Abstract—Ultrafast ultrasound (US) imaging based on plane wave (PW) insonification is a widely used modality nowadays. Two main types of approaches have been proposed for image reconstruction either based on classical delay-and-sum (DAS) or on Fourier reconstruction. Using a single PW, these methods lead to a lower image quality than DAS with multi-focused beams. In this paper we review recent beamforming approaches based on sparse regularization methods. The imaging problem, either spatial-based (DAS) or Fourier-based, is formulated as a linear inverse problem and convex optimization algorithms coupled with sparsity priors are used to solve the ill-posed problem. We describe two applications of the framework namely the sparse inversion of the beamforming problem and the compressed beamforming in which the framework is combined with sparse regularization methods. Based on numerical simulations and experimental studies, we show the advantage of the proposed methods in terms of image quality compared to classical methods.

Index Terms—Ultrasound, plane wave imaging, sparsity, compressed sensing

I. INTRODUCTION

In the last years, ultrafast ultrasound imaging (US) based on plane wave (PW) insonification has become a widely used modality in the US community. Indeed, while conventional beamforming methods rely on a sequential transmission of a number of focused waves equal to the number of scan lines, thus limiting the frame rate to several tens of frames per second, PW imaging only needs few insonifications to reconstruct an image, enabling to reach far higher frame rates, around thousands of frames per second, and opening a whole range of applications [1].

The ability to perform ultrafast imaging is closely linked to the possibility to implement efficient beamforming methods. Nowadays, two main approaches have been developed namely spatial-based approaches [2] where the image is reconstructed using delay-and-sum (DAS) technique, and Fourier-based approaches [3]–[5] where the Fourier spectrum of the received raw data is used to reconstruct the final image.

The use of few PW insonifications usually leads to a lower image quality than classical DAS approaches with multi-focused beams [2]. Indeed, instead of being focused to the region of interest, the energy carried by a PW is spread over all the medium which induces a decrease of the signal to noise ratio. Moreover, the information carried by a PW allows to retrieve partial information of the final image which makes the beamforming problem ill-posed. In addition, since all the beamforming methods are based on numerical approximations of continuous expressions, they implicitly make use of gridding and interpolation which lead to measurement inaccuracies and thus to a decrease of the image quality [6].

Compressed sensing (CS) is a mathematical framework to recover signals from incomplete information [7] that has been successfully applied to medical US imaging (see [8], [9] and references therein). In this paper we propose a general framework based on reformulating the beamforming process, either Fourier-based or spatial-based, as an inverse problem which is solved by exploiting sparsity of US images in a redundant dictionary. We detail two main applications of the proposed framework, only differing in the measurement model. The first one is the use of sparse regularization for image quality enhancement [8]. The second one is a compressed beamforming approach in which the desired image is retrieved from only few samples of the raw data.

The paper is organized as follows. Section II reviews the existing beamforming approaches. Section III describes the proposed framework and details a general formulation of the inverse problem. Sections IV and V illustrate the two applications of the proposed framework and show the benefits of the proposed approach through numerical simulations and in vivo experiments.

II. BACKGROUND

After a PW insonification, the received signal \( r(x_i, t) \) at the US probe (linear array of transducers), in a transducer at position \( x_i \) and at time \( t \), the so called raw data, consists of backscattered echoes from the medium. In order to be exploited, the positions of the inhomogeneities are inferred from these echoes in the process called beamforming. The signal obtained after beamforming is usually denoted as radio frequency (RF) signal and will be designated by \( s(x, z) \) in the following sections. The conventions used in the next sections are detailed in Figure 1.

This work was supported in part by the UltrasoundToGo RTD project (no. 20NA21 145911), funded by Nano-Tera.ch. This work was also performed within the framework of the LABEX PRIMES (ANR-11-LABX-0063) of Universite de Lyon.

©2016 IEEE
with angle \( \theta_t \), the raw data \( r(x_i, t) \) can be obtained using the following relationship [12]:

\[
r(x_i, t) = \int \int_{(x,z) \in \Omega} s(x, z) \, dx \, dz,
\]

where \( \Omega = \{(x, z) | (ct - z \cos(\theta_t) - z \sin(\theta_t))^2 - (x - x_i)^2 = 0\} \). If we consider a PW with normal incidence (\( \theta_t = 0 \)), then the integral (3) can be modified into the following equation:

\[
r(x_i, t) = \int_{\alpha \in \mathbb{R}} s(x(\alpha), z(\alpha)) \sqrt{1 + \frac{(\alpha - x_i)^2}{(ct)^2}} \, d\alpha,
\]

with \( x(\alpha) = \alpha \) and \( z(\alpha) = \frac{1}{ct} \left((ct)^2 - (\alpha - x_i)^2\right) \).

III. GENERAL FRAMEWORK

A. The inverse problem

Sparse regularization methods mainly rely on two pillars, namely the formulation of the considered problem as an inverse problem and the compressibility of the desired images in a given model. Formally, we firstly derive a linear measurement model \( H \in \mathbb{R}^{N \times M} \) such that:

\[
r = Hs + n
\]

with \( r \in \mathbb{R}^{N} \) the discretized raw data, \( s \in \mathbb{R}^{M} \) the RF image and \( n \in \mathbb{R}^{N} \) the measurement noise, and we secondly provide a model \( \Psi \in \mathbb{R}^{M \times P} \) such that \( s \) has a sparse decomposition \( \gamma \in \mathbb{R}^{P} \) on \( \Psi \), i.e. such that \( s = \Psi \gamma \) and \( \gamma \) has few non-zero elements. We retrieve \( s \) from \( r \) by solving the following convex problem (synthesis formulation):

\[
\min_{\bar{s} \in \mathbb{C}^{N}} \|\Psi^{\dagger}\bar{s}\|_1 \text{ subject to } \|r - H\bar{s}\|_2 \leq \epsilon,
\]

where \( \| \cdot \|_1 \) denotes the \( \ell_1 \)-norm, \( \Psi^{\dagger} \) denotes the adjoint operator of \( \Psi \) and \( \bar{s} \) designates the RF image. The choice of the \( \ell_1 \)-norm is a convex relaxation to promote sparsity. While several methods may be applied to automatically identify the best value of the regularization parameter \( \epsilon \), the choice of the optimal parameters is out of the scope of the paper. In the study, \( \epsilon \) is considered to be empirically tuned by the user based on visual quality.

B. The sparsifying model

In this paper, the average sparsity model proposed in [13] is used. This model has been previously studied in the context of US images in [8]. The dictionary, composed of a concatenation of several frames, enables to better capture image structures that are often sparse in several frames, thus leading to improved image reconstructions compared to single frame models.

In this study, the dictionary is composed of the concatenation of Daubechies wavelet bases from Daubechies 1 (Db1) to Daubechies 8 (Db8). Thus,

\[
\Psi = \frac{1}{\sqrt{q}}[\Psi_1, ..., \Psi_q]
\]

where \( q = 8 \) and \( \Psi_i \) denotes \( i \)-th Daubechies wavelet. Db1 is the Haar basis promoting piece-wise smooth signals while Db2
to Db8 provide smoother sparse decompositions. The sparsity prior used to promote average sparsity is thus:

\[ \| \Psi^\dagger s \|_1 = \sum_{i=1}^{q} \| \Psi_i^\dagger s \|_1. \]

IV. SPARSE REGULARIZATION FOR IMAGE QUALITY ENHANCEMENT

A. Measurement operators

The first application of the framework described in Section III aims at enhancing the quality of the reconstruction for PW imaging. In order to derive the measurement models associated with the methods described in Section II, we firstly introduce the following gridding of the continuous image space:

\[
\begin{align*}
 x & \in \{ (m-1) \frac{L}{N_x}, \forall m \in \{ -\frac{N_x}{2} + 1, ..., \frac{N_x}{2} \} \} \\
 z & \in \{ (l-1) \frac{N_x}{Z_{\text{max}}}, \forall l \in \{ 1, ..., N_Z \} \}
\end{align*}
\]

(7)

with \( L \) the width of the probe, \( N_x \) the number of image samples in the lateral direction, \( N_Z \) the number of image samples in the axial direction and \( Z_{\text{max}} \) the maximum depth. The gridding of the raw data is imposed by the number of transducers in the lateral direction and by the sampling frequency in the axial direction according to the following equations:

\[
\begin{align*}
 x_i & = \{ (m-1) p_t - \frac{N_x}{2} p_t, \forall m \in \{ 1, ..., N_1 \} \} \\
 t & = \{ (l-1) \frac{N_x}{f_s}, \forall l \in \{ 1, ..., N_r \} \}
\end{align*}
\]

(8)

with \( N_r \) the number of raw data samples in the axial direction, \( f_s \) the sampling frequency, \( p_t \) the probe’s pitch and \( N_1 \) the number of transducers in the probe. We derive the two matrices \( S = (S_{kl})_{k \in \{ 1, ..., N_2 \}, l \in \{ 1, ..., N_1 \}} \) and \( R = (R_{ij})_{i \in \{ 1, ..., N_r \}, j \in \{ 1, ..., N_1 \}} \) which are respectively the discretization of \( s \) and \( r \) on the grids (7) and (8) respectively. We define \( r \) and \( s \) as vectorized versions of \( R \) and \( S \) obtained by concatenating the columns of the matrices.

1) Fourier-based model: From equation (7), the corresponding grid of the image k-space can be deduced:

\[
\begin{align*}
 k_x & \in \{ \frac{2\pi (m-1)}{L} \}, \forall m \in \{ -\frac{N_x}{2} + 1, ..., \frac{N_x}{2} \} \\
 k_z & \in \{ \frac{2\pi (l-1)}{Z_{\text{max}}} \}, \forall l \in \{ 1, ..., N_Z \}
\end{align*}
\]

(9)

The same reasoning can be applied to discretize the k-space grid of the raw data \( (k_x, k) \). Given the discretized RF image \( s \), the k-space representation \( y \) of the raw data on the grid \( (k_x, k) \) is obtained by applying a composition of a 2D discrete Fourier transform (DFT) \( F_S \) on the grid \( (k_x', k_z') \) followed by an interpolation on \( (k_x, k) \). This operator, known as the non-uniform fast Fourier transform (NUFFT) [14], can be written as \( GF_S \) in which \( G \) is a matrix implementing the convolutional interpolation kernel. Then, \( r \) is obtained from \( y \) by a 2D inverse DFT \( F^\dagger_k \) on the grid \( (k_x, k) \) and the following equation holds:

\[ r = F^\dagger_k GF_S s + n = H_F s + n \]

(10)

with \( n \in \mathbb{R}^N \) representing the noise induced by the model inaccuracies and \( H_F \) the measurement model. It has to be noted that \( H_F \) is ill-posed since the frequency content of the raw data only allows to recover partial information in the k-space of the final image and vice-versa.

2) Spatial-based model: The spatial-based model is obtained by discretizing equation (4) using the grids described in equations (7) and (8). In order to do so, we firstly discretize the continuous variable \( \alpha \) by introducing the following vector:

\[
\begin{align*}
 \alpha & \in \mathbb{R}^J \\
 -\frac{L}{2} & \leq \alpha_1 < \alpha_2 < ... < \alpha_J \leq \frac{L}{2}
\end{align*}
\]

with \( J \in \mathbb{N} \). We thus have:

\[ R_{ij} = \sum_{k=1}^{J} s(x(\alpha_k), z(\alpha_k)) C(\alpha_k), \]

with \( C(\alpha_k) = \sqrt{1 + \frac{(\alpha_k - x_i)^2}{(at)^2}} \).

In a second step, we compute each \( s(x(\alpha_k), z(\alpha_k)) \) on the grid (7) using an interpolation with kernel \( K = (K_{ij})_{i \in \{ 1, ..., N_r \}, j \in \{ 1, ..., N_1 \}} \). The following equation holds:

\[ R_{ij} = \sum_{k=1}^{J} C(\alpha_k) \sum_{p=1}^{N_p} \sum_{l=1}^{N_l} K_{ijl} s_{p_{\alpha_k} + p - \frac{N_p}{2}, l_{\alpha_k} + l - \frac{N_l}{2}} \]

(11)

where \( [\cdot] \) denotes the floor function and \( (p_{\alpha_k}, l_{\alpha_k}) \) is the closest point to \( (x(\alpha_k), z(\alpha_k)) \) lying on the grid. In the following, we assume that \( \alpha = x \) and that \( K \) is a linear interpolation in the axial direction and can be written as \( K = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \). More elaborated quadrature rules (Simpson, Gauss-Legendre) and interpolation kernels may be used in (11) which should lead to better image quality. Following equation (11), \( R \) and \( S \) are related by a 4D–matrix \( H_S = (H_{S_{ijkl}}) \). \( H_S \) is then reduced to a 2D matrix by vectorizing both \( S \) and \( R \) and the following inverse problem is obtained:

\[ r = H_S s + n. \]

(12)

As for Fourier-based beamforming, it should be noticed that \( H_S \) is ill-posed since the raw data carry only partial information on the desired image and vice-versa. In equations (10) and (12), \( s \) is retrieved from \( r \) by solving problem (6).

B. Experiments

The discretization of the continuous NUFT in Fourier-based beamforming induces measurement inaccuracies and leads to a decrease of the image quality (contrast) [8]. The usual way to address this problem is by performing a zero-padding in the k-space of the echoes which improves the accuracy of the interpolation performed in the NUFT [6]. However, such method induces a non-negligible increase of the computational complexity and the gain in terms of image quality remains limited. Moreover, the measurement operator being ill-posed due to the partial information carried in the k-space of the raw data, the adjoint of \( H_F \) does not lead to the optimal reconstruction. For spatial-based method, the quality of the reconstruction is directly linked to the pitch and sampling frequency of the probe. As for Fourier-based
approach, since the raw data only contain partial information
(not for all the points lying in $\Omega$), the operator $H_S$ is ill-
posing. In order to achieve reasonable image reconstruction,
the axial direction is usually sampled at far higher frequency
than Nyquist frequency leading to both an higher data rate and
a bigger amount of delays to be calculated [15].

In order to show the benefits of the proposed method in
terms of image quality, a numerical study of the contrast is
performed. A standard linear probe whose settings are given in
Table I is implemented using Field II [16], [17]. Constant
speed of sound is assumed (1540 m.s$^{-1}$).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of elements ($N_r$)</td>
<td>128</td>
</tr>
<tr>
<td>Center frequency ($f_0$)</td>
<td>5 MHz</td>
</tr>
<tr>
<td>Wavelength ($\lambda$)</td>
<td>0.31 mm</td>
</tr>
<tr>
<td>Sampling frequency ($f_s$)</td>
<td>25 MHz</td>
</tr>
<tr>
<td>Pitch</td>
<td>0.1953 mm</td>
</tr>
<tr>
<td>Kerf</td>
<td>0.05 mm</td>
</tr>
</tbody>
</table>

**TABLE I: Probe characteristics.**

An anechoic cyst which is composed of a 8-mm diameter
anechoic occlusion at 4 cm depth embedded in a medium
with high density of scatterers (20 per resolution cell) is
insonified with one PW with normal incidence. No apodization
is used neither at transmission nor at reception. The RF image
is reconstructed with both the classical methods described in
Section II and the proposed framework. The envelope
image is extracted from the reconstructed RF image, gamma-
compressed using $\gamma = 0.3$ and finally converted to 8-bit gray
scale to get the B-mode image. The contrast metric defined in
the following equation is computed on the B-mode image [18]:

$$CR = 20 \log_{10} \frac{|\mu_t - \mu_b|}{\sqrt{\sigma_t^2 + \sigma_b^2}}$$  \hspace{1cm} (13)

where $\mu_t$ and $\mu_b$ ($\sigma_t^2, \sigma_b^2$) are the means (variances) of
respectively the target and the background. Table II displays the
contrast values of the classical method without upsampling
(column 1), with upsampling (column 2) and of the proposed
framework without upsampling (column 3). The upsampling
consists in a zero-padding of a factor 2 in the axial direction
and a factor 1.5 in the lateral direction for the Fourier-based
methods and an upsampling of a factor 4 compared to the
Nyquist frequency for the spatial-based approach. While the
benefit of the zero-padding on the Fourier-based approaches
has already been demonstrated [6] and is visible in Table II, the
increase of the sampling frequency is not so beneficial for the
spatial-based approach. Indeed, a higher sampling frequency
leads to a better estimation of the delays. However, since an
interpolation is already performed to improve the estimation of
the delays, such a high sampling frequency is not required. The
results show an increase of the contrast for the proposed frame-
work, for both Fourier-based and spatial-based approaches.
Since the contrast is measured in the anechoic area, it is
directly linked to the amount of noise in the reconstructed
image and since the images are obtained through simulation,
the only source of noise is induced by the measurement model
(gridding). This decrease can be observed in the anechoic area
of the B-mode images displayed on Figure 2.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Classic</th>
<th>Classic upsam.</th>
<th>Sparse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fourier-based approach</td>
<td>5.59</td>
<td>7.29</td>
<td>9.62</td>
</tr>
<tr>
<td>Spatial-based approach</td>
<td>7.04</td>
<td>7.48</td>
<td>9.10</td>
</tr>
</tbody>
</table>

**TABLE II: Contrast values (in dB) for the classical methods
and the sparse regularization framework.** The Fourier-based
approach used in the study is the Lu method [3].

Fig. 2: B-mode image of the anechoic phantom obtained
with 1 PW insonification and reconstructed with (a) classical
Lu method ($CR = 5.58$ dB), (b) classical Lu method with
upsampling ($CR = 7.29$ dB) and (c) Lu method with the
proposed framework ($CR = 9.62$ dB) without upsampling.

**V. COMPRESSED BEAMFORMING**

**A. Measurement operator**

The models described in equations (10) and (12) are
suitable to CS-based methods since they are formulated as
inverse problems. Formally, we consider an undersampled
measurement vector $r_u \in \mathbb{R}^P$ with $P \ll N$ and the
corresponding projection operator $P \in \mathbb{R}^{P \times N}$ such that,
$p_{ij} \in \{0,1\}, \forall (i,j) \in \{1,...,P\} \times \{1,...,N\}$ and $r_u = Pr$.
Retrieving $s$ given $r_u$ poses the inverse problem defined in
the following equation:

$$r_u = P(Hs + n) = H_n s + n_u$$  \hspace{1cm} (14)

with $H_n = PH \in \mathbb{R}^{P \times N}$, $n_u = Pn$ and $H$ is either $H_S$ or $H_F$
depending on the chosen approach. $s$ is retrieved from $r_u$ by
solving problem (6).

It can be noticed that problem (14) is close to the one
described in [12] with the difference that the inverse scattering
problem (not the beamforming) is formulated using Green’s
functions in the above mentioned paper.

**B. Experiments**

In the following section, the undersampling operator $P$
is designed by randomly selecting $N_c$ transducers at reception,
with $N_c < N_t$. This non-uniform spacing has proven to be
suited for CS in radar imaging [19].

The proposed method is evaluated on *in vivo* carotid images
acquired with a Verasonics ultrasound scanner (Redmond, WA,
USA) with a L12-5-50mm Verasonics linear probe whose
settings are given on Table I. The carotid is insonified with
one PW with normal incidence. No apodization is used neither
in transmit nor in receive. The raw data are undersampled according to the scheme defined above and the RF image is reconstructed using the different methods. The envelope image is extracted from the reconstructed RF image, gamma-compressed using $\gamma = 0.3$ and finally converted to 8-bit gray scale to get the B-mode image. The quality of the reconstruction is evaluated by the structural similarity index (SSIM) [20] and the normalized root mean square error (NRMSE) computed on the normalized envelope image [15]. The reference images are chosen as the one obtained with the classical method with full data. Table III displays the results. As expected the higher the number of transducers, the better the quality of the reconstruction. It can also be noticed that the quality of the reconstruction is comparable between both approaches. It seems that it is slightly better with the Fourier-based method. However, this increase is not visible on the B-mode images, displayed on Figure 3. Both speckle patterns, necessary for tracking tools, as well as thickness of the carotid artery plaque are well preserved with the proposed framework.

<table>
<thead>
<tr>
<th>SSIM</th>
<th>30</th>
<th>64</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fourier-based approach</td>
<td>0.78</td>
<td>0.86</td>
<td>0.92</td>
</tr>
<tr>
<td>Spatial-based approach</td>
<td>0.70</td>
<td>0.78</td>
<td>0.82</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NRMSE</th>
<th>30</th>
<th>64</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fourier-based approach</td>
<td>0.21</td>
<td>0.49</td>
<td>0.69</td>
</tr>
<tr>
<td>Spatial-based approach</td>
<td>0.25</td>
<td>0.43</td>
<td>0.55</td>
</tr>
</tbody>
</table>

TABLE III: (a) SSIM and (b) NRMSE values for the carotid images acquired with 30, 64 and 90 transducers (among 128) randomly chosen across the aperture.

![B-mode image of the carotid obtained with 1 PW insolation](image)

Fig. 3: B-mode image of the carotid obtained with 1 PW insolation and (a) 38 transducers and CS with Lu method, (b) 38 transducers and CS with spatial-based method and (c) 128 transducers and classic spatial-based approach.

VI. CONCLUSION

In this paper, we propose a beamforming framework for US imaging. In this framework, the beamforming process is expressed as an inverse problem and solved using a sparsity prior in a redundant dictionary. Two main applications follow on from this formulation. First it enables to increase the image quality by removing measurement artifacts induced by the gridding operation. Secondly, it is suitable to the CS framework and enables the reconstruction of high quality images from undersampled raw data acquired using only a few transducers. By alleviating data rates for acquisition, it opens the door to a whole range of applications and especially to 3D imaging where the amount of data to process remains a major limitation nowadays.

REFERENCES


