FAST NEAR-FIELD ACOUSTIC IMAGING WITH SEPARABLE ARRAYS

Flávio P. Ribeiro, Vítor H. Nascimento

Electronic Systems Engineering Dept., Universidade de São Paulo
{fr,vitor}@lps.usp.br

ABSTRACT

Acoustic imaging is a computationally intensive and ill-conditioned inverse problem, which involves estimating high resolution source distributions with large microphone arrays. We have recently shown how to significantly accelerate acoustic imaging under a far-field approximation using fast transforms. This paper generalizes our previous work to obtain computationally efficient and accurate transforms for near-field imaging with separable arrays. We show that under a suitable permutation, the imaging transform can be made nearly separable, even when modeling strong near-field effects. We exploit this quasi-separability to obtain a computationally efficient and accurate low-rank representation, allowing the design of fast transforms for near-field operation with arbitrary focal surfaces and arbitrary accuracy. We combine these transforms with calibration matrices, which compensate non-separable characteristics and allow one to quickly reshape focal surfaces without having to recompute the optimal transforms.

Index Terms— array processing, fast transform, acoustic imaging, low-rank approximation, kronecker approximation.

1. INTRODUCTION

Acoustic imaging with microphone arrays has become a standard tool for studying aeroacoustic sources. It is routinely used to measure the noise generated by engines, turbines, vehicles and aircraft for aerodynamic design and noise reduction purposes [1].

Due to its relatively low computational cost, beamforming remains a popular method for acoustic imaging. Unfortunately, it produces the source distribution of interest convolved with the array point spread function. Deconvolution algorithms [2] have been proposed to enhance beamformed images. More recently, regularized covariance fitting techniques [3] have been shown to deliver even better results. However, in their original formulations, these methods have high computational costs, because they have no means of efficiently transforming back and forth between the image under reconstruction and the measured data.

Motivated by this observation, we proposed a fast transform designed to enable computationally efficient and accurate imaging with separable[1] planar arrays [4]. Using this transform, one can perform deconvolutions an order of magnitude faster than with FFT-based methods. In related work [5], we use such transforms to recast acoustic imaging as regularized least-squares covariance fitting. While these formulations would be ordinarily intractable, by using fast transforms we are able to accelerate their solution by many orders of magnitude. Indeed, this approach delivers more accurate results than competing FFT-based deconvolution methods, and in a smaller amount of time.

To our knowledge, previous work for acoustic imaging involving fast transforms has always assumed sources located in the far-field of the array. Indeed, FFT-based deconvolution approaches such as DAMAS2 [2] require a shift-invariant point spread function, which in turn implies a far-field assumption.

Our transform for separable arrays was also derived under a far-field assumption [4]. Nevertheless, it does not impose any structure onto the array manifold vector other than its separability. Thus, we anticipated its use for near-field imaging, as long as one could produce suitable separable approximations to exact (non-separable) near-field manifold vectors.

In this paper we go a step further. We show that a suitable permutation and rearrangement of the transform matrix makes it easily approximable by a truncated series of Kronecker products, even when modeling strong near-field effects. This quasi-separability (in the Kronecker sense) is equivalent to the existence of a low-rank approximation, which we use to obtain a computationally efficient and accurate transform.

This paper is organized as follows: Section 2 provides definitions and reviews our fast transform for separable geometries. Section 3 generalizes the far-field transform by modeling spherical wavefronts and arbitrary focal surfaces. We show how to compute the optimal low-rank fast transform, and how to incorporate interpolation matrices for calibration and fast focusing. Section 4 features examples and Section 5 has our conclusions.

2. PRELIMINARIES

We use the superscripts $\cdot^T$, $\cdot^H$, and $\cdot^*$ to denote transposition, Hermitian transposition, and complex conjugation, respectively. The remainder of a/b, for $a, b \in \mathbb{Z}_+$ is written as $\mathrm{mod}\,(a, b)$. Round-off of $x \in \mathbb{R}$ towards $-\infty$ is denoted by $\lfloor x \rfloor$. Given suitably sized matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, we use $\mathbf{C} = \underline{\mathbf{A}}\,(\mathbf{B})$ to denote $\mathrm{vec}\,\{\mathbf{C}\} = \mathbf{A}\mathrm{vec}\,\{\mathbf{B}\}$.

Consider a planar array of N microphones with coordinates $\mathbf{p}_0$, ..., $\mathbf{p}_{N-1} \in \mathbb{R}^3$. Suppose the wave field of interest can be modeled by the superposition of M point sources with coordinates $\mathbf{q}_0, ..., \mathbf{q}_{M-1} \in \mathbb{R}^3$, where M is usually large. The $N \times 1$ array output for a frequency ω is modeled as

$$\mathbf{x}\,(\omega) = \mathbf{V}\,(\omega)\,\mathbf{f}\,(\omega) + \boldsymbol{\eta}\,(\omega)\,, \tag{1}$$

where $\mathbf{V}\,(\omega) = [\,\mathbf{v}\,(\mathbf{q}_0, \omega)\ \cdots\ \mathbf{v}\,(\mathbf{q}_{M-1}, \omega)\,]$ is the array manifold matrix, $\mathbf{f}\,(\omega) = [\,f_0\,(\omega)\ f_1\,(\omega)\ \cdots\ f_{M-1}\,(\omega)\,]^T$ is the frequency domain signal waveform and $\boldsymbol{\eta}\,(\omega)$ is uncorrelated noise.

The near-field array manifold vector models a spherical wavefront, and is given by

$$\mathbf{v}\,(\mathbf{q}_m, \omega) = \left[\ \frac{\mathrm{e}^{-j\left\{\frac{\omega}{c}\|\mathbf{p}_0 - \mathbf{q}_m\|\right\}}}{\|\mathbf{p}_0 - \mathbf{q}_m\|}\ \cdots\ \frac{\mathrm{e}^{-j\left\{\frac{\omega}{c}\left\|\mathbf{p}_{N-1} - \mathbf{q}_m\right\|\right\}}}{\left\|\mathbf{p}_{N-1} - \mathbf{q}_m\right\|}\ \right]^T. \tag{2}$$

Let $\mathbf{S}_{\mathbf{x}}\,(\omega) = \mathrm{E}\left\{\mathbf{x}\,(\omega)\,\mathbf{x}^H\,(\omega)\right\}$ be the array cross spectral matrix. If $\mathbf{x}_0\,(\omega)$, ..., $\mathbf{x}_{L-1}\,(\omega)$ correspond to L frequency domain snapshots, the spectral matrix can be estimated with

$$\mathbf{S}_{\mathbf{x}}\,(\omega) = \frac{1}{L}\sum_{l=0}^{L-1} \mathbf{x}_l\,(\omega)\,\mathbf{x}_l^H\,(\omega)\,. \tag{3}$$

[1] An array with manifold vector $\mathbf{v}\,(u_x, u_y)$ is said to be separable if there exist $\mathbf{a}\,(u_x)$ and $\mathbf{b}\,(u_y)$ such that $\mathbf{v}\,(u_x, u_y) = \mathbf{a}\,(u_x) \otimes \mathbf{b}\,(u_y)$ for all valid u_x, u_y, where $\otimes$ is the Kronecker product. For example, arrays with elements positioned over (not necessarily uniform) Cartesian grids are separable under a far-field assumption.

It is usually better to work with $\mathbf{S_x}(\omega)$ instead of directly with each $\mathbf{x}_l(\omega)$, because $\mathbf{S_x}(\omega)$ carries only the relative phase shifts between microphones and is the result of averaging, such that it has less noise content. We will assume narrow-band processing and omit the argument ω. To simplify the notation, $\mathbf{S_x}(\omega)$ will be written as $\mathbf{S}$.

Assuming that the noise is spatially white and uncorrelated with the sources of interest, we have

$$\mathbf{S} = \mathbf{V}\,\mathrm{E}\left\{\mathbf{f}\mathbf{f}^H\right\}\mathbf{V}^H + \sigma^2\mathbf{I}, \tag{4}$$

where $\sigma^2 = \mathrm{E}\{\eta_i\eta_i^*\}$, $0 \le i < N$. Assuming that the sources are uncorrelated (a nearly universal assumption for acoustic imaging) implies that $\mathrm{E}\left\{\mathbf{f}\mathbf{f}^H\right\}$ is diagonal.

For the purposes of acoustic imaging, the source coordinates $\{\mathbf{q}_i\}_{i=0}^{M-1}$ are chosen to discretize a surface of interest (the focal surface). We consider surfaces parameterized in two coordinates, such that $\mathbf{q}_i = \varphi(\mathbf{u}_i)$, for $\mathbf{u}_i \in \mathbb{R}^2$ and a parameterization φ. Furthermore, we discretize the focal surface using a Cartesian grid, such that $\{\mathbf{u}_i\}_{i=0}^{M-1} = \{u_{x_m}\}_{m=0}^{M_x-1} \times \{u_{y_n}\}_{n=0}^{M_y-1}$ and

$$\mathbf{u}_i = \left[\begin{array}{cc} u_{x_{\lfloor i/M_y \rfloor}} & u_{y_{\mathrm{mod}(i, M_y)}} \end{array}\right]^T. \tag{5}$$

Using this parameterization, the ideal acoustic image is given by $\mathbf{Y} \in \mathbb{R}^{M_y \times M_x}$ with $\mathbf{y} = \mathrm{vec}\{\mathbf{Y}\} = \left[\begin{array}{ccc} |f_0|^2 & \cdots & |f_{M-1}|^2 \end{array}\right]^T$. Thus, the pixel coordinates of $\mathbf{Y}$ correspond to source locations, and the pixel values correspond to source powers.

Note that $\mathrm{diag}\left[\mathrm{E}\left\{\mathbf{f}\mathbf{f}^H\right\}\right] = \mathbf{y}$, so we can rewrite (4) as

$$\mathbf{S} = \sum_{m=0}^{M-1} y_m \mathbf{v}\left(\varphi(\mathbf{u}_m)\right)\mathbf{v}^H\left(\varphi(\mathbf{u}_m)\right). \tag{6}$$

Many reconstruction algorithms iteratively compute an estimated image $\hat{\mathbf{Y}}$ and compare the corresponding $\hat{\mathbf{S}}$ obtained through (6) with the measured values obtained from (3). Unless the image and its iterates are very sparse, (6) becomes computationally intractable, motivating the need for fast transforms.

Let $\mathbf{s} = \mathrm{vec}\{\mathbf{S}\}$, $\mathbf{y} = \mathrm{vec}\{\mathbf{Y}\}$ and $\mathbf{v}_{\mathbf{u}_m} = \mathbf{v}\left(\varphi(\mathbf{u}_m)\right)$. Note that $\mathrm{vec}\left\{\mathbf{v}_{\mathbf{u}_m}\mathbf{v}_{\mathbf{u}_m}^H\right\} = \mathbf{v}_{\mathbf{u}_m}^* \otimes \mathbf{v}_{\mathbf{u}_m}$, where $\otimes$ is the Kronecker product. Thus, we can write (6) as (note that $\mathbf{A}$ below is $N^2 \times M$)

$$\mathbf{s} = \mathbf{A}\mathbf{y} = \left[\begin{array}{ccc} \mathbf{v}_{\mathbf{u}_0}^* \otimes \mathbf{v}_{\mathbf{u}_0} & \cdots & \mathbf{v}_{\mathbf{u}_{M-1}}^* \otimes \mathbf{v}_{\mathbf{u}_{M-1}} \end{array}\right]\mathbf{y}. \tag{7}$$

Let $N = N_x N_y$ be the number of array elements. In [4], we show that when the array manifold vector is separable such that $\mathbf{v}\left(\varphi(\mathbf{u})\right) = \mathbf{v}\left(\varphi(u_x, u_y)\right) = \mathbf{v}_x\left(\varphi_x(u_x)\right) \otimes \mathbf{v}_y\left(\varphi_y(u_y)\right)$, where $\mathbf{v}_x\left(\varphi_x(u_x)\right)$ and $\mathbf{v}_y\left(\varphi_y(u_y)\right)$ are $N_x \times 1$ and $N_y \times 1$ manifold vectors, the computation of (7) may be accelerated as follows. Let

$$\mathbf{S} = \left[\begin{array}{ccc} \mathbf{T}_{0,0} & \cdots & \mathbf{T}_{0,N_x-1} \\ \vdots & & \vdots \\ \mathbf{T}_{N_x-1,0} & \cdots & \mathbf{T}_{N_x-1,N_x-1} \end{array}\right], \quad \mathbf{T}_{i,j} \in \mathbb{C}^{N_y \times N_y}$$

$$\mathbf{Z} = \left[\,\mathbf{t}_{0,0}\ \mathbf{t}_{1,0}\ \ldots\ \mathbf{t}_{N_x-1,N_x-1}\,\right], \qquad \mathbf{t}_{i,j} = \mathrm{vec}\{\mathbf{T}_{i,j}\}.$$

Define $\mathbf{\Xi}$ such that $\mathrm{vec}\{\mathbf{S}\} = \mathbf{\Xi}\mathrm{vec}\{\mathbf{Z}\}$ (note that $\mathbf{\Xi}$ is a permutation). As shown in [4], the separability of $\mathbf{v}\left(\varphi(\mathbf{u})\right)$ implies that there exist $\mathbf{V_x} \in \mathbb{C}^{N_x^2 \times M_x}$ and $\mathbf{V_y} \in \mathbb{C}^{N_y^2 \times M_y}$ such that $\mathbf{A} = \mathbf{\Xi}\left(\mathbf{V_x} \otimes \mathbf{V_y}\right)$. Since $\left(\mathbf{A}^T \otimes \mathbf{B}\right)\mathrm{vec}\{\mathbf{C}\} = \mathrm{vec}\{\mathbf{BCA}\}$ whenever $\mathbf{BCA}$ is defined [6], $\mathbf{S} = \underline{\mathbf{A}}(\mathbf{Y})$ can be efficiently implemented as $\mathbf{S} = \underline{\mathbf{\Xi}}\left(\mathbf{V_y}\mathbf{Y}\mathbf{V_x}^T\right)$ (see Section 3 for the gain in computational cost). We remark that, under a far-field approximation, the manifold vector will be separable if the microphones are disposed in a (possibly nonuniform) Cartesian grid [4].

3. NEAR-FIELD IMAGING

3.1. Low-rank extension for non-separable arrays

Near-field manifold vectors are not in general separable. As we show in Section 4, in the absence of compensation, a (separable) far-field approximation can cause numerous artifacts. In this section we show how to approximate $\mathbf{A}$ using a sum of separable transforms, allowing us to dramatically improve reconstruction quality.

Recall that if $\mathbf{v}\left(\varphi(\mathbf{u})\right)$ is separable, there exist $\mathbf{V_x}$ and $\mathbf{V_y}$ such that $\mathbf{A} = \mathbf{\Xi}\left(\mathbf{V_x} \otimes \mathbf{V_y}\right)$. In the general case, we approximate $\mathbf{A}$ by $\breve{\mathbf{A}} = \mathbf{\Xi}\left(\sum_{k=1}^{K}\mathbf{C}_k \otimes \mathbf{D}_k\right)$, for small values of K and certain matrices $\mathbf{C}_k$ and $\mathbf{D}_k$, allowing $\mathbf{S} = \underline{\breve{\mathbf{A}}}(\mathbf{Y})$ to be efficiently implemented as $\mathbf{S} = \underline{\mathbf{\Xi}}\left(\sum_{k=1}^{K}\mathbf{D}_k\mathbf{Y}\mathbf{C}_k^T\right)$.

Let $\mathbf{B} \in \mathbb{C}^{m \times n}$ be a generic matrix with $m = m_1 m_2$ and $n = n_1 n_2$. Consider approximating $\mathbf{B}$ as a sum of Kronecker products in terms of $\mathbf{C}_k \in \mathbb{C}^{m_1 \times n_1}$ and $\mathbf{D}_k \in \mathbb{C}^{m_2 \times n_2}$, $1 \le k \le K$, obtained by solving

$$\min_{\{\mathbf{C}_k\},\{\mathbf{D}_k\}}\left\|\mathbf{B} - \sum_{k=1}^{K}\mathbf{C}_k \otimes \mathbf{D}_k\right\|_F.$$

In [7], this problem is shown to be equivalent to

$$\min_{\{\mathbf{C}_k\},\{\mathbf{D}_k\}}\left\|\mathcal{R}(\mathbf{B}) - \sum_{k=1}^{K}\mathrm{vec}\{\mathbf{C}_k\}\,\mathrm{vec}\{\mathbf{D}_k\}^T\right\|_F, \tag{8}$$

where $\mathcal{R}(\cdot)$ is a matrix rearrangement operator. This is a low-rank approximation problem which can be solved with the SVD of $\mathcal{R}(\mathbf{B})$.

For our purposes, we approximate $\mathbf{B} = \mathbf{\Xi}^{-1}\mathbf{A} = \mathbf{\Xi}^T\mathbf{A}$ (since $\mathbf{\Xi}$ is a permutation, $\mathbf{\Xi}^{-1} = \mathbf{\Xi}^T$). Note that $\mathbf{\Xi}$ is the key to a successful low-rank decomposition. As shown in Section 4, using $\mathbf{B} = \mathbf{A}$ is not useful, since $\mathcal{R}(\mathbf{A})$ has too many significant singular values.

3.2. Computing the low-rank fast transform

Computing the dominant singular values and vectors of $\mathcal{R}(\mathbf{B})$ is not trivial, since in practice $\mathcal{R}(\mathbf{B})$ is too large to be stored explicitly in memory. Nevertheless, one can use Lanczos methods [7, 8] which only require the implementation of the matrix-vector products $\mathcal{R}(\mathbf{B})\,\mathbf{u}$ and $\mathcal{R}(\mathbf{B})^H\mathbf{v}$ for arbitrary $\mathbf{u}$, $\mathbf{v}$. One can also use approximate SVD methods designed to require a small number of passes over $\mathcal{R}(\mathbf{B})$ (e.g. [9, 10]).

It can be shown that

$$\mathcal{R}(\mathbf{B})^T = \left[\begin{array}{ccc} \mathbf{Z}_{0,0} & \cdots & \mathbf{Z}_{M_x-1,0} \\ \vdots & & \vdots \\ \mathbf{Z}_{0,M_y-1} & \cdots & \mathbf{Z}_{M_x-1,M_y-1} \end{array}\right]$$

$$\mathbf{Z}_{m,n} = \underline{\mathbf{\Xi}}\left(\mathbf{v}\left(\varphi(u_{x_m}, u_{y_n})\right)\mathbf{v}^H\left(\varphi(u_{x_m}, u_{y_n})\right)\right).$$

Since $\mathbf{v}\left(\varphi(u_{x_m}, u_{y_n})\right)$ can be precomputed for $0 < m < M_x$ and $0 < n < M_y$ and $\mathbf{\Xi}$ is a very fast permutation, $\mathcal{R}(\mathbf{B})\,\mathbf{u}$ and $\mathcal{R}(\mathbf{B})^H\mathbf{v}$ can be evaluated with relative efficiency. Indeed, using the Lanczos method from [8], $N = 64$ and $M_x = M_y = 256$, we can solve (8) for $K = 8$ in 8 minutes on an Intel Core 2 Duo 2.4 GHz processor, using only one core. This is a very reasonable runtime for an offline procedure which must only be run once.

Even in the presence of strong near field effects $\mathcal{R}(\mathbf{B})$ can be well approximated by a low-rank decomposition (see Section 4). Even though the transform cost grows linearly with K, due to the Kronecker representation, the cost of applying each $\mathbf{C}_k \otimes \mathbf{D}_k$ is very small, so a transform with $K = 8$ outperforms an explicit matrix representation of $\mathbf{A}$ by several orders of magnitude.

In [4] we showed that $\mathbf{A} = \mathbf{\Xi}\left(\mathbf{V_x} \otimes \mathbf{V_y}\right)$ can be implemented with $\frac{1}{2}NM + N^2M^{1/2}$ complex MACs (multiply-accumulate operations). Thus, $\breve{\mathbf{A}} = \mathbf{\Xi}\left(\sum_{k=1}^{K}\mathbf{C}_k \otimes \mathbf{D}_k\right)$ can be implemented with $\frac{1}{2}KNM + KN^2M^{1/2}$ MACs, as opposed to the N^2M complex MACs required by an explicit matrix representation of $\mathbf{A}$. To our knowledge, there is no other fast transform that can model near-field propagation, so the explicit matrix representation is the only alternative to our proposal. Considering that the explicit matrix representation is too large to be stored in memory for realistic problem sizes, in practice its implementation is much slower, since its elements must be recomputed when applying $\mathbf{A}$.

As we show in Section 4, even when very strong near-field effects are present, $K = 8$ or $K = 16$ are sufficient for accurate reconstructions, such that this proposal is at least 100 times faster than explicit matrix multiplications for practical problem sizes.

3.3. Calibration and focusing

In practical applications, the array geometry might deviate slightly from an ideal Cartesian grid. Furthermore, microphones are seldom well matched, and require calibration. While these characteristics may be incorporated into the transform with a suitable choice of K, it is more computationally efficient to compensate departures from separability using a separate interpolation matrix. For the purposes of near-field imaging, this matrix can also be designed to make moderate changes to the focal surface without having to recompute the $\{\mathbf{C}_k, \mathbf{D}_k\}$ matrices. This approach is convenient for real-time imaging, since one can compute interpolation matrices in a few seconds, while obtaining $\{\mathbf{C}_k, \mathbf{D}_k\}$ requires several minutes.

Let $\mathbf{v_u}$ and $\check{\mathbf{v}}_\mathbf{u}$ be the ideal and desired manifold vectors, respectively. Assume that $\mathbf{v_u}$ is modeled by the fast transform, while $\check{\mathbf{v}}_\mathbf{u}$ potentially incorporates calibration data and changes to the focal surface. We propose designing an interpolation matrix $\mathbf{T}$ such that $\mathbf{T}\left[\underline{\breve{\mathbf{A}}}\left(\mathbf{Y}\right)\right]\mathbf{T}^H$ becomes the fast transform.

Previous methods for array interpolation [11, 12] design $\mathbf{T}$ such that $\mathbf{T}\mathbf{v}\left(\varphi\left(\mathbf{u}\right)\right) \approx \check{\mathbf{v}}\left(\varphi\left(\mathbf{u}\right)\right)$, for $\mathbf{u}$ in a region of interest. Defining

$$\mathbf{V} = \left[\, \mathbf{v}_{\mathbf{u}_1} \;\, \mathbf{v}_{\mathbf{u}_2} \;\, \cdots \;\, \mathbf{v}_{\mathbf{u}_M} \,\right], \quad \check{\mathbf{V}} = \left[\, \check{\mathbf{v}}_{\mathbf{u}_1} \;\, \check{\mathbf{v}}_{\mathbf{u}_2} \;\, \cdots \;\, \check{\mathbf{v}}_{\mathbf{u}_M} \,\right],$$

a traditional interpolation matrix is computed as

$$\arg\min_{\mathbf{T}} \left\|\mathbf{T}\mathbf{V} - \check{\mathbf{V}}\right\|_F = \check{\mathbf{V}}\mathbf{V}^+,$$

where $\left\|\cdot\right\|_F$ is the Frobenius norm and $\mathbf{V}^+$ is the Moore-Penrose pseudoinverse of $\mathbf{V}$.

We can improve this approach when imaging with transforms. It follows from (6) that $\mathbf{S} = \underline{\breve{\mathbf{A}}}\left(\mathbf{Y}\right)$ is a weighed sum of $\mathbf{v_u}\mathbf{v_u}^H$ outer products. Since $\mathbf{v_u}\mathbf{v_u}^H = \left(\alpha\mathbf{v_u}\right)\left(\alpha\mathbf{v_u}\right)^H$ for any $\alpha \in \mathbb{C}$ with $|\alpha| = 1$, it suffices that $\mathbf{T}\mathbf{v_u} \approx \alpha\check{\mathbf{v}}_\mathbf{u}$ for a conveniently chosen $\alpha \in \mathbb{C}$ with $|\alpha| = 1$. Thus, a more accurate $\mathbf{T}$ can be obtained by solving

$$\min_{\mathbf{T},\mathbf{U}} \left\|\mathbf{T}\mathbf{V} - \check{\mathbf{V}}\mathbf{U}\right\|_F = \min_{\mathbf{U}} \left\|\check{\mathbf{V}}\mathbf{U}\mathbf{V}^+\mathbf{V} - \check{\mathbf{V}}\mathbf{U}\right\|_F, \tag{9}$$

under the constraint that $\mathbf{U}$ is unitary diagonal.

Let $\operatorname{diag}\{\mathbf{U}\} = \boldsymbol{\alpha}$. We minimize (9) by solving for $\mathbf{U}$ and $\mathbf{T}$ alternatively, with

$$\begin{aligned} \mathbf{T}\left(0\right) &= \mathbf{I} \\ \alpha_i\left(n\right) &= \arg\min_{|\alpha|=1} \left\|\mathbf{T}\left(n\right)\mathbf{v}_{\mathbf{u}_i} - \alpha\check{\mathbf{v}}_{\mathbf{u}_i}\right\|_F \\ &= \check{\mathbf{v}}_{\mathbf{u}_i}^H\mathbf{T}\left(n\right)\mathbf{v}_{\mathbf{u}_i} / \left|\check{\mathbf{v}}_{\mathbf{u}_i}^H\mathbf{T}\left(n\right)\mathbf{v}_{\mathbf{u}_i}\right| \\ \mathbf{T}\left(n+1\right) &= \check{\mathbf{V}}\mathbf{U}\left(n\right)\mathbf{V}^+. \end{aligned}$$

Note (9) is not convex under the constraint that $\mathbf{U}$ is unitary diagonal, and may have multiple local minima. Nevertheless, in our simulations this approach converged to the global minimum or to solutions which are very close to it.

To increase the likelihood of converging to the global minimum, one can choose a more accurate $\mathbf{U}\left(0\right)$ at a higher computational cost. Assuming $\mathbf{U}$ is unitary diagonal,

$$\begin{aligned} \left\|\check{\mathbf{V}}\mathbf{U}\mathbf{V}^+\mathbf{V} - \check{\mathbf{V}}\mathbf{U}\right\|_F^2 &= \left\|\check{\mathbf{V}}\mathbf{U}\left(\mathbf{V}^+\mathbf{V} - \mathbf{I}\right)\right\|_F^2 \\ &= \left\|\left[\left(\mathbf{V}^+\mathbf{V} - \mathbf{I}\right)^T \otimes \check{\mathbf{V}}\right] \operatorname{vec}\{\mathbf{U}\}\right\|_2^2 \\ &= \boldsymbol{\alpha}^H\left[\left(\mathbf{V}^+\mathbf{V} - \mathbf{I}\right)^*\left(\mathbf{V}^+\mathbf{V} - \mathbf{I}\right)^T \odot \check{\mathbf{V}}\check{\mathbf{V}}^H\right]\boldsymbol{\alpha}, \end{aligned}$$

where $\odot$ is the pointwise (Hadamard) product. A nearly optimal $\boldsymbol{\alpha}\left(0\right)$ is the eigenvector associated with the smallest eigenvalue of $\left(\mathbf{V}^+\mathbf{V} - \mathbf{I}\right)^*\left(\mathbf{V}^+\mathbf{V} - \mathbf{I}\right)^T \odot \check{\mathbf{V}}\check{\mathbf{V}}^H$, normalized such that each of its coordinates lies on the unit circumference.

4. EXAMPLES

In this section we present simulation results for a rectangular focal surface, showing how reconstruction accuracy varies with K. We simulate a 64-microphone non-uniform planar array with a separable geometry, operating at 4 kHz and 8 kHz, with elements having x and y coordinates drawn from $\{\pm 2.8, \pm 7.8, \pm 12.2, \pm 15.0\}$ (cm). Each image has 256×256 pixels, and was reconstructed by solving the total-variation regularized least-squares problem given by $\min_{\hat{\mathbf{Y}}} \left\|\hat{\mathbf{Y}}\right\|_{BV} + \mu\left\|\operatorname{vec}\{\mathbf{S}\} - \mathbf{A}\operatorname{vec}\{\hat{\mathbf{Y}}\}\right\|_2^2$. We used the solver TVAL3 [13] with $\mu = 10^3$ and 100 iterations, and set σ in (4) to provide a 20 dB SNR, following the methodology from [5].

Fig. 1 shows examples reconstructing a checkerboard source distribution, which was chosen to illustrate how artifacts are influenced by source coordinates. We note that its symmetry does not provide any advantage to the proposed methods. The source distribution is located over a $0.5\,\mathrm{m} \times 0.5\,\mathrm{m}$ rectangle parallel to the array plane, simulated at 0.5 m from the array and parameterized by $\varphi\left(u_x, u_y\right) = \left[\, u_x \;\, u_y \;\, 0.5 \,\right]^T$, with $\left(u_x, u_y\right) \in \left[-0.5,\, 0.5\right]^2$.

Since the far-field parameterization models a complete hemisphere, the $0.5\,\mathrm{m} \times 0.5\,\mathrm{m}$ rectangle does not fill its respective acoustic image. For the far-field parameterization, we indicate the horizon using a white circumference. The source distribution is very close to the array, so reconstructing it under a plane wave assumption produces smeared images, motivating the use of our proposal.

Our benchmark is the exact (and slow) near-field transform. Since the array resolution decreases monotonically towards the horizon, all reconstructions present smearing toward the edges of the focal rectangle. This is not an artifact of the fast transform, and can also be observed when using the exact transform. Note that $\mathbf{A}$ is easier to approximate for low frequencies. For this particular example, $K = 4$ and $K = 8$ are sufficient to deliver accurate reconstructions at 4 kHz and 8 kHz, respectively.

Fig. 2 compares the first 100 singular values (out of a total of 16384) for $\mathcal{R}\left(\mathbf{A}\right)$ and $\mathcal{R}\left(\mathbf{\Xi}^T\mathbf{A}\right)$. We consider $\mathbf{A}$ modeling the rectangular focal surface defined previously, for 8 kHz sources. The sharp decay of the curve for $\mathcal{R}\left(\mathbf{\Xi}^T\mathbf{A}\right)$ highlights the importance of $\mathbf{\Xi}$ in enabling accurate low-rank approximations.

Fig. 3 shows how one can use reconstruction errors to estimate focal distances. Indeed, one obtains the best fit when using a transform which matches the true focal surface, with a correct focal distance. The dashed lines show that if one does not have an optimal transform designed for the true focal surface (in this case, a rectangle at 0.5 m), one can correct it using an interpolation matrix.

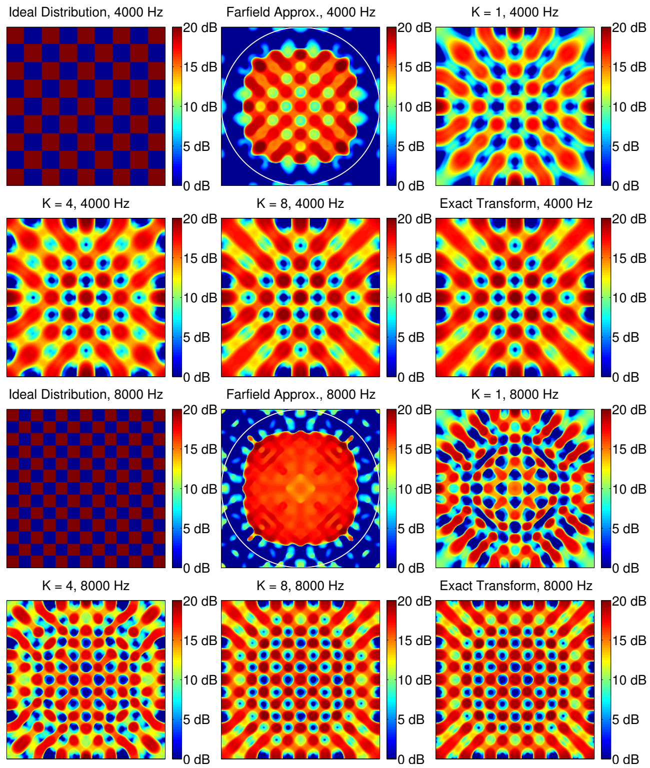

Fig. 1. TV-regularized reconstruction using a far-field approximation, the best Kronecker decomposition (figures with K) and the exact (slow) transform. Sources are positioned over a $0.5\,\mathrm{m} \times 0.5\,\mathrm{m}$ rectangle parallel to the array, located at a distance of $0.5\,\mathrm{m}$.

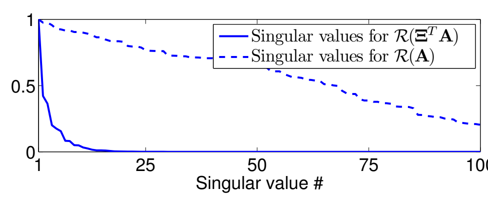

Fig. 2. First 100 singular values for $\mathcal{R}(\mathbf{A})$ and $\mathcal{R}\left(\boldsymbol{\Xi}^T\mathbf{A}\right)$, normalized to 1.

Reconstruction times for $K = 1$, 2, 4, 8 and 16 are approximately 4, 5, 6, 9 and 15 seconds per image, with MATLAB implementations running on an Intel Core 2 Duo T9400 processor in 64-bit mode, using only one core (exact reconstruction times usually vary by $\pm.5$ seconds, depending on the source distribution). In contrast, using an explicit matrix multiplication requires around 2000 seconds per image.

5. CONCLUSION

In this paper we have proposed a method for computationally efficient near-field imaging. As shown in Section 4, we can achieve very accurate reconstructions with computational costs at most one order of magnitude above those of the far-field fast transform. This is a notable result, since it still makes the proposed near-field transforms about as fast as FFT2-accelerated deconvolutions (see results from [4]), which degrade quality by using a far-field approximation. Furthermore, it enables the use of practical use of regularized least-squares methods, which produce better results than deconvolution methods.

Future work involves extending the Kronecker approximation for more diverse array geometries, for which other permutations $\boldsymbol{\Xi}$ will be required to promote separability.

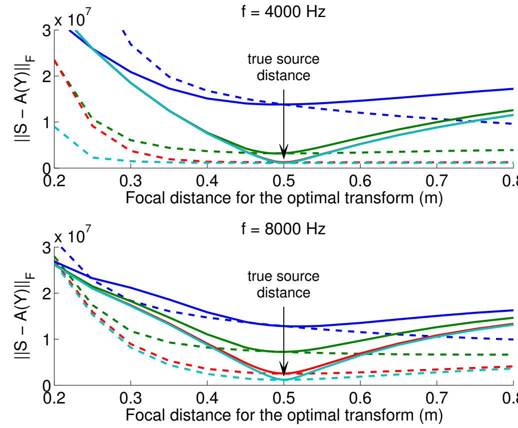

Fig. 3. Reconstruction errors for varying K. —: $K = 1$, no focusing; – –: $K = 1$, with $\mathbf{T}$; —: $K = 4$, no focusing; – –: $K = 4$, with $\mathbf{T}$; —: $K = 8$, no focusing; – –: $K = 8$, with $\mathbf{T}$; —: $K = 16$, no focusing; – –: $K = 16$, with $\mathbf{T}$.

6. REFERENCES

[1] T.J. Mueller, Ed., *Aeroacoustic measurements*, Springer Verlag, 2002.

[2] R.P. Dougherty, “Extensions of DAMAS and Benefits and Limitations of Deconvolution in Beamforming,” in *Proc. of the 11th AIAA/CEAS Aeroacoustics Conference*, 2005.

[3] 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, pp. 2631, 2008.

[4] F.P. Ribeiro and V.H. Nascimento, “A fast transform for acoustic imaging with separable arrays,” in *Proc. of ICASSP*, 2011.

[5] F.P. Ribeiro and V.H. Nascimento, “Computationally efficient regularized acoustic imaging,” in *Proc. of ICASSP*, 2011.

[6] R.A. Horn and C.R. Johnson, *Matrix analysis*, Cambridge University Press, 1990.

[7] C.F. Van Loan and N. Pitsianis, “Approximation with Kronecker products,” in *Linear Algebra for Large Scale and Real Time Applications*, M.S. Moonen and G.H. Golub, Eds., pp. 293–314. Kluwer Publications, 1992.

[8] J. Baglama and L. Reichel, “Augmented implicitly restarted Lanczos bidiagonalization methods,” *SIAM Journal on Scientific Computing*, vol. 27, no. 1, pp. 19–42, 2006.

[9] P. Drineas, R. Kannan, and M.W. Mahoney, “Fast Monte Carlo algorithms for matrices II: Computing a low-rank approximation to a matrix,” *SIAM Journal on Computing*, vol. 36, no. 1, pp. 158–183, 2007.

[10] N. Halko, P.G. Martinsson, and J.A. Tropp, “Finding structure with randomness: Stochastic algorithms for constructing approximate matrix decompositions,” Tech. Rep. ACM 2009-05, California Inst. Tech., Sept. 2009.

[11] J. Pierre and M. Kaveh, “Experimental performance of calibration and direction-finding algorithms,” in *Proc. of ICASSP*, 1991, pp. 1365–1368.

[12] M. Pesavento, AB Gershman, and Z.Q. Luo, “Robust array interpolation using second-order cone programming,” *IEEE Signal Processing Letters*, vol. 9, no. 1, pp. 8–11, 2002.

[13] C. Li, “An efficient algorithm for total variation regularization with applications to the single pixel camera and compressive sensing,” M.S. thesis, Rice University, 2009.