An efficient variational Bayesian inference approach via Studient's-t priors for acoustic imaging in colored noises

Ning Chu, Ali Mohammad-Djafari, José Picheral, Nicolas Gac

▶ **To cite this version:**

Ning Chu, Ali Mohammad-Djafari, José Picheral, Nicolas Gac. An efficient variational Bayesian inference approach via Studient's-t priors for acoustic imaging in colored noises. POMA - ICA 2013, Jun 2013, Montreal, Canada. 055031 (9p.). hal-00865786v2

HAL Id: hal-00865786

https://centralesupelec.hal.science/hal-00865786v2

Submitted on 10 Oct 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An efficient Bayesian Variational inference approach via sparsity enforcing priors for acoustic imaging in colored noises

Ning CHU[a,*], Ali MOHAMMAD-DJAFARI[a], José PICHERAL[b], Nicolas GAC[a]

[a] *Laboratoire des signaux et systèmes (L2S), CNRS-SUPELEC-UNIV PARIS SUD, 91192 GIF-SUR-YVETTE, FRANCE*
[b] *SUPELEC, Département du Signal et Systèmes Electroniques, 91192 GIF-SUR-YVETTE, FRANCE*

Abstract

Acoustic imaging is a powerful tool to localize and reconstruct source powers using microphone array. However, it often involves the ill-posed inversions and becomes too time-consuming to obtain high spatial resolutions. In this paper, we firstly propose a shift-invariant convolution model to approximate the forward model of acoustic power propagation. The convolution kernel is derived from the Symmetric Toepliz Block Toepliz (STBT) structure of propagation matrix. Then we propose a hierarchical Bayesian inference approach via Variational Bayesian Approximation (VBA) criterion in order to achieve robust acoustic imaging in colored background noises. For super spatial resolution and wide dynamic power range, we explore the Student's-t prior on the acoustic power distribution thanks to the sparsity and heavy tail of prior model. Colored noise distributions are also modeled by the Student's-t prior, and this does not excessively penalize large model errors as the Gaussian white prior does. Finally proposed 2D convolution model and VBA approach are validated through simulations and real data from wind tunnel compared to classical methods.

Keywords: Acoustic imaging, Variational Bayesian Approximation, Student's-t prior, colored noises, convolution, Teopliz matrix

1. Introduction

Acoustic imaging is an advanced technique for acoustic source localization and power reconstruction with microphone array, which can provide the insight into acoustic generating mechanisms and behavior properties. Since decades, it has been widely applied in vibration detection and acoustic comfort evaluation of transportation in wind tunnel tests etc. The classic forward model of power propagation for uncorrelated acoustic sources can be expressed by a determined linear system of equations obtained in frequency domain: $\mathbf{y} = \mathbf{C}\mathbf{x}$, where $\mathbf{y} \in \mathcal{R}^N$ denotes the measured beamforming power vector [1] at the microphone array; $\mathbf{x} \in \mathcal{R}^N$ denotes the unknown acoustic power vector on the source plane; $\mathbf{C} \in \mathcal{R}^{N \times N}$ denotes the propagation operation which depends on the relative geometric distance between the source plane and sensor array; N denotes the total number of discretization grids. Since $\mathbf{C}$ is usually a singular matrix, its direct inversion causes unstable and non-unique solutions. Recently, the Deconvolution Approach for Mapping of Acoustic Source (DAMAS) method [2] has been effectively applied in wind tunnel tests by NASA. DAMAS gives an iterative solution for $\mathbf{y} = \mathbf{C}\mathbf{x}$ under the non-negative constraint on $\mathbf{x}$. However, DAMAS is sensitive to background noise due to its noise-free assumption, and it also suffers from slow convergence due to its spatially variant Point Spread Function (PSF) of sensor array. To accelerate deconvolution, extended DAMAS [3] have been proposed to suppose the PSF to be shift invariant, so that Discrete Fourier Transportation (DFT) could be used to perform fast deconvolution. Nevertheless, this assumption inevitably affects spatial resolution. More recently, ℓ_1 sparse regularization has been investigated to greatly improve the spatial resolutions [4]. The authors have proposed several approaches by making good use of sparse distribution of source powers in Gaussian white noises, such as Robust DAMAS with Sparse Constraint (SC-RDAMAS), Bayesian inference approach via a sparsity enforcing prior [5].

*Principal corresponding author: Ning.CHU@lss.supelec.fr (Ning CHU). Tel. : +33 (0)1 69 85 1743. Fax : 00 33 (0)1 69 85 17 65

Email addresses: djafari@lss.supelec.fr (Ali MOHAMMAD-DJAFARI), jose.picheral@supelec.fr (José PICHERAL), Nicolas.GAC@lss.supelec.fr (Nicolas GAC)

In this paper, our motivation is to explore an efficient and robust approach for acoustic imaging in colored background noise on the application of detecting acoustic source on the surface of vehicle surface in wind tunnel tests. This paper is organized as follows: Section 2 introduces the proposed convolution approximation for acoustic power propagation model and the convolution kernel selections; then the Variational Bayesian Approximation (VBA) approach is proposed in Section 3; the validation of convolution approximation and proposed VBA approach are respectively shown in Section 4 and 5; finally we conclude this paper in Section 6.

The novelties are: to improve the efficiency, we modify the power propagation model by applying convolution approximation based on the spatially invariant PSF; and we derive appropriate PSF size and values from the Symmetric Toepliz Block Toepliz (STBT) structure of propagation matrix $\mathbf{C}$. To obtain super spatial resolution and colored noise suppression, we apply the Variational Bayesian Approximation (VBA) via Student's-t priors on source powers and colored noise distributions, thanks to the sparsity and long heavy tail of Student's-t distribution.

2. Proposed convolution approximation for power propagation

Before modeling, we make three necessary assumptions: Acoustic sources are uncorrelated monopole and sensors are omnidirectional with unitary gain. In Fig.1, our objectives are shown for the acoustic imaging on the surface of the vehicle in the wind tunnel S2A in Renault Company. The source plane is discretized by $N_r \times N_c$ identical grids, where N_r, and denote row number and column number, provided $N_r \geq N_c$. We define an acoustic power image $\mathbf{x}_0 = \{x_{r,c}\}_{N_r \times N_c}$. The acoustic power propagation from source plane to the microphone sensor array can be modeled as a multiply-and-add operator obtained in frequency domain[2], denoted by a power transferring matrix $\mathbf{C} = \{c_{i,j}\}_{N \times N}$, where array response $c_{i,j} \in \mathbf{C}$ can be modeled by the array geometry as

$$c_{i,j} = \frac{\|\mathbf{a}_i^H \mathbf{a}_j\|_2^2}{\|\mathbf{a}_i\|_2^2} . \tag{1}$$

where $\mathbf{a}_i = \left\{\frac{1}{r_{i,m}} \exp\left[-j2\pi f \tau_{i,m}\right]\right\}_M^T$ denotes the steering vector [6], in which, $r_{i,m}$ denotes the propagation distance from source signal at ith grid to mth microphone sensor; $\tau_{i,m} = r_{i,m}/c_0$ denotes the ideal propagation time during $r_{i,m}$, where c_0 denotes the acoustic speed in common air; f denotes the acoustic frequency and M denotes the total sensor number. We also apply the equivalent sources and mirror sources to correct the wind refraction and ground reflection respectively as discussed in authors' paper [5].

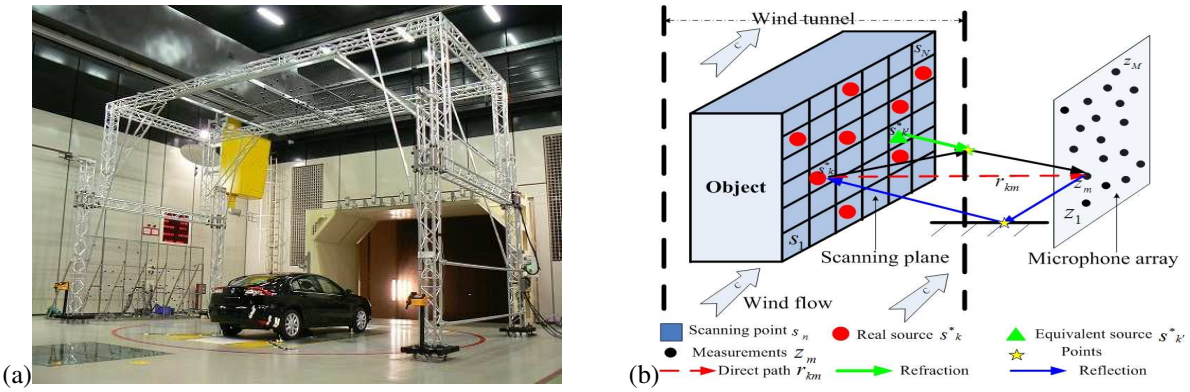

Figure 1: Wind tunnel experiments: (a) Wind tunnel S2A [7] and (b) Illustration of signal processing [5]

In Eq.(1), $N \times N$ propagation matrix $\mathbf{C}$ causes a computational complexity as heavy as $O(N^2)$. This motivates us to approximate Eq.(1) by a shift-invariant convolution model, so that the computational complexity could be significantly reduced into $O(N \log N)$.

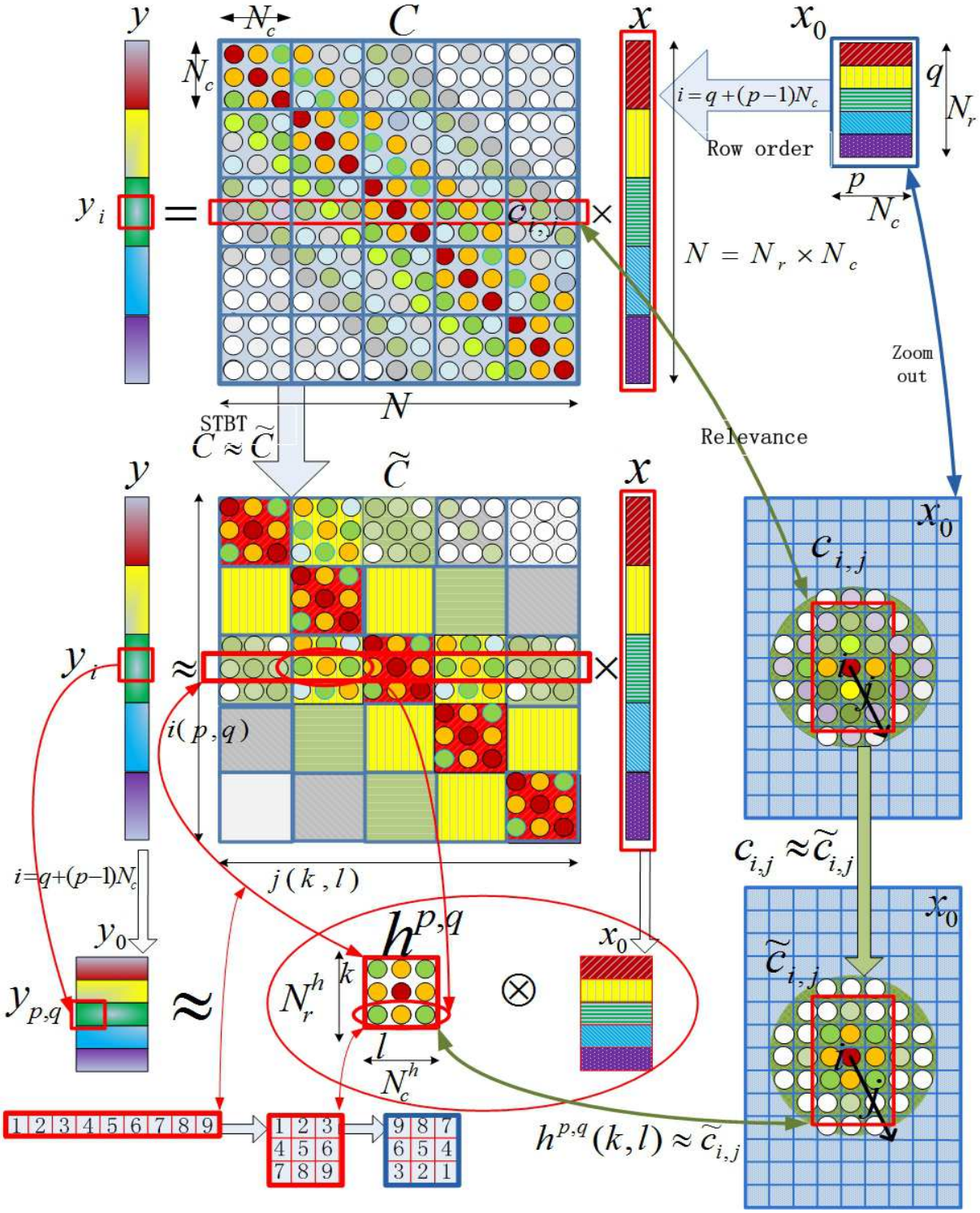

Figure 2: 2D convolution approximation model

2.1. STBT matrix approximation

From Eq.(1), $c_{i,j}$ mainly depends on geometric distance and $\mathbf{C}$ can be approximated into a Symmetric Toepliz Block Toepliz (STBT) $\tilde{\mathbf{C}}$, if the aperture of sensor array is relative small with respect to the source plane, and vice versa. Under this assumption, we have $\|\mathbf{a}_i\|_2^2 = \sum_{m=1}^{M} 1/r_{i,m}^2 \approx M/D^2$ for any $i, j \in \{1, \cdots, N\}$, with D being the averaged distance between sensor and source plane; And we also have $(r_{i,m}\, r_{j,m}) \approx D^2$ for any $m \in \{1, \cdots, M\}$. Then we get an approximated array response $\tilde{c}_{i,j}$ as:

$$c_{i,j} \approx \tilde{c}_{i,j} = \frac{1}{M} \left| \sum_{m=1}^{M} \exp\left[j \frac{2\pi f_l}{c_0} (r_{i,m} - r_{j,m}) \right] \right|^2 , \tag{2}$$

where c_0 denotes the acoustic speed in the common air; we get $\tilde{c}_{i,j} = \tilde{c}_{j,i}$, $\tilde{\mathbf{C}} = \tilde{\mathbf{C}}^T$, then $\tilde{\mathbf{C}}$ is a **symmetric** matrix as shown in Fig.(2).

Let $\mathbf{x}_0$ denote the unknown source power image with sizes $N_r \times N_c$, where N_r,N_c denotes the row and column number respectively, provided $N_r \geq N_c$. Source power vector $\mathbf{x}_0$ is composed of the pixels of $\mathbf{x}_0$ in the row-first order. If source powers $x_i, x_j \in \mathbf{x}_0$ are on the same row on power image $\mathbf{x}_0$, then we have $\tilde{c}_{i,j} \propto |r_{i,m} - r_{j,m}| \propto |i - j|$ for sensor m, therefore $\tilde{\mathbf{C}}$ is a **block** matrix:

$$\begin{cases} \tilde{c}_{k,l}^{(p,q)} = c_{i,j},\ \tilde{\mathbf{C}}_{p,q} = \{\tilde{c}_{k,l}^{(p,q)}\}_{N_c \times N_c},\ \tilde{\mathbf{C}} = \{\tilde{\mathbf{C}}_{p,q}\}_{N_r \times N_r} \\ i = k + (p-1)\, N_c,\ j = l + (q-1)\, N_c \end{cases} \tag{3}$$

where $c_{k,l}^{(p,q)}$ denotes the kth row, lth column item of the sub-matrix $\tilde{\mathbf{C}}_{p,q}$ which is the pth row, qth column sub-matrix of symmetric matrix $\tilde{\mathbf{C}}$. And there are N_r^2 square blocks $\tilde{\mathbf{C}}_{p,q}$ with the size N_c in $\tilde{\mathbf{C}}$, as shown in Fig.(2).

For any $\tilde{c}_{k,l}^{(p,q)}$, $\tilde{c}_{k+1,l+1}^{(p,q)} \in \tilde{\mathbf{C}}_{p,q}$, we have $\tilde{c}_{k,l}^{(p,q)} = \tilde{c}_{k+1,l+1}^{(p,q)} \propto |k - l|$, then $\tilde{\mathbf{C}}_{p,q}$ is a **Toeplitz** sub-matrix, as shown in Fig.(2).

For any $\tilde{c}_{k,l}^{(p,q)} \in \tilde{\mathbf{C}}_{p,q}$ and $\tilde{c}_{k,l}^{(p+1,q+1)} \in \tilde{\mathbf{C}}_{p+1,q+1}$, we get $\tilde{c}_{k,l}^{(p,q)} = \tilde{c}_{k,l}^{(p+1,q+1)} \propto |k - l|$, then $\tilde{\mathbf{C}}$ is a **Toeplitz block** matrix, as shown in Fig.(2).

Finally, $\mathbf{C}$ can be approximated by a **STBT** matrix $\tilde{\mathbf{C}}$, as shown in Fig.(2).

2.2. Convolution kernel

The above approximation make the classic forward model in Eq.(1) to be approximated by the convolution model as:

$$\mathbf{y}_0 = \mathbf{H}_0 \mathbf{x}_0 + \boldsymbol{\epsilon}\,, \tag{4}$$

where $\mathbf{y}_0$ denotes the beamforming power image with the same sizes as $\mathbf{x}_0$; $\boldsymbol{\epsilon}$ denotes the model errors, which are the combinations of background noises $\mathbf{e}$ at the sensors, forward model uncertainty caused by acoustic multipath propagations, as well as the convolution approximation errors caused in Eq.(2); $\mathbf{H}_0$ denotes the convolution matrix, satisfying $\mathbf{H}_0\mathbf{x}_0 = \{[\mathbf{h}_{p,q} * \mathbf{x}_0]_{p,q}\}_{N_r \times N_c}$, where operator $*$ denotes convolution, $[*]_{p,q}$ denotes the pth, qth item of the output convolution matrix in Fig.(2); And $\mathbf{h}_{p,q} = \{h_{p,q}(l,k)\}_{N_r^h \times N_c^h}$ denotes the convolution kernel matrix with the sizes $N_r^h \times N_c^h$; Owing to the STBT matrix $\tilde{\mathbf{C}}$, $\mathbf{h}_{p,q}$ is composed of $\tilde{c}_{i,j}$ from corresponding row of $\tilde{\mathbf{C}}$ as follows:

$$\begin{cases} h_{p,q}(l,k) = \tilde{c}_{i,j} \approx c_{i,j}\,,\ i = q + (p-1)\, N_c \\ j = i + (k - \lfloor \frac{N_r^h - 1}{2} \rfloor)\, N_c + l - \lfloor \frac{N_c^h - 1}{2} \rfloor \end{cases} , \tag{5}$$

where $\lfloor \cdot \rfloor$ denotes integer. In practice, we take $h_{p,q}(l,k) = c_{i,j}$, since $c_{i,j}$ can be directly calculated from Eq.(1). Finally, the real kernel should be rearranged as $\mathbf{h}^{p,q}(k,\, l) = \mathbf{h}^{p,q}(|l - \lfloor \frac{N_c^h - 1}{2} \rfloor|,\, |k - \lfloor \frac{N_r^h - 1}{2} \rfloor|)$ as shown in Fig.(2).

For fast and effective convolution, we should firstly select a shift-invariant kernel matrix. Owing to the STBT matrix $\tilde{\mathbf{C}}$, the items $\tilde{c}_{i,j}$ in the middle row ($i = \lfloor \frac{N_r - 1}{2} \rfloor$) can contain all the information of $\tilde{\mathbf{C}}$, so that it is reasonably to select the center kernel as $\mathbf{h}_{p,q} = \mathbf{h}_{\lfloor \frac{N_r - 1}{2} \rfloor, \lfloor \frac{N_c - 1}{2} \rfloor}$.

Secondly we should select a relatively small kernel sizes compared to the maximal size $2N_r - 1 \times 2N_c - 1$. Since STBT matrix $\tilde{\mathbf{C}}$ has the square blocks with the sizes $N_c \times N_c$, it is naturally to select a square kernel with the size N_c.

3. Proposed VBA inference approach

For the inverse problem in Eq.(4), some prior information (constraints) on both source powers $\mathbf{x}_0$ and colored noises $\boldsymbol{\epsilon}$ should be investigated in order to reduce the uncertainty of solutions. Let $\boldsymbol{\theta} = [\mathbf{x}_0, \boldsymbol{\epsilon}]^T$ denote the unknown parameters, where operator $(\cdot)^T$ denotes transpose, and $\mathbf{D}$ denotes observed data and known parameters. The inverse problem via priors can be effectively solved by the following Bayesian inference approaches: If we assign the specific prior probability $p(\boldsymbol{\epsilon})$ to noises $\boldsymbol{\epsilon}$, we can define the likelihood $p(\mathbf{D}|\boldsymbol{\theta})$ which is used classically by the Maximum Likelihood (ML) estimation as $\hat{\boldsymbol{\theta}}_{ML} = \arg\max_{\boldsymbol{\theta}} \{p(\mathbf{D}|\boldsymbol{\theta})\}$; In the Bayesian approach, we also assign the specific prior probabilities $p(\boldsymbol{\theta})$ to all unknown parameters. Then we use the Maximum A Posteriori (MAP) estimation as $\hat{\boldsymbol{\theta}}_{MAP} = \arg\max_{\boldsymbol{\theta}} \{ln\, p(\boldsymbol{\theta}|\mathbf{D})\} \propto \arg\min_{\boldsymbol{\theta}} \{-ln\, p(\mathbf{D}|\boldsymbol{\theta}) - ln\, p(\boldsymbol{\theta})\}$ according to Baye's rule. In fact, MAP can be seen as a regularization of ML, but MAP has the advantage of adaptively estimating the regularization parameter, compared to conventional regularization methods.

However in MAP, $ln\, p(\boldsymbol{\theta}|\boldsymbol{D})$ is often hardly to get an analytical form and usually a nonlinear function with respect to $\boldsymbol{\theta}$. Moreover, both ML and MAP are the point estimators. These difficulties can be overcome by the VBA [8, 9] estimation, in which, posterior $p(\boldsymbol{\theta}|\boldsymbol{D})$ is approximated by a family of basic easily handled probability distributions $q(\boldsymbol{\theta})$, namely $p(\boldsymbol{\theta}|\boldsymbol{D}) \approx q(\boldsymbol{\theta})$; and proper $q(\boldsymbol{\theta})$ are estimated by maximizing variational bound $\mathfrak{L}(\boldsymbol{\theta})$ as: $\hat{q}(\boldsymbol{\theta}) = \arg\max_{q(\boldsymbol{\theta})} \{\mathfrak{L}(\boldsymbol{\theta})\}$, where $\mathfrak{L}(\boldsymbol{\theta}) = \int q(\boldsymbol{\theta})\, ln \frac{p(\boldsymbol{D},\boldsymbol{\theta})}{q(\boldsymbol{\theta})}\, d\boldsymbol{\theta}$. Generally, $\boldsymbol{\theta}$ are supposed to be mutually independent: $q(\boldsymbol{\theta}) = \prod_i q(\theta_i)$. Then $\mathfrak{L}(\boldsymbol{\theta})$ is maximized by the mean field approximation as: $\hat{q}(\theta_i) = \frac{exp[I(\theta_i)]}{\int exp[I(\theta_i)]\, d\theta_i}$, where $I(\cdot)$ denotes the partition function, defined as $I(\theta_i) =< ln\, p(\boldsymbol{D}, \boldsymbol{\theta}) >_{q(\boldsymbol{\theta}_{-i})} = \int q(\boldsymbol{\theta}_{-i})\, ln\, p(\boldsymbol{D}, \boldsymbol{\theta}) d\boldsymbol{\theta}_{-i}$, where $\boldsymbol{\theta}_{-i}$ denotes the parameter vector except item θ_i. In fact, $I(\theta_i)$ could hardly be analytically computed, since it depends on $q(\boldsymbol{\theta}_{-i})$. But VBA inference can still obtain the approximating posterior $\hat{q}(\boldsymbol{\theta})$ owing to the conjugate distributions, in which, $\hat{q}(\boldsymbol{\theta})$ comes from the same family of the prior $p(\boldsymbol{\theta})$ based on the proper combination of the likelihood and priors.

3.1. Heavy tail prior on colored noises

In wind tunnel experiments, we model the colored noises $\boldsymbol{\epsilon}$ by the Student's-t prior $St(\boldsymbol{\epsilon})$ that has a long heavy tail, instead of Gaussian white ones whose sharp tail excessively penalizes the large errors of forward model. Another attractive superposition property is that $St(\boldsymbol{\epsilon})$ can be generated by marginalizing hidden variable $\boldsymbol{\nu}$ as $St(\boldsymbol{\epsilon}) = \int p(\boldsymbol{\epsilon}|\boldsymbol{\nu}) p(\boldsymbol{\nu})\, d\boldsymbol{\nu}$, in which, conditional prior $p(\boldsymbol{\epsilon}|\boldsymbol{\nu}) = \mathcal{N}(\boldsymbol{\epsilon}|0, \boldsymbol{\Sigma}_{\epsilon}^{-1})$ is the multivariate Gaussian distribution, with $\boldsymbol{\Sigma}_{\epsilon} = Diag\{\boldsymbol{\nu}\}$ being noise covariance matrix; $Diag(\cdot)$ denotes diagonal matrix; $\boldsymbol{\nu} = \{\nu_n\}_N$ denotes the noise precision vector; And $p(\boldsymbol{\nu}) = \prod_{n=1}^{N} Gama(\nu_n|a_\nu, b_\nu) = \prod_{n=1}^{N} \Gamma(a_\nu)^{-1} (b_\nu)^{a_\nu} \nu_n^{a_\nu - 1} e^{-b_\nu \nu_n}$, with a_ν, b_ν being the hyperparameters of $p(\boldsymbol{\nu})$ and $\Gamma(x) = \int t^{x-1} e^{-t} dt$.

According to proposed convolution forward model of Eq.(4), the likelihood is determined by the conditional prior $p(\boldsymbol{\epsilon}|\boldsymbol{\nu}) = \mathcal{N}(\boldsymbol{\epsilon}|0, \boldsymbol{\Sigma}_{\epsilon}^{-1})$ as:

$$p(\mathbf{y}_0|\mathbf{x}_0, \boldsymbol{\nu}) = \frac{|\boldsymbol{\Sigma}_{\epsilon}|^{1/2}}{(2\pi)^{N/2}}\, e^{-\frac{1}{2}(\mathbf{y}_0 - \mathbf{H}_0\mathbf{x}_0)^H \boldsymbol{\Sigma}_{\epsilon} (\mathbf{y}_0 - \mathbf{H}_0\mathbf{x}_0)}, \tag{6}$$

where operator $(\cdot)^H$ denotes conjugate transpose.

3.2. Sparse prior on acoustic power image

Acoustic source in wind tunnel experiments are generated by the wind collision on the specific parts of the vehicle surface. Therefore sources sparsely locate on some particular parts, while on the most of common parts, there are few sources. This is why acoustic powers $\mathbf{x}_0$ become K-sparsity signals when the source plane is discretized into N grids. Such a sparse distribution can be represented by the distribution that has a very high value around the original zero (sparsity) and a long heavy tail (dynamic range of source powers).

Here we apply the Student's-t priors $St(\mathbf{x}_0)$ [8] to enforce the sparsity and wide dynamic range of source power distribution. Owing to the superposition property, hidden variable $\boldsymbol{\gamma}$ is marginalized out for $St(\mathbf{x}_0) = \int p(\mathbf{x}_0|\boldsymbol{\gamma})\, p(\boldsymbol{\gamma}) d\boldsymbol{\gamma}$, where $p(\mathbf{x}_0|\boldsymbol{\gamma}) = \mathcal{N}(\mathbf{x}_0|0, \boldsymbol{\Sigma}_x^{-1})$ is assigned to multivariate Gaussian distribution, in which, $\boldsymbol{\Sigma}_x$ denotes power covariance matrix, defined as $\boldsymbol{\Sigma}_x = Diag\{\boldsymbol{\gamma}\}$ with $\boldsymbol{\gamma} = \{\gamma_n\}_N$ being the power precision vector; and $p(\boldsymbol{\gamma}) = \prod_{n=1}^{N} Gama(\gamma_n|a_\gamma, b_\gamma)$, where a_γ, b_γ denotes the hyperparameters of $p(\boldsymbol{\gamma})$. $0 < \gamma_n < 1$ greatly promotes the sparsity, while $\gamma_n \to \infty$ makes $St(x_n)$ to approach Gauss normal distribution, which has no sparsity. Compared to the Double Exponential (DE) prior in [5], $St(x_n)$ involves one hidden variable γ_n for each x_n, while DE prior requires two parameters in order to achieve the same sparsity and heavy tail distribution, as the red curve shown in Fig.3a.

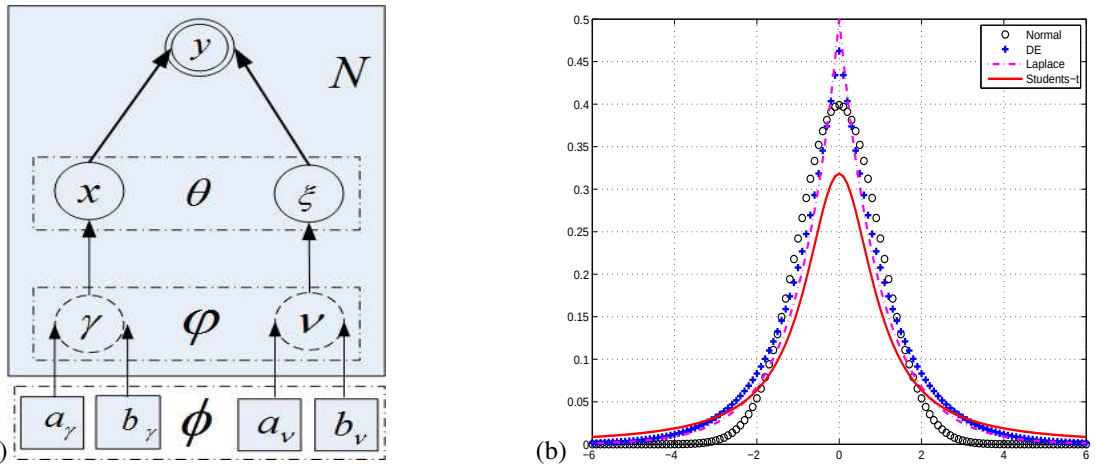

Figure 3: (a) Sparse priors on Gaussian normal, Laplace, DE and Student's-t (b) N dimension 3-hierarchical Bayesian Graphical model, Double circle: observed data; Single: unknown variables; Dash: hidden variables; Square: hyperparameters; Arrow: dependence.

3.3. VBA parameter estimations

In Fig.3b, the graphical model describes the dependencies between the observed data $\mathbf{y}_0$, unknown variables $\boldsymbol{\theta} = [\mathbf{x}_0, \boldsymbol{\epsilon}]^T$, their hidden variables $\boldsymbol{\varphi} = [\boldsymbol{\gamma}, \boldsymbol{\nu}]^T$ and hyperparameters $\boldsymbol{\phi} = [a_\gamma, b_\gamma, a_\nu, b_\nu]^T$. According to Baye's rule, we have $p(\boldsymbol{\theta}, \boldsymbol{\varphi}|\mathbf{y}_0) \propto p(\mathbf{y}_0|\boldsymbol{\theta}, \boldsymbol{\varphi})p(\boldsymbol{\theta}|\boldsymbol{\varphi})p(\boldsymbol{\varphi})$. Owing to the multivariate Gaussian likelihood in Eq.(6) and the superpositions of Student's-t priors on $\boldsymbol{\theta}$, we specifically obtain the posterior as

$$\begin{array}{ll} p(\boldsymbol{\theta}, \boldsymbol{\varphi}|\mathbf{y}_0) & = \mathcal{N}(\mathbf{x}_0|\mathbf{y}_0 - \mathbf{H}_0\mathbf{x}_0, \boldsymbol{\Sigma}_\epsilon^{-1})\,\mathcal{N}(\mathbf{x}_0|0, \boldsymbol{\Sigma}_x^{-1}) \\ & Gama(\boldsymbol{\gamma}|a_\gamma, b_\gamma)\mathcal{N}(\boldsymbol{\epsilon}|0, \boldsymbol{\Sigma}_\epsilon^{-1})Gama(\boldsymbol{\nu}|a_\nu, b_\nu) \end{array} . \tag{7}$$

Owing to the conjugate prior, approximating posterior belongs to Student's-t distribution which consists of multivariate Gaussian $\hat{q}(\mathbf{x}_0)$ and Gamma distributions $\hat{q}(\boldsymbol{\gamma})$, $\hat{q}(\boldsymbol{\nu})$ as:

$$\left\{ \begin{array}{l} \hat{q}(\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_0|\hat{\boldsymbol{\mu}}_x, \hat{\boldsymbol{\Sigma}}_x) \\ \hat{q}(\boldsymbol{\gamma}) = \prod_{n=1}^N Gama(\gamma_n|\hat{a}_\gamma, \hat{b}_\gamma^n) \\ \hat{q}(\boldsymbol{\nu}) = \prod_{n=1}^N Gama(\nu_n|\hat{a}_\nu, \hat{b}_\nu^n), \end{array} \right. , \tag{8}$$

where $\mathbf{x}_0$ and $\boldsymbol{\varphi}$ are supposed to be mutually independent; and expected variable estimations are as:

$$\left\{ \begin{array}{l} \hat{\boldsymbol{\mu}}_x = \hat{\boldsymbol{\Sigma}}_x\mathbf{H}_0^T < \boldsymbol{\Sigma}_\epsilon > \mathbf{y}_0 \\ \hat{\boldsymbol{\Sigma}}_x = (\mathbf{H}_0^T < \boldsymbol{\Sigma}_\epsilon > \mathbf{H}_0 + < \boldsymbol{\Sigma}_x >)^{-1} \\ \hat{a}_\gamma = a_\gamma + \frac{1}{2}\,,\ \hat{b}_\gamma^n = b_\gamma + \frac{1}{2} < \mathbf{x}_0\mathbf{x}_0^T >_{nn} \\ \hat{a}_\nu = a_\nu + \frac{1}{2}\,,\ \hat{b}_\nu^n = b_\nu + \frac{1}{2} < \boldsymbol{\epsilon}\boldsymbol{\epsilon}^T >_{nn} \end{array} \right. , \tag{9}$$

where operator $(\cdot)_{nn}$ denotes the nth diagonal item, and $< \cdot >$ denotes expectation, which are calculated as:

$$\left\{ \begin{array}{l} < \boldsymbol{\Sigma}_\epsilon >= Diag\{< \nu_n >\}_N = Diag\{< \hat{a}_\nu/\hat{b}_\nu^n >\}_N \\ < \boldsymbol{\Sigma}_x >= Diag\{< \gamma_n >\}_N = Diag\{< \hat{a}_\gamma/\hat{b}_\gamma^n >\}_N \\ < \mathbf{x}_0\mathbf{x}_0^T >= \hat{\boldsymbol{\mu}}_x\hat{\boldsymbol{\mu}}_x^T + \hat{\boldsymbol{\Sigma}}_x \\ < \boldsymbol{\epsilon}\boldsymbol{\epsilon}^T >= \mathbf{y}_0\mathbf{y}_0^T - 2\mathbf{H}_0\hat{\boldsymbol{\mu}}_x\mathbf{y}_0^T + \mathbf{H}_0 < \mathbf{x}_0\mathbf{x}_0^T > \mathbf{H}_0^T \end{array} \right. , \tag{10}$$

All the solutions in Eq.(9, 10) require the values of the hyperparameters $\boldsymbol{\phi} = [a_\gamma, b_\gamma, a_\nu, b_\nu]^T$. During the iterations of parameter estimations, variables $\boldsymbol{\theta} = [\mathbf{x}_0, \boldsymbol{\varphi}]$ are firstly computed, then the hyperparameters $\boldsymbol{\phi}$ can be alternatively estimated by making the first partial derivative of variational bound $\mathfrak{L}_{\boldsymbol{\theta}}(\boldsymbol{\phi})$ equal zero ($\frac{\partial \mathfrak{L}}{\partial \boldsymbol{\phi}_i} = 0$) as follows:

$$\left\{ \begin{array}{l} \frac{\partial \mathfrak{L}}{\partial a_\nu} = N\,ln\,b_\nu - N\,\digamma(a_\nu) + \sum_{n=1}^N < ln\,\nu_n > \\ \frac{\partial \mathfrak{L}}{\partial b_\nu} = N\,\frac{a_\nu}{b_\nu} - \sum_{n=1}^N < \nu_n > \end{array} \right. , \tag{11}$$

where $< \nu_n >$ and $< \gamma_n >$ are computed in Eq.(10), and $\digamma(\cdot)$ denotes the digamma function defined by $\psi(x) = \Gamma'(x)/\Gamma(x)$ where $\Gamma(\cdot)$ denotes the Gamma function. a_ν, b_ν can be simultaneously estimated from the same procedure. Parameter update can be done numerically by Matlab fzero function.

3.4. Computational analysis

From the solutions in Eq.(9), $\hat{\mathbf{\Sigma}}_x$ involves the matrix inversion which can not be calculated explicitly nor efficiently. We have to approximate $\hat{\mathbf{\Sigma}}_x$ with a circulant matrix as $\hat{\mathbf{\Sigma}}_x \approx (< \bar{\nu} > \mathbf{H}_0^H \mathbf{H}_0 + < \bar{\gamma} >)^{-1}$, where $\bar{\nu} = \sum_{n=1}^{N} \nu_n$, $\bar{\gamma} = \sum_{n=1}^{N} \gamma_n$ denote the arithmetic mean. Then the products of circulant matrices can be efficiently computed in the Discrete Fourier Transform (DFT) domain. In Eq.(9), the estimated expectation $\hat{\boldsymbol{\mu}}_x$ of source powers can be analytically expressed as $\hat{\mathbf{\Sigma}}_x^{-1} \hat{\boldsymbol{\mu}}_x = \mathbf{H}_0^H < \mathbf{\Sigma}_\epsilon > \mathbf{y}_0$. This linear system of equations is solved iteratively with the conjugate gradient algorithm, which requires $O(N \log N)$ computations to treat N dimension vector $\mathbf{x}_0$. If Q iterations are needed, total computations are of $O(Q\,N \log N)$, which remains moderate burden.

4. Simulations

The Simulation configurations are based on the wind tunnel experiments in Fig.1a: averaged distance between the sensors and source plane is D=4.50m; There are M=64 sensors forming a non-uniform array with $0.1m^2$ averaged aperture; Source plane with $1m^2$ surface is discretized by 5cm×5cm grids. In Fig.5a, input power image $\mathbf{x}_0$ consists of 4 monopoles and 5 extended sources with different patterns, whose powers are within 14dB dynamic range; And image size are of N_r=27, N_c=17. The output beamforming power image $\mathbf{y}_0$ is shown in Fig.5b. The colored noises are generated by using the Gaussian white noises via low pass filter (cut-off frequency 3000Hz), and the averaged Signal-to-Noise Ratio (SNR) is set as low as 0dB.

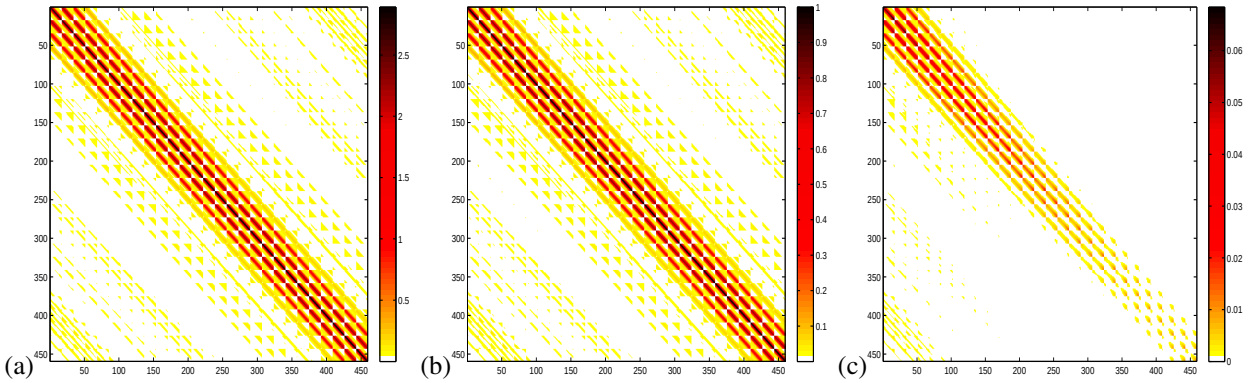

Figure 4: Propagation matrix: (a) Near field (b) Far field (c) Difference between Normalized near and far fields

We firstly show the structures of propagation matrix $\mathbf{C}$ in the cases of near and far fields in Fig.(4). It is seen that $\mathbf{C}$ can be approximated as a STBT matrix. Then we show the convolution approximating errors $\delta_y = \|\mathbf{y}_0 - \hat{\mathbf{y}}_0\|_2^2 / \|\mathbf{y}_0\|_2^2 \times 100\%$ VS various kernels in In Fig.6, where $\hat{\mathbf{y}}_0$ denotes convolution results. We examine six kernels with different forms and sizes. As we can see, the larger kernel size N_h is, the smaller δ_y become; Particularly when N_h approaches N_c, all δ_y are very close to each other and maintain 3%. Finally reconstruction results $\hat{\mathbf{x}}_0$ in Fig.5 are shown as: the beamforming merely gives blur image of strong soures; Though big kernel size inevitably deteriorate boundary pixels and definitely harms the deconvolution result, our Bayesian MAP inversion based on proposed convolution model actually approaches the high resolution result based on classic forward model, but it performs much faster owing to the shift-invariant convolution kernel.

5. Wind tunnel experiments

Wind tunnel experiments are designed to reconstruct the positions and acoustic powers on the traveling car surface. We suppose that all sources locate on the same plane, since the curvature of the car is relatively small compared to the

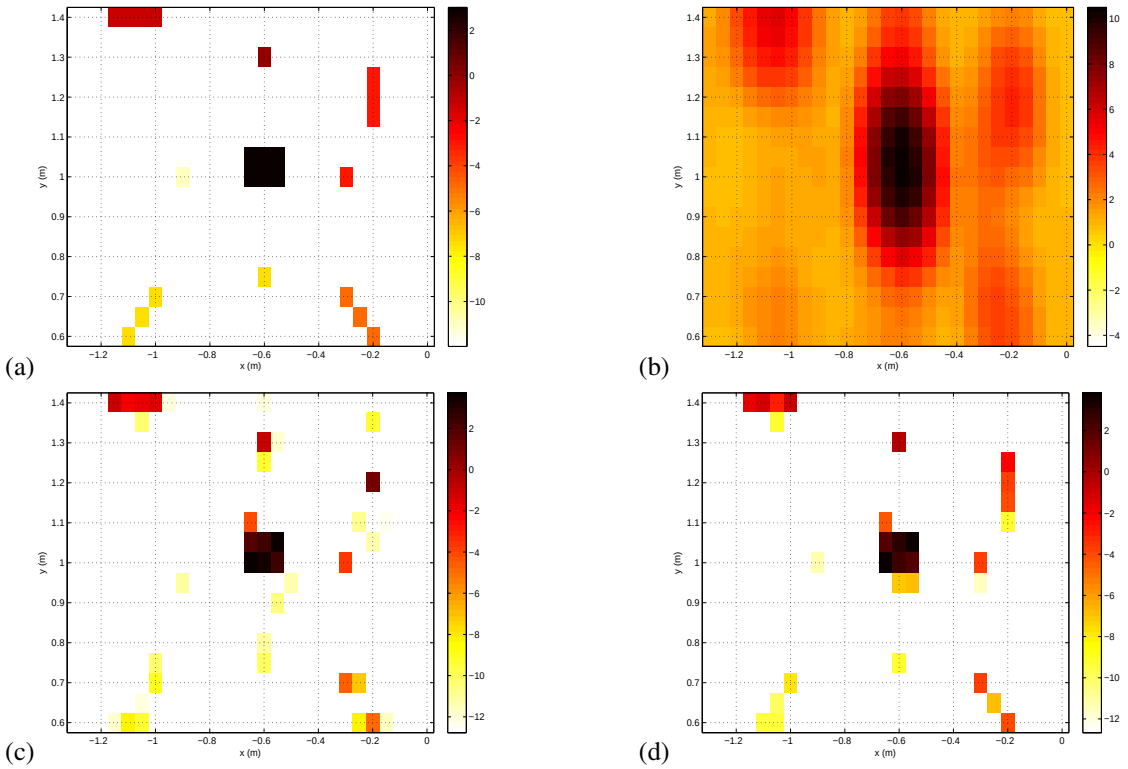

Figure 5: Simulation at 2500Hz, 0dB SNR in colored noises, 14dB display: (a) Source powers (b) Beamforming powers (c) Bayesian MAP inversion and (d) Proposed VBA inversion

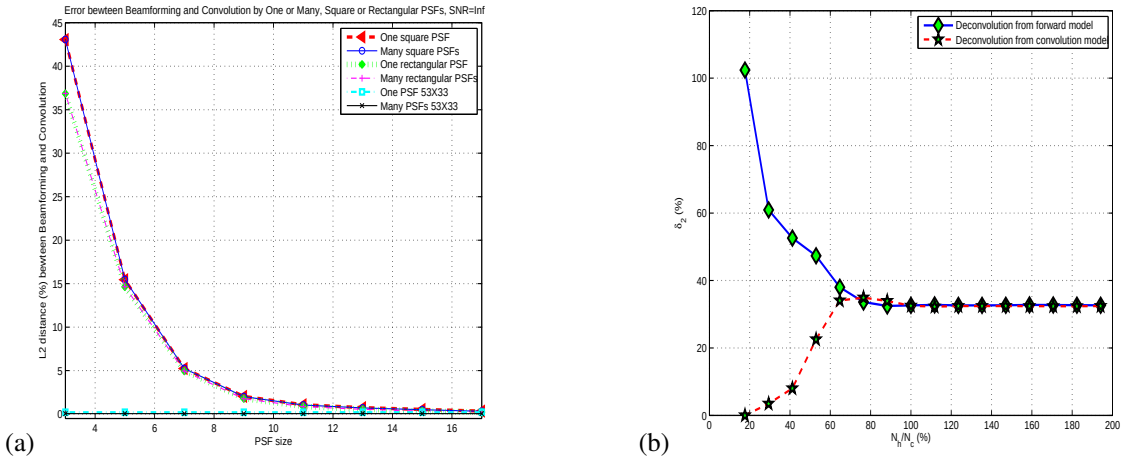

Figure 6: (a) Convolution errors VS Convolution kernel without noises. (b) Deconvolution errors VS kernel sizes without noises

distance 4.5m between the car and array plane. The grid is of $5cm$, and source plane is thus of 31×101 pixels. The wind speed is $160km/h$; there are 524288 samplings with the sampling frequency f_s=2.56×10^4 Hz. Total samplings are separated into I=204 blocks with 2560 samplings in each bloc. The working frequency is 2500Hz, which is sensitive to human being. The image results are obtained in frequency domain shown by normalized dB images with 10dB span. Propagation matrix $\mathbf{C}$ in Eq.(4) is rectified for the wind refraction and ground reflection respectively as discussed in [5].

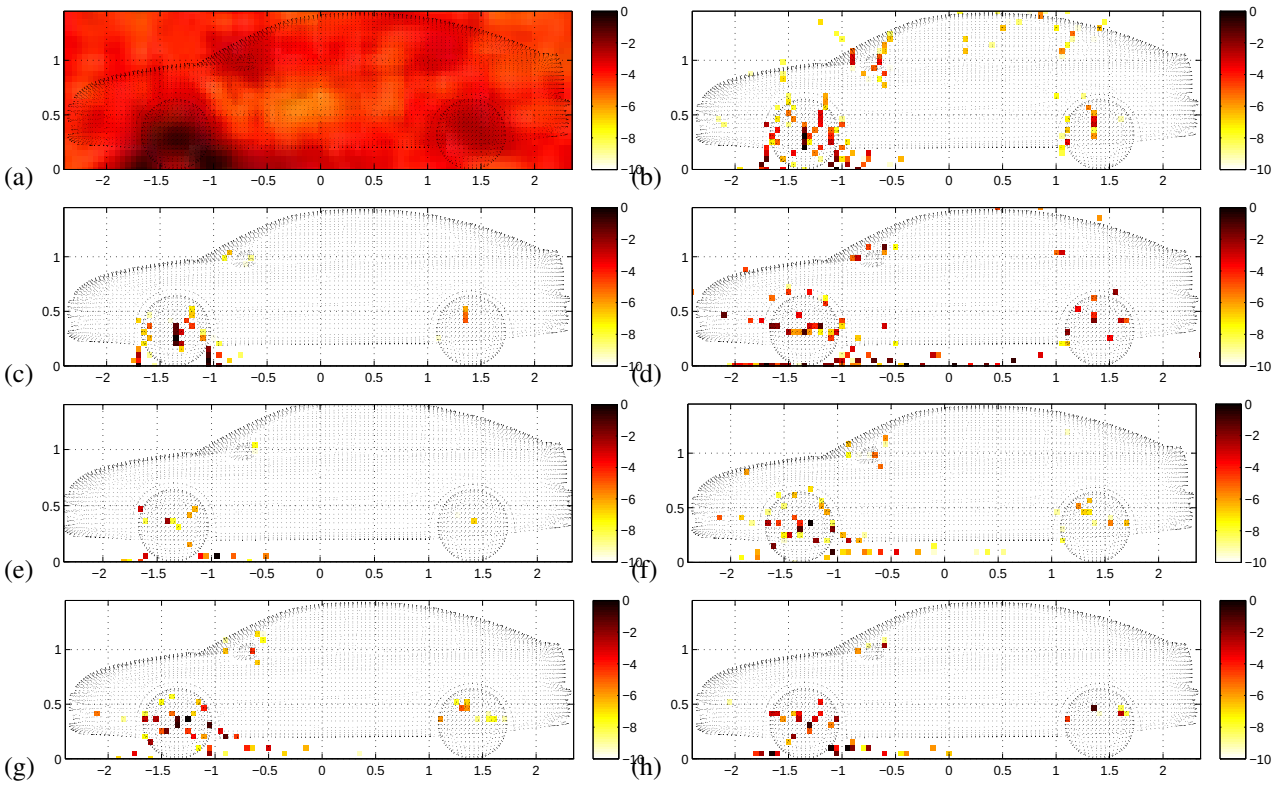

Figure 7: Acoustic imaging on car surface at 2500Hz: (a) Beamforming (b) DAMAS (5000i) (c) DR-DAMAS (5000i) (d) CLEAN (e) SC-DAMAS (f) SC-RDAMAS (g) Joint MAP inference and (h) Proposed VBA inference.

Figure.7 illustrates the estimated power images of mentioned methods at 2500Hz. Due to the high sidelobe effect, beamforming just gives a fuzzy image of strong sources in Fig.7a; DAMAS well deconvolves the beamforming image and discovers sources around the wheels and rearview mirror, however, many false targets are also detected on the air in Fig.7b; Diagonal Removal DAMAS in Fig.7 eliminates most of the artifacts, but it also harms weak sources. Figure.7d and 7e show that both CLEAN and SC-DAMAS overcome the drawbacks of the DAMAS, but unexpected strong points are detected on the ground due to the parameter selection for each use. In Fig.7f and g, the sparsity constraint and Joint MAP via sparsity prior not only manages to distinguish the strong sources around the two wheels, rearview mirror and side window, but also successfully reconstructs the week ones on the front cover and light. Finally, the proposed VBA inference effectively achieve super-resolutions and wide dynamic range on the two wheels and mirror; furthermore, the suppression of the background noise are much better than others thanks to the Student's-t prior on the colored noise. From table 1, proposed VBA via convolution model is more efficient realized compared to Bayesian MAP method via classical forward model.

Table 1: Computational cost for treating whole car: image 31×101 pixels, at 2500Hz, based on CPU:3.33GHz.

Methods	Beamforming	DAMAS	DR-DAMAS	CLEAN	SC-RDAMAS	MAP	Proposed
Time (s)	1	10	11	45	852	1012	578

6. Conclusion

In this paper, we have developed an efficient VBA inference approach via Student's-t priors on source powers and colored noises for efficient acoustic imaging with super spatial resolution, wide dynamic range and robustness to noises.

The main novelties are: 1) For efficiency, we have proposed convolution approximation with fixed kernel for acoustic power propagation. Appropriate kernel size and values have been determined from the Symmetric Toepliz Block Toepliz structure of propagation matrix. 2) For super resolution and wide dynamic range, we have proposed Student's-t prior on source power distribution owing to its sparsity and heavy tail. 3) For robustness to colored noises, we have proposed Student's-t prior on the colored noise, which can not excessively penalize large errors as the Gaussian prior does. 4) For hyperparameter estimations, we have applied the VBA optimization via conjugate distributions, in which, approximating posterior distributions are favorably composed of the multivariate Gaussian distribution for variables, and Gamma distributions for hidden variables.

For the validations, We have presented method comparisons based on the simulated and real data in wind tunnel experiments. For future work, we are investigating Graphical Processing Unit (GPU) to accelerate proposed the hierarchical VBA inference approach.

Acknowledgment

The authors are deeply grateful to Renault SAS, especially Mr. Jean-Luc Adam for offering real data and valuable discussions on our research.

References

[1] J. Chen, K. Yao, R. Hudson, Source localization and beamforming, Signal Processing Magazine, IEEE 19 (2) (2002) 30–39.

[2] T. Brooks, W. Humphreys, A Deconvolution Approach for the Mapping of Acoustic Sources (DAMAS) determined from phased microphone arrays, Journal of Sound and Vibration 294 (4-5) (2006) 856–879.

[3] R. Dougherty, Extensions of DAMAS and Benefits and Limitations of Deconvolution in Beamforming, in: 11th AIAA/CEAS Aeroacoustics Conference, Monterey, CA, USA; 23-25 May, 2005, pp. 1–13.

[4] D. Malioutov, M. Çetin, A. Willsky, A sparse signal reconstruction perspective for source localization with sensor arrays, IEEE Transactions on Signal Processing 53 (8) (2005) 3010–3022.

[5] N. Chu, A. Mohammad-Djafari, J. Picheral, Robust bayesian super-resolution approach via sparsity enforcing priors for near-field wideband aeroacoustic source imaging, in: Journal of Sound and Vibration, Elsevier Ltd, Submitted in May 2012, under minor revisions Dec.2012.

[6] Y. Wang, J. Li, P. Stoica, M. Sheplak, T. Nishida, Wideband RELAX and wideband CLEAN for aeroacoustic imaging, Journal of Acoustical Society of America 115 (2) (2004) 757–767.

[7] A. Menoret, N. Gorilliot, J.-L. Adam, Acoustic imaging in wind tunnel S2A, in: 10th Acoustics conference (ACOUSTICS2010), Lyon, France, 2010.

[8] D. Tzikas, A. Likas, N. Galatsanos, Variational bayesian sparse kernel-based blind image deconvolution with student's-t priors, IEEE Transactions on Image Processing 18 (4) (2009) 753–764.

[9] A. Mohammad-Djafari, Bayesian approach with prior models which enforce sparsity in signal and image processing, EURASIP Journal on Advances in Signal Processing 2012 (1) (2012) 52.