside the vasculature are imaged. Therefore, we take a different approach motivated by [16] where SVD is applied to the spatiotemporal matrix generated by flattening the whole DCEUS recording. We are motivated to model low rankness in spatial dimensions separately for several reasons. The voxels around a cancerous region show similar dispersion and perfusion characteristics [5], [28]. Significantly grown cancerous tumors will create spatial regions that can captured by the low-rank approximation. In addition, we postulate that the orthogonality assumption in the factor matrices will aid the separation between the tissue and the microvasculature, as well as the cancerous TICs and the benign TICs.

The paper is organized as follows. Section II describes the notation and a brief introduction to the tensor algebra and the decompositions. Section III formulates the problem and describes the tensor-based denoising algorithms. Section IV introduces the simulation, in vitro, and in vivo setups. Section V reports the simulation and in vitro results and the discriminative power of the CUDI features extracted from the DCEUS recordings of the prostate. We elaborate on the results in Section VI and conclude with possible future work in Section VII.

II. Notation

Tensors are represented by underlined boldface letters such as $\underline{\mathbf{Y}}$. Matrices are represented by boldface letters such as $\mathbf{U}^{(1)}$ and $\mathbf{I}$. The numbers given as superscripts in parentheses are used to refer to the different matrices that share a similar property. For example, the three-factor matrices of the multilinear singular value decomposition for a third-order tensor are denoted by $\mathbf{U}^{(1)}$, $\mathbf{U}^{(2)}$, and $\mathbf{U}^{(3)}$ [22]. Vectors are represented by lowercase boldface letters such as $\mathbf{b}$. Scalars are represented by lower case letters such as a_{ij} that represent element at the ith row and jth column of $\mathbf{A} \in \mathbb{R}^{I\times J}$. The MATLAB notation is used to describe the slicing of a tensor; for example, the first R_1, R_2, R_3 and R_4 elements of the tensor $\underline{\mathbf{S}}$ are $\underline{\mathbf{S}}(1 : R_1, 1 : R_2, 1 : R_3, 1 : R_4)$. The Hadamard product is shown with $\odot$. The Frobenius norm is the square root of the sum of each element and is shown by $\|\cdot\|_F$. The spectral norm of a matrix is shown with $\|\cdot\|$. The number of iterations is shown with superscript lowercase letters such as $\underline{\mathbf{L}}^k$ or $\mathbf{U}^{(1)^k}$. The initialization of an iteration is shown with $\underline{\mathbf{L}}^0$. The estimated variables are shown with a hat sign such as $\hat{\underline{\mathbf{L}}}$.

A. Tensor Notation and Preliminaries

The mode-n unfolding of $\underline{\mathbf{Y}} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$ is $\mathbf{Y}_{(n)} \in \mathbb{R}^{I_n \times I_1 \ldots I_{n-1} I_{n+1} \ldots I_N}$. The mode-n product of tensor $\underline{\mathbf{Y}} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$ and a matrix $\mathbf{U}^{(n)} \in \mathbb{R}^{J \times I_n}$ yields a tensor $\underline{\mathbf{Y}} \times_n \mathbf{U}^{(n)} = \underline{\mathbf{C}} \in \mathbb{R}^{I_1 \times \ldots \times I_{n-1} \times J \times I_{n+1} \times \ldots I_N}$ with entries $c_{i_1 \ldots i_{n-1} j i_{n+1} \ldots i_N} = \sum_{i_n=1}^{I_n} a_{i_1 \ldots . i_n \ldots i_N} b_{j i_n}$. Mode-(n,n+1) unfolding of a tensor is described further in the paper and shown with $\mathbf{Y}_{[n]} \in \mathbb{R}^{I_n \times D/(I_n I_{n+1}) \times I_{n+1}}$, where $D = \prod_{n=1}^{N} I_n$ is the number of elements in the tensor $\underline{\mathbf{Y}}$. A diagram representation of the tensors and matrices is used. Tensors are represented by circles with at least three lines, and matrices are represented by circles with two lines. The lines represent the different dimensions, and the sizes of the dimensions are written on the lines. For more information regarding tensor notations and decompositions, we refer to [29].

III. Low-Rank Approximation of DCEUS Sequences

A. Signal Model

The DCEUS recordings are four-dimensional recordings where the first three dimensions represent the spatial domain, and the last dimension represents the time domain. Due to the difference between the resolution cell and the size of the microbubbles, the recordings suffer from the speckle noise. Let $\underline{\tilde{\mathbf{M}}} \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4}$ represent the speckle noise where each entry is Rayleigh distributed with a scaling parameter of 1. The DCEUS recording, represented by $\underline{\tilde{\mathbf{Y}}} \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4}$, is the multiplication of the parameter of interest $\underline{\tilde{\mathbf{L}}} \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4}$ with $\underline{\tilde{\mathbf{M}}}$. This can be shown as

$$\underline{\tilde{\mathbf{Y}}} = \underline{\tilde{\mathbf{L}}} \odot \underline{\tilde{\mathbf{M}}}. \tag{1}$$

It has been shown in [4] and [11] that there is a linear relation between the microbubble concentration and $\underline{\tilde{\mathbf{L}}}$. Hence, the magnitude of each entry at $\underline{\tilde{\mathbf{L}}}$ can be used as an indirect measure of the bubble concentration.

Commonly, in ultrasound devices, logarithmic compression is applied for visualization purposes. This operation is given by

$$c \ln \underline{\tilde{\mathbf{Y}}}, \tag{2}$$

where

$$c = 255 \log_{10}(e) 10/DR \tag{3}$$

is a function of the dynamic range (DR). This operation affects the probability density function of the noise. We have

$$c \ln \underline{\tilde{\mathbf{Y}}} = c \ln \underline{\tilde{\mathbf{L}}} + c \ln \underline{\tilde{\mathbf{M}}}. \tag{4}$$

Let $c \ln \underline{\tilde{\mathbf{M}}} = \underline{\mathbf{M}}$, $c \ln \underline{\tilde{\mathbf{L}}} = \underline{\mathbf{L}}$ and $c \ln \underline{\tilde{\mathbf{Y}}} = \underline{\mathbf{Y}}$. With this notation, we will have

$$\underline{\mathbf{Y}} = \underline{\mathbf{L}} + \underline{\mathbf{M}}. \tag{5}$$

The log-transformed Rayleigh noise follows the Fisher-tippet distribution [9]. The goal of this paper is to recover the DCEUS recording $\underline{\mathbf{L}}$ from the signal model given in (5), assuming that the recording shows a low-rank structure. We introduce two algorithms and compare them with the literature on DCEUS denoising. These algorithms differ in the assumption of the noise and the low-rank structure. These algorithms are

- general tensor estimation framework (GTE) [23] that assumes $\underline{\mathbf{M}}$ follows Fisher-tippet distribution and $\underline{\mathbf{L}}$ is low rank in mode-n unfoldings,
- orientation invariant tubal nuclear norm (OITNN) [24] that assumes $\underline{\mathbf{L}}$ is low tubal rank in mode-(n,n+1) unfoldings and models the noise $\underline{\mathbf{M}}$ as the summation of a WGN and sparse outliers.

The truncated multilinear singular value decomposition [19], [21] is used as a warm initialization for the GTE algorithm. For completeness, we briefly introduce tr-MLSVD.

B. Tr-MLSVD

In this section, we estimate denoised recording $\hat{\underline{\mathbf{L}}}$ by the application of the MLSVD on $\underline{\mathbf{Y}}$, estimating the ranks in each mode, and finally truncating the core tensor and the factor matrices according to the estimated ranks and reconstructing the tensor.

1) MLSVD: We apply MLSVD to the preprocessed $\underline{\mathbf{Y}}$ to obtain

$$\underline{\mathbf{Y}} = \underline{\mathbf{S}} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)} \times_4 \mathbf{U}^{(4)}, \tag{6}$$

where $\underline{\mathbf{S}} \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4}$ is the core tensor and $\mathbf{U}^{(n)} \in \mathbb{R}^{I_n \times I_n}$ for $n \in \{1, \ldots, 4\}$ are the factor matrices. The MLSVD is the application of four SVDs to all the unfoldings $\mathbf{Y}_{(n)}$ for $n \in \{1, \ldots, 4\}$ and assigning the factor matrices $\mathbf{U}^{(n)}$ for $n \in \{1, \ldots, 4\}$ as the left singular vectors. In addition, the all-orthogonal core tensor $\underline{\mathbf{S}}$ can be found by multiplying the $\underline{\mathbf{Y}}$ with the transpose of the factor matrices in the corresponding mode. We estimate the tensor $\underline{\mathbf{L}}$ related to the microbubble movement by truncating the core tensor and the factor matrices according to the estimated ranks.

2) Rank Estimation: The mode-n singular values $\boldsymbol{\phi}^{(n)}$ are defined as the squared sum of the columns of the mode-n unfolding of $\underline{\mathbf{S}}$, that is,

$$\boldsymbol{\phi}^{(n)} = \frac{1}{(D/I_n)} diag(\mathbf{S}_{(n)}^T \mathbf{S}_{(n)}), \tag{7}$$

where D/I_n represents the multiplication of the sizes of all dimensions except the nth dimension. The most contributing columns are selected for the minimum description length estimation [30]. In each mode, the sparse representation is executed by the parameter ρ, which is suggested to be between 0.0001 and 0.01. Let $\mathbf{P}^{(n)} \in \mathbb{R}^{D/I_n \times D/I_n}$ denote the matrix that selects the $\rho D/I_n$ columns of $\mathbf{S}_{(n)}$ with the highest norm, while discarding the rest. We have the robust eigenvalues $\mathbf{s}^{(n)}$ at the diagonal of $\mathbf{S}_{(n)} \mathbf{P}^{(n)} \mathbf{P}^{(n)T} \mathbf{S}_{(n)}^T$, that is,

$$\mathbf{s}^{(n)} = \frac{1}{(\rho D/I_n)} diag(\mathbf{S}_{(n)} \mathbf{P}^{(n)} \mathbf{P}^{(n)T} \mathbf{S}_{(n)}^T).$$

The rank $\hat{R}_n$ for each mode $n \in \{1, \ldots, 4\}$ are estimated using the MDL criterion,

$$\hat{R}_n = \underset{r}{\operatorname{argmin}} - 2 \log \left(\frac{\prod_{i=r+1}^{I_n} (s_i^{(n)})^{1/(I_n - r)}}{\frac{1}{I_n - r} \sum_{i=r+1}^{I_n} s_i^{(n)}} \right)^{\rho (D/I_n)(I_n - r)} + r(2I_n - r) \log(\rho(D/I_n)). \tag{8}$$

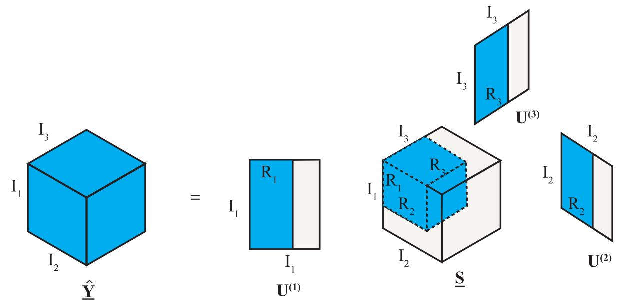

Fig. 1. The truncated multilinear singular value decomposition of a 3D tensor with ranks (R_1, R_2, R_3).

3) Truncation: After estimating the ranks, we reconstruct the tensor by truncating the core tensor and the factor matrices according to the estimated ranks. This is done by taking the first $\hat{R}_n$ columns of $\mathbf{U}^{(n)}$ for $n \in \{1, \ldots, 4\}$, i.e., $\hat{\mathbf{U}}n = \mathbf{U}^{(n)}(:, 1 : \hat{R}_n)$ and taking the corresponding elements of the core tensor, i.e., $\hat{\underline{\mathbf{S}}} = \underline{\mathbf{S}}(1 : \hat{R}_1, 1 : \hat{R}_2, 1 : \hat{R}_3, 1 : \hat{R}_4)$. The denoised DCEUS recording is given by

$$\hat{\underline{\mathbf{L}}} = \hat{\underline{\mathbf{S}}} \times_1 \hat{\mathbf{U}}1 \times_2 \hat{\mathbf{U}}2 \times_3 \hat{\mathbf{U}}3 \times_4 \hat{\mathbf{U}}4, \tag{9}$$

This algorithm is described in Algorithm 1. Previously, in [19] and [21], we implemented this algorithm and found improved performance in the classification of prostate cancer. In this paper, we extend the analysis to a higher number of patients and compare it with the results from the literature and the methods proposed in this paper.

Algorithm 1 The Tr-MLSVD Framework for Denoising DCEUS Sequences

Input : $\underline{\mathbf{Y}}$ ρ,
MLSVD : $\underline{\mathbf{Y}} \leftarrow \underline{\mathbf{S}} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)} \times_4 \mathbf{U}^{(4)}$
Rank Estimation :$\{\hat{R}_1, \hat{R}_2, \hat{R}_3, \hat{R}_4\}$ using (8)
Truncation : $\hat{\underline{\mathbf{S}}} \leftarrow \underline{\mathbf{S}}(1 : \hat{R}_1, 1 : \hat{R}_2, 1 : \hat{R}_3, 1 : \hat{R}_4)$
$\hat{\mathbf{U}}n \leftarrow \mathbf{U}^{(n)}(:, 1 : \hat{R}_n)$ for $n \in \{1, \ldots, 4\}$
Return :$\hat{\underline{\mathbf{L}}} \leftarrow \hat{\underline{\mathbf{S}}} \times_1 \hat{\mathbf{U}}1 \times_2 \hat{\mathbf{U}}2 \times_3 \hat{\mathbf{U}}3 \times_4 \hat{\mathbf{U}}4$

C. GTE

The authors in [23] proposed a low-rank tensor estimation framework that incorporates the prior distribution of the noise. They solve

$$\hat{\underline{\mathbf{L}}} = \underset{\underline{\mathbf{L}}\ is\ low\ rank}{\operatorname{argmin}} \; -\ln(p(\underline{\mathbf{Y}}; \underline{\mathbf{L}})) \tag{10}$$

where $p(\underline{\mathbf{Y}}; \underline{\mathbf{L}})$ is the probability distribution function of 4D DCEUS recording $\underline{\mathbf{Y}}$ parameterized by the low multilinear rank tensor $\underline{\mathbf{L}}$. The paper solves the problem for Poisson, Gaussian, and Binomial noise. The difference between [23] and the proposed method is the characteristics of the noise and the calculation of derivatives. We will extend the analysis to the Fisher-tippet noise in Section III-C1. In Section III-C2, we will calculate the derivatives of the loss function without the Kronecker products, which are shown in Appendix B and C.

1) Probability Distribution of the Fisher-Tippet Noise: Let the probability density distribution of $\underline{\mathbf{M}}$ be the Rayleigh distribution with a scaling parameter of 1, that is,

$$p(\underline{\mathbf{M}}; 1) = \prod_{i_1=1}^{I_1} \prod_{i_2=1}^{I_2} \prod_{i_3=1}^{I_3} \prod_{i_4=1}^{I_4} m_{i_1 i_2 i_3 i_4} \exp \frac{-m_{i_1 i_2 i_3 i_4}^2}{2}, \tag{11}$$

assuming that the noise is independent between voxels. The log compression, as described in (4), changes the probability density function of the noise. The Fisher-tippet distribution [9] with the log compression is

$$p(\underline{\mathbf{Y}}; \underline{\mathbf{L}}) = \prod_{i_1=1}^{I_1} \prod_{i_2=1}^{I_2} \prod_{i_3=1}^{I_3} \prod_{i_4=1}^{I_4} \frac{\exp(2(y-l)/c - \frac{\exp 2(y-l)/c}{2})}{c}. \tag{12}$$

Note that the indices $y_{i_1 i_2 i_3 i_4}$ and $l_{i_1 i_2 i_3 i_4}$ in (12) are dropped for notational convenience. The argument that minimizes the negative log-likelihood of (12) is the maximum likelihood estimate of $\underline{\mathbf{L}}$ with Fisher-tippet noise.

2) Gradient Descent: In addition to the maximum likelihood, the authors in [23] proposed to add the redundant term

$$\frac{a}{2} \sum_{n=1}^{4} \| (\mathbf{U}^{(n)})^T \mathbf{U}^{(n)} - b^2 \mathbf{I} \|_F^2, \tag{13}$$

to prevent the factor matrices from being singular throughout the gradient descent. Let the loss function be denoted by F. With the addition of these terms, the loss function F becomes

$$F(\underline{\mathbf{Y}}, \underline{\mathbf{L}}) = -\ln p(\underline{\mathbf{Y}}; \underline{\mathbf{L}}) + \frac{a}{2} \sum_{n=1}^{4} \| (\mathbf{U}^{(n)})^T \mathbf{U}^{(n)} - b^2 \mathbf{I} \|_F^2 . \tag{14}$$

The two regularization weights a and b given in (14) are selected using the spectral norm of the initial estimate $\underline{\mathbf{L}}^0$. The initialization, along with the estimation of the ranks, is done with the MLSVD-based algorithm from [19]. Four spectral norms can be defined using the four unfoldings of $\underline{\mathbf{L}}^0$. Let q denote the maximum spectral norm, that is,

$$q = \max(\| \mathbf{L}_{(1)}^0 \|, \| \mathbf{L}_{(2)}^0 \|, \| \mathbf{L}_{(3)}^0 \|, \| \mathbf{L}_{(4)}^0 \|). \tag{15}$$

Two regularization weights are assigned as $b = q^{1/4}$ and $a = q$. The initial factor matrices $\{{\mathbf{U}^{(1)}}^0 \ldots {\mathbf{U}^{(4)}}^0\}$ are multiplied by $q^{1/4}$, whereas the core tensor $\underline{\mathbf{S}}^0$ is divided by q. This is done to guarantee the local convergence with a high probability [23].

If we incorporate the probability distribution into (14), we have

$$\begin{aligned} F(\underline{\mathbf{Y}}, \underline{\mathbf{L}}) &= -\ln p(\underline{\mathbf{Y}}; \underline{\mathbf{L}}) + \frac{a}{2} \sum_{n=1}^{4} \| (\mathbf{U}^{(n)})^T \mathbf{U}^{(n)} - b^2 \mathbf{I} \|_F^2 \\ &= \sum_{i_1}^{I_1} \sum_{i_2}^{I_2} \sum_{i_3}^{I_3} \sum_{i_4}^{I_4} \ln c - 2\frac{y-l}{c} + \frac{\exp(2(y-l)/c)}{2} \\ &\quad + \frac{a}{2} \sum_{n=1}^{4} \| (\mathbf{U}^{(n)})^T \mathbf{U}^{(n)} - b^2 \mathbf{I} \|_F^2 . \end{aligned} \tag{16}$$

Define the element-wise derivative operator as $\frac{\partial F}{\partial \underline{\mathbf{L}}}$: $\mathbb{R}^{I_1\times I_2\times I_3\times I_4} \to \mathbb{R}^{I_1\times I_2\times I_3\times I_4}$. The derivative of F with respect to $\mathbf{U}^{(1)}$ follows the chain rule

$$\frac{\partial F}{\partial \mathbf{U}^{(1)}} = \frac{\partial F}{\partial \underline{\mathbf{L}}}\frac{\partial \underline{\mathbf{L}}}{\partial \mathbf{U}^{(1)}} \tag{17}$$

We can write

$$\frac{\partial F}{\partial \underline{\mathbf{L}}} = \frac{2-\exp\left(2(\underline{\mathbf{Y}}-\underline{\mathbf{L}})/c\right)}{c}. \tag{18}$$

We modified the derivatives $\frac{\partial \underline{\mathbf{L}}}{\partial \mathbf{U}^{(1)}}$ given in [23] as described in Appendix B. With this modification, the derivatives can be calculated for large tensors using less random access memory, which is further explained in Section VI.

An early stopping condition is defined as the relative change of the $\underline{\mathbf{L}}$ at each iteration to the change of the first iteration, that is,

$$\Delta\,\underline{\mathbf{L}}^k < \epsilon = \frac{\|\,\underline{\mathbf{L}}^k - \underline{\mathbf{L}}^{k-1}\,\|_F}{\|\,\underline{\mathbf{L}}^1 - \underline{\mathbf{L}}^0\,\|_F}, \tag{19}$$

where k is the iteration number, and the iterations stop when $\Delta\,\underline{\mathbf{L}}^k$ is smaller than ϵ or the maximum number of iterations K is reached. The pseudo-code for the algorithm is defined in Algorithm 2.

Algorithm 2 The GTE Framework for Estimating a Low-Rank Tensor With Fisher-Tippet Distribution

Input : $\underline{\mathbf{Y}}$, $\mathrm{E}\{\underline{\mathbf{M}}\}$, η, K, ϵ
Debias : $\underline{\mathbf{Y}} \leftarrow \underline{\mathbf{Y}} - \mathrm{E}\{\underline{\mathbf{M}}\}$
Initialization : $\underline{\mathbf{L}} \leftarrow \underline{\mathbf{S}} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)} \times_4 \mathbf{U}^{(4)}$ ▷ Using Algorithm 1
$\mathbf{U}^{(n)^0} \leftarrow q^{1/4}\,\mathbf{U}^{(n)}$ for $n \in \{1\ldots4\}$
$\underline{\mathbf{S}}^0 \leftarrow \underline{\mathbf{S}}/q$ see (15)
while $k < K$ or $\Delta\,\underline{\mathbf{L}}^k > \epsilon$ **do**
  $k \leftarrow k+1$
  $\mathbf{U}^{(n)^k} \leftarrow \mathbf{U}^{(n)^{k-1}} - \eta\frac{\partial F}{\partial \mathbf{U}^{(n)^{k-1}}}$ for $n \in \{1\ldots4\}$ see (34).
  $\underline{\mathbf{S}}^k \leftarrow \underline{\mathbf{S}}^{k-1} - \eta\frac{\partial F}{\partial \underline{\mathbf{S}}^{k-1}}$ see (35)
  $\underline{\mathbf{L}}^k \leftarrow \underline{\mathbf{S}}^{k-1} \times_1 \mathbf{U}^{(1)^{k-1}} \times_2 \mathbf{U}^{(2)^{k-1}} \times_3 \mathbf{U}^{(3)^{k-1}} \times_4 \mathbf{U}^{(4)^{k-1}}$
  if $k > 1$ **then**
    $\Delta\,\underline{\mathbf{L}}^k \leftarrow \frac{\|\underline{\mathbf{L}}^k - \underline{\mathbf{L}}^{k-1}\|_F}{\|\underline{\mathbf{L}}^1 - \underline{\mathbf{L}}^0\|_F}$ see (19)
  end if
  if $\Delta\,\underline{\mathbf{L}}^k > \Delta\,\underline{\mathbf{L}}^{k-1}$ **and** $k > 2$ **then**
    $\eta \leftarrow \eta/10$
    $k \leftarrow 0$
  end if
end while
Return : $\underline{\mathbf{L}}^k$

D. OITNN

In this section, we describe the OITNN algorithm for denoising DCEUS recordings. OITNN considers low rankness in mode-(n,n+1) unfoldings, and it is based on a framework called t-SVD [31]. We will first describe the basics of this framework.

Definition 1: The t-product: Given two 3D tensors $\underline{\mathbf{T}} \in \mathbb{R}^{I\times J\times I_3}$ and $\underline{\mathbf{H}} \in \mathbb{R}^{J\times M\times I_3}$, the t-product $\underline{\mathbf{G}} = \underline{\mathbf{T}} \star \underline{\mathbf{H}} \in \mathbb{R}^{I\times M\times I_3}$ is computed by taking the 1D Fourier transform of each tensor in the last axis, matrix multiplication of each frontal slice, and returning the 1D inverse Fourier transform in the last axis. The t-product is shown with the $\star$ sign.

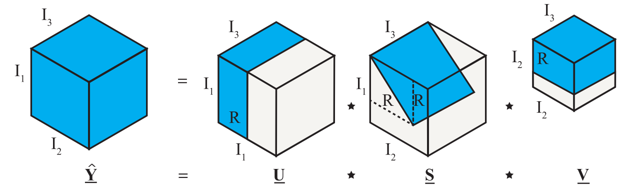

Fig. 2. The illustration of a low tubal rank t-SVD decomposition defined in Section III-D with rank R.

Definition 2: t-SVD: The t-SVD of the tensor $\underline{\mathbf{L}} \in \mathbb{R}^{I_1\times I_2\times I_3}$ is defined as $\underline{\mathbf{L}} = \underline{\mathbf{U}} \star \underline{\lambda} \star \underline{\mathbf{V}}^T$ with orthogonal $\underline{\mathbf{U}} \in \mathbb{R}^{I_1\times I_1\times I_3}$ and $\underline{\mathbf{V}} \in \mathbb{R}^{I_2\times I_2\times I_3}$. The real and positive valued $\underline{\lambda} \in \mathbb{R}^{I_1\times I_2\times I_3}$ has diagonal frontal slices. More information is provided regarding the tensor identity and orthogonality in [31]. The number of non-zero vectors $\underline{\lambda}(i_1, i_2, :)$ is called the tubal rank.

Definition 3: Tensor nuclear norm (TNN): Given $\underline{\mathbf{L}} = \underline{\mathbf{U}} \star \underline{\lambda} \star \underline{\mathbf{V}}^T$ the tubal nuclear norm is given by the average of the sum of the frontal slices of the core tensor, that is,

$$\|\,\underline{\mathbf{L}}\,\|_{\mathrm{TNN}} = \frac{1}{I_3}\sum_{i_3=1}^{I_3}\mathrm{Tr}(\underline{\lambda}(:,:,i_3)). \tag{20}$$

The authors in [24] extended the analysis to higher dimensions and solved the robust tensor denoising problem using invariant tubal nuclear norm (OITNN). Tensors that have dimensions greater than three are unfolded into 3D tensors using mode-(n,n+1) unfolding, which is defined as the following.

Definition 4: Mode-(n,n+1) unfolding: The mode-(n,n+1) unfolding of $\underline{\mathbf{L}} \in \mathbb{R}^{I_1\times I_2\times I_3\times I_4}$, creates a 3D tensor $\underline{\mathbf{L}}_{(n,n+1)} \in \mathbb{R}^{I_n\times D/I_nI_{n+1}\times I_{n+1}}$ by permuting the nth dimension of $\underline{\mathbf{L}}$ to the first, n+1th dimension to the last and grouping the rest. Here D is defined as $\prod_{n=1}^{N} I_n$. This unfolding is simply shown as $\underline{\mathbf{L}}_{[n]}$.

The OITNN is defined using the mode-(n,n+1) operation as

$$\|\,\underline{\mathbf{L}}\,\|_{\mathrm{OITNN}} = \frac{1}{4}\sum_{n=1}^{4}\|\,\underline{\mathbf{L}}_{[n]}\,\|_{\mathrm{TNN}}. \tag{21}$$

We refer back to the original problem formulation given in (5). The authors in [9] approximate the Fisher-tippet noise $\underline{\mathbf{M}}$ as the summation of WGN $\underline{\mathbf{W}} \in \mathbb{R}^{I_1\times\cdots\times I_4}$ and sparse outliers $\underline{\mathbf{O}} \in \mathbb{R}^{I_1\times\cdots\times I_4}$, that is,

$$\underline{\mathbf{Y}} = \underline{\mathbf{L}} + \underline{\mathbf{W}} + \underline{\mathbf{O}}\,. \tag{22}$$

The OITNN considers the low rankness in all orientations mode-(n,n+1) for $n \in \{1,2,3,4\}$ and solves the optimization problem that is defined as

$$\min_{\underline{\mathbf{L}},\underline{\mathbf{O}}}\left\{\frac{1}{2}\|\,\underline{\mathbf{L}}+\underline{\mathbf{O}}-\underline{\mathbf{Y}}\,\|_F^2 + \gamma_L\|\,\underline{\mathbf{L}}\|_{\mathrm{OITNN}} + \gamma_O\|\,\underline{\mathbf{O}}\,\|_1\right\}$$
$$\text{s.t. } \|\,\underline{\mathbf{L}}\,\|_\infty \le \alpha. \tag{23}$$

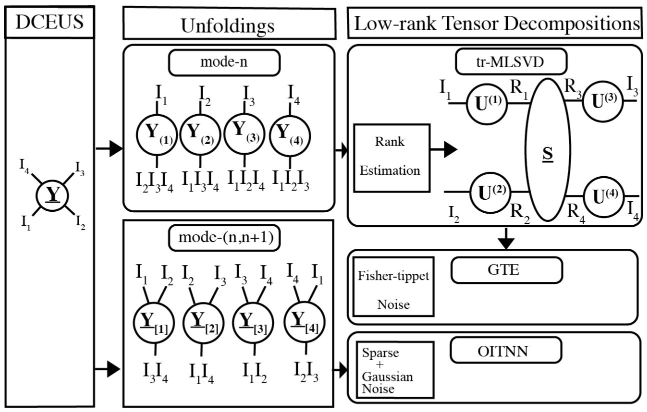

Fig. 3. Overview of the tensor-based DCEUS denoising methods.

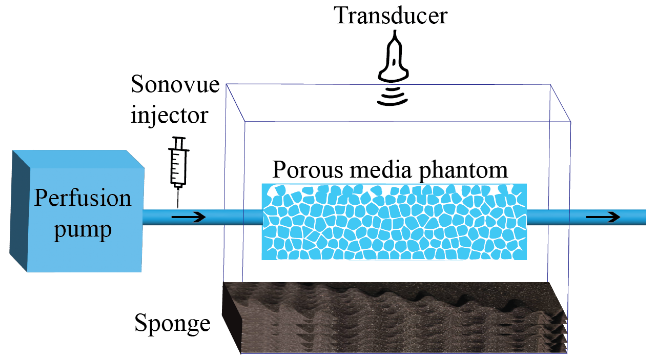

Fig. 4. Phantom setup.

Similar to GTE, the DCEUS recording is first preprocessed by subtracting the median of the first twelve seconds from the signal. The infinity norm α in (23) is assigned by median filtering preprocessed $\underline{\mathbf{Y}}$ with a window of 5 seconds and taking the maximum. This algorithm is solved by using the algorithm in [24] until the stopping convergence given in (19) or the maximum iteration number K is reached. An overview of the methods is shown in Fig. 3.

IV. Validation Methology

In this section, we introduce the setup for simulation, in vitro, and in vivo studies and report the results in Section V. We first compare the denoising performance of tr-MLSVD, GTE, and OITNN using synthetic data. Previously, in [19], we compared the performance of SVD and tr-MLSVD methods through simulation, where tr-MLSVD was found to perform better. For that reason, the SVD method is omitted from the simulation. Additionally, wavelet-based denoising [10] is omitted from the simulation because the assumption that the time evolutions are smooth does not hold for the general case. For this reason, we only compare the tensor-based denoising techniques. Following the simulations, we will conduct an in vitro study, report the model-fitting performance, and visualize the phantom recordings after denoising. Finally, we will compare the single-feature classification performance of the tensor-based methods with the state-of-the-art DCEUS denoising techniques.

A. Simulation

We generated three 4-dimensional tensors $\underline{\mathbf{L}} \in \mathbb{R}^{20\times 20\times 20\times 20}$ with three different ranks. First the core tensor was generated $\underline{\mathbf{S}} \in \mathbb{R}^{R_1\times R_2\times R_3\times R_4}$ from the normal distribution with sizes $[6, 6, 6, 6]$, $[4, 8, 12, 16]$ and $[12, 12, 12, 12]$. The core matrix was multiplied in each mode with the orthonormal matrices to get $\tilde{\underline{\mathbf{L}}} = \underline{\mathbf{S}} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)} \times_4 \mathbf{U}^{(4)}$. The orthonormal matrices $\mathbf{U}^{(1)} \in \mathbb{R}^{20\times R_1}$, $\mathbf{U}^{(2)} \in \mathbb{R}^{20\times R_2}$, $\mathbf{U}^{(3)} \in \mathbb{R}^{20\times R_3}$ and $\mathbf{U}^{(4)} \in \mathbb{R}^{20\times R_4}$ were generated according to the Haar measure as described in [32]. We selected such ranks to cover three cases that might occur in actual DCEUS recordings. An actual DCEUS recording might have low or high multilinear ranks in all dimensions, or it can have a lower rank in some of the dimensions than others. Since we expect the recording to be low-rank, we selected a value close to half of the tensor size for high-rank cases.

The $\tilde{\underline{\mathbf{L}}}$ values were scaled to the range $[10, 255]$. The noisy tensor $\tilde{\underline{\mathbf{Y}}}$ was obtained by multiplying $\tilde{\underline{\mathbf{L}}}$ with Rayleigh-shaped $\tilde{\underline{\mathbf{M}}}$ with scaling parameter 1 as shown in (11). The logarithmically transformed tensor $\underline{\mathbf{L}}$ was obtained using (4) with a dynamic range of 42. The denoising algorithm given in Algorithm 2 was applied to $\underline{\mathbf{L}}$. The theoretical noise statistics were taken from [33] and assigned as $\mathbf{E}[\underline{\mathbf{M}}] = 0.0579c$, where c was given in (3). The preprocessing step was skipped. The ranks were assumed to be known. The step size was $\eta = 10^{-7}$. The total number of iterations was defined as $K = 10^5$. The iterations stopped with the condition $\Delta \underline{\mathbf{L}}^k < \epsilon = 0.1$. The true ranks were used for tr-MLSVD and GTE. For OITNN, four values for γ_L and γ_O in the range $(1000, 4000)$ and $(10, 40)$, respectively, have been traversed. Furthermore, the true infinity norm of the original tensor is used. The minimum normalized mean error is reported.

B. In Vitro

We conducted a phantom study using the LOGIQ E10 scanner equipped with a RIC5-9-D endocavity transducer driven at 3.5 Mhz, a porous medium, and a sponge to prevent reverberations. The setup can be seen in Fig. 4. The porous media phantom was built by packing alginate beads of size 2.5 mm in a polyurethane tube with a diameter of 20 mm. The obtained alginate beads of the same size were packed into a polyurethane tube, and the cylinder shape was fixed with two circular nets on two sides of the phantoms. After that, we gently squeezed and shook the phantom to ensure a more homogeneous packing structure. The length of the phantoms was about 43 mm. The water pump is set to a flow rate of 0.22 mL/s. We mixed 1 mL of Sonovue with 100 mL of water and injected it into the tube before imaging. The model is fitted to each voxel using (25). The MFR-RMSE is calculated by taking the root mean square of the difference between the fitted and filtered TICs and multiplying it with a sigmoid-shaped weighting function that exponentially penalizes the error starting from 20 seconds after the peak time.

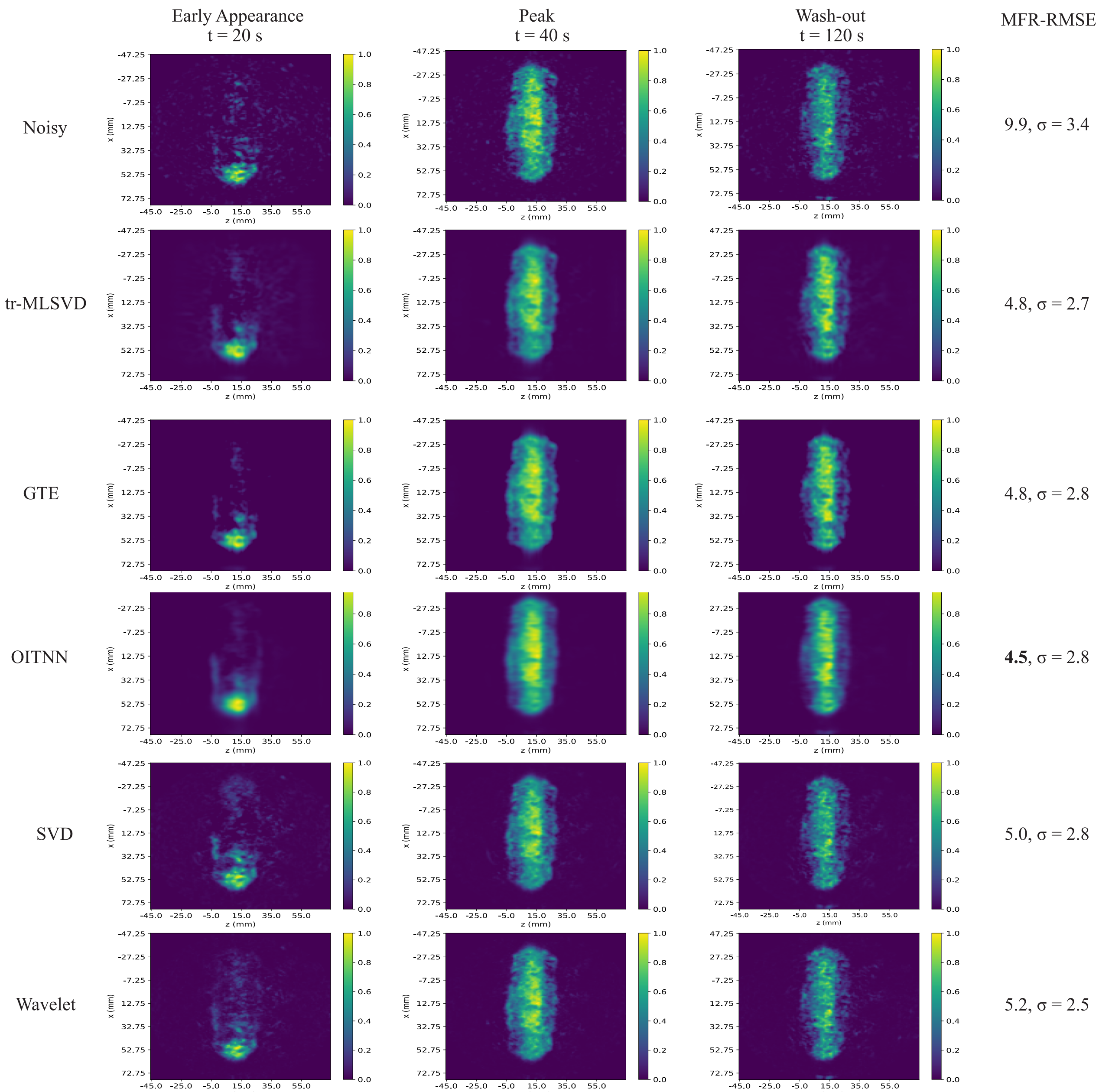

Fig. 5. The normalized amplitudes of noisy and denoised phantom recordings during the early appearance, peak, and wash-out times. The xz slice at the middle of the phantom is plotted.

C. In Vivo

Recordings of 32 patients from the Amsterdam University Medical Center and 62 patients from the Netherlands Cancer Institute were obtained. This study obtained IRB approval, and the patients provided written consent to be enrolled in the study. A 4D recording in contrast mode was obtained with the LOGIQ E10 scanner equipped with an RIC5-9-D endocavity transducer driven at 3.5 Mhz. The volume rate was fixed to 0.9 Hz by setting the image quality to BQMid1, and a low mechanical index of 0.1 was employed to minimize the bubble destruction. The patients went through radical prostatectomy after the recording because of biopsy-proven prostate cancer. The prostate was sliced with 4 mm thickness, and for each slice, an annotation was made by the pathologist. The annotations were registered back to the domain of the recording, and the ground truth was obtained. Significant malignant voxels, at least with a grade of $3 + 4$ are selected [34]. There were approximately seven million benign voxels and two hundred thousand malignant voxels.

The DCEUS recording was transformed from spherical to cartesian domain with a voxel size of 0.25 mm $\times$ 0.25 mm $\times$ 0.25 mm. The spatial resolution was regularized across space through a dedicated Wiener filter [14], and the data was downsampled by 3 such that a voxel size of 0.75 mm was obtained. The warm initialization is obtained using tr-MLSVD. The gain is estimated as the median of the first 12 seconds, which is subsequently subtracted from the TICs of both the tr-MLSVD and the noisy tensor. A step size of

$\eta = 10^{-7}$ was used. The maximum number of iterations was set to $K = 10e5$. The stopping condition was set at $\Delta \underline{\mathbf{L}}^k < \epsilon = 0.1$.

The CUDI features were extracted from each voxel after denoising with either of the five denoising methods, i.e., OITNN, GTE, tr-MLSVD, SVD, and wavelet-based denoising and without denoising. The extracted features were the following:

- model fitting features that quantify TIC perfusion and dispersion based on the local density random walk model denoted as MFR-κ and MFR-μ [35], [36], [37],
- similarity-based metrics, such as correlation (SA-ρ), spectral coherence (SA-r), and mutual information (SA-MI) [3], [5], [14], [28],
- the solution of the convective-dispersion equation [6], [7], denoted by CD-D and CD-v to represent dispersion and velocity, respectively,
- entropy and conditional entropy [38], denoted by VE-Ev and VE-CEv, respectively.

Among all the features, computing the model fitting took the most time. Therefore, the model fitting was approximated using the Exponential Linear Unit (ELU) function, which resulted in faster processing. The approximation is shown in Appendix A. Similarly to the process proposed in the literature [4], a windowing was applied where a higher weight was given to the first pass of the microbubbles. Adam optimizer [39] was used for fitting the model.

V. Results

A. Simulation

The performance metric was the normalized mean error given by

$$\mathrm{NME} = \frac{\|\hat{\underline{\mathbf{L}}} - \underline{\mathbf{L}}\|_F}{\|\underline{\mathbf{L}}\|_F}, \tag{24}$$

where $\hat{\underline{\mathbf{L}}}$ was the estimate of $\underline{\mathbf{L}}$. For both GTE and OITNN, the convergence is achieved with $\Delta \underline{\mathbf{L}}^k < \epsilon = 0.1$ defined in (19). We compared the performance of tr-MLSVD, GTE, and OITNN and reported the results in Fig. 6. We ran pairwise t-tests to compare the denoising performances. Only tr-MLSVD and OITNN for the rank (6,6,6,6) showed insignificant differences with $p < 0.0001$. For rank (6,6,6,6), the GTE gave the best NME. For ranks (4,8,12,16) and (12,12,12,12) OITNN performed better.

B. In Vitro

We applied the tensor-based speckle denoising algorithms to the phantom recordings. The rank estimation parameter for tr-MLSVD is selected to be $\rho = 10e-5$. The sparsity and low rankness related parameters γ_L and γ_O are selected after a 100 values between $(1, 20000)$ and $(1, 100)$, respectively, are swept. The value that gave the least MFR-RMSE is found to be 10000 and 30 for γ_L and γ_S, respectively. The GTE algorithm is run with $\eta = 10e-7$. In Fig. 5, the early appearance, peak, and wash-out times of the phantom recordings are shown on an xz slice at the middle of the phantom, and the averaged MFR-RMSE values are reported. Only the voxels inside the phantom are fitted with the model. The OITNN method resulted in the best fit of the model described in Appendix A. The tr-MLSVD and GTE performed similarly, while a better suppression of the speckle artifacts in the background surrounding the phantom was observed for GTE. Finally, SVD, Wavelet methods showed worse model-fitting performance than the tensor-based counterparts.

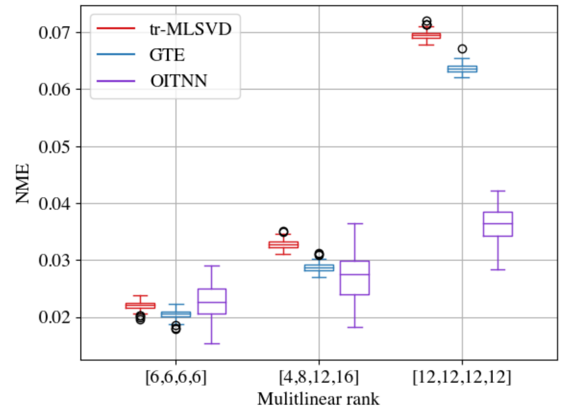

Fig. 6. The normalized mean error for random tensors of size $[20, 20, 20, 20]$ and multilinear rank $[6, 6, 6, 6]$, $[4, 8, 12, 16]$ and $[12, 12, 12, 12]$. The median NME over 10^5 Monte Carlo simulations are shown with a flat line, the box ranges represent the 25th and 75th quantiles, and the whiskers represent the inter-quartile ranges.

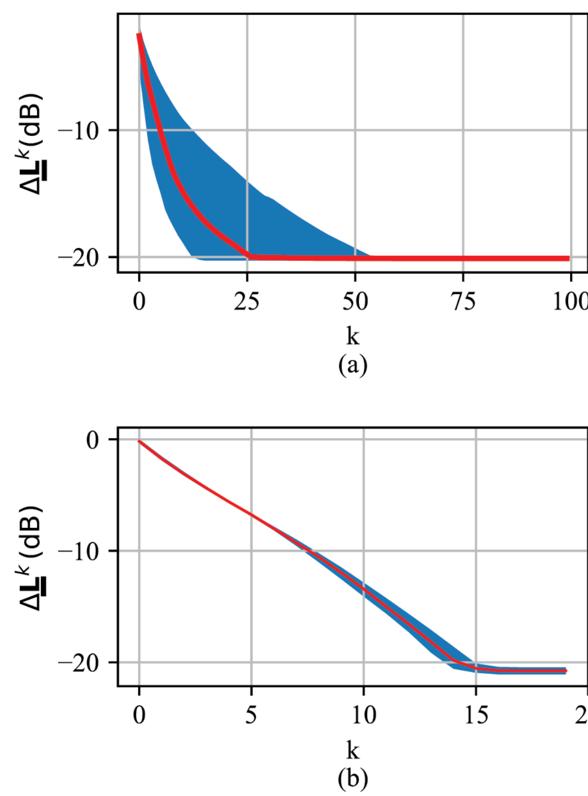

Fig. 7. The median, 25th and 75th percentiles of loss $\Delta \underline{L}^k$ that is defined in (19) calculated for all the patients when the GTE and OITNN algorithm given at Algorithm 2 is run. Subplot (a) represents the GTE, and subplot (b) represents the OITNN. The red line represents the median, and the blue shade represents the percentiles, and the y-axis is shown in dB.

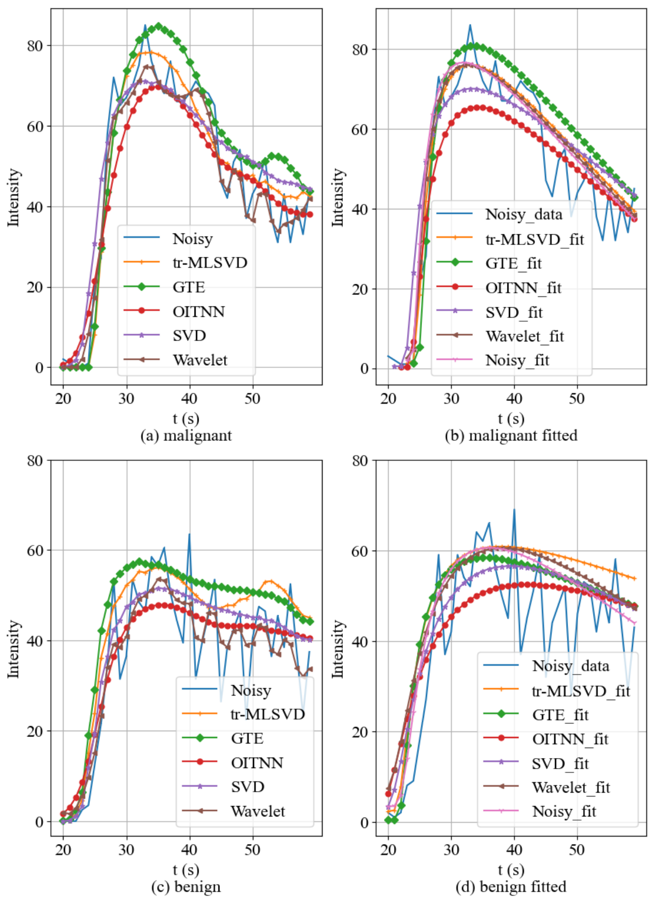

Fig. 8. The denoised signals for a random malignant and benign voxel and the corresponding fitted models. The upper row, defined with labels (a) and (b), represents the malignant voxels. The bottom row, defined with the labels (c) and (d), represents the benign voxels. Only the wash-in period is shown.

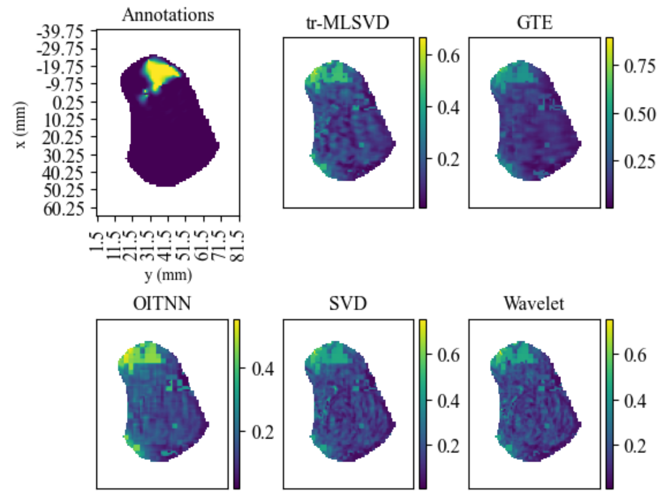

Fig. 9. The κ values for a z slice that has the most malignant voxels for various denoising schemes of a random patient.

TABLE I

THE STATISTICAL SIGNIFICANCE TEST RESULTS WITH ($p < 0.0001$) BETWEEN THE NOISY AND FILTERED CUDI FEATURE DISTRIBUTIONS. ONLY THE ONES THAT LACK STATISTICAL SIGNIFICANCE ARE SHOWN

	Noisy	tr-MLSVD	GTE	OITNN	SVD	Wavelet
Noisy		-	-	-	-	-
tr-MLSVD	-		-	-	-	-
GTE	-	-		-	CD-V, SA-MI	-
OITNN	-	-	MFR-κ		-	-
SVD	-	-	-	-		-
Wavelet	-	-	-	-	-	

We ran pairwise t-tests, and only tr-MLSVD and GTE gave statistically insignificant MFR-RMSE distributions inside the phantom with $p < 0.0001$.

C. In Vivo

We used the same parameters as the in vitro for the tensor-based denoising methods (MLSVD, GTE, OITNN), i.e. the rank selection parameter $\rho = 10e-5$ for tr-MLSVD, and the $\gamma_O = 30$, $\gamma_L = 10e4$ for the OITNN. The features extracted after tensor-based denoising methods were compared against features extracted after matrix-based (SVD) denoising and a state-of-the-art wavelet-based denoising, as well as no denoising. The denoising results are illustrated for an arbitrarily selected malignant and benign voxel in Fig. 8 (a) and (c). In the same figure, the results of the modified linear random walk model fitting described in Appendix A are shown in the subplots (b) and (d). Furthermore, in Fig. 9, we illustrate the model fit parameter MFR-κ on a slice with the highest number of malignant voxels of an arbitrarily selected patient.

The mean and the standard deviation of all features are given in Table II. The mean and the standard deviation of the model fit-related features did not change after the various denoising algorithms. To investigate further, we also compared the goodness of the model fit in terms of the weighted root mean square error of the fit (MFR-RMSE), which is explained in Section IV-B. The mean and standard deviation of MFR-RMSE are shown in the third column of Table II. MFR-RMSE is found to be the lowest for the OITNN method, with an average of 3.84 over all the malignant and benign voxels. The OITNN method is followed by GTE, SVD, MLSVD, and Wavelet methods with average RMSE of 4.11, 4.14, 4.80, and 5.12 respectively. The noisy RMSE fit has an average RMSE of 7.81. Although the RMSE changed, the model fit parameters stayed fairly close to each other, only with an increase in the variance of the benign features in GTE and OITNN. The similarity metrics SA-MI, SA-ρ, and SA-r increased significantly when a low-rank decomposition is applied. We ran pairwise t-tests to analyze the statistics of the CUDI features before and after filtering.

We ran t-tests to assess the two categories: the effect of filtering methods on the distribution of the features, and the difference between malignant and benign features. For the first one, we stacked the benign and malignant features and ran a pair-wise t-test to compare the filtering methods. All the filtering methods had statistically significant differences with $p < 0.0001$ except for two cases: CD-V and SA-MI features for GTE and SVD and MFR-κ feature for GTE and OITNN. This is shown in Table I. For the second assessment, we ran Welch's t-test between the malignant and benign features generated through each filtering method, including the noisy one. Welch's t-test was chosen due to the difference in the number of malignant and benign voxels. In all cases, we can differentiate the benign and malignant voxels with statistical significance ($p < 0.0001$).

TABLE II
THE MEAN AND THE STANDARD DEVIATION OF THE DCEUS FEATURES AFTER THE SIGNAL IS DENOISED BY WAVELET, SVD, MLSVD, GTE ALGORITHMS

		MFR-κ	MFR-μ	MFR-RMSE	CD-D	CD-v	SA-MI	SA-r	SA-ρ	VE-CEv	VE-Ev
Noisy	Malignant	0.15, $\sigma = 0.06$	25.99, $\sigma = 15.51$	9.11, $\sigma = 2.58$	0.05, $\sigma = 0.05$	0.28, $\sigma = 0.20$	1.19, $\sigma = 0.18$	0.45, $\sigma = 0.13$	0.76, $\sigma = 0.14$	2.00, $\sigma = 0.57$	2.92, $\sigma = 0.92$
	Benign	0.13, $\sigma = 0.09$	42.13, $\sigma = 18.83$	7.78, $\sigma = 2.87$	0.01, $\sigma = 0.02$	0.09, $\sigma = 0.11$	0.83, $\sigma = 0.48$	0.31, $\sigma = 0.19$	0.51, $\sigma = 0.30$	1.32, $\sigma = 0.39$	1.85, $\sigma = 0.60$
tr-MLSVD	Malignant	0.15, $\sigma = 0.06$	28.66, $\sigma = 15.42$	5.17, $\sigma = 2.88$	0.07, $\sigma = 0.08$	0.34, $\sigma = 0.26$	1.56, $\sigma = 0.38$	0.80, $\sigma = 0.19$	0.96, $\sigma = 0.06$	2.24, $\sigma = 0.67$	3.27, $\sigma = 1.06$
	Benign	0.13, $\sigma = 0.09$	44.83, $\sigma = 18.77$	4.79, $\sigma = 2.88$	0.02, $\sigma = 0.03$	0.10, $\sigma = 0.13$	0.92, $\sigma = 0.56$	0.58, $\sigma = 0.28$	0.85, $\sigma = 0.18$	1.37, $\sigma = 0.51$	1.93, $\sigma = 0.75$
GTE	Malignant	0.15, $\sigma = 0.07$	29.08, $\sigma = 15.94$	5.00, $\sigma = 4.77$	0.08, $\sigma = 0.10$	0.33, $\sigma = 0.26$	1.69, $\sigma = 0.40$	0.87, $\sigma = 0.14$	0.98, $\sigma = 0.06$	2.22, $\sigma = 0.68$	3.23, $\sigma = 1.07$
	Benign	0.13, $\sigma = 0.32$	45.65, $\sigma = 19.93$	4.09, $\sigma = 2.91$	0.02, $\sigma = 0.04$	0.10, $\sigma = 0.14$	0.99, $\sigma = 0.67$	0.62, $\sigma = 0.35$	0.75, $\sigma = 0.37$	1.38, $\sigma = 0.56$	1.96, $\sigma = 0.83$
OITNN	Malignant	0.15, $\sigma = 0.07$	33.83, $\sigma = 16.02$	4.33, $\sigma = 2.69$	0.08, $\sigma = 0.09$	0.35, $\sigma = 0.29$	1.92, $\sigma = 0.40$	0.91, $\sigma = 0.07$	0.99, $\sigma = 0.02$	2.07, $\sigma = 0.76$	3.08, $\sigma = 1.18$
	Benign	0.13, $\sigma = 0.35$	48.78, $\sigma = 18.86$	3.83, $\sigma = 2.67$	0.02, $\sigma = 0.04$	0.10, $\sigma = 0.14$	1.21, $\sigma = 0.72$	0.73, $\sigma = 0.30$	0.84, $\sigma = 0.27$	1.25, $\sigma = 0.54$	1.80, $\sigma = 0.82$
SVD	Malignant	0.16, $\sigma = 0.07$	30.05, $\sigma = 16.03$	4.55, $\sigma = 2.99$	0.07, $\sigma = 0.08$	0.33, $\sigma = 0.25$	1.65, $\sigma = 0.35$	0.86, $\sigma = 0.14$	0.98, $\sigma = 0.03$	2.25, $\sigma = 0.65$	3.26, $\sigma = 1.04$
	Benign	0.13, $\sigma = 0.09$	45.69, $\sigma = 18.67$	4.13, $\sigma = 3.00$	0.02, $\sigma = 0.03$	0.10, $\sigma = 0.14$	0.99, $\sigma = 0.62$	0.62, $\sigma = 0.32$	0.88, $\sigma = 0.20$	1.40, $\sigma = 0.53$	1.96, $\sigma = 0.77$
Wavelet	Malignant	0.15, $\sigma = 0.07$	31.32, $\sigma = 16.45$	5.77, $\sigma = 2.71$	0.05, $\sigma = 0.06$	0.28, $\sigma = 0.21$	1.33, $\sigma = 0.25$	0.64, $\sigma = 0.18$	0.93, $\sigma = 0.06$	2.04, $\sigma = 0.61$	2.96, $\sigma = 0.97$
	Benign	0.13, $\sigma = 0.09$	46.83, $\sigma = 18.63$	5.10, $\sigma = 2.77$	0.02, $\sigma = 0.03$	0.09, $\sigma = 0.11$	0.89, $\sigma = 0.55$	0.50, $\sigma = 0.29$	0.77, $\sigma = 0.24$	1.32, $\sigma = 0.43$	1.84, $\sigma = 0.65$

TABLE III
THE VOXEL-BASED AREA UNDER THE CURVE IS CALCULATED FOR EACH DENOISING METHOD WHEN A SINGLE FEATURE CLASSIFIER IS USED ACROSS 94 PATIENTS. THE THRESHOLD IS SELECTED SUCH THAT THE HIGHEST AUC IS ACHIEVED

	Noisy	tr-MLSVD	GTE	OITNN	SVD	Wavelet
MFR-κ	0.587	0.638	0.628	**0.649**	0.628	0.617
MFR-μ	0.761	**0.767**	0.765	0.749	0.761	0.752
CD-D	0.771	0.788	0.781	**0.793**	0.780	0.777
CD-v	0.832	**0.837**	0.819	0.826	0.828	0.829
SA-MI	0.728	**0.827**	0.803	0.798	0.824	0.744
SA-r	0.717	0.756	0.765	0.721	**0.779**	0.626
SA-ρ	0.774	0.801	0.783	**0.805**	0.782	0.752
VE-CEv	0.828	**0.840**	0.822	0.797	0.834	0.825
VE-Ev	0.831	**0.842**	0.822	0.803	0.836	0.827

TABLE IV
THE COMPUTATIONAL TIME OF RUNNING THE DENOISING METHODS AND THE MEMORY REQUIRED TO SAVE 94 PATIENT RECORDINGS, ALONG WITH THEORETICAL COMPUTATIONAL COMPLEXITIES. THE TOTAL ITERATIONS OF GTE AND OITNN ARE SET TO K

	Computation Time	Memory	Computational Complexity
tr-MLSVD	8.3 h	840 MB	$\mathrm{O}\Big(D(\sum_{n=1}^{4} I_n)\Big)$
GTE	17 h	840 MB	$\mathrm{O}\Big(KD\max\{I_1, \ldots, I_4\}\Big)$
OITNN	42 h	24.94 GB	$\mathrm{O}\Big(K(D\log(D)+ D(\sum_{n=1}^{4} I_n))\Big)$
SVD	7 h	2.87 GB	$\mathrm{O}\Big(DI_4\Big)$
Wavelet	9.5 h	68 GB	$\mathrm{O}\Big(D\log(I_4)\Big)$

To quantify the discriminative power of the features, the malignant and benign features from all the patients were stacked, and simple single-feature classification was done by varying a threshold over the entire range of each feature. The values that were greater or lower than the threshold were labeled as malignant or benign and the receiver operating characteristics (ROC) curve was generated. The area under the ROC curve (AUC) for the various methods and features is shown in Table III. Five out of nine CUDI features (MFR-μ, CD-v, SA-MI, VE-CEv, VE-EV) showed better separation between malignant and benign voxels for tr-MLSVD. Three out of the nine CUDI features (MFR-κ, CD-D, SA-ρ) had a better classification result for OITNN. Only one feature (SA-r) showed improved performance for SVD.

D. Memory and Computational Requirements

Low-rank decomposition methods enable the representation of any data with fewer parameters. For tr-MLSVD and GTE, compression is achieved by saving the factor matrices $\hat{\mathbf{U}}n$ for $n \in \{1, 2, 3, 4\}$ and the core tensor $\underline{\hat{\mathbf{S}}}$; for SVD, the left and right singular vectors, and the singular values; for OITNN t-compress algorithm from [31]. The t-compress algorithm is applied to mode-(1,2) unfolding. The other possible mode-(n,n+1) unfoldings resulted in similar results. Significantly small singular values and their corresponding factors are discarded, where the tolerance is set to the $1/100$th of the highest singular value. The highest compression is achieved for tr-MLSVD and GTE algorithms. This is followed by SVD and OITNN. Calculation of the features and the denoising is done on a server with an Intel 2×10 Xeon CPU and 256 GB RAM. SVD was the fastest algorithm, with 7 hours of computation time. Consecutively, tr-MLSVD, Wavelet, GTE and OITNN followed it with computation times of 8.3, 9.5, 17, and 42 hours, respectively. The full recordings of 94 patients were 68 GB. After the compression, the reduction is shown in Table IV.

In addition, the computational complexities of the algorithms are presented in Table IV. All algorithms except wavelet include SVD as its foundation. Here, we assume the Golub-Reinsch algorithm is used to calculate the SVD [40]. We point out that with known ranks, iterative algorithms such as [41] will have a reduced complexity. We consider the worst-case scenario, where the ranks are as high as the sizes of the corresponding mode. In [40], the complexity of SVD for an $I \times J$ matrix is described as $O(I^2 J)$ given $I < J$. Therefore, the SVD of the spatiotemporal matrix with size $I_4 \times I_1 I_2 I_3$ has the complexity $O(I_1 I_2 I_3 I_4^2) = O(DI_4)$. For tr-MLSVD, the SVD of the tensor $\underline{\mathbf{Y}}$ in the unfoldings dominates the complexity and results in a complexity of $\mathrm{O}(D(\sum_{n=1}^{4} I_n))$. For GTE, the algorithm is initialized by the tr-MLSVD output. The complexity per iteration is due to the tensor-matrix products that generate the low-rank tensor $\underline{\mathbf{L}}^k$ given in Algorithm 2. Each mode-n tensor-matrix multiplication is a matrix multiplication of sizes $D/I_n \times I_n$, and $I_n \times I_n$, which results in a complexity of $D\max\{I_1, \ldots, I_4\}$. For OITNN, the Fourier transform, inverse Fourier transform, and the SVDs of four unfoldings $\underline{\mathbf{Y}}_{[n]}$ dominate the complexity of each iteration. The Fourier transform of unfoldings $\underline{\mathbf{Y}}_{[n]}$ for $n \in \{1, \ldots, 4\}$ results in the complexity of $\mathrm{O}(D\log(D))$, and their SVD results in the complexity of $\mathrm{O}(D(\sum_{n=1}^{4} I_n))$. Finally, the wavelet decomposition has a complexity of $\mathrm{O}(D\log(I_4)$ due to the 1D Fourier transform of D/I_4 voxels with I_4 samples.

Avoiding the Kronecker products (described in Section III-C2) made it possible for the GTE method to be run with less random access memory (RAM). Assume the calculation of the derivative with respect to the first-factor matrix $\frac{\partial F}{\partial \mathbf{U}^{(1)}}$. Calculation of the Kronecker product given in (29) creates $(\mathbf{U}^{(4)} \otimes \mathbf{U}^{(3)} \otimes \mathbf{U}^{(2)}) \in \mathbb{R}^{I_2 I_3 I_4 \times R_2 R_3 R_4}$. For DCEUS recordings of size $100 \times 80 \times 100 \times 120$, and ranks of $30 \times 20 \times 30 \times 8$, the resulting matrix is of size 960000×4800, creating 34.3 GB of data. On the other hand, consecutive mode-n multiplication in each mode, that is, $\left(\frac{\partial F}{\partial \underline{\mathbf{L}}} \times_2 (\mathbf{U}^{(2)})^T \times_3 (\mathbf{U}^{(3)})^T \times_4 (\mathbf{U}^{(4)})^T\right.$ reduces the dimension to $100 \times 20 \times 30 \times 8$. The mode-1 unfolding of the resulting tensor is of size 100×4800, creating 0.00384 GB of data.

VI. Discussion

We introduced two tensor-based algorithms for denoising DCEUS data that incorporate two approaches to model the Fisher-tippet speckle noise. The GTE algorithm models low rankness in the mode-n unfoldings and uses the log-likelihood to reduce noise. On the other hand, the OITNN algorithm considers low rankness in the mode-(n,n+1) unfoldings and models the speckle noise as WGN and sparse outliers. In this section, we discuss the results and propose research directions for the future.

Inspired by the success of the application of SVD [16] for denoising DCEUS data, we extended the idea of low-rank approximation for denoising DCEUS data to multi-dimensions. Previously we introduced tr-MLSVD algorithm in [19] and [21] and showed an improved performance of prostate cancer classification. The tr-MLSVD algorithm is more suitable for denoising WGN. This introduces artifacts that can be seen in the phantom slice depicted at the early appearance time of tr-MLSVD in Fig. 5. The speckles around the phantom at tr-MLSVD and the noisy recording do not exist in GTE and OITNN. We believe this can aid the visualization of bubble movement. Furthermore, in the simulation given in Fig. 6, we observed that OITNN and GTE perform better than tr-MLSVD for denoising speckle noise in nearly all cases. This justifies the benefit of incorporating the prior distribution and the assumption regarding the orthogonality of the factor matrices.

On the other hand, improved performance was not observed for the in vivo classification. We made several assumptions that might affect the performance. Noise is assumed to be independent between voxels. We deconvolved the recording with a Wiener filter such that a resolution of 0.8 mm is obtained as proposed in [5]. We sampled the recording with a voxel size of 0.75 mm. Since each voxel is comparable to the resolution of the system, we assumed that the assumption of independence holds. There can be other factors, such as the movement of the probe, that can violate the independence.

A reason for the inferior performance of the GTE algorithm could be the mismatch of the assumed noise characteristics. The DCEUS recordings suffer from various noise types such as clutter, shadowing, ring-comet, reflection, refraction, and reverberation artifacts [42]. The GTE algorithm aimed to remove the speckle noise with the assumption of the Rayleigh noise. At lower bubble concentrations, this assumption is known to be invalid, as reported in [12]. Furthermore, distributions such as Gamma [10] or Nakagami [43] have been reported to be a better choice for describing the speckle statistics. Incorporating such a model could improve the results. On the other hand, OITNN assumes that the Fisher-tippet noise can be modeled as WGN with sparse outliers. The aforementioned artifacts might violate this assumption. However, the two regularization parameters γ_L and γ_O allow a direction of future work where each parameter can be tweaked for the removal of different artifacts. The reduced performance of the wavelet-based denoising could be due to the reduced temporal frequency. The results in the literature have been reported for 25 frames/s, whereas the 4D DCEUS recordings provide 1 frame/s. The assumption that the noise and the signal have separate subspaces in the wavelet domain might fail for such temporal frequencies.

The new model fitting algorithm described in Appendix A is found to be fairly robust to noise. As supported by Fig. 8 and Table II, the model fitting algorithm gave similar features for all the denoised signals. This is further supported by the close performance of the model fit features in Table III. When the noisy data is used to extract the model fit features, the performance of the voxel-based classifier is found to be similar to the features extracted from the denoised signals. In the literature, various fitting algorithms were compared for model fitting, such as the maximum likelihood and non-linear least squares [12]. In this paper, we used an ADAM optimizer for faster processing with the relaxation proposed in Appendix A. This could explain the closeness between the noisy and the denoised model fitting features. A significant increase in the similarity metrics (SA-ρ, SA-r, SA-MI) can be observed in Table II after the low-rank decompositions. This is expected since the low-rank decompositions explain the data using the components that describe the majority of the variance. Using a few components results in mostly similar TICs. This is especially observed in malignant regions. Finally, only temporal correlation SA-r gave better classification performance compared to tensor approaches. In SVD, low rankness is considered only in the fourth unfolding. In such a spatiotemporal matrix, temporal information gets smoothened out, resulting in similar TIC appearance times across the whole DCEUS recording. Hence, the temporal correlation SA-ρ is performing worse for SVD, and the spectral correlation SA-r, which only considers the frequency content, performs better.

In our study, we have used the same LOGIQ E10 and the probe in both Amsterdam University Medical Center and the Netherlands Cancer Institute. All the machine settings were fixed, and the outcome of the single feature classification was consistent across both datasets. We have not included a study regarding the generalizability of the proposed algorithms in diverse clinical settings. The generalizability of the proposed tensor-based techniques is dependent on several parameters, such as the variability in equipment settings, patient demographics, and disease presentations. We selected the low-rank-related γ_L and the sparsity-related γ_O hyperparameters defined in Section III-D, as well as the rank-related parameter ρ defined in Section III-B using phantom studies.

We recommend a study that explores variability between different machines, which can shed light on good parameter settings for all future users. An analysis of model-fitting performance, such as the one given in Fig. 5, can be conducted. If a phantom study is not available, we recommend an analysis of model fitting or the classification performance on a subset of the patient data. Equipment settings can be incorporated into the classifier so that the classifier can utilize the variability between different devices and settings.

The preprocessing steps, such as subtracting the median of the first few seconds and the varying step size in optimization algorithms, are still applicable to all DCEUS equipment. This is due to the similarity of protocols regarding DCEUS imaging [44]. The gain is increased such that the background noise is observed without microbubbles, and the intravenous injection is done after starting the imaging sequence. Therefore, the first few frames are expected to represent the gain of the system, which can be subsequently subtracted from the TICs. The subtraction of the gain helps the algorithm focus on learning relative changes related to microbubbles rather than an arbitrary baseline. The algorithm becomes more stable, and a faster convergence is achieved. Similarly, the varying step size η given in Algorithm 2 increases the convergence speed and improves the robustness. If the tensor unfoldings are well-conditioned, the algorithms will converge faster with a high step size. On the contrary, if the tensor unfoldings are ill-conditioned, a smaller step size is required to aid the convergence. In this paper, convergence is achieved for all the patient recordings. However, GTE is a non-convex algorithm, and the convergence is dependent on several factors, such as the amplitude of the noise, the condition number of the tensor unfoldings, and the regularity condition [23]. The OITNN is a convex algorithm with theoretical convergence guarantees [24]. We selected the stopping condition of the iterative algorithms as $\epsilon = 0.1$ given in (19). A lower stopping condition will improve the denoising performance while increasing the convergence time.

The patient demographics can affect the performance of the classification. Commonly, the maximum imaging depth is selected on the ultrasound device to cover all possible sizes of prostates. The patient age in our study ranged from 60 to 87. An interesting research direction can be to investigate the effect of age on the extracted DCEUS features. Additionally, the disease presentation might affect the low-rank assumption. Several factors signify prostate cancer's significance: the tumor size, Gleason score (or grade), and extracapsular extension [45]. Tumors require an increased supply of oxygen and nutrients beyond $1 - 2$ mm in diameter [46]. Clinically significant prostate cancer are tumors with at least size 0.5 cm^3 [45] and a grade of $3+4$ or higher. Subsequently, we selected tumor samples with at least a grade of $3 + 4$ Gleason score. We considered regions that are 2 mm in diameter accounting for the system's spatial resolution as described in Section IV-C, since the main focus of this paper was voxel-based classification. We propose a future study to analyze the effect of different Gleason scores and tumor sizes on the classification results. We expect the large regions to be identified more easily using low-rank tensor decomposition methods compared to early-stage tumors. This is due to their relative contribution to explaining the full DCEUS tensor. Regions that are small and have different TICs compared to the majority of the regions are expected to be captured in the smaller singular values. The signal subspace is assumed to be in the highest singular values in all the tensor-based denoising algorithms. Therefore, regions with low spatial structure and low temporal power will be ignored. Experimental validation of this hypothesis was not conducted, and we leave the comparison between the early-stage and significant prostate cancer classification as future work. Such a research direction requires annotations of malignant regions with different Gleason scores, the consideration of tumors with sizes varying between $2 - 25$ mm in diameter, and the inclusion of cases with extracapsular extension of the tumor. A possible research direction is to apply the aforementioned denoising techniques to subsets of the data rather than the full tensor so that the lower-grade or smaller tumor can be identified. Spatially, the input tensor can be divided into blocks, and the denoising can be applied to their time evolution. An analysis of the low-rank and sparsity-related hyperparameters, the selection of the ranks, and the effect of the block size on the classification results are recommended.

We calculated the DCEUS features by transforming the low-rank tensor back to the original size. Instead, the low-rank format could be kept for calculating the features without forming the tensor. This will further relax the RAM requirements and speed up the CUDI feature calculation time. In addition, low-rank decomposition allows the estimation of possibly missing or corrupted temporal frames [47].

VII. Conclusion

In this paper, we investigate low-rank tensor decomposition-based denoising of dynamic contrast-enhanced ultrasound data. Besides the use of MLSVD, we introduced a low-rank denoising algorithm suitable for Rayleigh-shaped multiplicative noise based on a gradient descent algorithm and a low-rank denoising algorithm based on the OITNN framework. The proposed algorithms perform better than the truncated MLSVD in the simulation. In the in vivo recordings, the same improvement was not observed for distinguishing benign and malignant voxels. However, low-rank tensor-based denoising using MLSVD outperformed other state-of-the-art approaches. In addition, the tr-MLSVD resulted in the best compression of the DCEUS recordings with a factor of 80. Although the addition of the noise distribution aids the visualization of DCEUS recordings, we did not find any improvement in the classification of prostate cancer. Considering these aspects and the added processing time of the GTE and OITNN methods, we believe approximating the noise as WGN, i.e. the use of tr-MLSVD is the best approach for denoising and compressing DCEUS recordings.

Appendix A
Relaxation of the Model Fitting With Exponential Linear Unit (ELU) Function

The model-fitting is the most time-consuming DCEUS feature, as described in Section IV-C. Here, we describe the relaxation of the model fitting through the ELU function, which resulted in a faster convergence. The modified local density random walk [4] was described as

$$\mathbf{l}_{xyz}(t) = c \ln \left(\alpha \sqrt{\frac{\kappa}{2\pi(t - t_0)}} e^{-\frac{\kappa(t-t_0-\mu)^2}{2(t-t_0)}} + 1 \right). \tag{25}$$

where c is the dynamic range-related scaling parameter given in (4), κ is the local dispersion-related parameter independent of the injection site's distance, μ is the convective time, and the t_0 is the injection time, α is the area under the time intensity curve. Using the ELU function, we could approximate the modified local density walk model as

$$\mathbf{l}_{xyz}(t) \approx ELU\left(\theta_1 - \frac{1}{2} c \ln t - t_0 - \frac{\theta_2(t - \theta_3)}{2(t - t_0)}\right), \tag{26}$$

with

$$ELU(x) = \begin{cases} x, & \text{if } x > 0, \\ e^x - 1, & \text{if } x \leq 0, \end{cases} \tag{27}$$

and

$$\theta_1 = c \ln \alpha + \frac{c}{2} \ln (\frac{\kappa}{2\pi}), \ \theta_2 = c\kappa, \ \theta_3 = \mu + t_0. \tag{28}$$

The model is fitted starting from the appearance time, which is estimated separately.

Appendix B
Modification of the Derivatives of GTE

The derivative of the loss function in (14) with respect to the first factor matrix $\frac{\partial F}{\partial \mathbf{U}^{(1)}}$ is described in [23] as

$$\frac{\partial F}{\partial \mathbf{U}^{(1)}} = \left(\frac{\partial F}{\partial \underline{\mathbf{L}}}\right)_{(1)} (\mathbf{U}^{(4)} \otimes \mathbf{U}^{(3)} \otimes \mathbf{U}^{(2)}) \mathbf{S}_{(1)}^T + a\, \mathbf{U}^{(1)} \left((\mathbf{U}^{(1)})^T \mathbf{U}^{(1)} - b^2 \mathbf{I}\right). \tag{29}$$

We describe the modification of the derivatives on the first-factor matrix, which can be applied to the other factor matrices by changing the unfoldings.

Proposition 1: The kronecker products given in (29) can be avoided by rewriting it as

$$\frac{\partial F}{\partial \mathbf{U}^{(1)}} = \left(\frac{\partial F}{\partial \underline{\mathbf{L}}} \times_2 (\mathbf{U}^{(2)})^T \times_3 (\mathbf{U}^{(3)})^T \times_4 (\mathbf{U}^{(4)})^T\right)_{(1)} \times (\mathbf{S}_{(1)})^T + a\, \mathbf{U}^{(1)} \left((\mathbf{U}^{(1)})^T \mathbf{U}^{(1)} - b^2 \mathbf{I}\right). \tag{30}$$

Proof: Using the relation [48] between

$$(\underline{\mathbf{S}} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)} \times_4 \mathbf{U}^{(4)})_{(k)}, \tag{31}$$

and

$$\mathbf{U}_{(k)} \mathbf{S}_{(k)} (\mathbf{U}^{(4)} \otimes \ldots \mathbf{U}^{(k+1)} \otimes \mathbf{U}^{(k-1)} \otimes \ldots \mathbf{U}^{(1)})^T, \tag{32}$$

the equality between the equations (29) and (34) can be proven. This can be shown by

$$\begin{aligned} &\left(\frac{\partial F}{\partial \underline{\mathbf{L}}} \times_2 (\mathbf{U}^{(2)})^T \times_3 (\mathbf{U}^{(3)})^T \times_4 (\mathbf{U}^{(4)})^T\right)_{(1)} \\ &= \left(\frac{\partial F}{\partial \underline{\mathbf{L}}} \times_1 \mathbf{I} \times_2 (\mathbf{U}^{(2)})^T \times_3 (\mathbf{U}^{(3)})^T \times_4 (\mathbf{U}^{(4)})^T\right)_{(1)}, \\ &= \left(\frac{\partial F}{\partial \underline{\mathbf{L}}}\right)_{(1)} (\mathbf{U}^{(4)} \otimes \mathbf{U}^{(3)} \otimes \mathbf{U}^{(2)}). \end{aligned} \tag{33}$$

In a similar fashion, the derivative of the loss function with respect to the other factor matrices can be shown. ■

Appendix C
Derivatives of GTE

Define the element-wise derivative operator as $\frac{\partial F}{\partial \underline{\mathbf{L}}} : \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4} \rightarrow \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4}$. The derivative of F with respect to the factor matrices are

$$\begin{aligned} \frac{\partial F}{\partial \mathbf{U}^{(1)}} &= \left(\frac{\partial F}{\partial \underline{\mathbf{L}}} \times_2 (\mathbf{U}^{(2)})^T \times_3 (\mathbf{U}^{(3)})^T \times_4 (\mathbf{U}^{(4)})^T\right)_{(1)} \\ &\quad \times (\mathbf{S}_{(1)})^T + a\, \mathbf{U}^{(1)} \left((\mathbf{U}^{(1)})^T \mathbf{U}^{(1)} - b^2 \mathbf{I}\right), \\ \frac{\partial F}{\partial \mathbf{U}^{(2)}} &= \left(\frac{\partial F}{\partial \underline{\mathbf{L}}} \times_1 (\mathbf{U}^{(1)})^T \times_3 (\mathbf{U}^{(3)})^T \times_4 (\mathbf{U}^{(4)})^T\right)_{(2)} \\ &\quad \times (\mathbf{S}_{(2)})^T + a\, \mathbf{U}^{(2)} \left((\mathbf{U}^{(2)})^T \mathbf{U}^{(2)} - b^2 \mathbf{I}\right), \\ \frac{\partial F}{\partial \mathbf{U}^{(3)}} &= \left(\frac{\partial F}{\partial \underline{\mathbf{L}}} \times_1 (\mathbf{U}^{(1)})^T \times_2 (\mathbf{U}^{(2)})^T \times_4 (\mathbf{U}^{(4)})^T\right)_{(3)} \\ &\quad \times (\mathbf{S}_{(3)})^T + a\, \mathbf{U}^{(3)} \left((\mathbf{U}^{(3)})^T \mathbf{U}^{(3)} - b^2 \mathbf{I}\right), \\ \frac{\partial F}{\partial \mathbf{U}^{(4)}} &= \left(\frac{\partial F}{\partial \underline{\mathbf{L}}} \times_1 (\mathbf{U}^{(1)})^T \times_2 (\mathbf{U}^{(2)})^T \times_3 (\mathbf{U}^{(3)})^T\right)_{(4)} \\ &\quad \times (\mathbf{S}_{(4)})^T + a\, \mathbf{U}^{(4)} \left((\mathbf{U}^{(4)})^T \mathbf{U}^{(4)} - b^2 \mathbf{I}\right). \end{aligned} \tag{34}$$

Finally, the derivative with respect to $\underline{\mathbf{S}}$ has the form,

$$\frac{\partial F}{\partial \underline{\mathbf{S}}} = \frac{\partial F}{\partial \underline{\mathbf{L}}} \times_1 (\mathbf{U}^{(1)})^T \times_2 (\mathbf{U}^{(2)})^T \times_3 (\mathbf{U}^{(3)})^T \times_4 (\mathbf{U}^{(4)})^T . \tag{35}$$

Acknowledgment

The authors would like to thank Angiogenesis Analytics for providing the resources and Auke Jager for conducting the clinical trials.

References

[1] S. Turco et al., "Contrast-enhanced ultrasound quantification: From kinetic modeling to machine learning," *Ultrasound Med. Biol.*, vol. 46, no. 3, pp. 518–543, Mar. 2020.
[2] J. Folkman, "Role of angiogenesis in tumor growth and metastasis," *Seminars Oncol.*, vol. 29, no. 6Q, pp. 15–18, Dec. 2002.
[3] M. Mischi, M. P. J. Kuenen, and H. Wijkstra, "Angiogenesis imaging by spatiotemporal analysis of ultrasound contrast agent dispersion kinetics," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 59, no. 4, pp. 621–629, Apr. 2012.
[4] M. P. J. Kuenen, M. Mischi, and H. Wijkstra, "Contrast-ultrasound diffusion imaging for localization of prostate cancer," *IEEE Trans. Med. Imag.*, vol. 30, no. 8, pp. 1493–1502, Aug. 2011.
[5] S. G. Schalk et al., "Contrast-enhanced ultrasound angiogenesis imaging by mutual information analysis for prostate cancer localization," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 3, pp. 661–670, Mar. 2017.
[6] R. R. Wildeboer et al., "Convective-dispersion modeling in 3D contrast-ultrasound imaging for the localization of prostate cancer," *IEEE Trans. Med. Imag.*, vol. 37, no. 12, pp. 2593–2602, Dec. 2018.
[7] R. J. van Sloun, L. Demi, A. W. Postema, J. J. de la Rosette, H. Wijkstra, and M. Mischi, "Ultrasound-contrast-agent dispersion and velocity imaging for prostate cancer localization," *Med. Image Anal.*, vol. 35, pp. 610–619, Jan. 2017.
[8] P. J. A. Frinking, A. Bouakaz, J. Kirkhorn, F. J. Ten Cate, and N. de Jong, "Ultrasound contrast imaging: Current and new potential methods," *Ultrasound Med. Biol.*, vol. 26, no. 6, pp. 965–975, Jul. 2000. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0301562900002295
[9] O. V. Michailovich and A. Tannenbaum, "Despeckling of medical ultrasound images," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 53, no. 1, pp. 64–78, Jan. 2006.

[10] G. Barrois, A. Coron, T. Payen, A. Dizeux, and L. Bridal, "A multiplicative model for improving microvascular flow estimation in dynamic contrast-enhanced ultrasound (DCE-U.S.): Theory and experimental validation," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 60, no. 11, pp. 2284–2294, Nov. 2013.

[11] A. D. Bar-Zion, C. Tremblay-Darveau, M. Yin, D. Adam, and F. S. Foster, "Denoising of contrast-enhanced ultrasound cine sequences based on a multiplicative model," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 8, pp. 1969–1980, Aug. 2015.

[12] M. P. J. Kuenen, I. H. F. Herold, H. H. M. Korsten, J. J. M. C. H. de la Rosette, H. Wijkstra, and M. Mischi, "Maximum-likelihood estimation for indicator dilution analysis," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 3, pp. 821–831, Mar. 2014.

[13] M. Mischi et al., "Contrast-ultrasound dispersion imaging of cancer neovascularization by mutual-information analysis," in *Proc. IEEE Int. Ultrason. Symp.*, Sep. 2014, pp. 1148–1151.

[14] M. P. J. Kuenen, T. A. Saidov, H. Wijkstra, and M. Mischi, "Contrast-ultrasound dispersion imaging for prostate cancer localization by improved spatiotemporal similarity analysis," *Ultrasound Med. Biol.*, vol. 39, no. 9, pp. 1631–1641, Sep. 2013.

[15] R. R. Wildeboer, R. J. G. van Sloun, P. Huang, H. Wijkstra, and M. Mischi, "3-D multi-parametric contrast-enhanced ultrasound for the prediction of prostate cancer," *Ultrasound Med. Biol.*, vol. 45, no. 10, pp. 2713–2724, Oct. 2019.

[16] R. R. Wildeboer et al., "Blind source separation for clutter and noise suppression in ultrasound imaging: Review for different applications," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 67, no. 8, pp. 1497–1512, Aug. 2020.

[17] A. G. Bruce, D. L. Donoho, H. Gao, and R. D. Martin, "Denoising and robust nonlinear wavelet analysis," *Wavelet Appl.*, vol. 2242, pp. 325–336, Mar. 1994.

[18] M. Mischi, T. Kalker, and E. Korsten, "Videodensitometric methods for cardiac output measurements," *EURASIP J. Adv. Signal Process.*, vol. 2003, no. 5, pp. 1–11, Apr. 2003.

[19] M. Calis, M. Mischi, A.-J. Van Der Veen, and B. Hunyadi, "Denoising of dynamic contrast-enhanced ultrasound sequences: A multilinear approach," in *Proc. 15th Int. Joint Conf. Biomed. Eng. Syst. Technol.*, 2022, pp. 192–199.

[20] T. Yokota, N. Lee, and A. Cichocki, "Robust multilinear tensor rank estimation using higher order singular value decomposition and information criteria," *IEEE Trans. Signal Process.*, vol. 65, no. 5, pp. 1196–1206, Mar. 2016.

[21] P. Chen, M. Calis, H. Wijkstra, P. Huang, B. Hunyadi, and M. Mischi, "Multiparametric ultrasound and machine learning for prostate cancer localization," in *Proc. 30th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2022, pp. 907–911.

[22] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1253–1278, Apr. 2000.

[23] R. Han, R. Willett, and A. R. Zhang, "An optimal statistical and computational framework for generalized tensor estimation," *Ann. Statist.*, vol. 50, no. 1, pp. 1–29, Feb. 2022.

[24] A. Wang, Q. Zhao, Z. Jin, C. Li, and G. Zhou, "Robust tensor decomposition via orientation invariant tubal nuclear norms," *Sci. China Technolog. Sci.*, vol. 65, no. 6, pp. 1300–1317, 2022.

[25] L. Zhu, C.-W. Fu, M. S. Brown, and P.-A. Heng, "A non-local low-rank framework for ultrasound speckle reduction," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5650–5658.

[26] S. V. M. Sagheer and S. N. George, "Denoising of medical ultrasound images based on non-local similarity: A low-rank approach," in *Proc. TENCON IEEE Region 10 Conf.*, Nov. 2017, pp. 176–181.

[27] S. V. M. Sagheer, S. N. George, and S. K. Kurien, "Despeckling of 3D ultrasound image using tensor low rank approximation," *Biomed. Signal Process. Control*, vol. 54, Sep. 2019, Art. no. 101595.

[28] M. P. J. Kuenen, T. A. Saidov, H. Wijkstra, J. J. M. C. H. De La Rosette, and M. Mischi, "Correspondence–spatiotemporal correlation of ultrasound contrast agent dilution curves for angiogenesis localization by dispersion imaging," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 60, no. 12, pp. 2665–2669, Dec. 2013.

[29] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev. Soc. Ind. Appl. Math.*, vol. 51, no. 3, pp. 455–500, Aug. 2009.

[30] J. Rissanen, "A universal prior for integers and estimation by minimum description length," *Ann. Stat.*, vol. 11, no. 2, pp. 416–431, 1983, doi: 10.1214/aos/1176346150.

[31] M. E. Kilmer, K. Braman, N. Hao, and R. C. Hoover, "Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 1, pp. 148–172, 2013.

[32] F. Mezzadri, "How to generate random matrices from the classical compact groups," 2006, *arXiv:math-ph/0609050*.

[33] H. Xie, L. E. Pierce, and F. T. Ulaby, "Statistical properties of logarithmically transformed speckle," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 3, pp. 721–727, Mar. 2002.

[34] B. Delahunt, R. J. Miller, J. R. Srigley, A. J. Evans, and H. Samaratunga, "Gleason grading: Past, present and future," *Histopathology*, vol. 60, no. 1, pp. 75–86, Jan. 2012, doi: 10.1111/j.1365-2559.2011.04003.x.

[35] J. Tang, J.-C. Yang, Y. Luo, J. Li, Y. Li, and H. Shi, "Enhancement characteristics of benign and malignant focal peripheral nodules in the peripheral zone of the prostate gland studied using contrast-enhanced transrectal ultrasound," *Clin. Radiol.*, vol. 63, no. 10, pp. 1086–1091, Oct. 2008.

[36] N. Elie, A. Kaliski, P. Péronneau, P. Opolon, A. Roche, and N. Lassau, "Methodology for quantifying interactions between perfusion evaluated by DCE-U.S. and hypoxia throughout tumor growth," *Ultrasound Med. Biol.*, vol. 33, no. 4, pp. 549–560, Apr. 2007.

[37] A. W. Postema et al., "Dynamic contrast-enhanced ultrasound parametric imaging for the detection of prostate cancer," *BJU Int.*, vol. 117, no. 4, pp. 598–603, Mar. 2015.

[38] R. J. van Sloun, L. Demi, A. W. Postema, J. J. De La Rosette, H. Wijkstra, and M. Mischi, "Entropy of ultrasound-contrast-agent velocity fields for angiogenesis imaging in prostate cancer," *IEEE Trans. Med. Imag.*, vol. 36, no. 3, pp. 826–837, Mar. 2017.

[39] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[40] G. H. Golub and C. Reinsch, "Singular value decomposition and least squares solutions," in *Handbook for Automatic Computation*, vol. 2. Berlin, Germany: Springer, 1971, pp. 134–151.

[41] C. Lanczos, "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators," *J. Res. Nat. Bur. Standards*, vol. 45, no. 4, p. 255, Oct. 1950.

[42] D. T. Fetzer, V. Rafailidis, C. Peterson, E. G. Grant, P. Sidhu, and R. G. Barr, "Artifacts in contrast-enhanced ultrasound: A pictorial essay," *Abdominal Radiol.*, vol. 43, no. 4, pp. 977–997, Apr. 2018.

[43] M. R. Lowerison, M. N. Hague, A. F. Chambers, and J. C. Lacefield, "Improved linear contrast-enhanced ultrasound imaging via analysis of first-order speckle statistics," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 63, no. 9, pp. 1409–1421, Sep. 2016.

[44] C. F. Dietrich et al., "How to perform contrast-enhanced ultrasound (CEUS)," *Ultrasound Int. open*, vol. 4, no. 1, pp. E2–E15, Jan. 2018.

[45] A. Matoso and J. I. Epstein, "Defining clinically significant prostate cancer on the basis of pathological findings," *Histopathology*, vol. 74, no. 1, pp. 135–145, Jan. 2019, doi: 10.1111/his.13712.

[46] G. Russo, M. Mischi, W. Scheepens, J. J. De la Rosette, and H. Wijkstra, "Angiogenesis in prostate cancer: Onset, progression and imaging," *BJU Int.*, vol. 110, no. 11c, pp. E794–E808, Dec. 2012.

[47] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208–220, Jan. 2013.

[48] H. G. Golub and C. F. van Loan, *Matrix Computations*. Baltimore, MD, USA: JHU Press, 2013.