Bayesian approach in acoustic source localization and imaging

Ning Chu

N° d'ordre: 2013PA112277

THÈSE

Présentée pour obtenir

LE GRADE DE DOCTEUR PHYSIQUE EN
SCIENCES DE L'UNIVERSITÉ PARIS-SUD XI

Spécialité: Traitement du Signal et l'Image

École doctorale: Sciences et Technologies de l'Information des Télécommunications et des Systèmes (STITS).

Lieu de soutenance: Laboratoire des Signaux et Systèmes (L2S) CNRS-SUPELEC-UNIV PARIS SUD, plateau de moulon, 3 rue de Joliot Curie, Gif-sur-yvette, 91192, France.

par

Ning CHU

Approche Bayésienne Pour La Localisation De Sources En Imagerie Acoustique

Bayesian Approach In Acoustic Source Localization And Imaging

Soutenue le 22 novembre 2013 devant la Commission d'examen:

Mr.	Jérôme ANTONI	(Rapporteur, Prof. INSA Lyon, France)
Mr.	Alain BERRY	(Rapporteur, Prof. Univ. Sherbrooke, Canada)
Mr.	Andrea MASSA	(Rapporteur, Prof. Univ. Trento, Italie)
Mr.	Daniel BLACODON	(Examinateur Invité, ONERA, France)
Mr.	Rémy BOYER	(Examinateur, HDR. Univ. Paris Sud, France)
Mr.	Alfred HERO	(Président, Prof. Univ. Michigan, USA)
Mr.	Ali MOHAMMAD-DJAFARI	(Directeur de thèse, DR. CNRS, France)
Mr.	José PICHERAL	(Co-encadrant, MF. SUPELEC, France)
Mr.	Nicolas GAC	(Co-encadrant, MF. Univ. Paris Sud, France)

.

- In theory, theory and practice are the same.
- In practice, they are not!

Albert Einstein

Thèse préparée au
Laboratoire des signaux et systémes (L2S)
UMR 8506 CNRS-SUPELEC-UNIV PARIS-SUD
SUPELEC, plateau de Moulon, 3 rue Joliot-Curie
91,192 GIF-SUR-YVETTE Cedex (France)

Abstract

Acoustic imaging is an advanced technique for acoustic source localization and power reconstruction using limited measurements at microphone sensor array. This technique can provide meaningful insights into performances, properties and mechanisms of acoustic sources. It has been widely used for evaluating the acoustic influence in automobile and aircraft industries. Acoustic imaging methods often involve in two aspects: a forward model of acoustic signal (power) propagation, and its inverse solution. However, the inversion usually causes a very ill-posed inverse problem, whose solution is not unique and is quite sensitive to measurement errors. Therefore, classical methods cannot easily obtain high spatial resolutions between two close sources, nor achieve wide dynamic range of acoustic source powers.

In this thesis, we firstly build up a discrete forward model of acoustic **signal** propagation. This signal model is a linear but under-determined system of equations linking the measured data and unknown source signals. Based on this signal model, we set up a discrete forward model of acoustic **power** propagation. This power model is both linear and determined for source powers. In the forward models, we consider the measurement errors to be mainly composed of background noises at sensor array, model uncertainty caused by multi-path propagation, as well as model approximating errors.

For the inverse problem of the acoustic power model, we firstly propose a robust super-resolution approach with the sparsity constraint, so that we can obtain very high spatial resolution in strong measurement errors. But the sparsity parameter should be carefully estimated for effective performance. Then for the acoustic imaging with large dynamic range and robustness, we propose a robust Bayesian inference approach with a sparsity enforcing prior: the double exponential law. This sparse prior can better embody the sparsity characteristic of source distribution than the sparsity constraint. All the unknown variables and parameters can be alternatively estimated by the Joint Maximum A Posterior (JMAP) estimation. However, this JMAP suffers a non-quadratic optimization and causes huge computational cost. So that we improve two following aspects: In order to accelerate the JMAP estimation, we investigate an invariant 2D convolution operator to approximate acoustic power propagation model. Owing to this invariant convolution model, our approaches can be parallelly implemented by the Graphics Processing Unit (GPU). Furthermore, we consider that measurement errors are spatially variant (non-stationary) at different sensors. In this more practical case, the distribution of measurement errors can be more accurately modeled by Students-t law which can express the variant variances by hidden parameters. Moreover, the sparsity enforcing distribution can be more conveniently described by the Student's-t law which can be decomposed into multivariate Gaussian and Gamma laws. However, the JMAP estimation risks to obtain so many unknown variables and hidden parameters. Therefore, we apply the Variational Bayesian Approximation (VBA) to overcome the JMAP drawbacks. One of the fabulous advantages of VBA is that it can not only achieve the parameter estimations, but also offer the confidential interval of interested parameters thanks to hidden parameters used in Students-t priors.

To conclude, proposed approaches are validated by simulations, real data from wind tunnel experiments of Renault S2A, as well as the hybrid data. Compared with some typical state-of-the-art methods, the main advantages of proposed approaches are robust to measurement errors, super spatial resolutions, wide dynamic range and no need for source number nor Signal to Noise Ration (SNR) beforehand.

Keywords : Acoustic source Localization, Bayesian approach, super resolution, sparsity, Students-t, acoustic imaging, deconvolution, wind tunnel, JMAP, VBA, GPU.

Approche Bayésienne Pour La Localisation De Sources En Imagerie Acoustique

Résumé

L'imagerie acoustique est une technique performante pour la localisation et la reconstruction de puissance des sourses acoustiques en utilisant des mesures limitées au réseau des microphones. Cette technique peut fournir des indications significatives sur les performances, les propriétés et les mécanismes de sources acoustiques. Elle a été largement utilisée pour évaluer l'influence acoustique dans l'industrie automobile et aéronautique. Les méthodes d'imagerie acoustique impliquent souvent deux aspects : un modèle direct de propagation du signal (la puissance) acoustique ; l'inversion de ce modèle direct. Cependant, cette inversion provoque généralement un problème inverse mal-posé, dont la solution n'est pas unique, et sensible aux erreurs des mesures. Par conséquent, les méthodes classiques ne peuvent pas obtenir facilement la haute résolution spatiale entre deux sources proches, ni de la large dynamique de puissance acoustique.

Dans cette thèse, nous avons tout d'abord construit un discret modèle direct sur la propagation du signal. Ce modèle est un système d'équations, linéaire mais sous-déterminé, qui permet de relier les données mesurées et les signaux inconnus. Basé sur ce modèle du signal, nous avons créé un discret modèle direct de la puissance acoustique qui est devenu à la fois linéaire et déterminé pour les puissances acoustique. Visant à la robustesse dans le modèle de la puissance, nous considérons que les erreurs des mesures se composent principalement de trois parties : les bruits de fond au réseau de capteurs, l'incertitude de modèle causée par les propagations à multi-trajets (réflection et réfraction), ainsi que les erreurs de l'approximation de la modélisation.

Pour le problème inverse du modèle direct de puissance, nous avons d'abord proposé une approche de hyper-résolution en utilisant la contrainte de la parcimonie, de sorte que nous pouvons obtenir très haute résolution spatiale au sein de fortes erreurs des mesures. Mais le paramètre de la parcimonie doit être estimé attentivement pour un rendement efficace. Ensuite, pour l'obtention de la large dynamique et la robustesse, nous avons proposé une approche robuste de l'inférence bayésienne avec un a priori parcimonieux : la loi double exponentielle. Toutes les variables et paramètres inconnus peuvent être estimées alternativement par l'estimation de maximum a posteriori conjoint (JMAP). Toutefois, le JMAP souffre d'optimisation non-quadratique et provoque d'énorme coût de calcul. Nous avons donc amélioré deux aspects suivants : d'un coté pour accélérer le JMAP, nous avons étudié une approximation pour le modèle direct de puissance en utilisant le 2D convolution avec un noyau invariant. Grâce à ce modèle, nos approches peuvent être réalisées parallèlement par Graphics Processing Unit (GPU). De l'autre coté, on considère que les erreurs des mesures sont spatialement non-stationnaires dans différents capteurs. La distribution des erreurs peut être modélisée par la loi de Students-t qui peut modéliser les variances variées via les paramètres cachés. Par ailleurs, la distribution parcimonieuse de la puissance acoustique peut être raffinée plus convenablement par la loi de Students-t qui peut être décomposée par la loi normale multidimensionnelle pour les puissances, et la loi de gamma pour les paramères cachés. Mais le JMAP rend difficulité aux estimations de large dimension des variables inconnues. Nous avons donc également appliqué l'Approximation Variationnelle Bayésienne (VBA) qui peut permettre non seulement d'obtenir toutes les estimations des inconnues, mais aussi de founir des intervalles confiants grâce aux paramètres cachés untilisés par les lois de Students-t.

Pour conclure, nos approches ont été comparées avec des méthodes d'état-de-l'art sur des simulations, données réelles provenant d'essais en soufflerie chez Renault S2A, ainsi que les données hybrides. Les principaux avantages des approches proposées sont de la robustesse au bruit de fond, d'une large dynamique, hyper-résolutions spatiales et efficaces à utiliser, également pas besoin de la connaissance préalable du numbre des sources ni du Rapport Signal sur Bruit (RSB).

Mots-clefs : Localisation de sources acoustiques, approche bayésienne, hyper-résolution, parcimonie, Students-t, imagerie acoustique, déconvolution, soufflerie, JMAP, VBA, GPU.

.

- Il faut avoir la qualité morale noble;
- Il faut s'entendre franchement avec les autres;
- Il faut avoir des comportements courtois.

Charles-Louis de Secondat (**Montesquieu**)

Acknowledgement

How time flies! It has been one year of master and three years of PhD studies in the Laboratory of the Signals and Systems (L2S), Ecole Supérieur d'Electricité (SUPELEC), University of Paris Sud (XI), France. When I began to write this thesis, the past ceaselessly hit into my eyes, just like yesterday once more:

Step by step, I have learned profoundly from my supervisors and my colleagues...

Little by little, I have made many friends during studying, sporting and traveling...

Day after day, I have understood comprehensively the French language, colorful culture, interesting customs, splendid history and French uniqueness etc, as well as the other European countries that I visited...Finally I find out that I have fallen in love with the French 'vin et fromage'!

Indeed, there are still countless exciting and interesting things for me to experience, to confront, to ponder during my PhD and afterwards, but first of all, I would like to express my infinite gratitude and respect to all who have taken care of me, to all who have been taking care of me.

How much gratitude can I express to my kind supervisors: Prof. Ali MOHAMMAD-DJAFARI, Dr. José PICHERAL and Dr. Nicolas GAC! I want to firstly thank Prof. DJAFARI, who admitted me in 2009, and suggested me 1-year Master then 3-years PhD study on a promising scientific research field, especially for his open mind to allow me to choose the research topic that I prefer, for his erudite to make me well comprehend the charming of Bayesian methods. I want to gratefully thank Dr. PICHERAL, who is always trying his best to help me solve lots of challenging technical problems during my PhD and on real data from wind tunnel experiments, especially for his deep understanding on my puzzles, for his strong support on my publications, for his nice encouragement on my advancements. I want to sincerely thank Dr. GAC, who have looked after me since my Master study in L2S, especially for his kindness to help me manage many trivial but important issues that concern my studies and daily life in France, for his recommendation for me to teach two undergraduate courses in the IUT of University of Paris 11, Cachan, for his arrangements for me to co-supervise the trainee in the L2S. Above all, I want to strongly express my great gratefulness to my three supervisors on cultivating me how to become a respectable scholar: to

be curious, to be rigorous, to be honest and to be respectful on what we did and what we are doing, as well as on other people's works and contributions.

I would like to gratefully thank Dr. Jean-Luc ADAM in Renault SAS for offering real data and valuable discussions on my PhD researches. I appreciate Prof. Sylvie MARCOS introduce Dr. PICHERAL to be my co-supervisor.

Thanks a lot to the L2S lab director Prof. Silviu-Iulian NICULESCUD, our kind colleges Mme Laurence STEPHEN, Mme Giron MARYVONNE, Mme Baverel MYRIAM, M. Frédéric DESPEREZ and M. Franck GUIMONET etc. in the L2S. I thank greatly to the director Stéphane FONT in the SEE SUPELEC, and Mme Anne BATALIE, M.Luc BATALIE, M. Huu Hung VUONG for their indispensable supports. I also appreciate the nice helps from the STITS director Prof.Véronique VEQUE and the assistant Laurence STEPHEN. I never forget the kind comprehension from M. Marine VALOIS from the library of Univ. Paris Sud.

To my colleagues in the Groupe de Problm̀e Inverse (GPI), they have offered me so much encouragement and pleasure that I feel that we are living in a big, warm and international family: Dr. Thomas RODET, Dr. Matthieu KOWALSKI, Dr. Fraysse AURELIA, CAI Caifang, CHEN Long, ZHENG Yuling, Thomas BOULAY, Leila GHARSALLI, Mitrea DUMITRU etc, my previous colleagues Dr. Hashem ayasso, Dr. Doriano Poveaza, Dr ZHU Sha and Dr. Diarra Fall, etc, as well as my other friends in L2S and SUPELEC: Dr. DAI Jing, Dr. ZHONG Yu, Dr. LI Yuanfu, Dr SONG Li, WU Jingwen, WANG Jing, CHEN Yijing, HUANG Fan, LI Changyou, MA Wenchao, Marcel-Stefan GEAMANU, Bien HOANG etc. I appreciate the valuable discussions with my two trainees in L2S and SUPELEC respectively: GAN Xiangyang and XIONG Wenmeng; thanks them very much to assist my research works.

To my friends who are always beside me and the ones who have left France, I want to share my joys and achievements with you, and say:'Take care of yourself, guys! I am always beside you too!'

To my dear cousin ZHOU Ren who also studies in Paris France, I want to thank her delicious cooking and nice accompaniment which relieve my feelings of homesick. To aunt REN Haiwei, many thanks for her kindness and nice visits in France.

In particular, I want to give my gratitude and respect to Prof. ZHOU

Zhimin in the National University of Defense Technology (NUDT), China. Without his support and concern, I could hardly insist on studying abroad. I respect him as my mentor, whose erudite, character and outlook have set up a great example for me.

I also deeply thank my dear friends from NUDT for their great supports: ZHU Sha, LIU Zhongxun, LI Erbao, HU Canbin, WU Jianfei, JU Jinchuan, FENG Dawei, ZENG cong, LIU Jiagui, TANG yangming etc. I will never forget the days and nights we studied and lived together in France. I thank very much to the professors such as Prof. WANG Xuesong, HUANG Xiaotao, XIAO Shunping etc. from NUDT who have visited the L2S. I will always remember your warm courages.

I want to thank the researchers whom I met in conferences, the unknown reviewers, especially the jury member of my PhD defense. Thank you all for your great interests and piercing comments on my PhD researches.

I never forget my first house owner Mr. and Mme Francis MOUGEOT when I lived in Les Ulis in my first year in France. They are so warm-hearted to teach me the French and help me adapt to the French life. I would be greatly happy if Mme MOUGEOT could see my PhD defense, since she passed away last year (2012). I want to thank Mr and Mme Claude OTTINGER for accommodating me for three years and offering me a convenient, beautiful and friendly environment in La Hacquiniére. I will miss the fascinating garden in front of the room window, which makes me released and relaxed in spite of the July sunshine or the January snowstorm.

Great thanks to the China Scholarship Council (CSC) and EGIDE-SUPELEC-FRANCE for their scholarships to support my Master and PhD studies in France.

I want to thank... so many people!

But the most importantly, I devote this thesis for my parents. I ask myself in the depth of my heart, how many words can I express 'I love them'? The answer is complicated but simple: I can do nothing but devote myself to love them with my heart and soul in my life! At last, this thesis should be devoted to Ms JIANG Miao for her endless supports!

It is the time to farewell; it is the time to memorize...

It is the time to keep up smiling; it is the time to set up sailing...

Since 'C'est la vie!'

Contents

List of Figures

List of Tables

Notations

Notations

$\mathbf{a}_n \in \mathbb{C}^M$	steering vector for $\mathbf{p}_n$
$\mathbf{a}_k^* \in \mathbb{C}^K$	steering vector for $\mathbf{p}_k^*$
$\tilde{\mathbf{a}}_n \in \mathbb{C}^M$	beamforming vector for $\mathbf{p}_n$
$\mathbf{A} \in \mathbb{C}^{M\times N}$	steering matrix for $\mathbf{P}$
$\mathbf{A}^* \in \mathbb{C}^{M\times K}$	steering matrix for $\mathbf{P}^*$
$c_{i,j}$	impulse response of sensor array, $c_{i,j} \in \mathbf{C}$
$\mathbb{C}$	complex number domain
$\mathbf{C} \in \mathbb{R}^{N\times N}$	power propagation matrix
d	length of sensor array
D	distance from array to source plane
e_m	noise at sensor m
$\mathbf{e} \in \mathbb{C}^M$	background noises at all M sensors
f_l	frequency at lth bin
I	number of time sampling block
$\mathbf{I}_M \in \mathbb{R}^{M\times M}$	identity matrix with size M× M
K	number of real sources
L	number of frequency bins
M	number of microphone sensors
N	number of scanning points
$\mathbf{p}_n \in \mathbb{R}^3$	coordinates of discrete source n
$\mathbf{P} \in \mathbb{R}^{N\times 3}$	coordinate matrix of N discrete sources
$\overline{\mathbf{p}}_m \in \mathbb{R}^3$	coordinates of sensor m
$\overline{\mathbf{P}} \in \mathbb{R}^{M\times 3}$	coordinate matrix of M sensors
$\mathbf{p}_k^* \in \mathbb{R}^3$	coordinates of real source k
$\mathbf{P}^* \in \mathbb{R}^{K\times 3}$	coordinate matrix of K real sources
$r_{n,m}$	distance from source n to sensor m
$r_{-n,m}$	distance from source -n to sensor m
$\mathbb{R}$	real number domain
$\mathbf{R} \in \mathbb{C}^{M\times M}$	measured cross-spectrum matrix
s_n	discrete signal of source n
$\mathbf{s} \in \mathbb{C}^N$	discrete signals of N sources
s_k^*	signal of real source k
$\mathbf{s}^* \in \mathbb{C}^K$	signal vector of K real sources
T	total number of time samplings
v	wind speed
x_n	power of discrete source n
x_k^*	power of real source k
$\mathbf{x} \in \mathbb{R}^N$	power vector of N discrete sources
$\mathbf{x}^* \in \mathbb{R}^K$	power vector of K real sources
$\mathbf{X} \in \mathbb{R}^{N\times N}$	cross spectrum matrix of N discrete sources
$\mathbf{y} \in \mathbb{R}^N$	measured power vector of N sources
$\tilde{\mathbf{y}} \in \mathbb{R}^N$	beamforming power vector of N sources
$z_{i,m}$	measured signal in ith block at sensor m
$\boldsymbol{z}_i \in \mathbb{C}^M$	measured signals ith block at M sensors
α	regularization parameter
β	shape parameter in $\mathcal{DE}(x)$ model
δ_i	relative error of image reconstruction
γ	scale parameter in $\mathcal{DE}(x)$ model
λ	wavelength
μ	gradient step
ρ	reflecting coefficient
σ_e^2	background noise variance
$\tau_{n,m}$	time for distance $r_{n,m}$
$\tau_{-n,m}$	time for distance $r_{-n,m}$
θ	hyperparameters to be estimated
θ_0	known parameters
θ_1	hyperparameters in forward model
θ_2	hyperparameters in prior models
$\xi \in \mathbb{R}^N$	forward model uncertainty vector
ΔB	beamforming spatial resolution
Δp	scanning step (grid size of discretization)
$\overline{\Delta x^*}$	averaged estimation error of $\mathbf{x}^*$
$\mathbf{1}_N \in \mathbb{R}^N$	vector with N elements of 1

Operator

e	natural exponential
exp{}	natural exponential
$\mathbb{E}\{\cdot\}$	mathematical expectation
diag [·]	diagonal values of matrix
ln{·}	natural log
tr {·}	trace of matrix
$\nabla(\cdot)$	gradient
$\|\| \cdot \|\|$	spectral norm of square matrix
$\|\| \cdot \|\|_l$	l norm of a vector
$(\cdot)^T$	transpose
$(\cdot)^\dagger$	conjugate transpose
$(\cdot)^{(k)}$	kth iteration
$*$	convolution

Abbreviation

AGWN	Additive Gaussian White Noise
CMF	Covariance Matrix Fitting method
dB	decibel
DAMAS	Deconvolution Approach for Mapping of Acoustic Source
$\mathcal{DE}$	Double Exponential
DFT	Discrete Fourier Transform
DR	Diagonal Removal
$\mathcal{GG}$	Generalized Gaussian
i.i.d	independent and identically distributed
MAP	Maximum A Posterior
PDF	Probability Density Function
PSF	Point Spread Function
VBA	Variational Bayesian Approximation

Subscripts

m, n	term associated with m and n
-n	mirror source n
n′	equivalent source n
wb	wide-band

.

- Tous pour un, un pour tous.

Alexandre Dumas

1 Introduction

The acoustic imaging is an advanced technique for acoustic source localization and power reconstruction using limited measurements at microphone sensor array. This technique can provide the insights into the performance, properties and mechanisms of acoustic sources. Nowadays, high-resolution acoustic imaging has been widely studied and applied in reconstructing the acoustic source distributions on the stationary, moving and rotating objects etc. [74, 88, 81, 25, 50]. For example, figure 1.1(a) shows the acoustic map on the airplane surface, which can be used to measure the noise pollution near the airport when airplane is taking off and landing. In the imaging result, the red color represents the strong power noise, while the blue represents the weak power. And the imaging shape indicates the noise distribution and positions. Figure 1.1(b) shows the wind noise on the stator and rotor [121], which serves to design an efficient wind power system. Figure 1.1(c) shows the flap noise imaging on the aircraft gear, which indicates its aerodynamic performance. Figure 1.1(d) shows the acoustic imaging on train and truck, which reflects the acoustic influence of the transportation.

1.1 Problem statement

Unfortunately, acoustic imaging often causes a very ill-conditioned inverse problem, in which the solutions are not unique. So that it is hard to exactly reconstruct the source powers and positions. This is because the only things

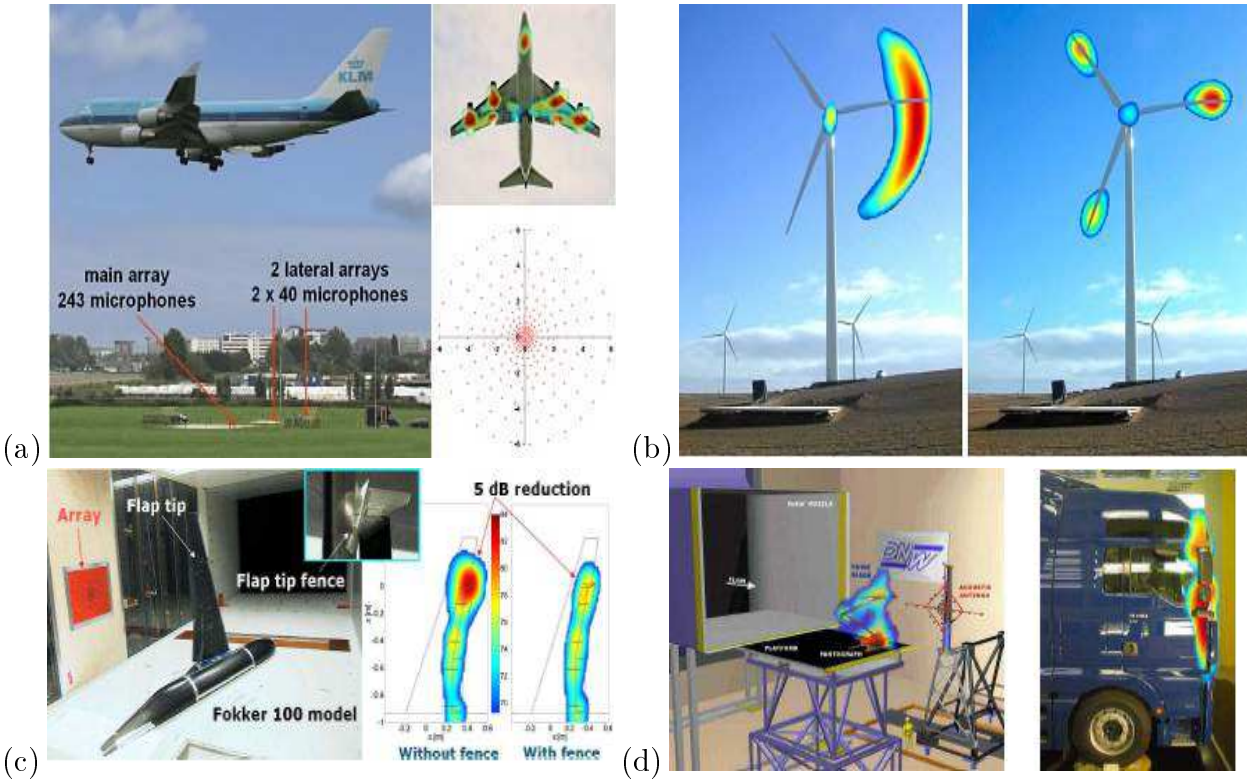

Figure 1.1: Examples of acoustic imaging researches at National Aerospace Laboratory (NLR) Holland [121]: (a) Airplane noise imaging. (b) Wind noise imaging on blade and rotor. (c) Flap noise imaging. (d) Acoustic imaging on train and truck.

we know are the limited number of measured signals, and microphone sensor configurations. But there remain at least 5 constraints to be investigated:

- Sensor measurements are limited. Due to the expensive cost, the total number of the microphones is usually few. Even if sensor number is large enough, the size of the sensor array should not be too large or too small due to the experiment requirements. Moreover, the number of measured time-samples (samplings) are sometimes very few due to the very short snap-shooting time. For example, there is a relatively very high speed between the moving objects and sensor platform. Even if the speed is quite small, the acoustic source instantaneously emit several signals in quite a while.

- Acoustic propagation paths are not easy to model exactly. Due to the acoustic reflection in the closed wind tunnel or non-anechoic chamber, the propagation paths usually consist in both direct and indirect parts. The latter ones might come from multi-reflections on the ground, ceiling and walls, thus become hard to describe their real multi-paths. And multi-path propagation often causes the acoustic reverberation effect, which is well known in the 'cocktail-party' problem [27]. Moreover, even

the direct path might be also difficult to determine due to the acoustic refraction in the non-uniform media. The propagation path is one of the essential factors which affect the acoustic power attenuation. Without a good propagation path model, it is difficult to accurately reconstruct the acoustic powers and positions.

- Background noise interference. There are at least two parts of it: the inherent noises at microphone sensors and the background noise of chamber (wind tunnel, room, concert hall for examples). The famous cocktail-party problem [27] can also well present this difficulty: background noises are too loud that acoustic signals are almost buried by noises.

- Spatial resolution. The term 'resolution' refers to the minimum distance between two distinguishable objects [65], and the spatial resolution of acoustic source is proportional to the acoustic wave length if other factors are fixed [125, 46]. In our common life, the acoustic signals whose frequencies are sensitive to human hearing are within $20 - 20000$Hz [16]. Since the acoustic propagation speed is $340m/s$ in the uniform of common air, the corresponding spatial resolution of are within the $0.8 - 800$cm. In other words, the lower the frequencies are, the worse the spatial resolution becomes.

- Acoustic dynamic range. It refers to the ratio between the largest and smallest powers of acoustic signals [16]. A good acoustic imaging result should achieve as wide as possible dynamic range, since it is better for us to hear and distinguish both loudest and the quietest sounds at the same time.

There are also other important issues on the acoustic sources, such as the directivity patterns, physical structures (monopole, extended or distributed) etc. But in this thesis, we just take main considerations on the above five aspects.

1.2 General methods

According to the acoustic physics and applications, acoustic imaging methods could be loosely classified as: Time-reversal acoustic imaging [77, 9], Near-field Acoustic Holography (NAH) [85] and inverse problems [111]. The inverse problem consists of using the measurements of forward physical model to

estimate parameters that characterize this forward model as pointed out in article [111]. Moreover, the inverse problems can be well developed by signal processing techniques and mathematical tools.

In this thesis, we focus on the inverse problems applied in acoustic imaging. The relevant methods usually involve in the following two aspects:

- A forward model of acoustic propagation. This model should consider the acoustic source model (signal or power model), propagation paths (direct or indirect path), as well as the background noise interference (Gaussian white or spatially non-stationary noises) as mentioned in article [74].
- Its inverse problem. The inverse solution often depends on measured acoustic data, the source spatial distribution (monopole or distributed), sensor array topology (uniform or random), as well as propagation types (near-field or far-field) as discussed in article [5].

According to these two aspects, many methods with advance signal processing techniques have been widely studied since decades:

- Spatial filter methods such as the Beamforming, Capon and MUSIC [118, 118, 22, 98] aim to solve the forward model of acoustic **signal** propagation. They can offer a direct and fast estimation for source localization. However, it is not easy to construct the source power distribution with high spatial resolution in background noises.
- Deconvolution methods such as the DAMAS and CLEAN [18, 8, 103, 123] are developed based on spatial filters, and they aim to solve the forward model of acoustic **power** propagation. These methods can iteratively deconvolve the blurry imaging results of spatial filters, and achieve the high resolution for both source localization and power reconstruction. However, deconvolution methods are still not robust to strong background noises. And some of deconvolution parameters should be carefully tuned for each application case.
- Regularization [112, 70, 38, 126, 107, 28] methods can further improve the spatial resolution of deconvolution methods in strong noise interference. They use the meaningful regularization term as the additional constraint into the deconvolution methods, and effectively improve the robustness to background noises. But the regularization term inevitably

increases the optimization complexity, and regularization parameter selection becomes an essential issue for practical use.

- Bayesian inference methods [115, 84, 87, 5, 24, 130, 29, 3] have been recently applied in acoustic imaging and successfully overcome the drawbacks of deconvolution and regularization methods. Bayesian inference is a powerful methodology for solving ill-posed inverse problem. It aims to estimate unknown random variables by applying the Bayes' rule to update the probability law: a posterior probability is obtained based on the likelihood and prior models. The likelihood can be derived from forward model and measured data. The prior models describe the knowledge or regularization on the unknown variables. Comparing to the classical deterministic methods, there are at least three advantages of Bayesian inference method:

  - Prior models perform like the regularization term. In fact, the Bayesian method via proper prior model is identical to the regularization methods. This is because prior models can appropriately translate the physical and statistical characteristics of unknown quantities into a concrete mathematical model. The prior models can bring not only novel information into the under-determined forward model, but also effectively regularize the ill-posed inverse problem and greatly reduce the solution uncertainty.
  - Important parameters of the deconvolution or regularization methods can also be seen as unknown quantities in Bayesian methods. These parameters can be embodied by the hyper-parameters in the prior and likelihood models. Some other unknown variables such as forward model uncertainty caused by multi-path propagation, as well as background noises at the sensor array, can also be modeled by proper prior models according to their statistical distributions, physical nature, upper or lower bonds and even our subjective presumption.
  - All unknown quantities can be estimated by well developed optimization methods such as the Maximum Likelihood (ML), Joint Maximum A Posterior (JMAP), Variational Bayesian Approximation (VBA) which are the analytical methods, as well as the Gibbs Sampling and sequential Monte Carlo methods which are the numerical methods. Not only the mean values of unknown random variables can be estimated by Bayesian inference, but also their un-

certainties (covariance matrix) in the same time. Indeed, the full posterior can be estimated automatically.

– However, one of the biggest limitations of Bayesian inference methods is the tremendous computational burden to get a global or subglobal optimization results due to non-quadratic or non-convex optimization. And some of optimizations cannot be implemented for practical use. So that it is a worthy work to balance the estimation performance and the calculating time.

Above all, there is no universal method for acoustic imaging which fit all purposes. We will discuss some of widely used the state-of-the-art methods in this thesis, and propose several effective approaches for the specific application: acoustic imaging on the vehicle surface in the wind tunnel experiment.

2

Motivation

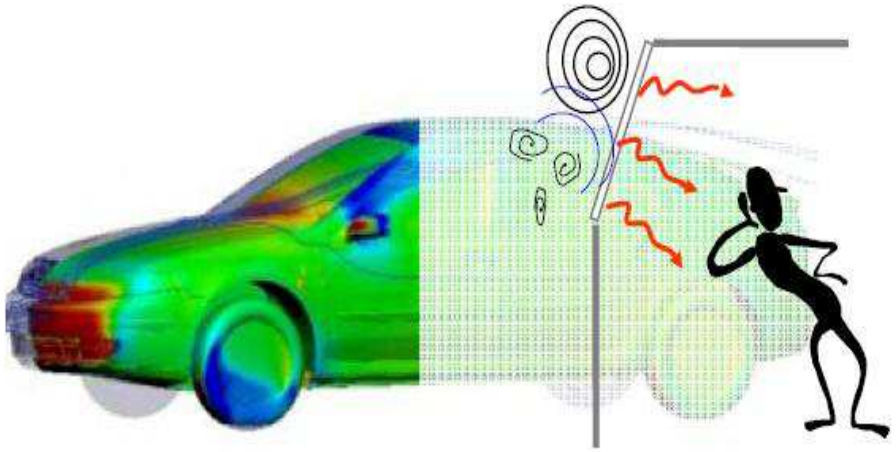

Figure 2.1: Acoustic imaging on vehicle surface [1].

In this thesis, we are going to obtain an acoustic imaging result on the surface of a fast moving car. So that we can tell on which car parts exist the acoustic sources, and how noisy they are. This imaging result will also contribute to design a more comfortable and more silent car as shown in Fig.2.1. There are many successful research works carried out in industry applications. For example, in Fig.2.2(a), we show the static vehicles (no engine noise) and microphone sensor arrays in the wind tunnel S2A [2] in Renault Automobile Company France. One of the objectives of this wind tunnel is to generate the wind flow at the speed of $160km/h$, so that this wind tunnel can simulate a fast running car on the high-way, and receive the acoustic signals produced by the conflicts between the air flow and moving

car. The results in Fig.(b) demonstrate the acoustic imaging results obtained by article [2], so that we can easily evaluate the acoustic influence to the drivers and passengers. While in Fig.2.2(c), another sensor array is provided in the same wind tunnel, so that the 3D acoustic imaging result on the vehicle surface can be obtained by article [86] in Fig.2.2(d).

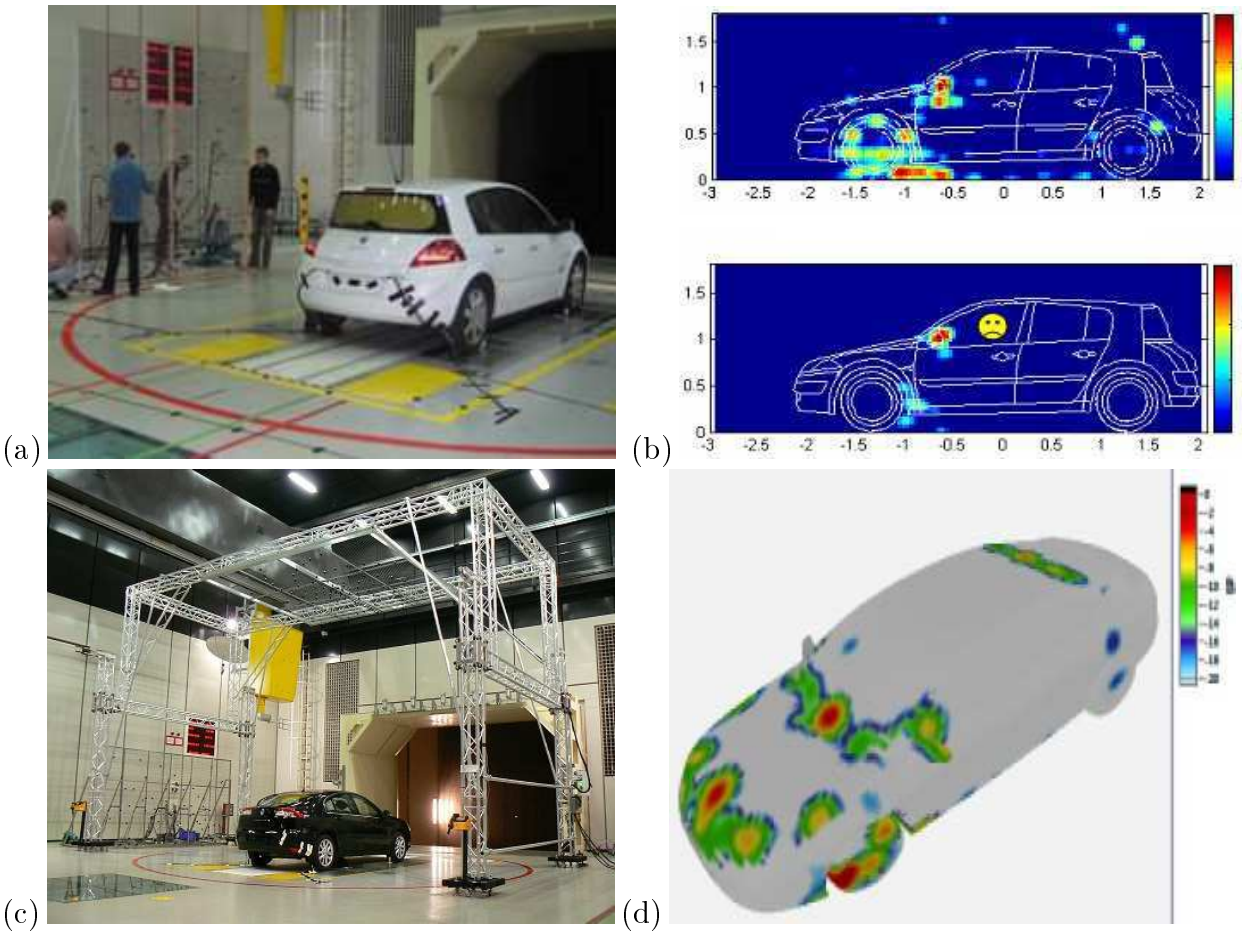

Figure 2.2: Acoustic imaging applications for acoustic influence of vehicles: (a) Wind tunnel S2A [2] (b) Acoustic imaging on vehicle surface [2] (c) Two sets of microphone sensor array in wind tunnel S2A [86] (d) 3D acoustic imaging on vehicle surface [86]

However, it is not an easy work to simultaneously obtain a robust acoustic imaging results with high spatial resolution and wide dynamic range of source powers. As discussed in the problem statement in Chapter 1.1, we should take at least five challenges into considerations:

- Limited measurements. There are limited microphone sensors to capture the acoustic data in this wind tunnel experiment, so that we can only build up a few equations between measured data and unknown source positions and powers on the vehicle surface. If the total number of these unknown variables are much larger than the equation number, the

acoustic imaging would be a very ill-conditioned inverse problem. In fact, it is often the case.

- Uncertainty of the multi-path propagation. In the wind tunnel, there inevitably exists ground reflection, since the car chassis is close to the ground. Moreover, the car is put inside the wind tunnel, but the sensor array are fixed outside. So that the air media from the car to sensors is not uniform, and there must be wind refraction at the interface between the wind flow and static air. Both the ground reflection and wind refration cause the uncertainty of the multi-path propagation. And this uncertainty can greatly affect the reconstruction of the source powers and positions.

- Noisy background. The wind tunnel generates wind flow as fast as $160km/h$, which itself might become a strong background noise. For simplicity, we suppose the background noise to be of Gaussian White distribution. Furthermore, microphone sensors might be different to each other, even if they come from the same products. Therefore, the inherent noise at each sensor could not be the same and might be of spatially non-stationary distribution. Moreover, although the sensor array is not located inside the wind flow, they might be still disturbed by the ground and wall reflection of wind flow. These unknown factors thus cause the noisy background and measurement errors.

- Super spatial resolution. In this wind tunnel experiment, we take the acoustic source at the frequency of $2500Hz$ for instance. Its spatial resolution caused by the corresponding wave length is about $20cm$, which is too low to accurately locate acoustic sources on rear-view mirrors, or on other small gears of a car. Classical methods for source localization can get high resolution to some extent, but still depend on wave length or sensor array geometry. In this thesis, we try to overcome the physical constraints and achieve much higher resolution than the classical methods, in which, we call it 'super resolution'. Here we have to point out the term 'super-resolution' in this thesis is not the same as the one in image processing.

- Acoustic dynamic range. For the issue of acoustic influence of a vehicle, both strong sources and weak ones at certain frequencies can greatly affect the hearing comfort. Therefore, a useful acoustic imaging result should reveal simultaneously the distributions of the loudest and the

quietest acoustic sources which both have potential impacts to drivers and passengers.

The above limitations motivate us to improve the acoustic imaging results in Fig.2.2(c) and (d), and provide several robust approaches with high resolution and wide dynamic range of acoustic powers in the strong background noises.

2.1 Main contributions

In this thesis, we first build up a discrete forward model of acoustic **signal** propagation. This signal model is an underdetermined linear system of equations linking measured data and unknown source signals and positions. Based on signal model, we set up a discrete forward model of acoustic **power** propagation [28, 29]. This power model is a determined and linear system of equations which can directly reflect the relationship between measured signals, source powers and positions. In the power model, we consider the measurement errors to be mainly composed of background noises at sensor array, model uncertainty caused by multi-path propagation, as well as model approximating errors.

For the inverse problem of the acoustic power model, we first propose a robust super-resolution approach with the sparsity constraint [28, 34, 30], so that we can obtain very high spatial resolution in strong measurement errors. But the sparsity parameter should be carefully estimated for effective performance. Then for the acoustic imaging with large dynamic range and super resolution, we propose a robust Bayesian inference approach with a sparsity enforcing prior [29, 32, 31, 33]. This sparse prior can better embody the sparsity characteristic of source distribution than the sparsity constraint. All the unknown variables and parameters can be automatically estimated by the Joint Maximum A Posterior (JMAP) estimation. However, this JMAP confronts a non-quadratic optimization and needs huge computation cost. In order to reduce the processing time of the JMAP, we investigate an invariant 2D convolution operator to approximate acoustic power propagation model [35]. Furthermore, we consider more realistic cases: the measurement errors are spatially non-stationary at different sensors. Also, the sparsity enforcing prior can be more accurately modeled by the Student-t priors. In these cases, the JMAP confronts more limitations than advantages. Therefore, we apply the Variational Bayesian Approximation (VBA) [35] to overcome the

drawbacks of the JMAP.

2.2 Thesis structure

This thesis is organized as follows: Chapter 3 briefly introduces the forward model of acoustic signal propagation. In Chapter 4, we present some of the-state-of-art methods for acoustic imaging: spatial filter, subspace decomposition, deconvolution and regularization methods. For super spatial resolution, we propose an effective approach using the sparsity constraint in Chapter 5. Chapter 6 presents the proposed Bayesian inference approach via a sparsity enforcing prior. In Chapter 7, we propose an invariant 2D convolution model to accelerate the proposed Bayesian method. So as to maintain the acceptable imaging results and moderate cost, we apply the VBA method based on the 2D convolution model in Chapter 8. In each chapter, simulations on monopole and complex acoustic source imaging will be made to reveal performance comparisons of the proposed approaches with state-of-the-art methods. Results of the real data in wind tunnel experiments are also used for method validation. To further prove the effectiveness of proposed approaches, we demonstrate method performance comparison on hybrid data, in which some known synthetic sources are added to the real data. Finally the conclusions and perspectives of the thesis are summarized in Chapter 9.

.

- Nous devons avoir la persévérance, surtout la confiance en nous-mêmes. Notre talent est utilisé pour réaliser quelque chose et on doit la réaliser à n'importe quel prix.

Marie Curie

3

Forward Model of Acoustic Signal Propagation

As we have stated in Introduction of Chapter 1, acoustic imaging refers to the inverse problem of source localization and power reconstruction from limited observations at microphone sensors. Before dealing with this inverse problem, in this chapter, we first introduce the acoustic source model, then build up the forward model of acoustic signal propagation, and this model can reflect the links between the known sensor measurements and unknown sources (signals and positions).

This chapter is organized as follows: important assumptions of this thesis are presented in Section 3.1. The acoustic signal model and propagation are introduced in Section 3.2. Then we present the forward model of acoustic signal propagation in Section 3.3. In order to linearize this forward model, we present a **discrete** forward model of acoustic signal propagation in Section 3.4. Finally, Section 3.5 concludes this chapter.

3.1 Assumptions

Before modeling, we have to make some necessary assumptions on acoustic sources, microphone sensors, reverberations in the wind tunnel and background noises as follows:

- Acoustic sources are simply supposed to be spatially punctual and uncorrelated monopoles as considered in many literature [37, 102, 126, 123, 96, 104]. This acoustic monopole model is used to simplify the physical properties and explicitly build up the forward model of acoustic propagation. We also use complex sources which are composed of several uncorrelated monopoles forming different spatial patterns as discussed in literature[15, 18]. All sources are located on the same plane.

- We consider the wave front of acoustic signal propagation to be a spherical wave in the near-field, and planar wave in the far-field. We don't account for the highly dispersive Rayleigh scattering sources.

- Microphone sensors are omni-directional with unitary gain, and are located on the same plane.

- Complex reverberations are negligible in wind tunnel, but we just consider the first order reflection on the ground.

- Since the sensors are located outside the wind tunnel, the wind refraction is supposed to take place in the interface between the wind flow inside the wind tunnel and the static air outside the tunnel.

- Background noises refer to the inherent noises at microphone sensors. The noises are independent of acoustic sources. For simplicity, we first model the background noise by an Additive Gaussian White Noise (AGWN), mutually independent and identically distributed (i.i.d) in Chapter 3-7. For more realistic applications, we propose a noise model to account for spatial non-stationarity at different sensors in Chapter 8.

- Acoustic propagation uncertainty is mainly caused due to the ground reflection and wind refraction in the wind tunnel. This uncertainty could not be exactly removed off, and we consider it to be also an i.i.d AGWN noise, independent of the background noises and acoustic sources.

- Other necessary assumptions will be specifically declared in the corresponding parts.

Owing to these necessary hypothesis, we can simplify the acoustic physical procedure, set up the analytical models and make good use of signal processing techniques.

3.2 Acoustic source signal model and propagation

Based on the monopole source assumption, we will introduce:

- The acoustic source signal model.
- The measured source signal model at the microphone sensor.

3.2.1 Acoustic signal model

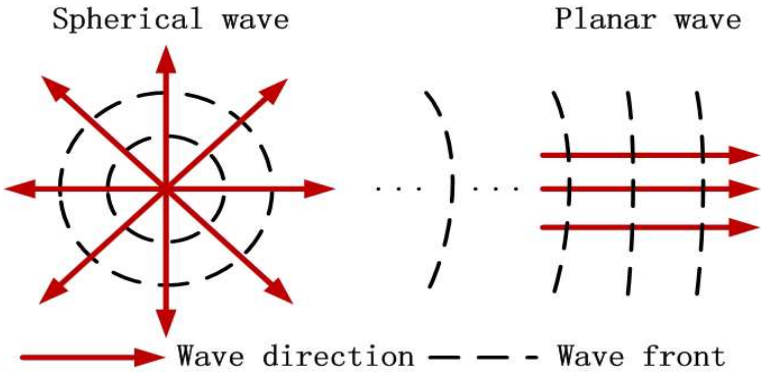

Figure 3.1: Spherical and planar wave fronts of a acoustic signal.

It is known that an acoustic signal propagates along the spatial-temporal 4D domain in the uniform media [65, 45] as shown in Fig.3.1: the acoustic wave front is in the form of the spherical wave in the near-field, and it gradually becomes the planar wave in the far-field. We first set up a Descartes coordinate system and note $\mathbf{O} = [0, 0, 0]^T$ to be the original point, where $(\cdot)^T$ denotes transpose operator. Suppose an acoustic monopole source fixed (static) at the position $\mathbf{p} = [p_x, p_y, p_z]^T$, and let r_s denote the distance from $\mathbf{p}$ to $\mathbf{O}$, defined as

$$r_s = ||\mathbf{p} - \mathbf{O}|| = \sqrt{p_x^2 + p_y^2 + p_z^2}\,. \tag{3.1}$$

Let $s(\mathbf{p}, t)$ represent the acoustic signal of monopole source. In the uniform media, the acoustic wave equation [125] at the position $\mathbf{p}$ and time t can be given as:

$$\frac{\partial^2 s(\mathbf{p},t)}{\partial p_x^2} + \frac{\partial^2 s(\mathbf{p},t)}{\partial p_y^2} + \frac{\partial^2 s(\mathbf{p},t)}{\partial p_z^2} = \frac{1}{c^2}\frac{\partial^2 s(\mathbf{p},t)}{\partial t^2}\,, \tag{3.2}$$

where c_0 denotes the acoustic propagation speed in the uniform media, taking the standard pressure and common temperature (20°C) for examples, $c =$

$340m/s$ in air and $c = 1481m/s$ in water; and $\frac{\partial^2(\cdot)}{\partial(\cdot)^2}$ denotes the second partial derivative.

For the acoustic signal $s(\mathbf{p}, t)$ on the spherical wave front, one of the solutions of mono-acoustic wave for Eq.(3.2) can be expressed as:

$$\begin{aligned} s(\mathbf{p}, t) =& A_0\, e^{-\jmath(\omega t - k\, r_s)} \\ =& A(t)\, e^{\jmath(\omega\, \Phi_0)}\,, \end{aligned} \tag{3.3}$$

where

- A_0 denotes the maximal signal amplitude;
- ω denotes the angular frequency, defined as $\omega = 2\pi\, f$ with f being the signal frequency;
- k denotes the wave number; in order to make signal model of Eq.(3.3) to satisfy Eq.(3.2), k should be defined as:

$$k = \frac{\omega}{c_0}\,, \tag{3.4}$$

  where the scalar value of k indicates the wave propagates omni-directionally at the speed of c_0, so that the wave front is spherical.
- $A(t) = A_0\, e^{-\jmath\omega\, t}$ denotes the signal amplitude, which is time-variant;
- $\Phi_0 = k\, r_s$ denotes the original signal phase, which only depends on source position $\mathbf{p}$ (the distance r_s).

3.2.2 Measured signal model at microphone sensors

Suppose there is a microphone sensor located at the 3D coordinates $\bar{\mathbf{p}} = [\bar{p}_x, \bar{p}_y, \bar{p}_z]^T$ as shown in Fig.3.2. Let τ denote the time delay from $\mathbf{p}$ to $\bar{\mathbf{p}}$; and Δr denotes the propagation distance between $\mathbf{p}$ and $\bar{\mathbf{p}}$, define as

$$\Delta r = ||\bar{\mathbf{p}} - \mathbf{p}|| = \sqrt{(p_x - \bar{p}_x)^2 + (p_y - \bar{p}_y)^2 + (p_z - \bar{p}_z)^2}\,, \tag{3.5}$$

where in the uniform media, it yields $\Delta r = c_0\, \tau$.

We suppose that amplitude attenuation is inversely proportional to the propagation distance, so that the maximal amplitude A_0 at the position $\mathbf{p}$ becomes $\frac{A_0}{\Delta r}$ at the sensor position $\bar{\mathbf{p}}$. According to the acoustic signal model in Eq.(3.3), signal amplitude $A(t) = A_0\, e^{-j\omega t}$ becomes $A(t-\tau) = \frac{A_0}{\Delta r}\, e^{-j\omega(t-\tau)}$

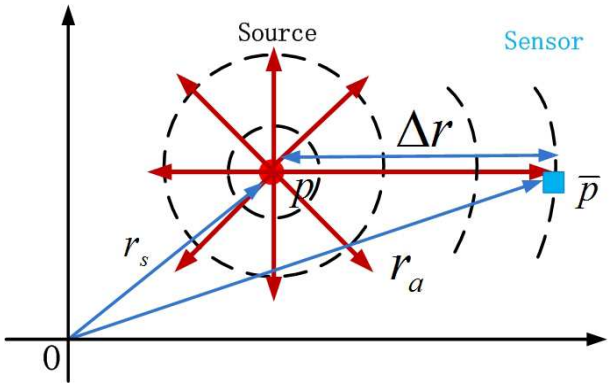

Figure 3.2: Measured signal at the sensor.

accounting for time delay τ. But the original signal phase $\Phi_0 = k\, r_s$ is unchanged, since Φ_0 does not depend on propagation distance Δr or delay time τ, but only depends on the original source position $\mathbf{p}$ as shown in Eq.(3.3).

Let $z(\Delta r, t)$ denote the noise-free measured signal at the sensor for the acoustic signal $s(\mathbf{p}, t)$. According to Eq.(3.3), $z(\Delta r, t)$ can be modeled as:

$$\begin{aligned} z(\Delta r, t) =& \frac{1}{\Delta r}\, A(t-\tau)\, e^{-\jmath \Phi_0} \\ =& \frac{1}{\Delta r}\, A(t)\, e^{-\jmath \omega \tau}\, e^{-\jmath \Phi_0} \\ =& \frac{1}{\Delta r}\, e^{-\jmath \omega \tau}\, s(\mathbf{p}, t) \\ =& \frac{1}{\Delta r}\, e^{-\jmath \omega \frac{\Delta r}{c_0}}\, s(\mathbf{p}, t)\,, \end{aligned} \tag{3.6}$$

where compared with source signal $s(\mathbf{p}, t)$, the measured signal $z(\Delta r, t)$ can be obtained by attenuating source signal amplitude by the factor of $\frac{1}{\Delta r}$, and by shifting the phase by $\omega\tau$, or $\omega \frac{\Delta}{r} c_0$ in the uniform media.

The Eq.(3.6) can be used with Fourier Transformation (FT) in the time domain, then we get

$$z(\Delta r, \omega) = a(\Delta r, \omega)\, s(\mathbf{p}, \omega)\,, \tag{3.7}$$

where $a(\Delta r, \omega)$ denotes the steering factor, defined as

$$a(\Delta r, \omega) = \frac{1}{\Delta r}\, e^{-\jmath \omega \tau}\,, \tag{3.8}$$

where $a(\Delta r, \omega)$ mainly depends on the source position $\mathbf{p}$ if the sensor position $\bar{\mathbf{p}}$ and angular frequency ω are both given. For the viewpoint of source

localization, $z(\Delta r, \omega)$ and $a(\Delta r, \omega)$ are both functions of unknown source position $\mathbf{p}$ if the sensor position $\bar{\mathbf{p}}$ and signal frequency are both given. So that we note $z(\mathbf{p}, f) = z(\Delta r, \omega)$ and $a(\mathbf{p}, f) = a(\Delta r, \omega)$ in the following parts of this thesis.

To sum up, equation (3.7) sets up an explicitly model of the measured signal $z(\Delta, \omega)$ and the source signal $s(\mathbf{p}, \omega)$ in the frequency domain. Based on Eq.(3.7), we can build up the forward model of signal propagation to the sensor array in the coming section.

3.3 Forward model of acoustic signal propagation

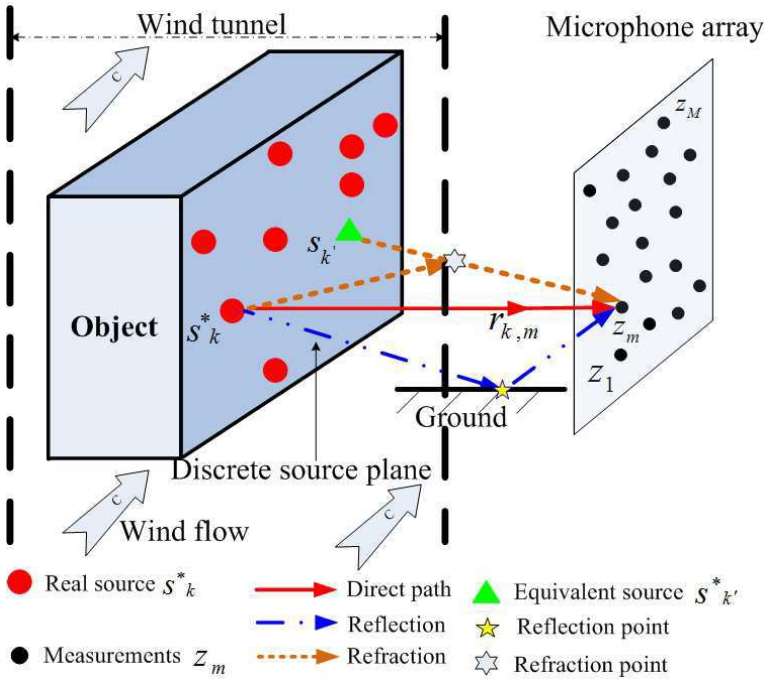

Figure 3.3: Illustration of acoustic signal propagation in wind tunnel[29].

Figure 3.3 illustrates the acoustic signal propagation from the source plane to the microphone sensor array in the wind tunnel, where sensors are installed outside the wind flow. On the source plane, we suppose K unknown original source signals $\mathbf{s}^* = [s_1^*, \cdots, s_K^*]^T$ at unknown positions $\mathbf{P}^* = [\mathbf{p}_1^*, \cdots, \mathbf{p}_K^*]^T$, where $\mathbf{p}_k^*$ denotes the 3D coordinates of kth original source signal s_k^*, and $(\cdot)^*$ represents notation of the original source. On the sensor plane, we consider M sensors at the known positions $\bar{\mathbf{P}} = [\bar{\mathbf{p}}_1, \cdots, \bar{\mathbf{p}}_M]^T$.

The signal processing procedure is illustrated in Fig.(3.4). For the mth microphone sensor with $m \in [1, \cdots, M]$, there are T samplings of acoustic

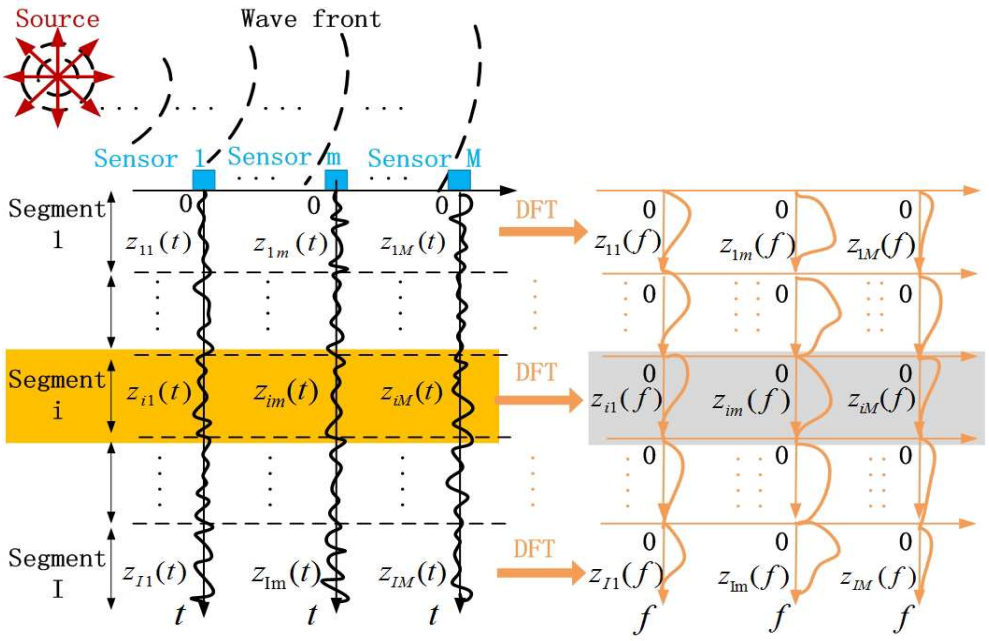

Figure 3.4: Illustration of the signal processing procedure in Eq.(3.9).

signals in time domain. Then these T temporal samplings are divided into I blocks with L samplings in each block, thus we note $\mathbf{z}_{i,m}(t)$ as the received signal of the ith sampling block ($i \in [1, \cdots, I]$) at the mth sensor ($m \in [1, \cdots, M]$) in the sampling time $t \in [(i-1)\,L+1, \cdots, i\,L-1]$, and total number of samplings is expressed as $T = I \times L$. Since the original source signals are usually of wide-band, we apply the Discrete Fourier Transform (DFT) in time domain to treat measured signals $\mathbf{z}_{i,m}(t)$ at each block in order to obtain L narrow frequency bins f_l (l$\in [1, \cdots, L]$). This procedure is made under the assumption of Gaussian Wide-Sense Stationary (WSS) process, so that we can independently make signal processing in each narrow frequency bin.

Let $\mathbf{z}_i(f_l) = [z_{i,1}(f_l), \cdots, z_{i,M}(f_l)]^T$ denote all measured signals at M sensors. According to the measured monopole signal at one sensor of Eq.(3.7) in the frequency domain, $\mathbf{z}_i(f_l)$ can be modeled [123] as shown in Fig.(3.5):

$$\mathbf{z}_i(f_l) = \mathbf{A}(\mathbf{P}^*, f_l)\, \mathbf{s}_i^*(f_l) + \mathbf{e}_i(f_l)\,, \tag{3.9}$$

where

- i denotes the ith sampling block;
- $\mathbf{e}_i(f_l) = [e_{i,1}(f_l), \cdots, e_{i,M}(f_l)]^T$, $\mathbf{e}_i(f_l) \in \mathbb{C}^M$ denotes i.i.d AGWN noise;
- $\mathbf{s}_i^*(f_l) = [s_{i,1}^*(f_l), \cdots, s_{i,K}^*(f_l)]^T$, $\mathbf{s}_i^*(f_l) \in \mathbb{C}^K$ denotes the DFT of the source signals;

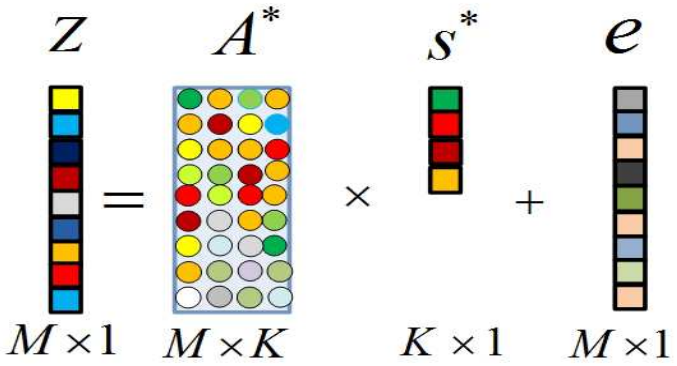

Figure 3.5: Illustration of the original signal equation in Eq.(3.9)

- $\mathbf{A}(\mathbf{P}^*, f_l) = [\mathbf{a}(\mathbf{p}_1^*, f_l) \cdots \mathbf{a}(\mathbf{p}_K^*, f_l)]$, $\mathbf{A}(\mathbf{P}^*, f_l) \in \mathbb{C}^{M \times K}$ denotes the steering matrix of original signal propagation.

Therefore, we get the forward model of original signal propagation in Eq.(3.9).

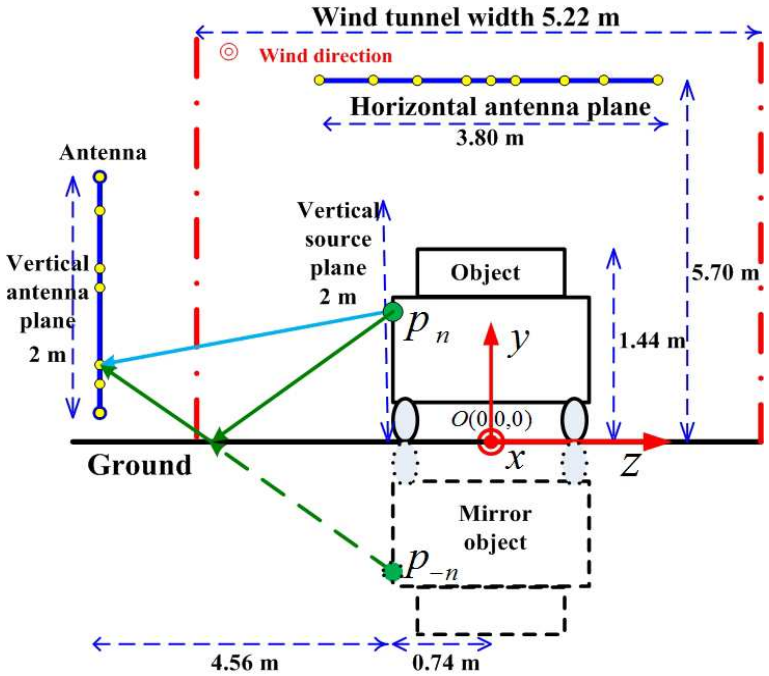

Figure 3.6: Ground reflection in the wind tunnel.

In particularly, we have to note that the conventional definition of $\mathbf{A}(\mathbf{P}^*, f_l)$ only models the direct propagation paths as discussed in many literature [123, 18, 128, 103]. Here, we also account for indirect propagation paths: one is the wind flow refraction happened on the interface between high

speed wind flow and normal static air as shown in Fig.3.3; if this refraction could not be considered, an equivalent source will be detected beside the original one due to the wind flow (direction and speed). The other phenomenon is the ground reflection as shown in Fig.3.6; since the object vehicle is very close to the ground, the ground reflection can not be negligible in wind tunnel experiments; if this reflection is not taken into account, ghost shadows will be overlapped on the original sources. Therefore the steering vector $\mathbf{a}(\mathbf{p}_k^*, f_l)$ is not only composed of the direct propagation vector $\mathbf{a}_d(\mathbf{p}_k^*, f_l)$, but also the ground reflection vector $\mathbf{a}_r(\mathbf{p}_{-k}^*, f_l)$ as follows:

$$\mathbf{a}(\mathbf{p}_k^*, f_l) = \mathbf{a}_d(\mathbf{p}_k^*, f_l) + \rho\, \mathbf{a}_r(\mathbf{p}_{-k}^*, f_l)\,, \tag{3.10}$$

where

- ρ denotes the reflecting coefficient ($0 \leq \rho \leq 1$), whose value mainly depends on ground conditions (material, temperature, humidity, etc.). For the real data used in this thesis, $\rho = 0.8$ is empirically estimated for the wind tunnel S2A experiments [2], thanks to the contribution of researchers in Renault Lab.

- $\mathbf{a}_d(\mathbf{p}_k^*, f_l)$ denotes the direct propagation vector in Eq.(3.10), defined as:

$$\mathbf{a}_d(\mathbf{p}_k^*, f_l) = \left\{ \frac{1}{r_{k,1}} \exp\left[-\jmath(2\pi f_l \tau_{k,1})\right], \cdots, \frac{1}{r_{k,M}} \exp\left[-\jmath(2\pi f_l \tau_{k,M})\right] \right\}^T, \tag{3.11}$$

  where $\tau_{k,m}$ is the propagation time from s_k^* to sensor m, and $r_{k,m}$ is the propagation distance during $\tau_{k,m}$; in the uniform media, $r_{k,m} = \tau_{k,m}\, c_0$ is the geometric distance from the original source position $\mathbf{p}_k^*$ to the known sensor position $\bar{\mathbf{p}}_m$.

- $\mathbf{a}_r(\mathbf{p}_{-k}^*, f_l)$ denotes the indirect steering vector of ground reflection, defined as:

$$\mathbf{a}_r(\mathbf{p}_{-k}^*, f_l) = \left\{ \frac{1}{r_{-k,1}} \exp\left[-\jmath(2\pi f_l\, \tau_{-k,1})\right], \cdots, \frac{1}{r_{-k,M}} \exp\left[-\jmath(2\pi f_l\, \tau_{-k,M})\right] \right\}^T, \tag{3.12}$$

  where $\mathbf{p}_{-k}^*$ denotes mirror positions of s_k^*, thus $\mathbf{p}_{-k}^*$ and $\mathbf{p}_k^*$ are symmetric to the ground.

- Wind refraction effect in the steering vector $\mathbf{a}(\mathbf{p}_k^*, f_l)$. Since the sensor array locates outside the wind flow, the medium for acoustic propagation is not uniform, thus the refraction will happen on the interface between

the common air and wind flow. In Appendix A and B, the actual $r_{k,m}$, $r_{-k,m}$ and $\tau_{k,m}$, $\tau_{-k,m}$ are formulated in details: we apply the concepts of equivalent sources and mirror sources to deal with the wind refraction and ground reflection respectively.

In the followings of this thesis, the signal processing will be carried out in the frequency domain. So that we omit f_l in all the functions and variables for simplicity.

3.4 Discrete forward model of acoustic signal propagation

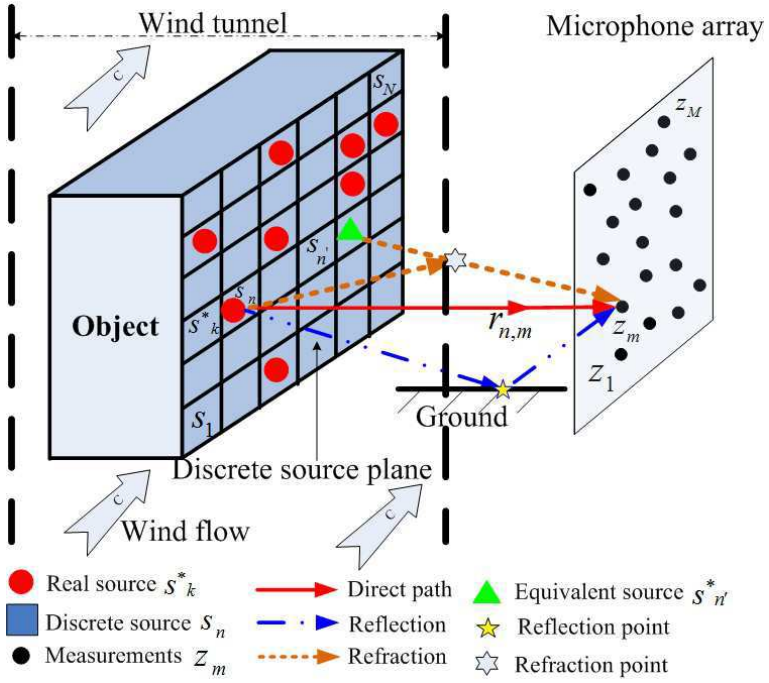

Figure 3.7: Discretization of acoustic signal propagation in wind tunnel [29].

In Eq.(3.9), it is thus a nonlinear system of equations, since source positions $\mathbf{P}^* = [\mathbf{p}_1^*, \cdots, \mathbf{p}_K^*]^T$ and source signals $\mathbf{s}^* = [s_1^*, \cdots, s_K^*]^T$ are both unknown. The relationship between measured data $\mathbf{z}$ and $\mathbf{s}^*$, $\mathbf{P}^*$ is not linear. If $\mathbf{P}^*$ is fixed and known, equation (3.9) is linear to $\mathbf{s}^*$. But for given $\mathbf{s}^*$, equation (3.9) is still non-linear with respect to $\mathbf{P}^*$. In order to transform Eq.(3.9) into a linear system, the classical inverse problem is based on the discretization of the source plane, as illustrated in Fig.3.7. The source plane is equally discretized into N grid at the known positions $\mathbf{P} = [\mathbf{p}_1, \cdots, \mathbf{p}_N]^T$,

and we assume that K original sources $\mathbf{s}^*$ sparsely distribute on these grids, satisfying $N > M > K$ and $\mathbf{P}^* \subset \mathbf{P}$, where $\subset$ denotes that the set $\mathbf{P}^*$ belongs to the set $\mathbf{P}$. For grid size Δp, too large value cannot reach high spatial resolution; too small, it will cause great of computation complexity. According to recent studies, literature [18] points out that Δp should satisfy $\Delta p/\Delta B < 0.2$, where ΔB denotes the spatial resolution of conventional Beamforming method, which will be introduced in Chapter 4.

Then we get N discrete source signals $\mathbf{s} = [s_1, \cdots, s_N]^T$ at known positions $\mathbf{P}$; and we suppose the relationship between s_n and s_k^* in the space domain as:

$$\begin{cases} s_n = s_k^* & \text{for} \quad \mathbf{p}_n = \mathbf{p}_k^* \\ s_n = 0 & \text{for} \quad \mathbf{p}_n \neq \mathbf{p}_k^* \end{cases} . \tag{3.13}$$

where the smaller grid is discretized, the higher spatial resolution can be achieved. Based on the assumptions, since the number K of original source signals $\mathbf{s}^*$ is limited, and the number N of source signals $\mathbf{s}$ is much larger (K<N), so that $\mathbf{s}$ is a sparse signal with K sparsity in the space domain. Therefore, to reconstruct original source signals $\mathbf{s}^*$ is transferred to reconstruct K-sparsity signals $\mathbf{s}$. Meanwhile, original source position $\mathbf{p}_k^*$ can be deprived from the discrete position $\mathbf{p}_n$, where the source power of discrete signal s_n is not zero-value.

To be clear, we state in the following parts that $\mathbf{s}^* = [s_1^*, \cdots, s_K^*]^T$ denotes the original source signals, while $\mathbf{s} = [s_1, \cdots, s_N]^T$ denotes the discrete source signals. The term 'source n' means the discrete source signal s_n at the discrete position $\mathbf{p}_n$. The term 'original source k' means the original source signal s_k^* at the original position p_k^*.

Considering the Eq.(3.13) and wide-band signal processing in Fig.3.4, the discrete signals $\mathbf{s}_i$ of the ith sampling block can be also described as

$$\mathbf{s}_i = [0, \cdots, s_{i1}^*, 0, \cdots, s_{ik}^*, 0, \cdots, s_{iK}^*, 0, \cdots]_{N\times 1} . \tag{3.14}$$

where $\mathbf{s}_i$ of the ith sampling block is also a sparse signal with K-sparsity in the spatial domain. After DFT transformation in time domain, $\mathbf{s}_i(f_l)$ still maintain the sparsity in space domain. The signal processing will be made independently for each frequency bin, thus in the following, we omit the frequency notation f_l for simplicity.

In Eq.(3.9), replacing $\mathbf{P}^*$ by $\mathbf{P}$ and $\mathbf{s}_i^*$ by $\mathbf{s}_i$, we get the discrete forward model of signal propagation based on the discretized source plane as follows:

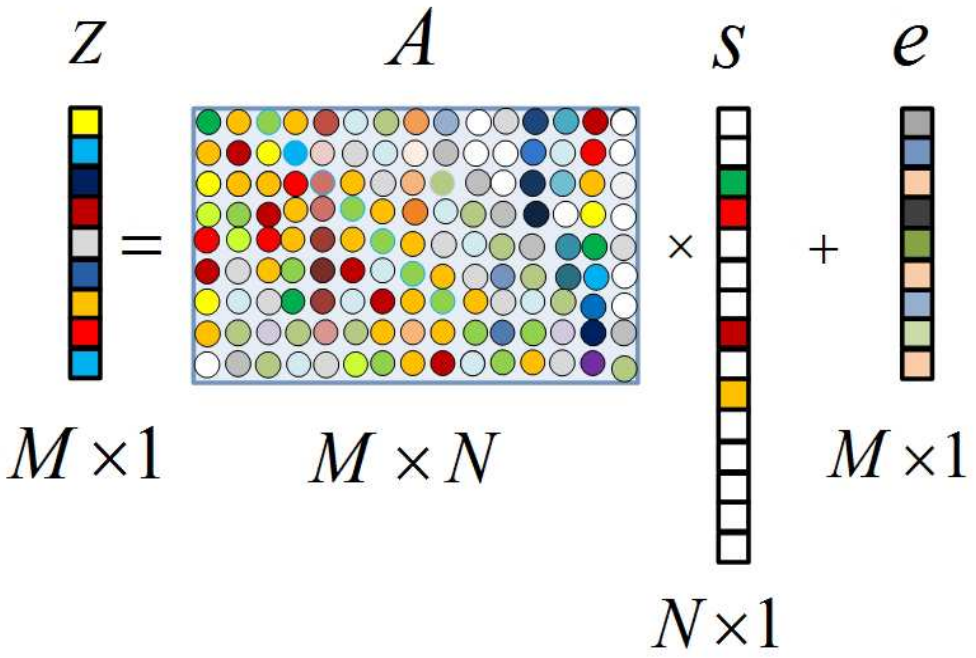

Figure 3.8: Illustration of the sparse signal equation in Eq.(3.15)

$$\mathbf{z}_i = \mathbf{A}(\mathbf{P})\,\mathbf{s}_i + \mathbf{e}_i\,, \tag{3.15}$$

where Eq.(3.15) is a sparse signal equation as shown in Fig.(3.8); $i = 1, \cdots, I$ denotes the index of sampling blocks; and $\mathbf{A}(\mathbf{P}) = [\mathbf{a}(\mathbf{p}_1) \cdots \mathbf{a}(\mathbf{p}_N)]$, $\mathbf{A}(\mathbf{P}) \in \mathbb{C}^{M\times N}$ consists of N steering vectors $\mathbf{a}(\mathbf{p}_n)$, which is similarly defined from Eq.(3.10) as follows:

$$\mathbf{a}(\mathbf{p}_n) = \mathbf{a}_d(\mathbf{p}_n) + \rho\,\mathbf{a}_r(\mathbf{p}_{-n})\,, \tag{3.16}$$

where $\mathbf{a}_d(\mathbf{p}_n)$ and $\mathbf{a}_r(\mathbf{p}_{-n})$ are similarly defined from Eq.(3.11) and Eq.(3.12) respectively as:

- $\mathbf{a}_d(\mathbf{p}_n)$ denotes the direct steering vector, defined as:

$$\mathbf{a}_d(\mathbf{p}_n) = \left\{ \frac{1}{r_{n,1}} \exp\left[-\jmath(2\pi f_l \tau_{n,1})\right], \cdots, \frac{1}{r_{n,M}} \exp\left[-\jmath(2\pi f_l \tau_{n,M})\right] \right\}^T, \tag{3.17}$$

  where $\tau_{n,m}$ is the propagation time from sourice n to sensor m, and $r_{n,m}$ is the propagation distance during $\tau_{n,m}$; in the uniform media, it yields $\tau_{n,m} = \frac{r_{n,m}}{c}$ and $r_{n,m}$ is the geometric distance from the discrete source position $\mathbf{p}_n$ to the known sensor position $\bar{\mathbf{p}}_m$.

- $\mathbf{a}_r(\mathbf{p}_{-n})$ denotes the ground reflection vector, defined as:

$$\mathbf{a}_r(\mathbf{p}_{-n}) = \left\{ \frac{1}{r_{-n,1}} \exp\left[-\jmath(2\pi f_l \tau_{-n,1})\right], \cdots, \frac{1}{r_{-n,M}} \exp\left[-\jmath(2\pi f_l \tau_{-n,M})\right] \right\}^T , \tag{3.18}$$

  where $\mathbf{p}_{-n}$ denotes mirror positions of source n, thus $\mathbf{p}_{-n}$ and $\mathbf{p}_n$ are symmetric to the ground.

- Wind refraction. Since the sensor array locates outside the wind flow, the medium for acoustic propagation is not uniform, the actual $r_{n,m}$, $r_{-n,m}$ and $\tau_{n,m}$, $\tau_{-n,m}$ are formulated in Appendix A and B.

For simplicity, $\mathbf{a}_n$ is short for $\mathbf{a}(\mathbf{p}_n)$; and we take similar short forms for the other steering vectors in Eq.(3.10) and Eq.(3.16).

Therefore, the forward model of Eq.(3.15) is the linear system of equations for source signals $\mathbf{s}$. The original source positions $\mathbf{P}^*$ can be estimated from discrete positions $\mathbf{P}$ where the corresponding $\mathbf{s}$ are non-zero values. However, since the source number N is usually larger than the sensor number M, equation (3.15) thus is quite underdetermined, so that some extra constraints should be involved in order to obtain stable and unique solutions of $\mathbf{s}$.

In short, forward model of signal propagation in Eq.(3.15) is an underdetermined linear system of equations for discrete source signals $\mathbf{s}_i$, since the measured signals $\mathbf{z}_i$ are known, and signal propagation matrix $\mathbf{A}(\mathbf{P})$ can be calculated from Eq.(3.16) owing to the known discrete positions $\mathbf{P}$ and known sensor positions.

If $\mathbf{s}$ could be well estimated from $\mathbf{z}$, the discrete source powers $\mathbf{x} = [x_1, \cdots, x_N]^T$ of uncorrelated sources are thus calculated as the diagonal elements of the source cross-spectrum matrix $\mathbf{R}_s$:

$$\mathbf{x} = \mathrm{diag}\left[\mathbf{R}_s\right] , \tag{3.19}$$

where $\mathrm{diag}\left[\cdot\right]$ denotes the diagonal items; $\mathbf{R}_s$ denotes the cross-spectrum matrix of the discrete source signals $\mathbf{s}$, defined as

$$\mathbf{R}_s = \mathbb{E}[\mathbf{s}\mathbf{s}^\dagger] , \tag{3.20}$$

where $\mathbb{E}[\cdot]$ denotes the mathematical expectation. But in practice, $\mathbf{R}_s$ can be approximated by $\hat{\mathbf{R}}_s$ as

$$\hat{\mathbf{R}}_s = \frac{1}{I} \sum_{i=1}^{I} \mathbf{s}_i \mathbf{s}_i^\dagger , \tag{3.21}$$

where $\mathbf{s}_i$ is the discrete source signal in the ith sampling block number as defined in Eq.(3.14); I is the total number of sampling blocks. The larger I is, the more closely $\hat{\mathbf{R}}_s$ approaches $\mathbf{R}_s$. Similarly, we can define the real source powers $\mathbf{x}^*$ and cross-spectrum matrix $\mathbf{R}_s^*$ of real source signals.

Finally, we can obtain original source powers $\mathbf{x}^*$ and positions $\mathbf{P}^*$ from the estimated discrete powers $\mathbf{x}$ and positions $\mathbf{P}$ as follows: according to the sparsity in Eq.(3.14), the non-zero element $x_n \in \mathbf{x}$ and its discrete position $\mathbf{p}_n \in \mathbf{P}$ could be the estimations of the real source power $x_n^* \in \mathbf{x}^*$ and its position $\mathbf{p}_n^* \in \mathbf{P}^*$

3.5 Conclusions and perspectives

In this thesis, using a monopole (punctual) acoustic source model, we explicitly build up the forward model of acoustic signal propagation from the source plane to the microphone sensor array. In this model, we consider the multi-path propagation effect such as ground reflection and wind refraction in the wind tunnel. However, this model is a non-linear function of unknown source signals and their unknown positions. For linear transformation, we set up the discrete forward model by discretizing the source plane, in which, the discrete source positions are known and are supposed to be able to contain the original source positions.

Based on this linear discrete signal model, we transfer it into acoustic power propagation model, and directly estimate the power distribution, which will be introduced in Chapter 4. Moreover, we make good use of the source sparse distribution so as to reconstruct the source powers and positions in our proposed approaches in the Chapter 5-8.

.

- Pas de grand désir, pas de grand talent.

Honoré de Balzac

4

State-of-the-art Methods in Source Localization

In Chapter 3, the forward model of acoustic signal propagation in Eq.(3.9) reveals the relation between known observations $\mathbf{z}$ at sensors and unknown source signals $\mathbf{s}^*$ and positions $\mathbf{P}^*$ in the frequency domain. Unfortunately, Eq.(3.9) is a **non-linear** system of equations. Thus we have to transfer Eq.(3.9) into a **linear** model by giving the **discrete** forward model in Eq.(3.15), in which, discrete source positions $\mathbf{P}$ are known. But this discrete forward model is usually quite under-determined.

In this chapter, we want to present several state-of-the-art methods: some of them aim to first estimating source powers and then obtain source positions. Others first focus on the source localization and then power reconstruction. No matter what kind of the estimation strategies are adopted, we will show the advantages and limitations of these classical methods. In the following chapters, we will propose our approaches based on these classical methods so as to achieve an effective acoustic imaging on the vehicle surface in the wind tunnel tests.

This chapter is organized as: In Section 4.1, we present the necessity of transformation from signals to powers in acoustic imaging. Acoustic signal model and signal propagation are introduced in Section 3.2. Then some State-of-the-art methods are introduced, such as the spatial filter methods in Section 4.2. Based on beamforming method, we present the forward model of

acoustic power propagation in Section 4.3. According to this power model, we introduce the deconvolution methods in Section 4.4 and the regularization in 4.5. Method comparisons are shown on simulations in Section 4.7, real data of wind tunnel experiments in Section 4.8, as well as the hybrid data 4.9, in which, some known simulated source are added to the real data. Finally, Section 4.10 concludes this chapter.

4.1 From signal reconstruction to power reconstruction

The direct estimation of discrete signals $\mathbf{s}$ from the discrete forward model of signal propagation in Eq.(3.15) often confronts two difficulties:

- The under-determined equation (3.15) causes the very ill-posed inverse problem, since the total number M of linear equations is often much less than total number N of the unknown discrete source signals $\mathbf{s}$. In particular, when very large number of grids are used to discretize the source plane in order to achieve the high spatial resolution imaging, there will be huge dimension ($N >> M$) of discrete source signals $\mathbf{s}$.

- The discrete source signals $\mathbf{s}$ modeled by Eq.(3.3) are of complex values, which contains three kinds of unknown variables such as amplitudes, phases and characteristic frequencies. Even though we can apply the DFT to analyze the frequency spectrum over the interested discrete frequencies f_l, with $l \in [1, \cdots, L]$ as shown in Fig.3.4, it still remains the unknown amplitudes and phases to be estimated. Even though the complex signals $\mathbf{s}$ can be transformed into two groups of real signals, this procedure inevitably increases the problem complexity and computational burden.

Moreover, as stated in Introduction of Chapter 1, the acoustic imaging issue mainly involves on the source power reconstruction and localization, instead of estimating the signal phases, characteristic frequencies or other signal parameters. Therefore, it is highly necessary to directly obtain the source powers $\mathbf{x}$ from measured signals $\mathbf{z}$ at sensor array, which will be discussed in Section 4.3

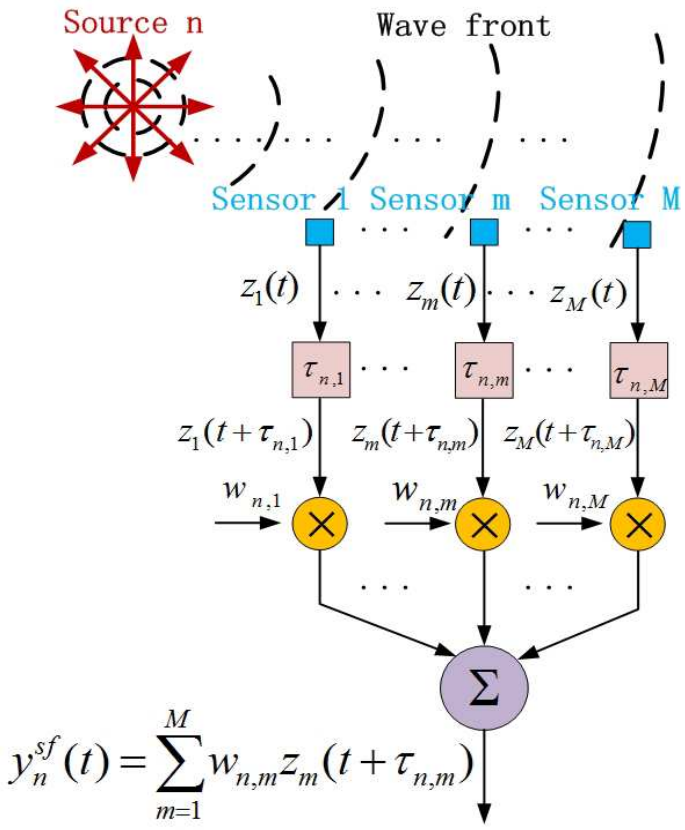

Figure 4.1: Illustration of the spatial filter in Eq.(4.1).

4.2 Spatial filter

Let us look back at Eq.(3.15) and the steering vector $\mathbf{a}_n$ of Eq.(3.16)-Eq.(3.18), the measured signals $\mathbf{z}(f) = [z_1(f), \cdots, z_m(f), \cdots, z_M(f)]^T$ at sensors are modeled by discrete signals $\mathbf{s}(f) = [s_1(f), \cdots, s_n(f), \cdots, z_N(f)]^T$ with amplitude attenuations and phase shifts in the frequency domain. Correspondingly in the time domain, $\mathbf{z}(t)$ can be modeled by $\mathbf{s}(t)$ with the same amplitude attenuation and time delays which are caused by propagation distances from sources to sensors. The above procedure is analogous to designing a spatial filter [36, 117, 60, 80] which can be used to compensate the time delays and amplitude attenuation, and reconstruct discrete signals $\mathbf{s}(f)$ as shown in Fig.(4.1):

$$s_n(t) \simeq y_n^{sf}(t) = \sum_{m=1}^{M} w_{n,m}\, z_m(t + \tau_{n,m})\,, \tag{4.1}$$

where $y_n^{sf}(t)$ denotes the spatial filter output, which can be an estimation for the nth discrete source signal $s_n(t)$ at the position $\mathbf{p}_n$, namely $s_n(t) \simeq y_n^{sf}(t)$; and $z_m(t)$ denotes the measured signal at the mth sensor; $\tau_{n,m}$ is the time delay from the nth source to the mth sensor, defined in Eq.(3.17); $w_{n,m}$ with $m \in [1, \cdots, M]$ denotes the filter coefficient at the mth sensor for the source signal s_n. After DFT in time domain, equation (4.1) becomes

$$\begin{aligned} s_n(f_l) \simeq y_n^{sf}(f_l) &= \sum_{m=1}^{M} w_{n,m}\, z_m(f_l)\, e^{\jmath\, 2\pi\, f_l\, \tau_{n,m}} \\ &= \mathbf{w}^{\dagger}(\mathbf{p}_n, f_l)\, \mathbf{z}(f_l)\,, \end{aligned} \tag{4.2}$$

where f_l is the lth frequency bin defined in Eq.(3.9); $\mathbf{z} = \mathbf{A}(\mathbf{P})\,\mathbf{s} + \mathbf{e}$ is sensor measurements modeled by Eq.(3.15); $\mathbf{w}(\mathbf{p}_n, f_l)$ denotes the spatial filter vector for the source s_n located at the position $\mathbf{p}_n$, which is defined as

$$\mathbf{w}(\mathbf{p}_n, f_l) = \left[w_{n,1}\, e^{-\jmath\, f_l\, \tau_{n,1}}, \cdots w_{n,m}\, e^{-\jmath\, f_l\, \tau_{n,m}}, \cdots, w_{n,M}\, e^{-\jmath\, f_l\, \tau_{n,M}}\right]^T\,, \tag{4.3}$$

where the spatial filter $\mathbf{w}(\mathbf{p}_n, f_l)$ serves to offer the gains $w_{n,m}$ and compensate the phase shift due to the time delay $\tau_{n,m}$ for the nth discrete source signal s_n coming from the position $\mathbf{p}_n$ at a given frequency f_l. For simplicity, we omit f_l and take $\mathbf{w}_n$ for $\mathbf{w}(\mathbf{p}_n, f_l)$ afterwards.

In order to get an effective output y_n^{sp} for the source signal s_n estimation, the expected spatial filter $\mathbf{w}_n$ should meet the following constraints:

- P1: $\mathbf{w}_n$ can make the source signal s_n coming from the position $\mathbf{p}_n$ to well pass the spatial filter without distortions. In the forward model of signal propagation in Eq.(3.15), the steering vector $\mathbf{a}_n$ in Eq.(3.16) embodies the signal amplitude attenuation and phase shift. In order to meet P1, $\mathbf{w}_n$ should satisfy $\mathbf{w}_n^{\dagger}\,\mathbf{a}_n = 1$.
- P2: $\mathbf{w}_n$ can attenuate all the other sources coming from the other positions.

In order to interpret P2, one of the criteria to optimize $\mathbf{w}_n$ is to obtain the maximal output power only if the source signal comes from the position $\mathbf{p}_n$. Based on output signal of Eq.(4.2), let us first define the output power y_n^{sf} as:

$$y_n^{power} = \mathbb{E}[|y_n^{sf}|^2] = \mathbf{w}_n^{\dagger}\,\mathbf{R}_z\,\mathbf{w}_n \tag{4.4}$$

where $|\cdot|$ denotes the scalar modulus; $\mathbf{R}_z$ denotes the cross-spectrum matrix of measured signals $\mathbf{z}$ at the sensor array, defined as:

$$\mathbf{R}_z = \mathbb{E}[\mathbf{z}\,\mathbf{z}^{\dagger}]\,, \tag{4.5}$$

where measured signals $\mathbf{z} = [z_1, \cdots, z_M]^T$ are modeled by $\mathbf{z}_i = \mathbf{A}(\mathbf{P})\,\mathbf{s}_i + \mathbf{e}_i$ in the discrete forward model of Eq.(3.15). In practice, $\mathbf{R}_z$ is calculated by

$$\hat{\mathbf{R}}_z = \frac{1}{I}\sum_{i=1}^{I} \mathbf{z}_i\,\mathbf{z}_i^{\dagger}\,, \tag{4.6}$$

where $i \in [1, \cdot, I]$ is the sampling block number, and I is total number of sampling blocks; $\mathbf{z}_i$ is measured signal vector in the ith sampling block at all M sensors. It is seen that the greater I is, the more closely $\hat{\mathbf{R}}_z$ approaches $\mathbf{R}_z$.

Therefore, according to the P1 and P2 properties, the expected spatial filter $\mathbf{w}_n$ should satisfy

$$\begin{cases} P1: & \mathbf{w}_n^{\dagger}\,\mathbf{a}_n = 1 \\ P2: & \mathbf{w}_n = \arg\max_{\mathbf{w}_n}\{y_n^{power}\} = \arg\max_{\mathbf{w}_n}\left\{\mathbf{w}_n^{\dagger}\,\hat{\mathbf{R}}_z\,\mathbf{w}_n\right\} \end{cases}, \tag{4.7}$$

where $\mathbf{a}_n$ is the steering vector defined in Eq.(3.16) for the nth discrete source signal at the position $\mathbf{p}_n$, with $\mathbf{p}_n \in \mathbf{P}$, $n \in [1, \cdots, N]$ on the discrete source plane.

4.2.1 Conventional Beamforming

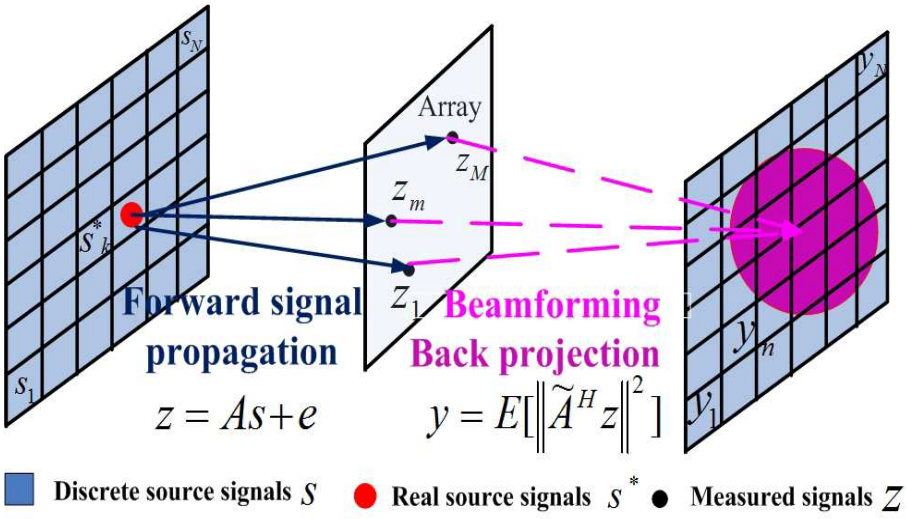

Figure 4.2: Illustration of beamforming.

Equation (4.7) uses the criterion that the maximal output power of the expected spatial filter should only come from the source position $\mathbf{p}_n$. In other words, the expected filter should decrease filter output powers coming

from the all other positions. One of the most commonly used spatial filters is the Conventional BeamForming (CBF) method [118, 25], which offers a simple, direct and coarse estimation for the source powers and positions as shown in Fig.(4.2). The CBF assumes background noise to be i.i.d AGWN and independent of the source signal. According to spatial filter in Eq.(4.7), to maximize the filter output power of the source signal is equivalent to minimize the output noise power as follows:

$$\begin{cases} P1: & \tilde{\mathbf{a}}_n^{\dagger}\,\mathbf{a}_n = 1 \\ P2: & \tilde{\mathbf{a}}_n = \arg\min_{\tilde{\mathbf{a}}_n}\left\{\tilde{\mathbf{a}}_n^{\dagger}\,\sigma_e^2\,\mathbf{I}\,\tilde{\mathbf{a}}_n\right\} \end{cases}, \tag{4.8}$$

where $\mathbf{a}_n$ is the signal propagation steering vector at the position $\mathbf{p}_n$, defined in Eq.(3.16); $\tilde{\mathbf{a}}_n$ denotes the beamforming steering vector (spatial filter coefficients) for the source signal coming from the position $\mathbf{p}_n$; and $\sigma_e^2 = \mathbb{E}[\mathbf{e}^{\dagger}\mathbf{e}]$ is the power (variance) of the i.i.d AGWN noise $\mathbf{e}$, defined in Eq.(3.9); σ_e^2 is a constant compared with $\tilde{\mathbf{a}}_n$ in Eq.(4.8). Using the Lagrange coefficient method [11], the above optimization can be achieved by

$$\tilde{\mathbf{a}}_n = \frac{\mathbf{a}_n}{\|\mathbf{a}_n\|^2}\,, \tag{4.9}$$

where $\|\cdot\|$ denotes the vector norm; $\mathbf{a}_n$ denotes the propagation steering vector at the position $\mathbf{p}_n$ defined in Eq.(3.16).

From Eq.(4.8-4.9), one of the characteristics is that the beamforming steering vector (spatial filter coefficients) does not depend on the measured data, namely the CBF method is data-independent as point out in article[79]. That is why the CBF is a direct and simple method for source localization. Then we should consider the following two performances of the CBF methods:

- Beamforming power estimation .
- Beamforming spatial resolution.

For the first aspect, let y_n define the beamforming power at the position $\mathbf{p}_n$. According to the filter output power in Eq.(4.4), y_n can be calculated as:

$$y_n = \tilde{\mathbf{a}}_n^{\dagger}\mathbf{R}_z\tilde{\mathbf{a}}_n\,, \tag{4.10}$$

where $\mathbf{R}_z$ is the cross-spectrum matrix of measured signals $\mathbf{z}$, defined in Eq.(4.5). Owing to the assumptions that sources are uncorrelated and they are independent of noises, $\mathbf{R}_z$ can be further expressed as

$$\mathbf{R}_z = \mathbb{E}[\mathbf{z}\mathbf{z}^{\dagger}] = \mathbf{A}\,\mathbf{R}_s\,\mathbf{A}^{\dagger} + \sigma_e^2\,\mathbf{I}_M = \sum_{n=1}^{N} x_n\,\mathbf{a}_n\mathbf{a}_n^{\dagger} + \sigma_e^2\,\mathbf{I}_M\,, \tag{4.11}$$

where $\mathbf{R}_s$ is the covariance matrix of discrete source signals, defined in Eq.(3.20); $x_n \in \mathbf{x}$ is the source power at position $\mathbf{p}_n$, defined in Eq.(3.19); $\mathbf{I}_M$ denotes the $M \times M$ identity matrix.

Inserting $\mathbf{R}_z = \mathbb{E}[\mathbf{z}\mathbf{z}^\dagger]$ of Eq.(4.11) into Eq.(4.10), we can rewrite the beamforming in a vector form as:

$$\begin{aligned} \mathbf{y} &= \tilde{\mathbf{A}}^\dagger \, \mathbb{E}[\mathbf{z}\mathbf{z}^\dagger] \, \tilde{\mathbf{A}} \\ &= \mathbb{E}[|\tilde{\mathbf{A}}^\dagger \, \mathbf{z}|^2] \,, \end{aligned} \tag{4.12}$$

where $\tilde{\mathbf{A}} = [\tilde{\mathbf{a}}_1, \cdots, \tilde{\mathbf{a}}_n, \cdots, \tilde{\mathbf{a}}_N]$ denotes the beamforming steering matrix, with $\tilde{\mathbf{a}}_n$ being the beamforming steering vector obtained in Eq.(4.9). From Eq.(4.12), the beamforming outputs can be regarded as the back-projecting powers of measured signal powers on the source plane as shown in Fig.(4.2). So that the beamforming method can be a simple and direct estimation for source powers and positions.

Besides, inserting $\mathbf{R}_z = \sum_{n=1}^{N} x_n \, \mathbf{a}_n \mathbf{a}_n^\dagger + \sigma_e^2 \, \mathbf{I}_M$ of Eq.(4.11) into Eq.(4.10), we can rewrite the beamforming as:

$$\begin{aligned} y_n &= \tilde{\mathbf{a}}_n^\dagger \sum_{q=1}^{N} x_q \, \mathbf{a}_q \mathbf{a}_q^\dagger \, \tilde{\mathbf{a}}_n + \sigma_e^2 \, \tilde{\mathbf{a}}_n^\dagger \tilde{\mathbf{a}}_n \\ &= \sum_{q=1}^{N} |\tilde{\mathbf{a}}_n^\dagger \, \mathbf{a}_q|^2 \, x_q + \sigma_e^2 \, \frac{1}{||\mathbf{a}_n||^2} \,, \end{aligned} \tag{4.13}$$

where the beamforming steering vector $\tilde{\mathbf{a}}_n = \frac{\mathbf{a}_n}{||\mathbf{a}_n||^2}$ in Eq.(4.9), so that $\tilde{\mathbf{a}}_n^\dagger \tilde{\mathbf{a}}_n = \frac{1}{||\mathbf{a}_n||^2}$. From Eq.(4.12), the beamforming outputs can be regarded as the linear combinations (overlapping effect) of all discrete source powers. So that the beamforming estimated powers would be greater than real source powers (over-estimated), and beamforming method can be regarded as a coarse estimation of source power.

In particular, suppose that there is only one source signal which is of the i.i.d AWGN distribution with the power σ_s, independent of the noise. So that we get $N = 1$, $x_1 = \sigma_s$ and $\mathbf{R}_z = \sigma_s \, \mathbf{a}_1 \mathbf{a}_1^\dagger + \sigma_e^2 \, \mathbf{I}_M$. Then the beamforming power in Eq.(4.10) becomes:

$$y_1 = \sigma_s^2 + \frac{\sigma_e^2}{\|\mathbf{a}_1\|^2} \,, \tag{4.14}$$

where it indicates that in the ideal case, the CBF can maintain the original source power and attenuate the background noise by the factor of $\|\mathbf{a}_1\|^2$,

which is equivalent to improve the Signal-to-Noise Ratio (SNR). According to the definition of propagation steering vector in Eq.(3.16), if there is no ground reflection ($\rho = 0$), we can calculate $\|\mathbf{a}_1\|^2$ as:

$$\|\mathbf{a}_1\|^2 = \sum_{m=1}^{M} \frac{1}{r_{1,m}^2} \,. \tag{4.15}$$

where M is the total number of sensors, $M \geq 1$; and $r_{1,m}$ is the propagation distance from source to the mth sensor. Here we discuss the Eq.(4.15) in the far-field and near-field respectively as follows:

- For simplicity, we define the far-field propagation satisfying $\bar{r}_1 >> c_1\, M$, where $\bar{r}_1 = \frac{1}{M}\sum_{m=1}^{M} r_{1,m}$ denotes the arithmetic mean of all the propagation distances, and $c_1 \geq 1$ denotes a constant with unit. And the near-filed is defined as $\bar{r}_1 << c_2\, M$, where $c_2 \leq 1$ denote a constant with unit.

- In the far-field, many literature [22, 79, 126, 127, 128, 123] model $\mathbf{a}_1 = [\exp[-\jmath(2\pi f_l\, \tau_{1,1})], \cdots, \exp[-\jmath(2\pi f_l\, \tau_{1,M})]]^T$ and suppose $\|\mathbf{a}_1\|^2 = M$ in Eq.(4.15). So that Eq.(4.14) becomes $y_1 = \sigma_s^2 + \frac{\sigma_e^2}{M}$. It is seen that the background noise can be attenuated by the factor M by the CBF. Therefore, beamforming power y_1 can be an estimation of the source power x_1, especially when the number of sensor array is very large ($M >> 1$).

- In the near-field, Eq.(4.15) is approximated by $\|\mathbf{a}_1\|^2 = \frac{M}{\bar{r}_1}$. According to the mentioned near-field assumption $\bar{r}_1 << M$, it derives $\|\mathbf{a}_1\|^2 >> 1$. So that in Eq.(4.14), the background noise can be also attenuated by the factor $\|\mathbf{a}_1\|^2$ by the CBF method. Therefore, beamforming power y_1 can be also a coarse estimation of the source power x_1, especially when averaged distance is very small ($\bar{r}_1 << 1$).

Therefore, no matter of the near or far-field, the CBF can maintain the original source power and attenuate the background noise by the factor of $\|\mathbf{a}_n\|^2$.

Now we discuss the beamforming resolution. In the beamforming power estimation of Eq.(4.10), $\mathbf{R}_z$ is practically calculated by $\hat{\mathbf{R}}_z$ in Eq.(4.6), thus

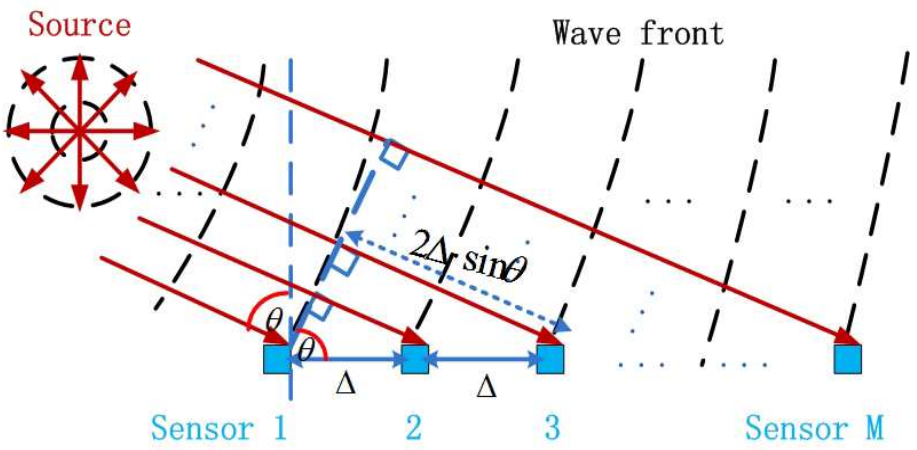

Figure 4.3: Illustration of linear sensor measurements in Eq.(4.10).

it yields:

$$y_n = \frac{1}{I}\sum_{i=1}^{I} \left|\tilde{\mathbf{a}}_n^{\dagger}\mathbf{z}_i\right|^2 \tag{4.16}$$
$$= \frac{1}{||\mathbf{a}_n||^2}\frac{1}{I}\sum_{i=1}^{I}\left|\sum_{m=1}^{M}\frac{1}{r_{n,m}}\exp\left[\jmath(2\pi f_l\,\tau_{n,m})\right] z_{i,m}\right|^2,$$

where beamforming steering vector $\tilde{\mathbf{a}}_n$ is obtained from Eq.(4.9), signal steering vector $\mathbf{a}_n$ is modeled from Eq.(3.16) with ground reflecting coefficient $\rho = 0$; and $||\mathbf{a}_n||^2$ can be similarly calculated from Eq.(4.15).

To calculate the Eq.(4.16), we take the uniform linear array of sensors for example as shown in Fig.(4.3): let Δ denote the spatial interval between sensors, satisfying $\Delta \leq \frac{\lambda}{2}$ for the spatial sampling condition; $\lambda = \frac{c_0}{f_l}$ denotes the acoustic wave length; θ denotes the Direction Of Arrival (DOA). So that the delay time $\tau_{n,m}$ can be expressed as

$$\tau_{n,m} = \tau_{n\,0} + \frac{1}{c_0}(m-1)\,\Delta\,sin\theta \tag{4.17}$$

where $\tau_{n\,0}$ denotes the time delay from the source n to the first sensor. Replacing the Eq.(4.17) into Eq.(4.16), and assuming $r_{n,m} \approx \bar{r}_n = \frac{1}{M}\sum_{m=1}^{m} r_{n,m}$, we can get an approximating expression of Eq.(4.16) as:

$$y_n \approx \frac{1}{||\mathbf{a}_n||^2}\frac{1}{\bar{r^2}_n}\frac{1}{I}\sum_{i=1}^{I}\left|\sum_{m=1}^{M}\exp\left[\jmath\frac{2\pi\Delta sin\theta}{\lambda}(m-1)\right] z_{i,m}\right|^2 \tag{4.18}$$

where for $I = 1$, equation (4.18) becomes the periodogram estimation [91] of the spectral density of measured signals $\mathbf{z}$; for $I > 1$, the approximated CBF

method in Eq.(4.18) can be regarded as the averaged (weighted) periodogram estimation, so that the CBF has the advantage of decreasing the standard deviation of periodogram estimation by the factor of $\frac{1}{\sqrt{I}}$. In Eq.(4.18), taking $z_{i,m} = 1$ for instance, we get

$$\begin{aligned} y_n &\approx \frac{1}{||\mathbf{a}_n||^2} \frac{1}{\bar{r^2}_n} \left| \sum_{m=1}^{M} \exp\left[\jmath \frac{2\pi \Delta sin\theta}{\lambda} (m-1) \right] \right|^2 \\ &= \frac{1}{||\mathbf{a}_n||^2} \frac{1}{\bar{r^2}_n} \left| \frac{sin[\pi \frac{(M-1)\Delta sin\theta}{\lambda}]}{sin[\pi \frac{\Delta sin\theta}{\lambda}]} \right|^2 \end{aligned} \tag{4.19}$$

where equation (4.19) indicates the similar form of periodic Sinc function. So that the angular resolution of the periodogram [91] can be given as:

$$\delta_\theta^{CBF} = \frac{\lambda}{D_a \, cos\theta} \; (rad) \, , \tag{4.20}$$

where $D_a = (M-1)\,\Delta$ denotes the total length of the linear array; δ_θ^{CBF} is with the unit of radian (rad); the angle (degree) is calculated as $\delta^{\circ CBF} = \delta_\theta \frac{180}{\pi}$. Based on Eq.(4.18), the spatial resolution of the CBF can be obtained as:

$$\delta_p^{CBF} = \delta_\theta \, \bar{r} = \frac{\lambda \, \bar{r}}{D_a \, cos\theta} = \frac{c_0 \, \bar{r}}{f \, D_a \, cos\theta} \, , \tag{4.21}$$

where $\bar{r}$ denotes the averaged distance from source plane to sensor array. Eq.(4.21) indicates that δ_p^{CBF} is proportional to the wave length and averaged distance, inversely proportional to the frequency and array length; and δ_p^{CBF} also depends on the DOA: when the incident wave is perpendicular to the linear array ($\theta = 0°$), $\delta_p^{CBF} = \frac{c_0 \, \bar{r}}{f \, D_a}$ reach its best spatial resolution. When the incident wave is parallel to the array ($\theta = 90°$), $\delta_p^{CBF} = \infty$ is too big to localize any sources. For instance, when $\theta = 60°$, $D_a = 3m$, $\bar{r} = 4.5m$, $c_0 = 340m/s$, we have $\delta^{\circ CBF} = 13°$ and $\delta_p^{CBF} = 102cm$ for $f = 1000Hz$; $\delta^{\circ CBF} = 4.3°$ and $\delta_p^{CBF} = 34cm$ for $f = 3000Hz$.

In brief, the CBF method is an easy, direct and robust estimation of source powers. But its spatial resolution is very limited for source localization due to the array topology and acoustic signal frequency.

4.2.2 Capon

In order to improve the spatial resolution of the CBF, the Capon filter [22], named the Minimum Variance Distortionless Response (MVDR), is often

used to maintain the signal at the source position and minimize the output power of received signals at the other positions as follows:

$$\begin{cases} P1: & \check{\mathbf{a}}_n^\dagger \, \mathbf{a}_n = 1 \\ P2: & \check{\mathbf{a}}_n = \arg\min_{\check{\mathbf{a}}_n} \left\{ \check{\mathbf{a}}_n^\dagger \, \mathbf{R}_z \, \check{\mathbf{a}}_n \right\} \end{cases}, \tag{4.22}$$

where $\check{\mathbf{a}}_n$ denotes the Capon steering vector for the source n at the position $\mathbf{p}_n$. Using the Lagrange coefficient method [11], the optimization is achieved by

$$\check{\mathbf{a}}_n = \frac{\mathbf{R}_z^{-1} \, \mathbf{a}_n}{\mathbf{a}_n^\dagger \, \mathbf{R}_z^{-1} \, \mathbf{a}_n}, \tag{4.23}$$

where $\mathbf{R}_z$ is supposed to be invertible and $\mathbf{R}_z^{-1}$ denotes the inverse of $\mathbf{R}_z$. Let $\check{y}_n$ define the Capon power at the position $\mathbf{p}_n$. According to the filter output power in Eq.(4.4) and Capon steering vector $\check{\mathbf{a}}_n$ in Eq.(4.23), the Capon filter output power $\check{y}_n$ can be calculated as:

$$\check{y}_n = \frac{1}{\mathbf{a}_n^\dagger \, \mathbf{R}_z^{-1} \, \mathbf{a}_n}, \tag{4.24}$$

where $\mathbf{R}_z^{-1}$ is practically estimated by $\hat{\mathbf{R}}_z^{-1}$ in Eq.(4.6).

Comparing with the CBF steering vector in Eq.(4.9), the Capon criterion in Eq.(4.23) depends on the measured data, namely it is a data-dependent method. This explains why the Capon method can get better spatial resolution than the CBF. Moreover, the over-estimation in the CBF due to its overlapping effect shown in Eq.(4.13) can be avoided to some extend in the Capon thank to MVDR criterion.

In particular, suppose that there is only one source signal, and it is of the i.i.d AWGN distribution with the power of σ_s, independent of the noise. So that we get $N = 1$, $x_1 = \sigma_s$ and $\mathbf{R}_z = \sigma_e^2 \, \mathbf{I}_M + \mathbf{a}_1 \, \sigma_s \, \mathbf{I}_M \, \mathbf{a}_1^\dagger$. Then the Capon power in Eq.(4.24) becomes:

$$\begin{aligned} \check{y}_1 &= \frac{1}{\mathbf{a}_1^\dagger \, \mathbf{R}_z^{-1} \, \mathbf{a}_1} \\ &= \frac{1}{\mathbf{a}_1^\dagger (\sigma_e^2 \, \mathbf{I}_M + \mathbf{a}_1 \, \sigma_s \, \mathbf{I}_M \, \mathbf{a}_1^\dagger)^{-1} \mathbf{a}_1}. \end{aligned} \tag{4.25}$$

According to the matrix inversion lemma, if matrix $\mathbf{A} \in \mathbb{C}^{M \times M}$, $\boldsymbol{C} \in \mathbb{C}^{M \times M}$ are both non-singular matrice (invertible), for any $\mathbf{B} \in \mathbb{C}^{M \times N}$, $\boldsymbol{C} \in \mathbb{C}^{M \times N}$, the inversion of matrix $(\mathbf{A} + \mathbf{B} \, \boldsymbol{C} \, \mathbf{D})$ is as:

$$(\mathbf{A} + \mathbf{B} \, \boldsymbol{C} \, \mathbf{D})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \, \mathbf{B} \, (\mathbf{D} \, \mathbf{A}^{-1} \, \mathbf{B} + \boldsymbol{C}^{-1})^{-1} \, \mathbf{D} \, \mathbf{A}^{-1}. \tag{4.26}$$

where $\mathbf{A} = \sigma_e^2 \mathbf{I}_M$, $\mathbf{B} = \mathbf{a}_1$, $\boldsymbol{C} = \sigma_s \mathbf{I}_M$ and $\mathbf{D} = \mathbf{a}_1^\dagger$. So that Eq.(4.25) becomes

$$\begin{aligned} \check{y}_1 &= \frac{1}{\mathbf{a}_1^\dagger (\sigma_e^{-2} \mathbf{I}_M - \sigma_e^{-4} \frac{\mathbf{a}_1 \mathbf{a}_1^\dagger}{\sigma_s^{-2} + \sigma_e^{-2} ||\mathbf{a}_1||^2}) \mathbf{a}_1} \\ &= \sigma_s^2 + \frac{\sigma_e^2}{||\mathbf{a}_1||^2}, \end{aligned} \tag{4.27}$$

where Eq.(4.27) shows that the Capon can get the similar denoising performance as the CBF does in Eq.(4.14): attenuate the background noise and maintain the source power. But for more than one source localization, the Capon can well adapt to the signals of interest and get high spatial resolution of DOA owing to its MVDR criterion in Eq.(4.22), while the CBF could not do so due to its data-independent criterion in Eq.(4.9).

In brief, the Capon filter is also a simple and direct method for source localization. It can get better spatial resolution than the CBF thanks to the MVDR criterion on the received signals. But the Capon is less robust method than the CBF, since it requires quite a lot of samplings to obtain $\hat{\mathbf{R}}_z^{-1}$ in Eq.(4.24). Moreover, since the Capon has strong relationship with the CBF, its spatial resolution could not be greatly improved.

4.2.3 MUSIC

The Capon method can improve the spatial resolution of the CBF to some extent, but in order to achieve very high resolution, one of the most commonly used methods is called the MUltiple SIgnal CLassification (MUSIC) [98, 12]. The MUSIC is derived from the subspace decomposition of the covariance matrix of the measured signals at the sensors.

We introduce the subspace decomposition in the following. Let's recall the measured signals $\mathbf{z}$ for original source signals $\mathbf{s}^*$ in Eq.(3.9) as shown in Fig.4.4:

$$\mathbf{z} = \mathbf{A}^* \mathbf{s}^* + \mathbf{e}, \tag{4.28}$$

where $\mathbf{A}^* \in \mathbb{C}^{M \times K}$ is short for $\mathbf{A}(\mathbf{P}^*)$ in Eq.(3.9). The other variable dimensions are: $\mathbf{z}, \mathbf{e} \in \mathbb{C}^M$ with M being the sensor number; $\mathbf{s}^* \in \mathbb{C}^K$ with K being the uncorrelated original source signals, providing $M > K$.

In order to make subspace separation of measured signals, we have to suppose that in Eq.(3.9), $\mathbf{A}^* \in \mathbf{C}^{M \times K}$ has the matrix rank as $rank(\mathbf{A}^*) = K$.

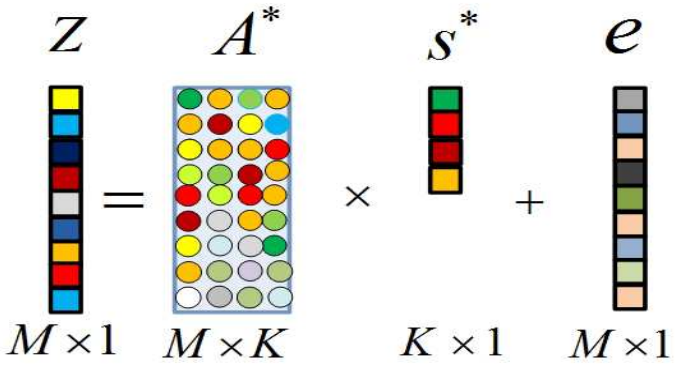

Figure 4.4: Illustration of the original signal equation in Eq.(4.28)

Since $\mathbf{A}^* = [\mathbf{a}_1^*, \cdots, \mathbf{a}_K^*]$ with $\mathbf{a}_k^*$, $k \in [1, \cdots, K]$ being short for $\mathbf{a}(\mathbf{p}_k^*)$, in other words, it is equivalent to suppose that K columns of $\mathbf{A}$ are uncorrelated.

Based on Eq.(4.28), the covariance matrix $\mathbf{R}_z$ of measured signals is obtained as

$$\mathbf{R}_z = \mathbb{E}[\mathbf{z}\mathbf{z}^\dagger] = \mathbf{A}^* \mathbf{R}_{s^*} \mathbf{A}^{*\dagger} + \sigma_e^2 \mathbf{I}_M \,, \tag{4.29}$$

where $\mathbf{R}_{s^*}$ denotes the covariance matrix of original signals $\mathbf{s}^*$, defined as

$$\mathbf{R}_{s^*} = \mathbb{E}[\mathbf{s}^* \mathbf{s}^{*\dagger}] \,, \tag{4.30}$$

where $\mathbf{R}_{s^*} \in \mathbb{C}^{K \times K}$ is a Hermitian matrix ($\mathbf{R}_{s^*}^\dagger = \mathbf{R}_{s^*}$). Since all the K source signals $\mathbf{s}^*$ are uncorrelated to each other, the rank of $\mathbf{R}_{s^*}$ is K, noted as $rank(\mathbf{R}_{s^*}) = K$. In Eq.(4.29), $\mathbf{R}_z \in \mathbb{C}^{M \times M}$ is also a Hermitian matrix ($\mathbf{R}_z^\dagger = \mathbf{R}_z$), whose eigenvectors are orthonormal to each other and eigenvalues are of real values.

On the left of Eq.(4.29), we take the EigenValue Decomposition (EVD) on $\mathbf{R}_z$ as:

$$\mathbf{R}_z = \mathbf{U}_z \, \Lambda_z \, \mathbf{U}_z^\dagger \tag{4.31}$$

where $\mathbf{U}_z = [\mathbf{u}_1, \cdots, \mathbf{u}_M]^T$ is a unitary matrix ($\mathbf{U}\,\mathbf{U}^\dagger = \mathbf{U}^\dagger\,\mathbf{U} = \mathbf{I}_M$) that is composed of the eigenvectors $\mathbf{u}_m$, $m \in [1, \cdots, M]$ of $\mathbf{R}_z$; and $\Lambda_z = Diag[\lambda_1, \cdots, \lambda_M]$ is a diagonal matrix that is composed of the eigenvalues λ_m of $\mathbf{R}_z$, satisfying $\lambda_1 \geq \cdots \geq \lambda_m \geq \cdots \geq \lambda_M$.

On the right of Eq.(4.29), $\mathbf{A}^* \mathbf{R}_{s^*} \mathbf{A}^{*\dagger}$ is also a Hermitian matrix. Since $rank(\mathbf{A}^*) = K$ and $rank(\mathbf{R}_{s^*}) = K$, we have $rank(\mathbf{A}^* \mathbf{R}_{s^*} \mathbf{A}^{*\dagger}) = K$. Then taking the EVD on $\mathbf{A}^* \mathbf{R}_{s^*} \mathbf{A}^{*\dagger}$, we define its eigenvalue matrix as:

$$\Lambda_{s^*} = Diag[\tilde{\lambda}_1, \cdots, \tilde{\lambda}_K] \,, \tag{4.32}$$

where Λ_{s^*} is diagonal matrix, whose diagonal items are composed of K number of the non-zero eigenvalues. So that Λ_{s^*} is invertible.

According to Eq.(4.29), Λ_{s^*} and Λ_z have the relationship as:

$$\begin{cases} \lambda_m = \tilde{\lambda}_m + \sigma_e^2 & \text{for} \quad m = 1, 2, \cdots, K \\ \lambda_m = \sigma_e^2 & \text{for} \quad m = K+1, K+2, \cdots, M \end{cases}, \tag{4.33}$$

where it can be rewritten in a matrix form as

$$\Lambda = \begin{bmatrix} \Lambda_{s^*} + \sigma_e^2 \mathbf{I}_K & \mathbf{0} \\ \mathbf{0} & \sigma_e^2 \mathbf{I}_{M-K} \end{bmatrix}., \tag{4.34}$$

where $\mathbf{I}_{M-K}$ is a $M-K$ dimensions identity matrix.

Therefore, the signal subspace of $\mathbf{R}_z$ is generated by the K eigenvectors $[\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_K]$ which correspond to K eigenvalues $[\lambda_1, \lambda_2, \cdots, \lambda_K]$. And noise subspace of $\mathbf{R}_z$ is generated by the other $M-K+1$ eigenvectors. Therefore, we can get the subspace decomposition from the Eq.(4.29, 4.31, 4.34) as:

$$\begin{cases} \mathbf{E}_{s^*} &= [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_K] \\ \mathbf{E}_e &= [\mathbf{u}_{K+1}, \mathbf{u}_{K+2}, \cdots, \mathbf{u}_M] \end{cases}, \tag{4.35}$$

where $\mathbf{E}_s$ denotes the basis of signal subspace, $rank(\mathbf{E}_{s^*}) = K$; its corresponding eigenvalues are $(\Lambda_{s^*} + \sigma_e^2 \mathbf{I}_K)$ in Eq.(4.34); and $\mathbf{E}_e$ denotes the basis of noise subspace, which is orthonormal to $\mathbf{E}_{s^*}$, namely

$$\mathbf{E}_{s^*}^{\dagger} \mathbf{E}_e = \mathbf{0}, \tag{4.36}$$

where the above equation reflects an importance property of orthogonality between the source subspace and noise subspace.

Then taking $\mathbf{E}_{s^*}$ and corresponding eigenvalues $(\Lambda_{s^*} + \sigma_e^2 \mathbf{I}_K)$ into Eq.(4.31), we get

$$\mathbf{R}_z \mathbf{E}_{s^*} = \mathbf{E}_{s^*} (\Lambda_{s^*} + \sigma_e^2 \mathbf{I}_K). \tag{4.37}$$

Replacing $\mathbf{R}_z$ by Eq.(4.29) into Eq.(4.37), we get:

$$\mathbf{A}^* \mathbf{R}_{s^*} \mathbf{A}^{*\dagger} \mathbf{E}_s + \sigma_e^2 \mathbf{E}_s = \mathbf{E}_{s^*} (\Lambda_{s^*} + \sigma_e^2 \mathbf{I}_K), \tag{4.38}$$

where it is the linear system of equations of $\mathbf{E}_{s^*}$, which can be solved as:

$$\mathbf{E}_{s^*} = \mathbf{A}^* \mathbf{T}, \tag{4.39}$$

where $\mathbf{T} \in \mathbf{C}^{K \times K}$ is a full rank matrix with $rank(\mathbf{T}) = K$, since it can be expressed as:

$$\mathbf{T} = \mathbf{R}_{s^*} \mathbf{A}^{*\dagger} \mathbf{E}_{s^*} \Lambda_{s^*}^{-1}, \tag{4.40}$$

where

- Source signal covariance matrix $\mathbf{R}_{s^*} \in \mathbf{C}^{K\times K}$ has $rank(\mathbf{R}_{s^*}) = K$, since K original source signals are supposed to be independent;
- Signal propagation matrix $\mathbf{A}^* \in \mathbf{C}^{M\times K}$ with $M > K$ is supposed to be $rank(\mathbf{A}^*) = K$;
- Signal subspace $\mathbf{E}_{s^*}$ has $rank(\mathbf{E}_{s^*}) = K$;
- $\Lambda_{s^*}^{-1} \in \mathbf{R}^{K\times K}$ has $rank(\Lambda_{s^*}^{-1}) = K$, since Λ_{s^*} is convertible according to its definition in Eq.(4.32).

Therefore, in Eq.(4.39), the source signal subspace $\mathbf{E}_{s^*}$ can be represented by the invertible linear combination ($\mathbf{T}$) of the columns of $\mathbf{A}^*$. So that $\mathbf{E}_{s^*}$ is equivalent to the space $esp\{\mathbf{A}^*\}$ generated by the columns of $\mathbf{A}^*$, namely $\mathbf{E}_{s^*} = esp\{\mathbf{A}^*\}$. Since $\mathbf{E}_{s^*}$ and $\mathbf{E}_e$ are orthonormal to each other, we can obtain an importance equation as follows:

$$\mathbf{A}^{*\dagger}\,\mathbf{E}_e = \mathbf{0}\,, \tag{4.41}$$

where it indicates that columns of $\mathbf{A}^*$ can generate the null space of $\mathbf{E}_e$, defined as

$$Null(\mathbf{E}_e) = \left\{\mathbf{a}_k^* \in \mathbf{A} \;:\; \mathbf{a}_k^{*\dagger}\,\mathbf{E}_e = \mathbf{0},\; k \in [1, \cdots, K]\right\} \tag{4.42}$$

where $\mathbf{a}_k^*$ is short for $\mathbf{a}(\mathbf{p}_k^*)$ defined in Eq.(3.10). In Eq.(4.42), it shows that the steering vector $\mathbf{a}_k^*$ at the source position $\mathbf{P}_k^*$ is orthonormal to the noise space $\mathbf{E}_e$.

Based on the orthonormal property of the subspace decomposition in Eq.(4.41), MUSIC criterion is to maximize the pseudo-power y_k^{music} at the position $\mathbf{P}_k^*$ as:

$$y_k^{music} = \frac{1}{\|\mathbf{a}_k^{*\dagger}\mathbf{E}_e\|^2}\,, \tag{4.43}$$

where original steering vector $\mathbf{a}_k^*$ at the position $\mathbf{p}_k^*$ is approximated by discrete steering vector $\mathbf{a}_n$ at the position $\mathbf{p}_n$ defined in Eq.(3.16), in which, we discretize the source plane into N grids at the positions $\mathbf{P} = [\mathbf{p}_1, \cdots, \mathbf{p}_N]$, supposing $\mathbf{p}_k^* \in \mathbf{P}$. Moreover, the noise subspace is calculated by $\hat{\mathbf{E}}_e$ which is obtained from the estimated covariance matrix $\hat{\mathbf{R}}_z$ defined in Eq.(4.6).

In brief, MUSIC can achieve infinite spatial resolution if the subspace decomposition could be exactly made (noiseless cases). To achieve correct subspace decomposition, it is important to know the noise level or the total number of sources. However, MUSIC requires as many as possible samplings

to well estimate the covariance matrix of received signals, such as the Capon. It also needs to know the number of original sources or the SNR level to effectively decompose the signal and the noise subspace. Therefore, MUSIC is less robust than the CBF. Besides, MUSIC can achieve high resolution in source localization, but could not directly offer the source power estimation.

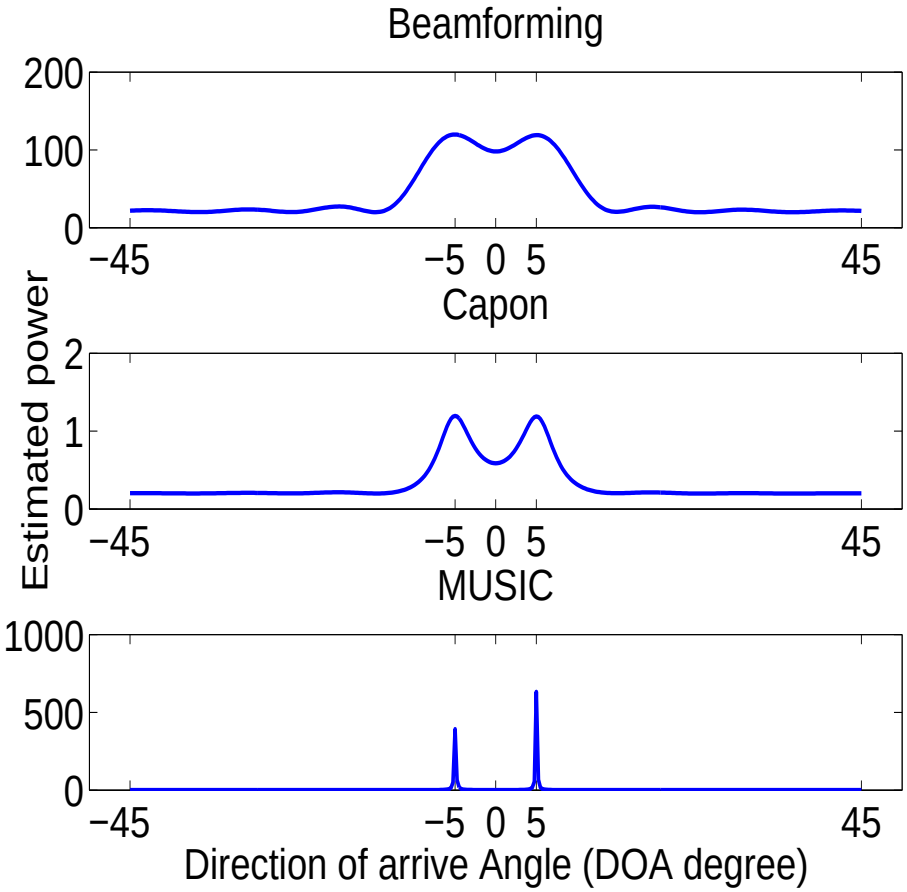

Figure 4.5: Performance comparison of the CBF, Capon and MUSIC.

In Fig.4.5, we show the performance comparison of the CBF, Capon and MUSIC. There are $K = 2$ source signals from the DOAs of $-5°$ and $5°$, with a single frequency of $f = 200$Hz and a power of $x_1^* = x_2^* = 1$. The background noise is supposed to be i.i.d AGWN. The SNR is as low as 0dB. There are $T = 10000$ time samplings. The source plane is discretized into $N = 361$ grids between $[-45°, 45°]$. It is seen that the CBF could hardly localize the two sources, since its angular resolution is as low as $\delta°_{200Hz}^{CBF} = 11°$ according to Eq.(4.20). The beamforming estimated powers are both over-estimated due to its overlapping effect as discussed in Eq.(4.13). The Capon can well detect the sources owing to its MVDR criterion in Eq.(4.22). The Capon estimated powers are much closer to source powers than the CBF as discussed in Eq.(4.24). The MUSIC can achieve the best resolution among

the three thanks to the signal subspace decomposition. But MUSIC could not reconstruct the source powers due to its pseudo-power estimation as shown in Eq.(4.43), nor estimate the relative amplitudes due to background noise interference.

4.3 Discrete forward model of acoustic power propagation

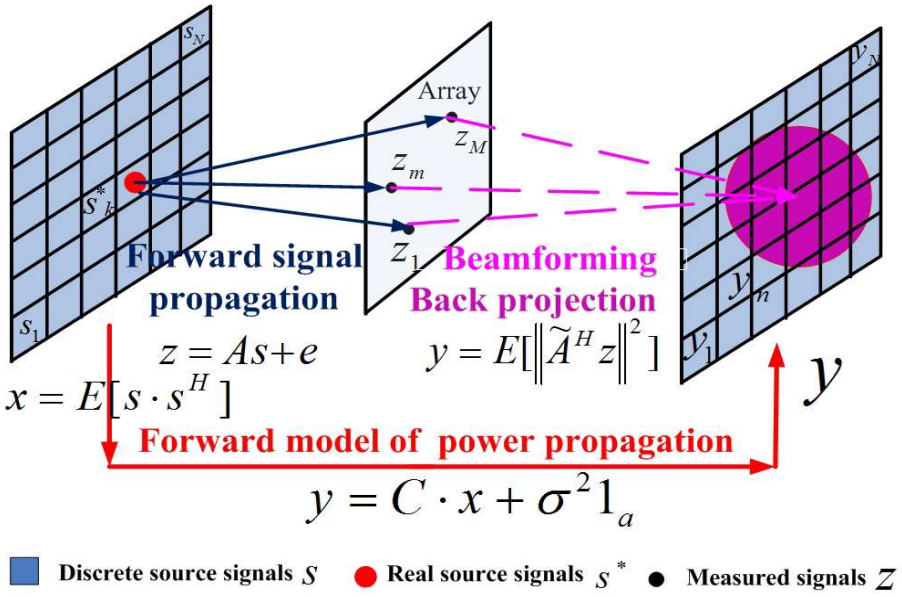

Figure 4.6: Illustration of forward model of acoustic power propagation.

The beamforming, Capon and MUSIC methods are the direct estimation of source DOA or source powers (pseudo-powers), and they are all based on the forward model of signal propagation in Eq.(3.15). In order to overcome their limitations in source localization, we will introduce some typical iterative methods (deconvolution and regularization). Before doing that, we should present the forward model of power propagation as shown in Fig.(4.6), since the acoustic imaging involves in the source power reconstruction, rather than signal estimation. Owing to the fact that the beamforming powers in Eq.(4.10) can provide a direct but coarse power estimation, it is necessary to deconvolve the blurry beamforming result for high spatial resolution,and deconvolution need to account for sampling errors.

Let's recall the beamforming power $y_n = \tilde{\mathbf{a}}_n^\dagger \mathbf{R}_z \tilde{\mathbf{a}}_n$ in Eq.(4.10), where measured covariance matrix $\mathbf{R}_z = \mathbb{E}[\mathbf{z}\,\mathbf{z}^\dagger]$ in Eq.(4.5) is defined by measured signal model $\mathbf{z}_i = \mathbf{A}(\mathbf{P})\,\mathbf{s}_i + \mathbf{e}_i$ in Eq.(3.15). As long as the sampling number

is big enough (sampling block number $I >> 1$), the measured covariance matrix $\mathbf{R}_z$ is practically approximated by $\hat{\mathbf{R}}_z = \frac{1}{I}\sum_{i=1}^{I}\mathbf{z}_i\,\mathbf{z}_i^{\dagger}$ in Eq.(4.6), where $\mathbf{z}_i$ is the measured signal vector in the ith sampling block at all M sensors. So that on the left of $y_n = \tilde{\mathbf{a}}_n^{\dagger}\mathbf{R}_z\tilde{\mathbf{a}}_n$ in Eq.(4.10), we actually calculate the beamforming power as

$$y_n = \tilde{\mathbf{a}}_n^{\dagger}\hat{\mathbf{R}}_z\tilde{\mathbf{a}}_n\,, \tag{4.44}$$

where y_n is the measured beamforming power at the position of $\mathbf{p}_n$ on the source plane; $\tilde{\mathbf{a}}_n$ is the beamforming steering vector for the position $\mathbf{p}_n$, defined in Eq.(4.9). On the right of Eq.(4.44), we replace $\hat{\mathbf{R}}_z$ by using $\mathbf{R}_z = \sum_{q=1}^{N} x_q\,\mathbf{a}_q\mathbf{a}_q^{\dagger} + \sigma_e^2\,\mathbf{I}_M$ of Eq.(4.11), then we get

$$\begin{aligned} y_n &= \tilde{\mathbf{a}}_n^{\dagger}\sum_{q=1}^{N} x_q\,\mathbf{a}_q\mathbf{a}_q^{\dagger}\,\tilde{\mathbf{a}}_n + \sigma_e^2\,\tilde{\mathbf{a}}_n^{\dagger}\tilde{\mathbf{a}}_n \\ &= \sum_{q=1}^{N}|\tilde{\mathbf{a}}_n^{\dagger}\,\mathbf{a}_q|^2\,x_q + \sigma_e^2\,\frac{1}{||\mathbf{a}_n||^2} \\ &= \sum_{q=1}^{N} c_{nq}\,x_q + \sigma_e^2\,\frac{1}{||\mathbf{a}_n||^2}\,, \end{aligned} \tag{4.45}$$

where the beamforming steering vector $\tilde{\mathbf{a}}_n = \frac{\mathbf{a}_n}{||\mathbf{a}_n||^2}$ in Eq.(4.9), so that $\tilde{\mathbf{a}}_n^{\dagger}\tilde{\mathbf{a}}_n = \frac{1}{||\mathbf{a}_n||^2}$.

For N positions of the beamforming outputs, equation (4.45) can be rewritten in a vector form as shown in Fig.4.7, and we obtain the forward model of source power propagation as follows:

$$\mathbf{y} = \mathbf{C}\,\mathbf{x} + \sigma_e^2\,\mathbf{1}_a\,, \tag{4.46}$$

where

- $\mathbf{y} = [y_1, \cdots, y_N]^T$ denotes the calculated beamforming powers on the source plane, which can be a primary estimation of the source powers;
- $\mathbf{x}$ is the unknown discrete source power vector defined in Eq.(3.19);
- σ_e^2 is the i.i.d AGWN noise power (unknown);
- $\mathbf{1}_a = [\frac{1}{\|\mathbf{a}_1\|^2}, \cdots, \frac{1}{\|\mathbf{a}_N\|^2}]^T$ serves to attenuate noise power for different positions;

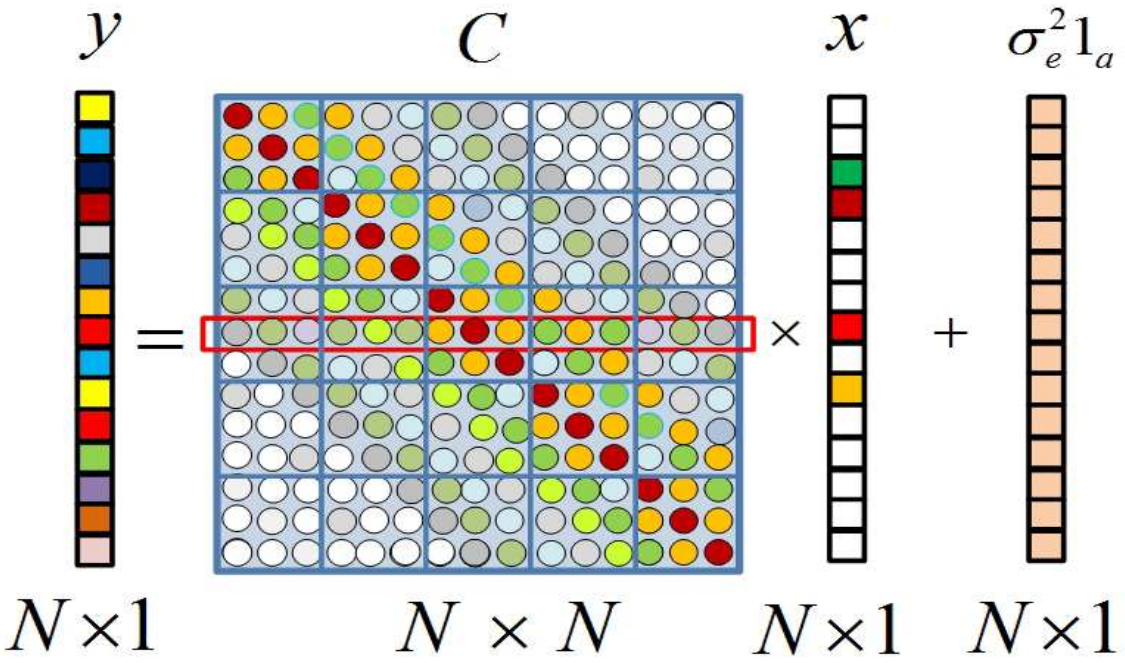

Figure 4.7: Illustration of the forward model power propagation in Eq.(4.46).

- $\mathbf{C} = [c_{n,q}]$ with $n, q \in [1, \cdots, N]$ denotes power propagation matrix, whose item $c_{n,q}$ is defined as:

$$c_{n,q} = |\tilde{\mathbf{a}}_n^{\dagger} \mathbf{a}_q|^2 \,, \tag{4.47}$$

where beamforming steering vector $\tilde{\mathbf{a}}_n$ and signal propagation steering vector $\mathbf{a}_n$ are defined in Eq.(4.9) and Eq.(3.16) respectively; $|\cdot|$ denotes the modulus; and $c_{n,q}$ represents the power contribution rate of the qth source to the nth position on the source plane, in which, we will give its geometrical interpretation on the bottom of Fig.7.8 in Chapter 7. According to Eq.(4.47), we have $c_{n,n} = 1$, namely, the diagonal value of power propagation matrix $\mathbf{C}$ is always 1. Moreover, $c_{n,q}$ in Eq.(4.47) can be seen as the Point Spread Function (PSF) of the actual sensor array. And this PSF is determined by two factors: the sensor array topology and the distance from the source plane. If the microphone array is ideal enough, $c_{n,q}$ becomes the Dirac function as

$$c_{n,q} = \delta_{n,q} = \begin{cases} 1 & \text{for} \quad n = q \\ 0 & \text{for} \quad n \neq q \end{cases} \,. \tag{4.48}$$

Then Eq.(4.46) becomes $\mathbf{y} = \mathbf{x} + \sigma_e^2 \, \mathbf{1}_a$, which reveals that the source powers $\mathbf{x}$ can be directly estimated from the beamforming powers $\mathbf{y}$. In particular when $N = 1$ and $K = 1$, we also get $y_1 = x_1 + \frac{\sigma_e^2}{\|\mathbf{a}_1\|^2}$ as discussed in Eq.(4.14) of the beamforming method.

Based on Eq.(4.47), the power propagation matrix $\mathbf{C}$ can be expressed as:

$$\mathbf{C} = |\tilde{\mathbf{A}}^{\dagger}\,\mathbf{A}|.^{2}\,, \tag{4.49}$$

where $\tilde{\mathbf{A}} = [\tilde{\mathbf{a}}_1 \cdots \tilde{\mathbf{a}}_N]$, $\tilde{\mathbf{A}} \in \mathbb{C}^{M\times N}$ denotes the beamforming steering matrix, defined in Eq.(4.12); $\mathbf{A} = [\mathbf{a}_1, \cdots, \mathbf{a}_N]$, $\mathbf{A} \in \mathbb{C}^{M\times N}$ denotes the signal propagation steering matrix, defined in Eq.(3.15); operator $|\cdot|.$ denotes the absolute modulus of each item of a matrix; $().^2$ denotes the square of each item of a matrix; $\mathbf{C} \in \mathbb{R}^{N\times N}$ is a square matrix full of non-negative values. Moreover, $\mathbf{A} \in \mathbb{C}^{M\times N}$ and $\tilde{\mathbf{A}} \in \mathbb{C}^{M\times N}$ with $M < N$, then their matrix ranks meet $rank(\tilde{\mathbf{A}}) \leq M < N$ and $rank(\mathbf{A}) \leq M < N$. In Appendix E, we give a detailed proof that $rank(\mathbf{C}) \leq M < N$, and $\mathbf{C} \in \mathbb{R}^{M\times N}$ is a singular matrix and cannot be invertible.

In brief, the underdetermined non-linear signal propagation model in Eq.(3.15) is transformed into the forward model of power propagation in Eq.(4.46) which is the determined linear system of equations for source powers $\mathbf{x}$. In Eq.(4.46), the beamforming power $\mathbf{y}$ and power propagation matrix $\mathbf{C}$ are calculated from Eq.(4.44) and Eq.(4.47) respectively. Unfortunately, $\mathbf{C}$ is very ill-conditioned and cannot be invertible, so that we can not directly obtain $\mathbf{x}$ from Eq.(4.46).

4.4 Deconvolution

As we have discussed that the beamforming powers in Eq.(4.10) can provide a direct and coarse power estimation, it is reasonable to deconvolve the blurry beamforming result to obtain high spatial resolution. Many deconvolution methods have been successfully developed since decades [7, 44, 13, 115, 4, 108, 8, 103, 123, 18, 19]. We introduce two of the most widely used for acoustic source localization and imaging: the DAMAS and CLEAN methods respectively.

4.4.1 DAMAS

Eq.(4.46) can be interpreted as a kind of deconvolution problem. One of the recently developed deconvolution methods is called: the Deconvolution Approach for Mapping of Acoustic Source (DAMAS) [18, 19]. The DAMAS

assumes $\sigma_e^2 = 0$ in Eq.(4.46) and try to solve

$$\mathbf{y} = \mathbf{C}\,\mathbf{x}\,. \tag{4.50}$$

The original DAMAS tries to solve Eq.(4.50) in a recursive manner as follows:

$$\begin{cases} x_1^{(i+1)} = y_1 - \sum_{q=2}^{N} c_{1,q}\, x_q^{(i)} \\ x_n^{(i+1)} = y_n - \sum_{q=1}^{n-1} c_{n,q}\, x_q^{(i+1)} - \sum_{q=n+1}^{N} c_{n,q}\, x_q^{(i)} \\ x_N^{(i+1)} = y_N - \sum_{q=1}^{N-1} c_{N,q}\, x_q^{(i+1)} \end{cases} . \tag{4.51}$$

where (i) denotes the ith iteration; and $\mathbf{x}^0 = [x_1^0, \cdots, x_N^0]^T$ can be initialized by 0 vector or the measured beamforming powers $\mathbf{y}$.

In brief, the DAMAS of Eq.(4.51) provides an iterative solution for the power propagation model of Eq.(4.46). It thus refines the CBF results and improve the spatial resolution. However, the most important drawback of this recursive algorithm is that the results may be very sensitive to the errors due to the noise free assumption, since the power propagation matrix $\mathbf{C}$ is very ill-conditioned. Moreover, DAMAS also suffers from the slow convergence problem. Many efforts have been done to improve the DAMAS performance. The most commonly used is called the Diagonal Removal (DR) DAMAS [18]. The DR-DAMAS sets diag $\left[\hat{\mathbf{R}}_z\right] = 0$ in Eq.(4.6) to suppress the noises, but DR technique inevitably harms weak sources whose powers are lower than the noise.

4.4.2 CLEAN

CLEAN [102, 123, 127] is another well developed deconvolution method. It can also iteratively extract strong sources from the blurred beamforming image, but depends on parameter selection for good performance.

Soppose there are K known original sources at the unknown positions $\mathbf{P}^* = [\mathbf{p}_1^*, \cdots, \mathbf{p}_K^*]$ on the source plane. After source plane discretization, it yields the grid (discrete source) positions $\mathbf{P} = [\mathbf{p}_1, \cdots, \mathbf{p}_N]^T$. Then the first source position $\mathbf{p}_1^* = \mathbf{p}_{n_1}$, $n_1 \in [1, \cdots, N]$ is initialized for the estimated position of the strongest source power x_{n_1}. Since x_{n_1} can be estimated by the strongest beamforming power $y_{n_1} \in \mathbf{y}$ as discussed in the beamforming method, the initialization can be expressed as:

$$\begin{cases} \mathbf{p}_1^* & = \mathbf{p}_{n_1}, \quad n_1 \in [1, \cdots, N] \\ \mathbf{p}_{n_1} & = \arg\max_{\mathbf{p}_{n_1} \in \mathbf{P}_n} \left\{ \mathbf{y} = [y_1, \cdots, y_N]^T \right\}, \; with \; y_n = \tilde{\mathbf{a}}_n^{\dagger} \hat{\mathbf{R}}_z \tilde{\mathbf{a}}_n \,, \end{cases} \tag{4.52}$$

where y_n is the beamforming power calculated in Eq.(4.44); $\mathbf{P}$ is grid positions on the source plane. According to the spatial filter definition in Eq.(4.2) and beamforming steering vector in Eq.(4.9), the estimated source signal $\hat{s}_1$ can be obtained from the measured data $\mathbf{z} = [z_1, \cdots, z_N]$ by

$$\hat{s}_1 = \tilde{\mathbf{a}}_1^{\dagger}\, \mathbf{z}\,, \tag{4.53}$$

where $\tilde{\mathbf{a}}_1$ refer to the beamforming steering vector at the source position $\mathbf{p}_1^*$ in Eq.(4.9). According to the discrete forward model of signal propagation in Eq.(3.15), the measured signal vector $\hat{\mathbf{z}}_1$ at the sensor for $\hat{s}_1$ is calculated as:

$$\hat{\mathbf{z}}_1 = \mathbf{A}\, \hat{s}_1 + \mathbf{e}\,, \tag{4.54}$$

where $\mathbf{A} = [\mathbf{a}_1, \cdots, \mathbf{a}_N]$ is the propagation steering vector as defined in Eq.(3.16). Then we can substract part of this strongest source power from the beamforming power $\mathbf{y}$ and get the residual powers $\mathbf{q}_1 = [q_{11}, \cdots, q_{1N}]^T$ (consisted of other sources and noises) as:

$$\mathbf{q}_1 = \mathbf{y} - \alpha\ \tilde{\mathbf{A}}^{\dagger}\hat{\mathbf{R}}_1\tilde{\mathbf{A}}\,, \tag{4.55}$$

where

- $\mathbf{y} = [y_1, \cdots, y_N]$ is the beamforming power vector obtained from Eq.(4.44);
- α denotes the attenuation factor. As pointed out in article[123], α is a key issue in CLEAN; $0 < \alpha < 1$ is used to avoid over-estimation and efficient convergence, $\alpha = 0.15$ is used for the proper compromise between convergence and estimation performance;
- $\tilde{\mathbf{A}} = [\tilde{\mathbf{a}}_1 \cdots \tilde{\mathbf{a}}_N]$ is the beamforming steering matrix as discussed in Eq.(4.49);
- $\hat{\mathbf{R}}_1 = \frac{1}{I}\sum_{i=1}^{I} \hat{\mathbf{z}}_{1\,i}^{\dagger}\, \hat{\mathbf{z}}_{1\,i}$ is the estimated covariance matrix of measured signals $\hat{\mathbf{z}}_1$ in Eq.(4.54).

Then the estimated position of the second strongest source power is given by:

$$\begin{cases} \mathbf{q}_1 & = \mathbf{y} - \alpha\ \tilde{\mathbf{A}}^{\dagger}\hat{\mathbf{R}}_1\tilde{\mathbf{A}} \\ \hat{\mathbf{p}}_{n_2} & = \arg\max_{\mathbf{p}_{n_2}\in\mathbf{P}} \{\mathbf{q}_1\} \end{cases} \tag{4.56}$$

where if $\hat{\mathbf{p}}_{n_2} = \mathbf{p}^*{}_1$, repeat Eq.(4.53)-Eq.(4.56); this means that we still locate the same strongest source, thus we have to furthermore subtract parts of this strongest source power from the beamforming powers, then try to find out

the second strong source in the residual powers. If $\hat{\mathbf{p}}_{n_2} \neq \mathbf{p}^*{}_1$, that means we have found out the second strongest source, so that we get the second strongest source position as $\mathbf{p}_2^* = \hat{\mathbf{p}}_{n_2}$, $n_2 \neq n_1 \in [1, \cdots, N]$, as well as the source signal $\hat{s}_2 = \mathbf{w}_n^\dagger \mathbf{z}$. Then we can subtract these two estimated sources from the total beamforming powers $\mathbf{y}$ and try to find the third strongest source in the actual residual powers $\mathbf{q}_2$ as:

$$\begin{cases} \mathbf{q}_2 & = \mathbf{y} - \alpha \ (\tilde{\mathbf{A}}^\dagger \hat{\mathbf{R}}_1 \tilde{\mathbf{A}} + \tilde{\mathbf{A}}^\dagger \hat{\mathbf{R}}_2 \tilde{\mathbf{A}}) \\ \hat{\mathbf{p}}_{n_3} & = \arg\max_{\mathbf{p}_{n_3} \in \mathbf{P}} \{\mathbf{q}_2\} \end{cases} \tag{4.57}$$

where $\hat{\mathbf{R}}_2 = \frac{1}{I} \sum_{i=1}^{I} \hat{\mathbf{z}}_{2\,i}^\dagger \, \hat{\mathbf{z}}_{2\,i}$ is the estimated covariance matrix of measured signals $\hat{\mathbf{z}}_2$, which can be obtained as $\hat{\mathbf{z}}_2 = \mathbf{A}\hat{s}_2 + \mathbf{e}$ in Eq.(4.54).

This procedure of Eq.(4.53)-Eq.(4.57) is repeated until all K sources are discovered. However, CLEAN depends on some important parameters such as the known source number K, selected attenuation factor α and iteration number etc. CLEAN tends to eliminate weak sources due to the background noise interference.

4.5 Regularizations

All the mentioned spatial filter and deconvolution methods merely consider the forward model of source signal (power) propagation. In order to estimate $\mathbf{x}$ from the ill-posed determined linear system of Eq.(4.46), we need to bring in useful constraints on these unknown variables. Since source powers $\mathbf{x}$ are as K-sparsity as their source signals $\mathbf{s}$ in Eq.(3.14) in the spatial domain, even though source number K is unknown, we can use the sparse power distributions, as well as the non-negative property to reconstruct source powers $\mathbf{x}$ and background noise power σ_e^2 for robust acoustic imaging. The regularization methods can be usually expressed as follows:

$$\begin{cases} \hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \left\{ \underbrace{||\mathbf{y} - \mathbf{C}\,\mathbf{x}||_2^2}_{Data\ fitting} + \underbrace{\alpha\, \mathcal{F}(\mathbf{x})}_{Regularization} \right\} \\ \text{s.t.} \quad \mathbf{x} \succeq 0 \end{cases}, \tag{4.58}$$

where $\succeq$ denotes each item of a vector is non-negative; the first ℓ_2-norm $||\cdot||_2$ represents the data fitting part, which aims to minimize the errors between the modeled powers from Eq.(4.46) and the measured beamforming powers from Eq.(4.10). The second term $\mathcal{F}(\cdot)$ usually takes the following forms:

1. $\mathcal{F}(\mathbf{x}) = ||\mathbf{x}||_0$ refers to the number of non-zero elements of $\mathbf{x}$, which can give the sparsest solution; unfortunately, the optimization of $||\mathbf{x}||_0$ becomes intractable for large dimension $\mathbf{x}$, even if the Iterative Hard Threshholding (IHT) method have been proposed recently [38].

2. $\mathcal{F}(\mathbf{x}) = ||\mathbf{x}||_1$ involves the ℓ_1 regularization which can enforce sparse solution; and the optimization of $||\mathbf{x}||_1$ can be solved by the LASSO [112] and atomic decomposition by basis pursuit [26]. Thanks to sparsity constraint on source powers which is equivalent to using $||\mathbf{x}||_1$ regularization, the DAMAS with sparsity constraint (SC-DAMAS) [126] can better improve the spatial resolution than the CBF and DAMAS.

3. $\mathcal{F}(\mathbf{x}) = ||\mathbf{w}\,\mathbf{x}||_l$ refers to the weighted sparse regularization, where $\mathbf{w}$ denotes the weighted coefficients and $0 < l < 1$; the optimization of $||\mathbf{wx}||_l$ can be solved by the Iterative Reweighed algorithm [21], and it can obtain the sparser reconstruction of source powers than the ℓ_1 does, but $0 < l < 1$ involves the difficult non-convex optimization problem.

4. $\mathcal{F}(\mathbf{x}) = ||\mathcal{D}\,\mathbf{x}||_2^2$ refers to the Tikhonov regularization [70], where the operator $\mathcal{D}$ can suppress the noise interference, but $||\cdot||_2^2$ can not offer as sparse solutions as the ℓ_1 regularization does.

The term α in Eq.(4.58) denotes regularization parameter. Tuning α is an important issue, and various empirical approaches [48, 26, 70] have been proposed. For example, the optimal parameter for the Tikhonov regularization is selected via Generalized Cross Validation[52] and L-curve method [70]. However, we expect that α should be automatically determined during the estimations in order to approach the global optimization.

The sparse regularization in Eq.(4.58) with $\mathcal{F}(\mathbf{x}) = ||\mathbf{x}||_1$ is equivalent to the sparsity constraint as:

$$\begin{cases} \hat{\mathbf{x}} = \arg\min_{(\mathbf{x})} \left\{ ||\mathbf{y} - \mathbf{C}\mathbf{x}||_2^2 \right\} \\ \text{s.t.} \quad ||\mathbf{x}||_1 = \beta, \quad \mathbf{x} \succeq 0 \end{cases}, \tag{4.59}$$

where β denotes the total source power; $||\mathbf{x}||_1 = \beta$ serves the sparsity constraint; $\mathbf{x} \succeq 0$ denotes $x_n \in \mathbf{x} \geq 0$. Recently, many effective methods have been proposed to solve Eq.(4.59), such as the DAMAS with sparsity constraint (SC-DAMAS) [126]. But β selection is the key issue for good performance. In SC-DAMAS [126], β is often fixed for simplicity.

4.6 Other robust or high resolution methods

Based on the mentioned classical methods, we introduce two other methods which are recently well developed for source localization, such as

- The Covariance Matrix Fitting (CMF) method [128] can directly estimate the covariance matrix of uncorrelated source signals using the sparsity constraint on the source powers as follow:

$$\begin{cases} (\hat{\mathbf{x}}, \hat{\sigma}^2) = \arg\min_{(\mathbf{x},\sigma_e^2)} \left\{ ||\hat{\mathbf{R}}_z - \mathbf{A}\,\mathbf{R}_s\,\mathbf{A}^\dagger - \sigma_e^2 \mathbf{I}_M||_2^2 \right\} \\ \text{s.t.} \quad tr[\mathbf{R}_s] = \beta,\ \mathbf{x} \succeq 0,\ \sigma_e^2 \geq 0 \end{cases}, \tag{4.60}$$

  where $tr[\cdot]$ denotes the matrix trace. For uncorrelated source signals, the the covariance matrix $\mathbf{R}_s$ is a diagonal matrix, whose diagonal items are composed of the source powers. So that we get $tr[\mathbf{R}_s] = \|\mathbf{x}\|_1$. The CMF can estimate the correlated sources, but it has much larger dimension of variables to be estimated than regularization methods in Eq.(4.58)-Eq.(4.59).

- Recently, the Spectral Estimation Method (SEM) [15, 14] has been proposed to improve the robustness in acoustic imaging. The SEM often pre-estimates background noise power in the wind tunnel without objects, then subtracts this noise power from the measured data with objects. This reference noise could be estimated by measuring the observed signals without any object in wind tunnel. However, the estimated noise power might not be the same one when the object exists in the wind tunnel.

4.7 Simulations of monopole and complex sources

This section shows the typical simulations on source localization and power reconstruction. To focus on method comparisons, we do not consider the ground reflection and wind refraction in simulations, but we should reconsider these multi-path propagation effects in wind tunnel experiments.

4.7.1 Simulation configurations

In Fig.3.3, simulation configurations come from the wind tunnel experiments carried out by Renault SAS [86] as shown in Fig.4.8:

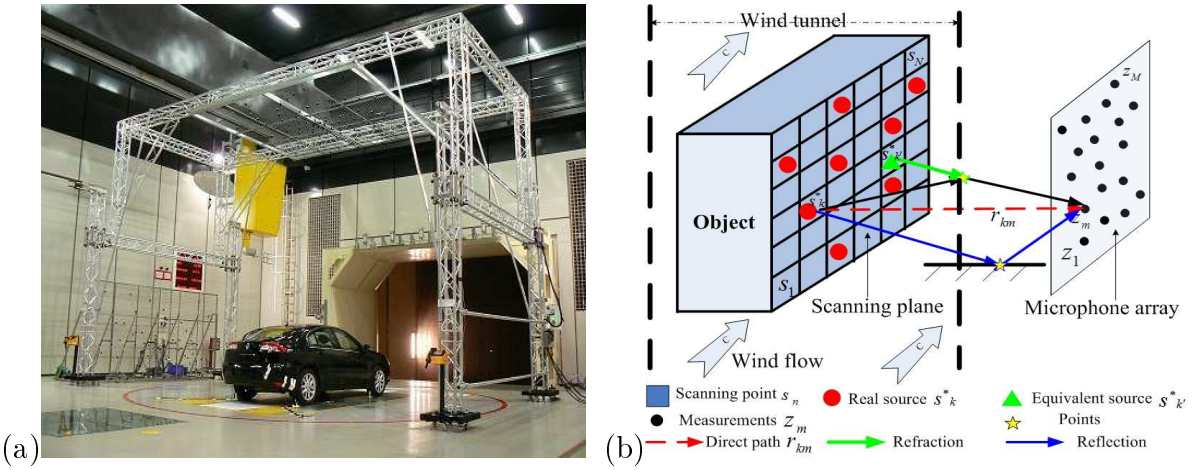

Figure 4.8: Illustration of simulations (a) Wind tunnel S2A [86] (b) Illustration of simulations and experiments.

- For the sensor array, there are $M = 64$ non-uniform sensors locating on the vertical plane. This Non-Uniform sensor Array (NUA) has a longer horizontal size than the vertical size [86]. And it has the $d = 2$m averaged size. The advantage of NUA array is that it can yield almost the same performance but less computation burden than the uniform array with the same sensors as discussed in article[68].

- For the wind tunnel, we do not consider the ground reflection and wind refraction in order to focus on method comparisons. But we reconsider the multi-path propagation effects in the real data of wind tunnel experiments in the following chapters.

- For the simulated sources in Fig.4.9(a), we have simulated 4 monopoles and 5 complex sources with different patterns; and the total number of monopole sources is $K = 23$. Sources are spaced at least 20cm. Original source powers $\mathbf{x}^*$ are within [0.08,2] ([-10.3,3.7]dB) and 14dB dynamic range.

- For the image result, there are 4 parts on Fig.4.9(a): the center image shows the source positions, patterns and powers. On its right, the colormap shows the dynamic range of source powers, in which, the dark-red colors represent strong powers, while light white colors represent the weak. On the left and bottom, 2 profile figures reveal the positions of 4 monopoles and complex sources on the center.

- Background noises: to simulate the very noisy background, the i.i.d AGWN noise power is set $\sigma_e^2 = 0.86$ (-0.7dB), thus the averaged SNR is 0dB.

- For other parameters, $D = 4.50\text{m}$ is the distance between the sensor plane and source plane. $c_0 = 340\text{m/s}$ is the acoustic speed in the common air. $T = 10000$ is the total number of samplings, which is large enough to calculate the covariance matrix in Eq.(4.6).

In order to make a fair comparison with other classical methods, the following simulation parameters should be selected carefully:

- According to Eq.(4.21), the spatial resolution of the CBF at $f = 2500Hz$ in Eq.(4.21) is $\delta_p^{CBF} = \frac{R\,c_0}{d\,f} = 31\text{cm}$.
- In order to avoid the spatial aliasing problem as discussed in the DAMAS [18], the discrete grid is set $\Delta p = 5\text{cm}$ and the frequency should be $f < 3100\text{Hz}$, so that they satisfy $\frac{\Delta p}{\delta_p^{CBF}} < 0.2$.
- To simulate a sparse distribution of discrete source signals, $\Delta p = 5\text{cm}$ is used to discretize the $100 \times 150\ cm^2$ source plane, so that the power image is of 21×31 pixels. Since total grid number $N = 651$ is much more larger than original source number $K = 23$, the discrete source signals $\mathbf{s}$ and their source powers $\mathbf{x}$ in Fig.4.9(a) are both K-sparsity signals.

Reconstruction results are presented on images which can directly show the estimated source powers (dB) and positions. Then we use 3 criteria to quantitatively evaluate estimation performance as follows:

- The first one is the averaged estimation error of original source powers, defined as

$$\overline{\Delta x^*} = \frac{1}{K}\sum_{k=1}^{K} |\hat{x}_k^* - x_k^*|\,, \tag{4.61}$$

where original source powers $\mathbf{x}^* = [x_1^*, \cdots, x_K^*]^T$ are defined as:

$$\mathbf{x}^* = \text{diag}\,[\mathbf{R}_{s^*}]\ , \quad \mathbf{R}_{s^*} = \mathbb{E}[\mathbf{s}^*\mathbf{s}^{*\dagger}]\,, \tag{4.62}$$

where $\mathbf{R}_{s^*}$ denotes the covariance matrix of the original source signals $\mathbf{s}^*$, which are defined in Chapter 3. The results are shown in Table 4.1 and 4.3.

- The second is the reconstruction error of source power image, defined as

$$\delta_2 = \frac{\|\hat{\mathbf{x}} - \mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2}\,, \tag{4.63}$$

where compared with $\overline{\Delta x^*}$, the value of δ_2 not only depends on the estimated positions and powers, but also on the suppression of the background noises. The results are shown in Table 4.1.

- The third is the averaged position error, defined as

$$\overline{\Delta p^*} = \frac{\sqrt{\sum_{k=1}^{K}(\|\mathbf{p}_k^* - \hat{\mathbf{p}}_k^*\|^2)}}{\sqrt{\sum_{k=1}^{K}\|\mathbf{p}_k^*\|^2}}, \tag{4.64}$$

where $\hat{\mathbf{p}}_k^*$ denotes the estimated positions of kth original source. The results are shown in Table 4.2.

4.7.2 Simulation results

Firstly we show the method comparisons at 2500Hz, since this frequency is very sensitive to human hearing and affect acoustic comfort. In Fig.4.9, the CBF gives a blurred image of source power distributions, since its spatial resolution in Eq.(4.21) is $\delta_p^{CBF} = 31$cm. The Capon (MVDR) obtains the better resolution than the CBF, owing to its MVDR criterion in Eq.(4.24). The MUSIC achieves much better resolution than CBF and Capon thanks to exact subspace decomposition with source number $K = 23$ known, but it can not provide the accurate source power estimations. The DR-DAMAS with 5000 iterations (5000i) greatly improves the resolution and offer better source power estimations than the Capon and MUSIC, but the Diagonal Removal technique cause to eliminate some weak sources. The CLEAN has the similar performance as the DAMAS, but it is sensitive to the noises. The SC-DAMAS well detects some of strong sources, but they do not provide reliable estimation of weak sources in strong background noises. The CMF achieves better estimation on the noise power and distinguishes most of sources, however, it fails to reconstruct some patterns of weak sources.

4.8 Real data of wind tunnel experiments

Figure 4.8 shows the static vehicle (no engine noise), microphone sensor array and the wind flow at the speed of 160km/h in the wind tunnel S2A [86]. One of objects of this wind tunnel is to detect acoustic powers and positions on the car surface. This wind tunnel can simulate a traveling car on the high-way and measure its acoustic influence to the passengers-by.

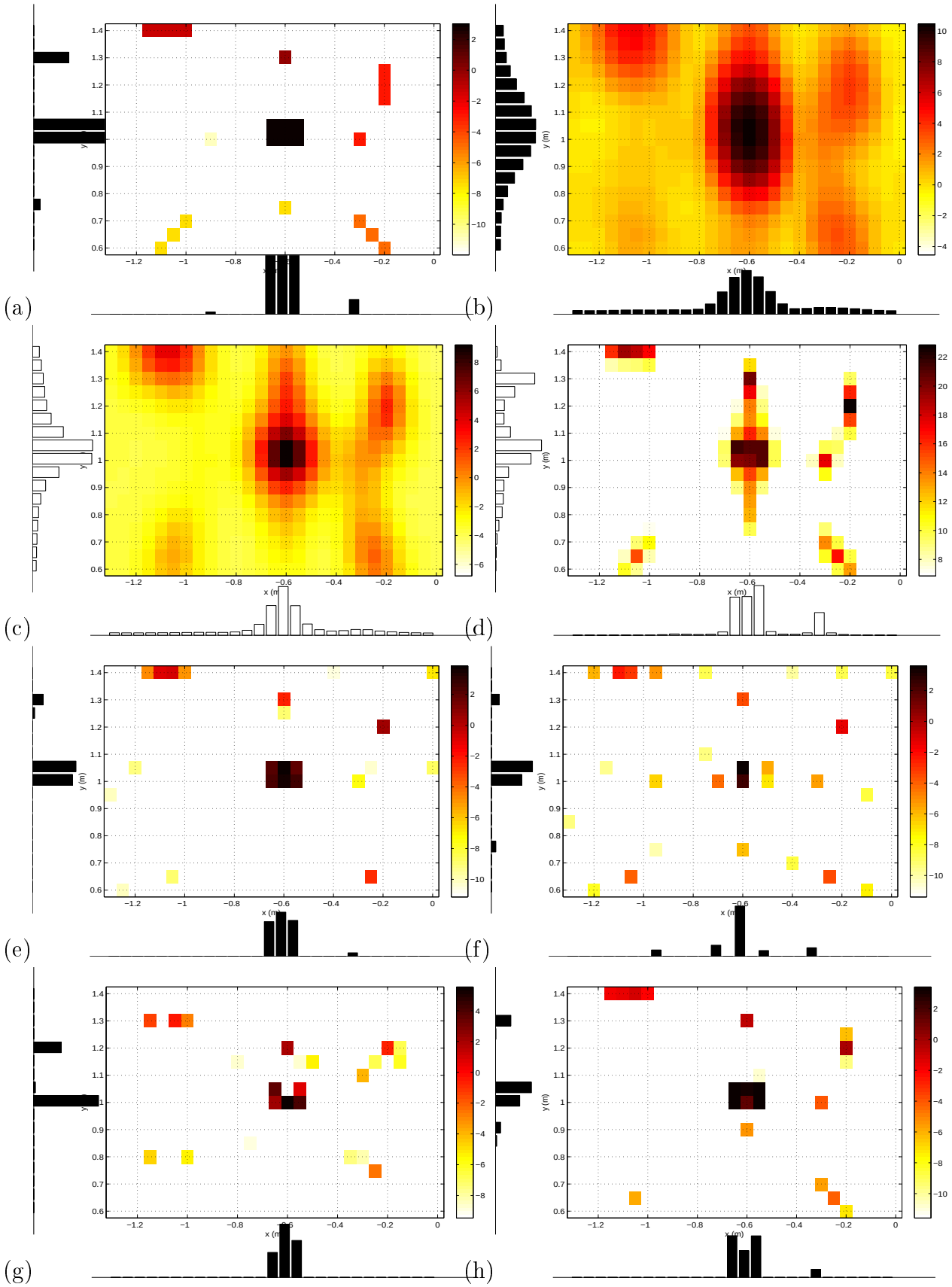

Figure 4.9: Simulation on monopole sources with 14dB power dynamic range at 2500Hz, $\sigma_e^2 = 0.86$, SNR=0dB and 15dB display: (a) Sources (b) CBF (c) Capon (d) MUSIC with $K = 23$ (e) DR-DAMAS with 5000 iterations (5000i) (f) CLEAN (5000i) (g) SC-DAMAS and (h) CMF

Table 4.1: Power estimations of 4 monopole sources by average power estimation error $\overline{\Delta x^*}$, relative errors of power image reconstruction δ_2 and estimated noise power $\hat{\sigma^2}_e$ at 2500Hz, SNR=0dB, simulated $\sigma_e^2 = 0.86$; '-' means unavailable.

Source power	0.08	0.18	0.98	0.50	$\overline{\Delta x^*}$	δ_2	$\hat{\sigma^2}_e$
CBF	1.57	11.28	3.51	2.02	69.64	121.9	-
Capon	0.50	0.73	1.50	0.98	0.49	14.62	-
MUSIC	3.52	9.37	108.13	51.42	42.60	6000	-
DR-DAMAS	-	-	0.77	0.23	0.30	0.08	-
CLEAN	-	0.25	0.44	0.28	0.87	0.67	-
SC-DAMAS	-	-	-	-	1.03	0.58	-
CMF	0.09	-	0.80	0.40	0.31	0.10	0.89

Table 4.2: Position estimations of 4 monopole sources by averaged position errors $\overline{\Delta p^*}$ at 2500Hz, SNR=0dB; '-' means unavailable.

Source position	(-0.9,1)	(-0.6,0.75)	(-0.3,1)	(-0.6,1.3)	$\overline{\Delta p^*}$
CBF	-	-	-	-	1
Capon	-	(-0.6,0.85)	-	(-0.6,1.1)	0.52
MUSIC	-	(-0.6,0.80)	(-0.3,1)	(-0.6,1.3)	0.20
DR-DAMAS	-	-	(-0.3,1)	(-0.6,1.3)	0.43
CLEAN	(-0.95,1)	(-0.6,0.75)	(-0.3,1)	(-0.6,1.3)	0.01
SC-DAMAS	-	-	-	(-0.6,1.2)	0.63
CMF	-	(-0.6,0.9)	(-0.3,1)	(-0.6,1.3)	0.29

Table 4.3: Power estimations of the complex source on the center of image by power estimation error $\overline{\Delta x^*}$ at 2500Hz, SNR=0dB; '-' means unavailable.

Source power	2.00	2.00	2.00	2.00	2.00	2.00	$\overline{\Delta x^*}$
CBF	2.64	9.60	9.70	9.64	11.34	9.77	6.78
Capon	5.18	8.31	4.99	5.01	8.21	4.98	4.13
MUSIC	143	127	127	97	98	126	117
DR-DAMAS	2.15	2.05	1.82	1.83	2.50	1.45	0.27
CLEAN	2.29	0.37	1.69	-	0.27	0.34	1.27
SC-DAMAS	1.68	2.49	1.16	0.10	2.23	0.65	0.75
CMF	1.36	2.86	2.07	2.09	1.92	1.05	0.45

4.8.1 Experiment configurations

The parameter configurations and multi-path propagation effects are shown in Fig.(4.10). The experiment configurations are introduced as follows:

- Source plane. We suppose that all acoustic sources locate on the same

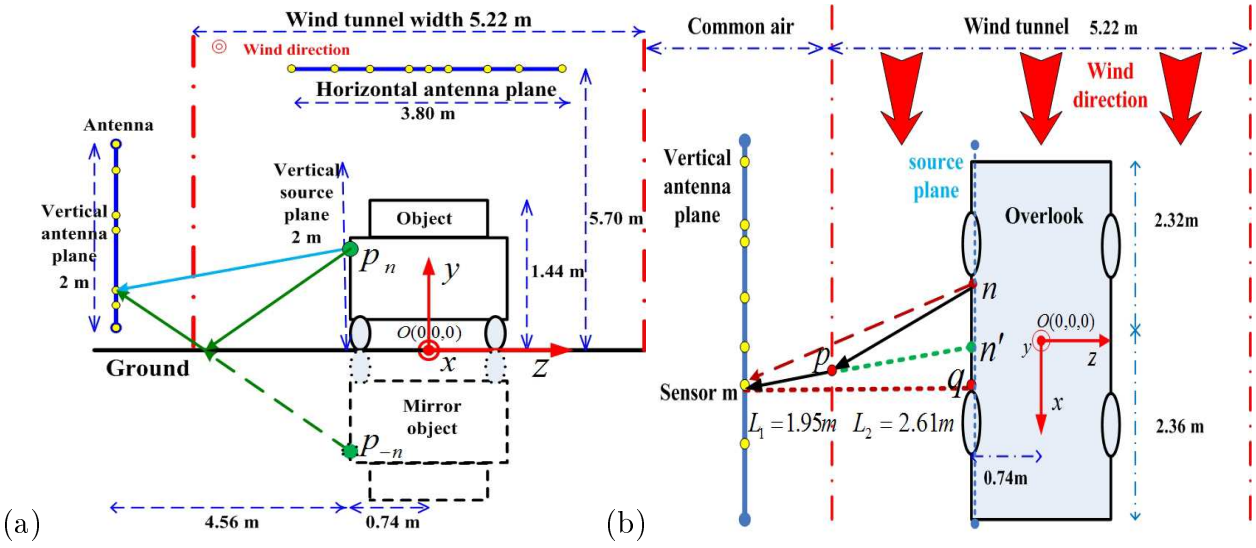

Figure 4.10: Configurations of wind tunnel experiment: (a) Front-look and ground reflection (b) Overlook and wind refraction.

2D plane, since the curvature of the car side is relatively small compared with the distance $D = 4.5m$ between the car and array plane. The surface of car side is of 150× 500 cm^2, and we discretize this source plane into 31×101 pixels by using identical grid $\Delta p = 5cm$. And we also focus on a small region of the rear-view mirror: 1×1.5 m^2 (21×31 pixels), as shown in Fig.4.11(a).

- Sparsity of the source. For the acoustic imaging on the vehicle surface in wind tunnel tests in Fig.(4.8), acoustic sources often sparsely locate on the rear-view mirrors and around the wheels, while on the rests of the parts, there are few significant sources. Therefore, the discrete source signals $\mathbf{s}$ and their powers $\mathbf{x}$ are both sparse signals.

- Sensor plane is the same as discussed on Section 4.7.

- Result images are shown by normalized dB images with 10dB span. And the colormap indicates the power range.

- Signal processing parameters. In the real data, there are $T = 524288$ samplings with the sampling frequency f_s=2.56×10^4 Hz. As discussed in Chapter 3, we separate these samplings into $I = 204$ blocks with $L = 2560$ samplings per block. The working frequency band is chosen as [2400,2600]Hz, which is sensitive to acoustic comfort of human being.

- Wind tunnel multi-path effects. We consider the ground reflection in Eq.(3.16) and wind refraction [29] in both synthetic and real data. For the actual propagation time $\tau_{n,m}$ and distance $r_{n,m}$ in Eq.(3.11), we

apply equivalent source to make refraction correction. As shown in Fig.4.10(b), for sensor m, it seems to receive the signal from equivalent source signal $s_{n'}$, instead of original source signal s_n, along a direct path $r_{n',m}$ during the same propagation time $\tau_{n',m}$, as if there is no wind influence in the wind tunnel. For $\tau_{-n,m}$ and $r_{-n,m}$ in Eq.(3.11), we use the mirror source signal s_{-n} to correct the ground reflection as shown in Fig.3.3(b). The details of the propagation corrections are discussed in A and B.

4.8.2 Experiment results

Figure.4.11 illustrates the estimated power images of mentioned methods at 2500Hz. In Fig.4.11(b), the CBF merely gives a blurred image of strong sources around the front wheel, rear-view mirror and back wheel. In Fig.4.11(c), the DAMAS well deconvolves the beamforming image, and discovers weak sources on the front light, front cover and side window. However, many false targets are also detected in the air. In Fig.4.11(d), DR-DAMAS eliminates most of the artifacts, but it also removes off some of weak sources. Figure.4.11(e) shows that the CLEAN overcomes the drawbacks of the DAMAS, but we have to carefully select the parameters for this good performance. In Fig.4.11(f), the SC-DAMAS has a better noise suppression than the DAMAS and CLEAN owing to the sparsity parameter selection, but SC-DAMAS overwhelms too much both the noises and the sources, so that it does not provide a wide dynamic range of source power estimations.

Table 4.4: Computational cost for treating whole car: image 30×100 pixels, at 2500Hz, based on CPU: 3.33GHz, '-' means unavailable.

Methods	CBF	DAMAS (5000i)	DR-DAMAS (5000i)	CLEAN	SC-DAMAS	CMF
Time (s)	1	10	11	45	1254	-

Furthermore, we give the computational costs in Table 4.4: the CBF is the quickest method; the deconvolution methods such as DAMAS and CLEAN have a moderate computations; the SC-DAMAS takes much more time than the above due to the sparsity constraint in Eq.(4.59), but it is still applicable; however, the CMF is hardly to work due to the huge dimensions of variables in covariance matrix in Eq.(4.60).

Based on the acoustic imaging on the car side, we investigate a small part

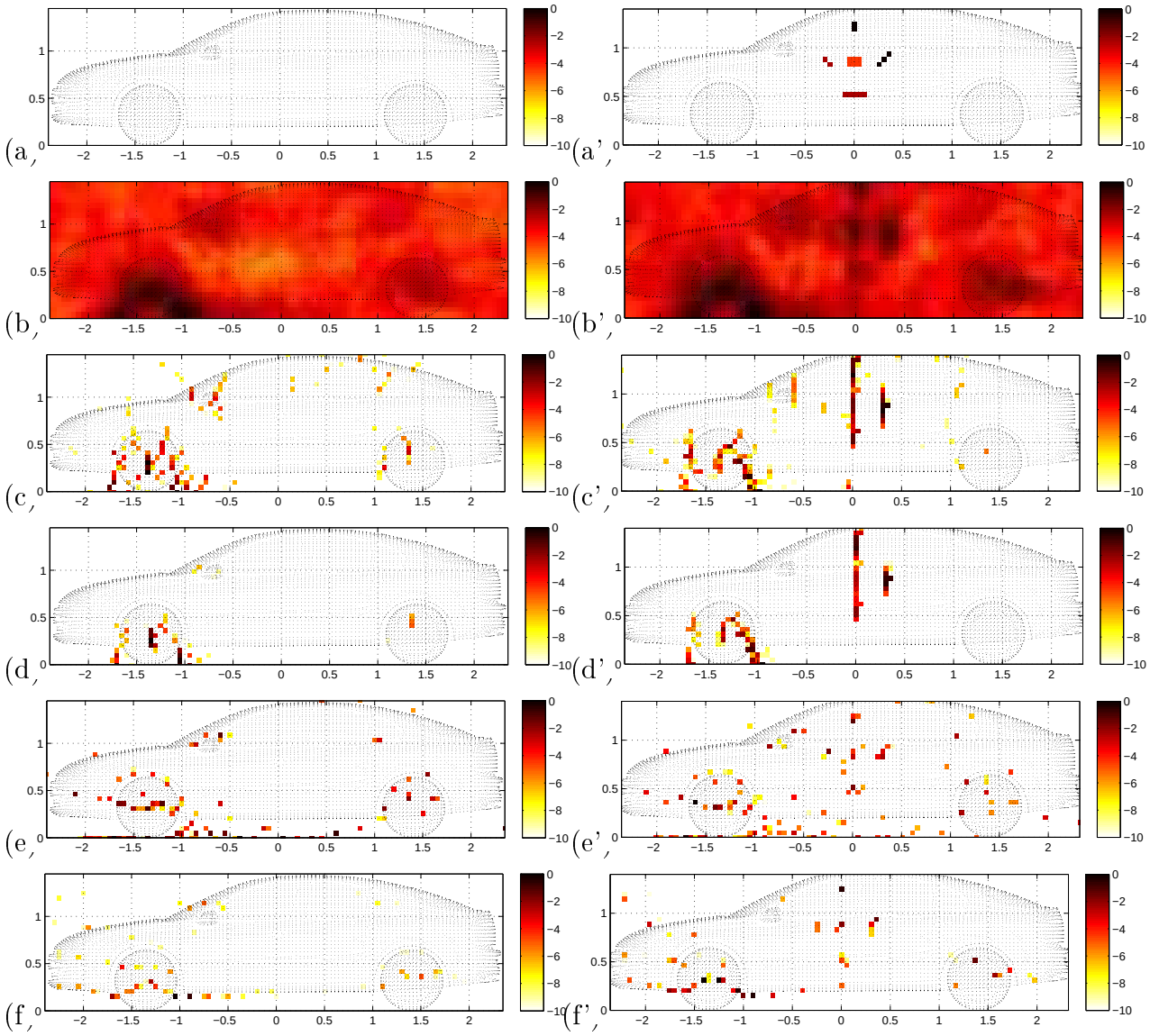

Figure 4.11: Acoustic imaging on the vehicle side at 2500Hz.
Left: Real data (a) Vehicle surface (b) Beamforming (c) DAMAS (5000i) (d) DR-DAMAS (5000i) (e) CLEAN (f) SC-DAMAS
Right: Hybrid data (a') Simulated sources (b') Beamforming (c') DAMAS (5000i) (d') DR-DAMAS (5000i) (e') CLEAN (f') SC-DAMAS.

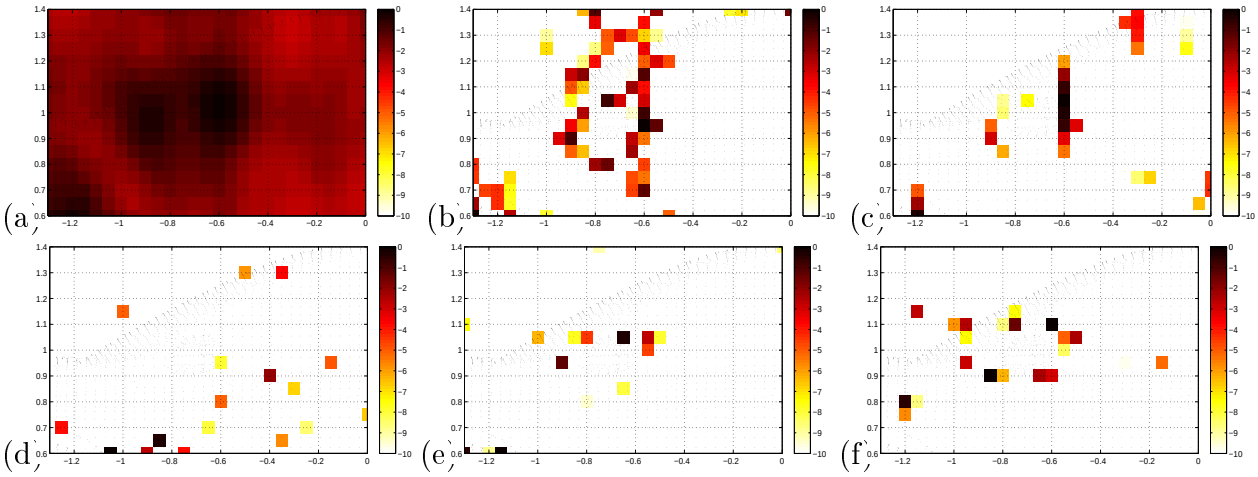

Figure 4.12: Acoustic imaging of rear-view mirror at 2500Hz: (a) CBF (b) DAMAS (5000i) (c) DR-DAMAS (5000i) (d) CLEAN (e) CMF and (f) SC-DAMAS

of the rear-view mirror. In Fig.4.12(a), the CBF detects strong sources on the corner of the front wheel and rear-view mirror. The DAMAS in Fig.4.12(b) improves the spatial resolutions, but it causes some unexpected spots. In Fig.4.12c, the DR-DAMAS eliminates most of false spots. In Fig.4.12(e)-(f), the CMF and SC-DAMAS achieve much better resolutions on the rear-view mirror.

4.9 Results with hybrid data

Even though classical method can obtain good performance on real data from wind tunnel experiments, it is not sufficient for performance validation. This is because the exact acoustic source distributions on the vehicle are not known beforehand. To further verify the performances of these methods, we use the hybrid data which composes of known synthetic sources and the real data. In order to avoid overlapping the original sources, the synthetic sources are set on the region where there are no significant sources powers. In Fig.4.11(a'), five synthetic complex sources with different patterns are generated at 2500Hz, whose powers are within $[-4.5, 0]$dB. We expect that effective methods can detect both the synthetic and original source powers from the hybrid data. If these known synthetic sources are successfully recovered, these methods can be able to effectively reconstruct the original acoustic sources on the vehicle surface.

4.9.1 Synthetic sources model

Based on the assumptions in Section 3.1, we suppose K' Gaussian white variables $\mathbf{w}'(t) = \{w'_k(t),\ k = 1, \cdots, K'\}$, with $w'_k(t) \sim \mathcal{N}(0, \sigma_k^2)$, and σ_k^2 is the variance of the variable $w'_k(t)$. In order to generate wide-band source signals $\{s'_k(t),\ k = 1, \cdots, K'\}$, $w'_k(t)$ is convoluted by the impulse response $h(t)$ (for instance Blackman filter). Thus the synthetic source signals are modeled by

$$s'_k(t) = w'_k(t) * h(t), \quad k = 1, \cdots, K', \tag{4.65}$$

where $*$ denotes convolution operation. Since the positions $\mathbf{P}'$ of synthetic $\mathbf{s}'$ are set to be known, the measurements $\boldsymbol{z}'$ at the sensor array can be calculated by the forward model of acoustic signal propagation in Eq.(3.15). Then we generate the hybrid data by adding $\boldsymbol{z}'$ to the real measured data $\boldsymbol{z}$.

4.9.2 Acoustic imaging results with hybrid data

For the reconstructed synthetic sources in Fig.4.11(b')-(f'), the CLEAN and SC-DAMAS get better estimations than the CBF and DAMAS. Thus their source reconstructions from the real data are more acceptable. But the some important parameters such as the attenuation factor in CLEAN and sparsity parameter in SC-DAMAS should be carefully selected for better use.

4.10 Conclusions and perspectives

Table 4.5: General performance of classical methods.

Methods	CBF	Capon	MUSIC	CLEAN	DAMAS	DR-DAMAS	SC-DAMAS	CMF
Resolutions	Low	Normal	High	Normal	Normal	Normal	High	High
Dynamic Range	Narrow	Narrow	Normal	Normal	Normal	Normal	Normal	Wide
Noise	Robust	Sensitive	Sensitive	Sensitive	Sensitive	Normal	Sensitive	Robust
Computation	Least	Normal	Normal	Normal	Normal	Normal	High	Higher
Samples	Normal	More	More	Normal	Normal	Normal	Normal	More
Source number	No	No	Required	Required	No	No	Required	Required

In Table 4.5, we give a brief summary for the advantages and drawbacks of mentioned classical methods. To summarize, all the above methods have good performances on their focused aspects:

The classical spatial filter methods aim to solve the forward model of acoustic **signal** propagation in Eq.(3.15), and they offer a direct estimation for source localization.

- Conventional beamforming (CBF) is a data-independent method, and offers a simple, direct and robust estimation for source powers; but the CBF cannot easily obtain high spatial resolution or exact acoustic imaging due to its low spatial resolution, especially in low frequencies.
- Capon method is a data-dependent method, and provides a better spatial resolution and source power estimation than the CBF; but the Capon is less robust, since it requires more measured samples to calculate the inversion of covariance matrix of measured signals.
- MUSIC is also a data-dependent method, and achieves very high spatial resolution than the CBF and Capon; but MUSIC is much less robust, since it relies on not only the large number of measured samples, but also the subspace decomposition which requires to known the source number or SNR level beforehand. Moreover, MUSIC cannot directly provide the source power estimation, but just pseudo-power.

The deconvolution methods aim to solve the forward model of acoustic **power** propagation in Eq.(4.46), and they iteratively deconvolve the blurry beamforming result to obtain a high spatial resolution for source power reconstruction and localization (acoustic imaging).

- Original DAMAS refines the CBF result and improves the spatial resolution. However, it is sensitive to background noises, since the power propagation matrix $\mathbf{C}$ is ill-conditioned.
- CLEAN also iteratively extracts strong sources from the blurred beamforming image, but it depends on parameter selection for good performance, such as known source number K, selected attenuation factor α and iteration number.

Regularization methods further improve the spatial resolution of deconvolution methods by using the sparse regularizations (ℓ_0, ℓ_1 or ℓ_l with $0 < l < 1$ norm), and improve the robustness to noises owing to the Tikhonov regularization. But the regularization term inevitably increases method complexity, and regularization parameter selection becomes an essential issue for effective performance.

Covariance Matrix Fitting (CMF) method directly estimates the covariance matrix of (un)correlated source signals via sparsity constraint, but it has huge dimension of variables to be estimated.

In brief, there is no one-fits-all method. Most of the mentioned suffer one of the following drawbacks: poor spatial resolution, sensitivity to background noises, narrow dynamic range or high computational cost. In addition, some methods need to set important parameters that must be tuned accurately for each case to obtain effective performance. These limits motivate us to investigate the robust approaches for acoustic imaging with high resolution and wide dynamic range, which will be discussed in the following chapters.

.

- Je pense, donc j'existe.

René Descartes

5

Super-Resolution Approach with Sparsity Constraint

In Chapter 4, we have pointed out that the mentioned state-of-the-art methods suffer one of the following drawbacks: poor spatial resolution, sensitivity to background noises, narrow dynamic range, high computational cost and parameter dependence. Motivated by these limitations, we propose a method to improve both the spatial resolution and robustness to noises in this chapter [28, 34, 30]. There are three aspects to be considered:

- Problem of the model discrepancy of propagation steering vector. In the forward model of acoustic signal propagation of Eq.(3.15), we have modified the propagation steering vector of Eq.(3.16) by correcting wind tunnel effects, such as the ground reflections and wind refractions as discussed in A and B respectively. These corrections can be reduced the signal propagation uncertainty to some extent, but unfortunately as pointed out by article [79], the real propagation steering vector could hardly be exactly modeled, thus the signal propagation uncertainty could not be completely removed. Therefore, we take signal model uncertainty into the Eq.(3.15) as shown in Fig.(5.1):

$$\mathbf{z}_i = \mathbf{A}(\mathbf{P})\mathbf{s}_i + \mathbf{e}_i + \boldsymbol{\varepsilon}_i \,, \tag{5.1}$$

  where $i \in [1, \cdot, I]$ is the sampling block number; $\boldsymbol{\varepsilon}_i = [\varepsilon_{i\,1}, \cdots, \varepsilon_{i\,M}]$ denotes the signal model uncertainty variables caused by the acoustic ground reflection and wind refraction in wind tunnel. So that it

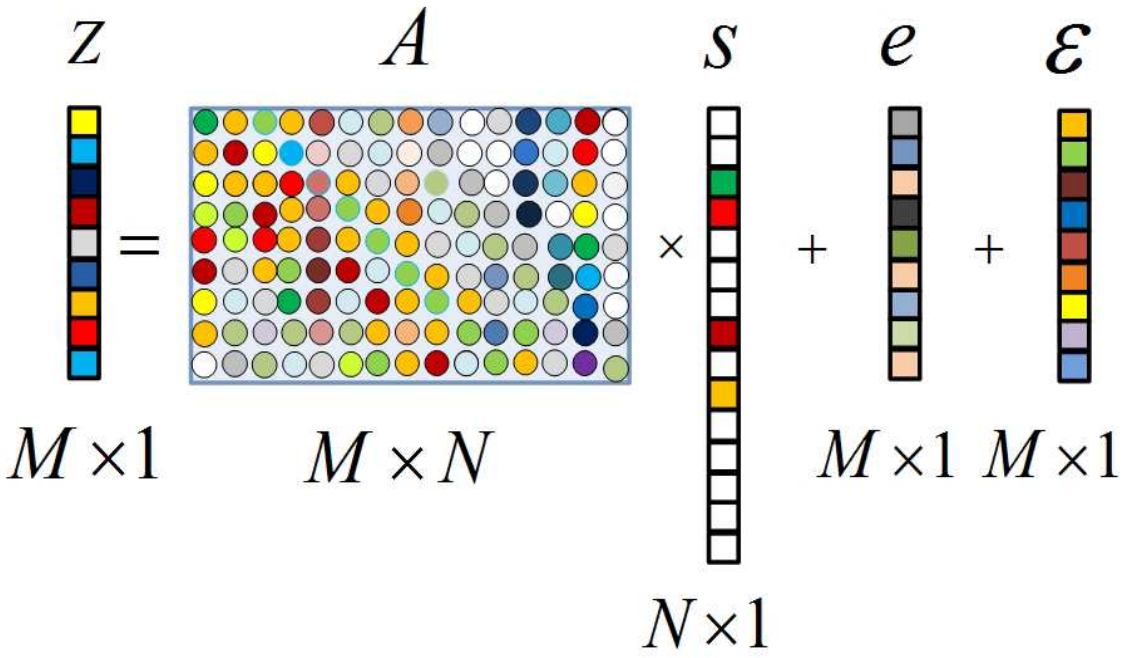

Figure 5.1: Illustration of sparse power equation with model uncertainty $\boldsymbol{\xi}$ in Eq.(5.1).

is necessary to minimize the forward model uncertainty for the model robustness. We also assume the mutual independence among forward model uncertainty, source signals and background noises.

- Super spatial resolution. The spatial resolution of the Conventional BeamForming (CBF) method depends on the wave length and array geometry, so that we call it the low resolution method. The MUSIC and deconvolution methods can achieve high resolution, which overcome the physical constraints in the CBF, so that we call them high resolution methods. The sparse constraint (regularization) methods such as the SC-DAMAS in Eq.(4.59), CMF in Eq.(4.60), as well as the proposed approach in this Chapter, they can furthermore improve the spatial resolution by adding useful constraints on source signals or effectively eliminating the background noises, so that we call the the super-resolution methods. Here we have to point out that so called 'super-resolution' is different from the one in image processing. .

- Robustness to noises. We then want to jointly estimate the source powers and background noise power. So that the noises can be removed automatically during the source power estimation.

This chapter is organized as: In order to improve the model robustness, we propose an improved model of acoustic power propagation in Section 5.1,

and we consider the model uncertainty caused by multi-path propagation in wind tunnel. Then our regularization approach with sparsity constraint is investigated in Section 5.2. Then, method comparisons with classical method are shown on simulations in Section 5.3 and real data in Section 5.4 respectively. To further confirm the effectiveness of proposed approach, Section 5.5 demonstrates the performance comparisons on the hybrid data, in which, some known synthetic sources are added to the real data. Finally, Section 5.6 concludes this chapter.

5.1 Proposed forward model of acoustic power propagations

Let's recall the beamforming power $y_n = \tilde{\mathbf{a}}_n^\dagger \mathbf{R}_z \tilde{\mathbf{a}}_n$ in Eq.(4.10). Here, the measured cross-spectrum matrix $\mathbf{R}_z = \mathbb{E}[\mathbf{z}\,\mathbf{z}^\dagger]$ defined in Eq.(4.5) is renewed by the improved model of measured signals $\mathbf{z}$ in Eq.(5.1):

$$\begin{aligned}\mathbf{R}_z &= \mathbb{E}[\mathbf{z}\mathbf{z}^\dagger] \\ &= \mathbf{A}\,\mathbf{R}_s\,\mathbf{A}^\dagger + \sigma_e^2\,\mathbf{I}_M + \mathbf{R}_\varepsilon \\ &= \sum_{n=1}^{N} x_n\,\mathbf{a}_n\mathbf{a}_n^\dagger + \sigma_e^2\,\mathbf{I}_M + \mathbf{R}_\varepsilon\,,\end{aligned} \tag{5.2}$$

where $\mathbf{R}_\varepsilon = \mathbb{E}[\boldsymbol{\varepsilon}_i^\dagger\,\boldsymbol{\varepsilon}_i]$ denotes the cross-spectrum matrix of signal model uncertainty $\boldsymbol{\varepsilon}_i$ in Eq.(5.1).

As long as the sampling number is big enough (sampling block number $I >> 1$), the improved cross-spectrum matrix $\mathbf{R}_z$ of measured signals in Eq.(5.2) is practically approximated by $\hat{\mathbf{R}}_z = \frac{1}{I}\sum_{i=1}^{I}\mathbf{z}_i\,\mathbf{z}_i^\dagger$ in Eq.(4.6), where $\mathbf{z}_i$ is the measured signal vector in the ith sampling block at all M sensors. So that on the left of $y_n = \tilde{\mathbf{a}}_n^\dagger \mathbf{R}_z \tilde{\mathbf{a}}_n$ in Eq.(4.10), we actually calculate the beamforming power as

$$y_n = \tilde{\mathbf{a}}_n^\dagger \hat{\mathbf{R}}_z \tilde{\mathbf{a}}_n\,,$$

where y_n is the measured beamforming power at the position of $\mathbf{p}_n$ on the source plane; $\tilde{\mathbf{a}}_n$ is the beamforming steering vector for the position $\mathbf{p}_n$, defined in Eq.(4.9). On the right of above equation, we replace $\hat{\mathbf{R}}_z$ by using

$\mathbf{R}_z = \sum_{q=1}^{N} x_q \, \mathbf{a}_q \mathbf{a}_q^\dagger + \sigma_e^2 \, \mathbf{I}_M + \mathbf{R}_\varepsilon$ in Eq.(5.2), it yields

$$\begin{aligned} y_n &= \tilde{\mathbf{a}}_n^\dagger \sum_{q=1}^{N} x_q \, \mathbf{a}_q \mathbf{a}_q^\dagger \, \tilde{\mathbf{a}}_n + \sigma_e^2 \, \tilde{\mathbf{a}}_n^\dagger \tilde{\mathbf{a}}_n + \tilde{\mathbf{a}}_n^\dagger \, \mathbf{R}_\varepsilon \, \tilde{\mathbf{a}}_n \\ &= \sum_{q=1}^{N} |\tilde{\mathbf{a}}_n^\dagger \, \mathbf{a}_q|^2 \, x_q + \sigma_e^2 \, \frac{1}{||\mathbf{a}_n||^2} + \tilde{\mathbf{a}}_n^\dagger \, \mathbf{R}_\varepsilon \, \tilde{\mathbf{a}}_n \\ &= \sum_{q=1}^{N} c_{nq} \, x_q + \sigma_e^2 \, \frac{1}{||\mathbf{a}_n||^2} + \tilde{\mathbf{a}}_n^\dagger \, \mathbf{R}_\varepsilon \, \tilde{\mathbf{a}}_n \, . \end{aligned} \tag{5.3}$$

where the beamforming steering vector $\tilde{\mathbf{a}}_n = \frac{\mathbf{a}_n}{||\mathbf{a}_n||^2}$ in Eq.(4.9), so that $\tilde{\mathbf{a}}_n^\dagger \tilde{\mathbf{a}}_n = \frac{1}{||\mathbf{a}_n||^2}$.

For N discrete sources, equation (5.3) can be rewritten in a vector form as shown in Fig.(5.2). Finally, we obtain the improved forward model of acoustic power propagation as:

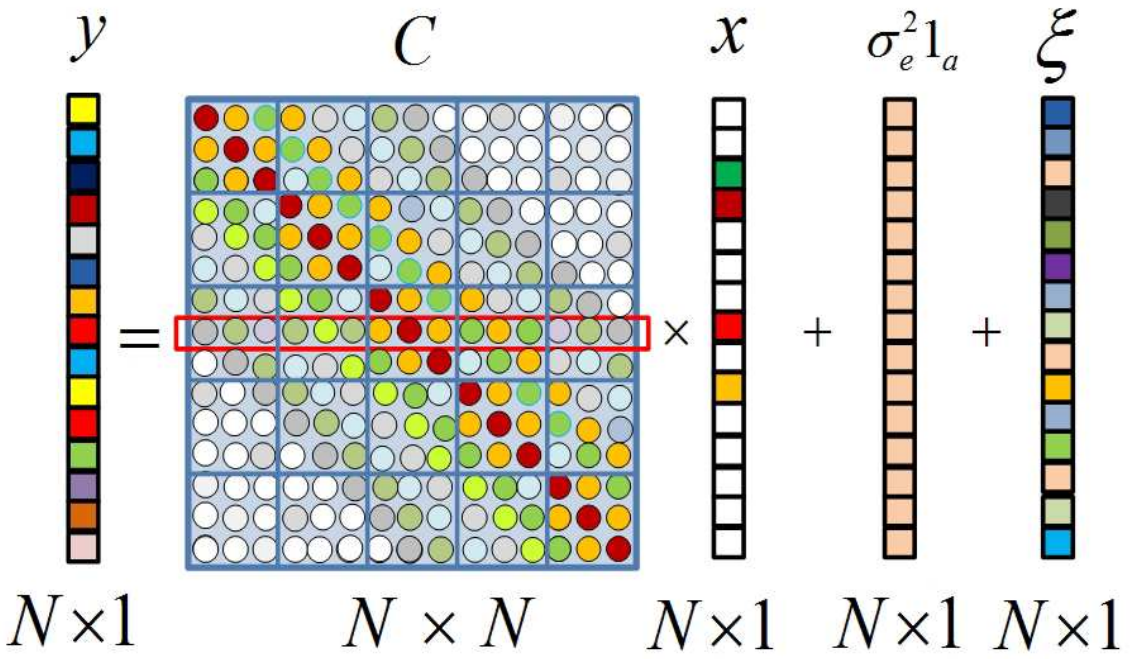

Figure 5.2: Illustration of the improved forward model power propagation with multi-path uncertainty in Eq.(5.4).

$$\mathbf{y} = \mathbf{C}\,\mathbf{x} + \sigma_e^2 \, \mathbf{1}_a + \boldsymbol{\xi} \, , \tag{5.4}$$

where

- $\mathbf{y} = [y_1, \cdots, y_N]$ is the beamforming power vector measured at sensor array, calculated from Eq.(4.10);
- $\mathbf{1}_a = [\frac{1}{\|\mathbf{a}_1\|^2}, \cdots, \frac{1}{\|\mathbf{a}_N\|^2}]^T$ is a constant vector; $\mathbf{a}_n$ is the steering vector calculated form Eq.(3.16).
- $\boldsymbol{\xi} = [\xi_1, \cdots, \xi_N]^T$ denotes the power model uncertainty caused by multi-path propagations; and ξ_n is modeled by

$$\xi_n = \tilde{\mathbf{a}}_n^{\dagger} \mathbf{R}_{\varepsilon} \tilde{\mathbf{a}}_n \,. \tag{5.5}$$

In Fig.5.3, it shows the relationship between the improved power propagation model in Eq.(5.4) and the convectional one in Eq.(4.46).

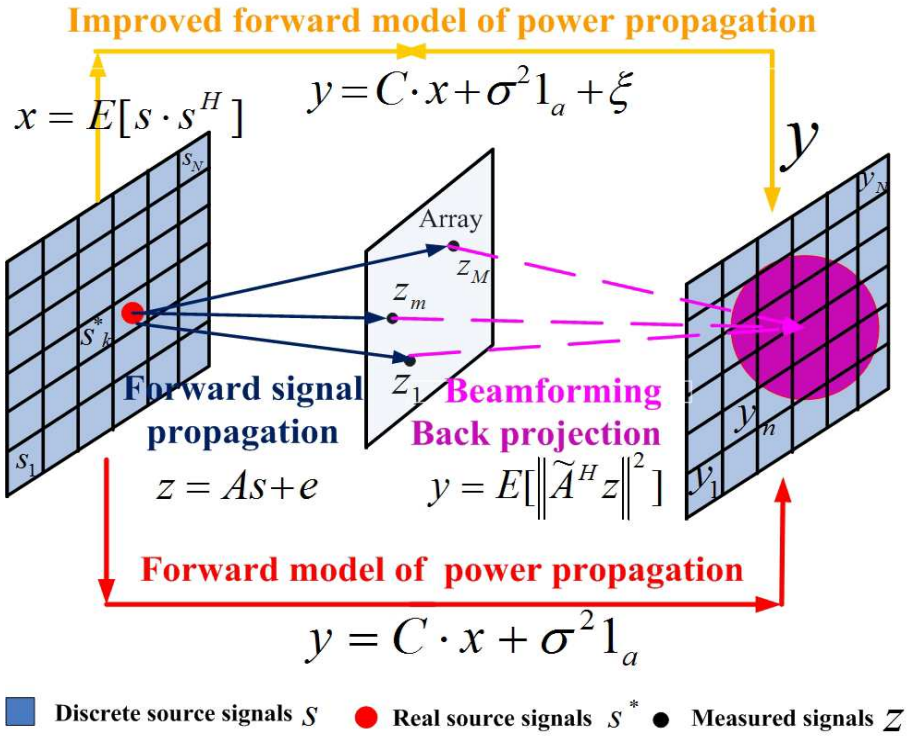

Figure 5.3: Illustration of improved forward model of acoustic power propagation.

5.2 Proposed approach using sparsity constraint

Compared with classical sparse regularization method in Eq.(4.58-4.60), we want to jointly estimate the source powers $\mathbf{x}$ and background noise power σ_e^2 by minimizing the propagation uncertainty $\boldsymbol{\xi}$. In order to obtain super resolution in strong background noises, we adaptively estimate the sparsity parameter β on total source power. Therefore, proposed robust super-resolution

approach with sparsity constraint (SC-RDAMAS) is expressed as:

$$\begin{cases} (\hat{\mathbf{x}}, \hat{\sigma_e^2}) = \arg\min\limits_{(\mathbf{x},\sigma_e^2)} \left\{ ||\mathbf{y} - \mathbf{C}\mathbf{x} - \sigma_e^2\, \mathbf{1}_a||_2^2 \right\} \\ \quad \text{s.t.} \mathbf{x} \succeq 0, \quad \|\mathbf{x}\|_1 = \beta, \quad \sigma_e^2 \geq 0 \end{cases}, \tag{5.6}$$

where sparsity parameter β is the total power of source signals, so that β is defined as:

$$\beta = \sum_{k=1}^{K} x_k^* = \|\mathbf{x}^*\|_1 = \sum_{n=1}^{N} x_n = \|\mathbf{x}\|_1 = tr[\mathbf{X}]\,, \tag{5.7}$$

where K is the total number of original source signals $\mathbf{s}^*$; and $\mathbf{x}^* = \text{diag}\left[\mathbb{E}[\mathbf{s}^*\mathbf{s}^{*\dagger}]\right]$ denotes the original source powers; $\mathbf{x} = \text{diag}\left[\mathbb{E}[\mathbf{s}\mathbf{s}^{\dagger}]\right]$ denotes the (discrete) source powers; $\mathbf{X} = \mathbb{E}[\mathbf{s}\mathbf{s}^{\dagger}]$ denotes the source power cross-spectrum matrix. If β in Eq.(5.7) is modeled too large, the estimated $\hat{\mathbf{x}}$ from Eq.(5.6) would be more dispersed than expected. If β too small, some of weak sources would be left out. Therefore, the adaptive estimation of sparsity parameter β is an essential issue in the proposed approach.

5.2.1 Adaptive estimation of sparsity parameter

Compared with the Covariance Matrix Fitting (CMF) method[128] in Eq.(4.60), we give an Adaptive estimation procedure of sparsity parameter. According to the definition of measured cross-spectrum matrix $\mathbf{R}_z$ in Eq.(4.5), we take the matrix trace as:

$$\begin{aligned} \text{tr}\,[\mathbf{R}] &= tr[\mathbf{A}\,\mathbf{X}\,\mathbf{A}^{\dagger}] + M\,\sigma_e^2 \\ &= \textstyle\sum_{n=1}^{N} \|\mathbf{a}_n\|^2\, x_n + M\,\sigma_e^2 \end{aligned}\,, \tag{5.8}$$

where $\mathbf{a}_n$ is the nth column of signal propagation matrix $\mathbf{A}$, defined in Eq.(3.16). Let $\|\mathbf{a}\|_{min}$ and $\|\mathbf{a}\|_{max}$ respectively denote the minimum and maximum vector norms within $\mathbf{a}_n$, $n = [1, \cdots, N]$. From Eq.(5.8), we have

$$\frac{1}{\|\mathbf{a}\|_{max}^2} \left(\text{tr}\,[\mathbf{R}_z] - M\,\sigma^2\right) \leq \|\mathbf{x}\|_1 \leq \frac{1}{\|\mathbf{a}\|_{min}^2} \left(\text{tr}\,[\mathbf{R}_z] - M\,\sigma^2\right)\,, \tag{5.9}$$

where $\|\mathbf{x}\|_1 = \sum_{n=1}^{N} x_n$ and $x_n \geq 0$ in Eq.(5.7). Since $\mathbf{R}_z$ is also a Hermitian matrix, it can be diagonalized as $\text{tr}\,[\mathbf{R}_z] = \text{tr}\,[\mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\dagger}] = \text{tr}\,[\mathbf{\Lambda}]$, where $\mathbf{U}$ is the unitary matrix, whose columns are eigenvectors of $\mathbf{R}_z$; and $\mathbf{\Lambda}$ is the eigenvalue matrix of $\mathbf{R}_z$. According to Eq.(5.7–5.9), we can obtain β estimation as:

$$\hat{\beta} = \frac{1}{2}\left(\frac{1}{\|\mathbf{a}\|_{max}^2} + \frac{1}{\|\mathbf{a}\|_{min}^2}\right) \left(\text{tr}\,[\hat{\mathbf{\Lambda}}] - M\,\hat{\sigma}^2\right)\,, \tag{5.10}$$

where $\hat{\mathbf{\Lambda}}$ is the eigenvalue matrix of $\hat{\mathbf{R}}_z$ in Eq.(4.6). And $\hat{\sigma}^2$ can be estimated as [128]:

$$\hat{\sigma_e^2} = \frac{1}{M - \hat{K}} \sum_{m=\hat{K}+1}^{M} \hat{\lambda}_m \,, \tag{5.11}$$

where $\hat{\lambda}_m$ denotes the eigenvalue of $\hat{\mathbf{R}}$, satisfying $\hat{\lambda}_1 \geq \cdots \geq \hat{\lambda}_K \geq \hat{\lambda}_{K+1} = \cdots = \hat{\lambda}_M = \hat{\sigma}^2$; and M is the total sensor number; $\hat{K}$ denotes the estimated source number, provided $\hat{K} \in [1, \cdots, M]$. When $\hat{K} = M$, we have noise power $\hat{\sigma}^2 = 0$.

In Eq.(5.10), source number K estimation plays an importance role in determining β. There are many methods for source number estimation such the SVD[82] and Bayesian framework [94]. Our article [34, 28] gives a fast and rough estimation on K as: let $\mathcal{F}(\lambda_m) = \mathcal{F}[\lambda_1, \cdots, \lambda_M]$ with $m \in [1, \cdots, M]$ denote the eigenvalue distribution function, where $\mathcal{F}[\cdot]$ denotes the interpolation. The first K big eigenvalues should contain the K original source powers, while the latter $M-K$ eigenvalues just equal the noise power σ_e^2. So that $\mathcal{F}(\lambda_m)$ has a sparse distribution. Suppose $\mathcal{F}(\lambda_m)$ to be second derivative, when its curvature $\kappa(\hat{K}) \approx 0$ within $\hat{K} \in [1, \cdots, M]$, we can thus get $K \approx \hat{K}$. However, since the under-estimation of source number could eliminate the weak sources, it is better to initialize K as a relative big value. Therefore, we can use the upper bound of source number as discussed in compressed sensing [42, 20, 10] as:

$$K = ||\mathbf{x}||_0 \leq \frac{1}{2}(1 + \frac{1}{\mu}) \,, \tag{5.12}$$

where $\mu = \max_{(1 \leq i \neq j \leq M)} \frac{\hat{\mathbf{R}}_i^\dagger \hat{\mathbf{R}}_j}{||\hat{\mathbf{R}}_i|| \, ||\hat{\mathbf{R}}_j||}$ denotes the incoherence of the measured cross-spectrum matrix $\hat{\mathbf{R}}_z$, where $\hat{\mathbf{R}}_i$ denotes its ith column vector. According to Eq.(4.5) and independence assumption between sources and noises, μ can reflect the incoherence of source power cross-spectrum matrix $\mathbf{X}$.

5.2.2 Proposed adaptive estimation procedure

In Eq.(5.6), proposed approach is a convex quadratic minimization under linear matrix constraints, which can be solved by interior point algorithms using MATLAB toolbox SeDuMi [109]. In order to improve the robustness to background noises σ_e^2 and sparsity parameter β, we propose an adaptive estimating algorithm as depicted in Algorithm 1: Let $\mathcal{J}(\mathbf{x}, \sigma_e^2) = ||\mathbf{y} - \mathbf{C}\mathbf{x} - \sigma_e^2 \, \mathbf{1}_a||_2^2$ define the cost function. We firstly initialize source number K

Algorithm 1 Proposed adaptive estimation procedure

1. **Input**:
   Signal propagation matrix $\mathbf{A}$ in Eq.(3.15);
   measured signal cross-spectrum matrix $\hat{\mathbf{R}}$ in Eq.(4.6);
   measured beamforming powers $\mathbf{y}$ in Eq.(4.10);
   power propagation matrix $\mathbf{C}$ in Eq.(4.47);

2. **Initialization**:
   Iteration number i=1;
   source number $\hat{K}^{(1)} = \frac{1}{2}(1 + \frac{1}{\mu})$ in Eq.(5.12);
   variables $\hat{\mathbf{x}}^{(1)}$=0; $\hat{\sigma_e^2}^{(1)}$=0;
   criterion $\mathcal{J}(\hat{\mathbf{x}}^{(1)}, \hat{\sigma_e^2}^{(1)})$=0, $\mathcal{J}(\hat{\mathbf{x}}^{(2)}, \hat{\sigma_e^2}^{(2)})$=1;

3. **Iterations**:
   While $|\mathcal{J}(\hat{\mathbf{x}}^{(i+1)}, \hat{\sigma_e^2}^{(i+1)}) - \mathcal{J}(\hat{\mathbf{x}}^{(i)}, \hat{\sigma^2}^{(i)})|$ is not small enough;
   3.1 Update: $\hat{\sigma_e^2}^{(i)}$ by Eq.(5.11), $\hat{\beta}^{(i)}$ by Eq.(5.10);
   3.2 Optimize:
   Solve Eq.(5.6) by interior point algorithm [109]:
   $(\hat{\mathbf{x}}^{(i+1)}, \hat{\sigma_e^2}^{(i+1)}) = \arg\min_{(\mathbf{x},\sigma_e^2)} \left\{ \mathcal{J}(\hat{\mathbf{x}}^{(i)}, \hat{\sigma_e^2}^{(i)}) \right\}$
   s.t. $||\mathbf{x}||_1 \leq \hat{\beta}^{(i)}$, $\hat{\mathbf{x}}^{(i+1)} \geq 0$, $\hat{\sigma_e^2}^{(i+1)} \geq 0$;
   3.3 Iterate: $\hat{K}^{(i+1)} = \hat{K}^{(i)} - 1$, i=i+1;

4. **Output**: $\hat{\mathbf{x}}$, $\hat{\sigma^2}_e$, $\hat{K}$, $\hat{\beta}$.

by using the matrix incoherence μ of Eq.(5.12). Then $\hat{\sigma}^2$ is obtained from Eq.(5.11) and $\hat{\beta}$ from (5.10). And then we simultaneously estimate source powers $\mathbf{x}$ and σ_e^2 by using the interior point algorithm [109]. Finally we update $\hat{K}^{(i+1)} = \hat{K}^{(i)} - 1$ for a new estimation.

5.3 Simulations of monopole and complex sources

This section shows the typical simulations on source power reconstruction and localization of monopole sources. To focus on method comparisons, we do not consider the ground reflection and wind refraction in simulations. But we should reconsider the multi-path propagation effects in wind tunnel experiments. The proposed SC-RDAMAS approach is compared with the CBF, DAMAS, CLEAN, CMF and SC-DAMAS methods mentioned discussed in Chapter 4.

The simulation configurations are also the same as the one in Chapter 4.7 as shown in above figure: there are $M = 64$ non-uniform sensors locating

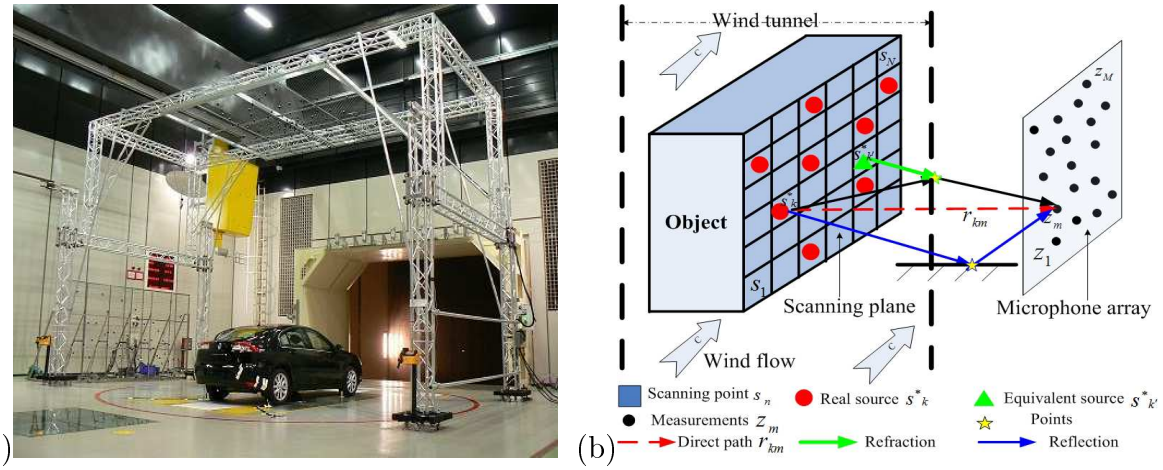

Illustration of simulations (a) Wind tunnel S2A [86] (b) Illustration of simulations and experiments.

on the vertical plane. $d = 2$m is the averaged size of sensor array. $D = 4.50$m is the distance between the sensor plane and source plane. $c_0 \approx 340$m/s is the acoustic speed in the common air. $T = 10000$ is the total number of samplings. For the simulated sources in Fig.5.4(a), there are simulated 4 monopoles and 5 complex sources, spaced at least 20cm from each other. Original source powers $\mathbf{x}^*$ are within [0.08,2] ([-10.3,3.7]dB) and 14dB dynamic range. The noise power is set $\sigma_e^2 = 0.86$ (-0.7dB), thus the averaged SNR is 0dB. To simulate a sparse distribution of discrete source signals, $\Delta p = 5$cm is used to discretize the 100×150 cm^2 source plane, so that the power image is of 21×31 pixels. Since total grid number $N = 651$ is much more larger than the maximal original source number $K = 23$, the discrete source signals $\mathbf{s}$ and their source powers $\mathbf{x}$ are both K-sparsity signals.

5.3.1 Simulation results

In Fig.5.4(h), proposed SC-RDAMAS approach not only detects most of the complex sources, but also well reconstructs source powers and positions in poor SNR situation. According to the adaptive estimation procedure in Algorithm 1, $\hat{K} = 25$ is better initialized owing to $\mu \approx 0.02$ from Eq.(5.12). According to the three quantitative criteria $\overline{\Delta x^*}$, $\overline{\Delta p^*}$ and δ_2 in the Table 5.1, 5.2 and 5.3 respectively, the proposed approach works much better than the others:

- Compared with the low spatial resolutions of the CBF ($\delta_p^{CBF} = 31$cm at 2500Hz), proposed approach makes good use of sparse distribution of source powers and it achieves the resolution as high as 5cm in both

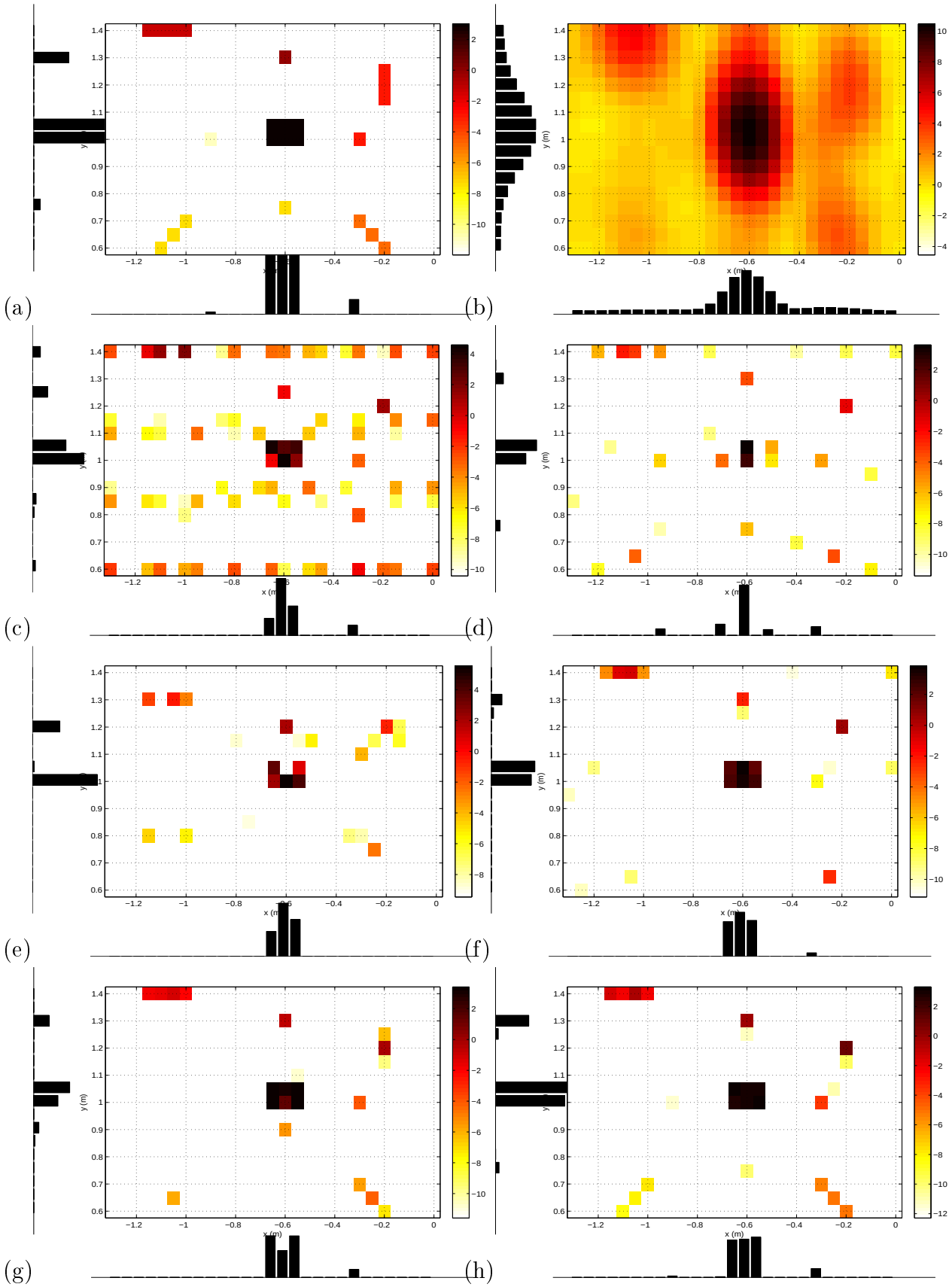

Figure 5.4: Simulation on monopole sources with 14dB power dynamic range at 2500Hz, $\sigma_e^2 = 0.86$, SNR=0dB and 15dB display: (a) Monopole sources (b) CBF (c) DAMAS with 5000 iterations (5000i) (d) CLEAN (e) SC-DAMAS (f) DR-DAMAS (5000i) (g) CMF and (h) Proposed SC-RDAMAS

horizontal and vertical directions. Since the horizontal aperture of NUA array is larger than the vertical, all the results of classical methods obtain better horizontal resolution. But proposed approach still achieves the high resolution in vertical direction.

- Compared with the sensitiveness of deconvolution methods, proposed approach greatly improves the robustness by jointly estimating the background noise as well as the source powers.
- Compared with sparse regularization methods, the sparsity parameter on the total source power is adaptively estimated. So that fits well for the strong i.i.d AGWN noise.
- To make a fair comparison, we realize the CMF, SC-DAMAS and proposed SC-RDAMAS based on Matlab toolbox SeDuMi [109].

In Fig.6.3, we show the relative error of power image reconstruction δ_2 of mentioned methods within SNR [-6, 18]dB at 2500Hz. Proposed approach is more robust to background noises than other classical methods.

In Fig.6.4, we show reconstruction errors δ_2 versus different frequency bins within [1600, 2600]Hz which affects the acoustic comfort of human being. The SNR is set 3dB. At high frequencies, proposed approach provides the most significant improvements. At low frequencies, proposed approach still maintains small reconstruction errors.

Table 5.1: Power estimations of 4 monopole sources by average power estimation error $\overline{\Delta x^*}$, relative errors of power image reconstruction δ_2 and estimated noise power $\hat{\sigma_e^2}$ at 2500Hz, SNR=0dB, simulated $\sigma_e^2 = 0.86$; '-' means unavailable.

Source power	0.08	0.18	0.98	0.50	$\overline{\Delta x^*}$	δ_2	$\hat{\sigma^2}_e$
CBF	1.57	11.28	3.51	2.02	69.64	121.9	-
DAMAS	-	-	-	0.44	3.14	1.33	-
CLEAN	-	0.25	0.44	0.28	0.87	0.67	-
SC-DAMAS	-	-	-	-	1.03	0.58	-
DR-DAMAS	-	-	0.77	0.23	0.30	0.08	-
CMF	0.09	-	0.80	0.40	0.31	0.10	0.89
SC-RDAMAS	0.09	0.10	1.05	0.43	0.06	0.06	0.85

5.3.2 Over-winnowing effects on simulations

Table 5.2: Position estimations of 4 monopole sources by averaged position errors $\overline{\Delta p^*}$ at 2500Hz, SNR=0dB; '-' means unavailable.

Position	(-0.9,1)	(-0.6,0.75)	(-0.3,1)	(-0.6,1.3)	$\overline{\Delta p^*}$
CBF	-	-	-	-	1
DAMAS	-	(-0.6,0.85)	(-0.3,1)	(-0.6,1.25)	0.28
CLEAN	(-0.95,1)	(-0.6,0.75)	(-0.3,1)	(-0.6,1.3)	0.01
SC-DAMAS	-	-	-	(-0.6,1.2)	0.63
DR-DAMAS	-	-	(-0.3,1)	(-0.6,1.3)	0.43
CMF	-	(-0.6,0.9)	(-0.3,1)	(-0.6,1.3)	0.29
SC-RDAMAS	(-0.9,1)	(-0.6,0.75)	(-0.3,1)	(-0.6,1.3)	0

Table 5.3: Power estimations of the complex monopole source on the center of image by power estimation error $\overline{\Delta x^*}$ at 2500Hz, SNR=0dB; '-' means unavailable.

Source power	2.00	2.00	2.00	2.00	2.00	2.00	$\overline{\Delta x^*}$
CBF	2.64	9.60	9.70	9.64	11.34	9.77	6.78
DAMAS	4.50	1.25	0.48	2.54	0.49	1.88	1.15
CLEAN	2.29	0.37	1.69	-	0.27	0.34	1.27
SC-DAMAS	1.68	2.49	1.16	0.10	2.23	0.65	0.75
CMF	1.36	2.86	2.07	2.09	1.92	1.05	0.45
DR-DAMAS	2.15	2.05	1.82	1.83	2.50	1.45	0.27
SC-RDAMAS	1.83	2.00	2.05	1.72	2.16	1.95	0.12

One of the common limitations in the SC-DAMAS, CMF and proposed SC-RDAMAS, is the over-winnowing effect which is well known in the compressed sensing[42, 20, 10]: the sparse results are often composed of discontinuous and unstructured (shapeless) points rather than continuous source distributions. Though these shapeless points could represent the monopole source power to some extent, they could hardly present distributed sources. Taking Fig.5.4(e) and (g) for example, the SC-DAMAS and CMF can only obtain discontinuous monopoles for complex sources. In Fig.5.4(h), though proposed SC-RDAMAS well detects most of the sources, it could barely discover one complex source whose continuous pattern is a vertical line. Moreover, the artifacts on the three figures are always monopole points which are near to sources.

The first reason for the above over-winnowing effects is the sparsity parameter influence. In Fig.5.7, we compare the sparsity parameter influence on δ_2 error between the SC-DAMAS and proposed SC-RDAMAS at 0dB and 2500Hz. The sparsity constraint is interpreted by the total source power β

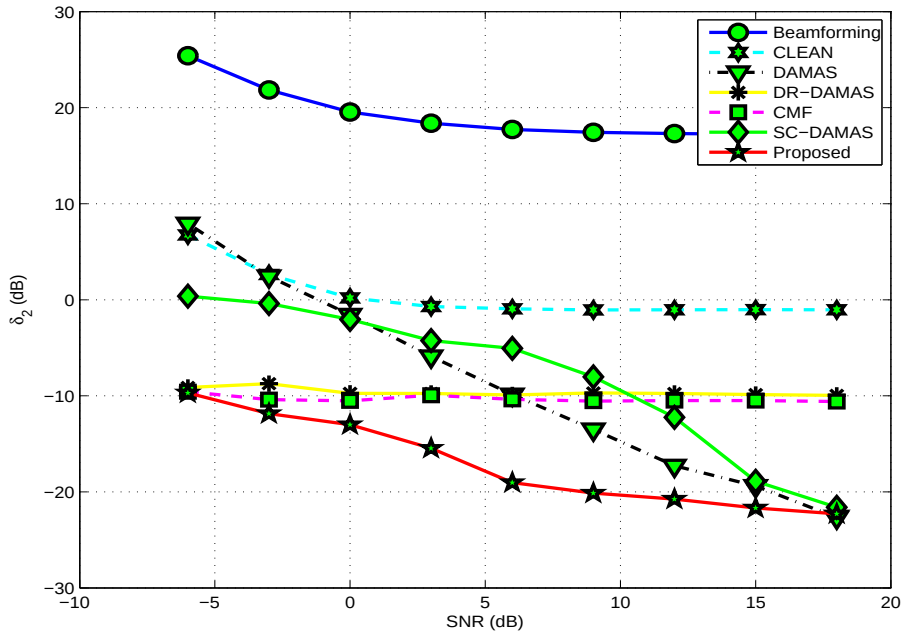

Figure 5.5: Performance comparison for relative errors of power image reconstruction δ_2 versus SNR [-6,18]dB on simulations at 2500Hz.

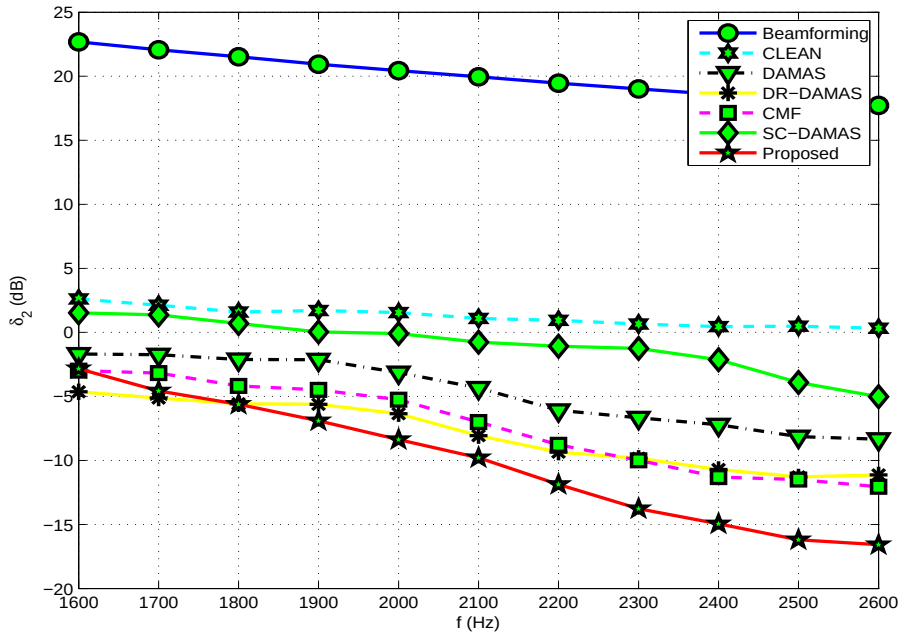

Figure 5.6: Performance comparison for relative errors of power image reconstruction δ_2 versus [1600, 2600]Hz on simulations at SNR=3dB.

so as to regularize data fitting errors in Eq.(4.59) and Eq.(5.6) respectively. And β in Eq.(5.10) mainly depends on the source number K estimation, when background noise is supposed to be i.i.d AGWN. Therefore, to evaluate the sparsity parameter is equivalent to evaluate the source number estimation.

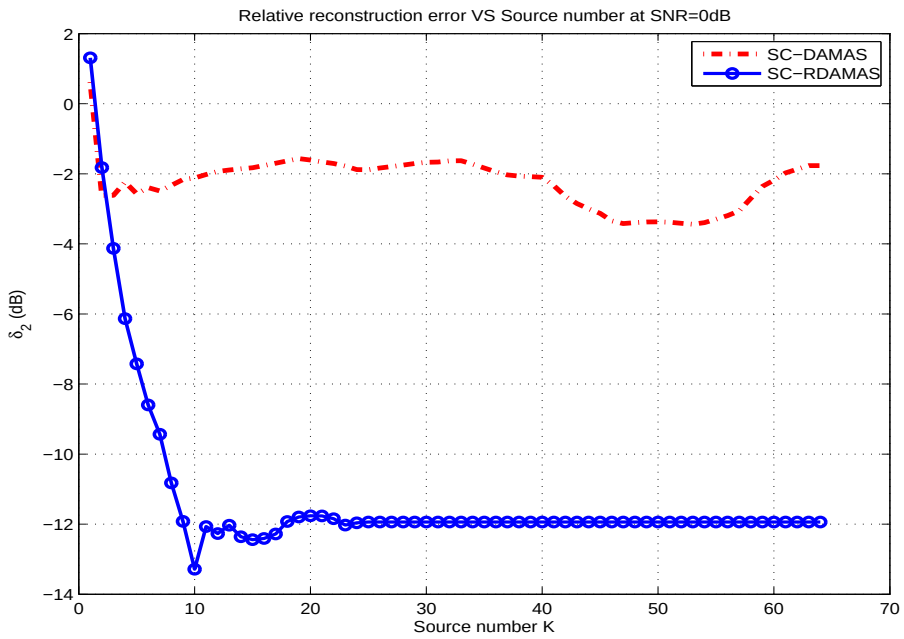

Figure 5.7: Sparsity parameter influence caused by estimated source number $\hat{K}$ versus power image reconstruction error δ_2 between the SC-DAMAS and proposed SC-RDAMAS at 0dB and 2500Hz.

Since complex sources in Fig.5.4(a) are supposed to be made of uncorrelated monopoles, it is reasonable to take $K \in [9, 23]$. In Fig.5.7, when $\hat{K} < 9$ is under-estimated, it is clear to see that δ_2 error of the two methods are both very sensitive to $\hat{K}$, so that neither of them could obtain a good reconstruction. When $9 \leq \hat{K} \leq 23$, proposed approach can obtain smaller δ_2 errors, but both of the two methods could hardly achieve stable results. When $\hat{K} > 23$ is over-estimated, proposed SC-RDAMAS steadily keeps much smaller δ_2 than the SC-DAMAS. Above all, proposed SC-RDAMAS can well initialize source number K and adaptively estimate sparsity parameter β compared with the SC-DAMAS.

Secondly, the sparsity constraint on total source power could not appropriately model the source sparse distributions nor source structures. Suppose two different source power distributions: $\mathbf{x}_1 = [1, 2, 3, 0, 0, 0, 0, 0, 0]^T$ and $\mathbf{x}_2 = [3, 2, 1, 0, 0, 0, 0, 0, 0]$ which have the same total source power $\beta = 6$. To reconstruct $\mathbf{x}_1$ and $\mathbf{x}_2$ from their beamforming data $\mathbf{y}_1$ and $\mathbf{y}_2$ respectively in the case of very strong background noises, it is highly necessary to investigate other sparsity prior models [10, 29, 61] instead of only using sparsity constraint $\beta = 6$. Therefore, in Fig.5.4(h), our approach can hardly detect all the complex sources with different sparse distributions.

The last but not least, the monopole assumption on acoustic model is too simple to model the complex sources with different patterns. Therefore, in

Fig.5.4(h), our approach tends to generate unexpected monopole artifacts near to sources, especially when the SNR is as small as 0dB. For the real data in Section 5.4, this drawback will become the dominant reason, and more obvious over-winnowing effects will be seen.

5.4 Real data of wind tunnel experiments

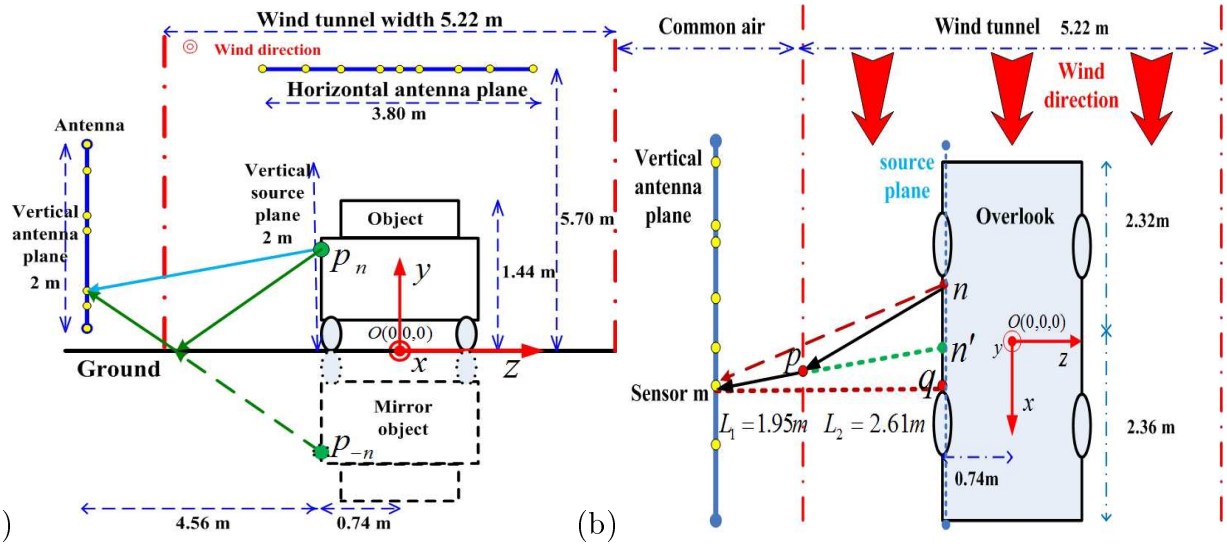

Figure 5.8: Configurations of wind tunnel experiment: (a) Frontlook and ground reflection (b) Overlook and wind refraction.

All the experiment configurations are the same as discussed in Chapter 4.8. But in real data, we do not know the exact source number or SNR beforehand. Therefore, it is necessary to investigate the adaptive estimation procedure in Algorithm 1 for robust acoustic imaging in wind tunnel tests.

The above figure shows the configurations of the wind tunnel S2A [86], object vehicle, NUA array and wind refraction. We suppose that all acoustic sources locate on the same plane. This assumption is almost satisfied, because the curvature of the car side is relatively small compared to the distance D=4.5m between the car and array plane. Since the scanning step is set by $\Delta p = 5cm$, the source plane of car side is of 1.5× 5 m^2 (31×101 pixels), and we also focus on a small region of the rear-view mirror: 1×1.5 m^2 (21×31 pixels). On the real data, there are T=524288 samplings with the sampling frequency f_s=2.56×10^4 Hz. We separate these samplings into I=204 blocks with L=2560 samplings in each bloc. The working frequency band is chosen as [2400,2600]Hz, which is sensitive to human being. The image results are shown by normalized dB images with 10dB span.

For the actual propagation time $\tau_{n,m}$ and distance $r_{n,m}$ in Eq.(3.11), we apply equivalent source to make refraction correction. For $\tau_{-n,m}$ and $r_{-n,m}$ in Eq.(3.11), we use the mirror source signal s_{-n} to correct the ground reflection. The details of the propagation corrections are discussed in A and B.

5.4.1 Results of single frequency data

In Fig.5.9(f), proposed SC-RDAMAS not only manages to distinguish the strong sources around the two wheels, rear-view mirror and side window, but also successfully reconstructs the week ones on the front cover and light. In fact, the proposed adaptive estimation procedure in Algorithm 1 inevitably increases more computational cost than the deconvolution methods such as DAMAS and CLEAN. But our approach still remains a moderate complexity compared with sparse regularization methods such as the SC-DAMAS as shown in Table 5.4. Due to the high dimension of variables in source power cross-spectrum matrix, we can not realize the original CMF method on real data.

Table 5.4: Computational cost for treating whole car: image 30×100 pixels, at 2500Hz, based on CPU:3.33GHz, '-' means unavailable.

Methods	CBF	DAMAS (5000i)	DR-DAMAS (5000i)	CLEAN	Proposed	SC-DAMAS	CMF
Time (s)	1	10	11	45	852	1254	Very Long

Based on the acoustic imaging on the car side, we investigate a small part of the rear-view mirror. In Fig.5.10(a), the CBF detects strong sources on the corner of the front wheel and rear-view mirror. The DAMAS in Fig.5.10(b) improves the spatial resolutions, but it causes some unexpected spots. In Fig.5.10(c), the DR-DAMAS eliminates most of false spots. In Fig.5.10(e)-(g), the CMF, SC-DAMAS and proposed approach achieve much better resolutions and offer more details of source power distributions on the rear-view mirror.

5.4.2 Results of wide-band data

In wind tunnel tests, acoustic sources are usually generated by wind frictions against the car surface. Different car parts produce different characteristic frequencies. Therefore, acoustic signals have the wide frequency band. In Chapter 3, we have taken DFT transformation and separated the wide-band

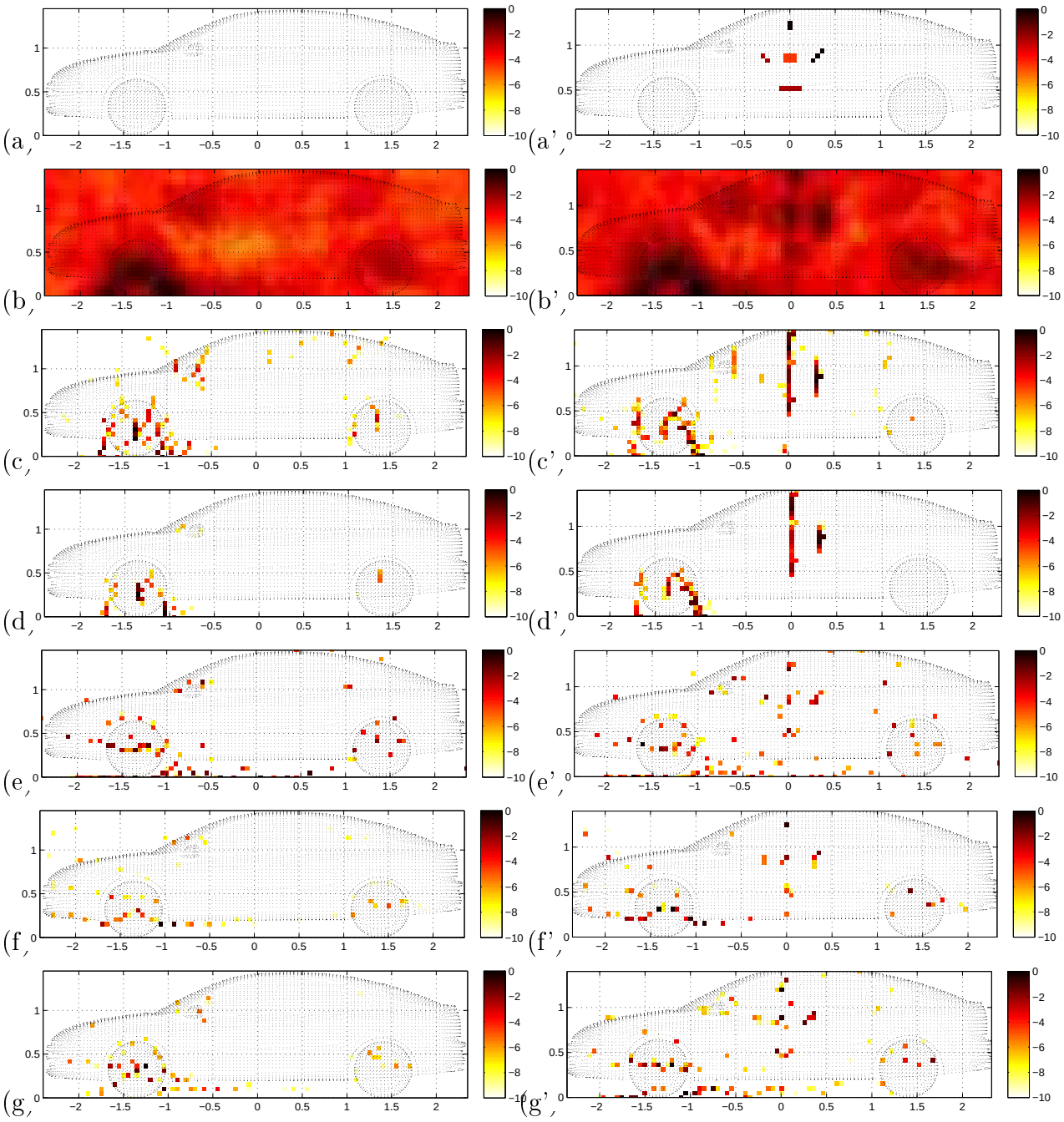

Figure 5.9: Acoustic imaging on the vehicle side at 2500Hz.
Left: Real data (a) Vehicle surface (b) Beamforming (c) DAMAS (5000i) (d) DR-DAMAS (5000i) (e) CLEAN (f) SC-DAMAS and (g) Proposed approach.
Right: hybrid data (a') Simulated sources (b') Beamforming (c') DAMAS (5000i) (d') DR-DAMAS (5000i) (e') CLEAN (f') SC-DAMAS and (g') Proposed SC-RDAMAS.

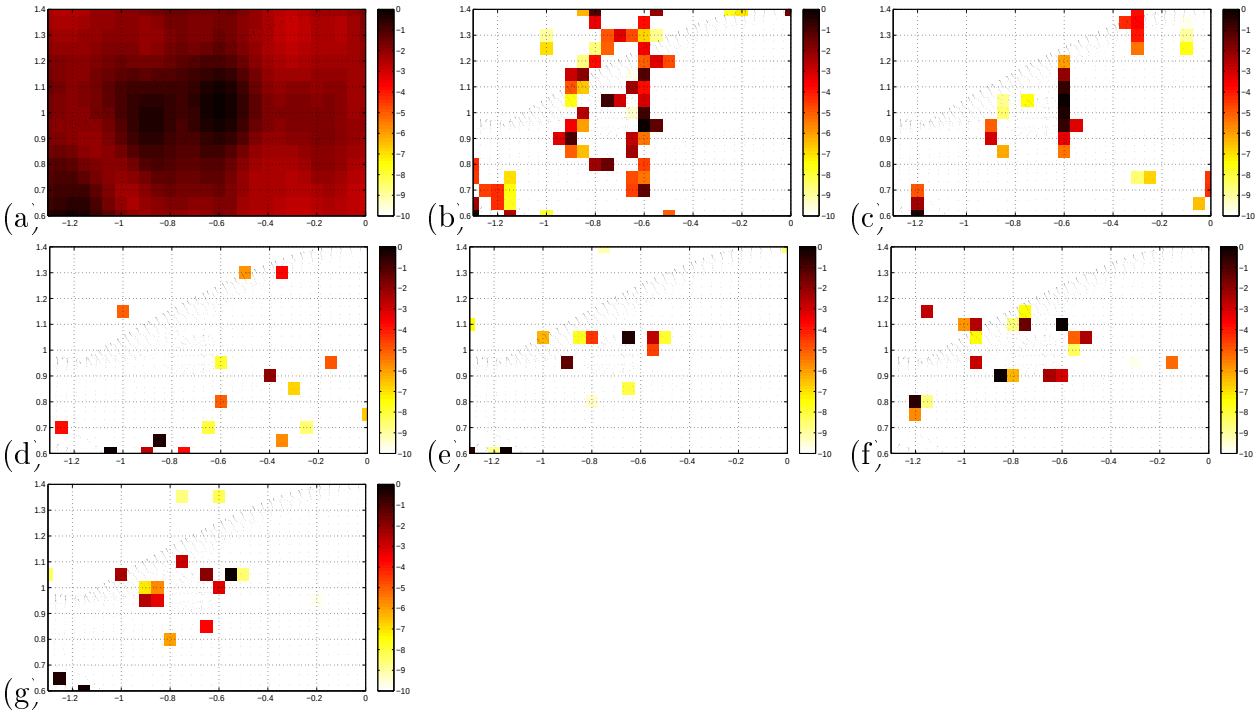

Figure 5.10: Acoustic imaging of rear-view mirror at 2500Hz: (a) CBF (b) DAMAS (5000i) (c) DR-DAMAS (5000i) (d) CLEAN (e) CMF and (f) SC-DAMAS and (g) Proposed SC-RDAMAS

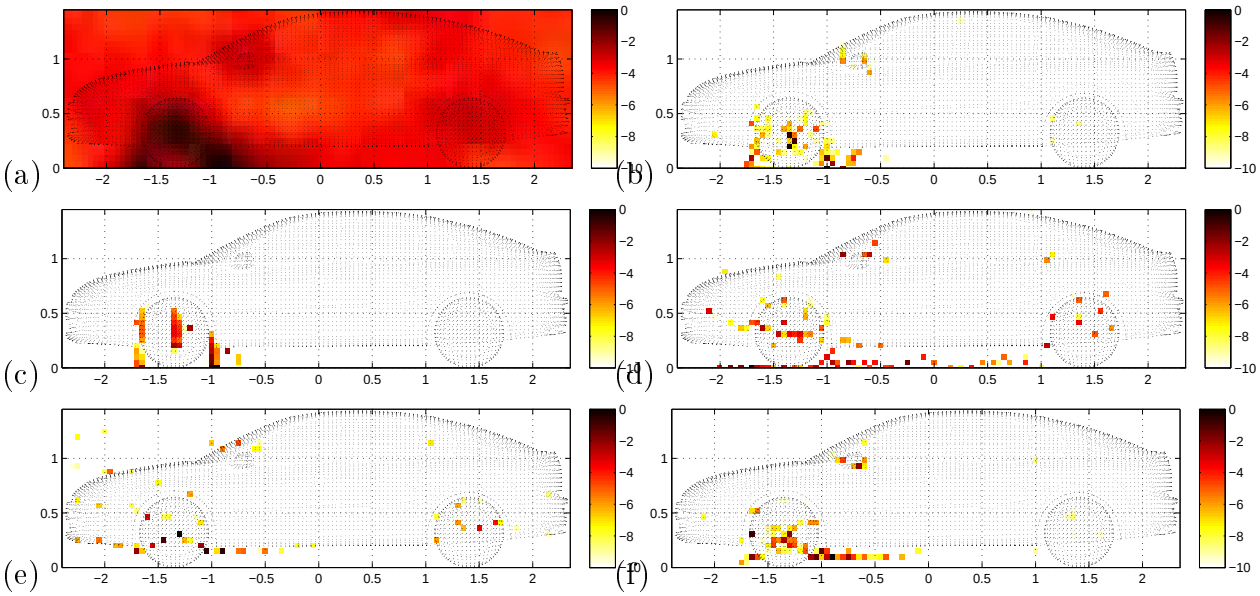

Figure 5.11: Wide-band data over [2400,2600]Hz: (a) CBF (b) DAMAS (c)DR-DAMAS (d) CLEAN (e) SC-DAMAS and (f) Proposed SC-RDAMAS

into L independent frequency bins, then we have engaged signal processing in each frequency bin. Using the proposed SC-RDAMAS approach in Eq.(5.6), we can obtain $\hat{\mathbf{x}}(f_l)$ as the estimation of source power $\mathbf{x}(f_l)$ at lth frequency bin. Finally, total power $\mathbf{x}_{wb}$ over the wide-band $[f_{min}, f_{max}]$ can be estimated

by averaging the summation of estimated result in each frequency bin as:

$$\hat{\mathbf{x}}_{wb} = \frac{1}{L} \sum_{f_l=f_{min}}^{f_{max}} \hat{\mathbf{x}}(f_l) \tag{5.13}$$

where in real data, we take $f_{min} = 2400$ and $f_{max} = 2600$.

Based on the imaging results at single frequency, we show performance comparisons of wide-band data within $[2400, 2600]$Hz which affects the acoustic comfort. In Fig.5.11, each method obtains a clearer result than the correspondent one at 2500Hz in Fig.5.9. This is because that source powers are enforced, but flashing false targets are suppressed over the wide-band average. The reconstruction of DAMAS in Fig.5.11(a) is reasonable, but its spatial resolution is not high enough on the front wheel and rear-view mirror. Figure.5.11(b) shows that the CLEAN greatly ameliorates the resolution, but unexpected points under the car caused by the ground reflection should be further eliminated. The SC-DAMAS in Fig.5.11(c) has the advantages of the CLEAN, but it could not detect the weak sources around the back wheel due to the sparsity parameter selection. Finally in Fig.5.11(d), the proposed approach provides the more acceptable reconstructions of source positions and powers for the strong sources on the mirror and the front wheel, as well as weak ones on the back wheel.

5.4.3 Over-winnowing effects on real data

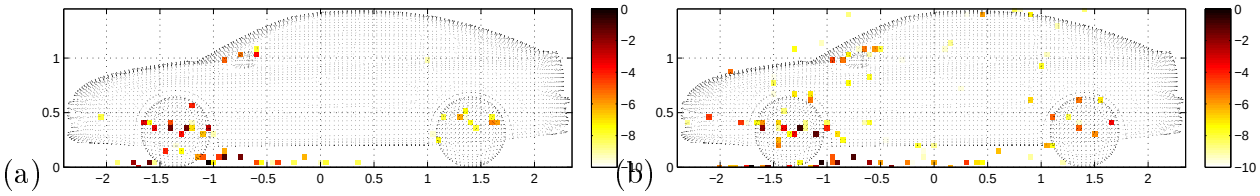

Figure 5.12: Sparsity parameter influence of proposed SC-RDAMAS on real data at 2500Hz: (a) Under-estimated $\hat{\beta}$ (b) Over-estimated $\hat{\beta}$.

The over-winnowing effects caused by the sparsity constraint in Fig.5.9(e) and (f) are more obvious than the simulations in Fig.5.4(e)(g)(h). This phenomenon on the real data could be explained by the following facts:

Source model problem. Some of acoustic sources on the car surface (wheels and rear-view mirrors) might be distributed sources which have structures and patterns. Thus the monopole source model used in this article could not

fit any more. Since the sparsity constraint (ℓ_1 norm) on total source power enforces the monopole reconstructions and neglects source structures to some extent, the mentioned sparsity methods prefer to offer the discontinuous point results, as typically shown on the back wheel in Fig.5.9(e) and (f).

Sparsity parameter problem. In proposed approach, the estimated sparsity parameter $\hat{\beta}$ in Eq.(5.10) depends on the source number and background noise power. On real data, however, it is hard to estimate source number K on the car surface. Moreover, background noises in the wind tunnel are not always i.i.d AGWN noises. Consequently, it is not easy to exactly derive $\hat{\beta}$ from Eq.(5.10). In Fig.5.12(a), if $\hat{\beta}$ is under-estimated, some weak sources would be inevitably lost on the headlight, top antenna and side window. And source patterns on the rear-view mirror and back wheel would be roughly expressed by several discontinuous points. If $\hat{\beta}$ is over-estimated, both weak sources and source patterns could be better detected, but the artifacts would be produced in the air and under the car body. Similarly over-winnowing effects can also be seen in Fig.5.10(d) (e)and (f) and Fig.5.11(c) and (d). But according to the adaptive sparsity parameter estimation procedure in Algorithm 1, proposed approach try to avoid under-estimate β.

5.5 Results with hybrid data

Even though proposed SC-RDAMAS approach obtain good performance on real data from wind tunnel experiments, it is not sufficient for performance validation. This is because the exact acoustic source distributions on the vehicle caused by wind flow are not known beforehand. To further verify the proposed method, we use the hybrid data which composes of known synthetic sources and the real data. The synthetic sources are generated in the same way as discussed in Chapter 4.9.

For the synthetic sources, figure 5.9(f') shows that proposed approach successfully detects most of the source powers and patterns. For the original sources in hybrid data, the proposed approach better discovers both strong and weak sources on two wheels and rear-view mirrors, as well as obtains a better noise suppression compared with mentioned methods in Fig.5.9(b')-(e').

In brief, the experiment results well agree with the simulations. The proposed SC-RDAMAS approach can achieve as good performance as the mentioned classical methods.

5.6 Conclusions and perspectives

In this chapter, we propose a robust super resolution approach with sparsity constraint for the acoustic imaging on the vehicle surface in wind tunnel experiments. The main contributions are:

For the robustness to background noises, we improve the forward model of power propagation by considering the noises at the sensors, as well as the propagation uncertainty caused by wind fraction and ground reflection in wind tunnel. The latter one is often ignored by classical methods. For the super spatial resolution, we adaptively estimate the sparsity parameter on source powers in the proposed Algorithm 1.

For the approach validation, we present performance comparisons with classical methods. The simulations show that proposed approach obtained the $5cm$ super resolution compared with the beamforming resolution 31cm at 2500Hz. It achieves 15dB dynamic range of power estimations, and well detects complex sources with different patterns. The real data results demonstrate that proposed approach effectively reconstructed strong sources on front wheels and rear-view mirrors, as well as the weak sources on back wheels. The hybrid data experiments furthermore confirm the effectiveness for reconstructing the known synthetic sources and original sources in the real data. We also show the moderate computational cost of our approach for the acoustic imaging in wind tunnel tests.

In Table 5.5, we give a brief summary for the advantages and drawbacks of mentioned classic methods and proposed SC-RDAMAS approach .

Table 5.5: General performance of classical methods and proposed SC-RDAMAS approach.

Methods	CBF	MUSIC	CLEAN	DAMAS	DR-DAMAS	SC-DAMAS	CMF	SC-RDAMAS
Resolutions	Low	High	Normal	Normal	Normal	High	Higher	Higher
Dynamic Range	Narrow	Normal	Normal	Normal	Normal	Normal	Wide	Wide
Noise	Robust	Sensitive	Sensitive	Sensitive	Normal	Sensitive	Robust	Robust
Computation	Least	low	Normal	Normal	Normal	High	Higher	High
Samples	Normal	More	Normal	Normal	Normal	Normal	More	Normal
Source number	No	Required	Required	No	No	Required	Required	Required

The main drawback of proposed approach is the over-winnowing effect existed in the compressed sensing methods. Due to the same sparsity con-

straint, it sometimes just obtains many unstructured or shapeless points and could not reconstruct the true source distribution, especially when acoustic sources could not be modeled by monopoles. To overcome this limitation, Chapter 6 investigates a Bayesian inference with a sparse prior [33, 29] which enforces the sparsity of source power distributions.

.

- L'homme qui sait réféchir est celui qui a la force illimitée.

Honoré de Balzac

6

Bayesian Approach with a Sparsity Enforcing Prior

As discussed in Chapter 4, the classical methods such as spatial filters (Beamforming, Capon, MUSIC), deconvolution and regularization methods suffer one of such limitations as: poor spatial resolution, sensitivity to background noises, narrow dynamic range, high computational cost or parameter dependence. In Chapter 5, we proposed a super resolution method with sparsity constraint (SC-RDAMAS), which overcomes most of the mentioned limitations to some extend. However, the sparsity parameter should be estimated carefully for good use. And the over-winnowing effect caused by sparsity constraint cannot always be avoided.

In this chapter, we are going to propose a Bayesian inference approach using a sparsity enforcing prior [29, 32, 31, 33]. So that we can get the robust acoustic imaging result in strong noise interference, and obtain high resolution and wide dynamic range of estimated powers. This sparse prior can better embody the sparsity characteristic of source distribution than the sparsity constraint. Furthermore, instead of empirically selecting important parameters, proposed Bayesian approach can automatically reconstruct source positions and powers, and alternatively estimates hyperparameters such as background noise power, the uncertainty of forward power model and other parameters in prior model. At the first step, we propose to use the Joint Maximum A Posterior (JMAP) estimation to jointly estimate all the unknown variables and parameters. However, this JMAP needs a non-

quadratic optimization and causes large computation cost.

This chapter is organized as follows: proposed Bayesian inference approach is presented in Section 6. On simulations, Section 6.4 demonstrates performance comparisons of the proposed approach with state-of-the-art methods for the monopole and complex source imaging. Results of the real data in wind tunnel experiments are illustrated in Section 6.5. In order to further prove the effectiveness of proposed Bayesian approach, Section 6.6 demonstrates its performances on hybrid data, in which some known synthetic sources are added to the real data. Finally the conclusions and perspectives are summarized in Section 6.7.

6.1 Bayesian inference methods

Bayesian inference methods [113, 115, 84, 87, 5, 24, 130, 29] have been recently applied in acoustic imaging and successfully overcome the drawbacks of deconvolution and regularization methods. Bayesian inference is a powerful methodology for solving ill-posed inverse problem. It aims to estimate the unknown random variables by applying the Bayes' rule to update the probability law: a posterior probability is obtained based on both the likelihood and prior models, in which, the previous one can be derived from the known data and forward model, while the latter can be adaptively imposed on the unknown random variables. Comparing to the classical determined methods, there are at least three advantages:

- The prior model acts like the regularization term, but work much better than it. This is because the prior model can appropriately translate the physical and statistical characteristics of unknown random variables into a concrete mathematical model, so that prior models can, not only bring in novel information into the under-determined forward model, but also can greatly reduce the solution uncertainty of ill-posed inverse problem.

- Important parameters of the deconvolution or regularization methods can be seen as the unknown quantities and estimated in the same way. Some of unknown quantities such as forward power model uncertainty caused by multi-path propagation, and the background noises at the sensor array can be modeled by proper prior models according to their statistical properties, physical nature, and even our presumptions.

- All unknown random variables can be estimated by well developed optimization methods such as the Maximum Likelihood (ML), joint Maximum A Posterior (MAP), Variational Bayesian Approximation (VBA) which are the analytical methods, as well as the Gibbs Sampling and sequential Monte Carlo methods which are the numerical methods. Not only the mean values of unknown random variables can be estimated by Bayesian inference, but also the error ranges (variance) of these variables can be also obtained in the same time.

- However, one of the biggest limitations of Bayesian inference methods is the tremendous computational burden to get a global or sub-global optimization results. And some of optimizations cannot be implemented for practical use. So that it is a worthy work to balance the estimation performance and the calculating time.

Starting with our improved forward model of acoustic power propagation in Eq.(5.4):

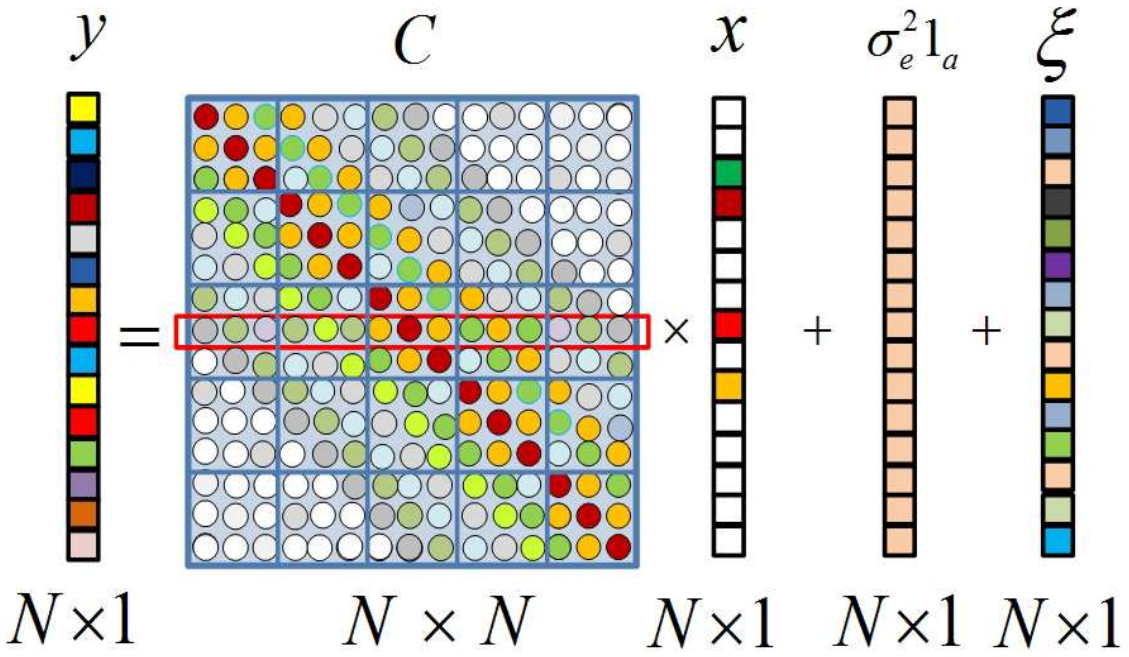

Illustration of sparse power equation with power model uncertainty $\boldsymbol{\xi}$ in Eq.(5.4).

$$\mathbf{y} = \mathbf{C}\,\mathbf{x} + \sigma_e^2\,\mathbf{1}_a + \boldsymbol{\xi}$$

we can assign the likelihood $p(\mathbf{y}|\mathbf{x})$ based on the basic information of power model uncertainty $\boldsymbol{\xi}$ ($\boldsymbol{\xi} \geq 0$). Generally, $\boldsymbol{\xi}$ is supposed to be Gaussian distribution as $\boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{\xi}|0, \sigma_\xi^2)$, where σ_ξ^2 represents the power model uncertainty.

Thus the likelihood $p(\mathbf{y}|\mathbf{x},\boldsymbol{\theta}_1)$ is obtained in narrow frequency band as follows:

$$p(\mathbf{y}|\mathbf{x},\boldsymbol{\theta}_1) = \frac{1}{(2\pi\sigma_\xi^2)^{N/2}} \exp\left[-\frac{\|\mathbf{y}-\mathbf{C}\mathbf{x}-\sigma_e^2\mathbf{1}_N\|^2}{2\sigma_\xi^2}\right] . \tag{6.1}$$

where unknown parameters $\boldsymbol{\theta}_1 = [\sigma_e^2, \sigma_\xi^2]^T$ denotes the hyperparameters of the above likelihood, and $\boldsymbol{\theta}_1$ are independent to source powers $\mathbf{x}$; and $p(\cdot)$ denotes the probability density function (PDF). In practice, $\mathbf{y}$ are known as the measured beamforming powers $\tilde{\mathbf{y}}$ from Eq.(4.10).

Let $\boldsymbol{\theta} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2]^T$ denote the hyperparameters to be estimated, with $\boldsymbol{\theta}_1$ being likelihood hyperparameters and $\boldsymbol{\theta}_2$ being the prior hyperparameters. And assume $\boldsymbol{\theta} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2]^T$ and source powers $\mathbf{x}$ are mutually independent to each other.

According to the assigned likelihood in Eq.(6.1), the Maximum Likelihood[40, 69] (ML) estimation is classically used as

$$\hat{\mathbf{x}}_{ML} = \arg\max_{\mathbf{x}} \{p(\mathbf{y}|\mathbf{x},\boldsymbol{\theta})\}$$

.

Now if we can assign a prior law $p(\mathbf{x}|\boldsymbol{\theta}_2)$ on $\mathbf{x}$, we can obtain

$$p(\mathbf{x}|\mathbf{y},\boldsymbol{\theta}) \propto p(\mathbf{y}|\mathbf{x},\boldsymbol{\theta}_1)\, p(\mathbf{x}|\boldsymbol{\theta}_2)$$

using the Bayes' rule $p(\mathbf{x}|\mathbf{y},\boldsymbol{\theta}) = p(\mathbf{y}|\mathbf{x},\boldsymbol{\theta}_1)\, p(\mathbf{x}|\boldsymbol{\theta}_2)/p(\mathbf{y})$ with $p(\mathbf{y})$ being a constant for the observed data $\mathbf{y}$.

Then if we apply the Maximum a Posterior (MAP) criterion, we can obtain

$$\hat{\mathbf{x}}_{MAP} = \arg\max_{(\mathbf{x})} \{p(\mathbf{x}|\mathbf{y},\boldsymbol{\theta})\} = \arg\min_{(\mathbf{x})} \{-\ln p(\mathbf{y}|\mathbf{x},\boldsymbol{\theta}_1) - \ln p(\mathbf{x}|\boldsymbol{\theta}_2)\} ,$$

where ln denotes Logarithm operator.

And then if $\boldsymbol{\theta}$ is unknown, we can also consider it as an extra unknown variable and try to estimate them by defining the joint posterior due to the full Bayes' rule:

$$\begin{cases} p(\mathbf{x},\boldsymbol{\theta}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x},\boldsymbol{\theta})\, p(\mathbf{x},\boldsymbol{\theta}) \\ \qquad\qquad = p(\mathbf{y}|\mathbf{x},\boldsymbol{\theta}_1)\, p(\mathbf{x}|\boldsymbol{\theta}_2)\, p(\boldsymbol{\theta}_2)\, p(\boldsymbol{\theta}_1) \end{cases} ,$$

where $\mathbf{x}$ and $\boldsymbol{\theta}$ are mutually independent. In this case, the proposed joint MAP (JMAP) criterion is thus given as follows:

$$\begin{cases} (\hat{\mathbf{x}}, \hat{\boldsymbol{\theta}})_{JMAP} = \arg\max\limits_{(\mathbf{x},\boldsymbol{\theta})} \{p(\mathbf{x}, \boldsymbol{\theta}|\mathbf{y})\} \\ \qquad\qquad = \arg\min\limits_{(\mathbf{x},\boldsymbol{\theta})} \{-\ln p(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta}_1) - \ln p(\mathbf{x}|\boldsymbol{\theta}_2) - \ln p(\boldsymbol{\theta}_2) - \ln p(\boldsymbol{\theta}_1)\} \end{cases}, \tag{6.2}$$

where $-\ln p(\mathbf{x}|\boldsymbol{\theta}_2)$ can be interpreted as the regularization form $\mathcal{F}(\cdot)$ in Eq.(4.58), and hyperparameter $\boldsymbol{\theta}_2$ takes the similar effect as the regularization parameter does in Eq.(4.58). Indeed, one of the advantages of Bayesian inference approach is that hyperparameters $\boldsymbol{\theta} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2]^T$ can be jointly estimated by JMAP in Eq.(6.2).

6.2 Proposed sparsity enforcing prior on source power distribution

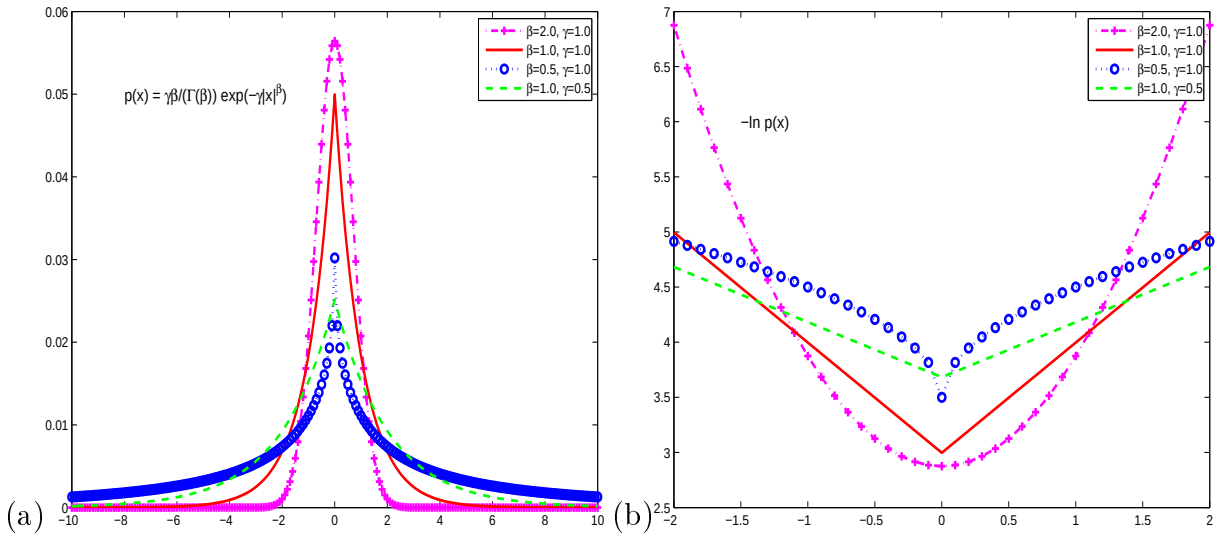

Figure 6.1: Generalized Gaussian family: (a) Probability density function $\mathcal{GG}(x_n)$ and (b) $-\ln[\mathcal{GG}(x_n)]$ function.

In Eq.(3.14), source signals $\mathbf{s}$ include just K non-zero signal $\mathbf{s}^*$ in the space domain, which is rather small with respect to $N - K$ zero values (N>>K). Therefore, the source powers $\mathbf{x}$ defined in Eq.(3.19) are also of the K-sparsity signals. For the PDF distribution of $\mathbf{x}$, most of the high probability values should concentrate around the original zero of $\mathbf{x}$. Meanwhile, $\mathbf{x}$ contains different items, among which, the difference between the maximal and minimal powers could be of 10dB dynamic range. This large dynamic range of powers

can be represented by a PDF distribution with a long heavy tail among the large values of $\mathbf{x}$. For the prior $p(\mathbf{x}|\boldsymbol{\theta}_2)$ on source powers $\mathbf{x}$, we reconsider the sparsity fact that acoustic sources sparsely lay on the surface of the object, and the source number K is relatively much smaller with respect to the total scanning points N on the source plane. Taking the vehicle in the wind tunnel test for example, most of sources mainly locate on the particular parts the rear-view mirrors and wheels, whereas for the rest parts, there are few sources existing. Such a sparse distribution can be represented by a centralized PDF function that has a very high value around the original zero (sparsity) and a long heavy tail (dynamic range of source powers). This kind of centralized PDF function can be selected among the Generalized Gaussian $\mathcal{GG}(x_n)$ family. The above sparse distribution can be defined from the Generalized Gaussian $\mathcal{GG}(x_n)$ family as discussed in [17, 87]. For the uncorrelated centralized $\mathbf{x}$, we take the Double Exponential $\mathcal{DE}(\mathbf{x}|\gamma,\beta)$ model as the sparsity enforcing prior as follows::

$$p(\mathbf{x}|\boldsymbol{\theta}_2)=\prod_{n=1}^{N}\mathcal{GG}(x_n|\gamma,\beta)=\left[\frac{\beta\gamma}{2\Gamma(1/\beta)}\right]^N\exp\left[-\gamma\sum_{n=1}^{N}|x_n|^\beta\right], \tag{6.3}$$

where $\boldsymbol{\theta}_2=[\gamma,\beta]^T$ and

$$\mathcal{GG}(x_n|\gamma,\beta)=\frac{\beta\gamma}{2\Gamma(1/\beta)}\exp\left[-\gamma|x_n|^\beta\right], \tag{6.4}$$

where $\Gamma(\cdot)$ denotes the Gamma function, and shape parameter β reflects the degree of sparsity, it controls the concentration of $p(\mathbf{x}|\boldsymbol{\theta}_2)$ at zero value. Thus the smaller β is, the sparser $p(\mathbf{x}|\boldsymbol{\theta}_2)$ becomes. And parameter γ reflects the inverse variance of x_n, it controls the tail of $p(\mathbf{x}|\boldsymbol{\theta}_2)$ and it affects the dynamic range of $\mathbf{x}$. Thus the smaller γ is, the longer and heavier the tail becomes.

In the case of $\beta=2$, we get Gaussian $\mathcal{N}(0,\frac{1}{2\gamma})$ model:

$$p(\mathbf{x}|\boldsymbol{\theta}_2)=\prod_{n=1}^{N}\mathcal{N}(x_n|0,\frac{1}{2\gamma})=\left(\frac{\gamma}{\pi}\right)^{N/2}\exp\left[-\gamma\|\mathbf{x}\|_2^2\right], \tag{6.5}$$

where $\boldsymbol{\theta}_2=\gamma$, and $\gamma=\frac{1}{2\sigma_x^2}$ reflects the inverse variance of $\mathbf{x}$, with σ_x^2 being the variance of $\mathbf{x}$.

When $\beta=1$, we get the Double Exponential $\mathcal{DE}(\mathbf{x}|\gamma,\beta)$ model from Generalized Gaussian $\mathcal{GG}(\mathbf{x})$ family as follows:

$$p(\mathbf{x}|\boldsymbol{\theta}_2)=\prod_{n=1}^{N}\mathcal{DE}(x_n|\gamma,\beta)=\left(\frac{\gamma}{2}\right)^N\exp\left[-\gamma\|\mathbf{x}\|_{\beta=1}\right], \tag{6.6}$$

where $\boldsymbol{\theta}_2 = \gamma$ is the scale parameter, which controls the dynamic range of $\mathbf{x}$, and $\|\mathbf{x}\|_{\beta=1}$ represents the l_1 norm, which can promote the sparsity of $\mathbf{x}$.

In Fig.6.1, four examples of $\mathcal{GG}(x_n)$ family and their $-\ln[\mathcal{GG}(x_n)]$ functions are illustrated. When $\gamma = 1$ is fixed, the smaller β is, the sparser $p(x_n)$ becomes. When $\beta = 1$ is set, the smaller γ is, the heavier tail of $p(x_n)$ becomes. The proper values of β and γ can balance the sparsity and dynamic range. For cases $0 < \beta < 1$, it is of great interest to enforce sparsity, but unfortunately, its $-\ln[\mathcal{GG}(x_n)]$ function is not convex. For the case $\beta = 1$ and proper γ, the Double Exponential $\mathcal{DE}(\mathbf{x}|\gamma, \beta)$ model can promote sparsity and obtain wide dynamic range, moreover, its $-\ln[\mathcal{GG}(x_n)]$ function is convex as well. Moreover, the symmetric PDF of sparsity distribution in Fig.6.4 could be explained that the negative source power distribution representing the mirror source powers due to the ground reflection effect in wind tunnel experiment.

In conclusion, we select the Double Exponential $\mathcal{DE}(\mathbf{x}|\gamma, \beta)$ model with $\beta = 1$ as the sparsity enforcing a prior. Here, we have to confine the non-negative constraint on source powers $\mathbf{x} \geq 0$, since the $\mathcal{DE}(\mathbf{x}|\gamma, \beta)$ model is the symmetry distribution.

For hyperparameters $\boldsymbol{\theta} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2]^T$, we have $\boldsymbol{\theta}_1 = [\sigma_e^2, \sigma_\xi^2]^T$, $\boldsymbol{\theta}_2 = \gamma$. Assuming the independence among hyperparameters, it yields $p(\boldsymbol{\theta}) = p(\sigma_e^2)\, p(\sigma_\xi^2)\, p(\gamma)$. Since these parameters are non-negative, we take Jeffreys priors [63] for simplicity as follows:

$$p(\sigma_e^2) \sim \frac{1}{\sigma_e^2}, \quad p(\sigma_\xi^2) \sim \frac{1}{\sigma_\xi^2}, \quad p(\gamma) \sim \frac{1}{\gamma}. \tag{6.7}$$

One of the advantages of Jeffreys priors is, taking $p(\gamma) \sim \frac{1}{\gamma}$ for instance, Jeffreys prior model mainly depends on γ, and it does not bring in other parameters which have to be selected carefully for good use.

6.3 Proposed Bayesian Joint Maximum A Posterior criterion

There are several classical methods [114, 49, 105, 4] to solve the joint posterior probability in Eq.(6.2). Here we choose the JMAP [4] estimation, since JMAP can build up a relationship with the classical regularization method in Eq.(4.58). Based on the selections of prior models in the above, we take the followings into the Eq.(6.2): the likelihood $p(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta}_1)$ in Eq.(6.1), the sparsity

enforcing prior $p(\mathbf{x}|\boldsymbol{\theta}_2)$ in Eq.(6.6), and Jeffreys priors $p(\boldsymbol{\theta})$ of parameters in Eq.(6.7). By omitting trivial value terms and considering the non-negative constraints of source powers and hyperparameters, our proposed Bayesian approach is obtained as follows:

$$\begin{cases} (\hat{\mathbf{x}}, \hat{\boldsymbol{\theta}}) = \arg\min\limits_{(\mathbf{x},\boldsymbol{\theta})} \{\mathcal{J}(\mathbf{x}, \boldsymbol{\theta})\} \\ \mathcal{J}(\mathbf{x}, \boldsymbol{\theta}) = \underbrace{\frac{1}{2\sigma_\xi^2}\|\mathbf{y} - \mathbf{C}\mathbf{x} - \sigma_e^2\,\mathbf{1}_a\|^2}_{Likelihood} + \underbrace{\gamma\,\|\mathbf{x}\|_1}_{Sparse\,prior} + \underbrace{\frac{N}{2}\ln\sigma_\xi^2 - N\,\ln\gamma}_{Hyperparameter\,prior}, \\ \quad \text{s.t.} \mathbf{x} \succeq 0, \sigma_e^2 \succeq 0, \sigma_\xi^2 \succeq 0, \gamma \succeq 0 \end{cases} \tag{6.8}$$

where hyperparameters $\boldsymbol{\theta} = [\sigma_e^2,\, \sigma_\xi^2,\, \gamma]^T$; $\mathbf{y}$ are known as the measured beamforming powers $\tilde{\mathbf{y}}$ from Eq.(4.10). In the cost function $\mathcal{J}(\mathbf{x}, \boldsymbol{\theta})$, the first term represents the data fitting. The second term represents the ℓ_1 sparse regularization, and γ performs as the regularization parameter. The rest terms are derived from the priors $p(\boldsymbol{\xi})$, as well as the sparse prior $p(\mathbf{x}|\gamma)$. Particularly, the JMAP estimation in Eq.(6.8) reveals that γ should not be neither too big nor too small, and its proper value can not only enforce the sparsity of source power distributions, but also can promote the dynamic range of estimated source powers. For hyperparameter estimations in Eq.(6.8), we use an alternative optimization procedure as: we first initialize $\mathbf{x}$, then at any iteration k+1, we obtain a new value for $\boldsymbol{\theta}$ by minimizing $\mathcal{J}(\hat{\mathbf{x}}^{(k)}, \boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$. At the next iteration we alternatively update $\mathbf{x}$ by minimizing $\mathcal{J}(\mathbf{x}, \hat{\boldsymbol{\theta}}^{(k)})$ with respect to $\mathbf{x}$. Finally we repeat this alternative optimization procedure until $\mathcal{J}(\mathbf{x}, \boldsymbol{\theta})$ converges.

We have to point out 4 advantages with respect to classical methods:

- Estimation of σ_ξ^2 and σ_e^2 improves the robustness to the power model uncertainty and noise interference.
- Applying ℓ_1 norm of $\|\mathbf{x}\|_1$ enforces sparsity of $\mathbf{x}$, and achieve a super spatial resolution of power distributions.
- Estimation of γ provides a wide dynamic range of source powers, moreover, γ takes the similar effect as the regularization parameter α does in Eq.(4.58).
- $\mathbf{x}$ and hyperparameters are alternatively estimated without knowing the source number or SNR.

6.3.1 Alternate optimization for hyperparameter estimations

For the JMAP criterion in Eq.(6.8), we alternatively estimate the source powers $\mathbf{x}$ and hyperparameters as follows:

$$\begin{cases} \hat{\boldsymbol{\theta}}^{(k)} = \arg\min_{(\boldsymbol{\theta})}\{\mathcal{J}(\hat{\mathbf{x}}^{(k)}, \boldsymbol{\theta})\} \\ \hat{\mathbf{x}}^{(k+1)} = \arg\min_{(\mathbf{x})}\{\mathcal{J}(\mathbf{x}, \hat{\boldsymbol{\theta}}^{(k)})\} \end{cases}, \tag{6.9}$$

where hyperparameters are $\boldsymbol{\theta} = [\sigma_e^2,\ \sigma_\xi^2,\ \gamma]^T$. The first iteration begins based on the simple initialization of $\mathbf{x} = 0$.

Moreover, we can facilitate the estimation of σ_ξ^2 as follows:

$$\widehat{\sigma_\xi^2}^{(k)} = \mathrm{tr}\,\{\hat{\mathbf{R}}_z\} - \|\hat{\mathbf{x}}^{(k-1)}\|_1 - M\widehat{\sigma^2}_e^{(k-1)}\,, \tag{6.10}$$

where $\hat{\mathbf{R}}_z$ is estimated from Eq.(4.6), and $\mathrm{tr}\,(\hat{\mathbf{R}}_z)$ represents the total power of measured signals. According to power conservation, the total measured power $\mathrm{tr}\,(\hat{\mathbf{R}}_z)$ consists of the source powers $\|\mathbf{x}\|_1$, total noise powers $M\sigma_e^2$ at M sensors, and the powers model uncertainty σ_ξ^2.

After hyperparameter estimation, $\hat{\mathbf{x}}^{(k)}$ are alternatively optimized by the steepest gradient algorithm based on the estimated hyperparameters $\hat{\boldsymbol{\theta}}^{(k)} = [\widehat{\sigma_e^2}^{(k)}, \widehat{\sigma^2}_\xi^{(k)}, \hat{\gamma}^{(k)}]^T$ as follows:

$$\hat{\mathbf{x}}^{(k)} = \hat{\mathbf{x}}^{(k-1)} + \mu\nabla J(\mathbf{x}, \hat{\boldsymbol{\theta}}^{(k)})\,, \tag{6.11}$$

where μ is the step size, which could be fixed as a small value, or optimally selected as discussed in paper [71]. Since $\mathbf{x} \geq 0$, $\|\mathbf{x}\|_1 = \sum_{n=1}^{N} x_n$, the gradient $\nabla J(\mathbf{x}, \hat{\boldsymbol{\theta}}^{(k)})$ is obtained as:

$$\nabla J(\mathbf{x}, \hat{\boldsymbol{\theta}}^{(k)}) = -\frac{1}{\widehat{\sigma^2}_\xi^{(k)}}\mathbf{C}^\dagger(\mathbf{y} - \mathbf{C}\hat{\mathbf{x}}^{(k-1)} - \widehat{\sigma_e^2}^{(k)}\mathbf{1}_N) + \hat{\gamma}^{(k)}\mathbf{1}_N\,. \tag{6.12}$$

When $\hat{\boldsymbol{\theta}}^{(k)} = [\widehat{\sigma_e^2}^{(k)}, \widehat{\sigma^2}_\xi^{(k)}, \hat{\gamma}^{(k)}]^T$ is fixed, the JMAP criterion in Eq. (6.8) is a convex minimization under linear matrix constraints. This optimization can also be solved by interior point methods using MATLAB toolbox SeMuDi [109].

6.3.2 Relation with regularization approach

Furthermore, comparing to the classical methods of Eq.(4.58), the regularization parameter can be expressed in function of hyperparameters in the proposed Bayesian approach of Eq.(6.8) as follows:

$$\alpha = 2\sigma_\xi^2 \gamma \,, \tag{6.13}$$

where we can see that α is proportional to σ_ξ^2 and γ. Particularly when Gaussian prior in Eq.(6.5) in 6.2 is used ($\gamma = \frac{1}{2\sigma_x^2}$) on source powers, $\alpha = \frac{\sigma_\xi^2}{\sigma_x^2}$ means the inverse of signal-to-uncertainty ratio. Since we can take the power model uncertainty as the background noises to some extent, α equals the inverse of SNR. Our conclusion agrees with the argument in paper [26]. Comparing to the classical regularization methods in Eq.(4.58), our proposed Bayesian approach has the advantage of automatically estimating the regularization parameter.

6.4 Simulations of monopole and complex sources

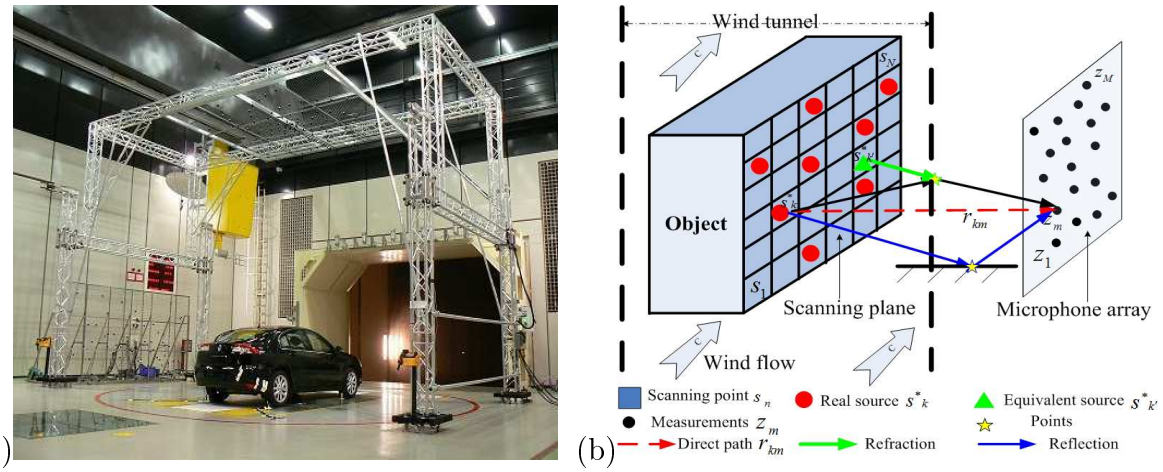

Illustration of simulations (a) Wind tunnel S2A [86] (b) Illustration of simulations and experiments.

This section shows the typical simulations on source power reconstruction and localization of monopole sources. To focus on method comparisons, we do not consider the ground reflection and wind refraction in simulations. But we should reconsider the multi-path propagation effects in wind tunnel experiments. The proposed Bayesian approach is compared with the CBF, DAMAS, CLEAN, CMF, SC-DAMAS and SC-RDAMAS methods.

The above figure shows the static vehicle (no engine noise), microphone sensor array and the wind flow at the speed of 160km/h in the wind tunnel

S2A [86]. One of objects of this wind tunnel is to detect acoustic powers and positions on the car surface. This wind tunnel can simulate a traveling car on the high-way and measure its acoustic comfort to the passengers-by.

The simulation configurations are also the same as the one in Chapter 4.7. In above figure: there are $M = 64$ non-uniform sensors locating on the vertical plane. $d = 2$m is the averaged size of sensor array. $D = 4.50$m is the distance between the sensor plane and source plane. $c_0 \approx 340$m/s is the acoustic speed in the common air. $T = 10000$ is the total number of samplings. For the simulated sources in Fig.6.2(a), there are simulated 4 monopoles and 5 complex sources, spaced at least 20cm from each other. Original source powers $\mathbf{x}^*$ are within [0.08,2] ([-10.3,3.7]dB) and 14dB dynamic range. The noise power is set $\sigma_e^2 = 0.86$ (-0.7dB), thus the averaged SNR is 0dB. To simulate a sparse distribution of discrete source signals, $\Delta p = 5$cm is used to discretize the 100×150 cm^2 source plane, so that the power image is of 21×31 pixels. Since total grid number $N = 651$ is much more larger than the maximal original source number $K = 23$, the discrete source signals $\mathbf{s}$ and their source powers $\mathbf{x}$ are both K-sparsity signals.

In this part, we first show the reconstruction results of source power images. Then we make quantitative comparisons of different methods. Finally we further validate our proposed Bayesian approach by simulations in the cases of various noise levels and different frequency bins.

Figure 6.2 shows the reconstruction results of simulated data at f=2500Hz and SNR=0dB. In Fig.6.2(b), the beamforming [25] method merely gives a very blurred result of strong sources due to its low spatial resolution ($\Delta B \approx 31$cm at 2500Hz). In Fig.6.2(c) and (d), both the DAMAS [18] with 5000 iterations (5000i) and CLEAN [123] could not provide reliable reconstructions in strong background noises. This can be explained by the fact that DAMAS omits the background noises as discussed in Eq.(4.51). As for CLEAN, it iteratively subtracts the maximal source power from the original beamforming result in Fig.6.2(b), but CLEAN might not easily detect the weak sources whose powers are lower than the noise, moreover, some important parameters such as the attenuation factor and iteration number have to be selected carefully by CLEAN. In Fig.6.2(e), DR-DAMAS [18] removes the noise influence, and roughly estimates the complex source, but it loses some of weak sources, this is because the diagonal removal operation distracts both the noise and weak sources. The CMF [128] aims to reconstruct both the spectrum cross-spectrum matrix $\mathbf{R}_z$ and the noise power in Eq.(4.5) under the sparsity constraint. In Fig.6.2(f), CMF well estimates the noise power

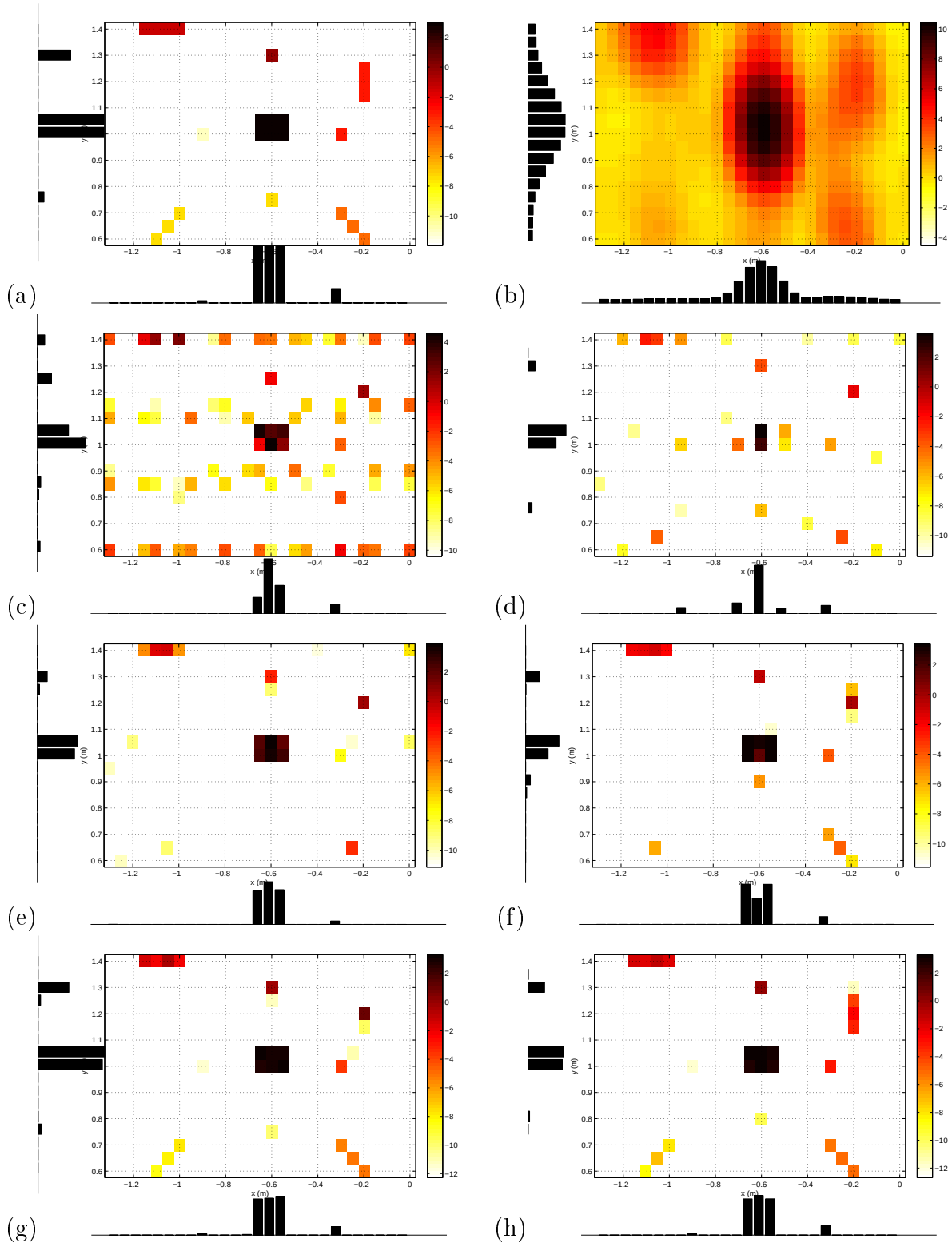

Figure 6.2: Simulation on complex sources with 14dB power dynamic range at 2500Hz, real $\sigma_e^2 = 0.86$, SNR=0dB and 15dB display : (a) 5 complex sources and 4 monopoles (b) Beamforming (c) DAMAS with 5000 iterations (5000i) (d) CLEAN. (e) DR-DAMAS (5000i) (f) CMF (g) SC-RDAMAS and (h) Proposed Bayesian JMAP approach

and distinguishes most of the sources, however, it also fails to find weak sources, the reason might be that the adaptiveness of CMF method depends on an important parameter of the sparsity constraint. Comparing to CMF, the SC-RDAMAS [34] in Fig.6.2(g) achieves better reconstructions, since SC-RDAMAS directly estimates the source powers and background noises, and it also adaptively estimates the sparsity parameter according to the noise level. But the performance of SC-RDAMAS still relies on the source number estimation. In order to make a fair comparison, both CMF and SC-RDAMAS are realized by the MATLAB toolbox SeMuDi [109].

In Fig.6.2(h), the proposed Bayesian inference approach works much better than the above methods regardless of source patterns and positions. For the monopole sources, proposed approach not only accurately estimate their positions, but also detects the weak sources as well as strong sources. For the complex sources, it successfully reconstructs both source powers and different patterns. The dynamic range of estimated powers is achieved as large as 14dB. Comparing to the beamforming resolution (31cm) [25], proposed approach achieves the 5cm super-resolution, thanks to the contribution of applying $\mathcal{DE}(\mathbf{x}|\gamma,\beta)$ sparse prior in Eq.(6.6). In addition, proposed approach is more robust to noise interference than the deconvolution methods (DAMAS, DR-DAMAS [18] and CLEAN [102]). This is owing to the simultaneous estimation of the background noises σ_e^2 and power model uncertainty σ_ξ^2, as well as source powers $\mathbf{x}$. These two important parameters are estimated as the hyperparameters in the JMAP criterion of Eq.(6.8). Comparing to the adaptiveness of sparse regularization methods (CMF [128] and SC-RDAMAS [34]), the scale parameter γ in proposed approach has the similar effect as the sparsity parameter. Since proposed Bayesian inference approach adaptively estimates γ from the the hyperparameter estimations, we provide a high resolution reconstruction with large dynamic range of estimated powers, even in the low SNR case.

The hyperparameter ($\theta = [\sigma_e^2, \sigma_\xi^2, \gamma]^T$) estimations offer the following results: the estimated variance of background noises is $\hat{\sigma^2}_e = 0.86$, almost the same as the simulated $\sigma^2 = 0.86$. And $\hat{\sigma}_\xi^2 = 0.69$ is the estimated power model uncertainty. This estimated result means that the propagation uncertainty is not very obvious, but indeed, $\hat{\sigma}_\xi^2 = 0.69$ can not be negligible with respect to the noise $\sigma_e^2 = 0.86$. The scale parameter $\hat{\gamma}$ in the sparse prior is the estimated as $\hat{\gamma} = 0.72$. As illustrated in Fig.6.1, parameter $0 < \hat{\gamma} < 1$ makes $\mathcal{DE}(\mathbf{x}|\gamma,\beta)$ model to have a much sparser distribution with a longer and heavier tail among the generated Gaussian distribution families,

so that our proposed Bayesian inference approach manages to obtain both the sparse reconstructions of source powers $\mathbf{x}$ and the wide dynamic ranges of the estimated $\mathbf{x}$ at the same time.

To quantitatively validate proposed Bayesian approach, Table 6.1, Table 6.2 and Table 6.3 show the power estimations for monopole and complex sources, as well as the average estimation error $\overline{\Delta x*}$, and the relative error of power image reconstruction δ_i. According to $\overline{\Delta x^*}$ and δ_i, it is seen that our proposed method not only successfully detects each pattern of complex sources, but also better estimates source powers and background noises in very poor SNR situations.

Table 6.1: Power estimations of 4 monopole sources by average power estimation error $\overline{\Delta x^*}$, relative errors of power image reconstruction δ_1, δ_2 and estimated noise cross-spectrum $\hat{\sigma^2}_e$ at 2500Hz, SNR=0dB, dynamic range 14dB, a cell containing '-' means unavailable.

Source powers	0.08	0.18	0.98	0.50	$\overline{\Delta x^*}$	δ_1	δ_2	$\hat{\sigma^2}_e$ ($\sigma_e^2 = 0.86$)
Beamforming	1.57	11.28	3.51	2.02	4.16	69.64	121.93	-
DAMAS	-	-	-	0.44	0.33	3.14	1.33	-
CLEAN	-	0.25	0.44	0.28	0.23	0.87	0.67	-
DR-DAMAS	-	-	0.77	0.23	0.19	0.30	0.08	-
CMF	0.09	-	0.80	0.40	0.12	0.31	0.10	0.89
SC-RDAMAS	0.09	0.10	1.05	0.43	0.06	0.21	0.06	0.85
JMAP	0.08	0.13	0.94	0.45	0.04	0.17	0.02	0.86

Table 6.2: Position estimations of 4 monopole sources by averaged position errors $\overline{\Delta p^*}$ at 2500Hz, SNR=0dB; '-' means unavailable.

Position	(-0.9,1)	(-0.6,0.75)	(-0.3,1)	(-0.6,1.3)	$\overline{\Delta p^*}$
CBF	-	-	-	-	1
DAMAS	-	(-0.6,0.85)	(-0.3,1)	(-0.6,1.25)	0.28
CLEAN	(-0.95,1)	(-0.6,0.75)	(-0.3,1)	(-0.6,1.3)	0.01
DR-DAMAS	-	-	(-0.3,1)	(-0.6,1.3)	0.43
CMF	-	(-0.6,0.9)	(-0.3,1)	(-0.6,1.3)	0.29
SC-RDAMAS	(-0.9,1)	(-0.6,0.75)	(-0.3,1)	(-0.6,1.3)	0
JMAP	(-0.9,1)	(-0.6,0.8)	(-0.3,1)	(-0.6,1.3)	0.001

Figure 6.3 and 6.4 show the simulations in the cases of various noise levels and different frequency bins. Each point on the curves is the averaged result based on 15 simulations. Firstly we fix the f=2500Hz. In Fig.6.3, the power image reconstruction errors δ_i with $i = 1, 2$ are plotted versus the

Table 6.3: Power estimations of the complex source on the center of image by power estimation error $\overline{\Delta x^*}$ at 2500Hz, SNR=0dB, a cell containing '-' means unavailable.

Source powers	2.00	2.00	2.00	2.00	2.00	2.00	$\overline{\Delta x^*}$
Beamforming	2.64	9.60	9.70	9.64	11.34	9.77	6.78
DAMAS	4.50	1.25	0.48	2.54	0.49	1.88	1.15
CLEAN	2.29	0.37	1.69	-	0.27	0.34	1.27
DR-DAMAS	2.15	2.05	1.82	1.83	2.50	1.45	0.27
CMF	1.36	2.86	2.07	2.09	1.92	1.05	0.45
SC-RDAMAS	1.83	2.00	2.05	1.72	2.16	1.95	0.12
JMAP	1.94	1.99	1.98	1.76	2.10	1.91	0.09

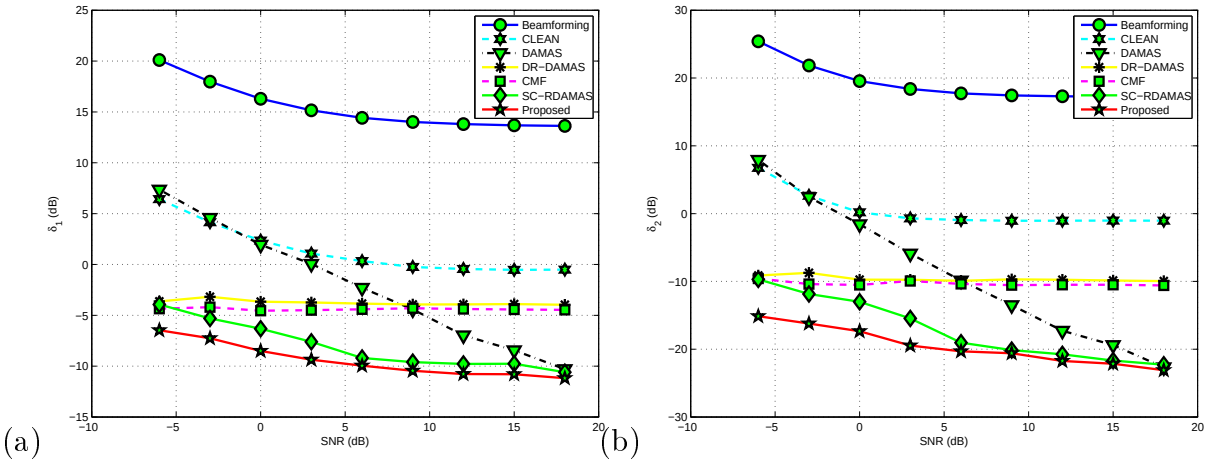

Figure 6.3: Performance comparison for relative errors of power image reconstruction δ_1 and δ_2 versus SNR [-6,18]dB on simulations at f=2500Hz (a) δ_1 VS SNR (dB) and (b) δ_2 VS SNR (dB).

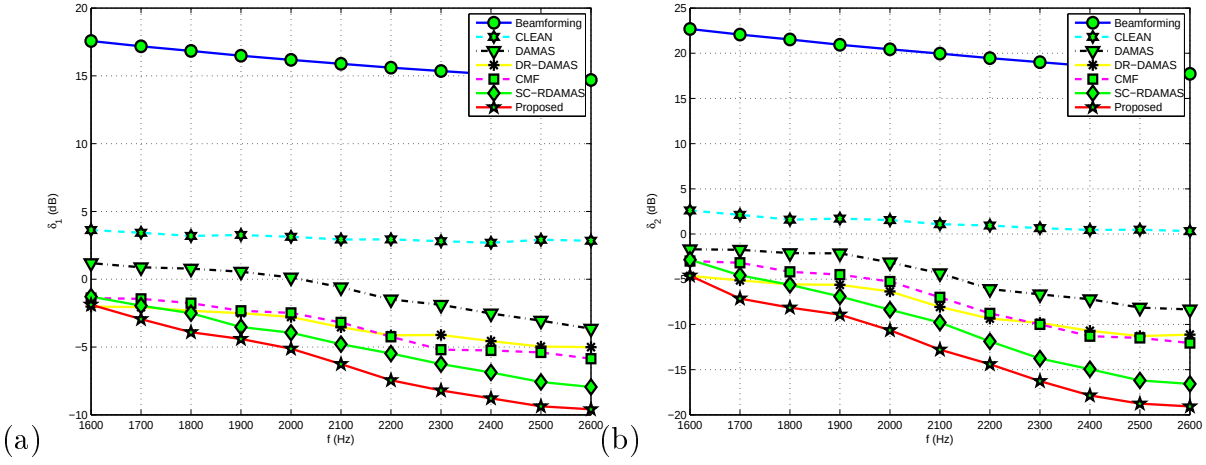

Figure 6.4: Performance comparison for relative errors of power image reconstruction δ_1 and δ_2 versus [1600,2600]Hz on simulations at SNR=3dB: (a) δ_1 (dB) VS f (Hz) and (b) δ_2 (dB) VS f (Hz).

SNR$\in [-6, 18]$dB, and it shows the robustness to background noises of the mentioned methods. As we can see, the proposed Bayesian approach is very robust to background noises, and outperforms the other methods in the cases from -6dB to 18dB. Particularly in very poor SNR cases, it still achieves very small δ_1 and δ_2. Then we fix the SNR=3dB. In Fig.6.4, δ_i are plotted versus frequencies $f \in [1600, 2600]$Hz. With higher frequencies, all the methods obtain better spatial resolution, indeed, our proposed method exceeds the others over the entire working frequency range.

6.5 Real data of wind tunnel experiments

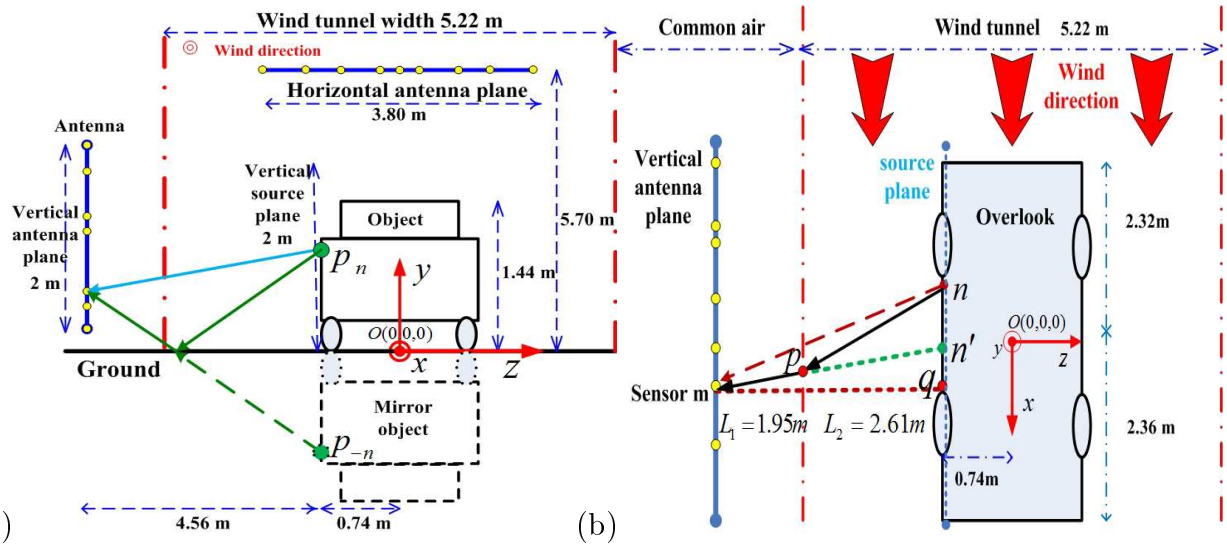

Figure 6.5: Configurations of wind tunnel experiment: (a) Frontlook and ground reflection (b) Overlook and wind refraction.

The wind tunnel experiments have been carried out by Renault SAS [86] as discussed in Chapter 4.8 and Chapter 5.4. This wind tunnel is designed to measure the acoustic source powers and their localization on the car surface, and it can simulate the situation when Renault cars travel fast on the highway.

The above figure shows the configurations of the wind tunnel S2A [86], object vehicle, NUA array and wind refraction. We suppose that all acoustic sources locate on the same plane. This assumption is almost satisfied, because the curvature of the car side is relatively small compared to the distance D=4.5m between the car and array plane. Since the scanning step is set by $\Delta p = 5cm$, the source plane of car side is of 1.5× 5 m^2 (31×101 pixels), and we also focus on a small region of the rear-view mirror: 1×1.5 m^2 (21×31 pixels). On the real data, there are T=524288 samplings with the sampling

frequency f_s=2.56×10^4 Hz. We separate these samplings into I=204 blocks with L=2560 samplings in each bloc. The working frequency band is chosen as [2400,2600]Hz, which is sensitive to human being. The image results are shown by normalized dB images with 10dB span.

For the actual propagation time $\tau_{n,m}$ and distance $r_{n,m}$ in Eq.(3.11), we apply equivalent source to make refraction correction. For $\tau_{-n,m}$ and $r_{-n,m}$ in Eq.(3.11), we use the mirror source signal s_{-n} to correct the ground reflection. The details of the propagation corrections are discussed in A and B.

6.5.1 Results of single frequency data

The proposed Bayesian approach provides the following hyperparameter estimations: the variance of background noises is $\hat{\sigma^2}_e = 25$. The power model uncertainty is as small as $\sigma_\xi^2 = 0.83$, thanks to propagation corrections in the wind tunnel, such as refraction and reflection as discussed in A and B. And the scale parameter γ in $\mathcal{DE}(\mathbf{x}|\gamma,\beta)$ prior model is estimated as $\hat{\gamma} = 0.33$, and this small value can insure the sparse distribution of source power with wide dynamic range, as discussed on the simulations in Section 6.4.

For acoustic imaging on the car side in Fig.6.6(a), the left side of Fig.6.6(b)-(g) illustrates the normalized estimated power images of mentioned methods. In Fig.6.6(b), the beamforming [25] hardly obtains a clear image of source powers, it just gives a very coarse image of strong powers around the front wheel, the rear-view mirror and the back wheel. In Fig.6.6(c), DAMAS [18] successfully deconvolves the beamforming image, and discovers weak sources in the front light, front cover and side windows. However, DAMAS also gets many false targets outside the car surface. In Fig.6.6(d), the DR-DAMAS [18] eliminates most of the false targets outside the car, but it also removes the sources on the rear-view mirror and back wheel. Figure 6.6(e) shows that CLEAN [123] overcomes drawbacks of the DAMAS and DR-DAMAS, but unexpected strong noises are detected on the ground. Moreover, we have to set carefully the parameters in CLEAN for each experiment. In Fig.6.6(f), the SC-RDAMAS [34] obtains a slightly better result than the CLEAN, but false alarms under the car body are still numerous. Finally, figure 6.6(g) reveals that proposed Bayesian approach achieves the best performance of all, especially for noise suppression and source reconstructions on the two wheels and the rear-view mirror. It removes most of the false targets under the car and on the air. The reasons of these good performances are mainly the same ones as discussed on the sim-

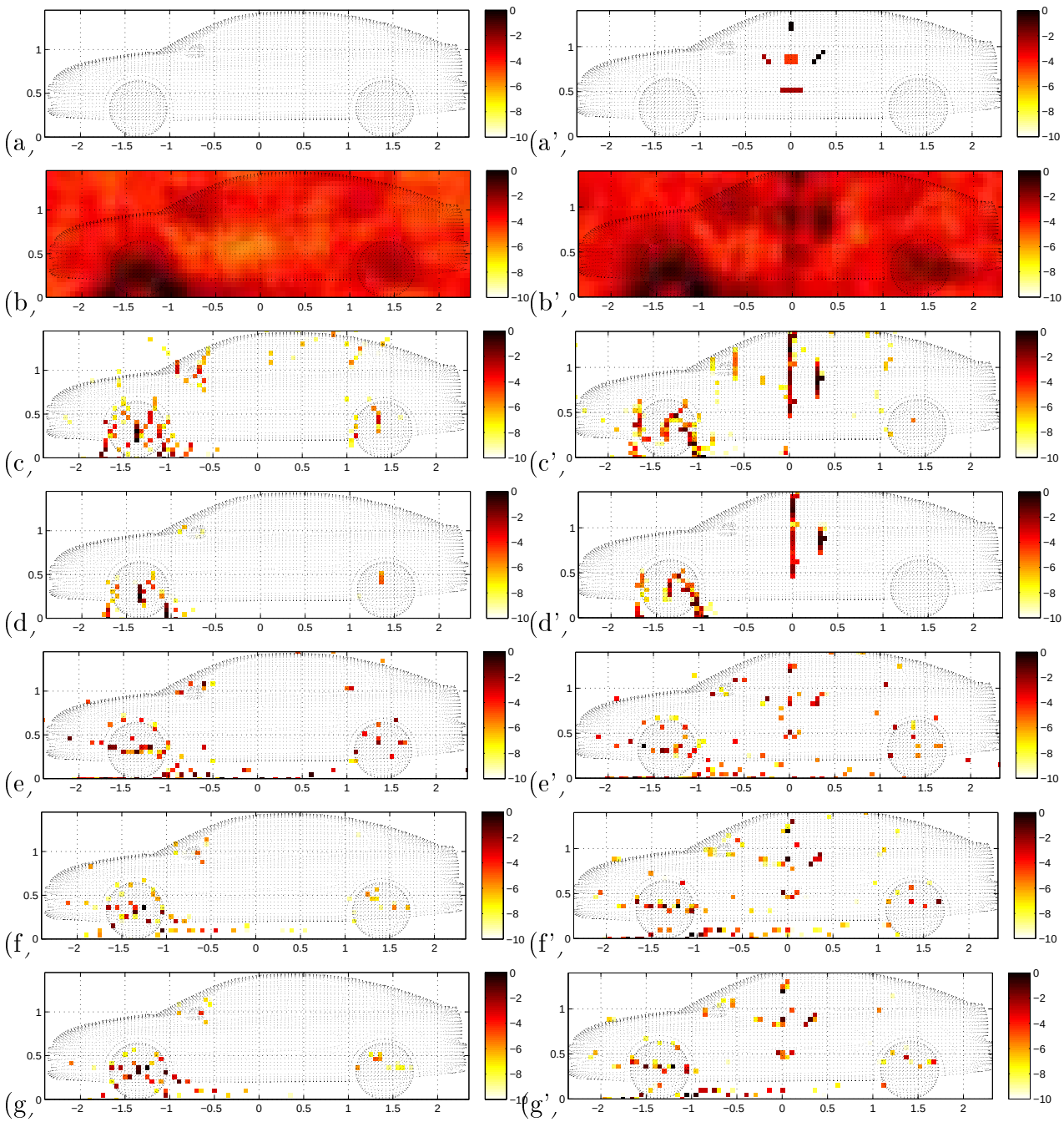

Figure 6.6: Acoustic imaging on the vehicle side at 2500Hz. Left: real data (a) vehicle surface (b) Beamforming (c) DAMAS (5000i) (d) DR-DAMAS (5000i) (e) CLEAN (f) SC-RDAMAS and (g) JMAP approach. Right: hybrid data (a') 5 simulated complex sources (b') Beamforming (c') DAMAS (5000i) (d') DR-DAMAS (5000i) (e') CLEAN (f') SC-RDAMAS and (g') JMAP approach.

ulations in Section 6.4. The computational times of the different methods are given in Table 6.4. Indeed, the hyperparameter estimations increase the computational cost, but our proposed approach remains feasible to realize it.

Based on the results of the car side, we further investigate weak sources on the rear-view mirror. In Fig.6.7, beamforming in Fig.6.7(a) mainly demon-

Table 6.4: Computational cost for treating whole car: image 31×101 pixels, at 2500Hz, based on CPU:3.33GHz, '-' means unavailable.

Methods	Beamforming	DAMAS	DR-DAMAS	CLEAN	SC-RDAMAS	JMAP	CMF
Time (s)	1	10	11	45	852	1012	-

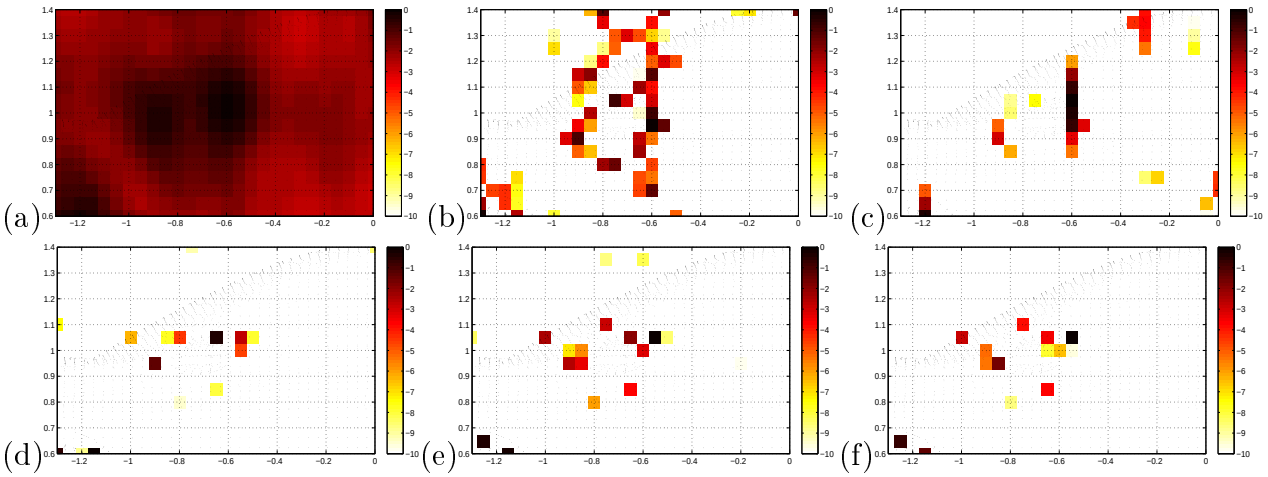

Figure 6.7: Acoustic imaging of real data on rear-view mirror part at 2500Hz: (a) Beamforming (b)DAMAS (5000i) (c) DR-DAMAS (5000i) (d) CMF (e) SC-RDAMAS and (f) Proposed Bayesian JMAP

strates 3 groups of sources: one on the corner of the front wheel, one on left of the rear-view mirror and the other on its right. DAMAS in Fig.6.7(b) improves the resolution of the beamforming, but also gets many dirty spots outside the car. In Fig.6.7(c), most of the false alarms are removed by DR-DAMAS, but the vertical resolution is not high enough. CMF in Fig.6.7(d) obtains better results than DR-DAMAS. The SC-RDAMAS in Fig.6.7(e) achieves a result as good as CMF does. But our proposed Bayesian approach in Fig.6.7(f) achieves more expected estimations of source positions and powers around the mirror, on the front wheel and on the front corner. This result is reasonable, since these car parts are highly probable to produce acoustic sources in wind tunnel.

Above all, the two experiment results agree closely with the simulation results in Fig.6.2 at 2500Hz. The proposed Bayesian approach is proved to achieve super-resolution, suppression of background noises and a wide dynamic range of power estimations.

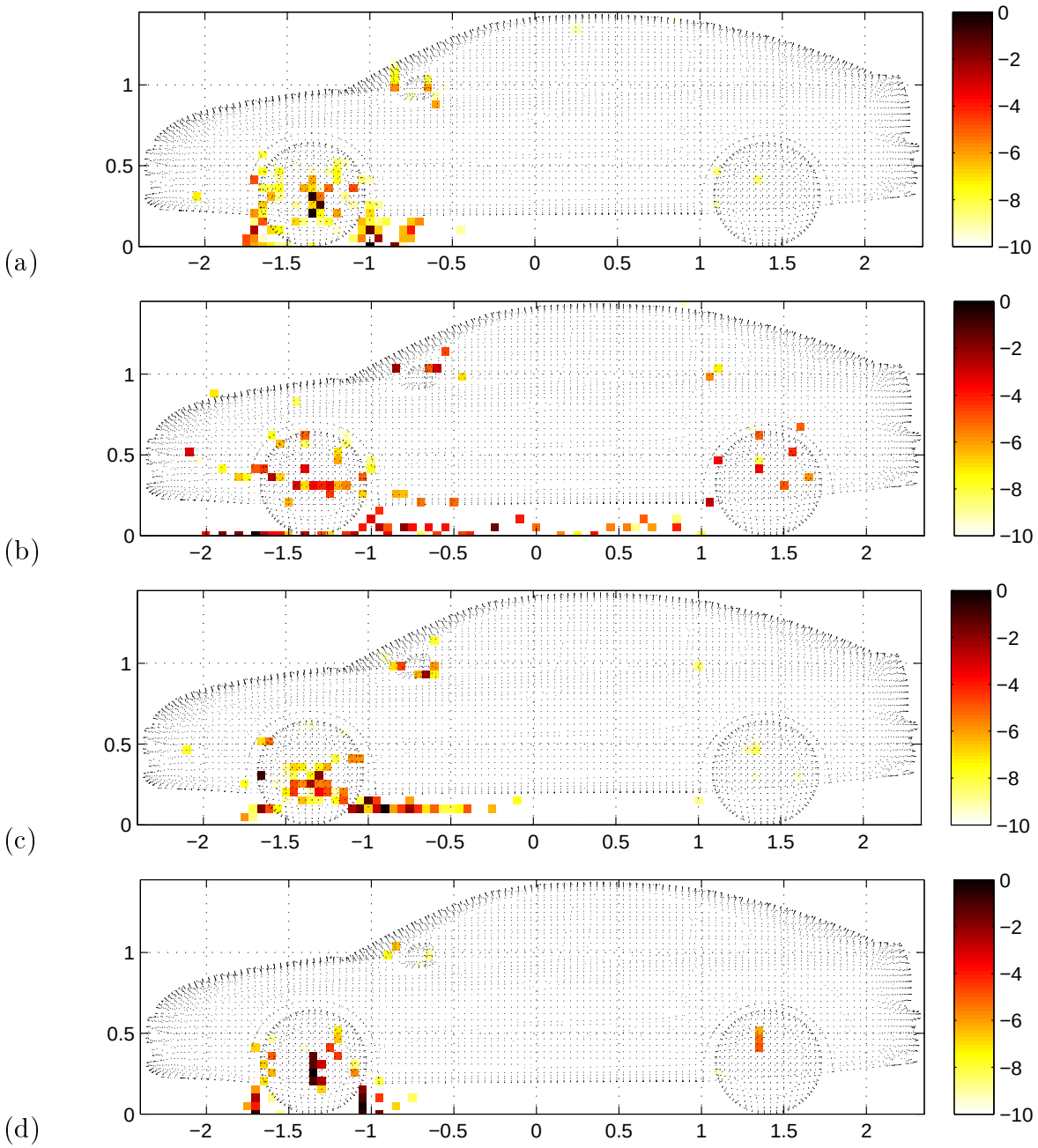

Figure 6.8: Wide-band data over [2400,2600]Hz: (a) DR-DAMAS (5000i) (b) CLEAN (c) SC-RDAMAS and (d) Proposed Bayesian JMAP

6.5.2 Results of wide-band data

Based on the effectiveness and feasibility at single frequency, we show the performance comparisons for wide-band data of [2400,2600]Hz, as illustrated in Fig.6.8. Each method obtains a better result than the corresponding one at 2500Hz in Fig.6.6. That is because the SNR is increased by averaging results over the working frequency band, and the flashing false alarms are suppressed over the wide-band average. The estimations of the DR-DAMAS

in Fig.6.8(a) are reasonable and acceptable, but the spatial resolution are not high enough on the front wheel and rear-view mirror. Fig.6.8(b) shows that the CLEAN greatly ameliorates the resolution, but it shows many unexpected spots under the car body. The SC-RDAMAS in Fig.6.8(c) has the advantages of the CLEAN, and it gets wide dynamic range of source powers around the front wheel, but it confront the same problem as CLEAN. Finally, our proposed approach in Fig.6.8(d) successfully extracts more expected source positions and powers than the above methods, for the weak ones on the mirror and back wheel, as well as the strong sources around the front wheel.

6.6 Results with hybrid data

Even though our proposed Bayesian approach obtains good performance on real data from wind tunnel experiments, it is not sufficient to validate our proposed methods. This is because the exact sources on the vehicle generated by wind flow are not known beforehand. To further verify our methods, we propose to generate hybrid data by adding synthetic sources to the real data. In order to avoid overlapping real sources, the synthetic sources are set on the region where there are no obvious sources, as shown in Fig.6.6(a'). We expect that our proposed approach could retrieve both the synthetic and potential real sources, respectively from the hybrid data. If these known synthetic sources are successfully detected, the proposed approach can be proved to be able to effectively recover the real sources on the vehicle surface. The synthetic sources are generated in the same way as discussed in Chapter 4.9.

Five synthetic complex sources with different patterns are generated as seen in Fig.6.6(a'). Their powers are among [-4.5,0]dB. The working frequency is 2500Hz. The right side of Fig.6.6(b)-(g) gives the results of mentioned methods. In fig.6.6b', the beamforming mainly shows the strong source distributions, but it reveals that there is no apparent overlaps of the measured powers between the synthetic sources and the potential ones in the real data, so that the synthetic sources are reasonably positioned beforehand. In Fig.6.6(c') and (d'), both the DAMAS and DR-DAMAS fail to provide good reconstructions for the synthetic sources, due to their sensitiveness to the background noises. In Fig.6.6(e') and (f'), both the CLEAN and SC-RDAMAS offer better estimations for the most of synthetic sources, and they also obtain comparable results on real data with respect to the left side of Fig.6.6(e) and (f). However, CLEAN also detects many artifacts under the car due to its parameter selection, and the performance of SC-RDAMAS

depends on the estimation of source number that is different between the real data and hybrid data.

In Fig.6.6(g'), the proposed Bayesian approach successfully detects almost all the synthetic sources, and it obtains more precise estimations of their powers and patterns than the other classical methods. Meanwhile, for the other sources in the real data, the proposed approach better reconstructs both the strong sources on the two wheels and rear-view mirror, and weak ones on the front cover and back window, and this result in Fig.6.6(g') is very close to the reconstruction for the real data in Fig.6.6(g). Moreover, proposed approach obtains simultaneously a better noise suppression and larger dynamic range of estimated powers, compared to the mentioned state-of-the-art methods.

But comparing to the result on real data in Fig.6.6(g), we notice the hybrid data in Fig.6.6(g'), there are additional sources except the synthetic ones, such as the ones between synthetic sources and the ones under the car body. This phenomenon could be explained as follows: first, comparing to the beamforming powers of real data in Fig.6.6(b), the beamforming of the hybrid data in Fig.6.6(b') have suffered the side effect of hybrid data, since some of the sidelobes of the synthetic sources have been more or less overlapped with the mainlobes of real sources, or vice versa. So that shadow sources are inevitably detected by proposed approach, especially for the sources on the vehicle bottom. Secondly, since the acoustic field is linear, the beamforming powers of hybrid data can be made as the superpositions of the synthetic and real data. However, the proposed JMAP estimation in Eq.(6.8) is not a linear estimator for source powers and hyperparameters, therefore small false detections would be made to some extent. But proposed JMAP approach still achieves the expected reconstructions for the synthetic sources in hybrid data, and it is reasonable to accept the source power recovery on the real data.

6.7 Conclusion and perspectives

In this chapter, we develop a robust Bayesian super-resolution approach via a sparsity enforcing prior for source localization and power reconstruction, as well as the hyperparameter estimations. Our motivation is to achieve the acoustic imaging with super spatial resolution, large dynamic range and robustness to background noises.

The main conclusions are:

- We first consider the power model uncertainty as a random variable, and we assign a probability prior to it.

- For the inverse problem, we propose to use the Bayesian inference approach via sparsity enforcing prior based on JMAP estimation.

- For the super resolution and wide dynamic range of source powers, we investigate the double exponential model for the sparse distribution of source powers, in which, $\beta = 1$ greatly improves the spatial resolution, and proper γ promotes the wide dynamic range of source powers.

- For the robust imaging in strong noises, we propose to jointly estimate γ, noise power and power model uncertainty, as well as the source powers via the JMAP criterion.

The validations of proposed forward model and JMAP estimation are presented both on the simulated, real data and hybrid data. Firstly, various simulations show that our proposed approach obtains 5cm ($\Delta B = \arg\tan\frac{\Delta p}{D} \approx 0.6^\circ$) super spatial resolution, 14dB wide dynamic range of power estimations in 0dB SNR cases for monopole and complex source imaging. Then wind tunnel experiments demonstrate that our approach has effectively detected the expected strong sources on the front wheels and mirrors, as well as weak sources on the back wheels. Finally, hybrid data further confirms that proposed approach not only well reconstructs the known synthetic sources, but also offers an expected results for real data. Moreover, proposed approach does not require the source number or SNR beforehand.

In Table 6.5, we give a brief summary for the advantages and drawbacks of mentioned classic methods, proposed SC-RDAMAS in Chapter 5 and proposed Bayesian JMAP approach in this Chapter.

Table 6.5: General performance of classical methods and proposed Bayesian JMAP approach.

Methods	CBF	CLEAN	DAMAS	DR-DAMAS	SC-DAMAS	CMF	SC-RDAMAS	JMAP
Resolutions	Low	Normal	Normal	Normal	High	Higher	Higher	Higher
Dynamic Range	Narrow	Normal	Normal	Normal	Normal	Wide	Wide	Wide
Noise	Robust	Sensitive	Sensitive	Normal	Sensitive	Robust	Robust	Robust
Computation	Least	Normal	Normal	Normal	High	High	High	Higher
Samples	Normal	Normal	Normal	Normal	Normal	More	Normal	Normal
Source number	No	Required	No	No	Required	Required	Required	No

However, one of the biggest problems of proposed Bayesian JMAP inference in Eq.(6.8) is that it confronts the non-quadratic optimization. Though we can apply alternate estimation in Eq.(6.9) for the unknown variables and hyper-parameters respectively, this procedure greatly increases the computational cost as shown in Table 6.4. But we apply the concept of power conservation in Eq.(6.10) to simplify the hyper-parameter estimation, so that proposed Bayesian JMAP approach can be still tractable.

In order to furthermore improve the efficiency of proposed Bayesian JMAP approach, we will investigate in two more directions:

- Forward model simplification in Chapter 7: to reduce the calculating time, we will use a 2D invariant convolution model to approximate the power propagation model in Eq.(5.4).

- A hierarchical prior model based on Student-t prior in Chapter 8: to improve the performance, we will first apply the Student-t priors on the sparse distribution of source powers, and on the non-stationarity of the measured errors. We will also Variational Bayesian Approximation (VBA) for the quadratic optimization.

.

- J'ai deux aides fidèles : ma patience et mes deux mains.

Michel Eyquem de Montaigne

7

2D Invariant Convolution Model of Power Propagation

In Chapter 6, we have proposed to apply the Bayesian inference approach via a sparsity enforcing prior to obtain a robust acoustic imaging. The Double Exponential model is selected as the sparse prior that contributes to achieve super spatial resolution and wide dynamic range of estimated source powers. The unknown variables and parameters can be jointly estimated by the JMAP optimization. Main steps of this optimization depend on the computation of forward and joint operators. However, the forward model of power acoustic propagation in Eq.(5.4) requires tremendous matrix multiplications due to large dimension of the power propagation matrix $\mathbf{C}$. This is one of the reasons that proposed Bayesian JMAP optimization is very time-consuming.

In this chapter, our objective is to approximate this forward model by using a 2D invariant convolution model, so that tremendous computational cost of matrix multiplication can be greatly reduced by 2D convolution. Owing to the quasi-Symmetric Toeplitz Block Toeplitz (STBT) structure of propagation matrix $\mathbf{C}$, the invariant convolution kernel (size and values) can be derived from the items on the middle row of $\mathbf{C}$. From the invariant kernel, two accelerations can be done:

- Algorithmic acceleration using a fixed kernel with a small and square size; further acceleration using a separable convolution kernel which can be separated into two vectors;

- Hardware acceleration using a 'many-core' processor such as the Graphics Processing Unit (GPU). Indeed, convolution is well suited to the parallelization structure of GPU.

This chapter is organized as follows: Section 7.1.3 presents the computational complexity of different approximation for power propagation model. Section 7.2 introduces the STBT matrix approximation of the power propagation model. Section 7.2.1 presents the 2D-convolution model with the variant and invariant kernels respectively. Section 7.3 and 7.4 validate the proposed 2D convolution approximation on simulations and real data respectively. Finally Section 7.5 concludes this chapter.

7.1 Conventional forward model via matrix multiplication

Here we firstly review our main points discussed from Chapter 3 to Chapter 5. We assume that acoustic sources are uncorrelated monopoles [18, 43]; microphones are omni-directional with unitary gain; background noises at the microphones are Additive Gaussian White Noise (AGWN), independent and identically distributed (i.i.d); complex reverberation in the open wind tunnel could be neglected.

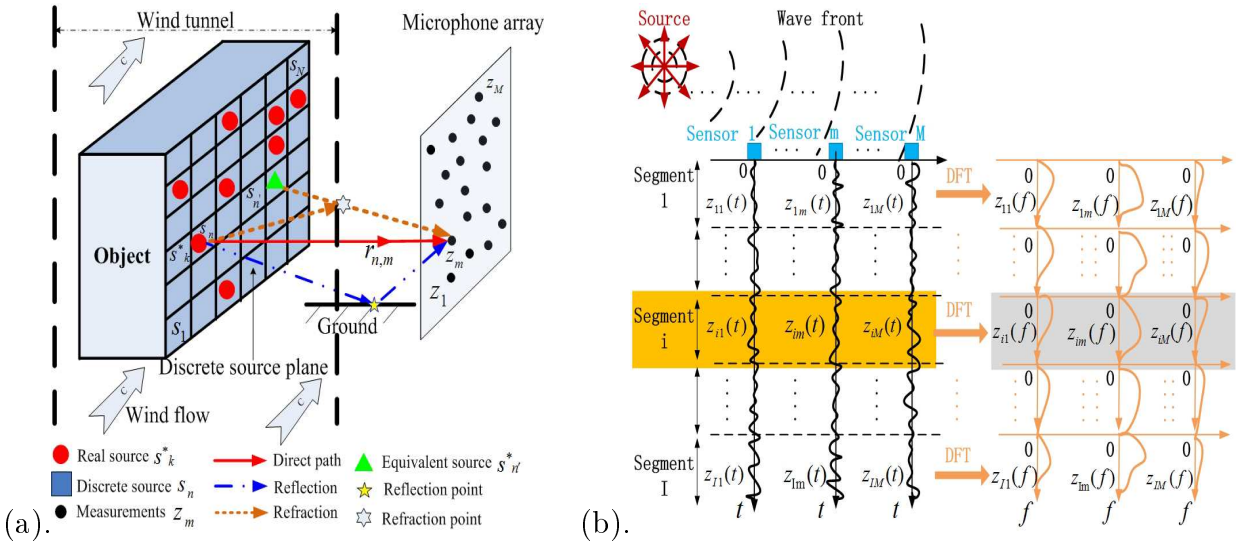

Figure 7.1: (a). Illustration of the acoustic signal propagation in wind tunnel[29] of Fig.3.7. (b). Illustration of the signal processing procedure in Eq.(3.15) of Fig.3.4.

Figure 3.7 illustrates the acoustic signal propagation from the source plane to the microphone array in the wind tunnel, where microphones are installed

outside the wind flow. On the source plane, we suppose K unknown original source signals $\mathbf{s}^* = [s_1^*, \cdots, s_K^*]^T$ at unknown positions $\mathbf{P}^* = [\mathbf{p}_1^*, \cdots, \mathbf{p}_K^*]^T$, where $\mathbf{p}_k^*$ denotes the 3D coordinates of kth original source signal s_k^*, notation $(\cdot)^*$ represents the original sources, and operator $(\cdot)^T$ denotes the transpose. On the microphone plane, we consider M microphones at known positions $\bar{\mathbf{P}} = [\bar{\mathbf{p}}_1, \cdots, \bar{\mathbf{p}}_M]^T$. The source plane is then equally discretized into N grids at known positions $\mathbf{P} = [\mathbf{p}_1, \cdots, \mathbf{p}_N]^T$. We assume that K original sources $\mathbf{s}^*$ sparsely distribute on these grids, satisfying $N > M > K$ and $\mathbf{P}$ including $\mathbf{P}^*$. We thus get N discrete source signals $\mathbf{s} = [s_1, \cdots, s_N]^T$ at known positions $\mathbf{P}$, satisfaying $s_n = s_k^*$, for $\mathbf{p}_n = \mathbf{p}_k^*$; $s_n = 0$ others. Since K<<N, $\mathbf{s}$ is full of zero, and it becomes a sparse signal with K-sparsity in the space domain. Therefore, to reconstruct $\mathbf{s}^*$ is to reconstruct K-sparsity signal $\mathbf{s}$. And $\mathbf{p}_k^*$ can be deprived from the discrete position $\mathbf{p}_n$, where s_n is non-zero.

7.1.1 Forward model of acoustic signal propagation

Signal processing procedure is illustrated in Fig.3.4. For the mth microphone with $m \in [1, \cdots, M]$, there are T samplings of acoustic signals in time domain. Then these T temporal samplings are divided into I blocks with L samplings in each block. We note $\mathbf{z}_{i,m}(t)$ as the received signal of the ith sampling block ($i \in [1, \cdots, I]$) at the mth microphone in the sampling time $t \in [(i-1)\,L + 1, \cdots, i\,L - 1]$, and total sampling number is noted by $T = I \times L$. Since original source signals are usually of wide-band, we apply the Discrete Fourier Transform (DFT) in time domain to treat measured signals $\mathbf{z}_{i,m}(t)$ at each block so as to obtain L narrow frequency bins f_l (l$\in$ $[1, \cdots, L]$). Let $\mathbf{z}_i(f_l) = [z_{i,1}(f_l), \cdots, z_{i,M}(f_l)]^T$ denote all measured signals in frequency domain. The signal processing is made independently for each frequency bin, thus in the following, we omit f_l for simplicity. Thus $\mathbf{z}_i$ can be modeled [123, 18, 29] in Eq.(3.15) as

$$\mathbf{z}_i = \mathbf{A}(\mathbf{P})\,\mathbf{s}_i + \mathbf{e}_i\,,$$

where $\mathbf{A}(\mathbf{P}) = [\mathbf{a}(\mathbf{p}_1) \cdots \mathbf{a}(\mathbf{p}_N)]$, $\mathbf{A}(\mathbf{P}) \in \mathbb{C}^{M \times N}$ consists of N steering vectors is defined in Eq.(3.10) with reflecting coefficient $\rho = 0$ as:

$$\mathbf{a}(\mathbf{p}_n) = \left\{ \frac{1}{r_{n,1}} \exp\left[-j(2\pi f_l \tau_{n,1})\right], \cdots, \frac{1}{r_{n,M}} \exp\left[-j(2\pi f_l \tau_{n,M})\right] \right\}^T,$$

where $r_{n,m}$ denotes the distance from source n to sensor m, $\tau_{n,m}$ propagation time during $r_{n,m}$. For $r_{n,m}$, we also consider the ground reflection and

wind refraction in authors' paper [29]. For simplicity, $\mathbf{a}(\mathbf{p}_n)$ is short as $\mathbf{a}_n$ afterwards.

In summary, the forward model of signal propagation in Eq.(3.15) is a linear but under-determined (M<N) system of equations for solving K-sparsity signal $\mathbf{s}$.

7.1.2 Forward model of acoustic power propagation

Based on Eq.(3.15), it is convenient to obtain the forward model of acoustic power propagation using Beamforming methods [25, 18, 29]:

Let us recall the forward model of acoustic power propagation of Eq.(5.4):

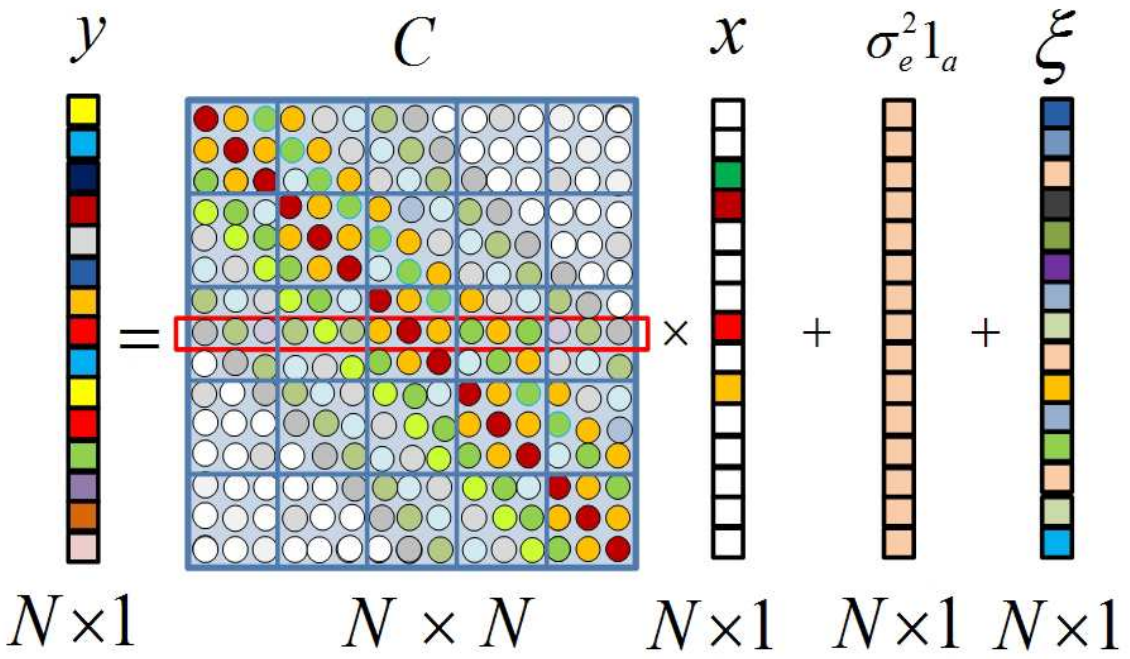

Figure 7.2: Illustration of forward power model considering multi-path uncertainty $\boldsymbol{\xi}$ in Eq.(5.4).

$$\mathbf{y} = \mathbf{C}\,\mathbf{x} + \sigma_e^2\,\mathbf{1}_a + \boldsymbol{\xi}$$

where $\mathbf{y} = \{y_n\}_N^T$ denotes the Beamforming power vector; y_n can be interpreted as the estimated source power at grid n. And $\mathbf{y} = \tilde{\mathbf{A}}^{\dagger}\,\mathbb{E}[\mathbf{z}\mathbf{z}^{\dagger}]\,\tilde{\mathbf{A}}$ can be directly obtained from Eq.(3.15), where $\tilde{\mathbf{A}} = [\tilde{\mathbf{a}}(\mathbf{p}_1)\cdots\tilde{\mathbf{a}}(\mathbf{p}_N)]$, $\tilde{\mathbf{A}}(\mathbf{P}) \in \mathbb{C}^{M\times N}$ denotes the Beamforming steering matrix, and $\tilde{\mathbf{a}}(\mathbf{p}_n) = \frac{\mathbf{a}_n}{||\mathbf{a}_n||_2^2}$, operator $(\cdot)^{\dagger}$ denotes conjugate transpose, $\mathbb{E}[\cdot]$ denotes mathematical expectation. In practice, $\mathbb{E}[\mathbf{z}\mathbf{z}^{\dagger}] \approx \frac{1}{I}\sum_i^I \mathbf{z}_i\mathbf{z}_i^{\dagger}$ is approximated. $\mathbf{x} = diag\,\{\mathbb{E}[\boldsymbol{s}\boldsymbol{s}^H]\}$ denotes the unknown source power vector, and $diag\{\cdot\}$ denotes diagonal

items; thus $\mathbf{x}$ is a signal as K-sparsity as $\mathbf{s}$. And σ_e^2 denotes the variance of i.i.d AGWN noises $\boldsymbol{e}$. Notation $\mathbf{1}_a = [\frac{1}{\|\mathbf{a}_1\|^2}, \cdots, \frac{1}{\|\mathbf{a}_N\|^2}]^T$ represents the noise attenuation for different grids. $\mathbf{C} = \{c_{i,j}\}_{N\times N}$ denotes the power propagation matrix, defined as:

$$\begin{aligned} c_{i,j} &= \frac{\|\mathbf{a}_i^H \mathbf{a}_j\|_2^2}{\|\mathbf{a}_i\|_2^2} \\ &= \left| \frac{1}{\sum_{m=1}^{M} \frac{1}{r_{i\,m}^2}} \sum_{m=1}^{M} \frac{1}{r_{i\,m}\, r_{j\,m}} e^{-\jmath 2\pi f_l (\tau_{j\,m} - \tau_{i\,m})} \right|^2 , \end{aligned} \tag{7.1}$$

where $\mathbf{a}_i$ is defined in Eq.(3.15); $r_{i\,m}$ denotes the propagation distance from ith discrete source (at the position $\mathbf{p}_i$ on the discrete source plane) to the mth microphone; f_l denotes the lth frequency bin; M is the total number of microphones. $r_{i\,m}$ denotes the propagation distance from ith discrete source (at the position $\mathbf{p}_i$ on the discrete source plane) to the mth sensor; f_l denotes the lth frequency bin; M is the total number of sensors; propagation time $\tau_{j\,m} = \frac{r_{j\,m}}{c_0}$ is obtained in the uniform media, with c_0 being the acoustic propagation speed.

According to Eq.(7.1), it yields $0 \leq c_{i,j} \leq 1$ and $c_{i,i} = 1$. In fact, $c_{i,j}$ can represent the power contribution (%) of the microphone array from the jth source to the ith position on the source plane. So that $c_{i,j}$ can also be seen as the Point Spread Function (PSF) of the microphone array. This PSF is determined by two factors: the microphone array topology and the distance from the source plane. In ideal case, $c_{i,j} = \delta_{i,j}$ becomes the Dirac function, and it derives the simple expression as $\mathbf{y} = \mathbf{x} + \sigma_e^2\, \mathbf{1}_a + \boldsymbol{\xi}$ from Eq.(5.4).

In brief, compared with signal propagation model of Eq.(3.15), the power propagation model of Eq.(5.4) is a linear and determined system of equations for solving K-sparsity source powers $\mathbf{x}$.

7.1.3 Computational complexity in forward model of power propagation

In Fig.7.3, we show one example of Eq.(5.4) in the way of image forms. N-length vector $\mathbf{x}$ and $N \times N$ power propagation matrix $\mathbf{C}$ cause a computational complexity as heavy as $O(N^2)$ to calculate matrix multiplications $\mathbf{C}\,\mathbf{x}$. Moreover, it seems that $\mathbf{C}$ is close to a quasi Symmetric Toeplitz Block Toeplitz matrix (STBT). These motivate us to use a 2D-convolution model

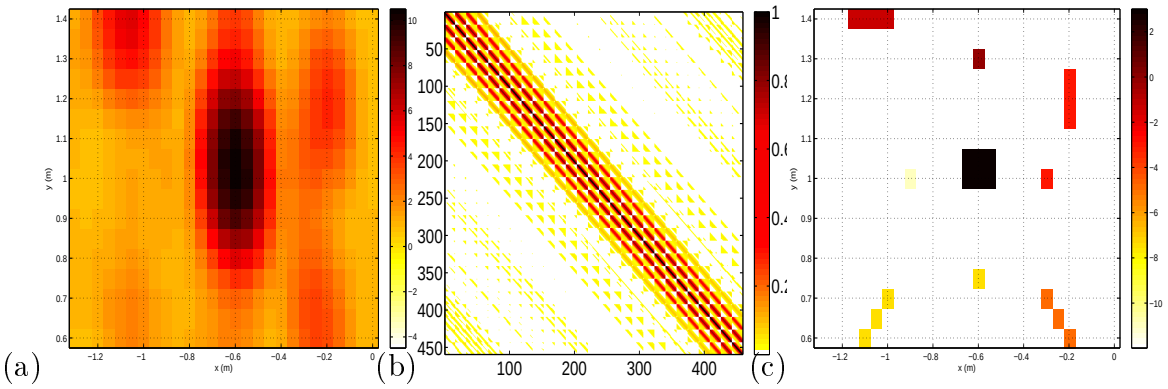

Figure 7.3: Simulation at 2500Hz on Eq.(5.4) in Chapter 5.3: (a) Measured beamforming power image $\mathbf{y}_0$ (17×27) (b) Power propagation matrix $\mathbf{C}$ (459×459) (c) Source power image $\mathbf{x}_0$ (17×27).

with a invariant kernel $\mathbf{h}$ to approximate $\mathbf{C}\,\mathbf{x}$ in Eq.(5.4):

$$\mathbf{C}\,\mathbf{x} \approx \mathbf{h} * \mathbf{x}_0\,, \tag{7.2}$$

where $\mathbf{x}_0$ denotes the source power image, which is matrix form of vector $\mathbf{x}$; and $\mathbf{h}$ denotes the 2D invariant kernel, with the size of $N_h \times N_h$ $(N_h^2 < N)$; operator $*$ denotes valid convolution: the output matrix of valid convolution consists of those overlap parts without zero-padded edges, so that the output matrix is the same size of input matrix.

In order to derive the 2D-convolution model, it is better to transform 1D vectors $\mathbf{x}$ and $\mathbf{y}$ in Eq.(5.4) into 2D matrix forms $\mathbf{x}_0$ and $\mathbf{y}_0$ respectively. We suppose that the source plane is discretized by $N = N_r \times N_c$ identical grids, where N_r and N_c denote row and column number respectively, provided $N_r \leq N_c$ for a rectangular plane. We then define a $N_r \times N_c$ source power image as $\mathbf{x}_0 = [x_{p,q}]$ with $p \in [1, \cdots, N_r]$ and $q \in [1, \cdots, N_c]$. Then $\mathbf{x}_0$ can be vectorized into the vector $\mathbf{x}$ in the manner of the column-first order as shown on the top of Fig.7.8. The relations between $x_j \in \mathbf{x}$ and $x_{p,q} \in \mathbf{x}_0$, $y_i \in \mathbf{x}$ and $y_{p,q} \in \mathbf{y}_0$ can be modeled as

$$\begin{cases} x_j = x_{p,q}\,, & j = p + (q-1)\,N_r \\ y_i = y_{p,q}\,, & i = p + (q-1)\,N_r \end{cases} \tag{7.3}$$

where $p \in [1, \cdots, N_r]$ and $q \in [1, \cdots, N_c]$.

In Eq.(7.2), owing to convolution approximation, the computational complexity can be significantly reduced from $O(N^2)$ into $O(N_h^2\,N)$, even further $O(N\,log_2\,N)$ using the Fast Fourier Transformation (FFT) .

In particular, if the 2D-convolution kernel can be separable into two 1D-convolution kernels as:

$$\mathbf{h} = \mathbf{h}_1 * \mathbf{h}_2^T\,, \tag{7.4}$$

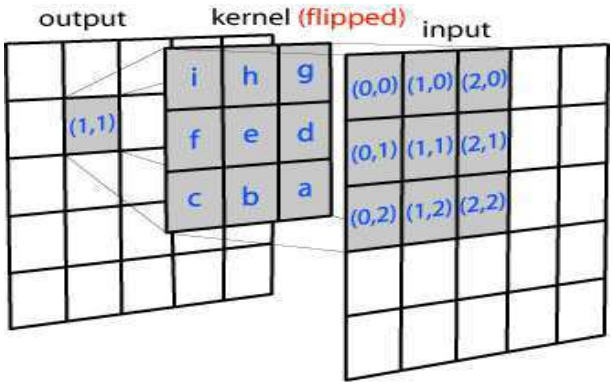

Figure 7.4: Illustration of valid 2D convolution (cited from site http://www.songho.ca/index.html).

where $\mathbf{h}_1$ and $\mathbf{h}_2$ denotes the 1D-convolution kernels, with the size of $N_h \times 1$. So that the computational complexity of matrix multiplications $\mathbf{C}\,\mathbf{x}$ can be greatly reduced into $O(2\,N_h\,N)$. The computational complexity comparison is shown in Table 7.1.

Table 7.1: The computational complexity comparison of operations.

Operation	Expression	Complexity	Speed gain
Matrix multiplication	$\mathbf{C}\,\mathbf{x}$	$O(N^2)$	1
2D invariant convolution	$\mathbf{h} * \mathbf{x}_0$	$O(N_h^2\,N)$	N/N_h^2
2D separable invariant convolution	$\mathbf{h}_1 * \mathbf{h}_2^T * \mathbf{x}_0$	$O(2\,N_h\,N)$	$N/2\,N_h$

7.2 Proposed convolution models via variant and invariant convolution kernels

The power propagation model in Eq.(5.4) reveals that the source power reconstruction can be seen as the image deconvolution from the blurred Beamforming result. However, the Beamforming often involves in the convolutions with spatially variant kernels. This effect is shown in Fig. 7.5(a-b): same sources produce different shapes of PSFs on different positions, and the center PSF has the smallest size, while the ones on the corners have much larger sizes. But in Fig. 7.5(c), all the PSFs look like similar to each other in the far-field.

We then derive convolution kernels from the power propagation matrix $\mathbf{C}$ [67, 54]. In the far-field condition, $\mathbf{C}$ can be separated into a STBT matrix

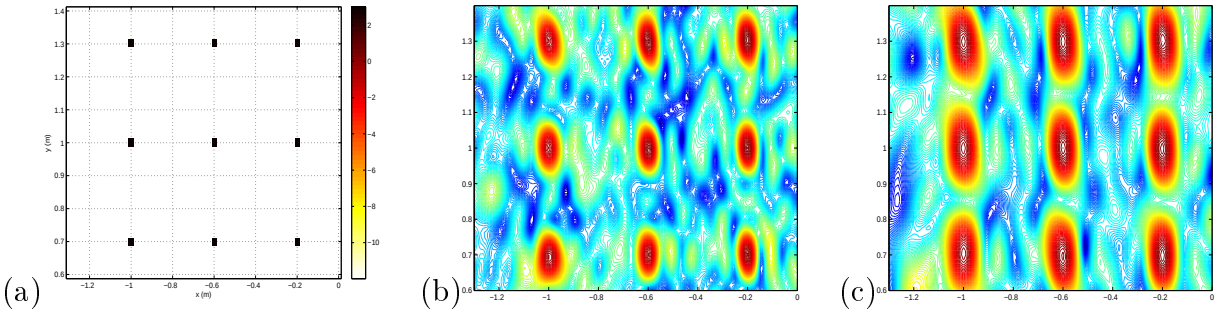

Figure 7.5: Simulation for 9 monopole sources, 0dB dynamic range, 15cm interval, 2.5cm grid, 12000 pixels, 64 sensors, 2500Hz, no background noise, no reflection nor refraction: (a). Source power image (b). Spatially variant PSF in near-field condition (c). Approximated spatially invariant PSF in far-field condition

$\tilde{\mathbf{C}}$, and two diagonal matrices $\mathbf{D}_1$, $\mathbf{D}_2$ as: (See Appendix for details)

$$\mathbf{C} \approx \mathbf{D}_1 \tilde{\mathbf{C}} \, \mathbf{D}_2 \,, \tag{7.5}$$

where $\mathbf{D}_1 = Diag\left[\bar{r}_i^2\right]$ and $\mathbf{D}_2 = Diag\left[\frac{1}{\bar{r}_j^2}\right]$ with $i, j \in [1, \cdots, N]$ denote the diagonal matrice; $\bar{r}_i = \frac{1}{M}\sum_{m=1}^{M} r_{i,m}$ denotes the averaged distance from the ith source to all microphone sensors; and $\tilde{c}_{i,j} \in \tilde{\mathbf{C}}$ is

$$\tilde{c}_{i,j} = \frac{1}{M^2} \left| \sum_{m=1}^{M} e^{j\frac{2\pi f_l}{c_0}(r_{i,m} - r_{j,m})} \right|^2 , \tag{7.6}$$

Therefore, using $\tilde{c}_{i,j}$ in Eq.(7.6) to replace $c_{i,j}$ in Eq.(7.1), the forward model of acoustic power propagation in Eq.(5.4) can be rewritten as:

$$y_i = \bar{r}_i^2 \sum_{j=1}^{N} \tilde{c}_{i,j} \, \frac{x_j}{\bar{r}_j^2} + \sigma_e^2 \, \frac{1}{||\mathbf{a}_i||^2} + \xi_i + \eta_i \,, \tag{7.7}$$

where $\boldsymbol{\eta} = [\eta_1, \cdots, \eta_N]^T$ denotes the approximation error vector. For N discrete sources, using Eq.(7.5) to replace $\mathbf{C}$ in the original forward model of Eq.(5.4), Eq.(7.7) can be rewritten in a vector form as:

$$\tilde{\mathbf{y}} = \tilde{\mathbf{C}} \, \tilde{\mathbf{x}} + \boldsymbol{\epsilon} \,, \tag{7.8}$$

where $\tilde{\mathbf{y}} = \mathbf{D}_1^{-1} \, \mathbf{y}$ denotes the measured Beamforming power vector with attenuation $\mathbf{D}_1^{-1}$; and $\tilde{\mathbf{x}} = \mathbf{D}_2 \, \mathbf{x}$ denotes the source power vector with attenuation $\mathbf{D}_2$; and $\boldsymbol{\epsilon} = \mathbf{D}_1^{-1}(\sigma_e^2 \, \mathbf{1}_a + \boldsymbol{\xi} + \boldsymbol{\eta})$ denotes the convolution model errors. Since the averaged distance $\bar{r}_i$ can be easily calculated beforehand, we take $\tilde{\mathbf{x}}$ as $\mathbf{x}$ and $\tilde{\mathbf{y}}$ as $\mathbf{y}$ for symbol simplicity in the followings.

According to the STBT matrix $\tilde{\mathbf{C}}$, we can rewrite Eq.(7.8) by using the 2D convolution model as:

$$\mathbf{y} = \mathbf{H}\,\mathbf{x} + \boldsymbol{\epsilon}\,, \tag{7.9}$$

where matrix $\mathbf{H}$ denotes valid convolution matrix, satisfying:

$$[\mathbf{H}\,\mathbf{x}]_i = [\mathbf{h} * \mathbf{x}_0]_{p,q}\,, \quad i = p + (q-1)\,N_r\,, \tag{7.10}$$

where index $[\cdot]_i$ represents the ith item of a vector; index $[\cdot]_{p,q}$ represent the pth row, qth column item of a matrix; N_r denotes row size of the source plane. To express convolution kernel $\mathbf{h}$, we will discuss the spatially-variant and spatially-invariant two cases in the following subsections.

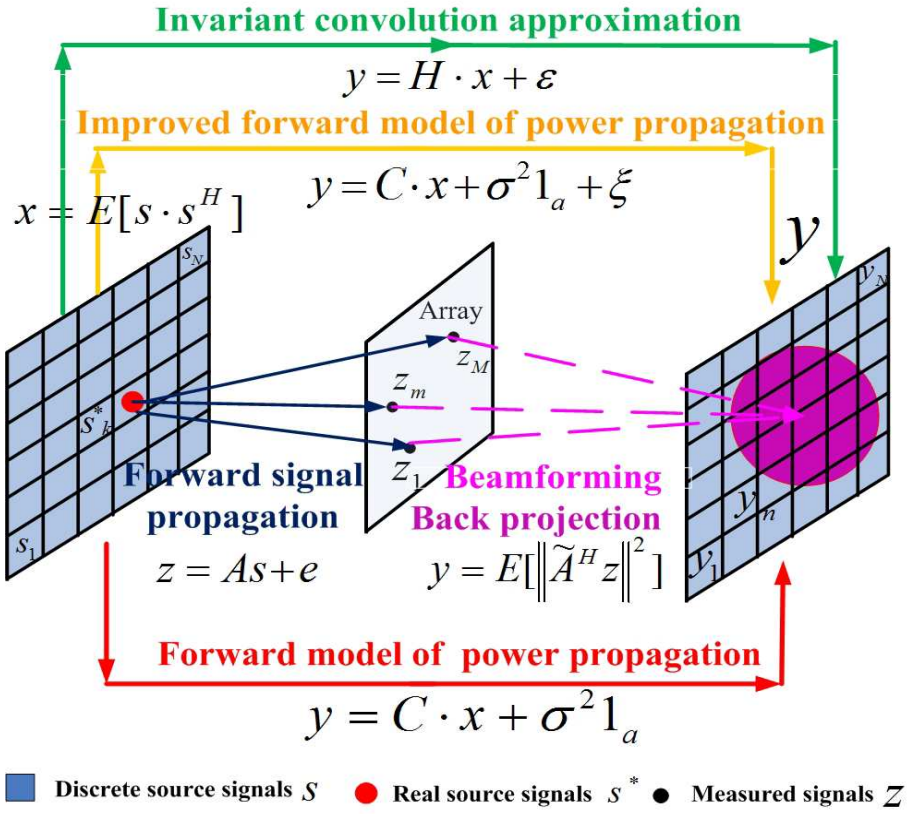

Figure 7.6: Illustration of 2D invariant convolution approximation model for acoustic power propagation.

In Fig.7.6, we show the relationships among the 2D invariant convolution approximation in Eq.(7.9), the improved power propagation model in Eq.(5.4) and the conventional one in Eq.(4.46).

7.2.1 2D spatially-variant kernel

According to the STBT matrix $\tilde{\mathbf{C}}$ in Eq.(7.8), spatially-variant kernels in Eq.(7.10) can be derived as [56, 57, 89]:

$$\mathbf{h} = \sum_{p=1}^{N_c} \sum_{q=1}^{N_c} \mathcal{D}^{p,q}\, \mathbf{h}^{p,\,q}\,, \tag{7.11}$$

where $\mathcal{D}^{p,q}$ denotes the piecewise constant interpolation function [89], which is non-negative diagonal matrix satisfying $\sum_{p=1}^{N_c} \sum_{q=1}^{N_c} \mathcal{D}^{p,q} = \mathbf{I}$ (identity matrix), and the lth diagonal item is 1 if the lth PSF is in the region of (p, q). We call $\mathbf{h}^{p,\,q}$ the spatially-variant kernel, since $\mathbf{h}^{p,\,q}$ varies along with the convolution output $y_i \in \mathbf{y}$, $i = p + (q-1)\, N_r$ in Eq.(7.9).

According to the expression of $\tilde{\mathbf{C}}$ in Eq.(7.6), each item $h^{p,\,q}(k,l) \in \mathbf{h}^{p,\,q}$ in Eq.(7.11) is obtained as:

$$\begin{cases} h^{p,\,q}(k,l) = \tilde{c}_{i,j}\,, \\ i = p + (q-1)\, N_r\,, \quad j = i + (\lfloor \frac{N_r^h+1}{2} \rfloor - k)\, N_r + \lfloor \frac{N_c^h+1}{2} \rfloor - l \end{cases}, \tag{7.12}$$

where $N_r^h \times N_c^h$ denotes the kernel size; $k \in [1, \cdots, N_r^h]$, $l \in [1, \cdots, N_c^h]$; operator $\lfloor \cdot \rfloor$ denotes integer part.

In brief, $h^{p,\,q}(k,l)$ is derived from $\tilde{c}_{i,j}$ in three steps: Firstly, $\mathbf{h}^{p,\,q}$ comes from the specific $\tilde{c}_{i,j}$ which are on the same row of $i = p + (q-1)\, N_r$ in matrix $\mathbf{C}$; Then, the item $h^{p,\,q}(k,l)$ is derived by specific $c_{i,j}$ on the column j which is determined by the known index i, k, l as shown in Eq.(7.12); Finally, $\mathbf{h}^{p,\,q}$ should be flipped up-down, left-right according to the definition of 2D valid convolution.

One example of variant convolution kernel derivation is shown in Fig 7.7, in which, $N_r = 17$, $N_c = 27$ and $N = N_r \times N_c = 459$, other simulation configurations are the same as Chapter 5.3.

7.2.2 2D invariant kernel

Owing to the STBT matrix $\tilde{\mathbf{C}}$, its middle row ($i = \lfloor \frac{N+1}{2} \rfloor$) contains most of the useful items of other rows in $\tilde{\mathbf{C}}$. According to variant kernels in Eq.(7.12), we can derive an invariant convolution kernel $\mathbf{h} = [h_{k,l}]$ with $k, l \in [1, \cdots, N_r]$ from the middle row of $\tilde{\mathbf{C}}$ as:

$$\begin{cases} h_{k,l} = \tilde{c}_{i,j}\,, \\ i = \lfloor \frac{N+1}{2} \rfloor\,, \quad j = i + (\lfloor \frac{N_r+1}{2} \rfloor - k)\, N_r + \lfloor \frac{N_r+1}{2} \rfloor - l \end{cases}, \tag{7.13}$$

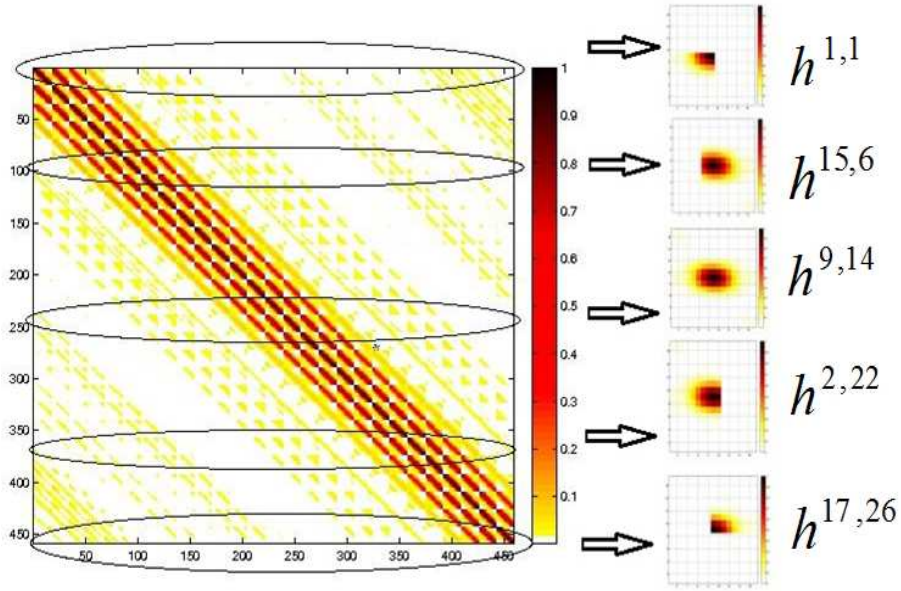

Figure 7.7: Illustration of variant kernel derivation at 2500Hz in Eq.(7.12).

where $\mathbf{h}$ is can be a $N_r \times N_r$ square matrix, since the STBT matrix $\tilde{\mathbf{C}}$ consists of $N_r \times N_r$ square subblocks. Compared with the 'variant' kernel in Eq.(7.12), the 'invariant' kernel in Eq.(7.13) does not change along with convolution output y_i, $i = p + (q-1)\,N_r$, but remains the same $i = \lfloor \frac{N+1}{2} \rfloor$.

On the top of Fig.7.8, we illustrate how power propagation matrix $\mathbf{C}$ can be approximated to a STBT matrix $\tilde{\mathbf{C}}$. As discussed in Eq.(4.47), $0 \leq c_{i,j} \leq 1$ with $c_{i,j} \in \mathbf{C}$ reflects the power contribution rate of the jth discrete source to the ith position on the source plane. On the bottom of Fig.7.8, we illustrate the geometrical interpretation of $c_{i,j} \in \mathbf{C}$ that for the fixed position i and all the source power x_j with $j \in [1, \cdots, N]$, the geometrical distribution of $c_{i,j}$ can form a circle with the center i and radius $|j-i|$, in which, the closer j is to i, the larger $c_{i,j}$ becomes. In particular $c_{i,j} = 1$ for $j = i$. Moreover, $c_{i,j}$ constitutes a quasi-symmetric Teoplitz matrix. Furthermore, since the index $i = p + (q-1)\,N_r$ for y_i and $j = p + (q-1)\,N_r$ for x_j are periodically varying due to the vectorization from matrix $x_{p,q} \in \mathbf{x}_0$ to vector $x_i \in \mathbf{x}$ in Eq.(7.3), it yields that $\mathbf{C}$ has a quasi-block structure, in which, each subblock is consisted of a quasi-symmetric Toeplitz matrix. In Appendix C, we discuss how $\tilde{\mathbf{C}}$ can be approximated into a STBT matrix.

According to Eq.(7.13) and Eq.(7.6), $\mathbf{h}$ is thus a real and non-negative symmetric matrix. In Appendix D, we give a detail discussion that $\mathbf{h}$ is also a separable 2D-convolution kernel which can be separated into two 1D-convolution kernels as discussed in Eq.(7.4). One of the advantages is that

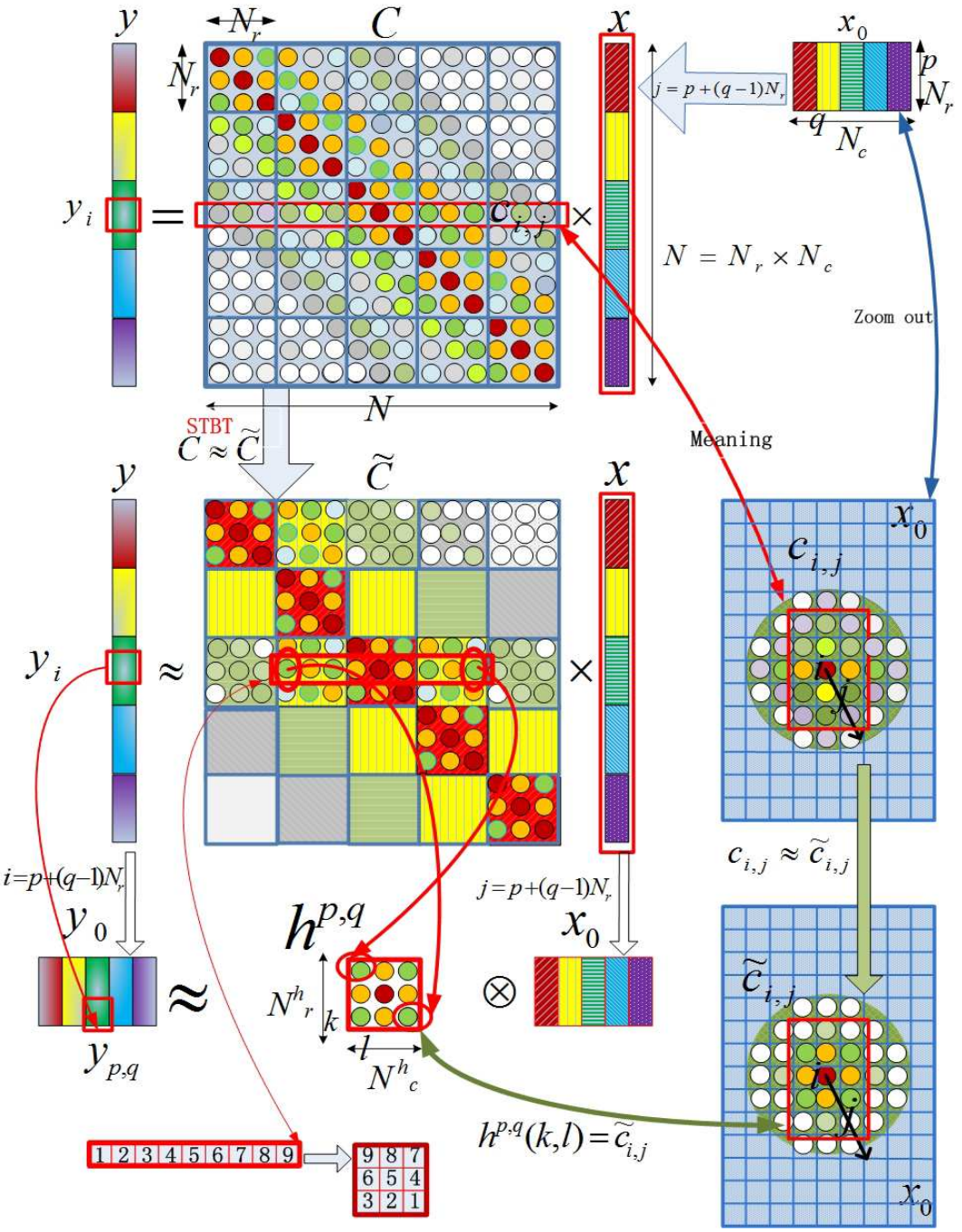

Figure 7.8: Procedure of 2D invariant convolution approximation.

the separable 2D-convolution can greatly reduce both the computational cost and memory requirements of conventional 2D-convolution operation.

In Fig.7.8, we show the whole proposed procedure of 2D invariant convolution approximation for the forward model of power propagation. However, two essential issues for selecting invariant convolution kernel $\mathbf{h} = [h_{k,l}]$ are kernel values $h_{k,l}$ and proper kernel size $N_r \times N_r$, so that the convolution operation can be accurately and efficiently performed. We will discuss these issues on simulations in Section 7.3.

7.3 Simulations

On simulations, we will consider the following aspects:

- Approximation errors between STBT matrix $\tilde{\mathbf{C}}$ and power propagation matrix $\mathbf{C}$;
- Convolution approximated errors for variant, invariant and separable kernels, as well as different kernel sizes and forms (square or rectangular);
- Convolution computational time for different kernels;
- Acoustic imaging results based on 2D invariant convolution model;
- Deconvolution performance of 2D separable kernel based on GPU.
- For the inverse problem of the 2D invariant convolution model in Eq.(7.9), we suppose that convolution approximated errors $\boldsymbol{\eta} = \mathbf{0}$ for simplicity, so that Eq.(7.9) becomes $\mathbf{y} = \mathbf{H}\,\mathbf{x} + \sigma_e^2\,\mathbf{1}_a + \boldsymbol{\xi}$. Therefore, proposed Bayesian JMAP approach in Chapter 6 can be easily applied on it. In the case of $\boldsymbol{\eta} \neq \mathbf{0}$, we will investigate the non-stationary distribution on model errors $\boldsymbol{\epsilon}$ and use the Variational Bayesian Approximation (VBA) to solve Eq.(7.9) in Chapter 8.

The simulation configurations are also the same as the one in Chapter 4.7 as shown in above figure: there are $M = 64$ non-uniform sensors locating on the vertical plane. $d = 2$m is the averaged size of sensor array. $D = 4.50$m is the distance between the sensor plane and source plane. $c_0 \approx 340$m/s is the acoustic speed in the common air. $T = 10000$ is the total number of samplings. For the simulated sources in Fig.7.13(a), there are simulated 4 monopoles and 5 complex sources, spaced at least 20cm from

each other. Original source powers $\mathbf{x}^*$ are within [0.08,2] ([-10.3,3.7]dB) and 14dB dynamic range. And power image size is of $N_c = 27$ and $N_r = 17$ as shown in Fig.7.3(c). The i.i.d AWGN noise power is set $\sigma_e^2 = 0.86$ (-0.7dB), thus the averaged SNR is 0dB.

7.3.1 Approximation errors of STBT matrix

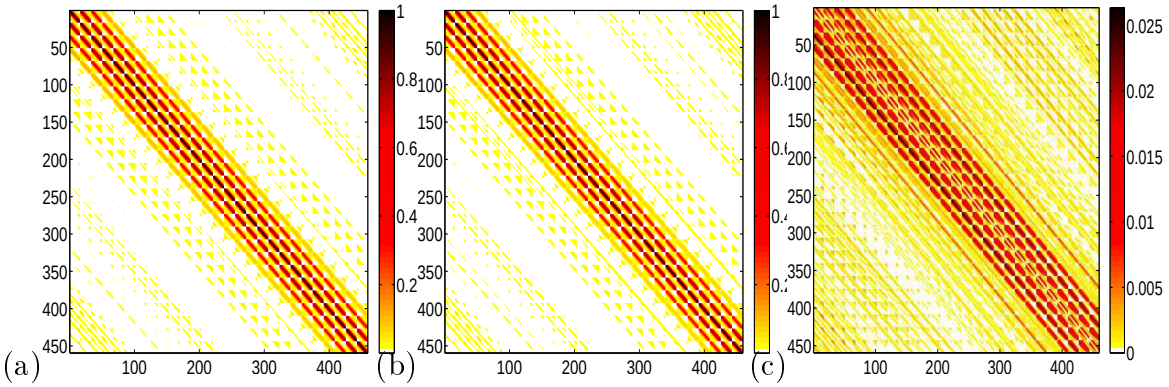

Figure 7.9: Power propagation matrix and its STBT approximation at 2500Hz: (a) $\mathbf{C} = [c_{i,j}]$ (b) $\tilde{\mathbf{C}} = [\tilde{c}_{i,j}]$ (c) Approximation error matrix $|c_{i,j} - \tilde{c}_{i,j}|$

In Fig.7.9, the matrix structures of $\mathbf{C} = [c_{i,j}]$ in Eq.(7.1) and $\tilde{\mathbf{C}} = [\tilde{c}_{i,j}]$ in Eq.(7.6) are quite similar to each other, and the error matrix of STBT approximation is full of relatively small values compared to $\mathbf{C}$. Therefore, the power propagation $\mathbf{C}$ can be effectively approximated by the STBT matrix $\tilde{\mathbf{C}}$.

7.3.2 Convolution approximated errors for different kernels

We define the convolution approximated errors as

$$\delta_y = \frac{\|\mathbf{y} - \hat{\mathbf{y}}\|_2^2}{\|\mathbf{y}\|_2^2} \times 100\% \tag{7.14}$$

where $\mathbf{y}$ refers to the beamforming result of the power propagation model in Eq.(4.12) as shown in Fig.7.3(a); $\hat{\mathbf{y}}$ refers to the convolution results respectively using variant kernels in Eq.(7.12), invariant kernel Eq.(7.9) and separable kernel in Eq.(D.5).

In Fig.7.10, we show convolution approximated errors δ_y versus various kernel sizes. We examine 7 types of kernels with different forms. Firstly, both the variant and invariant kernel with the largest size of 53×33 obtain the very small convolution errors, which validates our proposed (in)variant

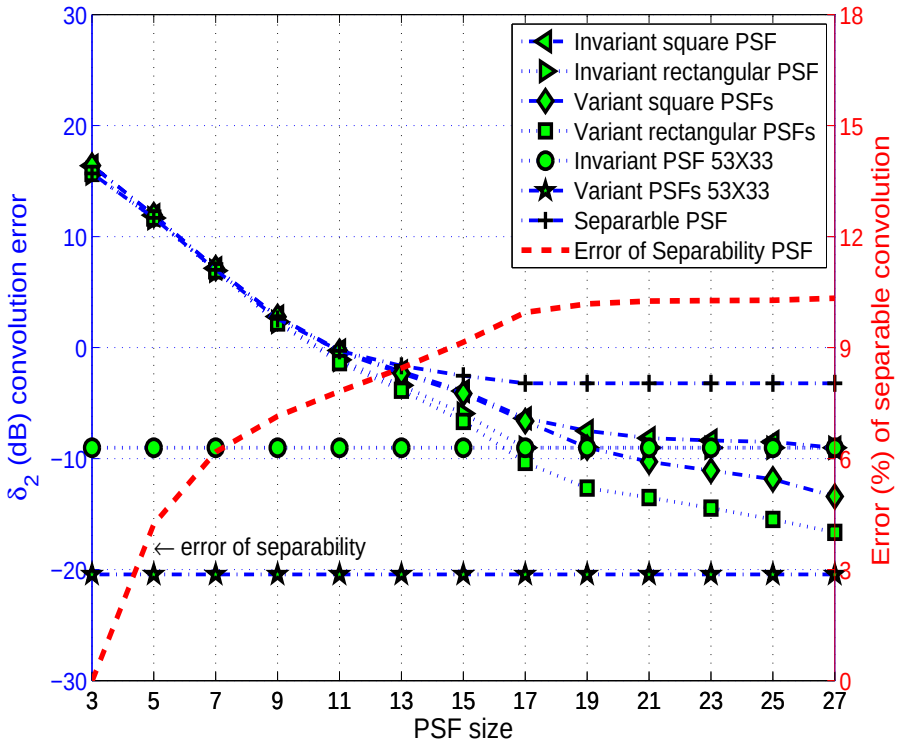

Figure 7.10: Convolution performance comparisons among variant, invariant and separable convolution kernels at 2500Hz

convolution models in Eq.(7.11) and Eq.(7.9). Secondly, the larger kernel size is, the smaller δ_y becomes, and both the square and rectangular kernels obtain similar δ_y for each case, so that we can choose the square kernel for simplicity. Thirdly, the invariant kernel obtains as small δ_y as those of variant kernels, so that we can use invariant model to effectively approximate power propagation model in Eq.(4.46). Fourthly, when kernel size N_h approaches source power image $N_r = 17$, all δ_y of 7 kernels becomes small and remains stable. Finally, the separable convolution with kernel size $N_r = 17$ also gets as similar δ_y as that of invariant kernel, meanwhile the separability error is less than 9%, so that we can use separable convolution to further approximate the invariant convolution for further acceleration.

7.3.3 Convolution computational time

In Fig.7.11, we show computation performance comparisons among invariant and separable convolution using CPU and GPU. The size of the source power image is enlarged as 30 times as that in Fig.7.3(c). Compared with CPU, one of the greatest advantages of GPU is the great number of parallel computational cores which contribute much more powerful computation capability than CPU, but the massive data with non-parallel processing can

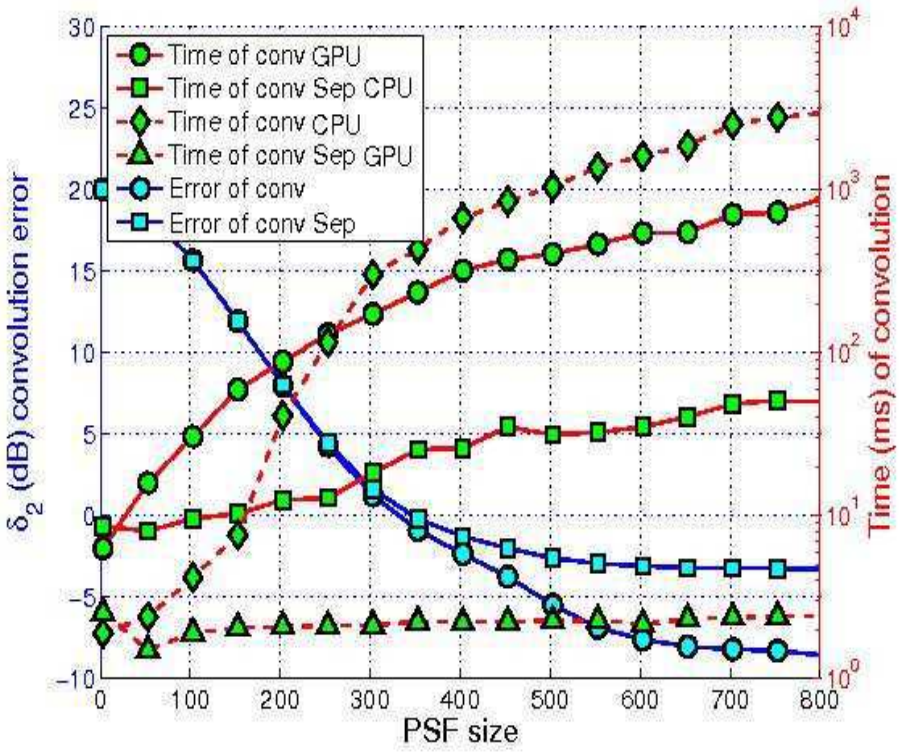

Figure 7.11: Computation performance comparisons at 2500Hz among invariant and separable convolution using CPU (3.33Hz clock) and GPU (Tesla C1060: 240 processing cores, 1.3G Hz clock, 622 GFLOPs (Peak); Using Parallel Computing Toolbox of MATLAB 2012b).

hardly be efficient performed by GPU [47, 100]. The structures of CPU and GPU are shown in Fig.7.12. For the computing time in Fig.7.11, all the red curves go up along with the large kernel size, but the separable convolution based on CPU or GPU keeps a slight increase and maintains the least computation burden, especially when the kernel size is very large. And GPU greatly increases the computation speed. For the convolution approximated errors, both the invariant kernel and separable kernel can obtain relatively small errors when the kernel size is near to the half size of the source power image.

7.3.4 Acoustic imaging via 2D invariant convolution model

In Fig.7.13, the beamforming merely gives some strong source powers. DR-DAMAS too much removes noises and fails to detect weak sources. CLEAN is sensitive to noise interference. Bayesian JMAP method via classical forward model well detects all source powers except for the weakest monopole source.

In Fig.7.13(f), the Bayesian JMAP method via convolution model can

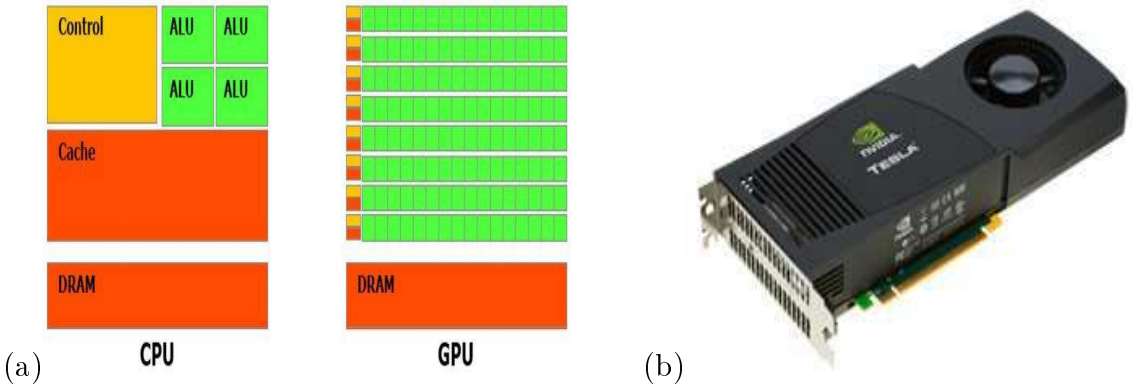

Figure 7.12: (a) Structures of CPU and GPU [47] (b) Used Tesla C1060 GPU: 240 processing cores, 1.3G Hz clock, 622 GFLOPs (Peak).

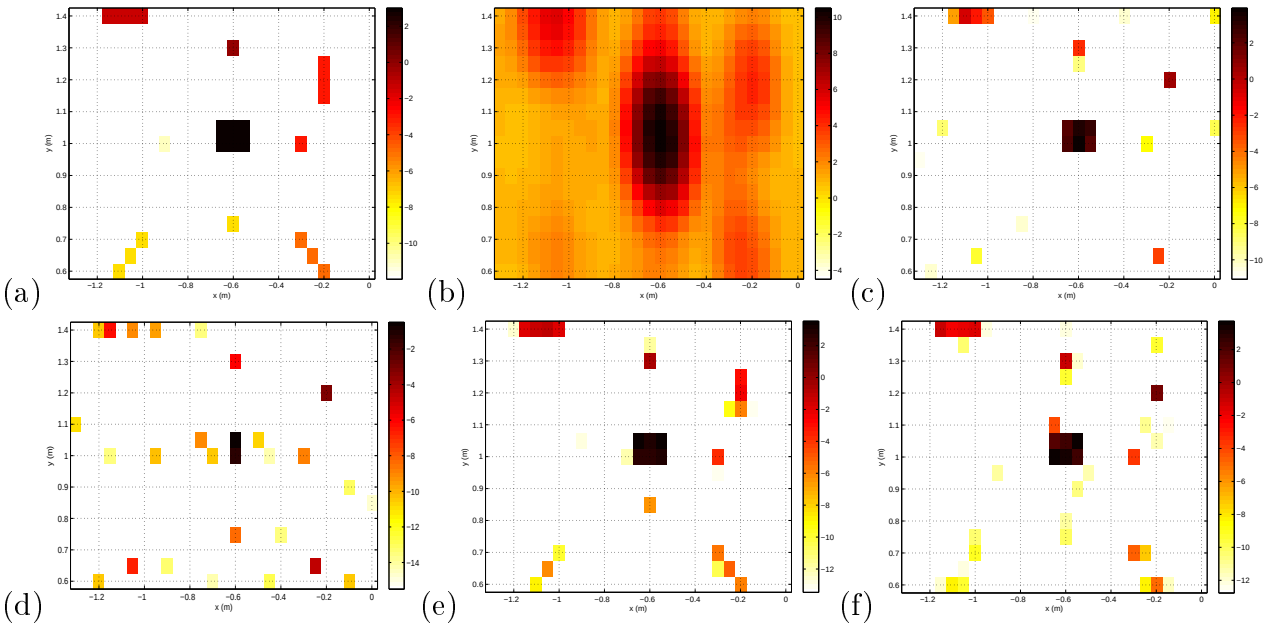

Figure 7.13: Simulations at 2500Hz, 0dB SNR, 15dB display: (a) Source powers (b) Beamforming (c) DR-DAMAS (d) CLEAN (e) Bayesian JMAP method via conventional forward model in Chapter 6 (f) Bayesian JMAP method via invariant convolution model

quickly reconstruct most of the sources, but the recovered source patterns are affected by measured errors.

7.3.5 Deconvolution performance of 2D separable kernel

In Table 7.2, we show the deconvolution performance of 2D separable kernel. In order to make a fair comparison, the Tikhonov regularization method in Eq.(4.58) is used for deconvolution, since the most time-consuming operation mainly depends on the convolution. The separable convolution operation $\mathbf{h}_1 * \mathbf{h}_2^T * \mathbf{x}$ is efficiently computed on GPU Tesla C1060 using the Parallel Computing Toolbox of MATLAB 2012b. In order to show the GPU

Table 7.2: Computational time of Tikhonov deconvolution via separable convolution kernels based on GPU Tesla C1060: 240 processing cores, 1.3G Hz clock, 622 GFLOPs (Peak); Using Parallel Computing Toolbox of MATLAB 2012b. Time results are averaged by 20000 iterations. $\delta_x = \frac{\|\mathbf{x}-\hat{\mathbf{x}}\|_2^2}{\|\mathbf{x}\|_2^2} \times 100\%$

Power image size	17×27	255×405	527×837
Invariant kernel size	13×13	215×215	415×415
Time (ms)/iteration	1.65	7.50	38.3
Deconvolution error δ_x (%)	19.6	29.6	30.6
Speed gain on $\mathbf{C}\,\mathbf{x}$ (CPU)	0.5	314.7	822.2
Speed gain on $\mathbf{h} * \mathbf{x}$ (CPU)	0.27	7.1	9.2
Speed gain on $\mathbf{h}_1 * \mathbf{h}_2^T * \mathbf{x}$ (CPU)	0.16	5.27	7.02

acceleration, we also use the CPU to completely implement the deconvolution methods using matrix multiplication $\mathbf{C}\,\mathbf{x}$, invariant convolution $\mathbf{h} * \mathbf{x}$ and separable convolution $\mathbf{h}_1 * \mathbf{h}_2^T * \mathbf{x}$ respectively. For the computational time, the bigger kernel size is, the greater the computational speed gain is obtained. For the deconvolution errors, though the Tikhonov regularization method does not offer very good results, it still shows the fact that deconvolution via separable convolution model can be efficiently solved based on GPU. In order to obtain a fast and effective deconvolution results, it is highly necessary to optimize the algorithm of proposed Bayesian JMAP based on GPU parallel structure.

However, for the small image size of 17×27, the GPU could not improve computational efficiency compared with CPU. This is because that the 240 cores of GPU Tesla C1060 can efficiently handle the large dimension of matrices, but small matrix cannot be well suited to the parallel structure of GPU. Convolution operation realized by MATLAB Parallel Toolbox could not completely use advantages of the GPU, since it still requires other operations on CPU and causes frequently data transfers between the GPU and CPU. Therefore, it is a promising work to implement the whole deconvolution algorithm (not only the convolution operation) completely based on GPU using the CUDA code library [90]. Moreover, we find out that calculating invariant convolution $\mathbf{h} * \mathbf{x}$ based on GPU merely makes use of about 14% of computational power of GPU, while separable convolution $\mathbf{h}_1 * \mathbf{h}_2^T * \mathbf{x}$ just occupies nearly 7%. So that there will be great potential to develop our own parallel separable convolution algorithm based on the GPU so as to make good use of GPU powerful peak computational capacity.

7.4 Wind tunnel experiments

The above figure shows the configurations of the wind tunnel S2A [86], object vehicle, NUA array and wind refraction. We suppose that all acoustic sources locate on the same plane. This assumption is almost satisfied, because the curvature of the car side is relatively small compared to the distance D=4.5m between the car and array plane. Since the scanning step is set by $\Delta p = 5cm$, the source plane of car side is of 1.5× 5 m^2 (31×101 pixels), and we also focus on a small region of the rearview mirror: 1×1.5 m^2 (21×31 pixels). On the real data, there are T=524288 samplings with the sampling frequency f_s=2.56×10^4 Hz. We separate these samplings into I=204 blocks with L=2560 samplings in each bloc. The working frequency band is chosen as [2400,2600]Hz, which is sensitive to human being. The image results are shown by normalized dB images with 10dB span. For the actual propagation time $\tau_{n,m}$ and distance $r_{n,m}$ in Eq.(3.11), we apply equivalent source to make refraction correction. For $\tau_{-n,m}$ and $r_{-n,m}$ in Eq.(3.11), we use the mirror source signal s_{-n} to correct the ground reflection. The details of the propagation corrections are discussed in A and B.

Figure.7.14 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.7.14(a). DAMAS well deconvolve the beamforming image and discovers sources around the wheels and rearview mirror, however, many false targets are also detected on the air in Fig.7.14(b). DR-DAMAS in Fig.7.14 eliminates most of the artifacts, but it also harms weak sources. Figure.7.14(d) show that CLEAN overcomes main drawbacks of the DAMAS, but unexpected strong points are detected on the ground. CLEAN has to carefully select some parameters for each case. In Fig.7.14(e) and (f), the proposed SC-RDAMAS with sparsity constraint and Bayesian JMAP method via sparse 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.

In Fig.7.14(g), the Bayesian JMAP method via proposed 2D invariant convolution model can achieve the source reconstruction as good as the JMAP via conventional forward model results. In Table 7.3, the computation speed is greatly improved by 2D invariant convolution model in Eq.(7.9) compared to conventional model in Eq.(5.4)

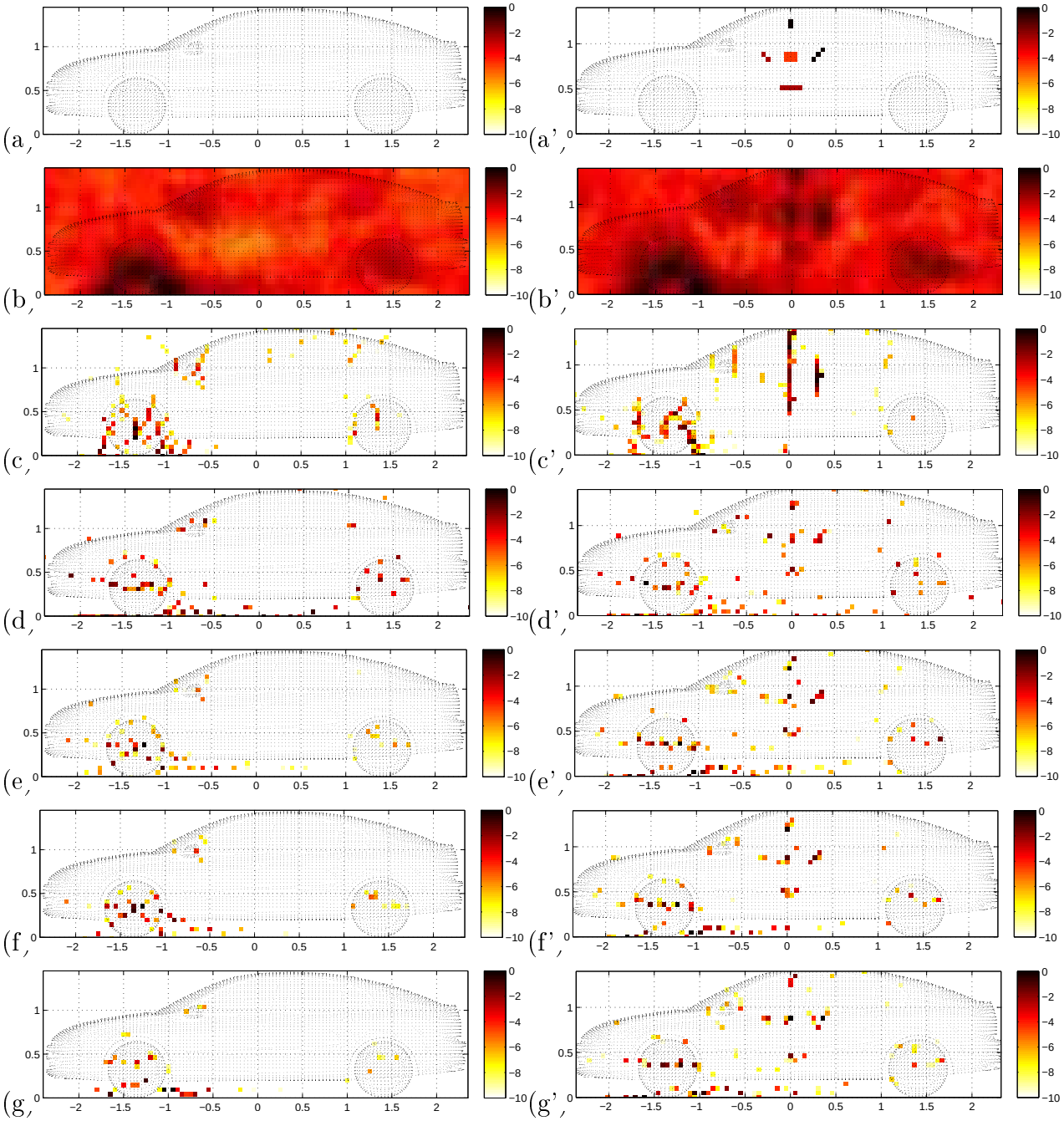

Figure 7.14: Left: real data at 2500Hz (a) vehicle surface (b) Beamforming (c) DAMAS (5000i) (d) CLEAN (e) SC-RDAMAS (f) Bayesian JMAP via classical forward model (g) JMAP via invariant convolution model. Right: hybrid data (a') 5 simulated complex sources (b')-(g') corresponding methods.

Table 7.3: Computational cost for treating real data of whole car: image 31×101 pixels, at 2500Hz, based on CPU: 3.33Hz. 'JMAP+Conv' is short for Bayesian JMAP method via 2D invariant convolution model

Methods	CB	DAMAS	DR-DAMAS	CLEAN	SC-RDAMAS	JMAP	JMAP+Conv
Time (s)	1	10	11	45	852	1012	180

7.5 Conclusions and perspectives

In this chapter, we propose a convolution model in Eq.(7.9) to approximate the forward model of source power propagation in Eq.(5.4), so that proposed Bayesian JMAP method in Chapter 6 is more quickly carried out.

We firstly discuss the 2D-convolution model using a variant kernel in Eq.(7.13), invariant kernel in Eq.(7.6) and separable convolution kernel in Eq.(D.5) respectively. Both the variant and invariant kernels (size and item values) are derived from the Symmetric Toeplitz Block Toeplitz (STBT) structure of power propagation matrix. Moreover, since the invariant kernel $\mathbf{h}$ is a real and non-negative symmetric matrix, $\mathbf{h}$ is separated into two 1D-convolution kernels as discussed in Appendix D. One of the advantages is that the separable 2D-convolution greatly reduces both computational cost and memory requirements of 2D-convolution operation. And it is well suited to the parallelization on many-core processors such as the GPU.

On simulations, the main conclusions are:

- There are relatively small approximation errors between STBT matrix $\tilde{\mathbf{C}}$ and power propagation matrix $\mathbf{C}$;
- 2D invariant convolution model successfully approximates the power propagation model;
- 2D invariant kernel whose size is just the half of the source power image efficiently performs the 2D convolution model.
- Bayesian JMAP method via 2D invariant convolution model obtains an acceptable imaging result compared to the conventional power propagation model;
- For acoustic image with very large size, deconvolution method via 2D separable convolution is very fast implemented using GPU. Compared

to CPU with matrix multiplication $\mathbf{C}\,\mathbf{x}$, the 2D separable convolution using GPU achieves the speed gain as high as three order of magnitude; compared to CPU with 2D non-separable convolution $\mathbf{h} * \mathbf{x}$, it gains two order of magnitude of speed gain; compared to GPU with separable convolution $\mathbf{h}_1 * \mathbf{h}_2^T * \mathbf{x}$, it still obtains nearly one order of magnitude of speed gain.

On real data and hybrid data, we demonstrate that using 2D invariant convolution model can greatly accelerate the Bayesian JMAP method and contribute a rapid implementation for industry application.

In Table 7.4, we give a brief summary for the advantages and drawbacks of mentioned classic methods, proposed SC-RDAMAS in Chapter 5, proposed Bayesian JMAP approach in Chapter 6, as well as the Bayesian JMAP approach via proposed 2D invariant convolution model in this Chapter.

Table 7.4: General performance of classical methods and proposed Bayesian JMAP approach. 'JMAP+Conv' is short for the Bayesian JMAP method via the 2D invariant convolution model.

Methods	CBF	CLEAN	DAMAS	SC-DAMAS	CMF	SC-RDAMAS	JMAP	JMAP+Conv
Resolutions	Low	Normal	Normal	High	Higher	Higher	Higher	High
Dynamic Range	Narrow	Normal	Normal	Normal	Wide	Wide	Wide	Normal
Noise	Robust	Sensitive	Sensitive	Sensitive	Robust	Robust	Robust	Robust
Computation	Least	Normal	Normal	High	High	High	Higher	Normal
Samples	Normal	Normal	Normal	Normal	More	Normal	Normal	Normal
Source number	No	Required	No	Required	Required	Required	No	No

However, there are at least three aspects to be further improved:

- The acoustic image quality using the 2D invariant or separable convolution model should be carefully refined, and we have to make a necessary balance between source reconstruction and computational cost, especially for real data treatment of wind tunnel experiments.

- To improve the source estimation results, it is highly necessary to make good use of the sparse distribution of source powers. In Chapter 8, we will investigate a more proper sparsity enforcing prior than the Double Exponential model used in Chapter 6. Furthermore, the model error $\boldsymbol{\epsilon}$ in the proposed convolution model might not be always Gaussian white noise distribution, but probably the spatially non-stationary Gaussian distribution on the different parts of source plane or on the different microphone sensors. This point will be also discussed in Chapter 8.

- For GPU implementation on the large scale of real data in tunnel experiments, it is quite worthy of optimizing advanced deconvolution algorithms (such as the Bayesian inference) via separable convolution model, so that the peak-power computation of GPU can be well utilized as possible as we can, and the drawbacks of limited local on-chip memory could be avoided to some extend. Instead of using MATALAB Parallel Toolbox, it is a promising work to implement the whole deconvolution algorithm (including 2D convolution) mainly based on GPU using the CUDA code library [90]

.

- Dans le monde il n'y a qu'une vérité : être fidèle à la vie et l'aimer.

Romain Rolland

8

Variational Bayesian Approximation Approach with Students-t prior for acoustic imaging in non-stationary noises

As discussed in Chapter 6, the proposed Bayesian JMAP approach in Eq.(6.8) suffers from tremendous computational cost and non-quadratic optimization. The first drawback can be solved by using a 2D invariant convolution approximation in Chapter 7. However, there are still three difficulties: proposed convolution approximation needs to balance the acoustic imaging quality and computational cost. Furthermore, the model errors $\boldsymbol{\epsilon}$ in the proposed approximated convolution model are no more spatially invariant but non-stationary on the different parts of source plane. And the sparse distribution of source powers $\mathbf{x} = [x_1, \cdots, x_N]^T$ modeled by the Double Exponential (DE) prior in Eq.(6.6) are improperly assumed to be the same scale parameter β which controls the same dynamic range of different x_n with $n \in [1, \cdots, N]$. In fact, this assumption may not be always satisfied in practice. For example, the source power on the rearview mirror probably has different dynamic range from the one on the wheels.

Motivated by the above three remarks, in Chapter 8, we propose the following changes in the modeling of priors on the noise ϵ and the unknown source powers $\mathbf{x}$ in the convolution forward model $\mathbf{y} = \mathbf{H}\,\mathbf{x} + \boldsymbol{\epsilon}$:

- Model error $\epsilon_n \in \boldsymbol{\epsilon}$ with $n \in [1, \cdots, N]$ is no longer assumed spa-

tially invariant. However, we still model it by a non-stationary Gaussian prior $p(\epsilon_n|\nu_n) = \mathcal{N}(\epsilon_n|0, \nu_n^{-1})$ with $\nu_n^{-1} = \mathbf{E}[\epsilon_n^2]$ denotes the variance of ϵ_n. For spatially invariant distribution, we suppose that ν_n depend on different position n and is not a constant. Furthermore, we assign the unknown variance ν_n with Gamma distribution as $p(\nu_n|, a_\nu, b_\nu) = \mathcal{G}(\nu_n|, a_\nu, b_\nu)$. This means that ϵ_n is modeled to be the Student-t prior. This is owing to the fact that

$$St(\epsilon_n|\alpha_\epsilon) = \int \mathcal{N}(\epsilon_n|0, \nu_n^{-1})\, \mathcal{G}(\nu_n|a_\nu, b_\nu)\, d\nu\,, \tag{8.1}$$

where α_ϵ denotes the degree of freedom, which controls the shape of $St(\epsilon_n|\alpha_\epsilon)$ distribution; $a_\nu = b_\nu = \frac{\alpha_\epsilon}{2}$; and ν_n denotes the inverse variance of ϵ_n. Indeed, ν_n is a hidden variable which can interpret the Student-t prior by the Normal-Gamma decomposition Owing to ν_n, this interpretation gives us the possibility to propose a hierarchical prior model for model error ϵ_n as given in Eq.(8.1). For all model errors $\boldsymbol{\epsilon} = [\epsilon_1, \cdots, \epsilon_N]^T$, their different variances $\boldsymbol{\nu}^{-1} = [\nu_1^{-1}, \cdots, \nu_N^{-1}]^T$ can model the spatially variant distribution.

Therefore, we use Student-t prior $St(\boldsymbol{\epsilon}|\alpha_\epsilon)$ to model a spatial non-stationary distribution of model errors $\boldsymbol{\epsilon}$.

- For source power $x_n \in \mathbf{x}$, we also use the similar hierarchical prior model as

$$St(x_n|\alpha_x) = \int \mathcal{N}(x_n|0, \gamma_n^{-1})\, \mathcal{G}(\gamma_n|a_\gamma, b_\gamma)\, d\gamma\,, \tag{8.2}$$

where

  - α_x denotes the degree of freedom, which controls the shape of $St(x_n|\alpha_x)$ distribution; $a_\gamma = b_\gamma = \frac{\alpha_x}{2}$. Knowing also that $\alpha_x = 1$ results to Cauchy prior, which can model a sparsity enforcing probability law with narrow summit and heavy tail; while $\alpha_x = \infty$ refers to Gaussian distribution; $1 < \alpha_x < \infty$ still remains the sparsity to some extend.
  - $\gamma_n^{-1} = \mathbf{E}[x_n^2]$ denotes the variance of x_n. For spatially invariant distribution, we suppose γ_n with $n \in [1, \cdots, N]$ is not a constant. For all $\mathbf{x} = [x_1, \cdots, x_N]^T$, their different variances $\boldsymbol{\gamma}^{-1} = [\gamma_1^{-1}, \cdots, \gamma_N^{-1}]^T$ model the spatially variant distribution.
  - Student-t distribution is a continuous function over all real values, while the DE used in Chapter 6 and Laplace are discontinuous at $x_n = 0$ point.

Therefore, using Student-t prior $St(\mathbf{x}|\alpha_x)$ in Eq.(8.2) as a prior model is a more suitable sparsity enforcing distribution of source powers $\mathbf{x}$.

- VBA estimation. Using the hierarchical prior models on $\boldsymbol{\epsilon}$ and $\mathbf{x}$ causes much more complexity for the estimations of $\mathbf{x}$ and other parameters based on JMAP estimation. This is because the joint posterior $p(\mathbf{x}, \boldsymbol{\gamma}, \boldsymbol{\nu}|\mathbf{y})$ of all the unknown quantities becomes more sophisticated, since hidden variables $\boldsymbol{\gamma}$ and $\boldsymbol{\nu}$ are both N-dimensional. An alternate optimization algorithm respectively with respect to $\mathbf{x}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\nu}$ may not converge. So that we need to use other efficient methods than the JMAP estimation. For this, we propose to use the Variational Bayesian Approximation (VBA) method which is a new way for doing more effective and more precise use of the joint posterior probability law of all the unknowns.

This chapter is organized as follows: Section 8.1 and 8.2 presents the non-stationary prior of model errors and more suitable sparsity enforcing prior of source powers respectively. Section 8.3 introduces the main principles of Variational Bayesian Approximation (VBA) such as VBA estimation and computational complexity. Section 8.4 and 8.5 validate the VBA approach via 2D invariant convolution model on simulations and real data respectively. Finally Section 8.6 concludes this chapter.

8.1 Non-stationary Gaussian prior of model errors

We reconsider the model errors in wind tunnel experiments. During the propagation from Acoustic sources to sensor array, the signals are inevitably deteriorated by complex unknown reverberations caused by multi-path propagation, inherent noises at microphone sensors, background noises in wind tunnel, as well as the model approximated errors in signal processing techniques. Therefore, model errors are the mixed effects of the above mentioned factors. Instead of the ideal i.i.d AGWN assumption in previous chapters, we use Student-t prior $St(\boldsymbol{\epsilon}|\alpha_\epsilon)$ to model a spatial non-stationary distribution of model errors $\boldsymbol{\epsilon}$ in Eq.(8.1).

Moreover, we suppose $\boldsymbol{\epsilon} = [\epsilon_1, \cdots, \epsilon_N]^T$ are mutually independent. We then get multivariate prior of model errors $\boldsymbol{\epsilon}$ as

$$St(\boldsymbol{\epsilon}|\alpha_\epsilon) = \prod_{n=1}^{N} St(\epsilon_n|\alpha_\epsilon)\,, \tag{8.3}$$

where

$$St(\epsilon_n|\alpha_\epsilon) = \frac{\Gamma(\frac{1+\alpha_\epsilon}{2})}{\sqrt{\alpha_\epsilon\,\pi}\,\Gamma(\frac{\alpha_\epsilon}{2})}\,(1+\frac{\epsilon_n^2}{\alpha_\epsilon})^{-\frac{1+\alpha_\epsilon}{2}} \tag{8.4}$$

where $\Gamma(\cdot)$ denotes the Gamma function, defined as $\Gamma(x) = \int_0^\infty t^{x-1}\,e^{-t}\,d\,t$ for any variable x; and $\Gamma(0) = \infty$, $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ and $\Gamma(1) = 1$.

We then use a hidden variable vector $\boldsymbol{\nu} = [\nu_1, \cdots, \nu_N]^T$ to interpret Eq.(8.3) multivariate $St(\boldsymbol{\epsilon}|\alpha_\epsilon)$. So that $St(\boldsymbol{\epsilon}|\alpha_\epsilon)$ can be generated by marginalizing $\boldsymbol{\nu}$ as

$$St(\boldsymbol{\epsilon}|\alpha_\epsilon) = \int_{\boldsymbol{\nu}} p(\boldsymbol{\epsilon},\boldsymbol{\nu}|\alpha_\epsilon)\,d\boldsymbol{\nu} = \int_{\boldsymbol{\nu}} p(\boldsymbol{\epsilon}|\boldsymbol{\nu})p(\boldsymbol{\nu}|\boldsymbol{\alpha}_\nu)\,d\boldsymbol{\nu}\,. \tag{8.5}$$

According to the Normal-Gamma decomposition of Student-t distribution $St(\epsilon_n|\alpha_\epsilon)$ in Eq.(8.1), we can get from Eq.(8.5) as

$$p(\boldsymbol{\epsilon}|\boldsymbol{\nu}) = \mathcal{N}(\boldsymbol{\epsilon}|\mathbf{0},\,\boldsymbol{\Sigma}_\nu^{-1})\,,\quad p(\boldsymbol{\nu}|\boldsymbol{\alpha}_\nu) = \prod_{n=1}^{N}\mathcal{G}(\nu_n|a_\nu,b_\nu)\,, \tag{8.6}$$

where $\mathbf{0} = [0]_N$ denotes N-dimension 0 value vector; $\boldsymbol{\alpha}_\nu = [a_\nu, b_\nu]^T$ denotes the hyperparameter of $p(\boldsymbol{\nu}|\boldsymbol{\alpha}_\nu)$; $\boldsymbol{\Sigma}_\nu^{-1} = Diag[\nu_n^{-1}]$ with $n \in [1, N]$ denotes the covariance matrix, which is a diagonal matrix whose diagonal items are supposed to be non-constant; and $\mathcal{G}(\nu_n|a_\nu,b_\nu)$ denotes the Gamma distribution, defined as

$$\mathcal{G}(\nu_n|a_\nu,b_\nu) = \prod_{n=1}^{N} \mathit{\Gamma}(a_\nu)^{-1}\,(b_\nu)^{a_\nu}\,\nu_n^{a_\nu-1}e^{-b_\nu\nu_n} \tag{8.7}$$

where a_ν denotes the shape parameter and b_ν denotes rate parameter of $\mathcal{G}(\nu_n|a_\nu,b_\nu)$. For $St(\boldsymbol{\epsilon}|\alpha_\epsilon)$ in Eq.(8.5), we have $a_\nu = b_\nu = \frac{\alpha_\epsilon}{2}$.

According to proposed convolution forward model $\mathbf{y} = \mathbf{H}\,\mathbf{x}+\boldsymbol{\epsilon}$ of Eq.(7.9), the likelihood $p(\mathbf{y}|\mathbf{x},\boldsymbol{\nu})$ is determined by the conditional prior $p(\boldsymbol{\epsilon}|\boldsymbol{\nu}) = \mathcal{N}(\boldsymbol{\epsilon}|\mathbf{0},\,\boldsymbol{\Sigma}_\nu^{-1})$ in Eq.(8.6) as:

$$p(\mathbf{y}|\mathbf{x},\boldsymbol{\nu}) = \mathcal{N}(\boldsymbol{\epsilon}|\mathbf{y}-\mathbf{H}\,\mathbf{x},\,\boldsymbol{\Sigma}_\nu^{-1}) = \frac{|\boldsymbol{\Sigma}_\nu|^{1/2}}{(2\pi)^{N/2}}\,e^{-\frac{1}{2}(\mathbf{y}-\mathbf{H}\,\mathbf{x})^\dagger\,\boldsymbol{\Sigma}_\nu\,(\mathbf{y}-\mathbf{H}\,\mathbf{x})}\,, \tag{8.8}$$

covariance matrix $\boldsymbol{\Sigma}_\nu^{-1} = Diag[\nu_n^{-1}]$ with $n \in [1, N]$ is a diagonal matrix whose diagonal items are supposed to be non-constant.

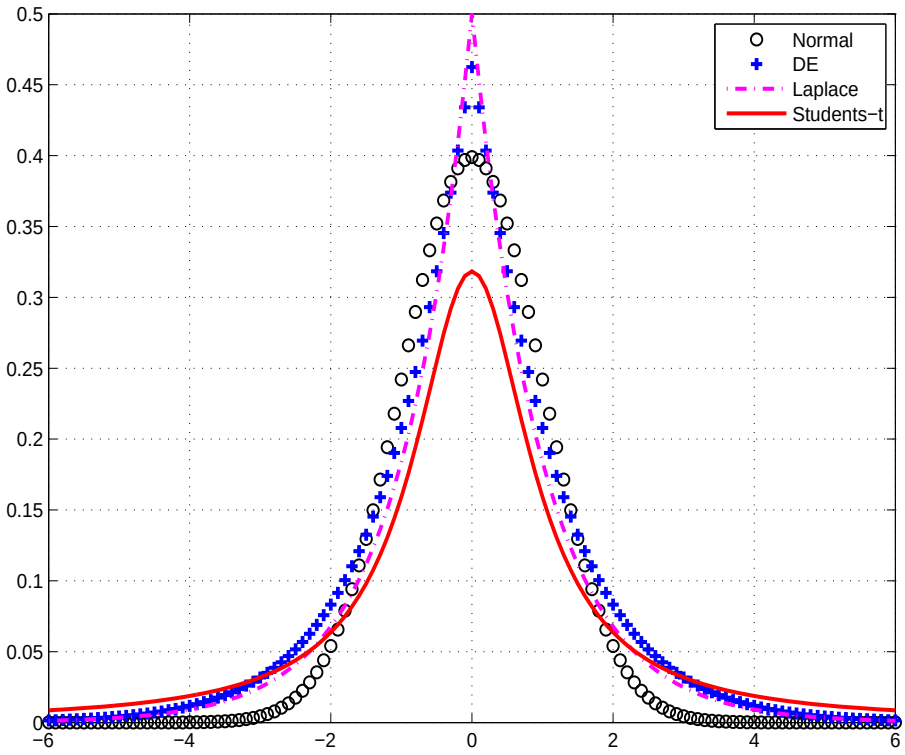

Figure 8.1: Sparse priors modeled by Gaussian normal (Normal), Laplace, Double Exponential (DE) and Student-t.

8.2 More suitable sparsity enforcing prior on source powers

As discussed in Chapter 6, 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. Suppose that there are just K acoustic source on the car surface. After source plane discretization by N grids with $N >> K$ provided, the discrete source powers $\mathbf{x} = [x_1, \cdots, x_N]^T$ become the K-sparsity signal whose K items are not zero, while other $N-K$ items are zero values.

Such a sparsity prior knowledge can be interpreted by a sparse distribution which has a very high value around the original zero (sparsity) and a long heavy tail (dynamic range of source powers). In Fig.8.1, we show several typical sparse priors modeled by Gaussian normal (Normal), Laplace, DE used in Chapter 6 and Student-t distribution respectively. It is seen that the Gaussian distribution cannot well model the sparsity since its tail attenuates very quickly along the big values. The Laplace and DE distribution can well reflect the sparsity, but they are both discontinuous functions at the zero

point, and this discontinuity inevitably causes problem in optimization. The Student-t distribution has three advantages:

- it can achieve the best sparsity enforcing effect among them, since it has the heaviest tails and relatively narrow summit around zero value;
- it is a continuous function for all values of source powers, and this avoids the discontinuous problem in DE and Laplace distribution;
- hidden variables $\boldsymbol{\gamma} = [\gamma_1, \cdots, \gamma_N]$ in Student-t prior can easily model the different variance of source powers at different positions, while the DE prior in Eq.(6.6) is improperly assumed to use the same scale parameter β for all the source powers.

Therefore, we choose the Student-t priors [115, 87] to properly enforce the sparsity and wide dynamic range of source power distribution which is spatially variant:

$$St(\mathbf{x}|\alpha_x) = \prod_{n=1}^{N} St(x_n|\alpha_x)\,. \tag{8.9}$$

Owing to the Normal-Gamma decomposition of Student-t distribution $St(x_n|\alpha_x)$ in Eq.(8.2), we use the hidden variables $\boldsymbol{\gamma} = [\gamma_1, \cdots, \gamma_N]$ to interpret $St(\mathbf{x}|\alpha_x) = \int_{\gamma} p(\mathbf{x}|\boldsymbol{\gamma})\, p(\boldsymbol{\gamma}|\boldsymbol{\alpha}_\gamma)\, d\boldsymbol{\gamma}$ as

$$p(\mathbf{x}|\boldsymbol{\gamma}) = \mathcal{N}(\mathbf{x}|\mathbf{0},\, \boldsymbol{\Sigma}_\gamma^{-1})\,, \quad p(\boldsymbol{\gamma}|\boldsymbol{\alpha}_\gamma) = \prod_{n=1}^{N} \mathcal{G}(\gamma_n|a_\gamma, b_\gamma)\,, \tag{8.10}$$

where $\boldsymbol{\alpha}_\gamma = [a_\gamma, b_\gamma]^T$ denotes the hyperparameter of $p(\boldsymbol{\gamma}|\boldsymbol{\alpha}_\gamma)$; covariance matrix $\boldsymbol{\Sigma}_\gamma^{-1} = Diag[\gamma_n^{-1}]$ with $n \in [1, N]$ is a diagonal matrix whose diagonal items are supposed to be non-constant; a_γ denotes the shape parameter and b_γ denotes rate parameter of $\mathcal{G}(\gamma_n|a_\gamma, b_\gamma)$. For $St(\boldsymbol{\gamma}|\alpha_x)$ in Eq.(8.5), we have $a_\gamma = b_\gamma = \frac{\alpha_x}{2}$.

8.3 Introduction of Variational Bayesian Approximation (VBA)

The ill-posed inverse problem in the convolution forward model $\mathbf{y} = \mathbf{H}\,\mathbf{x} + \boldsymbol{\epsilon}$ of Eq.(7.9) can be effectively solved by the Bayesian inference approaches [95, 41, 84, 87]. As stated in Chapter 6, Bayesian inference aims to estimate the unknown variables and model parameters by applying the Bayes'

rule to update the probability laws: the posterior probability $p(\mathbf{x}, \boldsymbol{\theta}|\mathbf{y})$ is derived from the likelihood $p(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta})$ and prior probability $p(\mathbf{x}, \boldsymbol{\theta})$, in which, the previous one can be derived from the known data $\mathbf{y}$ and forward model in Eq.(7.9). And the latter can be imposed properly to unknown variables, such as sparse priors. Compared with the deterministic regularization methods in Chapter 4 and proposed SC-RDAMAS approach in Chapter 5, the imposed priors can bring in novel information on the physical characteristics of unknown variables, which helps to reduce the uncertainty caused by the ill-posed inverse problem.

As discussed in Chapter 6, the Bayesian inference via JMAP estimation was expressed as:

$$\begin{cases} \mathbf{y} = \mathbf{H}\,\mathbf{x} + \boldsymbol{\epsilon} \\ (\hat{\mathbf{x}}, \hat{\boldsymbol{\theta}})_{JMAP} = \arg\min\limits_{(\mathbf{x},\boldsymbol{\theta})} \{p(\mathbf{x}, \boldsymbol{\theta}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta})\, p(\mathbf{x}, \boldsymbol{\theta})\} \end{cases}, \tag{8.11}$$

where

- $\boldsymbol{\theta}$ represent all the unknown parameters, or other parameters than $\mathbf{x}$. If we use the Gaussian distribution on model errors $\epsilon \sim \mathcal{N}(\epsilon|\mathbf{0}, \sigma_\epsilon^2\,\mathbf{I})$ and the DE prior on source power $\mathbf{x} \sim \mathcal{DE}(\mathbf{x}|\gamma, \beta)$, so that we get $\boldsymbol{\theta} = [\sigma_\epsilon^2, \gamma, \beta]$ which is 3-dimensional. If we use the Student-t priors on both $\boldsymbol{\epsilon}$ and $\mathbf{x}$ in Eq.(8.6) and Eq.(8.10), we then get $\boldsymbol{\theta} = [\boldsymbol{\gamma}, \boldsymbol{\nu}]$ which includes as large as 2N dimension of unknown hidden variables.
- According to the Student-t priors $St(\mathbf{x}|\alpha_x)$ in Eq.(8.10) and $St(\boldsymbol{\epsilon}|\alpha_\epsilon)$ in Eq.(8.6), as well as the likelihood $p(\mathbf{y}|\mathbf{x}, \boldsymbol{\nu})$ in Eq.(8.8), we obtain the JMAP estimation in Eq.(8.11) as

$$(\hat{\mathbf{x}}, \hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\nu}})_{JMAP} = \arg\max_{(\mathbf{x},\boldsymbol{\gamma},\boldsymbol{\nu})} \{p(\mathbf{x}, \boldsymbol{\nu}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x}, \boldsymbol{\gamma}, \boldsymbol{\nu})\, p(\mathbf{x}|\boldsymbol{\gamma})\, p(\boldsymbol{\gamma})\, p(\boldsymbol{\nu})\}, \tag{8.12}$$

  where joint posterior $p(\mathbf{x}, \boldsymbol{\theta}|\mathbf{y})$ can be expressed finite combination of multivariate Gaussian and Gamma distributions as:

$$\begin{aligned} p(\mathbf{x}, \boldsymbol{\gamma}, \boldsymbol{\nu}|\mathbf{y}) \propto & \mathcal{N}(\mathbf{y}|\mathbf{H}\,\mathbf{x},\, \boldsymbol{\Sigma}_\nu^{-1}) \prod_{n=1}^{N} \mathcal{G}(\nu_n|a_\nu, b_\nu) \\ & \mathcal{N}(\mathbf{x}|\mathbf{0},\, \boldsymbol{\Sigma}_\gamma^{-1}) \prod_{n=1}^{N} \mathcal{G}(\gamma_n|a_\gamma, b_\gamma) \end{aligned}. \tag{8.13}$$

However, in JMAP estimation in Eq.(8.12), it is very hard to estimate N unknown source powers $\mathbf{x}$ and 2N dimensional hidden variables $\boldsymbol{\theta} = [\boldsymbol{\gamma}, \boldsymbol{\nu}]$.

In Chapter 6, proposed Bayesian JMAP via DE prior could work well since we had $\mathbf{x}$ with only two parameters such background noise power σ_e^2 and scale parameter γ in DE prior ($\beta = 1$ was fixed).

For these reasons, we propose here a better way to use the joint posterior distribution. These above limitations can be overcome by the variational Bayesian approach (VBA) [66, 106, 115, 87]. In VBA, the posterior probability $p(\mathbf{x}, \boldsymbol{\theta}|\mathbf{y})$ is approximated by a family of basic easily handled probability distributions $q(\mathbf{x}, \boldsymbol{\theta})$, which can be estimated by minimizing the Kullback-Leibler (KL) divergence as

$$\begin{cases} \hat{q}(\mathbf{x}, \boldsymbol{\theta}) = \arg \min\limits_{q(\mathbf{x}, \boldsymbol{\theta})} \{KL(q:p)\} \\ KL(q:p) = \int q(\mathbf{x}, \boldsymbol{\theta}) \, \dfrac{q(\mathbf{x}, \boldsymbol{\theta})}{p(\mathbf{x}, \boldsymbol{\theta}|\mathbf{y})} \, d(\mathbf{x}, \boldsymbol{\theta}) \end{cases}, \tag{8.14}$$

where $KL(q:p)$ divergence can measure the distance between $q(\mathbf{x}, \boldsymbol{\theta})$ and $p(\mathbf{x}, \boldsymbol{\theta}|\mathbf{y})$.

In Eq.(8.14), minimizing the KL divergence is equivalent to maximizing the variational bound (free energy) $\mathfrak{L}(q)$ [66] as

$$\begin{cases} \hat{q}(\mathbf{x}, \boldsymbol{\theta}) = \arg \max\limits_{q(\mathbf{x}, \boldsymbol{\theta})} \{\mathfrak{L}(q)\} \\ \mathfrak{L}(q) = \int q(\mathbf{x}, \boldsymbol{\theta}) \, \ln \dfrac{p(\mathbf{x}, \boldsymbol{\theta}, \mathbf{y})}{q(\mathbf{x}, \boldsymbol{\theta})} \, d(\mathbf{x}, \boldsymbol{\theta}) \end{cases}, \tag{8.15}$$

where $p(\mathbf{x}, \boldsymbol{\theta}, \mathbf{y})$ denotes the total joint probability. To maximize the variational bound in Eq.(8.15), one of the simplest propositions is to suppose $\mathbf{x}, \boldsymbol{\theta}$ to be totally independent as

$$q(\mathbf{x}, \boldsymbol{\theta}) = q_1(\mathbf{x}) q_2(\boldsymbol{\theta}) = \prod_i q_{1i}(x_i) \prod_j q_{2j}(\theta_j) \tag{8.16}$$

where $x_i \in \mathbf{x}$, $\theta_j \in \boldsymbol{\theta}$. Then the simple probability functions in Eq.(8.16) can be obtained by the mean field approximation [124] as

$$\begin{cases} \hat{q}_{1i}(x_i) = \dfrac{\exp I(x_i)}{\int \exp I(x_i) \, d\, x_i} \\ I(x_i) = - < \ln p(\mathbf{y}, \mathbf{x}, \boldsymbol{\theta}) >_{q_1(\mathbf{x}_{-i}) \, q_2(\boldsymbol{\theta})} \end{cases}, \tag{8.17}$$

where $I(x_i)$ denotes the partition function for probability density function (PDF) $p(x_i)$; operator $< p(x,t) >_{q(t)} = \int q(t) \, p(x,t) \, d\, t$ denotes the integration for joint PDF $p(x,t)$, and $\mathbf{x}_{-i}$ denotes the parameter vector except the item x_i.

In Eq.(8.17), an analytical expression of $I(x_i)$ can be efficiently obtained with the help of the conjugate priors [119]. For example $\boldsymbol{\theta} = [\boldsymbol{\gamma}, \boldsymbol{\nu}]$, we call the conditional prior $p(\mathbf{x}|\boldsymbol{\gamma})$ as the 'conjugate prior' of the likelihood $p(\mathbf{y}|\mathbf{x}, \boldsymbol{\nu})$, the reason is that the approximated posterior $\hat{q}_1(\mathbf{x})$ can be derived from the same family as $p(\mathbf{x}|\boldsymbol{\gamma})$. Similarly, $\hat{q}_2(\boldsymbol{\gamma})$ can come from the same family as prior $p(\boldsymbol{\gamma})$, if $p(\boldsymbol{\gamma})$ is the conjugate prior of $p(\mathbf{x}|\boldsymbol{\gamma})$; and $\hat{q}_3(\boldsymbol{\nu})$ can come from the same family as prior $p(\boldsymbol{\nu})$, if $p(\boldsymbol{\nu})$ is the conjugate prior of likelihood $p(\mathbf{y}|\mathbf{x}, \boldsymbol{\nu})$.

8.3.1 VBA estimations

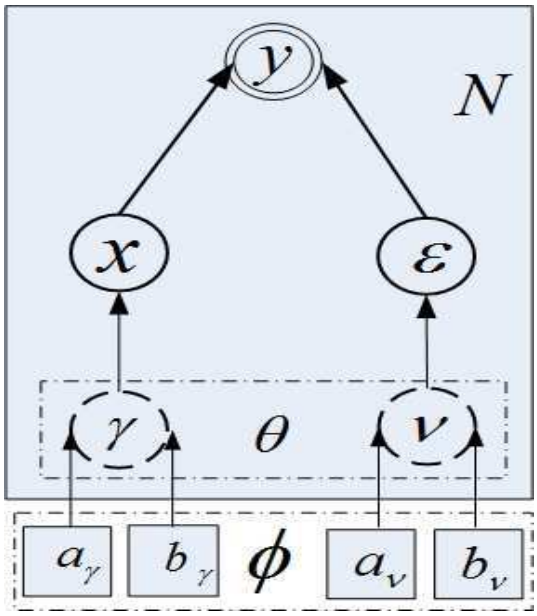

Figure 8.2: Three layers of hierarchical Bayesian Graphical model for N-dimension variable estimations. Double circle: Observed data; Single circle: Unknown variables; Dash circle: Hidden variables; Square: Parameters in hidden variable priors; Arrow: Dependency.

In Fig.8.2, the graphical model describes the dependencies between the observed data $\mathbf{y}$, unknown variables $\mathbf{x}$, their hidden variables $\boldsymbol{\theta} = [\boldsymbol{\gamma}, \boldsymbol{\nu}]^T$ in conditional priors $p(\mathbf{x}|\boldsymbol{\gamma})$, $p(\boldsymbol{\epsilon}|\boldsymbol{\nu})$, and parameters $\boldsymbol{\phi} = [a_\gamma, b_\gamma, a_\nu, b_\nu]^T$ in hidden variable priors.

According to above analysis, we give the VBA estimation framework as

follows:

$$
\begin{cases}
\mathbf{y} = \mathbf{H}\,\mathbf{x} + \boldsymbol{\epsilon} \\
\hat{q}(\mathbf{x}, \boldsymbol{\theta}) = \arg\min\limits_{q(\mathbf{x},\boldsymbol{\theta})} \left\{ KL(q:p) = \int q(\mathbf{x}, \boldsymbol{\theta}) \, \dfrac{q(\mathbf{x}, \boldsymbol{\theta})}{p(\mathbf{x}, \boldsymbol{\theta}|\mathbf{y})} \, d(\mathbf{x}, \boldsymbol{\theta}) \right\} \\
St(\mathbf{x}) = \int p(\mathbf{x}|\boldsymbol{\gamma}) \, p(\boldsymbol{\gamma}) \, d\boldsymbol{\gamma} = \int \mathcal{N}(\mathbf{x}|\mathbf{0}, \boldsymbol{\Sigma}_{\gamma}^{-1}) \, \mathcal{G}(\boldsymbol{\gamma}|a_{\gamma}, b_{\gamma}) \, d\boldsymbol{\gamma} \\
St(\boldsymbol{\epsilon}) = \int p(\boldsymbol{\epsilon}|\boldsymbol{\nu}) \, p(\boldsymbol{\nu}) \, d\boldsymbol{\nu} = \int \mathcal{N}(\boldsymbol{\epsilon}|\mathbf{0}, \, \boldsymbol{\Sigma}_{\nu}^{-1}) \, \mathcal{G}(\boldsymbol{\nu}|a_{\nu}, b_{\nu}) \, d\boldsymbol{\nu} \\
\boldsymbol{\theta} = [\boldsymbol{\gamma}, \boldsymbol{\nu}]^T \\
p(\mathbf{x}, \boldsymbol{\theta}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta}) \, p(\mathbf{x}|\boldsymbol{\gamma}) \, p(\boldsymbol{\gamma}) \, p(\boldsymbol{\nu}) \\
\qquad\qquad \propto \mathcal{N}(\mathbf{y}|\mathbf{H}\,\mathbf{x}, \, \boldsymbol{\Sigma}_{\nu}^{-1}) \, \mathcal{G}(\boldsymbol{\nu}|a_{\nu}, b_{\nu}) \, \mathcal{N}(\mathbf{x}|\mathbf{0}, \boldsymbol{\Sigma}_{\gamma}^{-1}) \, \mathcal{G}(\boldsymbol{\gamma}|a_{\gamma}, b_{\gamma}) \\
q(\mathbf{x}, \boldsymbol{\theta}) \propto \; q_1(\mathbf{x}) \, q_2(\boldsymbol{\gamma}) \, q_3(\boldsymbol{\nu})
\end{cases}
$$

According to joint posterior $p(\mathbf{x}, \boldsymbol{\gamma}, \boldsymbol{\nu}|\mathbf{y})$ in Eq.(8.13), we propose to use the following structure for the approximation:

$$
p(\mathbf{x}, \boldsymbol{\gamma}, \boldsymbol{\nu}|\mathbf{y}) \propto \; q_1(\mathbf{x}) \, q_2(\boldsymbol{\gamma}) \, q_3(\boldsymbol{\nu}) \, . \tag{8.18}
$$

Then, thanks to the conjugate priors, approximated joint posteriors $\hat{q}_1(\mathbf{x})$ of source powers is of multivariate Gaussian distribution, and $\hat{q}_2(\boldsymbol{\gamma})$, $\hat{q}_3(\boldsymbol{\nu})$ of hidden variables are of Gamma distributions as follows [115, 87]:

$$
\begin{cases}
\hat{q}_1(\mathbf{x}) = \mathcal{N}(\mathbf{x}|\hat{\boldsymbol{\mu}}_x, \hat{\boldsymbol{\Sigma}}_x) \\
\hat{q}_2(\boldsymbol{\gamma}) = \prod\limits_{n=1}^{N} \mathcal{G}(\gamma_n|\hat{a}_{\gamma}, \hat{b}_{\gamma}^n) \\
\hat{q}_3(\boldsymbol{\nu}) = \prod\limits_{n=1}^{N} \mathcal{G}(\nu_n|\hat{a}_{\nu}, \hat{b}_{\nu}^n)
\end{cases} , \tag{8.19}
$$

where $\hat{\boldsymbol{\mu}}_x$ denotes the mean values of $\mathbf{x}$, which is the objective as the estimated source powers; $\hat{\boldsymbol{\Sigma}}_x$ denotes the covariance matrix of $\mathbf{x}$, whose diagonal elements reflect the estimation uncertainty for each source power $x_n \in \mathbf{x}$. Compared with the JMAP estimation in Chapter 6, one of the advantages of VBA estimation is that the VBA can not only obtain the estimation of source powers $\hat{\boldsymbol{\mu}}_x$, but also achieve the estimation uncertainty $\hat{\boldsymbol{\Sigma}}_x$ of each source power.

In Eq.(8.19) all mentioned expectations are calculated as follows [115, 87]:

$$\begin{cases} \hat{\boldsymbol{\mu}}_x = \hat{\boldsymbol{\Sigma}}_x \, \mathbf{H}^T < \boldsymbol{\Sigma}_\nu > \mathbf{y} \\ \hat{\boldsymbol{\Sigma}}_x = (\mathbf{H}^T < \boldsymbol{\Sigma}_\nu > \mathbf{H} + < \boldsymbol{\Sigma}_\gamma >)^{-1} \\ \hat{a}_\gamma = a_\gamma + \dfrac{N}{2}, \quad \hat{b}_\gamma^n = b_\gamma + \dfrac{1}{2} < \mathbf{x}\mathbf{x}^T >_{nn} \\ \hat{a}_\nu = a_\nu + \dfrac{N}{2}, \quad \hat{b}_\nu^n = b_\nu + \dfrac{1}{2} < \boldsymbol{\epsilon}\boldsymbol{\epsilon}^T >_{nn} \end{cases}, \tag{8.20}$$

where operator $(\cdot)_{nn}$ denotes the nth diagonal item, and $< \cdot >$ denotes expectation, which are calculated as follows [115, 87]:

$$\begin{cases} < \boldsymbol{\Sigma}_\nu > = Diag\{< \nu_n >\}_N = Diag\{< \hat{a}_\nu / \hat{b}_\nu^n >\}_N \\ < \boldsymbol{\Sigma}_\gamma > = Diag\{< \gamma_n >\}_N = Diag\{< \hat{a}_\gamma / \hat{b}_\gamma^n >\}_N \\ < \mathbf{x}\mathbf{x}^T > = \hat{\boldsymbol{\mu}}_x \hat{\boldsymbol{\mu}}_x^T + \hat{\boldsymbol{\Sigma}}_x \\ < \boldsymbol{\epsilon}\boldsymbol{\epsilon}^T > = \mathbf{y}\mathbf{y}^T - 2\mathbf{H}\,\hat{\boldsymbol{\mu}}_x \, \mathbf{y}^T + \mathbf{H} < \mathbf{x}\mathbf{x}^T > \mathbf{H}^T \end{cases}, \tag{8.21}$$

All the solutions in Eq.(8.20, 8.21) require the values of these hyperparameters $\boldsymbol{\phi} = [a_\gamma, b_\gamma, a_\nu, b_\nu]^T$. We can consider their determination as a model selection. During the iterations of parameter estimations, variables $\boldsymbol{\theta} = [\mathbf{x}, \boldsymbol{\theta}]$ are first computed, then the parameters $\boldsymbol{\phi}$ can be alternatively estimated by making the first partial derivative of variational bound $\mathfrak{L}_{\boldsymbol{\theta}}(\boldsymbol{\phi})$ equal zero $(\frac{\partial \mathfrak{L}}{\partial \phi_i} = 0)$ as follows [115, 87]:

$$\begin{cases} \dfrac{\partial \mathfrak{L}}{\partial a_\gamma} = N \ln b_\gamma - N \digamma(a_\gamma) + \displaystyle\sum_{n=1}^{N} < \ln \gamma_n > \\ \dfrac{\partial \mathfrak{L}}{\partial b_\gamma} = N \dfrac{a_\gamma}{b_\gamma} - \displaystyle\sum_{n=1}^{N} < \gamma_n > \\ \dfrac{\partial \mathfrak{L}}{\partial a_\nu} = N \ln b_\nu - N \digamma(a_\nu) + \displaystyle\sum_{n=1}^{N} < \ln \nu_n > \\ \dfrac{\partial \mathfrak{L}}{\partial b_\nu} = N \dfrac{a_\nu}{b_\nu} - \displaystyle\sum_{n=1}^{N} < \nu_n > \end{cases}, \tag{8.22}$$

where $< \gamma_n >$ is computed in Eq.(8.21), and $\digamma(\cdot)$ denotes the Digamma function defined by $\psi(x) = \Gamma'(x)/\Gamma(x)$. a_ν, b_ν can be simultaneously estimated from the same procedure. Parameter update can be done numerically by Matlab fzero function.

8.3.2 VBA computation complexity

From the solutions in Eq.(8.20), $\hat{\mathbf{\Sigma}}_x$ 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 [115, 87]

$$\hat{\mathbf{\Sigma}}_x \approx (< \bar{\nu} > \mathbf{H}^T\mathbf{H} + < \bar{\gamma} > \mathbf{I})^{-1} ,$$

where $\bar{\nu} = \frac{1}{N}\sum_{n=1}^{N} \nu_n$, $\bar{\gamma} = \frac{1}{N}\sum_{n=1}^{N} \gamma_n$ denote the arithmetic mean. Then the products of circulant matrices can be efficiently computed using the Discrete Fourier Transform (DFT) domain. In Eq.(8.20), the estimated expectation $\hat{\boldsymbol{\mu}}_x$ of source powers can be analytically expressed as [115, 87]

$$\hat{\mathbf{\Sigma}}_x^{-1}\, \hat{\boldsymbol{\mu}}_x = \mathbf{H}^T < \mathbf{\Sigma}_\nu > \mathbf{y} .$$

This linear system of equations is solved iteratively with the conjugate gradient algorithm, which requires $O(N\, log_2\, N)$ computations to treat N dimension vector $\mathbf{x}$. If Q iterations are needed, total computations are of $O(Q\, N\, log_2\, N)$, which remains moderate burden.

8.4 Simulations

The simulation configurations are also the same as the one in Chapter 4.7 as shown in above figure: there are $M = 64$ non-uniform sensors locating on the vertical plane. $d = 2$m is the averaged size of sensor array. $D = 4.50$m is the distance between the sensor plane and source plane. $c_0 \approx 340$m/s is the acoustic speed in the common air. $T = 10000$ is the total number of samplings. For the simulated sources in Fig.7.13(a), there are simulated 4 monopoles and 5 complex sources, spaced at least 20cm from each other. Original source powers $\mathbf{x}^*$ are within [0.08,2] ([-10.3,3.7]dB) and 14dB dynamic range. And power image size is of $N_c = 27$ and $N_r = 17$ as shown in Fig.7.3(c). The non-stationary 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.

In Fig.8.3(b)-(d), the beamforming merely gives some strong source powers. DR-DAMAS too much removes non-stationary noises, but failed to distinguish weak sources. JMAP via classical forward model well detects all source powers except for the weakest monopole source. JMAP via convolution model quickly reconstruct most of the sources, but the complex source patterns are affected due to non-stationary noises. However, our proposed

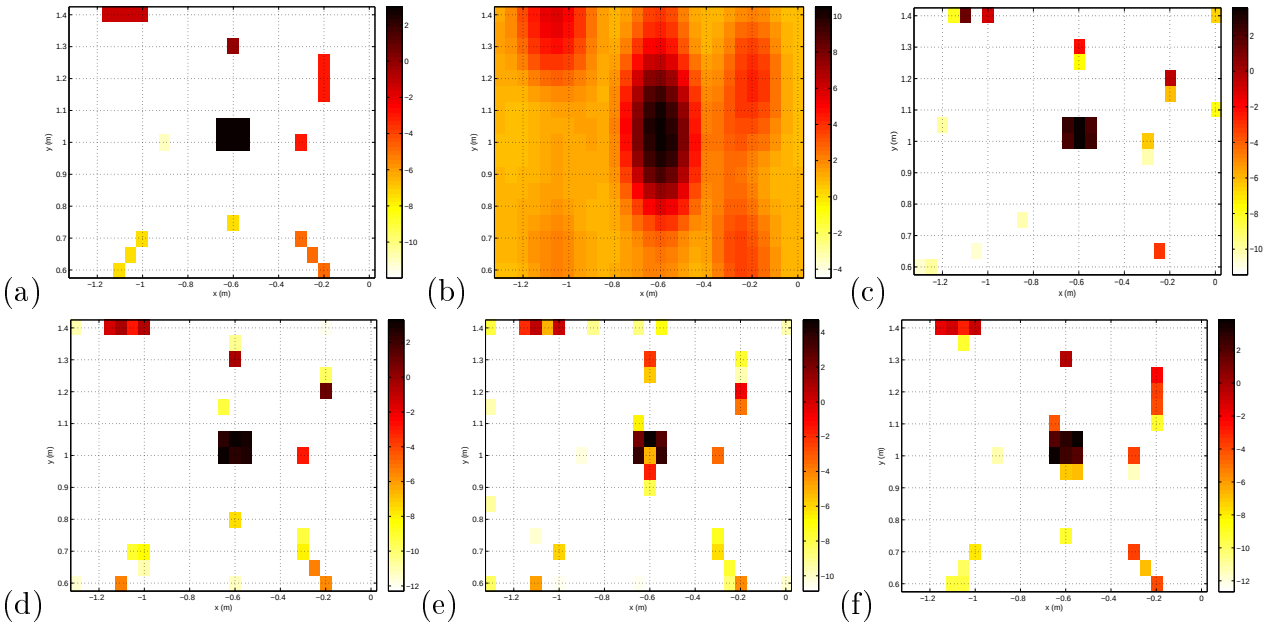

Figure 8.3: Simulation at 2500Hz, 0dB SNR in non-stationary noises, 14dB display: (a) Source powers (b) Beamforming powers (c) DR-DAMAS (d) JMAP via classical forward model (e) JMAP via invariant convolution model and (f) Proposed VBA via invariant convolution model.

VBA via convolution model outperforms the other methods, and efficiently offers more precise localization and pattern estimations, especially for the better non-stationary noise suppression.

8.5 Wind tunnel experiments

The above figure shows the configurations of the wind tunnel S2A [86], object vehicle, NUA array and wind refraction. We suppose that all acoustic sources locate on the same plane. This assumption is almost satisfied, because the curvature of the car side is relatively small compared to the distance D=4.5m between the car and array plane. Since the scanning step is set by $\Delta p = 5cm$, the source plane of car side is of 1.5× 5 m^2 (31×101 pixels), and we also focus on a small region of the rear-view mirror: 1×1.5 m^2 (21×31 pixels). On the real data, there are T=524288 samplings with the sampling frequency f_s=2.56×10^4 Hz. We separate these samplings into I=204 blocks with L=2560 samplings in each bloc. The working frequency band is chosen as [2400,2600]Hz, which is sensitive to human being. The image results are shown by normalized dB images with 10dB span. For the actual propagation time $\tau_{n,m}$ and distance $r_{n,m}$ in Eq.(3.11), we apply equivalent source to make refraction correction. For $\tau_{-n,m}$ and $r_{-n,m}$ in Eq.(3.11), we use the

mirror source signal s_{-n} to correct the ground reflection. The details of the propagation corrections are discussed in A and B.

Figure.8.4 illustrates the estimated power images of mentioned methods at 2500Hz. Due to the high sidelobe effect, beamforming just gives a blurred image of strong sources in Fig.8.4(a). 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.8.4(b). Diagonal Removal DAMAS in Fig.8.4 eliminates most of the artifacts, but it also harms weak sources. Figure.8.4(d) and 8.4(e) 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.8.4(f) 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 noises are much better than others thanks to the Student-t prior on the non-stationary noises. From table 8.1, proposed VBA via convolution model is more efficient realized compared with Bayesian MAP method via classical forward model.

In Fig.8.5, we show the estimation results of source powers $\hat{\mathbf{x}}$ and their variances (uncertainty) $\hat{\mathbf{\Sigma}}_x$, as well as the superposition of the two. In Fig.8.5(c), the height value denotes the estimated variance of each estimated source power. The smaller the variance is, the less uncertainty and more precise the estimation of the source power becomes. It is seen that all of the estimated variance values are very small (2.5%). So that proposed VBA estimation is validated to obtain the effective estimations of source powers. Moreover, the variances of source powers on the rearview mirror and back wheel are relatively big, while the variances on the front wheel are relatively small. These phenomena can be explained by the results in Fig.8.5(a) that stronger source powers on the front wheel are more easily estimated than the weak ones on mirror and back wheel part, so that the estimation uncertainty is smaller for strong sources than the weak ones.

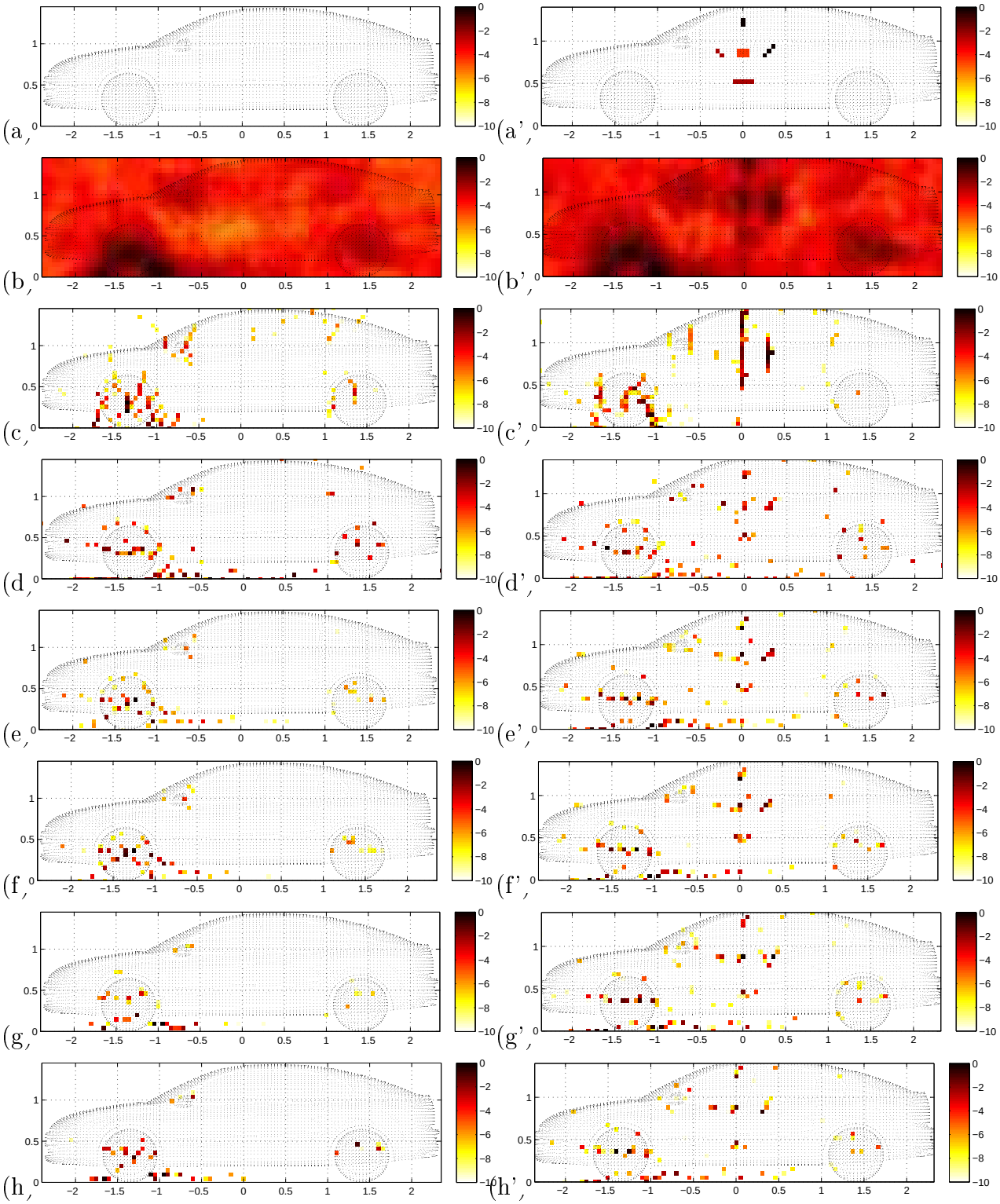

Figure 8.4: Left: real data (a) vehicle surface (b) Beamforming (c) DAMAS (5000i) (d) CLEAN (e) SC-RDAMAS (f) Joint MAP via classical forward model (g) Joint MAP via invariant convolution model and (h) Proposed VBA via invariant convolution model. Right: hybrid data (a') 5 simulated complex sources (b'-h') corresponding methods.

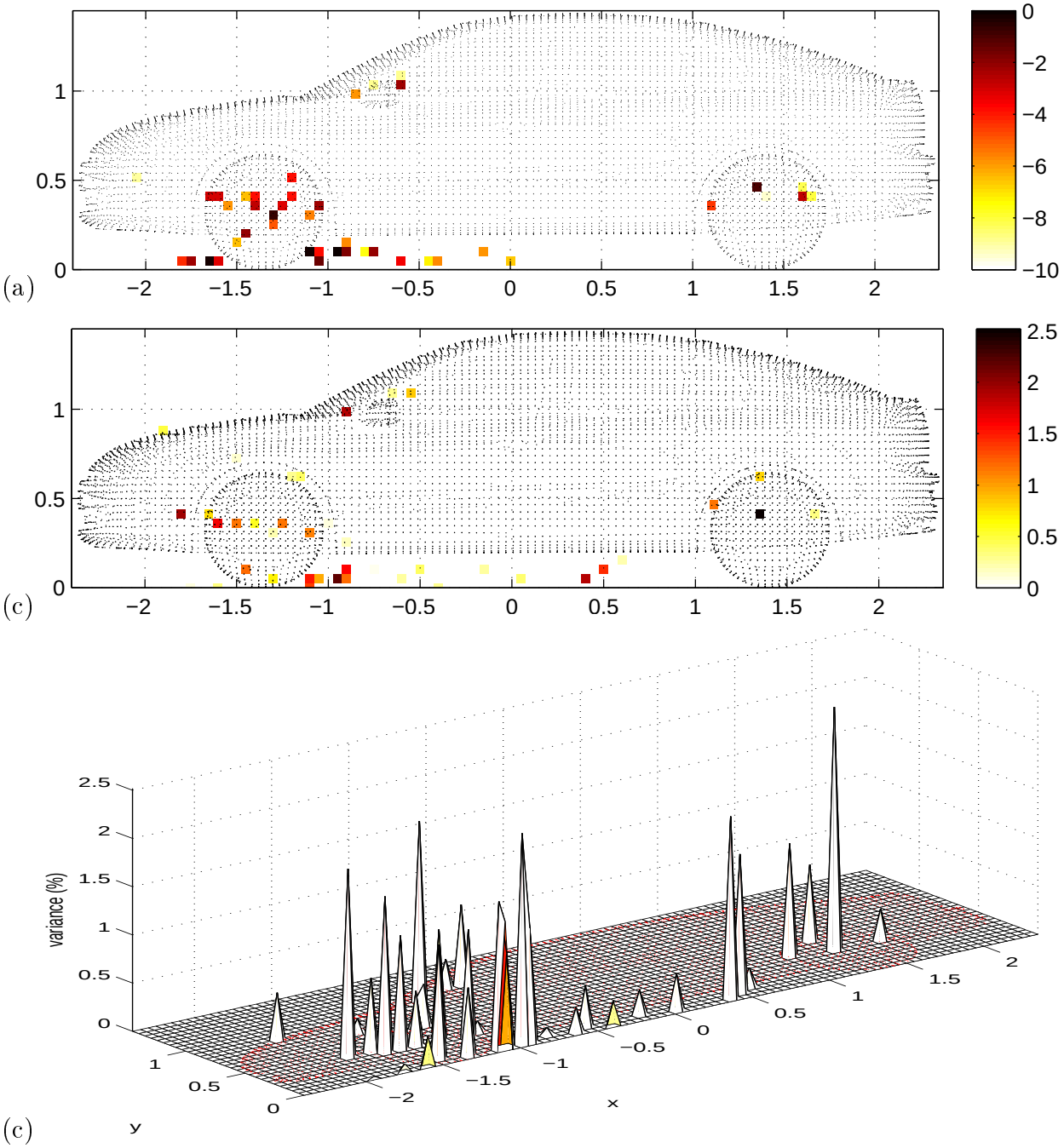

Figure 8.5: Estimation of variances of source powers on the vehicle surface on Simulation at 2500Hz, 0dB SNR in non-stationary noises: (a) Proposed VBA source power estimation via 2D invariant convolution model (b) Estimated variances (uncertainty) of source powers. (c) Superposition of source power and its uncertainty

Table 8.1: Computational cost for treating whole car: image 31×101 pixels, at 2500Hz, based on CPU:3.33GHz. 'JMAP+Conv' is short for Bayesian JMAP method via 2D invariant convolution model. 'VBA+Conv' is short for Variational Bayesian Approximation via 2D invariant convolution model.

Methods	CBF	DAMAS	DR-DAMAS	CLEAN	SC-RDAMAS	MAP	MAP+Con	VBA+Con
Time (s)	1	10	11	45	852	1012	180	578

8.6 Conclusions and perspectives

In this chapter, based on the 2D invariant convolution model, we develop an efficient VBA inference approach via Student-t priors on source powers and non-stationary noises. Proposed approach obtains efficient and effective acoustic imaging results with super spatial resolution, wide dynamic range and robustness to spatially variant model errors.

The main conclusions are:

- For a more suitable sparsity enforcing prior on source power distribution, we propose to use Student-t prior and bring in hidden variables to interpret the hierarchical structure of Student-t prior. Compared with the limitations of Double Exponential prior in Chapter 6 such as single model parameter and discontinuity at zero, there are three advantages in using Student-t prior: firstly, it has more heavier tail to enforce the sparsity and improve the dynamic range of source power distribution. Secondly, it is decomposed into multivariate Gaussian of source powers and Gamma distribution of hidden variables, and this decomposition not only offers the possibility of using conjugate priors, but also provides an alternate and easy way for parameter estimations (VBA), rather than the JMAP estimation. Finally, the hidden variables (non-constant) reflect the spatial variances of source powers, so that the estimation of each hidden variable offers the important result of the uncertainty (variance) of each estimated source power.

- For a more accurate prior of model errors rather then the i.i.d Gaussian distribution in the previous chapters, we propose to use the Student-t prior to model the spatially variant distribution of model errors, whose variances are not constant, but depending on the different positions. Owing to the Normal-Gamma decomposition of Student-t prior, we model the conditional prior of model errors by using a non-stationary multivariate Gaussian prior with the variance matrix (expressed by hid-

den variables) being modeled by Gamma prior.

- For a more effective and efficient way of parameter estimations, we apply the Variational Bayesian Approximation (VBA) estimation via conjugate priors. In VBA, the posterior probability is approximated by a family of basic easily handled probability distributions, which can be estimated by minimizing the Kullback-Leibler (KL) divergence. So that VBA estimation contributes a more precise use of the joint posterior probability laws of all the unknowns.

For the validations of proposed approach, we present method comparisons based on the simulated and real data in wind tunnel experiments. Proposed Bayesian achieves much better results than mentioned classical methods in Chapter 4. It gets more robust estimations than the proposed sparse regularization approach (SC-RDAMAS) in Chapter 5, the JMAP estimation via proposed 2D convolution forward model ($\mathbf{y} = \mathbf{H}\,\mathbf{x} + \boldsymbol{\epsilon}$) in Chapter 7. And it obtains as good results (a little better) as the proposed JMAP estimation via original forward model of power propagation ($\mathbf{y} = \mathbf{C}\,\mathbf{x} + \sigma_e^2\,\mathbf{1}_a + \boldsymbol{\xi}$) in Chapter 6.

However, even though using 2D convolution forward model, the computation cost of proposed VBA approach rests still very high for practical use. For the original forward model , we do not apply the VBA estimation for the huge computational cost of matrix multiplication.

In Fig.8.6, we show the main principles and estimation procedures of the JMAP and VBA methods. The common things between two methods are: (1) they are both developed based on joint posterior of all unknown quantities; (2) they all use useful priors on source powers and model parameters such as model errors, variances etc. But their difference are obvious: (1) JMAP is based on the alternate optimization of joint posterior, while VBA aims to approximate the joint posterior by using a family of basically, easily handled probability density functions (PDF) of unknown quantities; (2) JMAP is the point estimator of the unknown, while VBA first compute the approximated separable joint posterior and then estimate the mean values and variance matrix of the unknown. (3) JMAP is used when simple priors of unknown variables, while VBA is not only used with simple priors, but also fits well to the hierarchical priors. (4) JMAP estimates the objective unknown variables such as source powers, while VBA furthermore estimates the hidden variables in hierarchical priors and offers the credible interval (variance) estimations for more complex cases.

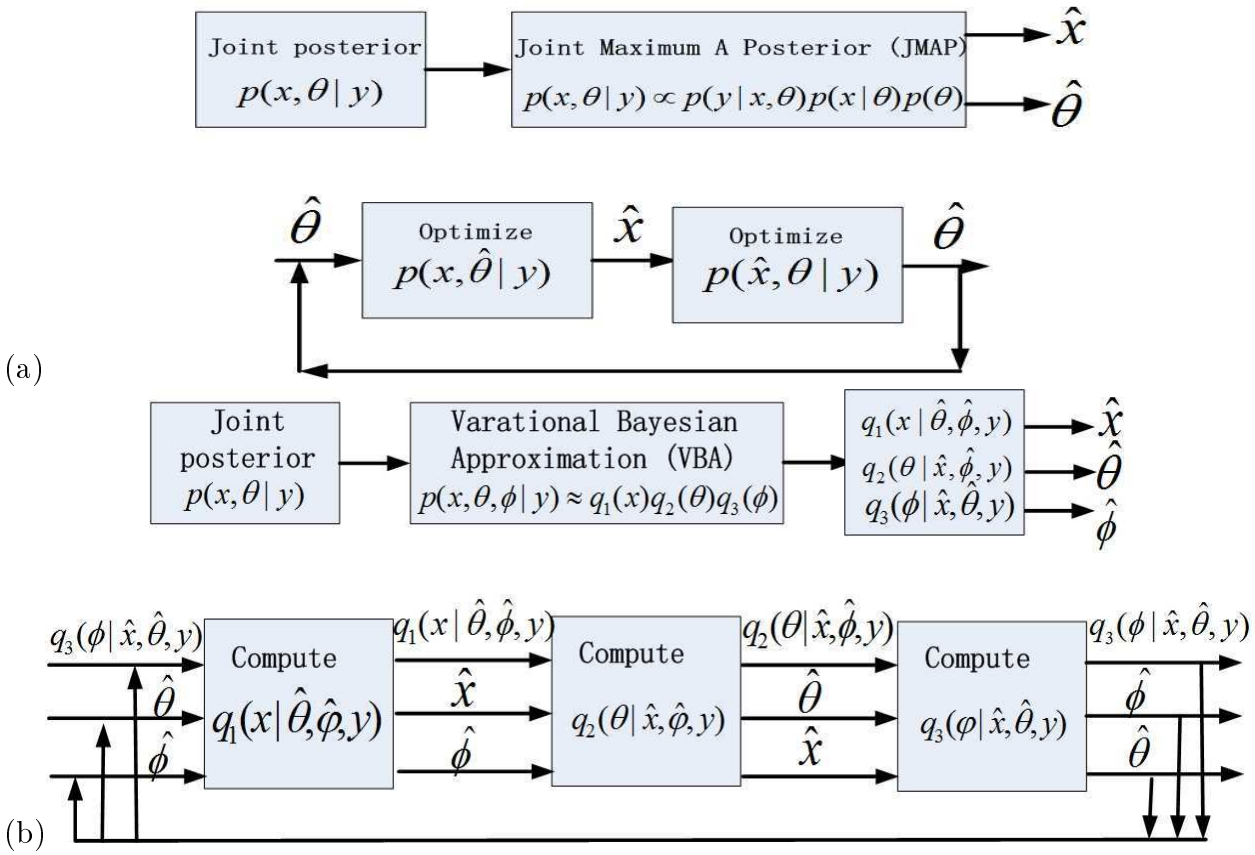

Figure 8.6: Main principles and estimation procedures of the JMAP and VBA methods [87]: (a) Joint Maximum A Posterior (JMAP) (b) Variational Bayesian Approximation (VBA)

For future works, to overcome the computational complexity of VBA, we propose to use the Graphical Processing Unit (GPU) to implement the VBA estimation via convolution forward model, so that the computational speed would be greatly improved. It is also import to investigate the other hierarchical priors on correlated sources and colored model errors so as to improve the VBA performance for real applications.

In Table 8.2, we give a brief summary for the advantages and drawbacks of mentioned classic methods, proposed SC-RDAMAS in Chapter 5, proposed Bayesian JMAP approach in Chapter 6, as well as the Bayesian JMAP approach via proposed 2D invariant convolution model in Chapter 7, and VBA approach in this chapter.

Table 8.2: General performance of classical methods and proposed Bayesian JMAP approach. 'JMAP+Conv' is short for Bayesian JMAP method via 2D invariant convolution model. 'VBA+Conv' is short for Variational Bayesian Approximation via 2D invariant convolution model.

Methods	CBF	CLEAN	DAMAS	CMF	SC-RDAMAS	JMAP	JMAP+Conv	VBA+Conv
Resolutions	Low	Normal	Normal	Higher	Higher	Higher	High	Higher
Dynamic Range	Narrow	Normal	Normal	Normal	Wide	Wide	Normal	Wider
Noise	Robust	Sensitive	Sensitive	Robust	Robust	Robust	Robust	Robust
Computation	Least	Normal	Normal	High	High	Higher	Normal	High
Samples	Normal	Normal	Normal	More	Normal	Normal	Normal	High
Source number	No	Required	No	Required	Required	No	No	No

.

- Ce qui est le plus pitoyable pour l'homme, c'est de ne pas avoir la connaissance et de ne pas pouvoir se contraindre.

Michel Eyquem de Montaigne

9

Conclusions and Perspectives

Acoustic imaging is a standard technique for mapping the positions and powers of acoustic sources with microphone arrays, which provides insights into the mechanisms and properties of acoustic sources, especially for evaluating the acoustic influence in automobile and aircraft industries. However, acoustic imaging often involves an ill-posed problem and becomes sensitive to measured errors. Motivated by these challenges, in this thesis, we develop several effective approaches for acoustic imaging on the vehicle surface in the wind tunnel experiments.

9.1 Conclusions

The main conclusions of this thesis can be summarized as follows:

- In Chapter 3, we build up a forward model in the frequency domain for acoustic **signal** propagation from source plane to microphone sensor array in Eq.(3.9):

$$\mathbf{z}(f_l) = \mathbf{A}(\mathbf{P}^*, f_l)\,\mathbf{s}^*(f_l) + \mathbf{e}(f_l)$$

  In this signal model, the novelty is that we improve the propagation matrix $\mathbf{A}$ by considering multi-path propagation effects such as ground reflection and wind refraction in the wind tunnel. However, this model is a non-linear function of unknown source signals $\mathbf{s}^*$ and their unknown

positions $\mathbf{P}^*$. For linearization, we set up a **discrete** forward model of acoustic **signal** propagation based on discretization of the source plane in Eq.(3.15):

$$\mathbf{z} = \mathbf{A}(\mathbf{P})\,\mathbf{s} + \mathbf{e}$$

In this **discrete** signal model, the discrete source positions $\mathbf{P}$ are known; and $\mathbf{P}^*$ and $\mathbf{s}^*$ are supposed to be included in $\mathbf{P}$ and $\mathbf{s}$ respectively. Therefore, this model becomes a linear system of equations for discrete source signals $\mathbf{s}$, but it is an underdetermined model which cause an ill-posed inverse problem, especially when discrete source number N is much greater than sensor number M, namely $M < N$.

- In Chapter 4, based on Eq.(3.15) and beamforming method in Eq.(4.12), we set up a forward model of acoustic **power** propagation in Eq.(4.46):

$$\mathbf{y} = \mathbf{C}\,\mathbf{x} + \sigma_e^2\,\mathbf{1}_a\,,$$

where beamforming powers $\mathbf{y} = \mathbb{E}[|\tilde{\mathbf{A}}^\dagger\,\mathbf{z}|^2]$ with $\tilde{\mathbf{A}}$ being the beamforming steering matrix which is obtained from $\mathbf{A}$; power propagation matrix $\mathbf{C} = |\tilde{\mathbf{A}}^\dagger\,\mathbf{A}|.^2$; source power $\mathbf{x} = \text{diag}\left[\mathbb{E}[\mathbf{s}\mathbf{s}^\dagger]\right]$; background noise power $\sigma_e^2 = \mathbb{E}[\mathbf{e}^\dagger\mathbf{e}]$. This model is a linear and determined system of equations for source power $\mathbf{x}$. Since acoustic imaging mainly involves source power reconstruction and localization, the **power** propagation model in Eq.(4.46) can be thus more easily handled than **signal** propagation model in Eq.(3.15). This is because the signal model contains more unknown variables such as signal phases than the power model.

We also present the state-of-the-art methods in source localization and power reconstruction. The classical spatial filter methods, such as the beamforming, Capon and MUSIC, aim to solve the forward model of acoustic **signal** propagation in Eq.(3.15). They can offer a direct estimation for source localization. The deconvolution methods aim to solve the forward model of acoustic **power** propagation in Eq.(4.46). They can iteratively deconvolve the blurry beamforming power result, and can obtain a high spatial resolution. The regularization methods can further improve the spatial resolution of deconvolution methods by using the sparse regularizations (ℓ_0, ℓ_1 or ℓ_l with $0 < l < 1$ norm), or improve the robustness to noises owing to the Tikhonov regularization. But the regularization term inevitably increases method complexity, and selecting regularization parameter becomes an essential issue for good use. Moreover, Covariance Matrix Fitting (CMF) method can directly estimate the cross-spectrum matrix of (un)correlated source signals via sparsity

constraint, but CMF has huge dimension of variables to be estimated. Therefore, there is no one-fits-all methods. Most of the mentioned suffer one of the following drawbacks: poor spatial resolution, sensitivity to background noises, narrow dynamic range or high computational cost. In addition, some methods need to set important parameters that must be tuned carefully for each case.

- In Chapter 5, based on Eq.(4.46), we propose an improved power propagation forward model for the model robustness to measured errors as given in Eq.(5.4):

$$\mathbf{y} = \mathbf{C}\,\mathbf{x} + \sigma_e^2\,\mathbf{1}_a + \boldsymbol{\xi}$$

In this improved power model, the novelty is that we consider the model uncertainty $\boldsymbol{\xi}$ caused by multi-path propagation effects in the wind tunnel. For this inverse problem, we propose a robust super-resolution approach with sparsity constraint (SC-RDAMAS) in Eq.(5.6):

$$\begin{cases} (\hat{\mathbf{x}}, \hat{\sigma_e^2}) = \arg\min\limits_{(\mathbf{x},\sigma_e^2)} \left\{ ||\mathbf{y} - \mathbf{C}\mathbf{x} - \sigma_e^2\,\mathbf{1}_a||_2^2 \right\} \\ \quad \text{s.t.} \mathbf{x} \succeq \mathbf{0}, \quad \|\mathbf{x}\|_1 = \beta, \quad \sigma_e^2 \geq 0 \end{cases}$$

The sparsity parameter β is interpreted by the total source power. We adaptively estimated β in the proposed Algorithm 1 in Section 5.2.2. For the approach validation, we present performance comparisons with mentioned classical methods. The simulations show that proposed approach obtained the $5cm$ super resolution compared with the beamforming resolution 31cm at 2500Hz. It achieved 15dB dynamic range of power estimations, and well detected complex sources with different patterns. The real data results demonstrate that proposed approach effectively reconstructed strong sources on front wheels and rear-view mirrors, as well as the weak sources on back wheels. The hybrid data experiments have furthermore confirmed the effectiveness for reconstructing the known synthetic sources and original sources in the real data. We also show the moderate computational cost of our approach for the acoustic imaging in wind tunnel tests. The main drawback of proposed approach is the over-winnowing effect that over-estimation or under-estimation of sparsity parameter β could causes many unstructured or shapeless points, and could not to reconstruct the true source distribution, especially when acoustic sources model could not be modeled by monopoles.

- In Chapter 6, to overcome the limitations of proposed SC-RDAMAS approach, we propose to use a Bayesian inference approach with a spar-

sity enforcing prior. We first considered the model uncertainty as a random variable $\boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{\xi}|\mathbf{0}, \sigma_\xi^2)$, so that the likelihood $p(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta}_1) = \mathcal{N}(\mathbf{y}|\mathbf{C}\,\mathbf{x} + \sigma_e^2\,\mathbf{1}_a, \sigma_\xi^2)$ can be derived from it ($\boldsymbol{\theta}_1 = [\sigma_e^2, \sigma_\xi^2]^T$). We then modeled the sparse distribution of source powers by using the Double Exponential $\mathcal{DE}(\mathbf{x})$ prior model as:

$$p(\mathbf{x}|\boldsymbol{\theta}_2) = \left(\frac{\gamma}{2}\right)^N \exp\left[-\gamma \|\mathbf{x}\|_\beta^\beta\right]$$

$\beta = 1$ greatly improved the spatial resolution, and proper γ promoted the wide dynamic range of source powers; $\boldsymbol{\theta}_2 = [\gamma, \beta]^T$. For robust solution of this inverse problem, we apply the Joint Maximum A Posterior (JMAP) estimation to jointly estimated source powers $\mathbf{x}$, sparsity parameter γ, noise power σ_e^2 and model uncertainty variance σ_ξ^2 ($\beta = 1$ is fixed). Therefore, propose Bayesian JMAP approach can be summarized as: (see Eq.(6.2) and Eq.(6.8))

$$\begin{cases} \mathbf{y} = \mathbf{C}\,\mathbf{x} + \sigma_e^2\,\mathbf{1}_a + \boldsymbol{\xi} \\ (\hat{\mathbf{x}}, \hat{\boldsymbol{\theta}}) = \arg\min\limits_{(\mathbf{x},\boldsymbol{\theta})} \{\mathcal{J}(\mathbf{x}, \boldsymbol{\theta})\}\ , \quad \mathcal{J}(\mathbf{x}, \boldsymbol{\theta}) = -\ln\, p(\mathbf{x}, \boldsymbol{\theta}|\mathbf{y}) \\ p(\mathbf{x}, \boldsymbol{\theta}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta}_1)\, p(\mathbf{x}|\boldsymbol{\theta}_2)\, p(\boldsymbol{\theta}_2)\, p(\boldsymbol{\theta}_1) \\ \mathcal{J}(\mathbf{x}, \boldsymbol{\theta}) = \underbrace{\frac{1}{2\sigma_\xi^2}\|\mathbf{y} - \mathbf{C}\mathbf{x} - \sigma_e^2\,\mathbf{1}_a\|^2}_{Likelihood} + \underbrace{\gamma\,\|\mathbf{x}\|_1}_{Sparse\,prior} + \underbrace{\frac{N}{2}\ln\sigma_\xi^2 - N\,\ln\gamma}_{Hyperparameter\,prior} \\ \text{s.t.} \quad \mathbf{x} \succeq \mathbf{0},\ \sigma_e^2 \succeq 0,\ \sigma_\xi^2 \succeq 0,\ \gamma \succeq 0 \\ \boldsymbol{\theta}_1 = [\sigma_e^2, \sigma_\xi^2]^T, \quad \boldsymbol{\theta}_2 = \gamma, \quad \boldsymbol{\theta} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2]^T \end{cases}$$

$\gamma\,\|\mathbf{x}\|_1$ can be interpreted as the regularization form $\mathcal{F}(\mathbf{x})$ in Eq.(4.58), and hyperparameter γ takes the similar effect of the regularization parameter. This optimization is carried out by alternate estimations for the unknown variables and hyperparameters in Eq.(6.9). By simulations, real data and hybrid data, the proposed Bayesian JMAP approach achieved better performance than proposed SC-RDAMAS. However, JMAP estimation confronts the non-quadratic optimization and requires large computational cost.

- In Chapter 7, to reduce the computational time of proposed Bayesian JMAP approach, we propose to use the 2D invariant convolution to approximate the power propagation model as given in Eq.(7.9):

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

We suppose that model errors $\boldsymbol{\epsilon}$ consist of background noise power σ_e^2, multi-path uncertainty $\boldsymbol{\xi}$, as well as convolution approximated errors $\boldsymbol{\eta}$ in Eq.(7.7). $\mathbf{H}$ denotes the valid convolution matrix. The 2D invariant convolution kernel $\mathbf{h}$ (size and values) was derived from the quasi Symmetric Toeplitz Block Toeplitz (STBT) structure of power propagation matrix $\mathbf{C}$ as given in Eq.(7.13):

$$\begin{cases} h_{k,l} = \tilde{c}_{i,j}\,, \\ i = \lfloor \frac{N+1}{2} \rfloor\,, \quad j = i + (\lfloor \frac{N_r+1}{2} \rfloor - k)\, N_r + \lfloor \frac{N_r+1}{2} \rfloor - l \end{cases}$$

Moreover, $\mathbf{h}$ was regarded as a separable 2D-convolution which could be separated into two 1D-convolution kernels. One of the advantages is that 2D separable convolution greatly reduces both computational cost and memory requirements of 2D invariant convolution operation. We also realized the 2D (in)variant and separable convolutions using MATLAB Parallel Toolbox based on many-core processors such as the Graphics Processing Unit (GPU). On simulations and real data, the Bayesian JMAP method via proposed 2D invariant convolution model obtained an efficient and acceptable imaging result. For acoustic imaging with very large size, deconvolution method via 2D separable convolution can be very fast implemented using GPU. Compared to CPU with matrix multiplication $\mathbf{C}\,\mathbf{x}$, the 2D separable convolution using GPU achieves the speed gain as high as three order of magnitude; compared to CPU with 2D non-separable convolution $\mathbf{h} * \mathbf{x}$, it gains two order of magnitude of speed gain; compared to GPU with separable convolution $\mathbf{h}_1 * \mathbf{h}_2^T * \mathbf{x}$, it still obtains nearly one order of magnitude of of speed gain.

- In Chapter 8, based on 2D invariant convolution model in Eq.(7.9), we considered the model errors to be spatially non-stationary Gaussian distribution, rather than i.i.d AGWN used in previous chapters. we propose to use the Student-t prior to model this kind of errors as given in Eq.(8.4):

$$St(\boldsymbol{\epsilon}|\alpha_\epsilon) = \prod_{n=1}^{N} \frac{\Gamma(\frac{1+\alpha_\epsilon}{2})}{\sqrt{\alpha_\epsilon\,\pi}\,\Gamma(\frac{\alpha_\epsilon}{2})} (1 + \frac{\epsilon_n^2}{\alpha_\epsilon})^{-\frac{1+\alpha_\epsilon}{2}}$$

One of the attractive properties of the Student-t distributions is that $St(\boldsymbol{\epsilon}|\alpha_\epsilon) = \int p(\boldsymbol{\epsilon}|\boldsymbol{\nu})\, p(\boldsymbol{\nu}|\alpha_\epsilon)\, d\boldsymbol{\nu}$ can be decomposed into multivariate Gaussian Normal distribution $p(\boldsymbol{\epsilon}|\boldsymbol{\nu}) = \mathcal{N}(\boldsymbol{\epsilon}|\mathbf{0}, \boldsymbol{\Sigma}_\nu^{-1})$ with $\boldsymbol{\Sigma}_\nu = Diag[\nu_n]$, and Gamma distribution $p(\boldsymbol{\nu}|\alpha_\epsilon) = \prod_{n=1}^{N} \mathcal{G}(\nu_n|a_\nu, b_\nu)$ with $a_\nu = b_\nu = \frac{\alpha_\epsilon}{2}$. Here $\boldsymbol{\nu} = [\nu_1, \cdots, \nu_N]^T$ denotes hidden variable vector, which presents the precision of variable $\boldsymbol{\epsilon}$ (credible interval).

Moreover, we investigated a more proper sparsity enforcing prior on source power distribution than the Double Exponential prior used in Chapter 6. The Student-t prior $St(\mathbf{x}|\alpha_x)$ is more suitable to represent a sparse distribution owing to its narrow peak and heavy tail. Again here, using the Normal-Gamma integration property of $St(\mathbf{x}|\alpha_x) = \int p(\mathbf{x}|\boldsymbol{\gamma})\, p(\boldsymbol{\gamma}|\alpha_x)\, d\boldsymbol{\gamma}$, we propose to use a hierarchical model as $p(\mathbf{x}|\boldsymbol{\gamma}) = \mathcal{N}(\mathbf{x}|\mathbf{0}, \boldsymbol{\Sigma}_\gamma^{-1})$ with $\boldsymbol{\Sigma}_\gamma = Diag[\gamma_n]$ and $p(\boldsymbol{\gamma}|\alpha_x) = \prod_{n=1}^{N} \mathcal{G}(\gamma_n|a_\gamma, b_\gamma)$, with $a_\gamma = b_\gamma = \frac{\alpha_x}{2}$. Here $\boldsymbol{\gamma} = [\gamma_1, \cdots, \gamma_N]^T$ denotes hidden variable vector, which presents the precision of variable $\mathbf{x}$.

For joint estimations of unknown variables ($\mathbf{x}$) and hyperparameters ($\boldsymbol{\gamma}$, $\boldsymbol{\nu}$), we could not easily apply the Bayesian JMAP as used in Chapter 6, the reason is that there are too many unknown quantities for JMAP criterion to get a quadratic optimization. So that we applied the Variational Bayesian Approximation (VBA) estimation. In VBA, the joint posterior probability is approximated by a family of basic easily handled probability distributions, which are estimated by minimizing the Kullback-Leibler (KL) divergence between the two probability functions. We summarize the proposed VBA approach via 2D invariant convolution model as follows:

$$
\left\{
\begin{array}{rl}
\mathbf{y} &= \mathbf{H}\,\mathbf{x} + \boldsymbol{\epsilon} \\
\hat{q}(\mathbf{x}, \boldsymbol{\theta}) &= \arg\min_{q(\mathbf{x},\boldsymbol{\theta})} \left\{ KL(q:p) = \int q(\mathbf{x}, \boldsymbol{\theta})\, \dfrac{q(\mathbf{x}, \boldsymbol{\theta})}{p(\mathbf{x}, \boldsymbol{\theta}|\mathbf{y})}\, d(\mathbf{x}, \boldsymbol{\theta}) \right\} \\
St(\mathbf{x}) &= \int p(\mathbf{x}|\boldsymbol{\gamma})\, p(\boldsymbol{\gamma})\, d\boldsymbol{\gamma} = \int \mathcal{N}(\mathbf{x}|\mathbf{0}, \boldsymbol{\Sigma}_\gamma^{-1})\, \mathcal{G}(\boldsymbol{\gamma}|a_\gamma, b_\gamma)\, d\boldsymbol{\gamma} \\
St(\boldsymbol{\epsilon}) &= \int p(\boldsymbol{\epsilon}|\boldsymbol{\nu})\, p(\boldsymbol{\nu})\, d\boldsymbol{\nu} = \int \mathcal{N}(\boldsymbol{\epsilon}|\mathbf{0},\, \boldsymbol{\Sigma}_\nu^{-1})\, \mathcal{G}(\boldsymbol{\nu}|a_\nu, b_\nu)\, d\boldsymbol{\nu} \\
\boldsymbol{\theta} &= [\boldsymbol{\gamma}, \boldsymbol{\nu}]^T \\
p(\mathbf{x}, \boldsymbol{\theta}|\mathbf{y}) &\propto p(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta})\, p(\mathbf{x}|\boldsymbol{\gamma})\, p(\boldsymbol{\gamma})\, p(\boldsymbol{\nu}) \\
&\propto \mathcal{N}(\mathbf{y}|\mathbf{H}\,\mathbf{x},\, \boldsymbol{\Sigma}_\nu^{-1})\, \mathcal{G}(\boldsymbol{\nu}|a_\nu, b_\nu)\, \mathcal{N}(\mathbf{x}|\mathbf{0}, \boldsymbol{\Sigma}_\gamma^{-1})\, \mathcal{G}(\boldsymbol{\gamma}|a_\gamma, b_\gamma) \\
q(\mathbf{x}, \boldsymbol{\theta}) &\propto\ q_1(\mathbf{x})\, q_2(\boldsymbol{\gamma})\, q_3(\boldsymbol{\nu})
\end{array}
\right.
$$

This optimization was efficiently solved by using mutually conjugate priors, that was because the conditional prior $p(\mathbf{x}|\boldsymbol{\gamma})$ and the likelihood $p(\mathbf{y}|\mathbf{x}, \boldsymbol{\nu})$ were multivariate Gaussian distributions, and $p(\boldsymbol{\gamma})$, $p(\boldsymbol{\nu})$ were Gamma distribution, they thus became the 'conjugate priors', so that the approximated posterior $\hat{q}_1(\mathbf{x})$ derived from the same family as $p(\mathbf{x}|\boldsymbol{\gamma})$, and $\hat{q}_2(\boldsymbol{\gamma})$ was as Gamma distribution as $p(\boldsymbol{\gamma})$, similarly, $\hat{q}_3(\boldsymbol{\nu})$ was as Gamma distribution $p(\boldsymbol{\nu})$. Suppose that source powers $\mathbf{x}$ and

hidden variables $\boldsymbol{\gamma}$, $\boldsymbol{\nu}$ are mutually independent, we easily obtained the solutions as given in Eq.(8.19):

$$\begin{cases} \hat{q}_1(\mathbf{x}) = \mathcal{N}(\mathbf{x}|\hat{\boldsymbol{\mu}}_x, \hat{\boldsymbol{\Sigma}}_x) \\ \hat{q}o(\boldsymbol{\gamma}) = \prod_{n=1}^{N} \mathcal{G}(\gamma_n|\hat{a}_\gamma, \hat{b}_\gamma^n) \\ \hat{q}_3(\boldsymbol{\nu}) = \prod_{n=1}^{N} \mathcal{G}(\nu_n|\hat{a}_\nu, \hat{b}_\nu^n) \end{cases}$$

All mentioned expectations and parameters were calculated in Eq.(8.20) and Eq.(8.21). On simulations and real data in wind tunnel experiments, we showed the proposed VBA via convolution model achieved much better results than Bayesian JMAP via convolution model, and it also obtained the results as good as the Bayesian JMAP via conventional model. However, the computation cost of VBA approach are still high for practical use.

To sum up, proposed approaches are compared with some of the state-of-the-art methods on simulations, real data from wind tunnel experiments of Renault S2A, as well as the hybrid data. The main advantages are robustness to background noises, wide dynamic range of acoustic powers, super spatial resolutions, efficient to use, and not need for the prior knowledge of the source number or SNR.

In Table 9.1, we show the brief summary of proposed methods for acoustic imaging.

9.2 Perspectives

Although we discuss several advantages of proposed approaches, there exist a great number of important works to correct the drawbacks and refine our researches of this thesis in the coming future. It is highly necessary to consider the following aspects:

- Practical applications:
  - 3D acoustic imaging on the vehicle surface. This work can be carried out based on the measured data of vertical and horizontal microphone sensors in wind tunnel S2A. The 3D imaging results can offer

more detailed acoustic source distribution than 2D results offered in this thesis.

- High-resolution acoustic imaging in lower frequencies ([20,1000]Hz) should be carefully studied, since some of noisy car sounds are generated by low frequency acoustic sources on car windows, wheels and rear-view mirrors etc. But low frequency signals confronts very low spatial resolution. There might be two ways to do: (1) using high resolution methods (Capon, MUSIC) to build up a power propagation model, instead of using beamforming method, since beamforming has low spatial resolution at low frequencies; (2) directly dealing with the signal propagation model in Eq.(3.15) rather than power model, since signal model can provide not only signal amplitudes and positions, but also signal phase and characteristic frequencies.
- Fast implementation using GPU for real-time acoustic imaging for industry application. The biggest drawback of the proposed Bayesian approaches is the huge computation cost. So that it is an essential issue to optimize proposed deconvolution algorithms so as to efficiently fit the parallel structure of GPU. Moreover, we find out that calculating invariant convolution $\mathbf{h} * \mathbf{x}$ based on GPU merely makes use of about 14% of computational power of GPU, while separable convolution $\mathbf{h}_1 * \mathbf{h}_2^T * \mathbf{x}$ just occupies nearly 7%. So that there will be great potential to develop our own parallel separable convolution algorithm based on the GPU so as to make good use of GPU powerful peak computational capacity.

- Accurate forward model of acoustic signal propagation:
  - Sophisticated acoustic source model. The full-wave acoustic propagation models should be investigated [125, 97] in order to get a more precise forward model in the near-filed according to boundary conditions. In fact, uncorrelated monopole source model is an over-simplified source model, which can not describe the non-compact sources, source correlation, pattern or directivity. It is better to consider the (in)coherent distributed source model that is modeled by a parametric angular cross-correlation kernel [99, 76, 92].
  - Direct acoustic signal reconstruction. In this thesis, we mainly focus on the acoustic power reconstruction. However, signal propagation model should be further studied by considering signal amplitude at-

tenuation and phase shift due to the propagation distance, as well as the characteristic frequency that contributes to distinct different acoustic sources. Moreover, the source directivity reconstruction is also an important and practical issue [73, 23, 92] in theory and industrial application, since the real source itself has directivity distribution, and the microphone sensors are not omni-directional but has a directivity pattern too.

  - Complex reverberation effect should be studied for real data processing in wind tunnel. In this thesis, we merely model the ground reflection and wind refraction effects, but there are still unknown multi-path propagation in wind tunnel, especially for a closed non-anechoic chamber. In acoustics, reverberation is the convolution of the original sound with echos from objects surrounding the sound source. We can apply the blind deconvolution methods [129, 115, 75, 83] for dereverberations.
  - Measured signal errors should account for non-stationary non-Gaussian noises in temporal-spatial domain [6, 53], or colored noises in frequency spectrum domain [51, 78]. In this thesis, we just consider the i.i.d AWGN noise and spatially non-stationary Gaussian noise for simplicity. This might be because measured signals are usually deteriorate by complex reverberations, inherent sensor noises and signal processing models.

- Effective inverse problem methods:

  - Direct results of source localization should be further studied. In this thesis, we just focus on source power estimation, and merely give a discrete estimation results for source positions. The limit or Cramér-Rao lower bound (CRLB) of spatial resolution [39, 46] should be a worthwhile work to measure the method performance for acoustic imaging.
  - More appropriate sparse priors of source power distribution should be also investigated with physical meaning, such as fractional Laplacian prior [110] which much enforces the sparsity, the Chi-square distribution [72] which is non-negative sparse distribution of source powers, as well as the group sparsity priors [64, 61] which can properly model a sparse distribution of a group of closely spaced monopole sources on the very large source plane.
  - Fast VBA methods such as Tree-structure compressed sensing [101]

and automatic relevance determination [58] could become promising works to overcome VBA computation burden. In this thesis, we merely apply the VBA method to improve the Bayesian estimation, but it is still the great matter to well study the VBA calculation complexity and improve its performance and efficiency.

– Acoustic source separation in reverberations is an important extension of source localization and imaging. "cocktail party problem" is one of typical example of source separation, in which, several acoustic signals are mixed together into a combined signal, and the objective is to recover all the original signal components from the mixed signals which are often deteriorated by serious reverberations in the non-anechoic chamber. The acoustic imaging on the car surface in wind tunnel is as similar as the cocktail party case. We want to not only distinct the different sources from their different positions on the car, but also from their physical characteristics such as frequencies, spectrum, directivity or powers. Several approaches have been proposed to solve this problem such as the principal components analysis (PCA) and independent components analysis (ICA), as well as Bayesian source separation [122, 55, 62, 120]. It is also a very promising work to exploit advanced source separation methods by using some of acoustic imaging techniques discussed in this thesis.

Table 9.1: Brief summary of proposed methods for acoustic imaging.

Chapter	Chapter 5	Chapter 6	Chapter 7	Chapter 8
Forward model	$\mathbf{y} = \mathbf{C}\,\mathbf{x} + \sigma_e^2\,\mathbf{1}_a + \boldsymbol{\xi}$	$\mathbf{y} = \mathbf{C}\,\mathbf{x} + \sigma_e^2\,\mathbf{1}_a + \boldsymbol{\xi}$	$\mathbf{y} = \mathbf{H}\,\mathbf{x} + \sigma_e^2\,\mathbf{1}_a + \boldsymbol{\xi} + \boldsymbol{\eta}$	$\mathbf{y} = \mathbf{H}\,\mathbf{x} + \boldsymbol{\epsilon}$
Main operation	Matrix multiplication	Matrix multiplication	2D convolution	2D convolution
Error factor	σ_e^2: Background noise; $\boldsymbol{\xi}$: Multi-path propagation uncertainty	σ_e^2, $\boldsymbol{\xi}$	$\boldsymbol{\eta}$: Model approximated error; σ_e^2, $\boldsymbol{\xi}$	$\boldsymbol{\epsilon}$ includes $\boldsymbol{\eta}$, σ_e^2, $\boldsymbol{\xi}$
Inverse criterion	Minimizing $\boldsymbol{\xi}$ under sparsity constraint	Joint Maximum A Posterior (JMAP) $(\hat{\mathbf{x}}, \hat{\sigma^2}_e) = \arg\max\{p(\mathbf{x}, \boldsymbol{\theta}\|\mathbf{y})\}$	Joint Maximum A Posterior (JMAP) $(\hat{\mathbf{x}}, \hat{\sigma^2}_e) = \arg\max\{p(\mathbf{x}, \boldsymbol{\theta}\|\mathbf{y})\}$	Variational Bayesian Approximation (VBA) $\hat{q}(\mathbf{x}, \boldsymbol{\theta}) \approx p(\mathbf{x}, \boldsymbol{\theta}\|\mathbf{y})$
Cost function	$\begin{cases} (\hat{\mathbf{x}}, \hat{\sigma^2}_e) = \arg\min\{\|\|\mathbf{y} - \mathbf{C}\,\mathbf{x} - \sigma_e^2\,\mathbf{1}_a\|\|_2^2\} \\ \text{s.t.} \quad \mathbf{x} \succeq \mathbf{0}, \quad \|\mathbf{x}\|_1 = \beta, \quad \sigma_e^2 \geq 0 \end{cases}$	$\begin{cases} (\hat{\mathbf{x}}, \hat{\boldsymbol{\theta}}) & = \arg\min\{-\ln p(\mathbf{x}, \boldsymbol{\theta}\|\mathbf{y})\} \\ & = \frac{1}{2\sigma_\xi^2}\|\mathbf{y}\mathbf{C}\,\mathbf{x} - \sigma_e^2\,\mathbf{1}_a\|^2 \\ & \quad + \gamma\,\|\mathbf{x}\|_1 + \frac{N}{2}\,\ln\sigma_\xi^2 - N\,\ln\gamma \\ \boldsymbol{\theta} & = [\sigma_e^2, \sigma_\xi^2, \gamma]^T \end{cases}$	$\begin{cases} (\hat{\mathbf{x}}, \hat{\boldsymbol{\theta}}) & = \arg\min\{-\ln p(\mathbf{x}, \boldsymbol{\theta}\|\mathbf{y})\} \\ & = \frac{1}{2\sigma_\xi^2}\|\mathbf{y} - \mathbf{H}\,\mathbf{x} - \sigma_e^2\,\mathbf{1}_a\|^2 \\ & \quad + \gamma\,\|\mathbf{x}\|_1 + \frac{N}{2}\,\ln\sigma_\xi^2 - N\,\ln\gamma \\ \boldsymbol{\theta} & = [\sigma_e^2, \sigma_\xi^2, \gamma]^T, \quad \boldsymbol{\eta} = \mathbf{0} \end{cases}$	$\begin{cases} \hat{q}(\mathbf{x}, \boldsymbol{\theta}) & = \arg\min\{KL(q:p)\} \\ KL(q:p) & = \int q(\mathbf{x}, \boldsymbol{\theta})\,\frac{q(\mathbf{x}, \boldsymbol{\theta})}{p(\mathbf{x}, \boldsymbol{\theta}\|\mathbf{y})}\,d(\mathbf{x}, \boldsymbol{\theta}) \\ p(\mathbf{x}, \boldsymbol{\theta}\|\mathbf{y}) & \propto \mathcal{N}(\mathbf{y}\|\mathbf{H}\,\mathbf{x},\, \boldsymbol{\Sigma}_\nu^{-1})\,\mathcal{G}(\boldsymbol{\nu}\|a_\nu, b_\nu) \\ & \quad \mathcal{N}(\mathbf{x}\|\mathbf{0}, \boldsymbol{\Sigma}_\gamma^{-1})\,\mathcal{G}(\boldsymbol{\gamma}\|a_\gamma, b_\gamma) \\ q(\mathbf{x}, \boldsymbol{\theta}) & \propto q_1(\mathbf{x})\,q_2(\boldsymbol{\gamma})\,q_3(\boldsymbol{\nu}) \\ \boldsymbol{\theta} & = [\boldsymbol{\gamma}, \nu]^T \end{cases}$
Error prior	$\sigma_e^2 \geq 0$	i.i.d Gaussian: $\mathcal{N}(\boldsymbol{\xi}\|\mathbf{0}, \sigma_\xi^2\,\mathbf{I})$	i.i.d Gaussian: $\mathcal{N}(\boldsymbol{\xi}\|\mathbf{0}, \sigma_\xi^2\,\mathbf{I})$	Student-t prior: $St(\boldsymbol{\epsilon}\|\alpha_\epsilon) = \int \mathcal{N}(\boldsymbol{\epsilon}\|\mathbf{0}, \boldsymbol{\Sigma}_\nu^{-1})\,\mathcal{G}(\boldsymbol{\nu}\|\frac{\alpha_\epsilon}{2}, \frac{\alpha_\epsilon}{2})\,d\boldsymbol{\nu}$
Likelihood	-	$\mathcal{N}(\mathbf{y}\|\mathbf{C}\,\mathbf{x} + \sigma_e^2\,\mathbf{1}_a,\, \sigma_\xi^2\,\mathbf{I})$	$\mathcal{N}(\mathbf{y}\|\mathbf{H}\,\mathbf{x} + \sigma_e^2\,\mathbf{1}_a,\, \sigma_\xi^2\,\mathbf{I})$	$\mathcal{N}(\mathbf{y}\|\mathbf{H}\mathbf{x},\, \boldsymbol{\Sigma}_\nu^{-1})$
Power prior	$\mathbf{x} \succeq \mathbf{0}$	Double Exponential: $p(\mathbf{x}\|\gamma, \beta) = \left(\frac{\gamma}{2}\right)^N \exp\left[-\gamma\|\mathbf{x}\|_{\beta=1}\right]$	Double Exponential: $p(\mathbf{x}\|\gamma, \beta) = \left(\frac{\gamma}{2}\right)^N \exp\left[-\gamma\|\mathbf{x}\|_{\beta=1}\right]$	Student-t prior: $St(\mathbf{x}\|\alpha_x) = \int \mathcal{N}(\mathbf{x}\|\mathbf{0}, \boldsymbol{\Sigma}_\gamma^{-1})\,\mathcal{G}(\boldsymbol{\gamma}\|\frac{\alpha_x}{2}, \frac{\alpha_x}{2})\,d\boldsymbol{\gamma}$
Hidden variable	-	-	-	$\boldsymbol{\gamma}$, $\boldsymbol{\nu}$ (vector)
Parameter prior	Estimated bound on $\hat{\beta}$	σ_e^2, σ_ξ^2, γ (scalar) Jeffrey priors	σ_e^2, σ_ξ^2, γ (scalar) Jeffrey priors	$\boldsymbol{\gamma}$, $\boldsymbol{\nu}$ (vector) Gamma priors

A

Wind Flow Refraction

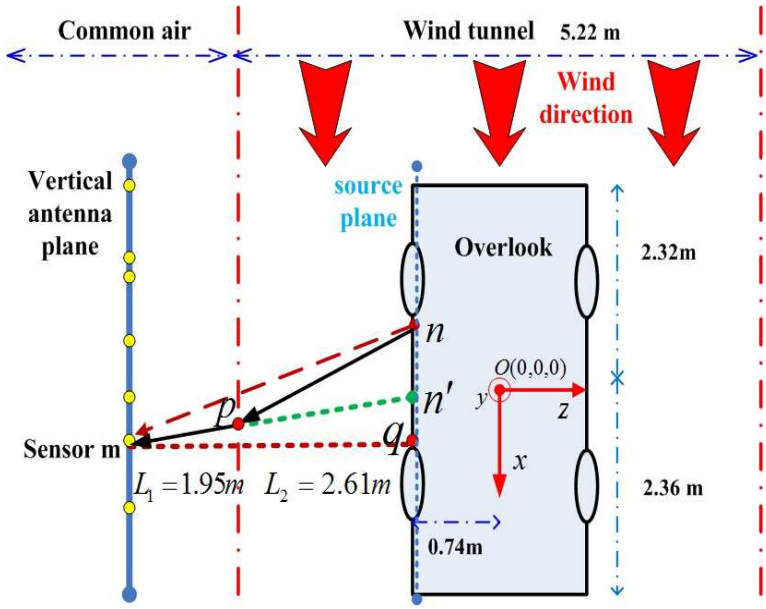

Figure A.1: Wind reflection in wind tunnel experiment.

As shown in Fig.A.1, for source n, we suppose the wind refraction takes place at point p on the interface between the common air and wind flow. We firstly calculate the actual propagation path $r_{n,m}$ and then obtain the actual propagation time $\tau_{n,m}$.

When the medium is uniform, $r_{n,m}$ is of a geometric distance calculated as follows:

$$r_{n,m} = \sqrt{(p_{nx} - \bar{p}_{mx})^2 + (p_{ny} - \bar{p}_{my})^2 + (p_{nz} - \bar{p}_{mz})^2}\,, \tag{A.1}$$

where $\bar{\mathbf{P}} = [\bar{\mathbf{p}}_1, \cdots, \bar{\mathbf{p}}_M]^T$ represents 3D coordinates of sensor array; and sensor m has $\bar{\mathbf{p}}_m = [\bar{p}_{mx}, \bar{p}_{my}, \bar{p}_{mz}]^T$. The corresponding propagation time is $\tau_{n,m} = r_{n,m}/c_0$, with c_0 being acoustic speed in the common air.

When the medium is not uniform due to the wind flow, $r_{n,m}$ is not a geometric distance, but composed of two geometric parts as follows:

$$r_{n,m} = d_{n,p} + d_{p,m}\,, \tag{A.2}$$

Where $d_{n,p}$ is the geometric distance from source n to refraction point p, and $d_{p,m}$ is the one from point p to sensor m.

Therefore, actual $r_{n,m}$ can be calculated by knowing the position of refraction point p.

$\tau_{n,m}$ subsequently is made up of two parts:

$$\tau_{n,m} = \tau_{n,p} + \tau_{p,m} = d_{n,p}/c_1 + d_{p,m}/c_0\,, \tag{A.3}$$

where $c_1 = ||\vec{c}_0 + \vec{v}||$, with $\vec{v}$ being the speed vector of the wind flow. Since it is complicated to obtain c_1, we use the equivalent source n' to calculate the equivalent $\tau_{n',p}$. We suppose that the sensor m seems to receive the signal from equivalent source n', instead of the source n, as if there were no wind influence. This means $\tau_{n,m} = \tau_{n',m} = d_{n',m}/c_0$. Taking the above equation into Eq.(A.3), we get

$$\tau_{n,m} = d_{n',p}/c_0 + d_{p,m}/c_0\,. \tag{A.4}$$

Therefore, actual $\tau_{n,m}$ depends on the positions of the equivalent source n' and refraction point p.

In Fig.A.1, when the positions of sensor m ($\bar{\mathbf{p}}_m = [\bar{p}_{mx}, \bar{p}_{my}, \bar{p}_{mz}]^T$) and source n ($\mathbf{p}_n = [p_{nx}, p_{ny}, p_{nz}]^T$) are given, we firstly locate the equivalent source n' ($\mathbf{p}_{n'} = [p_{n'x}, p_{n'y}, p_{n'z}]^T$) and then determine the refraction point p ($\mathbf{p}_p = [p_{px}, p_{py}, p_{pz}]^T$).

For equivalent source n', since the wind direction is parallel to the ground, the displacement exists only in horizontal direction. According to the coordinate in Fig.A.1, $\mathbf{p}_{n'}$ can be obtained as:

$$p_{n'x} = p_{xn} + d_{n,n'}\,, \quad p_{n'y} = p_{ny}\,, \quad p_{n'z} = p_{nz}\,, \tag{A.5}$$

where $d_{n',n}$ denotes the horizontal displacement, it can be calculated as

$$d_{n',n} = d_{n',p}\, v/c_0, \tag{A.6}$$

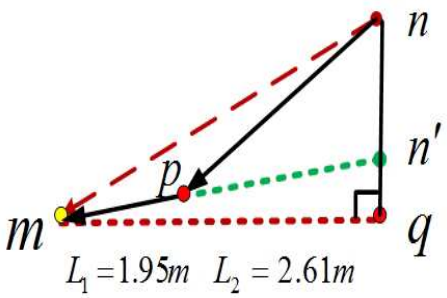

Figure A.2: Solutions of triangle problems.

where $d_{n',p}$ can be calculated by solving the triangles $\Delta pnn'$ and $\Delta mnn'$ in Fig.A.2 which is part of Fig.A.1:

$$d_{n',p} = \left[\frac{v}{c_0}(p_{nx} - p_{px}) + \sqrt{\frac{v^2}{c_0^2}(p_{nx} - p_{px})^2 + d_{n,p}^2(1 - \frac{v^2}{c_0^2})}\right] /(1 - \frac{v^2}{c_0^2}) , \tag{A.7}$$

where the wind speed naturally satisfies $v < c_0$. And we can get p_{px} by solving the triangle $\Delta mn'q$ on the horizontal plane in Fig.A.1, and similarly p_{py} on the vertical plane. Then $\mathbf{p}_p$ is calculated by

$$\begin{cases} p_{px} = \frac{L_1}{L_1+L_2} p_{n'x} + \frac{L_2}{L_1+L_2} \bar{p}_{mx} \\ p_{py} = \frac{L_1}{L_1+L_2} p_{n'y} + \frac{L_2}{L_1+L_2} \bar{p}_{my} \\ p_{pz} = \bar{p}_{mz} + L_1 \end{cases} , \tag{A.8}$$

where L_1 is the distance from the sensor plane to the wind interface, and L_2 is the distance from the wind interface to the vehicle plane as shown in Fig.A.1. Let $d'_{m,n'}$ denote the projection of $d_{m,n'}$ on the ground. If $d'_{m,n'} = L_1 + L_2$, then $\mathbf{p}_p$ can be directly calculated by Eq.(A.8). If $d'_{m,n'} > L_1 + L_2$, we solve the the triangle $\Delta mn'q$ based on solid geometry analysis, and finally we get the same solution as the Eq.(A.8). Above all, $\mathbf{p}_{n'}$ is obtained by substituting Eq.(A.8) into Eq.(A.7).

In conclusion, $r_{n,m}$ in Eq.(A.2) can be derived from the equivalent source n', and $\tau_{n,m}$ in Eq.(A.4) can be calculated from both equivalent source n' and refraction point p. Therefore the steering vector $\mathbf{a}(\mathbf{p}_n, f_l)$ in Eq.(3.16) can be corrected as follow:

$$a_{n,m} = \frac{1}{r_{n,m}} \exp\left[-j2\pi f_l \tau_{n',m}\right] + \rho \frac{1}{r_{-n,m}} \exp\left[-j2\pi f_l \tau_{-n',m}\right] , \tag{A.9}$$

where $a_{n,m} \in \mathbf{a}(\mathbf{p}_n, f_l)$ denotes the steering item from source n to sensor m, and time delay $\tau_{-n',m}$ of ground reflection will be discussed in B.

Our refraction correction can be extended for the other cases:

Case 1: No wind. Since $v = 0$ and $L_2 = 0$, we have $d_{n',p} = d_{n,p}$ from Eq.(A.7), and $\mathbf{p}_p = \mathbf{p}_{n'}$ from Eq.(A.8). That means that both equivalent source n' and refraction point p locate at the position of source n. This conclusion meets the common sense.

Case 2: Wind everywhere. Since $v \neq 0$ and $L_1 = 0$, we have $\mathbf{p}_p = \bar{\mathbf{p}}_m$ from Eq.(A.8). This means refraction point p is the same one as the sensor m, but horizontal displacement $d_{n',n}$ still takes place according to Eq.(A.7).

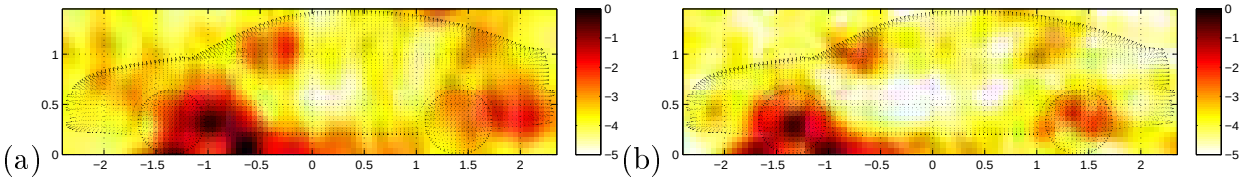

Figure A.3: Propagation corrections in wind tunnel: beamforming on the real data at 2500Hz: (a) without corrections and (b) with corrections

In Fig.A.3, we show the improvement of refraction correction. We use the real data in wind tunnel experiments, and the working frequency is 2500Hz. Take the beamforming method for example, it is seen that without corrections, all the sources are detected on the right side of their real positions, and there are the displacements along the wind direction. After corrections, we get the expected results.

B

Ground Reflection

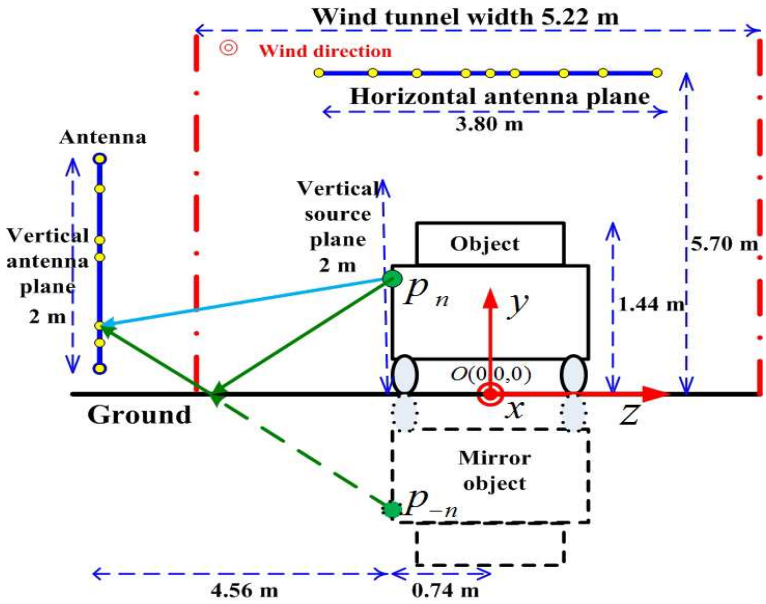

Figure B.1: Ground reflection in wind tunnel experiment.

In Fig.B.1, we show the ground reflection in wind tunnel. For the sensor m, its measured signals consists of two parts: the major one from the direct propagation, the other from the ground reflections. To correct ground reflection, here we apply the mirror source -n which is symmetric to the ground for source n. For the sensor m, it seems to receive the signal by way of the direct path $r_{-n,m}$ from the mirror source. Therefore, the steering vector $\mathbf{a}_r(\mathbf{p}_{-n}, f_l)$ of the ground reflection depends on the positions $\mathbf{p}_{-n}$ of mirror sources.

Considering $\mathbf{p}_n = [p_{nx}, p_{ny}, p_{nz}]^T$ are the coordinates of the source n, its mirror source n' has the coordinates $\mathbf{p}_{n'} = [p_{n'x}, 2p_{ny_0} - p_{ny}, p_{n'z}]^T$, where p_{ny_0} is the vertical coordinate of the ground as shown in Fig.B.1. Since the original O of coordinate system, we have $p_{ny_0} = 0$, thus the mirror position is $\mathbf{p}_{-n} = [p_{nx}, -p_{nx}, p_{nx}]^T$. Therefore the steering vector of ground reflection $\mathbf{a}_r(\mathbf{p}_{-n}, f_l)$ in Eq.(3.16) can be calculated as similarly as the steering vector of direct path $\mathbf{a}_d(\mathbf{p}_n, f_l)$ in Eq.(3.11). Then the correspond steering vector of wind refraction $\mathbf{a}_d(\mathbf{p}_{-n'}, f_l)$ can be corrected according to A.

Base on the analysis of ground reflection, we demonstrate the improvement of propagation corrections in Fig.A.3. As we can see, without correcting the ground reflection ($\rho = 0$), the estimated sources could be involved with the ghost sources outside and under the car. These ghosts are the influences of multi-path propagation. After corrections with $\rho = 0.8$, the beamforming result has much fewer ghosts in the air, on the center and rear of car body, and under the car. Furthermore, the potential strong sources are more concentrated. All the other methods are as improved as to the beamforming.

C

STBT Matrix Approximation

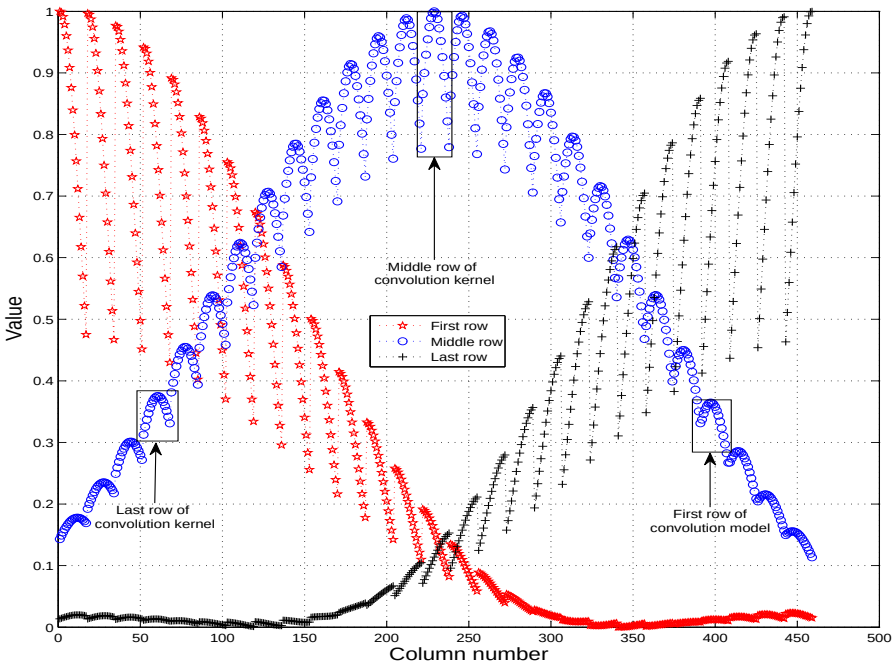

Figure C.1: Row item distributions of $\tilde{\mathbf{C}}$ and invariant convolution kernel $\mathbf{h}$. Simulated at 2500Hz, other configurations are the same with Chapter 5.3.

For the ith source $i \in [1, \cdots, N]$, we suppose that there exists an averaged distance $\bar{r}_i = \frac{1}{M}\sum_{m=1}^{M} r_{i,m}$ from sources to the sensor plane, satisfying $\bar{r}_i / r_{i,m} \approx 1$ for any sensor $m \in [1, \cdots, M]$.

According to the above assumption, each item $c_{i,j} \in \mathbf{C}$ in Eq.(7.1) can be approximated by:

$$\begin{aligned}\hat{c}_{i,j} &\approx \left| \frac{1}{\sum_{m=1}^{M} \frac{1}{r_{i\,m}^2}} \sum_{m=1}^{M} \frac{1}{r_{i\,m}\, r_{j\,m}} e^{\jmath \frac{2\pi f_l}{c_0}(r_{i\,m} - r_{j\,m})} \right|^2 \\ &= \bar{r}_i^2 \frac{1}{M^2} \left| \sum_{m=1}^{M} e^{\jmath \frac{2\pi f_l}{c_0}(r_{i,m} - r_{j,m})} \right|^2 \frac{1}{\bar{r}_j^2}, \end{aligned} \tag{C.1}$$

According to Eq.(7.6), we have

$$\tilde{c}_{i,j} = \frac{1}{M^2} \left| \sum_{m=1}^{M} e^{\jmath \frac{2\pi f_l}{c_0}(r_{i,m} - r_{j,m})} \right|^2, \tag{C.2}$$

where $r_{i\,m}$ denotes the propagation distance from ith discrete source (at the position $\mathbf{p}_i$ on the discrete source plane) to the mth sensor; f_l denotes the lth frequency bin; M is the total number of sensors; $i, j \in [1, \cdots, N]$; and c_0 is the acoustic propagation speed. In Fig.C.1, we show the item distribution on the rows of $\tilde{\mathbf{C}}$, and the relation between $\tilde{\mathbf{C}}$ and invariant convolution kernel $\mathbf{h}$. The middle row refers to $\mathbf{h}$, and some particular items of this row constitute the kernel elements.

According to Eq.(C.2), we get $\tilde{c}_{i,j} = \tilde{c}_{j,i}$, $\tilde{\mathbf{C}} = \tilde{\mathbf{C}}^T$. Therefore, $\tilde{\mathbf{C}}$ is a symmetric matrix as shown on the top of Fig.(7.8). And $\tilde{\mathbf{C}}$ can be expressed by subblock matrices $\tilde{\mathbf{C}}_{q,l}$ as follows:

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

where $\tilde{\mathbf{C}}_{q,l}$ with $q, l \in [1, \cdots, N_c]$ denotes the subblock matrix at qth-row and lth-column block of $\tilde{\mathbf{C}}$; and $\tilde{\mathbf{C}}$ has the number of $N_c \times N_c$ subbolcks as shown in the middle of Fig.7.8. $c_{p,k}^{(q,l)}$ with $p, k \in [1, \cdots, N_r]$ denotes the pth-row and kth-column item of $\tilde{\mathbf{C}}_{q,l}$, and $\tilde{\mathbf{C}}_{q,l}$ has the size of $N_r \times N_r$.

We then suppose that

$$|r_{i,m} - r_{j,m}| \approx |r_{i+1,m} - r_{j+1,m}|, \quad \lfloor \frac{i}{N_r} \rfloor = \lfloor \frac{j}{N_r} \rfloor \quad , (q = l) \tag{C.4}$$

where $\lfloor \cdot \rfloor$ denotes the integer part, which reflects that the ith and jth, $i+1$th and $j+1$th discrete sources are on the same column on the source power image. This approximation is shown in the right bottom of Fig.C.2.

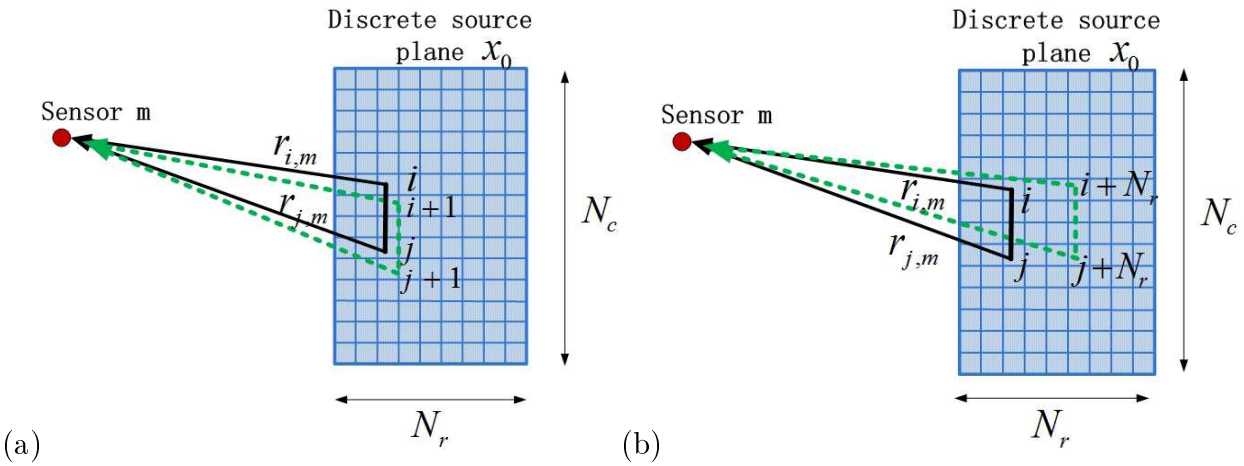

Figure C.2: Assumptions for STBT matrix $\hat{\mathbf{C}}$ approximation (a) Approximation for Toeplitz in a subblock (b) Approximation for block Toeplitz block

Based on Eq.(C.2), it yields $\tilde{c}_{i,j} = \tilde{c}_{i+1,j+1}$ for i and j belong to one subblock. Since index i and j are periodically changing, we then have $\tilde{c}_{i,j} = \tilde{c}_{i+N_r,j+N_r}$ in two subblock.

According to Eq.(C.3) and (C.4), for any $\tilde{c}_{p,k}^{(q,l)}$, $\tilde{c}_{p+1,k+1}^{(q,l)} \in \tilde{\mathbf{C}}_{q,l}$, we have $\tilde{c}_{p+1,k+1}^{(q,l)} = \tilde{c}_{i+1,j+1} = \tilde{c}_{i,j} = \tilde{c}_{p,k}^{(q,l)}$. Therefore, subblock $\tilde{\mathbf{C}}_{q,l}$ is a Toeplitz matrix, as shown in the middle of Fig.7.8.

For any $\tilde{c}_{p,k}^{(q,l)} \in \tilde{\mathbf{C}}_{q,l}$ and $\tilde{c}_{p,k}^{(q+1,l+1)} \in \tilde{\mathbf{C}}_{q+1,l+1}$, we get $\tilde{c}_{k,l}^{(q+1,l+1)} = \tilde{c}_{i+N_r,j+N_r} = \tilde{c}_{i,j} = \tilde{c}_{p,k}^{(q,l)}$, and $\tilde{\mathbf{C}}_{q,l} = \tilde{\mathbf{C}}_{q+1,l+1}$. Therefore, $\tilde{\mathbf{C}}$ is a block Toeplitz matrix, as shown in the middle of Fig.7.8.

Above all, $\tilde{\mathbf{C}}$ is proved ton be a STBT matrix, as shown in Fig.7.8.

D

Separable Convolution

In order to accelerate the 2D convolution, we here investigate the separability of the convolution kernel. Firstly, we introduce the **separability condition** of 2D convolution kernel [93, 116] as follows:

Let $r(\mathbf{h})$ denote the rank of a $N_r \times N_r$ convolution kernel $\mathbf{h}$. And $\mathbf{h}_1$, $\mathbf{h}_2$ denote two column vectors with the same length of N_r. If $r(\mathbf{h}) = 1$, we can get $\mathbf{h} = \mathbf{h}_1 * \mathbf{h}_2^T$. In other words, one operation of 2D convolution can be replaced by using two operations of 1D convolution, **as long as** the rank of the 2D convolution kernel equals one.

For an input vector with the length N, the computational complexity of 2D convolution (using $\mathbf{h}$) is $\mathrm{O}(N_r^2\, N)$, while the separable convolutions using $\mathbf{h}_1$, $\mathbf{h}_2$ can be greatly reduced into $\mathrm{O}(2N_r\, N)$ as shown in Table 7.1. Meanwhile, the storage of convolution kernels is also reduced from N_r^2 to $2N_r$.

Even if the separability condition can not be always guaranteed, the 2D convolution can be also approximated by separable convolution, as long as the kernel separability is big enough. We define the kernel separability as:

$$\rho = \frac{\lambda_1^2}{\sum_{i=2}^{N_r} \lambda_i^2}\,, \tag{D.1}$$

where $[\lambda_1, \cdots, \lambda_{N_r}]^T$ denotes eigenvalues of $\mathbf{h}$, with $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_{N_r} \geq 0$.

Then we want to derive the 1D convolution vectors $\mathbf{h}_1$ and $\mathbf{h}_2$. Since every

real symmetric matrix $\mathbf{h}$ can be diagonalized [59], we take the EigenValue Decomposition (EVD) of of $\mathbf{h}$ as:

$$\mathbf{h} = \mathbf{U}\Lambda\mathbf{U}^T = \sum_{i=1}^{N_r} \lambda_i \mathbf{u}_i \mathbf{u}_i^T \,, \tag{D.2}$$

where $\mathbf{U} = [\mathbf{u}_i]_{N_r}$ is a $N_r \times N_r$ orthogonal matrix, whose columns $\mathbf{u}_i$ with $i \in [1, \cdots, N_r]$ are eigenvectors of $\mathbf{h}$; and $\Lambda = Diag[\lambda_i]_{N_r}$ is a real and diagonal matrix, whose diagonal values λ_i are eigenvalues of $\mathbf{h}$.

From Eq.(D.2), we can approximate $\mathbf{h}$ by using the biggest eigenvalue λ_1 and its corresponding eigenvector $\mathbf{u}_1$ as $\mathbf{h} \approx \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T$. Then we define the approximating error as:

$$\varepsilon = \frac{||\mathbf{h} - \lambda_1 \mathbf{u}_1 * \mathbf{u}_1'^T||_2^2}{\sum_{i=1}^{N_r} \lambda_i^2} \times 100\% = \frac{||\sum_{i=2}^{N_r} \lambda_i \mathbf{u}_i \mathbf{u}_i^T||_2^2}{\sum_{i=1}^{N_r} \lambda_i^2} \times 100\% \,, \tag{D.3}$$

where the valid convolution satisfies $\mathbf{u}_1 * \mathbf{u}_1'^T = \mathbf{u}_1 \mathbf{u}_1^T$, and $\mathbf{u}_1' = [u_{1,N_r}, \cdots, u_{1,1}]$ denotes symmetric vector of $\mathbf{u}_1 = [u_{1,1}, \cdots, u_{1,N_r}]$. Since $\mathbf{U} = [\mathbf{u}_i]_{N_r}$ is an orthogonal matrix, we then have $\mathbf{u}_i \mathbf{u}_j^T = \mathbf{I}$ and $|\mathbf{u}_i^T \mathbf{u}_j| = 1$ for $i = j$, $|\mathbf{u}_i^T \mathbf{u}_j| = 0$ for $i \neq j$. According to Eq.(D.2), $||\sum_{i=2}^{N_r} \lambda_i \mathbf{u}_i \mathbf{u}_i^T||_2^2 \leq \sum_{i=2}^{N_r} \lambda_i^2 ||\mathbf{u}_i \mathbf{u}_i^T||_2^2 = \sum_{i=2}^{N_r} \lambda_i^2$. Finally, according to Eq.(D.3), the upper bound of approximating error is:

$$\varepsilon \leq \frac{\sum_{i=2}^{N_r} \lambda_i^2}{\lambda_1^2 + \sum_{i=2}^{N_r} \lambda_i^2} \times 100\% = \frac{1}{\rho + 1} \times 100\% \,, \tag{D.4}$$

where $\rho = \frac{\lambda_1^2}{\sum_{i=2}^{N_r} \lambda_i^2}$ denotes the separability degree. The bigger ρ is, the more separable $\mathbf{h}$ becomes. If $r(\mathbf{h}) = 1$, then $\mathbf{h} = \lambda_1 \mathbf{u}_1 * \mathbf{u}_1^T$, so that $\varepsilon = 0$ in Eq.(D.4) and $\rho \to \infty$ in Eq.(D.1). Therefore, we can obtain the 2D separable convolution kernel as

$$\mathbf{h} \approx \mathbf{h}_{sep} = \lambda_1 \mathbf{u}_1 * \mathbf{u}_1'^T \,, \tag{D.5}$$

where λ_1 denotes the biggest eigenvalue of $\mathbf{h}$; $\mathbf{u}_1$ denotes the corresponding eigenvector of $\mathbf{h}$; and $\mathbf{u}_1'$ denotes the symmetric vector of $\mathbf{u}_1$. The separability error ε has the upper bound in Eq.(D.4).

In conclusion, if the 2D convolution kernel $\mathbf{h}$ is a non-negative, real and symmetric matrix, it can be approximated by a separable convolution kernel which consists of the bigger eigenvalue and its corresponding eigenvector of $\mathbf{h}$. And the approximating error of separable convolution is relatively small, as long as the kernel separability is big enough.

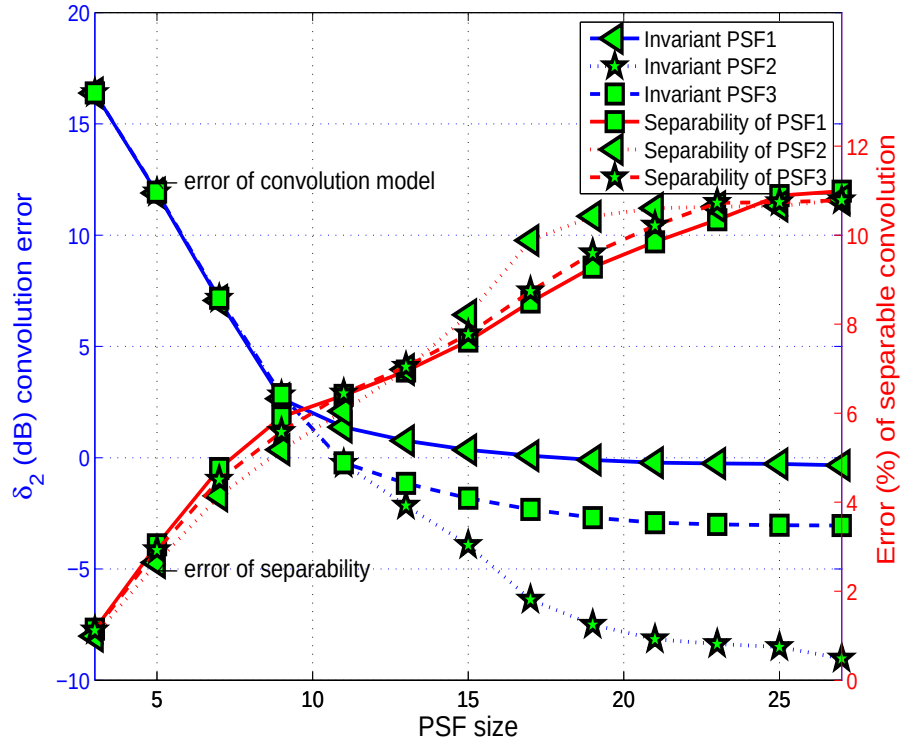

Figure D.1: Performance of 2D separable convolution VS 2D invariant convolution

In Fig.D.1, we show the kernel separabilities of three 2D convolution kernels (PSFs) versus different PSF sizes on the same simulation of Chapter 7.3. The input of source power image $\mathbf{x}_0$ has the size of $N_r = 17 \times N_c = 27$. Thus the propagation matrix $\mathbf{C}$ has the size of $N = 459 \times N = 459$. The PSF1 is obtained by the items of $c_{1,j}$, $j \in [1, \cdots, N]$ on the first line of $\mathbf{C}$; PSF2 is from the middle line ($i = 230$) of $\mathbf{C}$; PSF3 is from the line of $i = 300$. Compared with the 'variant' kernel in Eq.(7.12), the 'invariant' kernel is fixed in Eq.(7.13), and does not varies along with the index of output item.

For three red curves in Fig.D.1, we show the invariant convolution kernel separabilities of PSF1, PSF2 and PSF3 respectively. All of three separabilities share the same trend and are close to each other. The larger the size is, the bigger error of separability is. But separability error remains relatively small ($< 11\%$). Particularly, when kernel size is about 15×15 which is close to $N_r = 17$, the separability error is just around 5%. This is because the symmetric structures of PSF1, PSF2 and PSF3 can well meet the separable conditions in Eq.(D.4).

For three blue curves in Fig.D.1, we show the convolution approximating errors between the invariant convolution model in Eq.(7.13) and simulated beamforming result in Eq.(4.12). According to the separability errors (blue curves), the invariant convolutions can be effectively approximated by two

1D-convolutions. It is seen that the larger the kernel size is, the bigger the convolution approximating error becomes. But PSF2 can achieve the smallest convolution approximating errors using a relatively small kernel size (15×15), so that PSF2 can be well approximated by a 2D separable convolution.

In conclusion, figure D.1 tells that PSF2 can be a 2D separable convolution which contributes an efficient and effective convolution to approximate the forward model of source power propagation in Eq.(5.4).

E

Proof: Power propagation matrix $\mathbf{C}$ is non-invertible

As discussed in Eq.(4.49), the power propagation matrix $\mathbf{C}$ can be modeled as

$$\mathbf{C} = |\tilde{\mathbf{A}}^T \mathbf{A}|.^2 , \tag{E.1}$$

where $\tilde{\mathbf{A}} = [\tilde{\mathbf{a}}_1 \cdots \tilde{\mathbf{a}}_N]$, $\tilde{\mathbf{A}} \in \mathbb{C}^{M \times N}$ denotes the beamforming steering matrix, defined in Eq.(4.12); $\mathbf{A} = [\mathbf{a}_1 \cdots \mathbf{a}_N]$, $\mathbf{A} \in \mathbb{C}^{M \times N}$ denotes the signal propagation steering matrix, defined in Eq.(3.15); operator $|\cdot|$ denotes the absolute modulus of each item of a matrix; $(\cdot).^2$ denotes the square of each item of a matrix; $\mathbf{C} \in \mathbb{R}^{N \times N}$ is a square matrix full of non-negative values.

According to Eq.(E.1) and two non-linear operator $(\cdot).^2$ and$|\cdot|$, we define two matrices $\mathbf{B}$ and $\mathbf{D}$ as

$$\begin{cases} \mathbf{B} &= \tilde{\mathbf{A}}^T \mathbf{A} &= [\mathbf{b}_1 \cdots \mathbf{b}_N] \\ \mathbf{D} &= (\mathbf{B}).^2 &= [\mathbf{b}_1^2 \cdots \mathbf{b}_N^2] \\ \mathbf{C} &= |\mathbf{D}|. &= [|\mathbf{b}_1^2| \cdots |\mathbf{b}_N^2|] \end{cases} , \tag{E.2}$$

where $\mathbf{B} = [\mathbf{b}_1, \cdots, \mathbf{b}_N]$ is made of N column vectors $\mathbf{b}_n \in \mathbb{C}^{N \times 1}$ with $n \in [1, \cdots, N]$; and $\mathbf{B}, \mathbf{D} \in \mathbb{C}^{N \times N}$ are two square matrices; operator $().^2$ and $||.$ respectively denote the square value and absolute value of each item of matrix or vector. Let matrix rank of $rank(\mathbf{B}) = K$. Since $\mathbf{A} \in \mathbb{C}^{M \times N}$ and $\tilde{\mathbf{A}} \in \mathbb{C}^{M \times N}$ with $M < N$, their matrix ranks meet $rank(\tilde{\mathbf{A}}) \leq M < N$ and $rank(\mathbf{A}) \leq M < N$. According to the property of matrix rank [59],

$\mathbf{B} \in \mathbb{C}^{N\times N}$ has a matrix rank:

$$rank(\mathbf{B}) = K,\ \ 0 < K \leq min\{rank(\mathbf{A}), rank(\tilde{\mathbf{A}})\} \leq M < N\,. \tag{E.3}$$

Therefore, $\mathbf{B}$ is a singular matrix and cannot be invertible.

In the followings, we want to firstly prove that matrix $\mathbf{D} \in \mathbb{C}^{N\times N}$ is also singular and non-invertible, and then prove that so does matrix $\mathbf{C} \in \mathbb{R}^{N\times N}$.

According to Eq.(E.3), since $rank(\mathbf{B}) = K < N$, there exist K linear-independent column vectors in $\mathbf{B}$. Suppose that $[\mathbf{b}_1, \cdots, \mathbf{b}_K]$ are linear-independent, so that the column vector $\mathbf{b}_{K+1}$ can be a linear combination of $[\mathbf{b}_1, \cdots, \mathbf{b}_K]$:

$$\mathbf{b}_{K+1} = \sum_{m=1}^{K} \alpha_m\, \mathbf{b}_m\,, \tag{E.4}$$

where α_m with $m \in [1, \cdots, K]$ denotes the coefficient (real value), satisfying $\sum_{m=1}^{K} |\alpha_m| \neq 0$. Since $\mathbf{b}_m$ with $m \in [1, \cdots, K+1]$ is a column vector, we take the operator of $(\cdot).^2$ (item square of a matrix) on both sides of Eq.(E.4), then it yields

$$\begin{array}{ll} \mathbf{b}_{K+1}^2 & = (\sum_{m=1}^{K} \alpha_m\, \mathbf{b}_m).^2 \\ & = \sum_{m=1}^{K} \alpha_m^2\, \mathbf{b}_m^2 + 2\, \sum_{m,\, n=1,\, n\neq m}^{K} \alpha_m\, \alpha_n\, \mathbf{b}_m\, .\mathbf{b}_n \end{array}\,, \tag{E.5}$$

where operator . denotes the vector dot product of two vectors. We discuss Eq.(E.5) as follows:

- If linear-independent column vectors $[\mathbf{b}_1, \cdots, \mathbf{b}_K]$ are mutually orthogonal, it yields $\mathbf{b}_m\, .\mathbf{b}_n = \mathbf{0}_N$ for $m \neq n$, with $\mathbf{0}_N = [0, \cdots, 0]_N^T$ denotes the column zero-vector. In this case, equation (E.5) becomes

$$\mathbf{b}_{K+1}^2 = \sum_{m=1}^{K} \alpha_m^2\, \mathbf{b}_m^2 + \mathbf{0}_N\,, \tag{E.6}$$

  where α_m^2 with $m \in [1, \cdots, K]$ is a coefficient (real value); since satisfying $\sum_{m=1}^{K} |\alpha_m| \neq 0$, it yields $\sum_{m=1}^{K} |\alpha_m^2| \neq 0$. Equation (E.6) shows that column vector $\mathbf{b}_{K+1}^2$ is the linear combination of K number of column vectors $[\mathbf{b}_1^2, \cdots, \mathbf{b}_K^2]$. According to the definition of matrix $\mathbf{D} \in \mathbb{C}^{N\times N}$, it is proved that $rank(\mathbf{D}) \leq K < N$, so that $\mathbf{D} \in \mathbb{C}^{N\times N}$ is singular and non-invertible.

- If linear-independent column vectors $[\mathbf{b}_1, \cdots, \mathbf{b}_K]$ are not mutually orthonormal, $\mathbf{b}_m$ with $m \in [1, \cdots, K]$ can be modeled by K number of

standard orthonormal basis $\mathbf{e}_i$ with $i \in [1, \cdots, K]$ as:

$$\mathbf{b}_m = \sum_{i=1}^{K} \beta_{m\,i}\, \mathbf{e}_i\,, \tag{E.7}$$

where $\beta_{m\,i}$ with $i \in [1, \cdots, K]$ denotes the coefficient (real value), satisfying $\sum_{i=1}^{K} |\beta_{m\,i}| \neq 0$; and orthonormal basis $\mathbf{e}_i.\mathbf{e}_j = \mathbf{0}$ for $i \neq j$; $\mathbf{e}_i\,.\mathbf{e}_j = \mathbf{1}_i$ for $i = j$, with $\mathbf{1}_i = [0, \cdots, 1_i, \cdots, 0]_N^T$ denotes the column vector whose ith item is 1 and others are 0. Inserting $\mathbf{b}_m = \sum_{m=1}^{K} \beta_{m\,i}\,.\mathbf{e}_i$ of Eq.(E.7) into Eq.(E.4), it yields

$$\begin{array}{rl} \mathbf{b}_{K+1} & = \sum_{m=1}^{K} \alpha_m\, \mathbf{b}_m \\ & = \sum_{m=1}^{K} \alpha_m \sum_{i=1}^{K} \beta_{m\,i}\, \mathbf{e}_i \\ & = \sum_{i=1}^{K} (\sum_{m=1}^{K} \alpha_m\, \beta_{m\,i})\, \mathbf{e}_i \\ & = \sum_{i=1}^{K} \gamma_i\, \mathbf{e}_i \end{array}\,, \tag{E.8}$$

where $\gamma_i = \sum_{m=1}^{K} \alpha_m\, \beta_{m\,i}$ with $i \in [1, \cdots, K]$ denote the coefficient (real value); since $\sum_{m=1}^{K} |\alpha_m| \neq 0$ and $\sum_{i=1}^{K} |\beta_{m\,i}| \neq 0$, it yields $\sum_{m=1}^{K} |\gamma_i| \neq 0$. Then replacing Eq.(E.8) into Eq.(E.5), it yields

$$\begin{array}{rl} \mathbf{b}_{K+1}^2 & = (\sum_{i=1}^{K} \gamma_i\, \mathbf{e}_i).^2 \\ & = \sum_{i=1}^{K} \gamma_i^2\, \mathbf{e}_i^2 + 2 \sum_{i,\,j=1,\,i\neq j}^{K} \gamma_i\, \gamma_j\, \mathbf{e}_i\,.\mathbf{e}_j \\ & = \sum_{i=1}^{K} \gamma_i^2\, \mathbf{e}_i^2 + \mathbf{0}_N\,, \end{array}\,, \tag{E.9}$$

where K number of vectors $\mathbf{e}_i^2$ with $i \in [1, \cdots, K]$ are also the standard orthonormal basis; since $\sum_{m=1}^{K} |\gamma_i| \neq 0$, it yields $\sum_{m=1}^{K} |\gamma_i^2| \neq 0$. Therefore, equation (E.9) shows that column vector $\mathbf{b}_{K+1}^2$ is the linear combination of K number of standard orthonormal basis $[\mathbf{e}_1^2, \cdots, \mathbf{e}_K^2]$. According to the definition of matrix $\mathbf{D} \in \mathbb{C}^{N\times N}$, it is proved that $rank(\mathbf{D}) \leq K < N$, so that $\mathbf{D} \in \mathbb{C}^{N\times N}$ is singular and non-invertible.

Now we want to prove that matrix $\mathbf{C} \in \mathbb{R}^{N\times N}$ is also singular and non-invertible.

According to the definition of three operators $|\cdot|.$, $(\cdot).^2$ and $.$, they all manipulate each item of a vector or matrix. For the complex vector $\mathbf{b}_{K+1}$, we have the operator property as

$$|\mathbf{b}_{K+1}^2| = |\mathbf{b}_{K+1}|.^2 = \mathbf{b}_{K+1}.\mathbf{b}_{K+1}^T\,. \tag{E.10}$$

Replacing Eq.(E.5) into Eq.(E.10), we get

$$\begin{array}{rl} |\mathbf{b}_{K+1}^2| & = (\sum_{m=1}^{K} \alpha_m\, \mathbf{b}_m)\,.(\sum_{m=1}^{K} \alpha_m\, \mathbf{b}_m^T) \\ & = \sum_{m=1}^{K} \alpha_m^2\, |\mathbf{b}_m^2| + \sum_{m,\,n=1,\,m\neq n}^{K} \alpha_m\, \alpha_n\, \mathbf{b}_m\, \mathbf{b}_n^T \end{array}\,, \tag{E.11}$$

where equation (E.11) is quite similar with the Eq.(E.5). So that we can use the similar methods in Eq.(E.5)-Eq.(E.9) to prove the fact that matrix $\mathbf{C} \in \mathbb{R}^{N \times N}$ is also singular and non-invertible.

.

- La Chine est une fois réveillée, le monde tremblera pour elle.

Napoléon Bonaparte

F

Author's publications during 3 years of PhD

Articles published in journals:

- N. CHU, A. Mohammad-Djafari and J. Picheral, Robust Bayesian super-resolution approach via sparsity enforcing priors for near-field acoustic source imaging, **Journal of Sound and Vibration**, Vol. 332, No. 18, pp 4369-4389, Feb. 2013. DOI: 10.1016/j.jsv.2013.02.037.
- N. CHU, J. Picheral and A. Mohammad-Djafari, N. Gac, A robust super-resolution approach with sparsity constraint in acoustic imaging, **Applied Acoustics**, accepted, 08.2013. DOI: 10.1016/j.apacoust.2013.08.007, to be published vol.76, pp.197-208, 2014.

Preprints in journals:

- N. CHU, N. Gac, A. Mohammad-Djafari, and J. Picheral, A fast Bayesian acoustic imaging method based on convolution model and GPU realization, **Journal of Sound and Vibration**
- N. CHU, A. Mohammad-Djafari, N. Gac, and J. Picheral, A hierarchical variational Bayesian inference approach via Student's-t priors for acoustic imaging with non-stationary noises, **Journal of the Acoustical Society of America**

- N. CHU, N. Gac, A. Mohammad-Djafari and J. Picheral, A fast varational Bayesian method using 2D invariant convolution, **International Journal of Aeroacoustics**

Papers published in international conferences:

- N. CHU, A. Mohammad-Djafari, N. Gac, and J. Picheral, An efficient variational Bayesian inference approach via Student's-t priors for acoustic imaging in colored noises , **Journal of the Acoustical Society of America**, Vol. 133, No.5. Pt.2, POMA Vol 19, pp. 055031-40, 2013.
- N. CHU, A. Mohammad-Djafari and J. Picheral, A Bayesian sparse inference approach in near-field wideband aeroacoustic imaging, 2012 IEEE International Conference on Image Processing, Orlando (**ICIP2012**), USA, Sep. 30-Oct. 04, 2012. (EI)
- N. CHU, A. Mohammad-Djafari and J. Picheral, Bayesian sparse regularization in near-field wideband aeroacoustic imaging for wind tunnel test, 2012 IOA annual meeting and 11th Congrès Français d'Acoustique (**ACOUSTICS2012**), Nantes, France, Apr. 23-27, 2012, pp. 1391-1396.
- N. CHU, A. Mohammad-Djafari and J. Picheral, Two robust super-resolution approaches with sparsity constraint and sparse regularization for near-field wideband extended aeroacoustic source imaging, Berlin Beamforming Conference 2012 (**BeBeC2012**), Berlin, Germany, Feb. 22-23, 2012, pp. 29.
- N. CHU, J. Picheral and A.Mohammad-Djafari, A robust super-resolution approach with sparsity constraint for near-field wideband acoustic imaging, IEEE International Symposium on Signal Processing and Information Technology (**ISSPIT2011**), Bilbao, Spain, Dec. 14-17, 2011, pp. 310-315. (EI)

Talks in seminars

- An efficient Bayesian inference approach using 2D invariant convolution approximation in acoustic imaging. Seminar of Information Signal Image viSion (**GdRISIS**), Telecom ParisTech, 46 Rue Barrault 75013 Paris France, May. 2013

- Efficient Bayesian Variational Approximation method in acoustic imaging. Laboratoire d'Analyse Topologie Probabilités (**LATP**) UMR 7353, 39 rue F.Joliot Curie 13453 Marseille France, Mar. 2013

- An Invariant Convolution Model And Bayesian Inversion In Acoustic Imaging, La Journée de l'Image Optique Non-Conventionnelle (**JIONC**) 2013, Mars 2013

- Robust Bayesian super-resolution approach in acoustic imaging. Laboratoire de Mécanique et d'Acoustique (**LMA**) UPR 7051, 31 chemin Joseph-Aiguier 13402 Marseille France, Jan. 2013

- Bayesian super-resolution approach via sparsity enforcing a prior in aeroacoustic imaging. Seminar of **GdRISIS**, Telecom ParisTech, 46 Rue Barrault 75013 Paris France, Nov. 2012.

- Bayesian compressed sensing in near-field wideband aeroacoustic imaging, 1st International Workshop on Compressed Sensing applied to Radar (**CoSeRa2012**), Bonn, Germany, May 14-16, 2012.

- Bayesian methods in acoustic imaging on simulations and real data in wind tunnel. Seminar with researchers from Office National d'Etudes et de Recherches Aérospatiales (**ONERA**), Laboratoire des signaux et systèmes (L2S) UMR 8506, 3 rue Joliot-Curie 91192 Gif-sur-yvette France, Oct. 2011

- Deconvolution methods in high resolution acoustic imaging on real data from wind tunnel S2A. Seminar with researchers of **Renault SAS**, 1 avenue de golf Guyancourt 78288 France, Jul. 2011.

- Robust super-resolution methods with sparsity constraint and sparse regularization in acoustic imaging.**GRESTI** Summer School, Peyresq 04170 France, Aug. 2011

Bibliography

[1] Jean-Luc Adam, Denis Ricot, Flavien Dubief, and Christine Guy. Aeroacoustic simulation of automotive ventilation outlets. *Journal of the Acoustical Society of America*, 123(5):3250–3250, 2008.

[2] Jean-Luc Adam, Denis Ricot, Christophe Lambourg, and Arnaud Menoret. Correlated Beamforming Method for Relevant Aeroacoustic Sources Identification. In *SAE 2009 Noise and Vibration Conference and Exhibition*, pages 2009–01–2234, St. Charles, Illinois, United States, 20-23 May 2009. SAE.

[3] Arash Amini, Ulugbek S Kamilov, Emrah Bostan, and Michael Unser. Bayesian estimation for continuous-time sparse stochastic processes. *Signal Processing, IEEE Transactions on*, 61(4):907–920, 2013.

[4] H.S. Anderson and M.R. Gupta. Joint deconvolution and classification with applications to passive acoustic underwater multipath. *The Journal of the Acoustical Society of America*, 124(5):2973–2983, 2008.

[5] J. Antoni. A Bayesian approach to sound source reconstruction: optimal basis, regularization, and focusing. *The Journal of the Acoustical Society of America*, 131:2873–2890, 2012.

[6] Jérôme Antoni. The spectral kurtosis: a useful tool for characterising non-stationary signals. *Mechanical Systems and Signal Processing*, 20(2):282–307, 2006.

[7] GR Ayers and J Christopher Dainty. Iterative blind deconvolution method and its applications. *Optics letters*, 13(7):547–549, 1988.

[8] C.J. Bahr and L.N. Cattafesta III. Wavespace-based coherent deconvolution. pages NF1676L–13733, 18th AIAA/CEAS Aeroacoustics Conference; 4-6 Jun. 2012; Colorado Springs, CO; United States.

[9] Eric Bavu and Alain Berry. High-resolution imaging of sound sources in free field using a numerical time-reversal sink. *Acta Acustica united with Acustica*, 95(4):595–606, 2009.

[10] Zvika Ben-Haim, Yonina C Eldar, and Michael Elad. Coherence-based performance guarantees for estimating a sparse vector under random noise. *IEEE Transactions on Signal Processing*, 58(10):5030–43, 2010.

[11] Dimitri P Bertsekas. *Nonlinear programming (Second ed.)*. Athena Scientific, 1999.

[12] Georges Bienvenu and Laurent Kopp. Optimality of high resolution array processing using the eigensystem approach. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 31(5):1235–1248, 1983.

[13] D. Blacodon. Array Processing for Noisy Data: Application for Open and Closed Wind Tunnels. *AIAA J.*, 49(1):55–66, 2011.

[14] D. Blacodon. Array processing for noisy data: Application for open and closed wind tunnels. *AIAA journal*, 49(1):55–66, 2011.

[15] Daniel Blacodon. Spectral estimation method for noisy data using a noise reference. *Applied Acoustics*, 72(1):11–21, 2011.

[16] Jens Blauert. *Spatial hearing: the psychophysics of human sound localization*. MIT press, 1997.

[17] C. Bouman and K. Sauer. A generalized gaussian image model for edge-preserving map estimation. *Image Processing, IEEE Transactions on*, 2(3):296–310, 1993.

[18] 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*, 294(4-5):856–879, 2006.

[19] T.F. Brooks and W.M. Humphreys Jr. Extension of DAMAS Phased Array Processing for Spatial Coherence Determination (DAMAS-C). In *12th AIAA/CEAS Aeroacoustics Conference*, pages AIAA–2006–2654. American Institute of Aeronautics and Astronautics, Cambridge, MA, United States, 8-10 May 2006.

[20] E. Candes and J. Romberg. Sparsity and incoherence in compressive sampling. *Inverse problems*, 23(3):969, 2007.

[21] E.J. Candès, M.B. Wakin, and S.P. Boyd. Enhancing sparsity by reweighted ℓ_1 minimization. *Journal of Fourier Analysis and Applications*, 14(5):877–905, 2008.

[22] Jack Capon. High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, 57(8):1408–1418, 1969.

[23] Jacques Chatillon. Influence of source directivity on noise levels in industrial halls: Simulation and experiments. *Applied Acoustics*, 68(6):682–698, 2007.

[24] J.D. Chazot, E. Zhang, and J. Antoni. Acoustical and mechanical characterization of poroelastic materials using a bayesian approach. *The Journal of the Acoustical Society of America*, 131:3240–3250, 2012.

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

[26] S.S. Chen, D.L. Donoho, and M.A. Saunders. Atomic decomposition by basis pursuit. *SIAM journal on scientific computing*, 20(1):33–61, 1999.

[27] Colin Cherry and JA Bowles. Contribution to a study of the cocktail party problem. *The Journal of the Acoustical Society of America*, 32(7):884–884, 1960.

[28] N. Chu, A. Mohammad-Djafari J. Picheral, and N. Gac. A robust super-resolution approach with sparsity constraint in acoustic imaging. *Applied Acoustics*, 76:197–208, 2014.

[29] N. Chu, A. Mohammad-Djafari, and J. Picheral. Robust Bayesian super-resolution approach via sparsity enforcing a priori for near-field aeroacoustic source imaging. *Journal of Sound and Vibration*, 332(18):4369–4389, 2013.

[30] N. Chu, A. Mohammad-Djafari, and J. Picheral. Two robust super-resolution approaches with sparsity constraint and sparse regularization for near-field wideband extended aeroacoustic source imaging. In *Berlin Beamforming Conference 2012 (BeBeC2012)*, page No.29, Berlin, Germany, Feb.22-23,2012.

[31] N. Chu, A. Mohammad-Djafari, and J. Picheral. Bayesian compressed sensing in near-field wideband aeroacoustic imaging. In *1st Inter-*

national Workshop on Compressed Sensing applied to Radar, page No.1041, Bonn, Germany, May.14-16,2012.

[32] N. Chu, A. Mohammad-Djafari, and J. Picheral. Bayesian sparse regularization in near-field wideband aeroacoustic imaging for wind tunnel test. In *2012 IOA annual meeting and 11th Congrés Francais d'Acoustique*, pages 1391–1396, Nantes, France, Apr.23-27,2012.

[33] N. Chu, A. Mohammad-Djafari, and J. Picheral. A Bayesian sparse inference approach in near-field wideband aeroacoustic imaging. In *2012 IEEE International Conference on Image Processing*, pages 2529–2532, Orlando, USA, Sep.30-Oct.4, 2012.

[34] N. Chu, J. Picheral, and A. Mohammad-Djafari. A robust super-resolution approach with sparsity constraint for near-field wideband acoustic imaging. In *IEEE International Symposium on Signal Processing and Information Technology*, pages 310–315, Bilbao, Spain, Dec.14-17,2011.

[35] Ning Chu, Ali Mohammad-Djafari, José Picheral, and Nicolas Gac. An efficient variational Bayesian inference approach via studient's-t priors for acoustic imaging in colored noises. volume 19, page 055031. ASA, 2013.

[36] Ingvar Claesson and Sven Nordholm. A spatial filtering approach to robust adaptive beaming. *IEEE Transactions on Antennas and Propagation*, 40(9):1093–1096, 1992.

[37] David Lem Colton and Rainer Kress. *Inverse acoustic and electromagnetic scattering theory*, volume 93. Springer, 2013.

[38] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on pure and applied mathematics*, 57(11):1413–1457, 2004.

[39] John G Daugman et al. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Optical Society of America, Journal, A: Optics and Image Science*, 2(7):1160–1169, 1985.

[40] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39:1–38, 1977.

[41] N. Dobigeon, A.O. Hero, and J.Y. Tourneret. Hierarchical bayesian sparse image reconstruction with application to mrfm. *IEEE Transactions on Image Processing*, 18(9):2059–2070, 2009.

[42] David L Donoho, Michael Elad, and Vladimir N Temlyakov. Stable recovery of sparse overcomplete representations in the presence of noise. *IEEE Transactions on Information Theory*, 52(1):6–18, 2006.

[43] R.P. Dougherty. Extensions of DAMAS and Benefits and Limitations of Deconvolution in Beamforming. In *11th AIAA/CEAS Aeroacoustics Conference*, pages 1–13, Monterey, CA, USA, 23-25 May, 2005.

[44] K. Ehrenfried and L. Koop. Comparison of iterative deconvolution algorithms for the mapping of acoustic sources. *AIAA journal*, 45(7):1584–1595, 2007.

[45] Mohammed Nabil El Korso, Rémy Boyer, Alexandre Renaux, and Sylvie Marcos. Conditional and unconditional cramér–rao bounds for near-field source localization. *Signal Processing, IEEE Transactions on*, 58(5):2901–2907, 2010.

[46] Mohammed Nabil El Korso, Rémy Boyer, Alexandre Renaux, and Sylvie Marcos. Statistical resolution limit of the uniform linear co-centered orthogonal loop and dipole array. *Signal Processing, IEEE Transactions on*, 59(1):425–431, 2011.

[47] Randima Fernando. *GPU gems: programming techniques, tips, and tricks for real-time graphics*. 2004.

[48] J.J. Fuchs. Multipath time-delay detection and estimation. *IEEE Transactions on Signal Processing*, 47(1):237–243, 1999.

[49] A.E. Gelfand, A.F.M. Smith, and T.M. Lee. Bayesian analysis of constrained parameter and truncated data problems using Gibbs sampling. *Journal of the American Statistical Association*, 87(418):523–532, 1992.

[50] Anthony Gérard, Alain Berry, and Patrice Masson. Control of tonal noise from subsonic axial fan. part 1: reconstruction of aeroacoustic sources from far-field sound pressure. *Journal of sound and vibration*, 288(4):1049–1075, 2005.

[51] Jerry D Gibson, Boneung Koo, and Steven D Gray. Filtering of colored noise for speech enhancement and coding. *Signal Processing, IEEE Transactions on*, 39(8):1732–1742, 1991.

[52] Gene H Golub, Michael Heath, and Grace Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21(2):215–223, 1979.

[53] Neil J Gordon, David J Salmond, and Adrian FM Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. In *IEE Proceedings F (Radar and Signal Processing)*, volume 140, pages 107–113. IET, 1993.

[54] Robert M Gray. *Toeplitz and circulant matrices: II*. Information Systems Laboratory, Stanford Electronics Laboratories, Stanford University, 1977.

[55] Banu Gunel, H Hachabiboglu, and Ahmet M Kondoz. Acoustic source separation of convolutive mixtures based on intensity vector statistics. *Audio, Speech, and Language Processing, IEEE Transactions on*, 16(4):748–756, 2008.

[56] Per Christian Hansen. Deconvolution and regularization with Toeplitz matrices. *Numerical Algorithms*, 29(4):323–378, 2002.

[57] Jarvis Haupt, Waheed U Bajwa, Gil Raz, and Robert Nowak. Toeplitz compressed sensing matrices with applications to sparse channel estimation. *Information Theory, IEEE Transactions on*, 56(11):5862–5875, 2010.

[58] Lihan He, Haojun Chen, and Lawrence Carin. Tree-structured compressive sensing with variational bayesian analysis. *Signal Processing Letters, IEEE*, 17(3):233–236, 2010.

[59] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 2012.

[60] Osamu Hoshuyama, Akihiko Sugiyama, and Akihiro Hirano. A robust adaptive beamformer for microphone arrays with a blocking matrix using constrained adaptive filters. *IEEE Transactions on Signal Processing*, 47(10):2677–2684, 1999.

[61] Junzhou Huang and Tong Zhang. The benefit of group sparsity. *The Annals of Statistics*, 38(4):1978–2004, 2010.

[62] Yiteng Huang, Jacob Benesty, and Jingdong Chen. *Acoustic MIMO signal processing*, volume 1. Springer Berlin, 2006.

[63] Harold Jeffreys. An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 186(1007):453–461, 1946.

[64] Rodolphe Jenatton, Jean-Yves Audibert, and Francis Bach. Structured variable selection with sparsity-inducing norms. *The Journal of Machine Learning Research*, 12:2777–2824, 2011.

[65] Douglas Samuel Jones. Acoustic and electromagnetic waves. *Oxford/New York, Clarendon Press/Oxford University Press, 1986, 764 p.*, 1, 1986.

[66] M.I. Jordan, Z. Ghahramani, T.S. Jaakkola, and L.K. Saul. An introduction to variational methods for graphical models. *Machine learning*, 37(2):183–233, 1999.

[67] Mark Kac. Toeplitz matrices, translation kernels and a related problem in probability theory. *Duke Mathematical Journal*, 21(3):501–509, 1954.

[68] Carine El Kassis, José Picheral, and Chafic Mokbel. Advantages of nonuniform arrays using root-music. *Signal Processing*, 90(2):689–695, 2010.

[69] AK Katsaggelos and KT Lay. Maximum likelihood blur identification and image restoration using the EM algorithm. *IEEE Transactions on Signal Processing*, 39(3):729–733, 1991.

[70] Y. Kim and PA Nelson. Optimal regularisation for acoustic source reconstruction by inverse methods. *Journal of sound and vibration*, 275(3-5):463–487, 2004.

[71] K.C. Kiwiel. Convergence and efficiency of subgradient methods for quasiconvex minimization. *Mathematical programming*, 90(1):1–25, 2001.

[72] Kenneth J Koehler and Kinley Larntz. An empirical investigation of goodness-of-fit statistics for sparse multinomials. *Journal of the American Statistical Association*, 75(370):336–344, 1980.

[73] Martin Kompis and Norbert Dillier. Simulating transfer functions in a reverberant room including source directivity and head-shadow effects. *JOURNAL-ACOUSTICAL SOCIETY OF AMERICA*, 93:2779–2779, 1993.

[74] J. Lanslots, F. Deblauwe, and K. Janssens. Selecting Sound Source Localization Techniques for Industrial Applications. *Sound and Vibration*, 44(6):6–10, 2010.

[75] Katia Lebart, Jean-Marc Boucher, and PN Denbigh. A new method based on spectral subtraction for speech dereverberation. *Acta Acustica united with Acustica*, 87(3):359–366, 2001.

[76] Jooshik Lee, Iickho Song, Hyoungmoon Kwon, and Sung Ro Lee. Low-complexity estimation of 2D DOA for coherently distributed sources. *Signal processing*, 83(8):1789–1802, 2003.

[77] S.K. Lehman and A.J. Devaney. Transmission mode time-reversal super-resolution imaging. *The Journal of the Acoustical Society of America*, 113(5):2742–2753, 2003.

[78] Hanoch Lev-Ari and Yariv Ephraim. Extension of the signal subspace speech enhancement approach to colored noise. *Signal Processing Letters, IEEE*, 10(4):104–106, 2003.

[79] Jian Li, Petre Stoica, and Zhisong Wang. Doubly constrained robust capon beamformer. *IEEE Transactions on Signal Processing*, 52(9):2407–2423, 2004.

[80] Craig S MacInnes. Source localization using subspace estimation and spatial filtering. *IEEE Journal of Oceanic Engineering*, 29(2):488–497, 2004.

[81] M.B.S. Magalhaes and R.A. Tenenbaum. Sound sources reconstruction techniques: A review of their evolution and new trends. *Acta Acustica united with Acustica*, 90(2):199–220, 2004.

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

[83] Claude Marro, Yannick Mahieux, and Klaus Uwe Simmer. Analysis of noise reduction and dereverberation techniques based on microphone arrays with postfiltering. *Speech and Audio Processing, IEEE Transactions on*, 6(3):240–259, 1998.

[84] A. Massa and G. Oliveri. Bayesian compressive sampling for pattern synthesis with maximally sparse non-uniform linear arrays. *IEEE Transactions on Antennas and Propagation*, 59(10):467–681, 2011.

[85] J. D. Maynard, E. G. Williams, and Y. Lee. Nearfield acoustic holography: I. Theory of generalized holography and the development of NAH. 78(4):1395–1413, 1985.

[86] Arnaud Menoret, Nathalie Gorilliot, and Jean-Luc Adam. Acoustic imaging in wind tunnel S2A. In *10th Acoustics conference (ACOUSTICS2010)*, Lyon, France, 2010.

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

[88] A. B. Nagy. Aeroacoustics research in Europe: The CEAS-ASC report on 2010 highlights. *Journal of Sound and Vibration*, 330(21):4955–4980, 2011.

[89] James G Nagy and Dianne P O'leary. Fast iterative image restoration with a spatially varying psf. In *Optical Science, Engineering and Instrumentation'97*, pages 388–399. International Society for Optics and Photonics, 1997.

[90] John Nickolls, Ian Buck, Michael Garland, and Kevin Skadron. Scalable parallel programming with cuda. *Queue*, 6(2):40–53, 2008.

[91] Alan V Oppenheim, Ronald W Schafer, John R Buck, et al. *Discrete-time signal processing*, volume 5. Prentice hall Upper Saddle River, 1999.

[92] Dimitri Papamoschou. Imaging of directional distributed noise sources. *Journal of Sound and Vibration*, 330(10):2265–2280, 2011.

[93] Pietro Perona. Deformable kernels for early vision. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 17(5):488–499, 1995.

[94] Bill M Radich and Kevin M Buckley. Single-snapshot DOA estimation and source number detection. *IEEE Signal Processing Letters*, 4(4):109–111, 1997.

[95] F. Ronquist and J.P. Huelsenbeck. Mr Bayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19(12):1572–1574, 2003.

[96] E. Sarradj. A fast signal subspace approach for the determination of absolute levels from phased microphone array measurements. *Journal of Sound and Vibration*, 329(9):1553–1569, 2010.

[97] Henrik Schmidt and Finn B Jensen. A full wave solution for propagation in multilayered viscoelastic media with application to gaussian beam reflection at fluid-solid interfaces. *The Journal of the Acoustical Society of America*, 77(3):813–825, 1985.

[98] Ralph Schmidt. Multiple emitter location and signal parameter estimation. *Antennas and Propagation, IEEE Transactions on*, 34(3):276–280, 1986.

[99] Shahram Shahbazpanahi, Shahrokh Valaee, and Alex B Gershman. A covariance fitting approach to parametric localization of multiple incoherently distributed sources. *IEEE Transactions on Signal Processing*, 52(3):592–600, 2004.

[100] Le-Jeng Shiue, Ian Jones, and Jörg Peters. A realtime GPU subdivision kernel. In *ACM Transactions on Graphics (TOG)*, volume 24, pages 1010–1015. ACM, 2005.

[101] Dmitriy Shutin, Thomas Buchgraber, Sanjeev R Kulkarni, and H Vincent Poor. Fast variational sparse bayesian learning with automatic relevance determination for superimposed signals. *Signal Processing, IEEE Transactions on*, 59(12):6257–6261, 2011.

[102] P. Sijtsma. CLEAN based on spatial source coherence. *International Journal of Aeroacoustics*, 6(4):357–374, 2007.

[103] Pieter Sijtsma. CLEAN based on spatial source coherence. *International Journal of Aeroacoustics*, 6(4):357–374, 2008.

[104] Patrice Simard and Jerome Antoni. Acoustic source identification: Experimenting the l1 minimization approach. *Applied Acoustics*, 74(7):974 – 986, 2013.

[105] O. Siohan, C. Chesta, and C.H. Lee. Joint Maximum A Posteriori adaptation of transformation and HMM parameters. *IEEE Transactions on Speech and Audio Processing*, 9(4):417–428, 2001.

[106] V. Smídl and A. Quinn. *The Variational Bayes Method in Signal Processing (Signals and Communication Technology)*. Springer-Verlag New York, Inc., Secaucus, NJ, 2005.

[107] Kihyuk Sohn, Dae Yon Jung, Honglak Lee, and Alfred O Hero. Efficient learning of sparse, distributed, convolutional feature representations for

object recognition. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 2643–2650. IEEE, 2011.

[108] C. Soussen, J. Idier, D. Brie, and J. Duan. From Bernoulli Gaussian deconvolution to sparse signal restoration. *IEEE Transactions on Signal Processing*, 59(10):4572–4584, 2011.

[109] J.F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization methods and software*, 11(1-4):625–653, 1999.

[110] Qiyu Sun and Michael Unser. Left-inverses of fractional laplacian and sparse stochastic processes. *Advances in Computational Mathematics*, 36(3):399–441, 2012.

[111] A. Tarantola. *Inverse problem theory and methods for model parameter estimation*. Society for Industrial Mathematics, 2005.

[112] R. Tibshirani. Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288, 1996.

[113] Michael Ting, Raviv Raich, and Alfred O Hero. Sparse image reconstruction for molecular imaging. *Image Processing, IEEE Transactions on*, 18(6):1215–1227, 2009.

[114] D.G. Tzikas, A.C. Likas, and N.P. Galatsanos. The variational approximation for Bayesian inference. *Signal Processing Magazine, IEEE*, 25(6):131–146, 2008.

[115] D.G. Tzikas, A.C. Likas, and N.P. Galatsanos. Variational Bayesian sparse kernel-based blind image deconvolution with Student's-t priors. *IEEE Transactions on Image Processing*, 18(4):753–764, 2009.

[116] Michael Unser, Philippe Thevenaz, and Leonid Yaroslavsky. Convolution-based interpolation for fast, high-quality rotation of images. *Image Processing, IEEE Transactions on*, 4(10):1371–1381, 1995.

[117] W Van Drongelen, M Yuchtman, BD Van Veen, and AC Van Huffelen. A spatial filtering technique to detect and localize multiple sources in the brain. *Brain Topography*, 9(1):39–49, 1996.

[118] B.D. Van Veen and K.M. Buckley. Beamforming: A versatile approach to spatial filtering. *ASSP Magazine, IEEE*, 5(2):4–24, 1988.

[119] J-P Vila, Verene Wagner, and Pascal Neveu. Bayesian nonlinear model selection and neural networks: a conjugate prior approach. *Neural Networks, IEEE Transactions on*, 11(2):265–278, 2000.

[120] Tuomas Virtanen. Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria. *Audio, Speech, and Language Processing, IEEE Transactions on*, 15(3):1066–1074, 2007.

[121] HMM. V. Wal and P. Sijtsma. Source localization techniques with acoustic arrays at NLR. In *NAG/DAGA 2009 International Conference on Acoustics*. NAG/DAGA, 2009.

[122] DeLiang Wang, Guy J Brown, et al. *Computational auditory scene analysis: Principles, algorithms, and applications*, volume 147. Wiley interscience, 2006.

[123] Yanwei Wang, Jian Li, Petre Stoica, Mark Sheplak, and Toshikazu Nishida. Wideband RELAX and wideband CLEAN for aeroacoustic imaging. *Journal of Acoustical Society of America*, 115(2):757–767, 2004.

[124] Kazuho Watanabe and Sumio Watanabe. Stochastic complexities of gaussian mixtures in variational bayesian approximation. *The Journal of Machine Learning Research*, 7:625–644, 2006.

[125] Earl G Williams. *Fourier acoustics: sound radiation and nearfield acoustical holography*. Academic Press, 1999.

[126] 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*, 123(5):2631–42, 2008.

[127] T. Yardibi, J. Li, P. Stoica, M. Xue, and A.B. Baggeroer. Source localization and sensing: A nonparametric iterative adaptive approach based on weighted least squares. *IEEE Transactions on Aerospace and Electronic Systems*, 46(1):425–443, 2010.

[128] Tarik Yardibi, Jian Li, Petre Stoica, Nikolas S. Zawodny, and Louis N. Cattafesta. A covariance fitting approach for correlated acoustic source mapping. *Journal of The Acoustical Society of America*, 127(5):2920–31, 2010.

[129] Takuya Yoshioka, Armin Sehr, Marc Delcroix, Keisuke Kinoshita, Roland Maas, Tomohiro Nakatani, and Walter Kellermann. Making machines understand us in reverberant rooms: robustness against reverberation for automatic speech recognition. *Signal Processing Magazine, IEEE*, 29(6):114–126, 2012.

[130] Erliang Zhang, Jérôme Antoni, Bin Dong, and Hichem Snoussi. Bayesian space-frequency separation of wide-band sound sources by a hierarchical approach. *The Journal of the Acoustical Society of America*, 132(5):3240–3250, 2012.