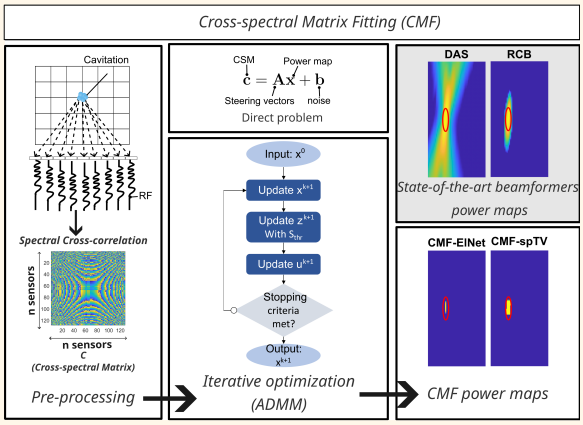
An inverse method using Cross spectral Matrix Fitting for passive cavitation imaging

Célestine Lachambre, Adrian Basarab, Jean-Christophe Béra, Barbara Nicolas, François Varray, Bruno Gilles

An inverse method using Cross spectral Matrix Fitting for passive cavitation imaging

Célestine Lachambre, Adrian Basarab, Jean-Christophe Béra, Barbara Nicolas, François Varray, Bruno Gilles

***Abstract*— High intensity focused ultrasound (HIFU) can produce cavitation, which requires monitoring for specific applications such as sonoporation, targeted drug delivery or histotripsy. Passive acoustic mapping has been proposed in the literature as a method for monitoring cavitation, but it lacks spatial resolution, primarily in the axial direction, due to the absence of a time reference. This is a common issue with passive imaging compared to standard pulse-echo ultrasound. In order to improve the axial resolution, we propose an adaptation of the Cross spectral Matrix Fitting (CMF) method for passive cavitation imaging, which is based on the resolution of an inverse problem with different regularizations that promote sparsity in the reconstructed cavitation maps: Elastic Net (CMF-ElNet) and sparse Total Variation (CMF-spTV). The results from both simulated and experimental data are presented and compared to state-of-the-art approaches, such as the frequential Delay-and-Sum (DAS) and the frequential Robust Capon Beamformer (RCB). We show the interest of the method for improving the axial resolution, with an axial Full Width Half Maximum (FWHM) divided by 3 and 5 compared to RCB and DAS, respectively. Moreover, CMF based methods improve Contrast-to-Noise Ratio (CNR) by more than 15 dB in experimental conditions compared to RCB. We also show the advantage of the sparse Total Variation prior over Elastic Net when dealing with cloud shaped cavitation sources, that can be assumed as sparse grouped sources.**



***Index Terms*— Cavitation imaging, Constrained optimization, Frequential beamforming, Inverse problems, Passive ultrasound imaging, Total Variation**

I. Introduction

HIGH intensity focused ultrasound (HIFU) finds diverse therapeutic applications relying on various underlying mechanisms [1]: for example, HIFU can generate thermal energy for tissue ablation or induce mechanical effects conducive to lithotripsy [2], [3]. The phenomenon of cavitation, where clouds of gas bubbles oscillate and potentially implode due to pressure variations caused by focused ultrasound waves, is utilized in numerous applications [1], such as sonoporation [4], [5], targeted drug delivery [6]–[8], opening of the blood-brain barrier [9], [10] and histotripsy [11]–[13].

In treatments involving cavitation effects, it is necessary to accurately locate and measure its activity, monitor its effects, and possibly establish closed-loop control. To avoid interference with HIFU signals, Gyöngy *et al.* adapted passive temporal acoustic mapping method from seismic imaging [14], [15] to cavitation localization, known as time-exposure acoustic or passive -and-Sum (DAS) [16]. However, unless very specific high aperture arrays are used, DAS is limited in its ability to accurately localize cavitation because it lacks a time reference, resulting in limited axial resolution. Moreover, passive temporal beamforming is computationally cumbersome as it requires delaying and adding signals for each pixel in the power map. Spatial resolution has been improved by various adaptive beamforming techniques, including the Robust Capon Beamformer (RCB) [17], [18], which has been adapted for passive cavitation imaging in the temporal domain [19], [20],

This work was supported by the LABEX CELYA (ANR-10- LABX-0060) and LABEX PRIMES (ANR-11-LABX-0063) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR), as well as the "CAVIIAR" Project (ANR-22-CE19-0006), operated by the French National Research Agency (ANR).

C. Lachambre is with LabTau, Inserm U1032, Université Claude Bernard Lyon 1, and INSA-Lyon, Universite Claude Bernard Lyon 1, CNRS, Inserm, CREATIS UMR5220, U1294, F-69100, LYON, France. (e-mail: celestine.lachambre@inserm.fr)

A. Basarab, B. Nicolas and F. Varray are with INSA-Lyon, Universite Claude Bernard Lyon 1, CNRS, Inserm, CREATIS UMR5220, U1294, F-69100, LYON, France. (e-mail: adrian.basarab@creatis.insa-lyon.fr, barbara.nicolas@creatis.insa-lyon.fr, francois.varray@creatis.insa-lyon.fr)

J-C. Béra and B. Gilles are with LabTau, Inserm U1032, Université Claude Bernard Lyon 1, Lyon, France. (e-mail: jean-christophe.bera@inserm.fr, bruno.gilles@inserm.fr)

Highlights

- **We introduce CMF-spTV: an inverse reconstruction in passive cavitation imaging with a Sparse Total Variation penalization.**
- **CMF-spTV reduces the axial lobe by 5 and 10 compared to RCB and DAS, respectively, and better reconstructs the cloud sources.**
- **Using inverse methods in passive cavitation imaging allows the introduction of prior information to the reconstruction.**

as well as phase coherence factor [21] and Pisarenko [22] beamformers, at the cost of increased computational time.

Therefore, frequency-based beamformers, widely developed in domains requiring array processing such as aeroacoustics [23], have been explored, initially with Passive Delay-and-Sum [24], as a means to reduce computational complexity while maintaining comparable spatial resolution to temporal beamformers. Adaptive beamformers, including Capon and Robust Capon, Pisarenko and MidWay beamformers [25]–[28], have then been adapted to frequency-based passive cavitation imaging. These frequency-based beamformers introduce the Cross-Spectral Matrix (CSM) [29] to quantify the correlation between the Radio-Frequency (RF) signals in the Fourier domain. In parallel, angular-spectrum based beamformers have been adapted to passive cavitation imaging to maintain a temporal resolution close to time-domain based methods while reducing computational cost [30], [31]. Despite these advances, these beamformers continue to face challenges in achieving better axial resolution, a requirement for clinical applications.

In other domains, such as aeroacoustics [23], some source localization methods based on regularized inverse problems are noteworthy for their ability to incorporate prior information into the reconstruction [32]. Regularization techniques are used to address the ill-posed nature of inverse problems, such as when they are underdetermined. Sparsity-based regularization appears suitable for cavitation imaging due to the inherent sparsity of the resulting maps and the limited number of sensors compared to the grid size, making the problem underdetermined. The Deconvolution Approach for the Mapping of Acoustic Sources (DAMAS) [33], which deconvolves RF signals in the Fourier domain with steering vectors, has been first developed but lacks stability. Another method which formulates beamforming as the inversion of a direct model relating the cross-spectral matrix to the power map [32], appears to be more stable due to the use of the cross-spectral matrix, which can be robustly estimated [18]. In a preliminary study [34], we introduced a method that focuses on fitting the CSM with Elastic Net regularization to promote sparsity in cavitation maps, thereby improving axial resolution in cavitation imaging [32], [35]. Elastic Net is derivated from the Least Absolute Shrinkage and Selection Operator (LASSO) method, originally introduced by Tibshirani [36]. The latter encourages strong sparsity, which usually results in point-wise solutions. Elastic Net is known for its ability to relax sparsity compared to LASSO and can therefore produce more spatially compact results. Its interest in active medical ultrasound beamforming was shown in [37], [38]. However, in some cases, such as for cavitation millimeter clouds, sparsity constraints tend to reduce the source to multiple sparse points that are not representative of the true source shapes. To mitigate this problem, we here propose to also introduce the use of Sparse Total Variation as a regularization [39], [40], [41].

The remainder of this paper is structured as follows: after a reminder of the DAS and RCB beamformers that will be used as references, we present the developed Cross-spectral Matrix Fitting (CMF) algorithm, followed by a comparative analysis of the different methods in various simulation and experimental settings, the results of which are then discussed.

II. Imaging Method: Cross-spectral matrix fitting (CMF)

This section presents the signal model generally considered in passive acoustic mapping and used in this work, with a reminder of the two state-of-the-art beamformers considered for the comparative study, as well as a description of the CMF and the resolution method used.

We consider the case of an N-element linear array operating in passive mode and recording N RF signals of T samples originating from a medium containing acoustic sources. We denote by $\mathbf{Y} \in \mathbb{R}^{N\times T}$ the matrix containing the N RF signals, and $\bar{\mathbf{Y}} \in \mathbb{C}^{N\times F}$ its corresponding version in the frequency domain, *i.e.*, in which each RF signal is replaced by its 1D discrete Fourier transform containing F frequencies. In the following, we will focus on a single frequency f and thus consider the acquired data as a vector $\bar{\mathbf{y}} \in \mathbb{C}^{N\times 1}$, which contains the frequency samples corresponding to the same frequency for each of the N RF signals.

A. Cross-spectral matrix

The CSM, denoted by $\mathbf{C} \in \mathbb{C}^{N\times N}$, is the cross-correlation of the Fourier domain representation of the RF signals, given by:

$$\mathbf{C} = \mathbb{E}[\bar{\mathbf{y}}\bar{\mathbf{y}}^*], \tag{1}$$

where $\cdot^*$ stands for the complex conjugate and $\mathbb{E}[\cdot]$ is the expected value. Note that the CSM can be viewed as the central point of cavitation imaging beamformers working in the frequency domain. In the present work, to robustly estimate the CSM, the temporal RF signals are divided in K time windows (further denoted snapshots) giving [18]:

$$\mathbf{C} = \frac{1}{K}\sum_{k=1}^{K} \bar{\mathbf{y}}_k \bar{\mathbf{y}}_k^*, \tag{2}$$

where $\bar{\mathbf{y}}_k$ are the RF signals in the Fourier domain in the temporal window corresponding to the k-th snapshot. Snapshots are generally overlapped in order to add stability. K is

therefore expressed as:

$$K = \frac{T_{rec}}{(1 - \%_{ov})T_{snap}}, \tag{3}$$

where T_{rec} is the duration of the whole sequence, T_{snap} is the duration of a snapshot and $\%_{ov}$ is the proportion of overlapping. In practice, a compromise between large K, leading to more stable solutions, and long T_{snap} insuring sufficient spectral resolution, is needed.

Passive DAS applied to cavitation mapping was first introduced in the time domain in [16]. The received raw signals are virtually focused on each point of the beamforming grid, which imposes relatively high computational resources. The spectral version of this beamformer substantially decreased the computational resources needed compared to temporal beamformers. Using the CSM, the estimated power at a particular point $\vec{r}$ given by passive DAS [17], [26] can be written as:

$$P_{DAS}(\vec{r}) = \frac{\mathbf{a}^*(\vec{r})\,\mathbf{C}\,\mathbf{a}(\vec{r})}{N^2}, \tag{4}$$

where $\mathbf{a}(\vec{r}) = e^{-2\pi f \mathbf{d}(\vec{r})/c_0}$, is the propagation vector of a given pixel $\vec{r}$, with c_0 the speed of sound in the medium and $\mathbf{d}(\vec{r})$ the estimated time-of-flight of the signal emitted by a source at location $\vec{r}$ to the different array elements. As in the case of passive DAS beamformer, passive Robust Capon Beamformer (RCB) can be expressed in both the temporal and frequency domains, the latter being much less cumbersome in terms of computational load. It aims to adapt the steering vectors for each pixel using the CSM data in order to minimize the noise contribution [20]. The estimated power can be expressed as:

$$P_{RCB}(\vec{r}) = \frac{1}{N}\frac{\|\mathbf{a}_0(\vec{r})\|^2}{\mathbf{a}_0^*(\vec{r})\mathbf{C}^{-1}\mathbf{a}_0(\vec{r})}, \tag{5}$$

with $\mathbf{a}_0^*(\vec{r})$ the adapted propagation vectors, whose expression can be found in [26].

B. CMF's forward model

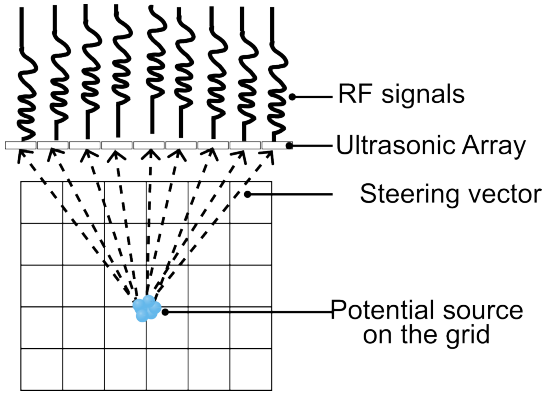


Fig. 1. Direct Model: potential sources on the grid of size M at position m are emitting ultrasonic signals received with a known delay by each element n of the probe. The delays are represented by the steering vectors $\mathbf{h}(n, m)$.

The proposed CMF method is based on the inversion of the model described below and illustrated in Fig. 1. We denote $\mathbf{s} \in \mathbb{R}^{M\times 1}$ a discrete vectorized version of the cavitation map to be beamformed, where M is the total number of potential source positions on a rectangular grid. The signals received from one source at a given location impinging on N elements are related to that source by the corresponding propagation vector. In the Fourier domain and considering the M potential source positions, a forward model can be established relating the frequency vector $\bar{\mathbf{y}}$ to the source map $\mathbf{s}$ using the model of wave propagation described in [32]:

$$\bar{\mathbf{y}} = \mathbf{H}\mathbf{s}, \tag{6}$$

where $\mathbf{H} \in \mathbb{C}^{N\times M}$ is a matrix whose columns represent the steering vectors relating one source position to all elements.

$$\mathbf{H} = \frac{1}{N}\begin{bmatrix} h(1,1) & \cdots & h(1,M) \\ \vdots & \ddots & \vdots \\ h(N,1) & \cdots & h(N,M) \end{bmatrix}, \tag{7}$$

with $h(n,m) = e^{-2\pi f d(n,m)/c_0}$ the steering vector of a source located at a given pixel m to the array element n and d the estimated distance between the signal emitted by a source located at a given pixel m to the array element n.

Injecting eq. (6) in the definition of the CSM in eq. (1) gives:

$$\mathbf{C} = \mathbf{H}\mathbb{E}[\mathbf{s}\mathbf{s}^*]\mathbf{H}^*. \tag{8}$$

Note that $\mathbb{E}[\mathbf{s}\mathbf{s}^*]$ represents the cross-correlation matrix between the sources located on the beamforming grid. Assuming that cavitation sources are uncorrelated, $\mathbb{E}[\mathbf{s}\mathbf{s}^*]$ becomes a diagonal matrix, where the main diagonal, denoted by $\mathbf{x} \in \mathbb{R}^{M\times 1}$, holds the power of each potential source on the beamforming grid.

Denoting by $x_m = s_m s_m^*$ the power of the m-th source, the model in eq. (8) becomes:

$$\mathbf{C} = \mathbf{H}\begin{bmatrix} x_1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & x_m & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & x_M \end{bmatrix}\mathbf{H}^*. \tag{9}$$

Finally, considering a vectorized version of the CSM, denoted by $\mathbf{c}$, and using simple matrix multiplication calculations, one can re-write this model as:

$$\mathbf{c} = \mathbf{A}\mathbf{x} + \mathbf{b}, \tag{10}$$

with $\mathbf{A}$, the reorganization of the steering vectors matrix multiplication (see appendix A), defined as:

$$\mathbf{A} = \frac{1}{N^2}\begin{bmatrix} h(1,1)h^*(1,1) & \cdots & h(1,M)h^*(1,M) \\ \vdots & \ddots & \vdots \\ h(1,1)h^*(N,1) & \cdots & h(1,M)h^*(N,M) \\ \vdots & \ddots & \vdots \\ h(N,1)h^*(1,1) & \cdots & h(N,M)h^*(1,M) \\ \vdots & \ddots & \vdots \\ h(N,1)h^*(N,1) & \cdots & h(N,M)h^*(N,M) \end{bmatrix} \tag{11}$$

where $\mathbf{c} \in \mathbb{C}^{N^2\times 1}$ is the vectorized version of the CSM, $\mathbf{A} \in \mathbb{C}^{N^2\times M}$ is a reshaped steering matrix, $\mathbf{b} \in \mathbb{C}^{N^2\times 1}$ is an additive white Gaussian noise accounting for measurement noise and model imperfections, and $\mathbf{x} \in \mathbb{R}^{M\times 1}$ is the vectorized form of the power map to be estimated.

C. Sparsity-based regularization

Inverting the model in eq. (10), *i.e.*, estimating the power map $\mathbf{x}$ from the CSM robustly estimated from the acquired raw RF data, can be treated as a typical inverse problem. Due to the ill-posedness of such problems, one common way to solve them is to express them as the minimization of a cost function:

$$\min_{\mathbf{x}} \left(\frac{1}{2}\|\mathbf{c} - \mathbf{A}\mathbf{x}\|_2^2 + g(\mathbf{x}) \right), \tag{12}$$

where the first term is a quadratic data fidelity term resulting from the assumption of additive Gaussian noise and $g(\mathbf{x})$ is a penalty function that contains prior information on the cavitation map.

In this work, the regularization functions proposed are based on the hypothesis that only a low number of cavitation sources occur in the field of view. Therefore, the cavitation map to be estimated may be considered sparse. The most classical way to impose sparsity is to use the ℓ_1-norm, turning the optimization in eq. (12) into the common LASSO problem [36]:

$$\min_{\mathbf{x}} \left(\frac{1}{2}\|\mathbf{c} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda_{l1}\|\mathbf{x}\|_1 \right), \tag{13}$$

where $\|\mathbf{x}\|_1$ is the ℓ_1-norm of $\mathbf{x}$ weighed by the hyperparameter λ_{l1}. The choice of λ_{l1} is guided by the trade-off between the data fidelity and the sparsity of the cavitation map. It is shown in [42] that its theoretical maximum value, before the solution degenerates to 0, is:

$$\lambda_{max} = \|\mathbf{A}^T\mathbf{c}\|_\infty. \tag{14}$$

where $\cdot^T$ stands for the transpose operator. In the remainder of the paper, this value is considered for all regularization terms as their maximum. The regularization terms are then weighted as a percentage of this maximum value.

However, LASSO tends to give pointwise solutions only, and, in order to be able to consider spread source configurations, we evaluated two different regularizations. The first one is a common relaxation of the LASSO regularization, called Elastic Net [35]. It considers, in addition to the ℓ_1-norm , an ℓ_2-norm, that tends to create less pointwise results. In this case, the optimization problem to solve is:

$$\mathbf{P}_{CMF-ElNet} = \min_{\mathbf{x}} \left(\frac{1}{2}\|\mathbf{c} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda_{l1}\|\mathbf{x}\|_1 + \frac{\lambda_{l2}}{2}\|\mathbf{x}\|_2^2 \right). \tag{15}$$

This method will be further denoted CMF-ElNet.

Although CMF-ElNet relaxes the sparsity constraints, it might not consider the group structure of cavitation bubble clouds, and reproduces, in its way, the axial bias present in classical DAS and RCB beamformers. Another class of regularization, called sparse Total Variation, penalizes the difference between neighboring pixels. In case of 2D anisotropic Total Variation, the penalization factor differs between vertical and horizontal axes, and could force the solution to spread less in the axial direction than in the lateral one. It is expressed as:

$$\begin{aligned} \mathbf{P}_{CMF-spTV} = \min_{\mathbf{x}} \Big(& \frac{1}{2}\|\mathbf{c} - \mathbf{A}\mathbf{x}\|_2^2 \\ & + \lambda_{TV_{ax}} \sum |\mathbf{x}_{i,j} - \mathbf{x}_{i+1,j}| \\ & + \lambda_{TV_{lat}} \sum |\mathbf{x}_{i,j} - \mathbf{x}_{i,j+1}| \\ & + \lambda_{l1}\|\mathbf{x}\|_1 \Big), \end{aligned} \tag{16}$$

where i and j represent the position of the pixels on a grid, $\lambda_{TV_{lat}}$ and $\lambda_{TV_{ax}}$ are respectively weighting lateral and axial Total Variation. This method will be further denoted CMF-spTV.

D. Inverse problem optimization

The Alternating Direction Method of Multipliers (ADMM) is used to solve the two optimization problems introduced in the previous section. eq. (15) and (16) can be expressed as a generalized LASSO problem [43], such that:

$$\min_{\mathbf{x}} \left(\frac{1}{2}\|\mathbf{c} - \mathbf{A}\mathbf{x}\|_2^2 + \|Diag(\boldsymbol{\lambda}_g)\mathbf{F}\mathbf{x}\|_1 + \frac{\lambda_{l2}}{2}\|\mathbf{x}\|_2^2 \right), \tag{17}$$

where $\mathbf{F} = [\mathbf{D}_{ax}; \mathbf{D}_{lat}; \mathbf{I}_M] \in \mathbb{R}^{P\times M}$ is a stack of $\mathbf{D}_{ax}$, the axial difference matrix, $\mathbf{D}_{lat}$, the lateral difference matrix, and $\mathbf{I}_M$ the identity matrix promoting sparsity. $\mathbf{F}$ is weighted by $\boldsymbol{\lambda}_g = [\boldsymbol{\lambda}_{TVax}, \boldsymbol{\lambda}_{TV_{lat}}, \boldsymbol{\lambda}_{l1}] \in \mathbb{R}^{1\times P}$, a vector allowing different weights for each penalty.
In the case of CMF-ElNet, $\boldsymbol{\lambda}_{TVax}$ and $\boldsymbol{\lambda}_{TV_{lat}}$ are set to zero while $\boldsymbol{\lambda}_{l1}$ and λ_{l2} are set to non-zero values. For CMF-spTV, $\boldsymbol{\lambda}_{TVax}, \boldsymbol{\lambda}_{TV_{lat}}$ and $\boldsymbol{\lambda}_{l1}$ are set to non-zero values while $\lambda_{l2} = 0$.

The general framework of ADMM is described in [44]. Applied to our problem, it becomes:

$$\begin{aligned} \min \quad & f(\mathbf{x}) + g(\mathbf{z}) \\ \text{s.t} \quad & \mathbf{F}\mathbf{x} = \mathbf{z}, \end{aligned} \tag{18}$$

where $f(\mathbf{x})$ represents the data fidelity term, $g(\mathbf{z})$ represents the penalty term, with z dependent of the variable x. By identification, $f(\mathbf{x})$ and $g(\mathbf{z})$ are defined by:

$$\begin{aligned} f(\mathbf{x}) &= \frac{1}{2}\|\mathbf{c} - \mathbf{A}\mathbf{x}\|_2^2 + \frac{\lambda_{l2}}{2}\|\mathbf{x}\|_2^2, \\ g(\mathbf{z}) &= \|Diag(\boldsymbol{\lambda}_g)\mathbf{z}\|_1. \end{aligned} \tag{19}$$

The augmented Lagrangian associated to eq. (18) and eq. (19) is:

$$\begin{aligned} L_p(\mathbf{x}, \mathbf{z}, \mathbf{u}) = \frac{1}{2}\|\mathbf{c} - \mathbf{A}\mathbf{x}\|_2^2 + \frac{\lambda_{l2}}{2}\|\mathbf{x}\|_2^2 \\ + \|Diag(\boldsymbol{\lambda}_g)\mathbf{z}\|_1 + \mathbf{u}^T(\mathbf{F}\mathbf{x} - \mathbf{z}) + (\frac{\rho}{2})\|\mathbf{F}\mathbf{x} - \mathbf{z}\|_2^2 \end{aligned} \tag{20}$$

Finally, the optimization problems in eq. (15) and eq. (16) can be solved with ADMM according to the scheme given in

eq. (21), following [45].

$$\begin{cases} \mathbf{x}^{k+1} & = \arg\min_{x} L_p(\mathbf{x}, \mathbf{z}^k, \mathbf{u}^k) \\ & = ((\mathbf{A}^t\mathbf{A}) + \rho(\mathbf{F}^t\mathbf{F}) + \lambda_{l2}\mathbf{I}_M)^{-1}(\mathbf{A}^t\mathbf{c} \\ & \quad +\rho\mathbf{F}^t(\mathbf{z}^k - \mathbf{u}^k)) \\ \mathbf{z}^{k+1} & = \arg\min_{z} L_p(\mathbf{x}^{k+1}, \mathbf{z}, \mathbf{u}^k) = S_{thr}(\mathbf{Fx}, \frac{\lambda_g}{\rho}) \\ \mathbf{u}^{k+1} & = \mathbf{u}^k + \mathbf{Fx}^{k+1} - \mathbf{z}^{k+1}, \end{cases} \quad (21)$$

where ρ is the penalty parameter of the Lagrangian described in [45]. $S_{thr}(\mathbf{Fx}, \frac{\lambda_g}{\rho})$ is the soft thresholding operator used to find the minimum of ℓ_1-norm, that are not differentiable [46]:

$$S_{thr}(\mathbf{Fx}, \frac{\lambda_g}{\rho}) = \max(0, \mathbf{Fx} - \frac{\lambda_g}{\rho}) - \max(0, -\mathbf{Fx} - \frac{\lambda_g}{\rho}). \quad (22)$$

The stopping criteria of the algorithm are chosen when the following conditions are met as explained by Boyd *et al.* [47].

$$\begin{cases} \|r^k\| = \|Fx^k - z^k\|_2 < \epsilon^{pri} \\ \|s^k\| = \| - \rho F^T(z^k - z^{k-1})\|_2 < \epsilon^{dual}, \end{cases} \quad (23)$$

where r^k and s^k are the primal and dual residuals at iteration k, respectively, and ϵ^{pri} and ϵ^{dual} their tolerances, expressed below as:

$$\begin{cases} \epsilon^{pri} = \sqrt{P}\epsilon^{abs} + \epsilon^{rel}\max(\|Fx^k\|_2, \| - z\|_2) \\ \epsilon^{dual} = \sqrt{P}\epsilon^{abs} + \epsilon^{rel}\|\rho F^T u^k\|_2, \end{cases} \quad (24)$$

with $\sqrt{P}$ the size of the first dimension of F, ϵ^{abs} the absolute residual, fixed at 10^{-6} and ϵ^{rel} the relative residual, fixed at 10^{-2} for the application.

III. Simulation and Experimental Setup

A. Simulation setup

To evaluate the performances of the different methods, two types of simulations are performed, one considering the sources as points and the other as elliptic clouds of different sizes. The coordinates for the described configurations are given in (lateral-, axial-) directions, with the origin taken at the center of the imaging probe.

The point source configurations aim to compare the PSF of the different beamforming methods, as well as their performance on close sources separation, in both axial and lateral directions. Therefore, we define two cases:

- Two inertial cavitation point sources laterally distributed at (−5,70) and (−3,70) mm;
- Two inertial cavitation point sources axially distributed at (−3,64) and (−3,72) mm.

The elliptic cloud source configurations aim to compare both the denoising capabilites and the fidelity to the true cavitation cloud shape for the different beamformers. Because of the asymetric extension of the axial and lateral lobes observed when imaging clouds, three cases are considered:

- A circular source 2 mm in diameter, with its center at (−7, 70) mm;
- An elliptic source extended in the axial direction, with principal axis dimensions of (1, 4) mm, and centered at (−7, 70) mm;
- An elliptic source extended in the lateral direction, with principal axis dimensions of (4, 1) mm, and centered at (−5, 70) mm;

For all configurations, each point source emitting is simulated as broadband noise for 200 µs, traveling to the probe elements as in [48]. In cloud configurations, we consider a density of 100 point sources/mm² randomly distributed in the cloud area. In all cases, additional white noise is added independently to each of the simulated RF signals. The linear probe simulated to obtain the received RF signals is similar to the L7-4 probe described in the next section, with $N = 128$ elements, a pitch of 298 µm, a center frequency of 5.2 MHz and a bandwidth at −6 dB ranging from 3.5 to 7.5 MHz.

B. Experimental setup

The experimental setup, represented in Fig. 2, consists of a degassed water tank in which is immersed a 1 MHz HIFU spherical transducer (Imasonics, France) and a needle placed at its focal point, 8 cm away from the transducer surface. The needle is slightly scratched on a paraffin bloc to trap cavitation nuclei at the tip. An imaging unit consists of an L7-4 probe (Verasonics, USA) with the same parameters as in simulation, placed above the needle. A high-speed camera (Phantom v12.1; Vision Research Inc., USA) is placed on the side of the tank to have the same imaging plane as the ultrasound imaging probe, with an exposure time of 4.68 µs and a frame rate of 18,000 frames per second. The imaging probe is controlled by a Vantage 256 system (Verasonics, USA). The HIFU transducer is connected to an amplifier unit (GN1000, PRÂNA, France) connected to a function generator to create the HIFU sequence (1.5 ms bursts of 1 MHz sinusoidal signal). For time reference, the Verasonics Vantage unit triggers the high-speed camera and the function generator at the beginning of the passive imaging acquisitions.

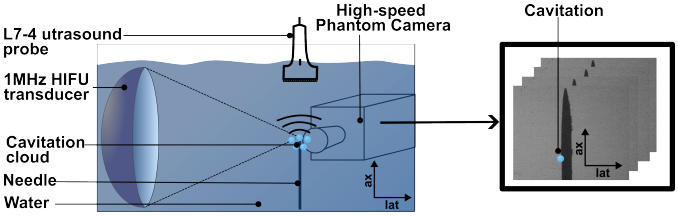


Fig. 2. Experimental setup. On the left side, the HIFU transducer focalizes on a needle placed at the focal spot. During the shot, the propagated ultrasound signals are saved through the L7-4 probe. The right image represents frame from the high speed camera video on which cavitation could be seen at the level of the blue circle (see video in the online supplemental material).

C. Imaging parameters and metrics

We compare the different CMF-based beamformers (CMF-ElNet and CMF-spTV) to the frequency-domain DAS of [24], [26] and to the frequency-domain RCB of [20], [26]. The CSM parameters are: $K = 130$ and $\%_{ov} = 0.9$.

The regularization parameters were manually set for each method, as a proportion of the maximum theoretical value of equation eq. (14), *i.e* $\lambda_{max} = \|\mathbf{A}^t\mathbf{c}\|_\infty$, with the circular

source configuration, and the proportions are kept constant across the different simulation and experimental configurations:

- For CMF-ElNet: $\lambda_{l1} = 0.5\lambda_{max}$ and $\lambda_{l2} = 0.3\lambda_{max}$ ($\lambda_{TVax} = \lambda_{TV_{lat}} = 0$).
- For CMF-spTV: $\lambda_{TVax} = 0.03\lambda_{max}$, $\lambda_{TV_{lat}} = 0.25\lambda_{max}$ and $\lambda_{l1} = 0.3\lambda_{max}$ ($\lambda_{l2} = 0$).
- For both CMF methods: $\rho = 4$.

Additional reconstructions illustrates the influence of the choice of λ values for the axial elliptic source configuration, as follows:

- Variation of λ_{l1} from 0 to $0.7\lambda_{max}$ with a fixed $\lambda_{TVax} = 0.03\lambda_{max}$ and $\lambda_{TV_{lat}} = 0.25\lambda_{max}$ ($\lambda_{l2} = 0$)
- Variation of $\lambda_{TV_{lat}}$ from 0 to $0.8\lambda_{max}$ with a fixed $\lambda_{TV_{ax}} = 0.12\lambda_{TV_{lat}}$ and $\lambda_{l1} = 0.3\lambda_{max}$ ($\lambda_{l2} = 0$)

For RCB, the adaptive parameter value is fixed at $\epsilon_{RCB} = 5$ for all configurations.

Imaging is done at 4 MHz for all the methods and the simulated configurations. It is set to 4.2 MHz for the experimental configuration in order to reconstruct only the inertial cavitation components. The pixel size for the reconstructed power maps is 200 µm for both simulated and experimental configurations. T_{rec} is set to 200 µs for all simulations, representing the duration of the emitted cavitation signal. In the experimental setup, we divide the emitted sequence into 8 sub-sequences of $T_{rec} = 166$ µs to have multiple reconstructions to analyze. The speed of sound is set to 1480 m/s for the experimental reconstruction.

In point source configurations, four metrics are estimated. On the one side, we compare the profile of the lateral and axial lobes of these maps with three metrics: the axial and lateral Full Width Half Maximum (FWHM), expressed in mm, and the power of separation of two sources by the minimum estimated power measured along the line between two sources, called P_{min}, expressed in dB relatively to the maximum of the power map. On the other side, we measure the position error as the absolute difference between the true position and the position of the barycenter of the estimated power for each cavitation spot, expressed in millimeters.

For simulated elliptic source and experimental configurations, we compare the Contrast-to-Noise Ratio (CNR) of the different methods, expressed as:

$$CNR = 10\log_{10}\left(\frac{|\mu_i - \mu_o|}{\sqrt{\sigma_i^2 + \sigma_o^2}}\right), \tag{25}$$

where μ_i and μ_o are respectively the average power within the signal and noise zones, respectively, and σ_i^2, σ_o^2 are the corresponding standard deviations. For simulated cases, the signal power zone is defined as the area inside the cavitation cloud. The noise power zone is defined as the area 2 mm around the cavitation cloud. For experimental configurations, the size of the cavitation cloud is difficult to estimate. Therefore, the noise power zone is located far from the cavitation area, where only noise should contribute, while the signal power zone is taken as a 0.4×0.7 mm rectangle, based on the high-speed camera visualization. In elliptic cloud simulations, to measure the fidelity of the reconstructed cavitation cloud shape, we also measure the Dice score at -3 dB. We consider true positives as pixels above or equal to -3 dB within the signal power zone, false negatives as pixels below -3 dB within the signal power zone, and false positives as pixels above -3 dB in the noise power zone. The Dice score is defined as

$$Dice = 2\frac{|X \cap Y|}{|X| + |Y|}, \tag{26}$$

where X is the binary image containing the true positive and the false negative (*i.e* the signal power zone) and Y is the binary image containing the true positive and the false positive (*i.e* the region above -3 dB).

IV. Results

A. Simulations results

Fig. 3 presents the reconstructions of inertial point source configurations with a SNR of 10 dB obtained with DAS, RCB, CMF-ElNet and CMF-spTV. In addition, the lateral and axial profiles and the position errors are shown in Fig. 4. Table I summarizes the different metrics. The position errors for the laterally distributed bubbles are below the wavelength at the reconstruction frequency (0.375 mm at 4 MHz) for both the RCB and CMF algorithms. For axially distributed sources, the position error is the lowest for CMF-ElNet (0.3 mm) and CMF-spTV (0.4 mm), when it is 0.6 mm for RCB and 1.7 mm for DAS. The axial FWHM is similar for CMF-spTV and CMF-ElNet (1.7 mm and 1.3 mm, respectively for the axially distributed case), while it reaches 3.5 mm for RCB and 14 mm for DAS. This affects the separation of the 2 lobes, which is better than -20 dB for CMF-ElNet and CMF-spTV, while for RCB and DAS this separation is -5.7 and -1.6 dB, respectively. On the other hand, due to its piecewise constant property, the lateral FWHM of CMF-spTV is 0.45 mm, larger than the ones of RCB and CMF-ElNet which are under 0.2 mm, but still better than the one of DAS.

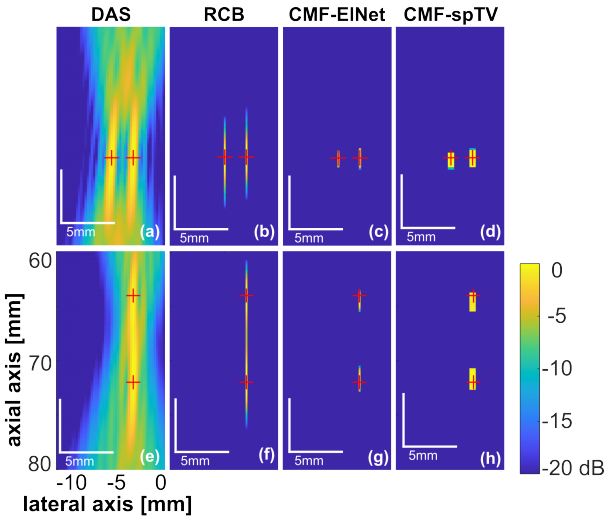


Fig. 3. Power maps of point source configurations. (a) to (d) represent the power maps of two laterally distributed inertial bubbles simulated respectively at (-5,70) mm and (-3,70) mm. (e) to (h) are the power maps of two axially distributed inertial bubbles simulated respectively at (-3,64) mm and (-3, 72) mm.

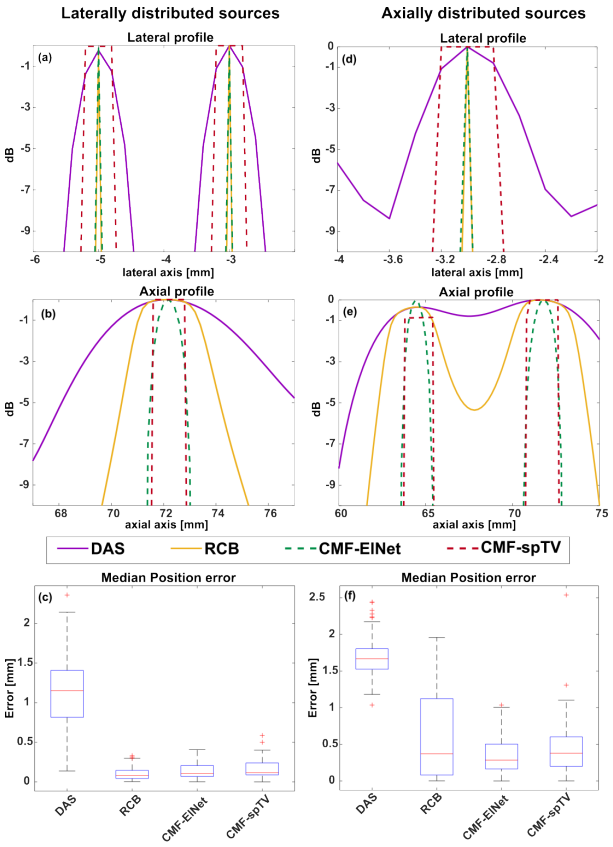


Fig. 4. Metrics for the point source configurations. The profiles are representative of Fig. 3. The position error is the difference between the ground truth and the position of the barycenter of the cavitation spots on the power maps, averaged on 50 different images for each configuration.

The purpose of simulating cloud configurations is to evaluate the ability of the different reconstruction algorithms to maintain shape consistency, as measured by the Dice score, and their resistance to noise as indicated by the CNR. Fig. 5 shows the reconstructed power maps using DAS (Fig. 5.a,e,i), RCB (Fig. 5.b,f,j), CMF-ElNet (Fig. 5.c,g,k) and CMF-spTV (Fig. 5.d,h,l). The size of the estimated source increases in the axial direction as the lateral size of the cloud increases for DAS and RCB. However, this behavior is partially corrected when CMF-based regularizations are used. It can also be seen that CMF-ElNet produces very sparse and ungrouped results, especially for laterally extended clouds (Fig. 5.k). Conversely, the piecewise constant property of CMF-spTV, combined with the anisotropy of the Total Variation, reconstructs clouds that are more similar to the ground truth. The metrics presented in Fig. 6 and Table II support these observations. For the circular source and the lateral elliptic source configurations, CMF-spTV has higher Dice scores than the other beamformers, with an advantage of 0.3 and 0.06 points over DAS, respectively, and 0.26 and 0.06 points over RCB, while, in the case of the circular source, CMF-ElNet has a lower Dice score, with a decrease of 0.29 and 0.34 point over DAS and RCB, respectively. For the axial elliptic configuration, however, the Dice score for CMF-spTV is lower than for RCB of 0.11 point, but still higher than for DAS and CMF-ElNet, with 0.10 and 0.49 point difference, respectively. CMF-spTV improves the

TABLE I

METRICS AVERAGED OVER 50 SIMULATION CASES FOR POINT SOURCE CONFIGURATIONS. STANDARD DEVIATION IS SHOWN IN PARENTHESES. BEST VALUES ARE IN BOLD TYPE.

Axial FWHM [mm]				
Beamformer	***DAS***	***RCB***	***CMF-ElNet***	***CMF-spTV***
Laterally distributed	7.4 (0.5)	3.3 (0.2)	**1.3 (0.2)**	1.4 (0.1)
Axially distributed	14 (2)	3.5 (0.5)	**1.3 (0.2)**	1.7 (0.5)
Lateral FWHM [mm]				
Beamformer	***DAS***	***RCB***	***CMF-ElNet***	***CMF-spTV***
Laterally distributed	0.61 (0.01)	< 0.2	< 0.2	0.44 (0.01)
Axially distributed	0.67 (0.005)	< 0.2	< 0.2	0.45 (0.01)
Separation power [dB]				
Beamformer	***DAS***	***RCB***	***CMF-ElNet***	***CMF-spTV***
Laterally distributed	-13.9 (0.7)	< -20	< -20	< -20
Axially distributed	-1.6 (0.7)	-5.7 (0.6)	< -20	< -20
Position error [mm]				
Beamformer	***DAS***	***RCB***	***CMF-ElNet***	***CMF-spTV***
Laterally distributed	1.1 (0.4)	**0.10 (0.01)**	0.1 (0.1)	0.2 (0.1)
Axially distributed	1.7 (0.3)	0.6 (0.6)	**0.3 (0.2)**	0.4 (0.3)

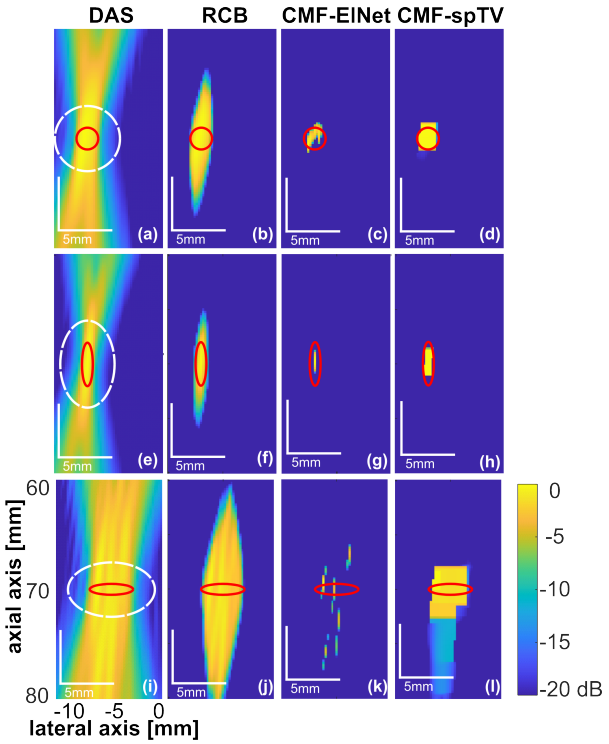


Fig. 5. Power map of clouds with the different reconstruction methods. First column is for DAS, second column is for RCB, third column is for CMF-ElNet and fourth column is for CMF-spTV. (a) to (d) are the power maps of the circular cloud 2 mm in diameter, (e) to (h) are the power maps of a elliptic cloud of size (1,4) mm and (i) to (l) of size (4,1) mm. The white dotted zone corresponds to the noise area and the red zone corresponds to the signal area for CNR and Dice computations.

CNR on average by 10 dB, 9 dB, and 3.7 dB compared to DAS, RCB, and CMF-ElNet, respectively. In general, CMF-ElNet has the second highest CNR after CMF-spTV, except for the lateral elliptic source configuration, where it completely fails to reconstruct a cloud type source, which results in the lowest CNR.

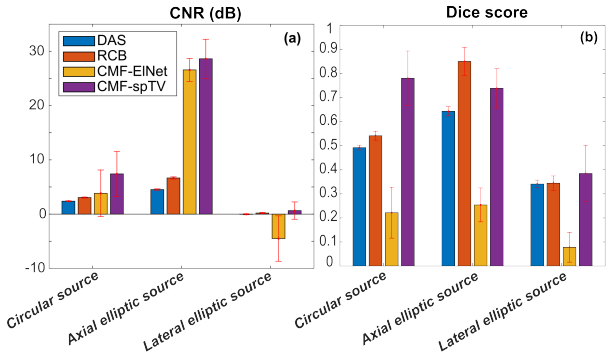


Fig. 6. Metrics for the cloud configurations. (a) shows the averaged CNR in dB, and (b) shows the averaged Dice score, in the three configurations imaged in Fig 5. Circular represents the configuration of Fig 5(a) to (d), axial ellipse of Fig 5(e) to (h), and longitudinal ellipse of Fig 5(i) to (l) Numerical values are presented in Table II.

TABLE II
METRICS OVER 50 SIMULATIONS FOR ELLIPTIC SOURCES CONFIGURATIONS. STANDARD DEVIATION IS SHOWN IN PARENTHESES. BEST VALUES ARE IN BOLD TYPE.

CNR [dB]				
Beamformer	***DAS***	***RCB***	***CMF-ElNet***	***CMF-spTV***
Circular source	2.39 (0.08)	3.06 (0.01)	4 (4)	**7 (4)**
Axial elliptic source	4.2 (0.1)	6.6 (0.2)	27 (2)	**29 (4)**
Lateral elliptic source	0 (0.1)	0.2 (0.1)	-4 (4)	**1 (2)**
Dice score				
Beamformer	***DAS***	***RCB***	***CMF-ElNet***	***CMF-spTV***
Circular source	0.49 (0.01)	0.54 (0.02)	0.2 (0.1)	**0.8 (0.1)**
Axial elliptic source	0.64 (0.02)	**0.85 (0.06)**	0.25 (0.07)	0.74 (0.08)
Lateral elliptic source	0.34 (0.02)	0.34 (0.03)	0.08 (0.06)	**0.4 (0.1)**

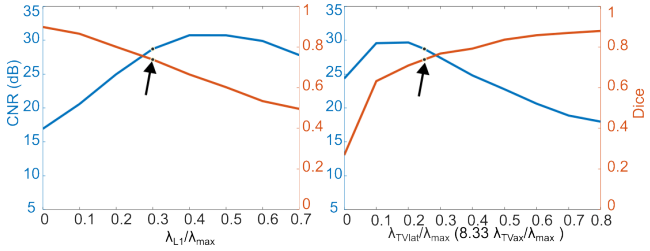


Fig. 7. CNR (blue curve) and Dice (orange curve) on axial elliptic clouds with varying (left) $\boldsymbol{\lambda_{l1}}$ and (right) $\boldsymbol{\lambda_{TV_{lat}}}$ and $\boldsymbol{\lambda_{TV_{ax}}}$. The arrows point the value of $\boldsymbol{\lambda}$ used for all the other configurations.

Fig 7 shows the CNR and Dice scores for the axial elliptic sources for different λ_{l1} or $\lambda_{TV_{ax}}$ and $\lambda_{TV_{lat}}$ values. When varying λ_{l1}, the CNR stays 20dB with λ_{l1} between 0.2 and $0.7\lambda_{max}$ while the Dice score decreases with the augmentation of λ_{l1}, to reach 0.5 when λ_{l1} reaches $0.5\lambda_{max}$. When $\lambda_{TV_{ax}}$ and $\lambda_{TV_{lat}}$ are varied, the CNR stays over 20dB as long as $\lambda_{TV_{lat}} < 0.6\lambda_{max}$ and that the Dice score increases with $\lambda_{TV_{lat}}$ and $\lambda_{TV_{ax}}$, to reach values larger than 0.5 when $\lambda_{TV_{lat}} > 0.1\lambda_{max}$.The regularization parameters we chose for the main configurations compromise between the best CNR and Dice scores.

TABLE III
MEAN COMPUTATION TIME FOR RECONSTRUCTION OVER 50 AXIAL ELLIPTIC SOURCE CONFIGURATION SIMULATIONS. THE * IS A SINGLE COMPUTATION TIME REQUIRED FOR THE INVERSION IN EQ. (21)

Reconstruction time [s]		
DAS	***RCB***	***CMF-based***
0.037	8.7	1.9 (+15*)

Optimization-based methods are known to be computationally expensive. In the present work, we accelerate the processing time by storing the result of the matrix inversion in eq. (21) (*i.e* $((\mathbf{A}^t\mathbf{A}) + \rho(\mathbf{F}^t\mathbf{F}) + \lambda_{l2}\mathbf{I}_M)^{-1}$), that is constant while imaging a given field-of-view. Table III regroups the average times for reconstruction over the 50 simulations of the axial elliptic cloud, with the inverse operator stored for CMF-based methods. The computation of this inversion takes several seconds ($15 - 20$s), thus motivating its storage. All the reconstructions are carried over Matlab (R2023a,The MathWorks, Natick, MA, USA), using the same configuration (12th Gen Intel(R) Core(TM) i7-12700H 2.30 GHz CPU and a 32 GB memory).

B. Experimental results

The cavitation is stationary and identifiable on the high-speed camera images. Fig. 8 presents the experimental cavitation maps overlaid with both a B-mode ultrasound frame in the first column and with a high-speed camera frame in the second column. The images show a reduction of the axial lobes for both CMF-ElNet and CMF-spTV compared to RCB and DAS. Additionally, CMF-ElNet tends to divide the cavitation cloud into multiple sources, while CMF-spTV only creates a single grouped source, as already observed in section IV-A. The region of interest (ROI) for the signal, located within the red square in Fig. 8.h, is chosen from the camera images where the cavitation is visible. The noise ROI, within the striped blue square of Fig. 8 is chosen in a zone where there is no cavitation visible on the camera images during the HIFU emission. Using this configuration, CMF-spTV achieves a median CNR of 33 dB, which represents an improvement of 2 dB, 21 dB, and 26 dB compared to CMF-ElNet, RCB, and DAS, respectively. This emphasizes the gain of contrast and noise filtering for CMF-based reconstructions compared to direct beamforming methods.

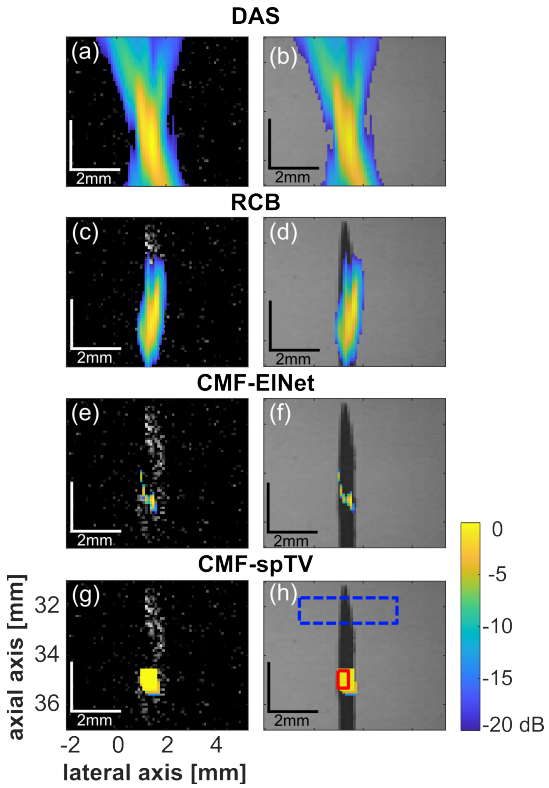


Fig. 8. Experimental results with cavitation on a needle. The passive cavitation images are overlaid with high active B-mode images taken before the pulse (first column), and with the high-speed camera images (second column) are reconstructed with DAS ((a), (b)), with RCB ((c), (d)), with CMF-ElNet ((e), (f)) and with CMF-spTV ((g), (h).

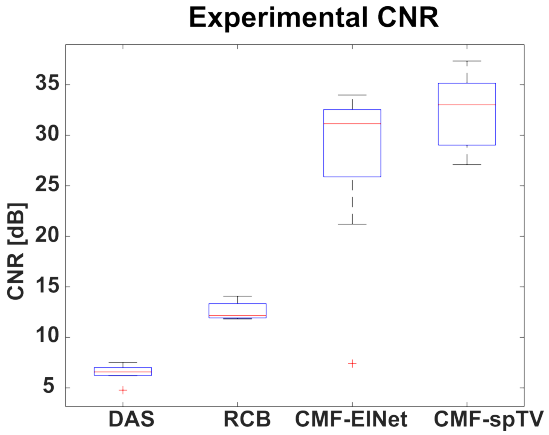


Fig. 9. Boxplot of the experimental CNR with cavitation on a needle, estimated over 8 reconstructed power maps.

V. Discussion

In this paper, we propose the first method for passive cavitation imaging using model based regularized inverse problems. This method, CMF [32], consists of the inversion of CSM throughout the steering vectors, along with a prior. The resolution of the problem is made using an ADMM [45] algorithm. The introduction of prior information is a true asset compared to state-of-art beamformers, as they can add different properties to the reconstruction, such as sparsity [35] (CMF-ElNet) and grouped sparsity (CMF-spTV) [39].

The main objective is to improve the axial resolution in passive cavitation imaging while preserving a grouped structure when imaging cavitation clouds.

Due to the sparsity property, CMF based methods enhance the axial resolution compared to state-of-the-art DAS and RCB. Indeed, in simulated point source configurations (Fig. 3), both CMF-spTV and CMF-ElNet exhibit improved axial resolution compared to state-of-the-art methods, with an axial FWHM reduction by a factor 5 and 3, compared to DAS and RCB, respectively.

This property allows as well a better separation of sources distributed along the axial axis, as it can be seen in Fig. 3.e to 3.h. With CMF-based methods, the axial separation power is better than -20 dB while it -5.7 dB for RCB and -1.6 dB for DAS (see Table I). This can be explained by the decrease of FWHM that increases the possibility of separating close axial sources and by the power concentration due to the sparsity constraint. This better separation is coupled with a better estimation of the source position in this configuration for CMF based method, with a lower error of position for CMF-ElNet and CMF-spTV than for RCB and DAS (see Table I).

Additionally, the sparsity property of CMF based methods improves the denoising of the passive cavitation maps. Indeed, in both simulated and experimental conditions, the CNR is improved for CMF based methods compared to RCB and DAS, as shown in Table II and Fig. 9.

When comparing Elastic Net and sparse Total Variation, particularly in the context of cloud configuration regularizations, it is evident that the latter better preserves the shape of the imaged clouds. Sparse Total Variation allows for denoising due to sparsity constraints and adds an anisotropic piecewise constant property to the reconstruction, which is a hypothesis that closely resembles a cavitation cloud. The results on cloud configurations confirms this, as shown in Fig. 5.l for CMF-spTV which displays a grouped cloud with a reduced axial lobe compared to state-of-the-art methods, while Fig. 5.k for CMF-ElNet shows dispersed points inscribed into an extended pattern similar to the main lobe of DAS and RCB. The higher Dice score and CNR of CMF-spTV demonstrate its superiority over CMF-ElNet in both simulated (Table II) and experimental (Fig. 9) configurations.

Finally, it is important to note the importance of accurately tuning the regularization parameters. Setting the sparsity regularization parameter too high, especially in the presence of multiple sources, may lead to underestimation of the weakest source, potentially erasing it. Conversely, setting it too low could result in denoising performances lower than that of state-of-the-art methods. Therefore, precise estimation of the numerous parameters required for CMF-based methods is crucial: understanding the theoretical maximum values of these parameters provides insight into their order of magnitude, but further investigations would be necessary to provide guidelines to choose those parameters.

VI. Conclusion

This study presents the potential application of Cross-Spectral Matrix Fitting (CMF) combined to two different

regularizations involving a sparsity a priori, in comparison to state-of-the-art beamformers for passive cavitation mapping. The results highlight the improved axial resolution achieved by CMF based methods for both point sources and millimeter bubble clouds compared to current state-of-the-art passive acoustic mapping strategies. Concerning the influence of the regularization, CMF-Elnet, based on elastic-net aims to promote only sparsity, while CMF-spTV, based on the sparse Total Variation, rather promotes grouped sparsity. The results show that CMF-spTV provides superior shape preservation compared to CMF-ElNet when imaging cavitation clouds, due to the Total Variation penalization incorporated into the regularization method. One difficulty of the developed model is the parameter tuning. To streamline this parameter tuning and avoid potential trade-offs between source resolvability and cloud reconstruction, advanced denoising methods could be explored while keeping the same inversion model. Additionally, an isotropic Total Variation regularization could be implemented in order to give the image of elliptic sources a more accurate shape. Furthermore, it has to be noted that, regarding the use of CMF methods in aeroacoustics [32] the proposed methods could be suited for quantitative analysis of the cavitation clouds. However, to confirm their potential in quantitative passive imaging, further investigations needs to be conducted, with a particular attention on the choice of the regularization parameters, that can lead to underestimation of the true power value.

Code Availability and Supplemental Material

The code associated to the developed method, as well as some example and the video of the experiment are available at https://github.com/creatis-ULTIM/CMF_sparseTV.

Appendix

This appendix develops how operator A (see eq. ((11))) is obtained from the steering vectors. If we vectorize $\mathbf{C}$, it gives:

$$\mathbf{c} = \frac{1}{N^2}\begin{bmatrix} x_1 h(1,1)h^*(1,1) + \cdots + x_M h(1,M)h^*(1,M) \\ \vdots \\ x_1 h(1,1)h^*(N,1) + \cdots + x_M h(1,M)h^*(N,M) \\ \vdots \\ x_1 h(N,1)h^*(1,1) + \cdots + x_M h(N,M)h^*(1,M) \\ \vdots \\ x_1 h(N,1)h^*(N,1) + \cdots + x_M h(N,M)h^*(N,M) \end{bmatrix} \tag{27}$$

We can write:

$$\mathbf{c} = \frac{1}{N^2}\begin{bmatrix} h(1,1)h^*(1,1) & \cdots & h(1,M)h^*(1,M) \\ \vdots & \ddots & \vdots \\ h(1,1)h^*(N,1) & \cdots & h(1,M)h^*(N,M) \\ \vdots & \ddots & \vdots \\ h(N,1)h^*(1,1) & \cdots & h(N,M)h^*(1,M) \\ \vdots & \ddots & \vdots \\ h(N,1)h^*(N,1) & \cdots & h(N,M)h^*(N,M) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_M \end{bmatrix} \tag{28}$$

Therefore:

$$\mathbf{A} = \frac{1}{N^2}\begin{bmatrix} h(1,1)h^*(1,1) & \cdots & h(1,M)h^*(1,M) \\ \vdots & \ddots & \vdots \\ h(1,1)h^*(N,1) & \cdots & h(1,M)h^*(N,M) \\ \vdots & \ddots & \vdots \\ h(N,1)h^*(1,1) & \cdots & h(N,M)h^*(1,M) \\ \vdots & \ddots & \vdots \\ h(N,1)h^*(N,1) & \cdots & h(N,M)h^*(N,M) \end{bmatrix} \tag{29}$$

References

[1] D. L. Miller, N. B. Smith, M. R. Bailey, G. J. Czarnota, K. Hynynen, I. R. S. Makin, and B. C. of the American Institute of Ultrasound in Medicine, "Overview of therapeutic ultrasound applications and safety considerations," *Journal of ultrasound in medicine*, vol. 31, no. 4, pp. 623–634, 2012.

[2] M. R. Bailey, L. A. Crum, O. A. Sapozhnikov, A. P. Evan, J. A. McAteer, T. Colonius, and R. O. Cleveland, "Cavitation in shock wave lithotripsy," *Journal of the Acoustical Society of America*, vol. 114, no. 4, pp. 2417–2417, 2003.

[3] S. Zhu, F. H. Cocks, G. M. Preminger, and P. Zhong, "The role of stress waves and cavitation in stone comminution in shock wave lithotripsy," *Ultrasound in medicine & biology*, vol. 28, no. 5, pp. 661–671, 2002.

[4] D. L. Miller, S. V. Pislaru, and J. F. Greenleaf, "Sonoporation: mechanical dna delivery by ultrasonic cavitation," *Somatic cell and molecular genetics*, vol. 27, pp. 115–134, 2002.

[5] C.-D. Ohl, M. Arora, R. Ikink, N. De Jong, M. Versluis, M. Delius, and D. Lohse, "Sonoporation from jetting cavitation bubbles," *Biophysical journal*, vol. 91, no. 11, pp. 4285–4295, 2006.

[6] S. Paliwal and S. Mitragotri, "Ultrasound-induced cavitation: applications in drug and gene delivery," *Expert opinion on drug delivery*, vol. 3, no. 6, pp. 713–726, 2006.

[7] C. D. Arvanitis, M. Bazan-Peregrino, B. Rifai, L. W. Seymour, and C. C. Coussios, "Cavitation-enhanced extravasation for drug delivery," *Ultrasound in medicine & biology*, vol. 37, no. 11, pp. 1838–1852, 2011.

[8] E. Stride and C. Coussios, "Nucleation, mapping and control of cavitation for drug delivery," *Nature Reviews Physics*, vol. 1, no. 8, pp. 495–509, 2019.

[9] N. McDannold, N. Vykhodtseva, and K. Hynynen, "Targeted disruption of the blood–brain barrier with focused ultrasound: association with cavitation activity," *Physics in Medicine & Biology*, vol. 51, no. 4, p. 793, 2006.

[10] T. Sun, G. Samiotaki, S. Wang, C. Acosta, C. C. Chen, and E. E. Konofagou, "Acoustic cavitation-based monitoring of the reversibility and permeability of ultrasound-induced blood-brain barrier opening," *Physics in Medicine & Biology*, vol. 60, no. 23, p. 9079, 2015.

[11] K. B. Bader, K. J. Haworth, H. Shekhar, A. D. Maxwell, T. Peng, D. D. McPherson, and C. K. Holland, "Efficacy of histotripsy combined with rt-pa in vitro," *Physics in Medicine & Biology*, vol. 61, no. 14, p. 5253, 2016.

[12] A. D. Maxwell, C. A. Cain, A. P. Duryea, L. Yuan, H. S. Gurm, and Z. Xu, "Noninvasive thrombolysis using pulsed ultrasound cavitation therapy–histotripsy," *Ultrasound in medicine & biology*, vol. 35, no. 12, pp. 1982–1994, 2009.

[13] A. D. Maxwell, T.-Y. Wang, C. A. Cain, J. B. Fowlkes, O. A. Sapozhnikov, M. R. Bailey, and Z. Xu, "Cavitation clouds created by shock scattering from bubbles during histotripsy," *The Journal of the Acoustical Society of America*, vol. 130, no. 4, pp. 1888–1898, 2011.

[14] S. J. Norton and I. Won, "Time exposure acoustics," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 3, pp. 1337–1343, 2000.

[15] S. Norton, I. Won, A. Witten, A. Oren, F. Funak, and G. L. R. NC, "Time-exposure acoustics for imaging underground structures," tech. rep., 2003.

[16] M. Gyongy and C.-C. Coussios, "Passive spatial mapping of inertial cavitation during hifu exposure," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 1, pp. 48–56, 2009.

[17] P. Stoica, Z. Wang, and J. Li, "Robust capon beamforming," in *Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers, 2002.*, vol. 1, pp. 876–880, IEEE, 2002.

[18] J. Li, P. Stoica, and Z. Wang, "On robust capon beamforming and diagonal loading," *IEEE transactions on signal processing*, vol. 51, no. 7, pp. 1702–1715, 2003.

[19] C. M. Coviello, S. R. Faragher, and C.-C. Coussios, "Robust capon beamforming for passive cavitation mapping during high-intensity focused ultrasound therapy.," *The Journal of the Acoustical Society of America*, vol. 128, no. 4_Supplement, pp. 2280–2280, 2010.

[20] C. Coviello, R. Kozick, J. Choi, M. Gyöngy, C. Jensen, P. P. Smith, and C.-C. Coussios, "Passive acoustic mapping utilizing optimal beamforming in ultrasound therapy monitoring," *The Journal of the Acoustical Society of America*, vol. 137, no. 5, pp. 2573–2585, 2015.

[21] P. Boulos, F. Varray, A. Poizat, A. Ramalli, B. Gilles, J.-C. Bera, and C. Cachard, "Weighting the passive acoustic mapping technique with the phase coherence factor for passive ultrasound imaging of ultrasound-induced cavitation," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 65, no. 12, pp. 2301–2310, 2018.

[22] M. Polichetti, F. Varray, J.-C. Béra, C. Cachard, and B. Nicolas, "Advanced beamforming techniques for passive imaging of stable and inertial cavitation," in *2018 IEEE International Ultrasonics Symposium (IUS)*, pp. 1–4, IEEE, 2018.

[23] R. Merino-Martínez, P. Sijtsma, M. Snellen, T. Ahlefeldt, J. Antoni, C. J. Bahr, D. Blacodon, D. Ernst, A. Finez, S. Funke, *et al.*, "A review of acoustic imaging methods using phased microphone arrays: Part of the "aircraft noise generation and assessment" special issue," *CEAS Aeronautical Journal*, vol. 10, pp. 197–230, 2019.

[24] K. J. Haworth, K. B. Bader, K. T. Rich, C. K. Holland, and T. D. Mast, "Quantitative frequency-domain passive cavitation imaging," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 64, no. 1, pp. 177–191, 2016.

[25] S. Lu, H. Hu, X. Yu, J. Long, B. Jing, Y. Zong, and M. Wan, "Passive acoustic mapping of cavitation using eigenspace-based robust capon beamformer in ultrasound therapy," *Ultrasonics sonochemistry*, vol. 41, pp. 670–679, 2018.

[26] M. Polichetti, F. Varray, B. Gilles, J.-C. Béra, and B. Nicolas, "Use of the cross-spectral density matrix for enhanced passive ultrasound imaging of cavitation," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 68, no. 4, pp. 910–925, 2020.

[27] L. Du, T. Yardibi, J. Li, and P. Stoica, "Review of user parameter-free robust adaptive beamforming algorithms," *Digital Signal Processing*, vol. 19, no. 4, pp. 567–582, 2009.

[28] A. Sivadon, F. Varray, J.-C. Béra, B. Nicolas, and B. Gilles, "3d passive cavitation imaging using adaptive beamforming and matrix array transducer with random apodization," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 71, no. 2, pp. 238–254, 2024.

[29] P. Stoica, R. L. Moses, *et al.*, *Spectral analysis of signals*, vol. 452. Pearson Prentice Hall Upper Saddle River, NJ, 2005.

[30] C. D. Arvanitis, C. Crake, N. McDannold, and G. T. Clement, "Passive acoustic mapping with the angular spectrum method," *IEEE transactions on medical imaging*, vol. 36, no. 4, pp. 983–993, 2016.

[31] M. Li, J. Gu, T. Vu, G. Sankin, P. Zhong, J. Yao, and Y. Jing, "Time-resolved passive cavitation mapping using the transient angular spectrum approach," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 68, no. 7, pp. 2361–2369, 2021.

[32] T. Yardibi, J. Li, P. Stoica, and L. N. Cattafesta III, "Sparsity constrained deconvolution approaches for acoustic source mapping," *The Journal of the Acoustical Society of America*, vol. 123, no. 5, pp. 2631–2642, 2008.

[33] T. F. Brooks and W. M. Humphreys, "A deconvolution approach for the mapping of acoustic sources (damas) determined from phased microphone arrays," *Journal of sound and vibration*, vol. 294, no. 4-5, pp. 856–879, 2006.

[34] C. Lachambre, A. Basarab, J.-C. Bera, B. Nicolas, F. Varray, and B. Gilles, "Cross-spectral matrix fitting for passive mapping of the ultrasonic cavitation based on elastic-net regularization," in *IEEE International Ultrasonics Symposium (IUS)*, pp. 1–4, 2023.

[35] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 67, no. 2, pp. 301–320, 2005.

[36] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 58, no. 1, pp. 267–288, 1996.

[37] S. Schlunk, K. Dei, and B. Byram, "Iterative model-based beamforming for high dynamic range applications," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 68, no. 3, pp. 482–493, 2020.

[38] T. Szasz, A. Basarab, M.-F. Vaida, and D. Kouamé, "Elastic-net based beamforming in medical ultrasound imaging," in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pp. 477–480, IEEE, 2016.

[39] B. Wahlberg, S. Boyd, M. Annergren, and Y. Wang, "An admm algorithm for a class of total variation regularized estimation problems," *IFAC Proceedings Volumes*, vol. 45, no. 16, pp. 83–88, 2012.

[40] Z. Chu, C. Chen, Y. Yang, and L. Shen, "Two-dimensional total variation norm constrained deconvolution beamforming algorithm for acoustic source identification," *IEEE Access*, vol. 6, pp. 43743–43748, 2018.

[41] L. Yu, Z. Gong, N. Chu, Y. Ning, Y. Zheng, and P. Hou, "Adaptive imaging of sound source based on total variation prior and a subspace iteration integrated variational bayesian method," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–17, 2021.

[42] J. J. Fuchs, "More on sparse representations in arbitrary bases," *IFAC Proceedings Volumes*, vol. 36, no. 16, pp. 1315–1320, 2003.

[43] R. J. Tibshirani, *The solution path of the generalized lasso*. Stanford University, 2011.

[44] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011. Chapter 3.1.

[45] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011. Chapter 6.4.

[46] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," *Fixed-point algorithms for inverse problems in science and engineering*, pp. 185–212, 2011.

[47] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011. Chapter 3.3.

[48] W. Lauterborn and T. Kurz, "Physics of bubble oscillations," *Reports on progress in physics*, vol. 73, no. 10, p. 106501, 2010.