Studies on Sparse Time-Frequency Representations

スパース時間周波数表現に関する研究

October, 2022

Tsubasa KUSANO
草野　翼

Studies on Sparse Time-Frequency Representations

スパース時間周波数表現に関する研究

October, 2022

Waseda University Graduate School of Fundamental Science and Engineering

Department of Intermedia Studies, Research on Communication Acoustics

Tsubasa KUSANO
草野　翼

Contents

List of Figures

List of Tables

Acknowledgements

I would first like to express my sincere gratitude to Prof. Yasuhiro Oikawa, the supervisor of my doctoral study. My research could not have pursued without his continuous and dedicated guidance and support.

Besides my advisor, I would like to thank the referees of my thesis, Prof. Tetsuya Ogata and Prof. Yukio Gunji, for their insightful advices and encouragement. I would also like to express my thanks to Associate Prof. Kohei Yatabe at Tokyo University of Agriculture and Technology for his very precise advice and discussions about my research.

Finally, I would like to express my deepest gratitude to everyone involved with me. I would like to express my profound gratitude to my family for their generous support and continuous encouragement throughout my Ph.D. study and as I was conducting my research and writing this thesis.

Nomenclature

Notations

$\mathbb{R}$	Set of all real numbers
$\mathbb{C}$	Set of all complex numbers
$\mathbb{R}_+$	Set of all non-negative real numbers
$L^2(\mathbb{R})$	Space of square integrable functions
$C^\rho(\mathbb{R})$	Set of all ρ times continuously differentiable functions
$\mathbb{S}_+^N$	Set of all symmetric positive semidefinite $N \times N$ matrices
i	Imaginary unit $\mathrm{i}^2 = -1$
$f(x)$	Value of a function f at x
$\mathbf{x}[n]$	nth element of a vector $\mathbf{x}$
$\mathbf{X}[n, m]$	(n, m)th element of a matrix $\mathbf{X}$
$\overline{(\cdot)}$	Complex conjugate
$(\cdot)^{\mathrm{T}}$	Transpose
$(\cdot)^{\mathrm{H}}$	Conjugate transpose
$(\cdot)^*$	Adjoint
$\langle \mathbf{x}, \mathbf{y} \rangle$	Inner product of $\mathbf{x}, \mathbf{y} \in \mathbb{C}^L$; $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* \mathbf{x}$
$\langle \mathbf{X}, \mathbf{Y} \rangle_{\mathrm{F}}$	Frobenius inner product of $\mathbf{X}, \mathbf{Y}$; $\langle \mathbf{x}, \mathbf{y} \rangle = \mathrm{Tr}(\mathbf{Y}^* \mathbf{X})$
$\lVert \cdot \rVert$	Euclidean norm
$\lVert \cdot \rVert_1$	ℓ_1 norm
$\lVert \cdot \rVert_2$	ℓ_2 norm
$\lVert \cdot \rVert_*$	Trace norm
$\lVert \cdot \rVert_{\mathrm{F}}$	Frobenius norm
$\mathbf{1}_N$	N-dimensional vector whose elements are all one
$\mathbf{0}_N$	N-dimensional vector whose elements are all zero
$\mathbf{I}_N$	$N \times N$ identity matrix

$\mathbf{O}_{N,M}$	$N \times M$ zero matrix
$\Re\{z\}$	Real part of z
$\Im\{z\}$	Imaginary part of z
$\mathrm{Arg}\{z\}$	(Principal value of) complex argument of z
$\mathrm{diag}(\mathbf{x})$	Diagonal matrix whose diagonal elements are $\mathbf{x}$
$\lfloor x \rceil$	Round function that rounds x to the nearest integer
$\lfloor x \rfloor$	Floor function
$\mathrm{sign}(x)$	Sign function; $\mathrm{sign}(x) = \begin{cases} \frac{x}{\lvert x \rvert} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$
$\mathrm{rank}(\mathbf{X})$	Rank of a matrix $\mathbf{X}$
$\mathrm{Tr}(\mathbf{X})$	Trace of a matrix $\mathbf{X}$
$\mathbf{X} \succeq 0$	$\mathbf{X}$ being positive semidefinite

Abbreviations

T-F	Time-frequency
STFT	Short-time Fourier transform
WVD	Wigner–Ville distribution
EMD	Empirical mode decomposition
DGT	Discrete Gabor transform
SST	Synchrosqueezing transform
FSST	Fourier synchrosqueezing transform
IF	Instantaneous frequency
DFT	Discrete Fourier transform
SE	Sidelobe energy
HSL	Highest sidelobe level
ADMM	Alternating direction method of multipliers
SDP	Semidefinite programming
ANM	Atomic norm minimization
RAM	Reweighted atomic norm minimization

Chapter 1

Introduction

Signal analysis and processing is an important element in obtaining information from observed data in various applications. Non-stationary signals including audio and speech signals are generally analyzed or processed in the time-frequency (T-F) domain because it is insufficient to analyze them in only the time or frequency domains. The short-time Fourier tranform (STFT) [1] is usually utilized to convert signals into the T-F domain, owing to its simplicity and easily understandable structure [2–4]. STFT is a method of obtaining a T-F representation by the Fourier transform of the signal multiplied by a time-shifted window function. The resolution of the T-F representation obtained by STFT depends on the window function and is limited in Heisenberg's uncertainty principle. For more detailed analysis and more accurate processing, a time-frequency representation with higher resolution than the Heisenberg uncertainty principle is required.

The reassignment method and the synchrosqueezing transform (SST) have been proposed to overcome Heisenberg's uncertainty principle [5]. Kodera first proposed the reassignment method to improve the readability of the T-F representation obtained by the STFT [6]. Then, Auger and Flandrin popularized the reassignment method by discovering an efficient computational method; furthermore, they generalized the reassignment method to T-F representations in Cohen's class and time-scale representations [7]. Then, the reassignment method was generalized for any filterbank [8,9]. It has been applied to acoustics, biomedics, and machine condition monitoring [10–12]. The reassignment method sharpens a T-F representation using the time and frequency derivative of its phase at the expense of invertibility.

In the context of audio signal analysis, Daubechies and Maes proposed the

SST [13, 14], which is a variant of the reassignment method. The SST performs frequency-only reassignment to a complex-valued T-F representation to sharpen the T-F representation while ensuring invertibility. Subsequently, the SST was also generalized for STFT [15, 16] and other representations [9, 17] and has been widely studied [18–22]. SST has been applied to various fields such as seismic waves and gravity waves [17, 21]. The main computations in the reassignment method and the SST are the STFTs with a window and with its derivatives. In addition, they can also be performed online using a finite-length window function in the time domain. On the other hand, it is not easy to reassign the components in their original positions when they are mixed by the influence of windowing.

In a different vein, sparsity-aware methods for estimating T-F representation have been proposed [23–32]. Sparsity-aware methods aim to find the coefficients of a predefined redundant dictionary such that the signal is represented with as few coefficients as possible. The direct formulation for finding the sparsest representation of representing the signal is to minimize the number of non-zero coefficients, called the ℓ_0 norm. Unfortunately, this problem is usually an intractable combinatorial optimization problem, and its solution is sensitive to noise. Mallat et al. have proposed the method of matching pursuit that finds an approximate solution to this problem with a greedy algorithm [23]. Chen et al. propose the method of basis pursuit that solves the minimization problem using the ℓ_1 norm, which is a convex relaxation of the ℓ_0 norm [24]. This formulation is a convex optimization problem, which can be solved by convex optimization algorithms efficiently. Furthermore, It has advantages in obtaining globally optimal solutions and robustness to noises and mixing components. However, it is computationally intensive due to the use of iterative algorithms. It is also difficult to apply to online processing because it solves the optimization problem using the entire signal. Thus, there is a trade-off between robustness and computational complexity for the reassignment method, the synchrosqueezing transformation, and the iterative algorithm, and the choice must be made according to the application. Thus, the reassignment method, SST, and the sparsity-aware methods should be selected according to the application.

The aim of this thesis is to explore and improve the STFT and the estimation methods of sparse T-F representations based on STFT. The organization of this thesis is illustrated in Fig. 1.1. Chapter 2 provides the general knowledge related to the STFT and Heisenberg's uncertainty principle. Next, the reassignment method,

SST, and sparsity-aware method are described as techniques for estimating sparse T-F represenations to overcome Heisenberg's uncertainty principle.

In Chapter 3, we propose a method for designing window functions for computing instantaneous frequency (IF), which used in the reassignment method and the SST. The IF is computed using the STFTs with a window and with its derivatives. These frequency responses affect computing the IF. The proposed method designs window functions to minimize the sidelobes of derivatives of the windows. This work was presented in [33, 34].

In Chapter 4, we propose an estimation of sparse T-F representation using the atomic norm. The formulation using the ℓ_1 norm requires the discretization of the frequency parameter onto a grid. If the grid and the actual frequencies do not match, a sparse solution may not be obtained. To avoid the grid mismatch, we propose a method for estimating sparse T-F representation using the atomic norm.

In Chapter 5, we propose a method to improve sparsity by combining the gridless sparse method using the atomic norm with nonconvex optimization. The estimator via the atomic norm minimization includes a bias to reduce the amplitude as well as the ℓ_1 norm. In the grid-based sparse optimization, nonconvex penalties induceing sparsiy have been proposed instead of the ℓ_1 norm. The proposed method combines of a gridless sparse optimization method using the atomic norm and a sparse optimization method using non-convex functions. Chapter 6 summarizes this thesis and discusses the remaining issues.

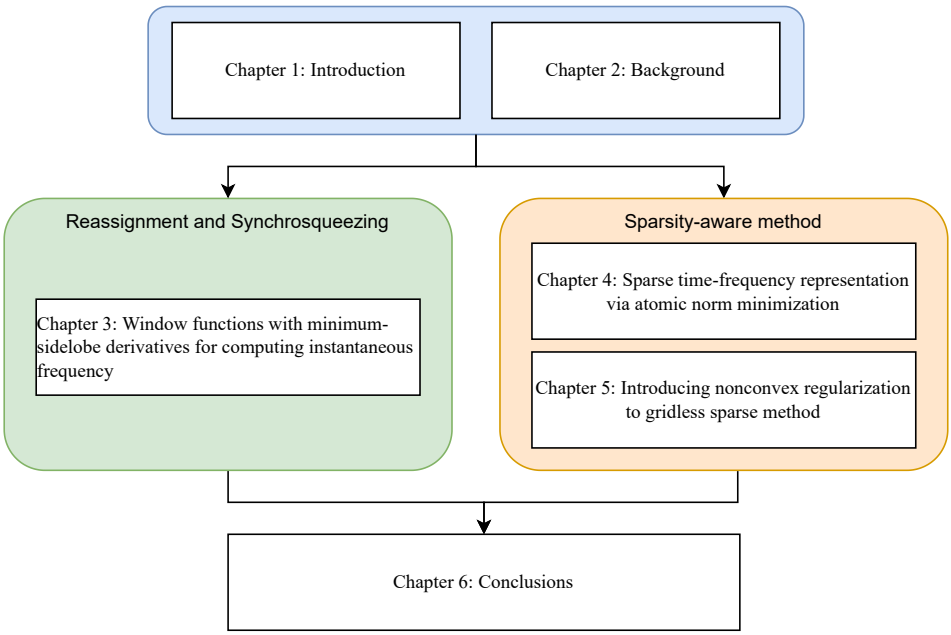

Figure 1.1: Organization of this thesis.

Chapter 2

Background

This chapter explains the STFT and the methods of estimating a sparse T-F representation. First, the classical T-F representations including the STFT, WVD, and Cohen's class are described. Next, we introduce reassignment method and sparse optimization-based methods as methods for estimating sparse T-F representations.

2.1 Classical time-frequency representation

2.1.1 Fourier transform

Before describing the STFT, we define some terms related to the Fourier transform. The Fourier transform maps a time-domain signal to the frequency domain. The Fourier transform $\hat{f} \in L^2(\mathbb{R})$ of a signal $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ is defined as

$$\hat{f}(\omega) \coloneqq \int_{\mathbb{R}} f(t) \mathrm{e}^{-\mathrm{i}2\pi\omega t} \mathrm{d}t. \tag{2.1}$$

The signal f can be recovered from $\hat{f}$ via the inverse Fourier transform:

$$f = \int_{\mathbb{R}} \hat{f}(\omega) \mathrm{e}^{\mathrm{i}2\pi\omega} \mathrm{d}\omega. \tag{2.2}$$

For a discrete signal $\mathbf{f} \in \mathbb{C}^L$, the discrete Fourier transform (DFT) $\mathbf{F} : \mathbb{C}^L \to \mathbb{C}^L$ is defined as

$$\hat{\mathbf{f}}[m] \coloneqq (\mathbf{F}\mathbf{f})[m] = \sum_{l=0}^{L-1} \mathbf{f}[l] \mathrm{e}^{-\mathrm{i}\frac{2\pi m l}{L}}, \tag{2.3}$$

and its inverse $\mathbf{F}^{-1} : \mathbb{C}^L \to \mathbb{C}^L$ is given by

$$\mathbf{f}[l] = (\mathbf{F}^{-1}\hat{\mathbf{f}})[l] = \frac{1}{L} \sum_{m=0}^{L-1} \hat{\mathbf{f}}[m] \mathrm{e}^{\mathrm{i}\frac{2\pi m l}{L}}. \tag{2.4}$$

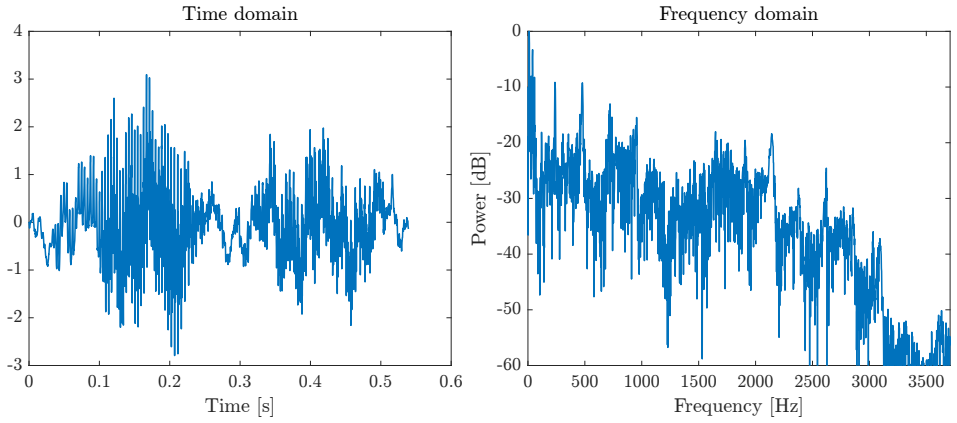

Figure 2.1: Time and frequency representations of a speech signal.

Then, we define the (continuous) frequency response of a discrete signal $\mathbf{f} \in \mathbb{C}^L$ as

$$(\mathcal{F}\mathbf{f})(\omega) \coloneqq \sum_{l=0}^{L-1} \mathbf{f}[l] \mathrm{e}^{-\mathrm{i}2\pi\omega l}. \tag{2.5}$$

2.1.2 Short-time Fourier transform (STFT)

When analyzing non-stationary signals such as audio signals, the time-domain and the frequency-domain representations are not sufficient. Fig. 2.1 plots time and frequency representations of a speech signal. These two representations show some information on the time interval of high energy and the main frequency components, but they do not tell us which frequency components are included at which time. STFT is a tool to obtain time and frequency information simultaneously [1–4].

The basic idea of STFT is to apply the Fourier transform to a localized signal obtained by multiplying a window function. The STFT of a signal $f \in L^2(\mathbb{R})$ with a window function $g \in L^2(\mathbb{R}) \cap C^1(\mathbb{R})$ is defined as

$$(\mathcal{V}_g f)(t, \omega) \coloneqq \int_{\mathbb{R}} f(\tau) \overline{g(\tau - t)} \mathrm{e}^{-\mathrm{i}2\pi\omega\tau} \mathrm{d}\tau, \tag{2.6}$$

where $t \in \mathbb{R}$ and $\omega \in \mathbb{R}$ represent time and frequency, respectively. The signal f can be recovered from its STFT via the inverse STFT with a window function h satisfying $\langle g, h \rangle \neq 0$:

$$f = \frac{1}{\langle g, h \rangle} \int_{\mathbb{R}^2} (\mathcal{V}_g f)(\tau, \omega) h(t - \tau) \mathrm{e}^{\mathrm{i}2\pi\omega\tau} \mathrm{d}\tau \mathrm{d}\omega. \tag{2.7}$$

Fig. 2.2 displays the T-F representations of the speech signal obtained by STFT with three Blackman windows of different lengths. These T-F representations adequately represent time-varying frequencies. The length of the window changes the

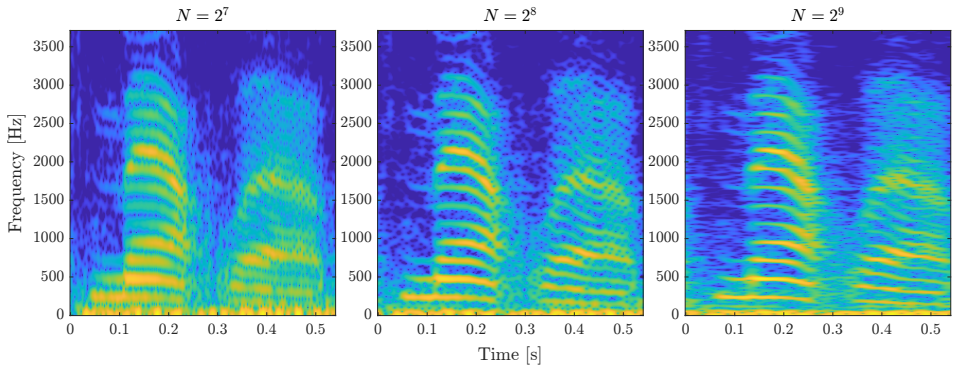

Figure 2.2: T-F representations of the speech signal obtained by STFT with the Blackman window. Each column shows (from left to right) the T-F representations using the Blackman windows of length $N = 2^7, 2^8, 2^9$.

localization of the obtained T-F representation. Specifically, shortening the window length increases the frequency spread instead of decreasing the time spread. By contrast, increasing the window length decreases the frequency spread instead of increasing the time spread. This inability to simultaneously decrease the time and frequency spread is referred to as the uncertainty principle. Gabor introduced Heisenberg's uncertainty inequality [1]:

$$\left(\int_{\mathbb{R}} (t-t_0)^2 \left|f(t)\right|^2 \mathrm{d}t\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}} (\omega-\omega_0)^2 \left|\hat{f}(\omega)\right|^2 \mathrm{d}\omega\right)^{\frac{1}{2}} \geq \frac{1}{4\pi} \int_{\mathbb{R}} \left|f(t)\right|^2 \mathrm{d}t, \tag{2.8}$$

for arbitrary points (t_0, ω_0). The equality holds if and only if f is a multiple of $\mathrm{e}^{\frac{-(t-t_0)^2}{\sigma^2}} \mathrm{e}^{\mathrm{i}2\pi\omega_0 t}$ $(\sigma > 0)$.

Since the property of the STFT is characterized by the window function, the choice of the window function affects the performance of the signal analysis and synthesis. To adjust the resolution of the STFT, many window functions have been proposed from various viewpoints, such as frequency responses [35–44] and numerical stability in signal processing [45–50]. For example, the Hann and Nuttall windows are popular windows designed to achieve a better sidelobe decay. The Kaiser window was proposed so that its frequency response was adjustable by a tuning parameter [38]. Furthermore, a method for determining the optimal window width using the Rényi entropy is proposed [51].

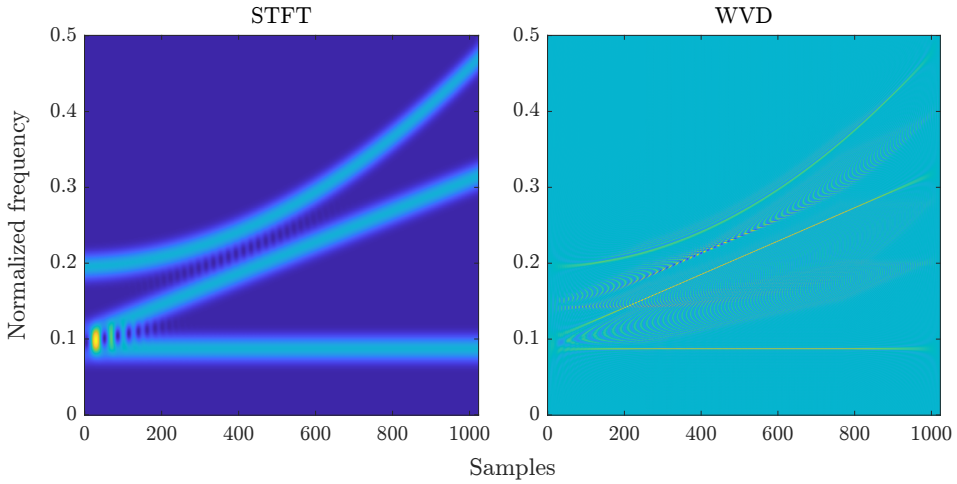

Figure 2.3: Comparison of STFT and WVD of the sum of a sinusoid, a linear chirp, and a quadratic chirp.

2.1.3 Wigner–Ville distribution (WVD) and Cohen's class

WVD was derived by Wigner in quantum mechanics [52], and then introduced in signal analysis by Ville [53]. The WVD of a signal $f(t)$ is defined by

$$\mathscr{W}_f(t,\omega) \coloneqq \int_{\mathbb{R}} f\left(t+\frac{\tau}{2}\right)\overline{f\left(t-\frac{\tau}{2}\right)}\mathrm{e}^{-\mathrm{i}2\pi\omega\tau}\mathrm{d}\tau. \tag{2.9}$$

The WVD depends only on the signal and achieves the perfect localization in the case of linear chirp signals. However, in the presense of multiple components, the WVD have the *cross terms* between the components.

Fig. 2.3 shows the T-F representations obtained by STFT and WVD. Each component in the WVD is more localized than that in the STFT. On the other hand, the WVD have cross terms not seen in the STFT. The cross terms in WVD make it difficult to analyze the more complex signals.

The way to reduce the interference terms is convolving the WVD with a smoothing kernel Π,

$$\mathcal{C}_f^{\Pi}(t,\omega) \coloneqq \int_{\mathbb{R}^2} \mathscr{W}_f(\tau,\nu)\Pi(\tau-t,\nu-\omega)\mathrm{d}\tau\mathrm{d}\nu, \tag{2.10}$$

which is referred to as Cohen's class [54]. The spectrogram $|(\mathcal{V}_g f)(t,\omega)|^2$ and the WVD $\mathscr{W}_f(t,\omega)$ correspond to $\mathcal{C}_f^{\Pi}(t,\omega)$ whose kernels are the WVD of the window $\mathscr{W}_g(t,\omega)$ and $\delta(t)\delta(\omega)$, respectively. There is a tradeoff between the concentration of the T-F representation and the attenuation of the cross terms.

2.1.4 Other representations

The classical STFT analyzes a signal with a single T-F resolution over the entire T-F domain. Some signals may are desirable to analyze at different resolutions. As the methods analysing with different resolutions in the frequency direction, wavelet transform [55], S-transform (Stockwell transform) [56], constant-Q transform [57], and the method using any filterbank [58] have been introduced. The method analysing in different resolutions over the T-F domain is referred to as the non-stationary Gabor transform and has been studied [59, 60].

As a different approach for non-stationary signal analysis from these representations, empirical mode decomposition (EMD) has been proposed [61]. EMD is a data-driven technique which decomposes a nonstationary signal into intrinsic mode functions (IMFs). To compute the time-frequency representation, the instantaneous frequency of each IMF is calculated using the Hilbert transform.

2.2 Reassignment and synchrosqueezing

The T-F representation obtained by the STFT is spread by the window function. The reassignment method aims to assign the energy spread to the original position using the information of phase derivatives [6].

The original position can be estimated from the local centroid of the spectrogram, which is written as

$$\widehat{t} = \frac{\int_{\mathbb{R}^2} \tau \mathscr{W}_f(\tau,\nu)\mathscr{W}_g(\tau-t,\nu-\omega)\mathrm{d}\tau\mathrm{d}\nu}{\int_{\mathbb{R}^2} \mathscr{W}_f(\tau,\nu)\mathscr{W}_g(\tau-t,\nu-\omega)\mathrm{d}\tau\mathrm{d}\nu}, \tag{2.11}$$

$$\widehat{\omega} = \frac{\int_{\mathbb{R}^2} \nu \mathscr{W}_f(\tau,\nu)\mathscr{W}_g(\tau-t,\nu-\omega)\mathrm{d}\tau\mathrm{d}\nu}{\int_{\mathbb{R}^2} \mathscr{W}_f(\tau,\nu)\mathscr{W}_g(\tau-t,\nu-\omega)\mathrm{d}\tau\mathrm{d}\nu}. \tag{2.12}$$

Then, the reassigned spectrogram is represented using $(\widehat{t},\widehat{\omega})$ as

$$\mathcal{T}_g^f(t,\omega) = \int_{\mathbb{R}^2} |(\mathcal{V}_g f)(\tau,\nu)|^2 \, \delta(t-\widehat{t}(\tau,\nu))\delta(\omega-\widehat{\omega}(\tau,\nu))\mathrm{d}\tau\mathrm{d}\nu, \tag{2.13}$$

where $\delta(t)$ denotes the Dirac delta distribution. The reassignment method is generalized to T-F representations in Cohen's class, time-scale representations [7], and any filterbank [8, 9].

Although the reassignment method provides the sparse T-F representation, the representation cannot be converted onto the time-domain signal. The synchrosqueezing transform (SST) estimates an invertible and sparse T-F representa-

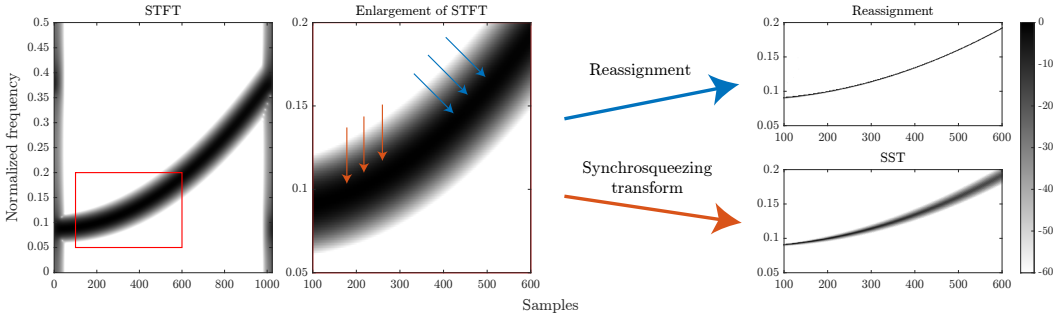

Figure 2.4: Diagram of reassignment method and SST.

tion by reassigning the complex-valued T-F representation in the frequency direction only [14, 15]. Using (2.12), the SST written as

$$\mathcal{S}_g^f(t,\omega) = \int_{\mathbb{R}} (\mathcal{V}_g f)(t,\omega)\delta(\omega - \hat{\omega}(t,\nu))\mathrm{d}\nu. \tag{2.14}$$

The diagram of the reassignment method and SST is shown in Fig. 2.4. Though the SST is less localized than the reassginment method since the SST reassigns only in the frequency direction, the SST retains invertibility. High-order SST [20–22] and time-reassigned synchrosqueezing [62] have been proposed for obtaining better-localized representation.

2.2.1 Relation between reassignment operators and phase derivatives

The relation between the centroid $(\widehat{t}, \widehat{\omega})$ and the derivative of the phase of STFT is shown in [6]. Let $\mathcal{M}_g^f := |\mathcal{V}_g f|$ and Φ_g^f be the STFT magnitude and phase, respectively. The STFT can be represented as

$$(\mathcal{V}_g f)(t,\omega) = \mathcal{M}_g^f(t,\omega) \cdot \mathrm{e}^{\mathrm{i}2\pi\Phi_g^f(t,\omega)}. \tag{2.15}$$

Then, (2.11) and (2.12) are also represented using the gradients of phase as

$$\widehat{t}(t,\omega) = -\frac{\partial}{\partial\omega}\Phi_g^f(t,\omega), \tag{2.16}$$

$$\widehat{\omega}(t,\omega) = \omega + \frac{\partial}{\partial t}\Phi_g^f(t,\omega). \tag{2.17}$$

The time and frequency derivatives of the STFT phase are referred to as the instantaneous frequency (IF) and group delay (GD), respectively, they are denoted

by [63, 64]

$$\mathrm{GD}_g^f(t,\omega) \coloneqq -t - \frac{\partial}{\partial\omega}\Phi_g^f(t,\omega), \tag{2.18}$$

$$\mathrm{IF}_g^f(t,\omega) \coloneqq \frac{\partial}{\partial t}\Phi_g^f(t,\omega). \tag{2.19}$$

For instance, let us consider the IF of a continuous sinusoid

$$s(t) = A_{\mathbf{s}}\mathrm{e}^{\mathrm{i}(2\pi\xi_{\mathbf{s}}t+\phi_{\mathbf{s}})}. \tag{2.20}$$

The STFT of $s(t)$ is explicitly expressed as

$$\begin{aligned}(\mathcal{V}_g s)(t,\omega) &= A_{\mathbf{s}}\mathrm{e}^{\mathrm{i}\phi_{\mathbf{s}}}\int_{\mathbb{R}}\overline{g(\tau-t)}\mathrm{e}^{-\mathrm{i}2\pi(\omega-\xi_{\mathbf{s}})\tau}\mathrm{d}\tau \\ &= A_{\mathbf{s}}\mathrm{e}^{\mathrm{i}\phi_{\mathbf{s}}-\mathrm{i}2\pi(\omega-\xi_{\mathbf{s}})t}\int_{\mathbb{R}}\overline{g(\tau)}\mathrm{e}^{-\mathrm{i}2\pi(\omega-\xi_{\mathbf{s}})\tau}\mathrm{d}\tau \\ &= \overline{\hat{g}(\xi_{\mathbf{s}}-\omega)}A_{\mathbf{s}}\mathrm{e}^{\mathrm{i}\phi_{\mathbf{s}}-\mathrm{i}2\pi(\omega-\xi_{\mathbf{s}})t},\end{aligned} \tag{2.21}$$

where $\hat{g}(\omega)$ is the Fourier transform of $g(t)$. Then, its phase is given by

$$\Phi_g^s(t,\omega) = \frac{\phi_{\mathbf{s}} + \mathrm{Arg}\left\{\overline{\hat{g}(\xi_s-\omega)}\right\}}{2\pi} - (\omega-\xi_{\mathbf{s}})t. \tag{2.22}$$

Consequently, the IF of $s(t)$ is calculated as

$$\mathrm{IF}_g^s(t,\omega) = \frac{\partial}{\partial t}\Phi_g^s(t,\omega) = \xi_{\mathbf{s}} - \omega. \tag{2.23}$$

This corresponds to the difference between the frequency of the sinusoid $\xi_{\mathbf{s}}$ and the frequency axis ω, which allows us to observe detailed frequency information from the spread T-F representation obtained by the STFT. Some studies refer to it as *the relative instantaneous frequency* [65][1].

2.2.2 Computing IF using window derivative

The straightforward approach for computing IF_g^f from (2.19) in the discrete setting is an approximation of the time derivative of the phase by finite differences. However, this suffers from the phase unwrapping problem [67]. Avoiding such problems, [7] proposed an alternative expression of the IF given by

$$\mathrm{IF}_g^f = \frac{1}{2\pi}\Im\left\{\frac{1}{\mathcal{V}_g f}\cdot\frac{\partial}{\partial t}\mathcal{V}_g f\right\} = -\frac{1}{2\pi}\Im\left\{\frac{\mathcal{V}_{g'}f}{\mathcal{V}_g f}\right\}, \tag{2.24}$$

[1]This comes from defining STFT as (2.6), and other definitions will change this expression. More details are shown in [66].

where $g' = \mathrm{d}g/\mathrm{d}t$ is the derivative of the window g. This is derived from the following calculation using the chain rule:

$$\begin{aligned}\frac{\partial}{\partial t}(\mathcal{V}_g f)(t,\omega) &= \mathrm{e}^{\mathrm{i}2\pi\Phi_g^f}\cdot\frac{\partial}{\partial t}\mathcal{M}_g^f + \mathcal{M}_g^f\mathrm{e}^{\mathrm{i}2\pi\Phi_g^f}\cdot\mathrm{i}2\pi\frac{\partial}{\partial t}\Phi_g^f \\ &= \mathrm{e}^{\mathrm{i}2\pi\Phi_g^f}\cdot\frac{\partial}{\partial t}\mathcal{M}_g^f + \mathcal{V}_g f\cdot\mathrm{i}2\pi\frac{\partial}{\partial t}\Phi_g^f. \end{aligned} \tag{2.25}$$

In addition, the derivative of the STFT with respect to time can be rewritten as

$$\begin{aligned}\frac{\partial}{\partial t}(\mathcal{V}_g f)(t,\omega) &= -\int_{\mathbb{R}} f(\tau)\overline{\frac{\mathrm{d}g}{\mathrm{d}t}(\tau-t)}\mathrm{e}^{-\mathrm{i}2\pi\omega\tau}\mathrm{d}\tau \\ &= -(\mathcal{V}_{g'}f)(t,\omega). \end{aligned} \tag{2.26}$$

According to (2.24), computing the IF requires the STFT using the window function g and its derivative g'. Hence, the computed IF depends on both windows.

From the similar manipulations, the GD can be expressed as

$$\begin{aligned}\mathrm{GD}_g^f(t,\omega) &= -t - \frac{1}{2\pi}\Im\left\{\frac{1}{\mathcal{V}_g f}\cdot\frac{\partial}{\partial\omega}\mathcal{V}_g f\right\} \\ &= -t - \frac{1}{2\pi}\Im\left\{-\mathrm{i}2\pi t - \mathrm{i}\frac{1}{\mathcal{V}_g f}\cdot\mathcal{V}_{g_\mathrm{T}} f\right\} \\ &= \frac{1}{2\pi}\Re\left\{\frac{\mathcal{V}_{g_\mathrm{T}} f}{\mathcal{V}_g f}\right\}, \end{aligned} \tag{2.27}$$

where $g_\mathrm{T}(t) = \frac{t}{2\pi}g(t)$. Hence, (2.11) and (2.12) can be rewritten as

$$\widehat{t}(t,\omega) = t + \frac{1}{2\pi}\Re\left\{\frac{\mathcal{V}_{g_\mathrm{T}} f}{\mathcal{V}_g f}\right\}, \tag{2.28}$$

$$\widehat{\omega}(t,\omega) = \omega - \frac{1}{2\pi}\Im\left\{\frac{\mathcal{V}_{g'} f}{\mathcal{V}_g f}\right\}. \tag{2.29}$$

From the above equations, the reassignment is calculated from the three STFTs: $\mathcal{V}_g f$, $\mathcal{V}_{g'} f$, and $\mathcal{V}_{g_\mathrm{T}} f$.

2.3 Sparsity-aware methods

Sparsity-aware methods aim to find a sparse solution of an underdetermined system. The sparsity-aware methods for estimating T-F representations prepare the overcomplete dictonaries localized in the T-F domain and estimate sparse coefficients representing the signal [23–32]. Matching pursuit [23] and basis pursuit [24] are known as methods to find sparse solutions. In particular, basis pursuit, which finds the solution minimizing the ℓ_1 norm, is widely employed in sparse optimizations and estimating sparse T-F representations. Although there are many possible

choices of dictionaries [29], this section describes the basis pursuit using the Gabor frame for discrete signals.

2.3.1 Gabor frame

Let $\mathbf{g} \in \mathbb{R}^L$ denote a window. A Gabor system is defined as a collection of windowed sinusoids [2, 3],

$$\mathcal{G}(\mathbf{g}, a, M) = \{\mathbf{g}_{m,n}\}_{m=0,\dots,M-1,\, n=0,\dots,N-1}\,, \tag{2.30}$$

where

$$\mathbf{g}_{m,n}[l] = \mathrm{e}^{\mathrm{i}\frac{2\pi m(l-an)}{M}}\mathbf{g}[l-an], \tag{2.31}$$

$a \in \mathbb{N}$ is the time-shifting width, and $M \in \mathbb{N}$ is the number of frequency channels. The discrete Gabor transform (DGT) and inverse DGT with respect to the Gabor system $\mathcal{G}(\mathbf{g}, a, M)$ are defined by

$$(\mathbf{G}_{\mathbf{g}}^{*}\mathbf{f})[m+nM] = \langle \mathbf{f}, \mathbf{g}_{m,n}\rangle\,, \qquad \mathbf{G}_{\mathbf{g}}\mathbf{c} = \sum_{m,n}\mathbf{c}[m+nM]\mathbf{g}_{m,n}, \tag{2.32}$$

where $\mathbf{c} \in \mathbb{C}^{MN}$ is a collection of the coefficients corresponding to a T-F representation. DGT can be viewed as the discrete and downsampled STFT. In the acoustic signal processing community, "STFT" often refers to DGT.

A system $\mathcal{G}(\mathbf{g}, a, M)$ is said to be a frame [2,3] if there exist $0 < A, B < \infty$ such that

$$A\,\|\mathbf{f}\|_2^2 \le \sum_{m,n}|\langle \mathbf{f}, \mathbf{g}_{m,n}\rangle|^2 \le B\,\|\mathbf{f}\|_2^2\,, \tag{2.33}$$

for all $\mathbf{f} \in \mathbb{C}^L$. A and B are called the lower and upper frame bound, respectively. If the Gabor system is a frame, there exists a dual frame $\mathcal{G}(\mathbf{h}, a, M) = \{\mathbf{h}_{m,n}\}$ which satisfies

$$\mathbf{f} = \sum_{m,n}\langle \mathbf{f}, \mathbf{g}_{m,n}\rangle\,\mathbf{h}_{m,n}. \tag{2.34}$$

That is, a T-F representation $\mathbf{c}$ satisfying $\mathbf{f} = \mathbf{G}_{\mathbf{g}}\mathbf{c}$ can be obtained by DGT with a dual window $\mathbf{h}$ associated with $\mathbf{g}$. When a Gabor frame is redundant, the corresponding dual window $\mathbf{h}$ is not unique. A standard construction of the dual window is the canonical dual window:

$$\tilde{\mathbf{h}} = (\mathbf{G}_{\mathbf{g}}\mathbf{G}_{\mathbf{g}}^{*})^{-1}\mathbf{g}. \tag{2.35}$$

The inverse DGT using the canonical dual window $\mathbf{G}_{\tilde{\mathbf{h}}}^{*}$ corresponds to the Moore–Penrose pseudo-inverse of $\mathbf{G}_{\mathbf{g}}$.

2.3.2 Sparse T-F representation using ℓ_1 norm

A T-F representation obtained by DGT $\mathbf{G}_{\mathbf{h}}^{*}$ is spread due to multiplication of a dual window $\mathbf{h}$. If $\mathcal{G}(\mathbf{g}, a, M)$ is a frame, the T-F representation $\mathbf{c}$ is a redundant representation of a time-domain signal $\mathbf{f}$, i.e., the T-F representation $\mathbf{c}$ satisfying

$$\mathbf{f} = \mathbf{G}_{\mathbf{g}}\mathbf{c} \tag{2.36}$$

is not unique. The direct formulation for finding the sparsest solution of this underdetermined system is to minimize the ℓ_0 norm. This problem is usually an intractable combinatorial optimization problem. Instead of the ℓ_0 norm, the ℓ_1 norm has been widely used as a cost function for promoting sparsity [24–32]. It is formulated as

$$\underset{\mathbf{c}}{\text{minimize}} \quad \|\mathbf{c}\|_1 \quad \text{subject to} \quad \mathbf{f} = \mathbf{G}_{\mathbf{g}}\mathbf{c}, \tag{2.37}$$

which is referred to as basis pursuit. This problem is convex and thus can be solved by convex optimization algorithms.

A global optimum of a convex optimization problem can be found independently of the choice of algorithm. Furthermore, Prior information other than sparsity can be introduced in the formulation using ℓ_1 norm. For example, if the signal contains additive noise, a sparse T-F representation can be estimated with a formulation that relaxes the constraint, which is written as

$$\underset{\mathbf{c}}{\text{minimize}} \quad \frac{1}{2}\|\mathbf{f} - \mathbf{G}_{\mathbf{g}}\mathbf{c}\|_2^2 + \lambda \|\mathbf{c}\|_1 . \tag{2.38}$$

The parameter λ adjusts for the sparsity and data fidelity. This formulation is called basis pursuit denoising or lasso. Furthermore, the formulation that take into account the relationship between neighboring dictionaries can also be made by group sparsity [28]. Since an iterative algorithm is required to solve the ℓ_1-norm minimization problem, its computational complexity is greater than those of the reassignment method and SST.

Chapter 3

Window functions with minimum-sidelobe derivatives for computing instantaneous frequency

3.1 Introduction

The short-time Fourier transform (STFT) is widely used to convert a signal into the T-F domain, owing to its simplicity and well-understood structure [1–4]. The resolution of a T-F representation obtained by the STFT is limited by Heisenberg's uncertainty principle. The reassignment method and the synchrosqueezing transform (SST) have been proposed to overcome Heisenberg's uncertainty principle [5]. The reassignment method sharpens a T-F representation using the time and frequency derivatives of its phase at the expense of invertibility. The time and frequency derivatives of the STFT phase are referred to as the instantaneous frequency (IF) and group delay, respectively. The SST performs frequency-only reassignment to a complex-valued T-F representation to sharpen the T-F representation while ensuring invertibility. The FSST reassigns the spread components using the IF, which is affected by the window function. Therefore, to improve the performance of the FSST, the window should be designed considering the computation of the IF.

Moreover, IF has also been employed in other applications, such as a phase vocoder [68–71], T-F mask estimation [72, 73], phase conversion [66, 74–76], and speech processing [77–79]. In the context of a phase vocoder, the effect of a window on the IF computation has been demonstrated by comparing several existing windows [70]. Therefore, designing a window for computing the IF can also improve the performance of its applications.

A method to compute the IF uses the STFTs with a window and with its (time-)derivative [7], which can compute accurately even in a discrete setting. That is, the computed IF depends on both the window and its derivative. Furthermore, interference of multiple signal components influences the IF computation. To reduce the interference of multiple signal components, the window and its derivative should be designed to reduce the T-F spreading under Heisenberg's uncertainty.

The main purpose of a window design for reducing spreading is its frequency response because the spread in the time direction can be controlled relatively easily by the support of the window. Hence, the sidelobes of the frequency responses of the window and its derivative need to be reduced. In particular, the sidelobes of the window derivative should be given more attention since the differential operator emphasizes high-frequency components. Several window functions are designed by considering the continuity at their edges, which is related to the sidelobe of their derivatives [43, 44]. However, no method has explicitly considered the frequency response of the window derivative. Designing a window function to minimize the sidelobes of the frequency response of its derivative is expected to obtain a window function that is more suitable for IF computation.

Therefore, in this chapter, we propose a framework for designing a window function for IF computation. The proposed method first designs the window derivative to minimize the sidelobes and then estimates the window function from the designed window derivative. The designed windows are evaluated by the IF computation and an application to the FSST.

The rest of this chapter is organized as follows. Sections 3.2 and 3.3 introduce the IF of the STFT and bandwidth-adjustable windows, respectively. Then, Section 3.4 explains our proposed method for designing windows. Section 3.5 presents the frequency responses of windows designed by the proposed method. Section 3.6 provides numerical experiments to evaluate the performance of the designed windows in terms of computing the IF. Section 3.7 presents the performance of applying

the designed window to the FSST, and the conclusion of this chapter is presented in Section 3.8.

3.2 Preliminary

3.2.1 Computing IF in Discrete STFT

The discrete version of the STFT for a discrete signal $\mathbf{f} \in \mathbb{C}^L$ with a discrete window function $\mathbf{g} \in \mathbb{C}^L$ is written as

$$(\mathbf{V_g f})[n,m] \coloneqq \sum_{l=0}^{L-1} \mathbf{f}[l]\overline{\mathbf{g}[l-n]}\mathrm{e}^{-\mathrm{i}\frac{2\pi ml}{L}} \tag{3.1}$$

where $n = 0, 1, \ldots, L-1$ is the time-shift index and $m = 0, 1, \ldots, L-1$ is the modulation index. $\mathbf{g}[l]$ outside the domain $[0, L-1]$ is evaluated as an L-periodic sequence [2], i.e.,

$$\mathbf{g}[l+L] = \mathbf{g}[l]. \tag{3.2}$$

As in (2.24), the IF of the discrete STFT can be computed by the window $\mathbf{g}$ and its spectral derivative $\mathbf{g}' = \mathbf{Dg}$,

$$\mathrm{IF}^{\mathbf{f}}_{\mathbf{g}}[n,m] = -\frac{1}{2\pi}\Im\left\{\frac{(\mathbf{V_{g'}f})[n,m]}{(\mathbf{V_g f})[n,m]}\right\}. \tag{3.3}$$

The spectral differentiation operator $\mathbf{D}$ is defined as

$$\begin{aligned}(\mathbf{Dg})[l] \coloneqq& \frac{1}{L}\sum_{m=0}^{L-1}\mathbf{d}[m]\hat{\mathbf{g}}[m]\mathrm{e}^{\mathrm{i}\frac{2\pi lm}{L}} \\ =&(\mathbf{F}^{-1}\,\mathrm{diag}(\mathbf{d})\mathbf{Fg})[l],\end{aligned} \tag{3.4}$$

where

$$\mathbf{d}[m] \coloneqq \begin{cases} \mathrm{i}2\pi m/L & \text{if } 0 \le m < L/2 \\ 0 & \text{if } m = L/2 \\ \mathrm{i}2\pi(m-L)/L & \text{if } L/2 < m \le L-1 \end{cases}. \tag{3.5}$$

The rationale for using the spectral derivative as a counterpart to the continuous-time derivative is given by [8][1].

[1] To be precise, [8] considered a counterpart of the continuous-time derivative for an infinite-length sequence. The rationale for using the spectral derivative was derived by a straightforward adaptation of [8] for a discrete and finite-length signal.

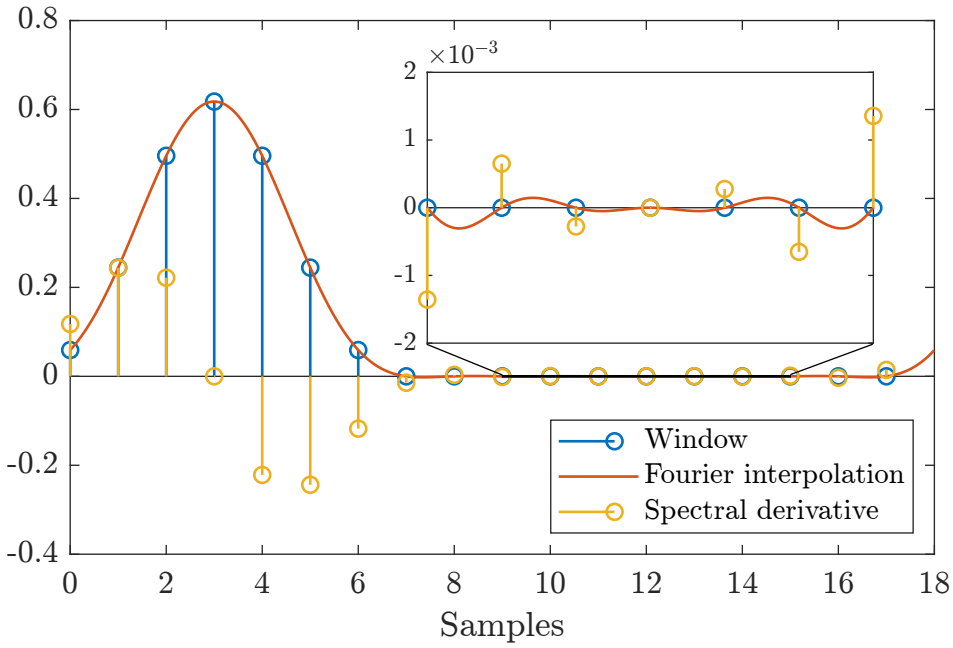

Figure 3.1: Spectral differentiation of a window function supported on $[0, N-1]$ ($N = 7$).

An example of a window function and its derivative obtained by spectral differentiation is illustrated in Fig. 3.1. In Fig. 3.1, the signal length is $L = 18$, and the window is supported on $[0, N-1]$ ($N = 7$). The spectral differentiation corresponds to the differentiation of a function obtained by the Fourier interpolation of a discrete signal. A function obtained by the Fourier interpolation of the window oscillates slightly outside of $[0, N-1]$, even if the window is supported on $[0, N-1]$. Therefore, the support of the window does not coincide with the support of its derivative. In general, the spectral derivative of a window supported on $[0, N-1]$ does not have the same support when $N < L$. When $N = L$, the STFT requires computing the DFT of the entire signal, which reduces its applicability to real-time processing or lengthy signal analysis [80–82]. In practice, the window derivative is truncated to have the same support as the original window.

3.2.2 Comparison of spectral and analytical derivatives

When the window function is initially defined as a differentiable continuous function, an alternative to spectral differentiation is to calculate the window derivative analytically and sample the derivative. Let us illustrate this method using the

Hann window, which is one of the most famous of such window functions. The Hann window supported on $[0, N-1]$ is defined as

$$g(t) = \begin{cases} \frac{1}{2}\left(1 - \cos\left(\frac{2\pi t}{N-1}\right)\right) & \text{if } t \in [0, N-1] \\ 0 & \text{otherwise} \end{cases}. \tag{3.6}$$

Its analytical derivative is given by

$$g'(t) = \begin{cases} \frac{\pi}{N-1}\sin\left(\frac{2\pi t}{N-1}\right) & \text{if } t \in [0, N-1] \\ 0 & \text{otherwise} \end{cases}. \tag{3.7}$$

Here, we compare the spectral and analytical derivatives in terms of the computed IF of a sinusoid. Considering a discrete complex sinusoid

$$\mathbf{s}[l] = A_{\mathbf{s}}\mathrm{e}^{\mathrm{i}(2\pi\xi_{\mathbf{s}}l + \phi_{\mathbf{s}})}, \tag{3.8}$$

for $l = 0, 1, \ldots, L-1$, we evaluated the error between the analytical IF of the sinusoid (2.23) and the computed IF

$$e[n, m] = \left(\xi_{\mathbf{s}} - \frac{m}{L}\right) - \mathrm{IF}^{\mathbf{s}}_{\mathbf{g}}[n, m]. \tag{3.9}$$

With the same manipulations in (2.21), the STFT of sinusoid $\mathbf{s}$ is calculated as

$$(\mathbf{V}_{\mathbf{g}}\mathbf{s})[n, m] = \overline{(\mathcal{F}\mathbf{g})\left(\xi_{\mathbf{s}} - \frac{m}{L}\right)}A_{\mathbf{s}}\mathrm{e}^{\mathrm{i}\phi_{\mathbf{s}} - \mathrm{i}2\pi\left(\frac{m}{L} - \xi_{\mathbf{s}}\right)n}. \tag{3.10}$$

Then, the IF of sinusoid $\mathbf{s}$ is calculated as

$$\begin{aligned} \mathrm{IF}^{\mathbf{s}}_{\mathbf{g}}[n, m] &= -\frac{1}{2\pi}\Im\left\{\frac{(\mathbf{V}_{\mathbf{g}'}\mathbf{s})[n, m]}{(\mathbf{V}_{\mathbf{g}}\mathbf{s})[n, m]}\right\} \\ &= -\frac{1}{2\pi}\Im\left\{\frac{\overline{(\mathcal{F}\mathbf{g}')(\xi_{\mathbf{s}} - \frac{m}{L})}}{\overline{(\mathcal{F}\mathbf{g})(\xi_{\mathbf{s}} - \frac{m}{L})}}\right\}, \end{aligned} \tag{3.11}$$

which is independent of the time index n. When $\mathbf{g}$ is real-valued, (3.11) is simplified as

$$\mathrm{IF}^{\mathbf{s}}_{\mathbf{g}}[n, m] = -\frac{1}{2\pi}\Im\left\{\frac{(\mathcal{F}\mathbf{g}')(\frac{m}{L} - \xi_{\mathbf{s}})}{(\mathcal{F}\mathbf{g})(\frac{m}{L} - \xi_{\mathbf{s}})}\right\}. \tag{3.12}$$

Thus, the error (3.9) is rewritten as

$$e[m] = \left(\xi_{\mathbf{s}} - \frac{m}{L}\right) + \frac{1}{2\pi}\Im\left\{\frac{(\mathcal{F}\mathbf{g}')(\frac{m}{L} - \xi_{\mathbf{s}})}{(\mathcal{F}\mathbf{g})(\frac{m}{L} - \xi_{\mathbf{s}})}\right\}. \tag{3.13}$$

Fig. 3.2 shows the frequency responses of the window derivatives and errors of IF computation using these window derivatives. The window length and the signal

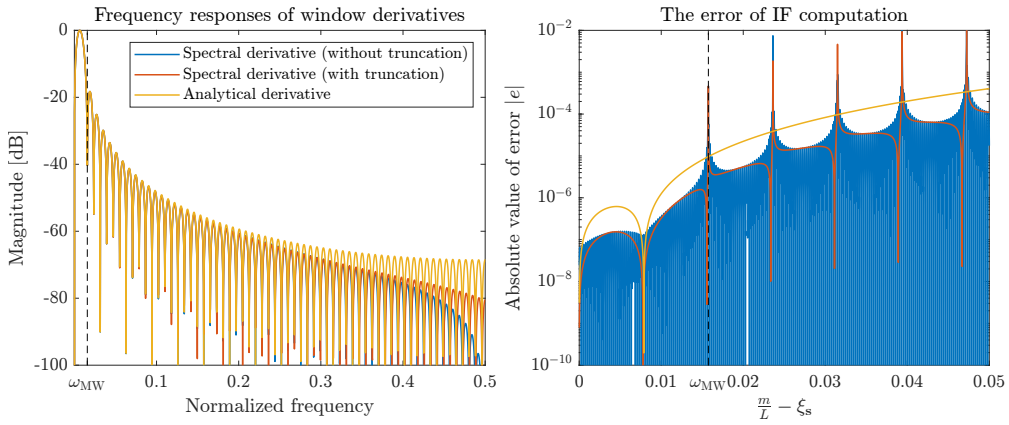

Figure 3.2: Comparison of derivatives of the Hann window. The left and right sides show the frequency responses of the window derivatives and the errors of the IF computation using these window derivatives.

length were set to $N = 2^7$ and $L = 2^{12}$. The mainlobe width ω_{MW} is defined as the first null point of the frequency response except for $\omega = 0$. The error in the IF computed by the analytical derivative is larger than that of the spectral derivative. Therefore, this chapter focuses on spectral differentiation, even if these windows are analytically differentiable.

3.3 Bandwidth-adjustable windows

One aim of the window design is to control T-F spreading under Heisenberg's uncertainty principle. The spread in the time direction can be controlled by setting the length of the window function. We assume that a window $\mathbf{g}$ is supported on $[0, N-1]$ and let $\mathbf{w} \in \mathbb{C}^N$ denote its nonzero part, i.e.,

$$\mathbf{g}[l] = \begin{cases} \mathbf{w}[l] & \text{if } l = 0, 1, \ldots, N-1 \\ 0 & \text{otherwise} \end{cases}. \tag{3.14}$$

To obtain a well-localized T-F representation, a window should be designed so that its frequency response has a narrow mainlobe and low sidelobe levels under the defined window length N. However, a window function has a trade-off between the mainlobe width and the sidelobe level. The mainlobe width is closely related to the appearance of the T-F representation and can be intuitively chosen according to the application. Thus, the sidelobe characteristics should be optimized under the mainlobe width ω_{MW} chosen according to the application. The sidelobe energy

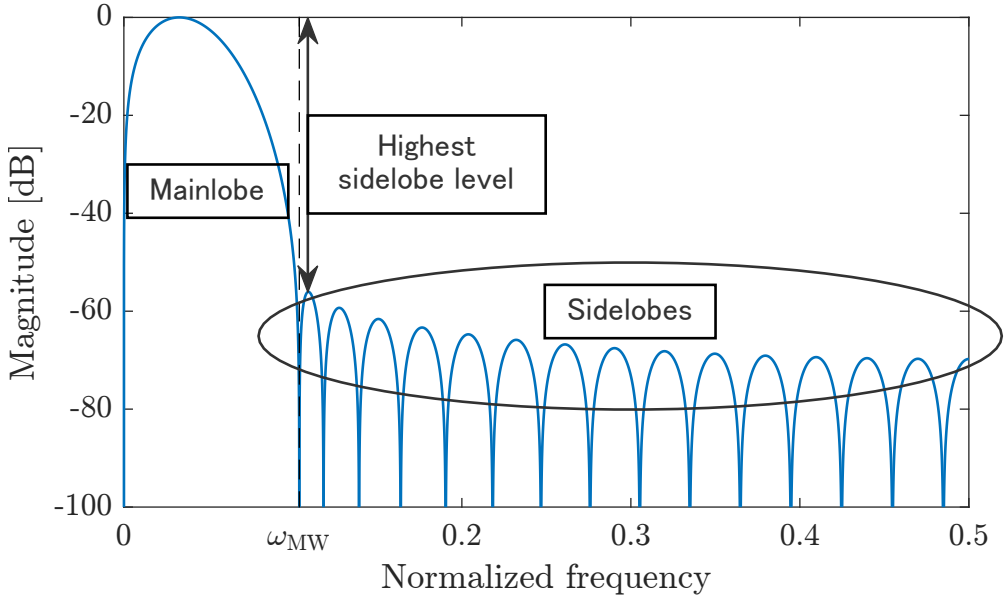

Figure 3.3: Sidelobe energy and highest sidelobe level of a window function.

(SE) and highest sidelobe level (HSL) are used to evaluate the localization of the frequency response of the window functions, which are defined as

$$\mathrm{SE} = 10\log_{10}\frac{\int_{-\frac{1}{2}}^{\frac{1}{2}} \mathcal{W}_{\omega_{\mathrm{MW}}}(\omega)\,|(\mathcal{F}\mathbf{w})(\omega)|^2\,\mathrm{d}\omega}{\int_{-\frac{1}{2}}^{\frac{1}{2}} |(\mathcal{F}\mathbf{w})(\omega)|^2\,\mathrm{d}\omega}, \tag{3.15}$$

$$\mathrm{HSL} = 10\log_{10}\frac{\max_\omega \mathcal{W}_{\omega_{\mathrm{MW}}}(\omega)\,|(\mathcal{F}\mathbf{w})(\omega)|^2}{\max_\omega |(\mathcal{F}\mathbf{w})(\omega)|^2}, \tag{3.16}$$

where $\mathcal{W}_{\omega_{\mathrm{MW}}}(\omega)$ is a weight function,

$$\mathcal{W}_{\omega_{\mathrm{MW}}}(\omega) = \begin{cases} 0 & \text{if } |\omega| < \omega_{\mathrm{MW}} \\ 1 & \text{if } |\omega| \geq \omega_{\mathrm{MW}} \end{cases}. \tag{3.17}$$

The SE and HSL of a window function are shown in Fig. 3.3.

Many windows, such as the rectangular window, Bartlett window, Hann window, Blackman window, and Nuttall window [43] depend only on the window length. However, some window functions contain additional parameters for adjusting the bandwidth; these are referred to as *adjustable windows*. The Dolph–Chebyshev window [35], the Slepian window [36], the Kaiser window [38], the Saramäki window [39], the ultraspherical window [40,41], the cosh window [42], and the Tukey window (tapered cosine window) [83] belong to this class. Among them, the Slepian window and the Dolph–Chebyshev window are characterized by the following optimization

problems:

$$\underset{\mathbf{w}}{\text{minimize}} \quad \frac{\int_{-\frac{1}{2}}^{\frac{1}{2}} \mathcal{W}_W(\omega)\,|(\mathcal{F}\mathbf{w})(\omega)|^2\,\mathrm{d}\omega}{\int_{-\frac{1}{2}}^{\frac{1}{2}} |(\mathcal{F}\mathbf{w})(\omega)|^2 \mathrm{d}\omega}, \tag{3.18}$$

$$\underset{\mathbf{w}}{\text{minimize}} \quad \frac{\max_{\omega\in[-\frac{1}{2},\frac{1}{2}]} \ \mathcal{W}_W(\omega)|(\mathcal{F}\mathbf{w})(\omega)|}{\max_{\omega\in[-\frac{1}{2},\frac{1}{2}]} \ |(\mathcal{F}\mathbf{w})(\omega)|}, \tag{3.19}$$

respectively, where $W \in (0, \frac{1}{2}]$. These windows are designed to have a well-localized frequency response in terms of SE and HSL, and they certainly show better characteristics than other window functions (as indicated in the left side of Fig. 3.10).

3.4 Proposed method

We now propose a window design method to reduce the influence of the sidelobes of the window derivative on the IF computation. A comparison between the conventional and proposed methods of computing a window function and a window derivative pair is illustrated in Fig. 3.4. In general, to obtain a window and its derivative, the window function is first designed, and then the window derivative is calculated by differentiating the designed window [Fig. 3.4 (a)]. By contrast, our method first designs the window derivative to minimize the sidelobes and then estimates the window function from the window derivative [Fig. 3.4 (b)].

3.4.1 Problem formulation

To restrict the spread in the time direction, we assume that the window derivative is supported on $[0, N-1]$, i.e.,

$$\mathbf{g}'[l] = \begin{cases} \mathbf{z}[l] & \text{if } l = 0, 1, \ldots, N-1 \\ 0 & \text{otherwise} \end{cases}, \tag{3.20}$$

where $\mathbf{z} \in \mathbb{C}^N$ corresponds to the nonzero part of $\mathbf{g}'$. Calculating the window derivative from the window function can be performed straightforwardly by spectral differentiation. Conversely, from $\mathbf{z}$ such that $\langle \mathbf{z}, \mathbf{1}_N \rangle = 0$, $\mathbf{g}$ can be calculated as the spectral integration:

$$\mathbf{g} = \mathbf{F}^{-1}\,\mathrm{diag}(\mathbf{b})\mathbf{F}\mathbf{P}_{L,N}\mathbf{z} + c\mathbf{1}_L, \tag{3.21}$$

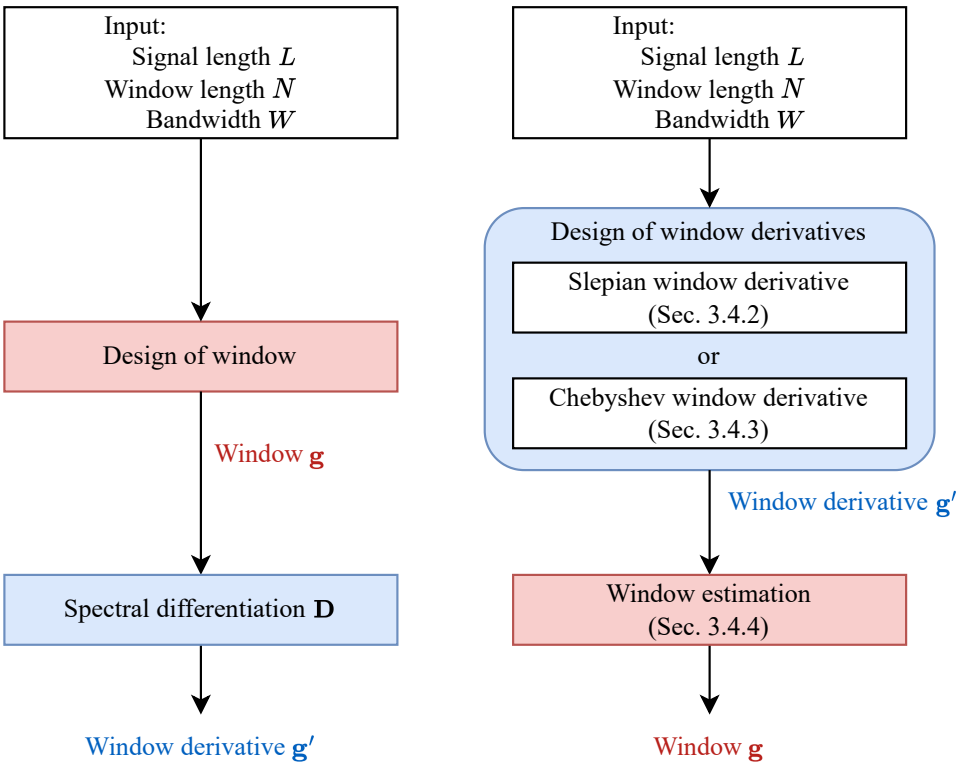

Figure 3.4: Block diagrams of (a) the conventional computation of a window function and a window derivative pair and (b) the proposed method.

where

$$\mathbf{b}[m] \coloneqq \begin{cases} 0 & m = 0 \\ -\mathrm{i}L/2\pi m & \text{if } 0 < m < L/2 \\ 0 & \text{if } m = L/2 \\ -\mathrm{i}L/2\pi(m-L) & \text{if } L/2 < m \leq L-1 \end{cases},$$

$\mathbf{P}_{L,N} \in \mathbb{R}^{L \times N}$ is the zero-padding matrix,

$$\mathbf{P}_{L,N} \coloneqq [\mathbf{I}_N, \mathbf{O}_{N,L-N}]^{\mathrm{T}}, \tag{3.22}$$

and c is an integral constant. $\langle \mathbf{z}, \mathbf{1}_N \rangle = 0$ constrains the integration of the Fourier series to also be a Fourier series. Therefore, we formulate the design of the window

derivative as

$$\begin{aligned} &\underset{\mathbf{z}}{\text{minimize}} \ \Theta(\mathbf{z}) \\ &\text{subject to } \langle \mathbf{z}, \mathbf{1}_N \rangle = 0, \end{aligned} \tag{3.23}$$

where $\Theta(\mathbf{z})$ is an objective function measuring the sidelobes of the frequency response of $\mathbf{z}$, which corresponds to the SE or HSL of $\mathbf{z}$. In summary, we first design the window derivative $\mathbf{z}$ by (3.23) and then estimate the window function from $\mathbf{z}$ by (3.21).

In the remainder of this section, Sec. 3.4.2 and Sec. 3.4.3, we explain the methods for designing the window derivative to minimize the SE and HSL, respectively. Estimating the original window $\mathbf{w}$ from the designed window derivative $\mathbf{z}$ is introduced in Sec. 3.4.4. Hereafter, the window derivatives minimizing the SE and HSL are referred to as the *Slepian window derivative* and the *Chebyshev window derivative*, respectively.

3.4.2 Slepian window derivative

This subsection explains the design of the Slepian window derivative. Considering the case where the cost function $\Theta(\mathbf{z})$ in (3.23) is the ratio of the energy outside $[-W, W]$ to the total energy, the design problem of the Slepian window derivative is formulated as

$$\begin{aligned} &\underset{\mathbf{z}}{\text{minimize}} \quad \frac{\int_{-\frac{1}{2}}^{\frac{1}{2}} \mathcal{W}_W(\omega) \left|(\mathcal{F}\mathbf{z})(\omega)\right|^2 \mathrm{d}\omega}{\int_{-\frac{1}{2}}^{\frac{1}{2}} \left|(\mathcal{F}\mathbf{z})(\omega)\right|^2 \mathrm{d}\omega} \\ &\text{subject to} \quad \langle \mathbf{z}, \mathbf{1}_N \rangle = 0. \end{aligned} \tag{3.24}$$

The cost function in Eq. (3.24) can be rewritten as the Rayleigh quotient:

$$\frac{\mathbf{z}^{\mathrm{T}}\mathbf{z} - \mathbf{z}^{\mathrm{T}}\mathbf{S}_N\mathbf{z}}{\mathbf{z}^{\mathrm{T}}\mathbf{z}} = 1 - \frac{\mathbf{z}^{\mathrm{T}}\mathbf{S}_N\mathbf{z}}{\mathbf{z}^{\mathrm{T}}\mathbf{z}}, \tag{3.25}$$

where $\mathbf{S}_N \in \mathbb{R}^{N \times N}$ is a real-symmetric matrix whose elements are given by

$$\mathbf{S}_N[m, n] = 2W \operatorname{sinc}(2W(m - n)), \tag{3.26}$$

$$\operatorname{sinc}(x) \coloneqq \begin{cases} \frac{\sin(\pi x)}{\pi x} & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases}. \tag{3.27}$$

Fixing $\mathbf{z}^{\mathrm{T}}\mathbf{z} = 1$ and reducing the constant 1 not relevant to minimization, (3.24) can be rewritten as

$$\begin{aligned} \underset{\mathbf{z}}{\text{minimize}} \quad & -\mathbf{z}^{\mathrm{T}}\mathbf{S}_N\mathbf{z} \\ \text{subject to} \quad & \mathbf{z}^{\mathrm{T}}\mathbf{z} = 1, \quad \mathbf{z}^{\mathrm{T}}\mathbf{1}_N = 0. \end{aligned} \tag{3.28}$$

Such a problem can be simplified to an eigenvalue problem [84].

Let the Lagrangian function associated with (3.28) be

$$\mathcal{L}(\mathbf{z}, \mu, \eta) = -\mathbf{z}^{\mathrm{T}}\mathbf{S}_N\mathbf{z} + \mu(\mathbf{z}^T\mathbf{z} - 1) + 2\eta\mathbf{z}^{\mathrm{T}}\mathbf{1}_N, \tag{3.29}$$

where μ and η are the Lagrange multipliers. The optimal solution $\mathbf{z}^\star$ to (3.28) satisfies the following necessary conditions:

$$\frac{\partial}{\partial \mathbf{z}}\mathcal{L}(\mathbf{z}^\star, \mu^\star, \eta^\star) = 2\left(-\mathbf{S}_N\mathbf{z}^\star + \mu^\star\mathbf{z}^\star + \eta^\star\mathbf{1}_N\right) = \mathbf{0}, \tag{3.30}$$

$$\frac{\partial}{\partial \mu}\mathcal{L}(\mathbf{z}^\star, \mu^\star, \eta^\star) = \mathbf{z}^{\star\mathrm{T}}\mathbf{z}^\star - 1 = 0, \tag{3.31}$$

$$\frac{\partial}{\partial \eta}\mathcal{L}(\mathbf{z}^\star, \mu^\star, \eta^\star) = \mathbf{z}^{\star\mathrm{T}}\mathbf{1}_N = 0. \tag{3.32}$$

Multiplying (3.30) on the left by $\mathbf{1}_N^{\mathrm{T}}$, $\eta^\star$ can be calculated as

$$\eta^\star = \frac{1}{N}\mathbf{1}_N^{\mathrm{T}}\mathbf{S}_N\mathbf{z}^\star. \tag{3.33}$$

Substituting $\eta^\star$ into (3.30), we obtain

$$\mathbf{T}_N\mathbf{S}_N\mathbf{z}^\star = \mu^\star\mathbf{z}^\star, \tag{3.34}$$

where $\mathbf{T}_N = \mathbf{I}_N - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^{\mathrm{T}}$ is the projection matrix onto $\left\{\mathbf{z} \in \mathbb{C}^N \,\middle|\, \mathbf{z}^{\mathrm{T}}\mathbf{1}_N = 0\right\}$. Hence, denoting $\mathbf{K}_N = \mathbf{T}_N\mathbf{S}_N$, the eigenvectors of $\mathbf{K}_N$ are candidates for the solution to the problem (3.28). Furthermore, since the eigenvectors of $\mathbf{K}_N$ satisfy (3.31) and (3.32), multiplying (3.34) on the left by $\mathbf{z}^{\mathrm{T}}$, we get

$$\begin{aligned} \mathbf{z}^{\mathrm{T}}\mathbf{T}_N\mathbf{S}_N\mathbf{z} &= \mu\mathbf{z}^{\mathrm{T}}\mathbf{z}, \\ \mathbf{z}^{\mathrm{T}}\mathbf{S}_N\mathbf{z} &= \mu. \end{aligned} \tag{3.35}$$

Therefore, the eigenvector corresponding to the largest eigenvalue μ_0 is the solution to the problem (3.28). However, finding the eigenvalues of $\mathbf{K}_N$ is numerically ill-conditioned; likewise, $\mathbf{S}_N$. This ill-conditionedness follows from the following fact and proposition.

Fact 1. *The eigenvalues of* $\mathbf{S}_N$ *are nondegenerate and take values between zero and one. Furthermore, most eigenvalues of* $\mathbf{S}_N$ *are clustered around* 1 *or* 0 *[36].*

Fact 2. *Let* $\lambda_0, \lambda_1, \ldots, \lambda_{N-1}$ *denote the eigenvalues of* $\mathbf{S}_N$ *such that*

$$1 > \lambda_0 > \lambda_1 > \cdots > \lambda_{N-1} > 0, \tag{3.36}$$

and their corresponding eigenvectors be $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{N-1}$*, whose norm and signs are determined so that*

$$\|\mathbf{v}_k\| = 1, \quad \sum_{n=0}^{N-1} \mathbf{v}_k[n] \geq 0 \quad \textit{for } k = 0, 1, \ldots, N-1. \tag{3.37}$$

The eigenvectors $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{N-1}$ *of* $\mathbf{S}_N$ *have the following properties [36]:*

$$\textbf{\textit{Orthogonality:}} \quad \langle \mathbf{v}_j, \mathbf{v}_k \rangle = 0 \quad \textit{for } j \neq k, \tag{3.38}$$

$$\textbf{\textit{Symmetry:}} \quad \mathbf{v}_k[n] = (-1)^k \mathbf{v}_k[N-n-1], \tag{3.39}$$

for $k = 0, 1, \ldots, N-1$.

Proposition 1. $\mathbf{v}_k$ *for* $k = 1, 3, \ldots, 2\lfloor N/2 \rfloor - 1$ *are the eigenvectors of* $\mathbf{K}_N$.

Proof. According to (3.39), $\mathbf{v}_k$ for $k = 1, 3, \ldots, 2\lfloor N/2 \rfloor - 1$ satisfies

$$\mathbf{T}_N \mathbf{v}_k = \mathbf{v}_k. \tag{3.40}$$

Then, the following relationship holds:

$$\begin{aligned} \mathbf{K}_N \mathbf{v}_k &= \mathbf{T}_N \mathbf{S}_N \mathbf{v}_k \\ &= \lambda_k \mathbf{T}_N \mathbf{v}_k \\ &= \lambda_k \mathbf{v}_k. \end{aligned} \tag{3.41}$$

Therefore, λ_k and $\mathbf{v}_k$ are the eigenvalues and eigenvectors of $\mathbf{K}_N$, respectively. □

Finding the eigenvalues of $\mathbf{S}_N$ is numerically ill-conditioned since most eigenvalues of $\mathbf{S}_N$ are clustered around 1 or 0, as shown by the blue line in Fig. 3.5. Even if the eigenvalues of $\mathbf{S}_N$ are nondegenerate, they behave as if they are degenerate because of rounding errors in the numerical computation. According to Proposition 1, its eigenvalues λ_k for $k = 1, 3, \ldots, 2\lfloor N/2 \rfloor - 1$ are also eigenvalues of $\mathbf{K}_N$. Hence, most eigenvalues of $\mathbf{K}_N$ are also clustered around 1 or 0, as shown by the red line in Fig. 3.5.

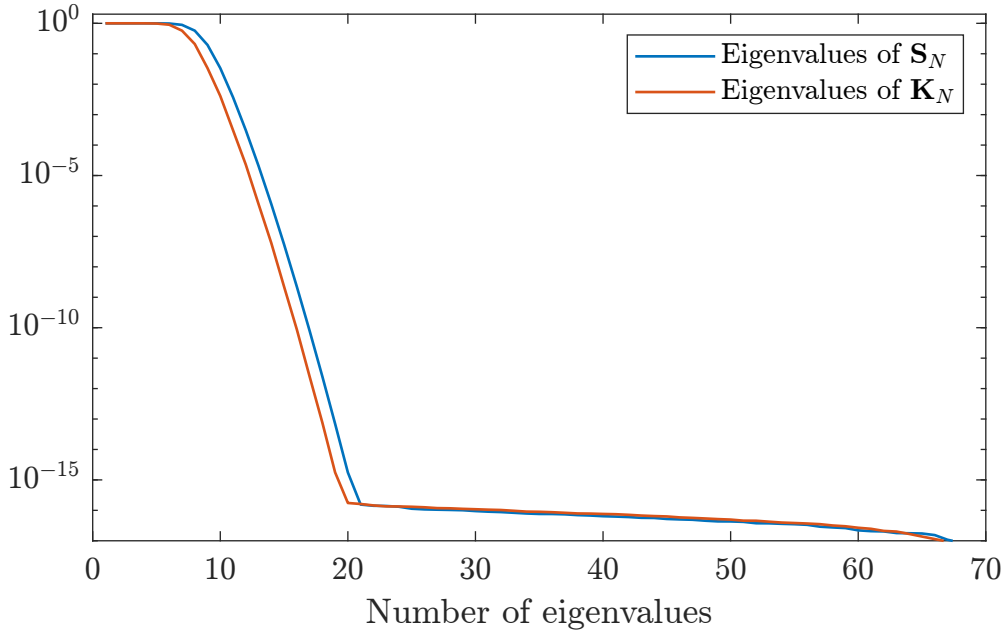

Figure 3.5: Eigenvalues of $\mathbf{S}_N$ and $\mathbf{K}_N$ for $N = 2^7$ and $W = 0.03$.

Here, $\mathbf{K}_N$ is centrosymmetric since both $\mathbf{T}_N$ and $\mathbf{S}_N$ are centrosymmetric [85]. Thus, the eigenvectors are symmetric or antisymmetric, but their order is unclear. When $\mathbf{z}$ are symmetric or antisymmetric, $(\mathcal{F}\mathbf{z})(\omega)$ can be represented as the cosine or sine series [86]. Fig. 3.6 shows $(\mathcal{F}\mathbf{z})(\omega)$ localized in $[-W, W]$ under $\langle \mathbf{z}, \mathbf{1}_N \rangle = 0$. Note that the linear phase of the frequency response is ignored to make the spectrum real-valued for display. The symmetric case contains an extra extremum in $[-W, W]$ compared with the antisymmetric case to satisfy the constraint $\langle \mathbf{z}, \mathbf{1}_N \rangle = 0$. From this observation, the antisymmetric window localizes the frequency response in $[-W, W]$ under the constraint $\langle \mathbf{z}, \mathbf{1}_N \rangle = 0$ more than the symmetric window. Since $\mathbf{v}_k$ for $k = 1, 3, \ldots, 2\lfloor N/2 \rfloor - 1$ are antisymmetric from Proposition 1, the eigenvector $\mathbf{v}_1$ corresponding to the second largest eigenvalue λ_1 is the solution to (3.24). To compute the eigenvector $\mathbf{v}_1$, efficient methods for computing the eigenvectors of $\mathbf{S}_N$ [87, 88] are available instead of computing the eigenvectors of $\mathbf{K}_N$ directly.

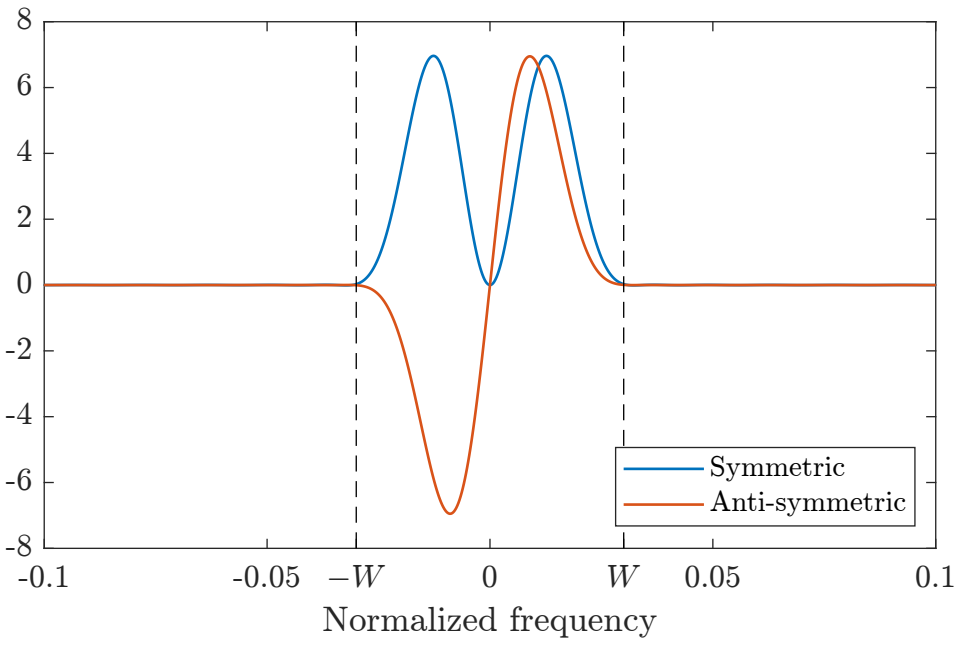

Figure 3.6: Frequency responses $(\mathcal{F}\mathbf{z})(\omega)$ localized in $[-W, W]$ under $\langle \mathbf{z}, \mathbf{1}_N \rangle = 0$ of symmetric and antisymmetric windows. Note that the linear phase of the frequency response is ignored for display.

3.4.3 Chebyshev Window Derivative

In this subsection, we consider the design problem of the Chebyshev window derivative. Similar to (3.24), the direct formulation of the design problem of the Chebyshev window derivative is

$$\begin{aligned} \underset{\mathbf{z}}{\text{minimize}} \quad & \frac{\max_{\omega\in[-\frac{1}{2},\frac{1}{2}]} \ \mathcal{W}_W(\omega)|(\mathcal{F}\mathbf{z})(\omega)|}{\max_{\omega\in[-\frac{1}{2},\frac{1}{2}]} \ |(\mathcal{F}\mathbf{z})(\omega)|} \\ \text{subject to} \quad & \langle \mathbf{z}, \mathbf{1}_N \rangle = 0. \end{aligned} \tag{3.42}$$

Considering the constraint $\langle \mathbf{z}, \mathbf{1}_N \rangle = 0$ and the symmetry of the cost function with (3.24), the solution to (3.42) should be antisymmetric.

When $\mathbf{z}$ is antisymmetric, the following decomposition of $(\mathcal{F}\mathbf{z})(\omega)$ is proposed by McClellan and Parks [86]:

$$(\mathcal{F}\mathbf{z})(\omega) = \mathrm{e}^{\mathrm{i}\left(\frac{\pi}{2}-(N-1)\pi\omega\right)} Q(\omega) P(\omega, \boldsymbol{\alpha}), \tag{3.43}$$

where

$$P(\omega, \boldsymbol{\alpha}) = \sum_{k=0}^{K-1} \boldsymbol{\alpha}[k] \cos(2\pi\omega k), \tag{3.44}$$

$$Q(\omega) = \begin{cases} \sin(2\pi\omega) & \text{if } N \text{ is odd} \\ \sin(\pi\omega) & \text{if } N \text{ is even} \end{cases}, \tag{3.45}$$

$\boldsymbol{\alpha} \in \mathbb{R}^K$, and $K = \lfloor N/2 \rfloor$. The derivation of (3.43) and the relation between $\boldsymbol{\alpha}$ and $\mathbf{z}$ are explained in Appendix A. The linear phase $\mathrm{e}^{\mathrm{i}\left(\frac{\pi}{2}-(N-1)\pi\omega\right)}$ has no effect on the cost function of (3.42). Since $Q(\omega)P(\omega, \boldsymbol{\alpha})$ is antisymmetric, finding the highest sidelobe level only needs to consider the positive frequency. Furthermore, since the constraint $\langle \mathbf{z}, \mathbf{1}_N \rangle = 0$ is always satisfied when $\mathbf{z}$ is antisymmetric, (3.42) can be rewritten as

$$\underset{\boldsymbol{\alpha}}{\text{minimize}} \quad \frac{\max_{\omega\in[0,\frac{1}{2}]} \ \mathcal{W}_W(\omega)|Q(\omega)P(\omega, \boldsymbol{\alpha})|}{\max_{\omega\in[0,\frac{1}{2}]} \ |Q(\omega)P(\omega, \boldsymbol{\alpha})|}. \tag{3.46}$$

Fixing the denominator to be 1 as in (3.28), (3.46) can be rewritten as

$$\begin{aligned} \underset{\boldsymbol{\alpha}}{\text{minimize}} \quad & \max_{\omega\in[0,\frac{1}{2}]} \ \mathcal{W}_W(\omega)|Q(\omega)P(\omega, \boldsymbol{\alpha})| \\ \text{subject to} \quad & \max_{\omega\in[0,\frac{1}{2}]} \ |Q(\omega)P(\omega, \boldsymbol{\alpha})| = 1. \end{aligned} \tag{3.47}$$

This optimization problem is still difficult to solve due to this equality constraint, hence it needs to be further rewritten. Here, the role of the constraint is to

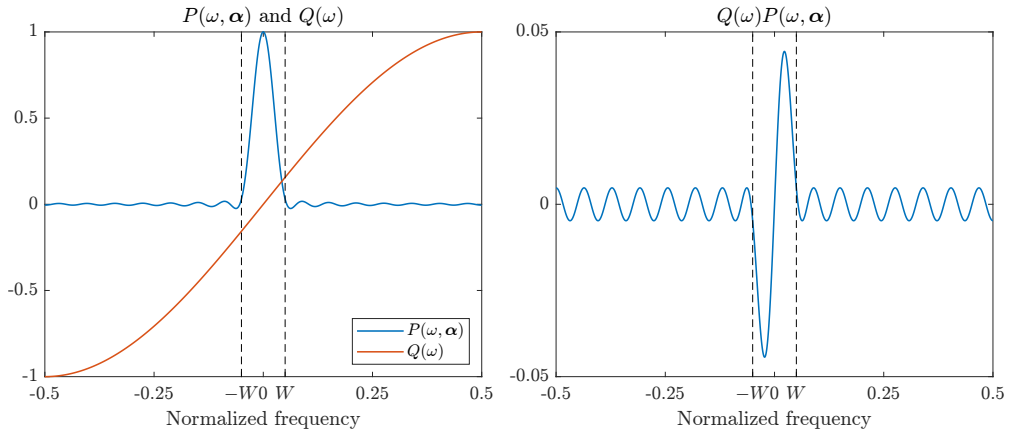

Figure 3.7: $P(\omega,\alpha)$ and $Q(\omega)$ (left) and their product $Q(\omega)P(\omega,\boldsymbol{\alpha})$ (right).

preserve the large peaks of $|Q(\omega)P(\omega,\boldsymbol{\alpha})|$ in $[-W,W]$ while reducing the sidelobe level. Fig. 3.7 shows $P(\omega,\alpha)$, $Q(\omega)$ and their product $Q(\omega)P(\omega,\boldsymbol{\alpha})$. According to Fig. 3.7, the peaks of $Q(\omega)P(\omega,\boldsymbol{\alpha})$ in $[-W,W]$ are composed of the product of $Q(\omega)$ and the mainlobe of $P(\omega,\boldsymbol{\alpha})$ in $[-W,W]$. Since $Q(\omega)$ is independent of $\boldsymbol{\alpha}$, constraining $P(0,\boldsymbol{\alpha})$ to a nonzero value will preserve the large peaks of $|Q(\omega)P(\omega,\boldsymbol{\alpha})|$ in $[-W,W]$. Based on these observations, we formulate the design problem of the Chebyshev window derivative as

$$\begin{aligned} &\underset{\boldsymbol{\alpha}}{\text{minimize}} \quad \max_{\omega\in[0,\frac{1}{2}]} \ \mathcal{W}_W(\omega)|Q(\omega)P(\omega,\boldsymbol{\alpha})| \\ &\text{subject to} \ \ P(0,\boldsymbol{\alpha})=1. \end{aligned} \tag{3.48}$$

(3.48) can be solved using the modified Remez (MRemez) algorithm [89], which is an extension of the Parks–McClellan algorithm [90] to deal with an equality-constrained minimax approximation problem. Introducing a weight function

$$\widetilde{\mathcal{W}}_W^{\epsilon}(\omega)=\begin{cases} \frac{1}{\epsilon} & \text{if } \omega=0 \\ 0 & \text{if } 0<|\omega|<W \\ Q(\omega) & \text{if } W\le|\omega| \end{cases}, \tag{3.49}$$

(3.48) is rewritten as the unconstrained optimization problem,

$$\underset{\boldsymbol{\alpha}}{\text{minimize}} \quad \max_{\omega\in[0,\frac{1}{2}]} \ |E(\omega,\boldsymbol{\alpha})|, \tag{3.50}$$

where

$$E(\omega, \boldsymbol{\alpha}) = \lim_{\epsilon \to 0} \widetilde{\mathcal{W}}_W^{\epsilon}(\omega) \left[P(\omega, \boldsymbol{\alpha}) - D(\omega) \right], \tag{3.51}$$

$$D(\omega) = \begin{cases} 1 & \text{if } \omega = 0 \\ 0 & \text{if } \omega \neq 0 \end{cases}. \tag{3.52}$$

Denoting the solution of (3.48) as $\boldsymbol{\alpha}^\star$, there exist $K+1$ frequencies $\omega_0^\star, \omega_1^\star, \ldots, \omega_K^\star$ satisfying $0 = \omega_0^\star < \omega_1^\star < \cdots < \omega_K^\star \leq \frac{1}{2}$ such that [91]

$$P(\omega_k^\star, \boldsymbol{\alpha}^\star) = \begin{cases} 1 & \text{if } k = 0 \\ \frac{(-1)^k}{\mathcal{W}(\omega_k^\star)} \delta^\star & \text{if } k = 1 \ldots, K \end{cases} \tag{3.53}$$

where

$$|\delta^\star| = \max_{\omega \in [0, \frac{1}{2}]} |E(\omega, \boldsymbol{\alpha}^\star)|. \tag{3.54}$$

The MRemez algorithm finds $\omega_0^\star, \omega_1^\star, \ldots, \omega_K^\star$ using the following procedure:

Step 1. Initialize reference frequencies $\omega_0, \omega_1, \ldots, \omega_K$ so that $\omega_0 = 0$ and $W \leq \omega_1 < \cdots < \omega_K \leq 1/2$.

Step 2. Compute δ and $P(\omega, \boldsymbol{\alpha})$ by the barycentric formula:

$$\delta = \lim_{\epsilon \to 0} \frac{\sum_{k=0}^K \beta_k D(\omega_k)}{\sum_{k=0}^K (-1)^k \beta_k \widetilde{\mathcal{W}}_W^{\epsilon}(\omega_k)}, \tag{3.55}$$

$$P(\omega, \boldsymbol{\alpha}) = \frac{\sum_{k=0}^K \frac{\beta_k}{\cos(2\pi\omega) - \cos(2\pi\omega_k)} p_k}{\sum_{k=0}^K \frac{\beta_k}{\cos(2\pi\omega) - \cos(2\pi\omega_k)}}, \tag{3.56}$$

where

$$p_k = \lim_{\epsilon \to 0} D(\omega_k) + \frac{(-1)^k}{\widetilde{\mathcal{W}}_W^{\epsilon}(\omega_k)} \delta, \tag{3.57}$$

$$\beta_k = \prod_{i=0, i \neq k}^K \frac{1}{\cos(2\pi\omega_k) - \cos(2\pi\omega_i)}. \tag{3.58}$$

Step 3. Compute $E(\omega, \boldsymbol{\alpha})$ by (3.51) and find the new reference frequencies $\omega_0^+, \omega_1^+, \ldots, \omega_K^+$ satisfying [92]

$$\operatorname{sign}\left[E(\omega_{k-1}^+, \boldsymbol{\alpha})\right] = -\operatorname{sign}\left[E(\omega_k^+, \boldsymbol{\alpha})\right], \tag{3.59}$$

$$|\delta| \leq |P(\omega_k^+, \boldsymbol{\alpha})|, \tag{3.60}$$

where at least one of the inequalities is strict.

Step 4. Repeat 2–3 until convergence.

After estimating $\omega_0^\star, \omega_1^\star, \ldots, \omega_K^\star$ using the MRemez algorithm, the Chebyshev window derivative $\mathbf{z}$ can be computed from its frequency response by sampling and performing the inverse DFT,

$$\mathbf{z} = \mathbf{F}^{-1}\hat{\mathbf{z}}, \tag{3.61}$$

$$\hat{\mathbf{z}}[m] = \mathrm{e}^{\mathrm{i}\left(\frac{\pi}{2} - \frac{N-1}{N}\pi m\right)} Q(m/N) P(m/N, \boldsymbol{\alpha}^\star), \tag{3.62}$$

for $m = 0, 1, \ldots, N-1$. $P(m/N, \boldsymbol{\alpha}^\star)$ can be calculated from $\omega_0^\star, \omega_1^\star, \ldots, \omega_K^\star$ by the barycentric formula (Step. 2).

3.4.4 Window estimation from the window derivative

In the previous subsections, two window derivatives were introduced as solutions to (3.23). This subsection presents the method for estimating the window from the obtained window derivative.

Denoting $\mathbf{g}_0 = \mathbf{F}^{-1}\,\mathrm{diag}(\mathbf{b})\mathbf{F}\mathbf{P}_{L,N}\mathbf{z}$, (3.21) is rewritten as

$$\mathbf{g} = \mathbf{g}_0 + c\mathbf{1}_L. \tag{3.63}$$

Recall that c is the integral constant, which has to be determined for estimating the window $\mathbf{g}$. In our conference paper [33], estimating c is formulated as the minimization problem of the ratio of the energy outside $[-W, W]$ to the total energy, similar to the Slepian window,

$$\underset{c}{\text{minimize}} \quad 1 - \frac{(\mathbf{g}_0 + c\mathbf{1}_L)^{\mathrm{T}}\,\mathbf{S}_N\,(\mathbf{g}_0 + c\mathbf{1}_L)}{(\mathbf{g}_0 + c\mathbf{1}_L)^{\mathrm{T}}\,(\mathbf{g}_0 + c\mathbf{1}_L)}, \tag{3.64}$$

to obtain the well-localized frequency response.

An example of the window obtained by the spectral integration from a window derivative supported on $[0, N-1]$ ($N = 7$) is illustrated in Fig. 3.8. As with the spectral differentiation, even if window derivatives are supported on $[0, N-1]$ ($N < L$), the integrated windows are not supported on $[0, N-1]$. The integrated windows are truncated to have the same support as the window derivatives in practice. This truncation may change the frequency response of the window and decrease the accuracy of the IF compuation. In this chapter, we propose a method for estimating c based on minimizing the truncation effect.

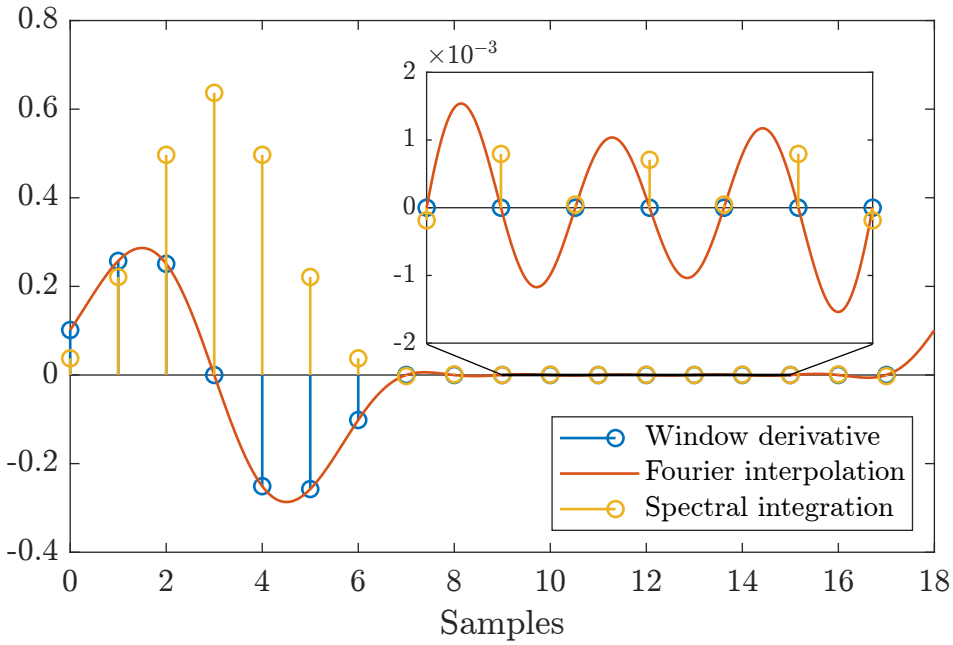

Figure 3.8: Spectral integration of a window derivative supported on $[0, N-1]$ ($N = 7$).

The truncated window outside $[0, N-1]$ can be expressed using the zero-padding matrix $\mathbf{P}_{L,N}$ as

$$\mathbf{P}_{L,N}\mathbf{P}_{L,N}^{\mathrm{T}}(\mathbf{g}_0 + c\mathbf{1}_L). \tag{3.65}$$

We estimate c by minimizing the squared error of the window before and after truncation, i.e., our proposed estimation problem is formulated as

$$\underset{c}{\text{minimize}} \quad \frac{1}{2}\left\|(\mathbf{I}_L - \mathbf{P}_{L,N}\mathbf{P}_{L,N}^{\mathrm{T}})(\mathbf{g}_0 + c\mathbf{1}_L)\right\|_2^2. \tag{3.66}$$

$\mathbf{I}_L - \mathbf{P}_{L,N}\mathbf{P}_{L,N}^{\mathrm{T}}$ can be interpreted as a projection onto the set of vectors that take a zero value in $[0, N-1]$. Note that the cost function of (3.66) does not go to zero since $\mathbf{g}_0[l]$ outside $[0, N-1]$ oscillates similarly to the spectral differentiation in Fig. 3.1. This problem is a linear least squares problem, so the solution $c^\star$ to (3.66) satisfies

$$\mathbf{1}_L^{\mathrm{T}}(\mathbf{I}_L - \mathbf{P}_{L,N}\mathbf{P}_{L,N}^{\mathrm{T}})^{\mathrm{T}}(\mathbf{I}_L - \mathbf{P}_{L,N}\mathbf{P}_{L,N}^{\mathrm{T}})(\mathbf{g}_0 + c^\star\mathbf{1}_L) = 0.$$

Because $\mathbf{I}_L - \mathbf{P}_{L,N}\mathbf{P}_{L,N}^{\mathrm{T}}$ is a diagonal projection matrix and $\mathbf{1}_L^{\mathrm{T}}\left(\mathbf{I}_L - \mathbf{P}_{L,N}\mathbf{P}_{L,N}^{\mathrm{T}}\right)\mathbf{1}_L = L - N$, the solution $c^\star$ to (3.66) is given by

$$c^\star = -\frac{1}{L-N}\mathbf{1}_L^{\mathrm{T}}(\mathbf{I}_L - \mathbf{P}_{L,N}\mathbf{P}_{L,N}^{\mathrm{T}})\mathbf{g}_0. \tag{3.67}$$

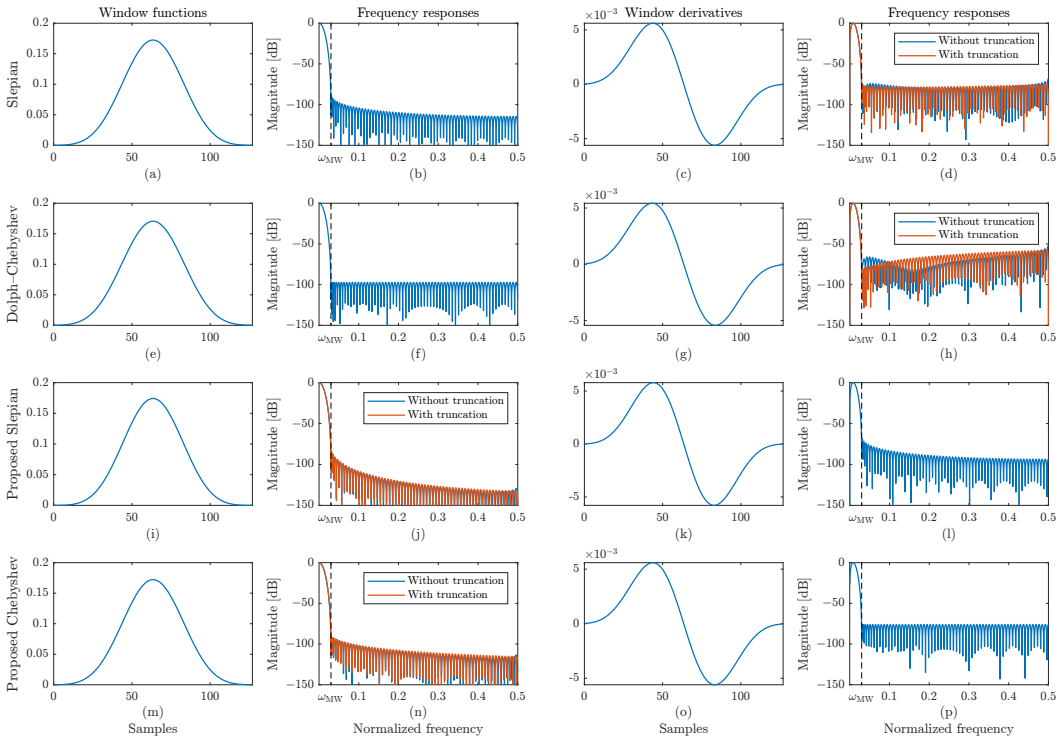

Figure 3.9: Results of designed windows. Each column shows (from left to right) the window functions, their frequency responses, the window derivatives, and the frequency responses of the window derivatives. Each row represents the results of a type of window function. The red lines show the frequency responses when the conventional window derivatives or the proposed windows are truncated outside $[0, N-1]$. The broken black lines indicate the bandwidths ω_{MW}.

Since this formulation minimizes the energy of the outer components, the proposed integration method is applicable for the case $N < L$. The proposed integration method is effective when the window width is limited, as in real-time processing, because the effect of truncation can be significant.

3.5 Frequency responses of designed windows

This section compares the proposed windows with the Slepian and Dolph–Chebyshev windows. Hereafter, the windows calculated from the Slepian window derivative and the Chebyshev window derivative are referred to as the proposed Slepian window and the proposed Chebyshev window, respectively.

First, Fig. 3.9 illustrates the shapes and frequency responses of the windows

and the window derivatives. The length of the windows was set to $N = 2^7$, and the signal length for the spectral differentiation/integration was set to $L = 2^{12}$. Each parameter W for the window was chosen such that $\omega_{\mathrm{MV}} = 0.03$. Note that the frequency responses in Figs. 3.9 were normalized such that the maxima are 0 dB. The derivatives of the Slepian and Dolph–Chebyshev windows have nonzero values outside $[0, N-1]$, as in Figs. 3.1, which are omitted in Figs. 3.9(c) and (g) since they are too small for illustration. The red lines in Fig. 3.9(d) and (h) show the frequency responses when these outside values are truncated. Conversely, the proposed Slepian and Chebyshev windows have nonzero values outside $[0, N-1]$ due to the spectral integration. The frequency responses of the truncated windows are plotted as the red lines in Fig. 3.9(j) and (n).

In the time domain, there is little difference between the shapes of the four windows. The derivative of the Dolph–Chebyshev window (Fig. 3.9(g)) has a slight oscillation at both ends compared with the others. According to Figs. 3.9(b) and (l), the Slepian window derivative has a similar sidelobe decay to the Slepian window. Likewise, the sidelobe decay of the Chebyshev window derivative in Fig. 3.9(f) resembles that of the Dolph–Chebyshev window in Fig. 3.9(p). Additionally, the proposed Slepian and Chebyshev windows have better sidelobe decays than the Slepian and Dolph–Chebyshev windows, respectively, although their highest sidelobe levels are slightly higher. This is because estimating a window from its derivative window in (3.21) suppresses the high frequencies. Furthermore, Fig. 3.9(d), (h), (j), and (n) show that the proposed Slepian and Chebyshev windows have smaller effects on the frequency responses caused by truncation compared with the Slepian and Dolph–Chebyshev windows. The result indicates that the proposed design method can reduce the effect of truncation.

Then, the sidelobe energy and the highest sidelobe level of each window at various bandwidths ω_{MW} are summarized in Fig. 3.10. Fig. 3.10 also shows the SE and HSL of four well-known windows: Hann, Blackman, Nuttall, and truncated Gaussian windows. The truncated Gaussian window is represented by

$$\mathbf{w}_\sigma[n] = \mathrm{e}^{-\frac{\{n-(L-1)/2\}^2}{2\sigma^2}}, \tag{3.68}$$

where σ is a parameter to control the mainlobe ω_{MW}. Although they were originally defined as continuous functions, we computed these derivatives by the spectral differentiation as mentioned at the end of Sec. 3.2. The jumps in the mainlobe

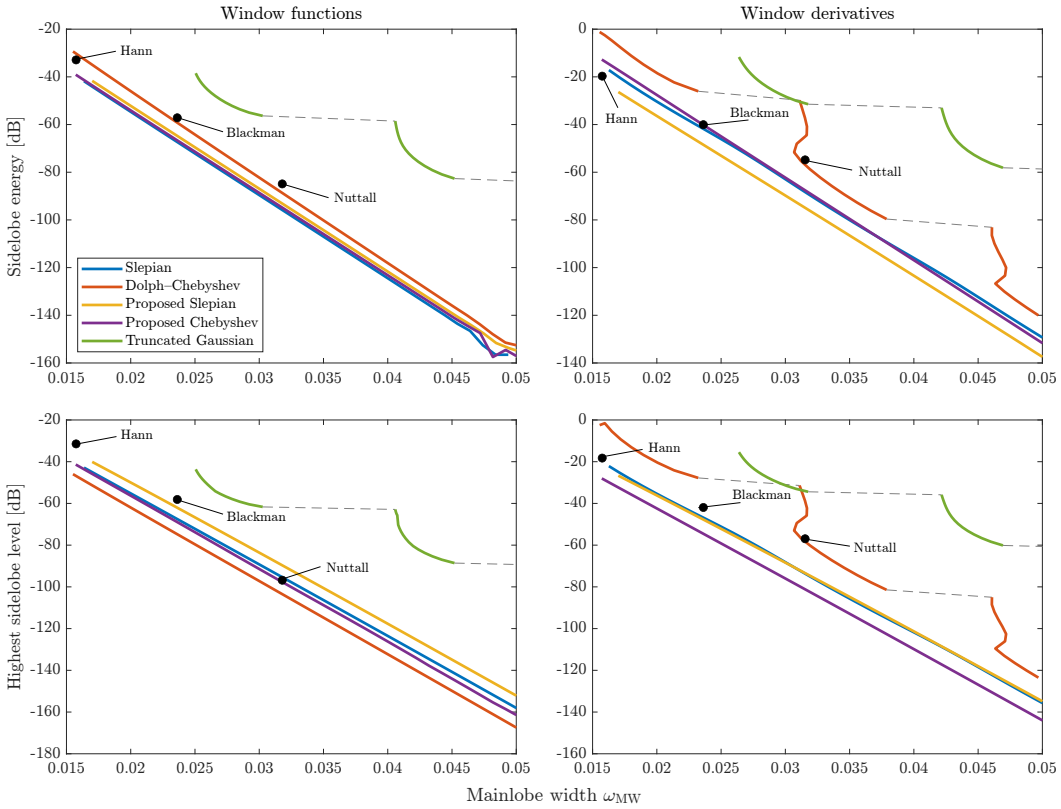

Figure 3.10: Sidelobe energy and highest sidelobe level of the designed windows. The top and bottom rows show the sidelobe energy and highest sidelobe level of the window functions, respectively. The left and right columns correspond to the results of the window functions and the window derivatives. Each line color represents a different type of window.

widths seen in the truncated Gaussian and Dolph–Chebyshev are due to nonzero minima in the frequency responses, as shown on the right side of Fig. 3.11.

When the bandwidth W was set to a higher value, the SE and HSL of the Slepian, Dolph–Chebyshev, proposed Slepian and proposed Chebyshev windows decreased. Fig. 3.10 confirms that the Slepian window derivative and the Chebyshev window derivative were correctly designed with the desired optimality.

3.6 Evaluation of IF computation

This section compares the four windows in terms of the IF computation. Throughout this section, the window length $N = 2^7$ and the signal length $L = 2^{12}$. Each parameter W for the window was chosen so that $\omega_{\mathrm{MV}} = 0.03$.

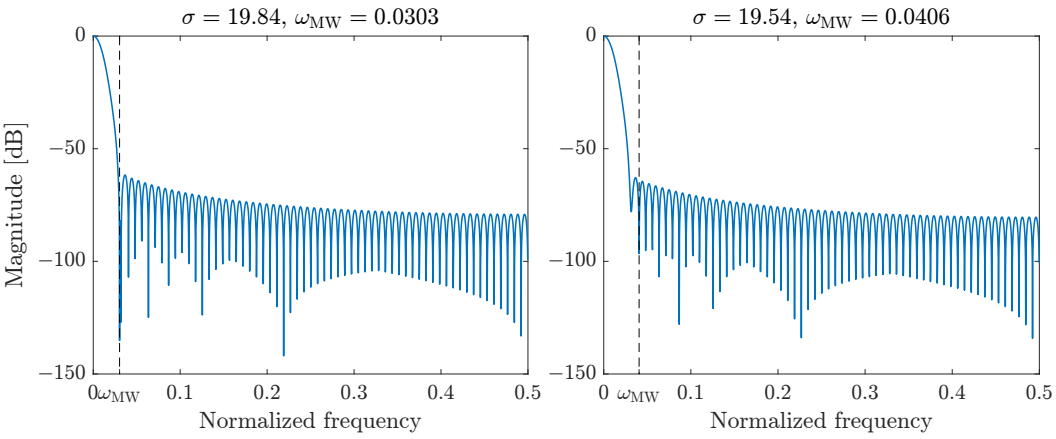

Figure 3.11: Truncated Gaussian windows with different σ.

3.6.1 Comparing window integration methods

First, the proposed integration method in (3.66) was compared with the conventional estimation method (3.64) [33]. They were assessed based on the error in computing the IF of the sinusoid in (3.12).

The frequency responses of the estimated windows from the Slepian and Chebyshev window derivatives using the two integration methods are shown in Fig. 3.12. The SE and HSL of each window are shown in the upper right corner of each subfigure. Although the conventional method determines c by minimizing the energy outside of $[-W, W]$ to reduce the SE, the frequency responses of the windows estimated by the two methods have comparable SE. However, for both window derivatives, the changes due to truncation of the windows estimated by the proposed method were smaller than those of the conventional method.

The errors of the IF using two integrated windows are shown in Fig. 3.13. The errors using the conventional method increase due to truncation, but this is not seen in the results of the proposed method. This suggests that the proposed method can estimate the window to avoid an increase in the error due to truncation effects. Additionally, with or without truncation, the errors using the proposed method are smaller than those using the conventional method in $[0, \omega_{\mathrm{MW}}]$. This may be because minimizing the outer energy reduces the oscillation of the frequency responses.

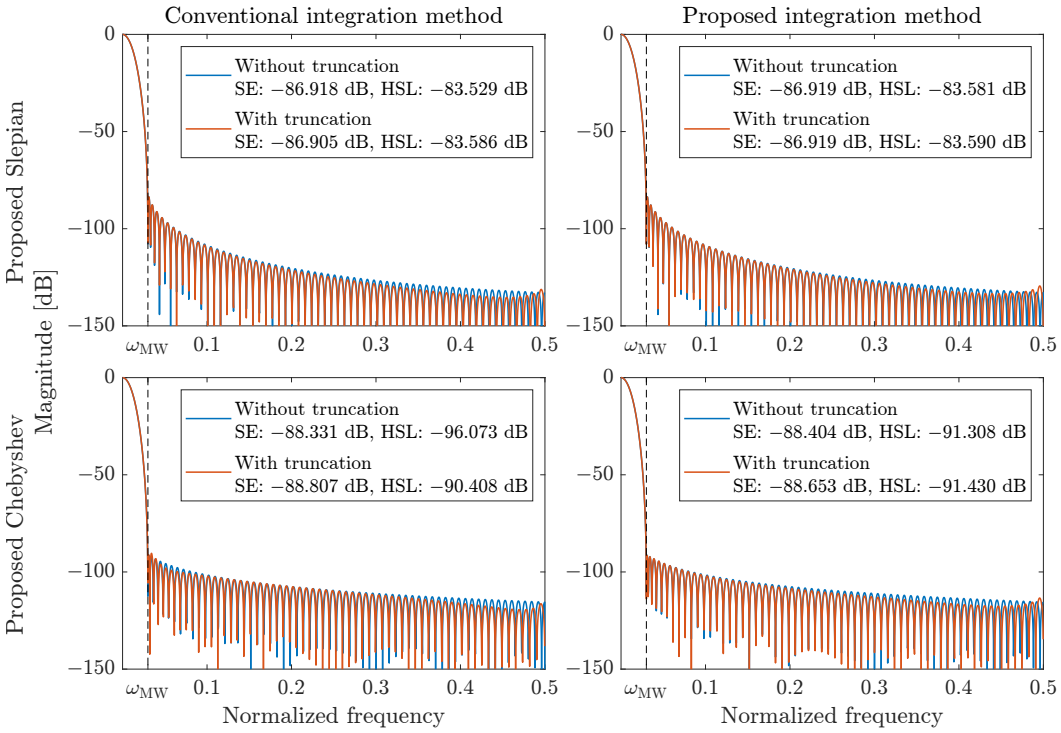

Figure 3.12: Frequency responses of estimated windows from Slepian and Chebyshev window derivatives using two integration methods. The left and right columns correspond to the frequency responses of the conventional and proposed methods, respectively.

3.6.2 Comparison of windows for IF computation of a sinusoid

Second, the four windows were compared in terms of computing the IF of a sinusoid. They were also evaluated by the error in (3.13).

The IF errors using the four windows are shown in Fig. 3.14. Regardless of truncation, the errors using the proposed windows are smaller than those using the conventional windows. In particular, the error using the proposed Chebyshev window was smaller than that using the Slepian window, shown in Fig. 3.14, even though they have similar frequency responses, shown in Fig. 3.9. The results indicate that the proposed window design method can improve the accuracy of the IF computation of the sinusoid.

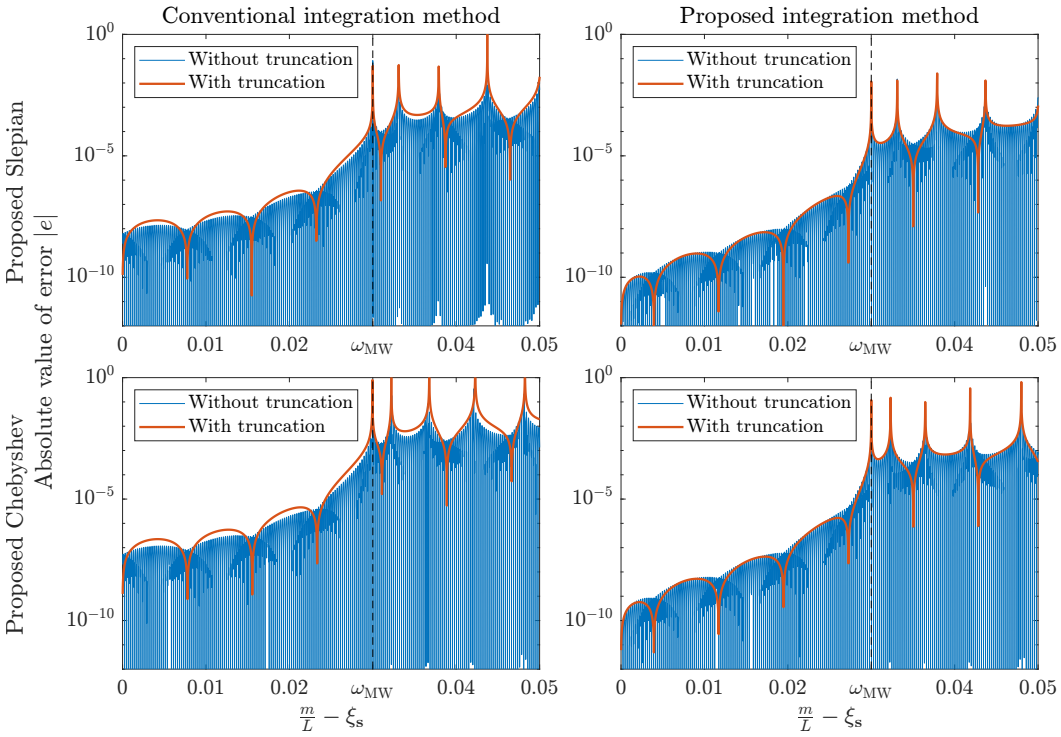

Figure 3.13: IF computation error in a sinusoid using the windows shown in Fig. 3.12. The left and right columns correspond to the results of the conventional and proposed integration methods, respectively.

3.6.3 Comparison of windows for IF computation in the presence of another sinusoid

Then, we consider estimating the IF of the sinusoid $\mathbf{s}$ from a signal composed of two complex sinusoids,

$$\mathbf{x} = \mathbf{s} + \boldsymbol{i}, \tag{3.69}$$

where

$$\boldsymbol{i}[l] = A_{\boldsymbol{i}} \mathrm{e}^{\mathrm{i}(2\pi\xi_{\boldsymbol{i}} l + \phi_{\boldsymbol{i}})}, \tag{3.70}$$

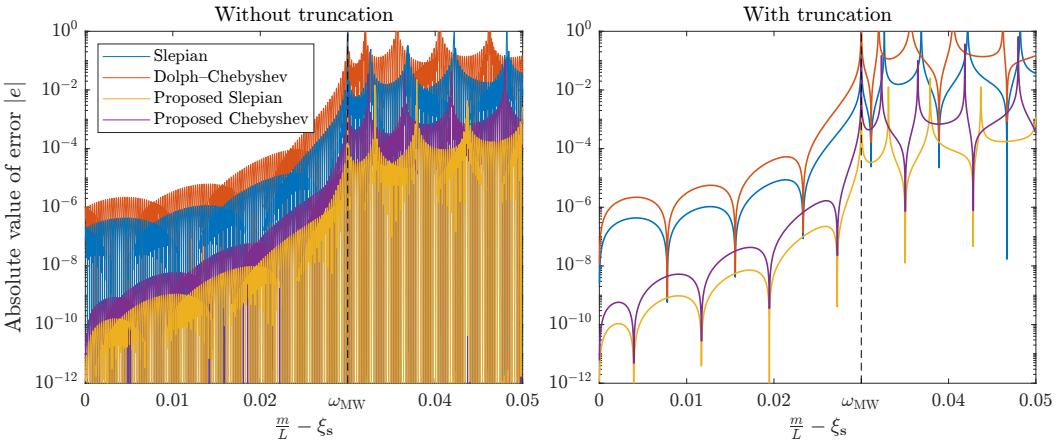

Figure 3.14: IF computation error in a sinusoid using the four windows. The left and right columns correspond to the results without and with truncation, respectively.

for $l = 0, 1, \ldots, L-1$. $\mathbf{s}$ and $\boldsymbol{i}$ correspond to the target and interference signals, respectively. The IF of $\mathbf{x}$ is expressed as

$$\begin{aligned}
&\mathrm{IF}_{\mathbf{g}}^{\mathbf{x}}[n,m] \\
&= -\frac{1}{2\pi}\Im\left\{\frac{(\mathbf{V}_{\mathbf{g}'}\mathbf{x})[n,m]}{(\mathbf{V}_{\mathbf{g}}\mathbf{x})[n,m]}\right\} \\
&= -\frac{1}{2\pi}\Im\left\{\frac{re^{\mathrm{i}\Delta\phi}(\mathcal{F}\mathbf{g}')(\frac{m}{L}-\xi_{\mathbf{s}}) + (\mathcal{F}\mathbf{g}')(\frac{m}{L}-\xi_{\boldsymbol{i}})}{re^{\mathrm{i}\Delta\phi}(\mathcal{F}\mathbf{g})(\frac{m}{L}-\xi_{\mathbf{s}}) + (\mathcal{F}\mathbf{g})(\frac{m}{L}-\xi_{\boldsymbol{i}})}\right\},
\end{aligned} \tag{3.71}$$

where

$$r = \frac{A_{\mathbf{s}}}{A_{\boldsymbol{i}}}, \qquad \Delta\phi = 2\pi\xi_{\mathbf{s}}n + \phi_{\mathbf{s}} - (2\pi\xi_{\boldsymbol{i}}n + \phi_{\boldsymbol{i}}).$$

The error of the IF between $\mathbf{s}$ and $\mathbf{x}$ is given by

$$\begin{aligned}
e &= \mathrm{IF}_{\mathbf{g}}^{\mathbf{x}}[n,m] - \mathrm{IF}_{\mathbf{g}}^{\mathbf{s}}[n,m] \\
&= -\frac{1}{2\pi}\Im\left\{\frac{\frac{(\mathcal{F}\mathbf{g}')(\frac{m}{L}-\xi_{\boldsymbol{i}})}{(\mathcal{F}\mathbf{g})(\frac{m}{L}-\xi_{\boldsymbol{i}})} - \frac{(\mathcal{F}\mathbf{g}')(\frac{m}{L}-\xi_{\mathbf{s}})}{(\mathcal{F}\mathbf{g})(\frac{m}{L}-\xi_{\mathbf{s}})}}{re^{\mathrm{i}\Delta\phi}\frac{(\mathcal{F}\mathbf{g})(\frac{m}{L}-\xi_{\mathbf{s}})}{(\mathcal{F}\mathbf{g})(\frac{m}{L}-\xi_{\boldsymbol{i}})} + 1}\right\}.
\end{aligned} \tag{3.72}$$

According to (2.25), the real part of the numerator of (3.72) is zero since the STFT magnitude of a complex sinusoid is time-invariant. Hence, (3.72) can be rewritten using $\mathrm{IF}_{\mathbf{g}}^{\boldsymbol{i}}$ and $\mathrm{IF}_{\mathbf{g}}^{\mathbf{s}}$ as

$$e = \Re\left\{\frac{\mathrm{IF}_{\mathbf{g}}^{\boldsymbol{i}}[n,m] - \mathrm{IF}_{\mathbf{g}}^{\mathbf{s}}[n,m]}{re^{\mathrm{i}\Delta\phi}\frac{(\mathcal{F}\mathbf{g})(\frac{m}{L}-\xi_{\mathbf{s}})}{(\mathcal{F}\mathbf{g})(\frac{m}{L}-\xi_{\boldsymbol{i}})} + 1}\right\}. \tag{3.73}$$

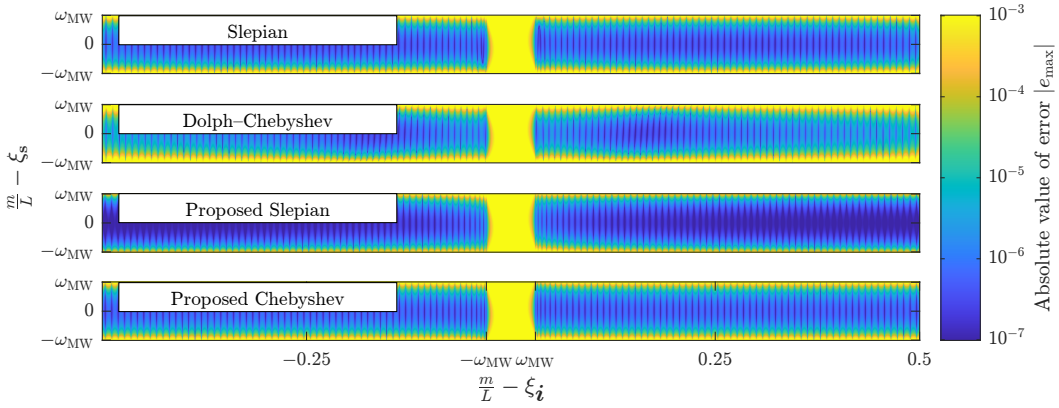

Figure 3.15: IF computation error in the sum of two sinusoids. Each row shows (from top to bottom) the results for the Slepian window, Dolph–Chebyshev window, proposed Slepian window, and proposed Chebyshev window.

The transitions of e associated with the initial phase $\phi_{\mathbf{s}}$, $\phi_{\boldsymbol{i}}$, and the time index n are irrelevant to the evaluation. Moreover, from (3.12), the IF of a complex sinusoid is constant regardless of the time index n. Therefore, we consider the worst case of e in any $\Delta\phi$. The error e becomes large regardless of sidelobe level when m/L is outside the mainlobe of the target signal or inside the mainlobe of the interference signal. Under these conditions, it is challenging to obtain a meaningful IF unless the amplitude ratio is quite large. Thus, we consider the case where m/L is inside the mainlobe of the target signal and outside the mainlobe of the interference signal. Assume that the amplitude ratio r satisfies $|(\mathcal{F}\mathbf{g})(m/L - \xi_{\mathbf{s}})/(\mathcal{F}\mathbf{g})(m/L - \xi_{\boldsymbol{i}}))| > 1/r$. This assumption is mild in the condition we considered because $|(\mathcal{F}\mathbf{g})(m/L - \xi_{\mathbf{s}})/(\mathcal{F}\mathbf{g})(m/L - \xi_{\boldsymbol{i}}))|$ should be large. Then, the worst case of e by choosing $\Delta\phi$ is given by

$$e_{\max} = \frac{\mathrm{IF}^{\boldsymbol{i}}_{\mathbf{g}}[n,m] - \mathrm{IF}^{\mathbf{s}}_{\mathbf{g}}[n,m]}{1 - r\left|\frac{(\mathcal{F}\mathbf{g})(\frac{m}{L} - \xi_{\mathbf{s}})}{(\mathcal{F}\mathbf{g})(\frac{m}{L} - \xi_{\boldsymbol{i}})}\right|}. \tag{3.74}$$

The deviation of (3.74) is shown in Appendix B. (3.74) indicates that as the amplitude ratio r decreases, the error of the IF increases.

Fig. 3.15 plots the error of the IF computation in (3.74) for $r = 1$. The bright central areas in the four results correspond to cases where the two mainlobes overlap. The top and bottom bright regions in the four results represent the errors of the IF outside the mainlobe of $\mathbf{s}$. In addition, the error at $m/L = \xi_{\mathbf{s}}$ in Fig. 3.15 is shown in Fig. 3.16.

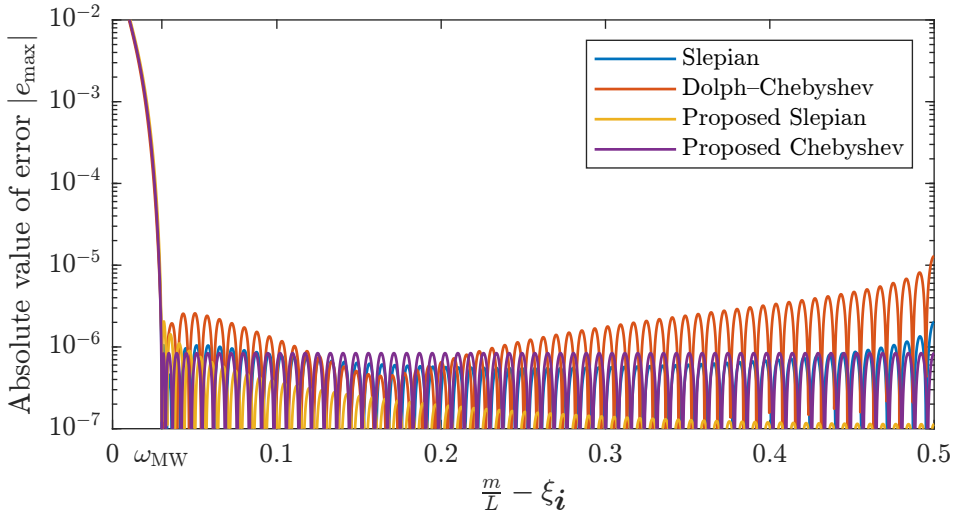

Figure 3.16: Error at $m/L = \xi_{\mathbf{s}}$ in Fig. 3.15. Each color represents the result of a different window.

The IF using the Dolph–Chebyshev window had the most significant error among the four windows. In contrast, the results for the proposed Slepian window had the smallest error. The results using the Slepian window and the proposed Chebyshev window had comparable errors. These results suggest that reducing the sidelobe of the window derivative decreases the error of the IF computation.

3.7 Application to FSST

As an application for IF computation, the proposed windows were applied to the FSSTs of an artificial signal and a speech signal. The FSST of a signal $\mathbf{f}$ with a window function $\mathbf{g}$ is defined as

$$\mathcal{S}^{\mathbf{f}}_{\mathbf{g}}[n,m] \coloneqq \sum_{k=0}^{L-1} \mathrm{e}^{\mathrm{i}\frac{2\pi kn}{L}} (\mathbf{V}_{\mathbf{g}}\mathbf{f})[n,k]\delta\left[m - \tilde{\mathbf{m}}[n,k]\right], \tag{3.75}$$

where

$$\tilde{\mathbf{m}}[n,m] \coloneqq \left\lfloor m + L \cdot \mathrm{IF}^{\mathbf{f}}_{\mathbf{g}}[n,m] \right\rceil, \tag{3.76}$$

and $\delta[l]$ is the Kronecker delta. We used the Rényi entropy as a metric of the energy concentration of the FSST spectrogram [93]. The Rényi entropy of the FSST spectrogram is given by

$$H_\alpha \coloneqq \frac{1}{1-\alpha} \log_2 \left(\sum_{n,m} \left(\frac{\left|\mathcal{S}^{\mathbf{f}}_{\mathbf{g}}[n,m]\right|}{\sum_{n,m}\left|\mathcal{S}^{\mathbf{f}}_{\mathbf{g}}[n,m]\right|} \right)^{\alpha} \right) - \log_2(L),$$

with $\alpha > 2$ being recommended for the T-F domain measures [93]. $\alpha = 3$ is chosen throughout the experiments.

As shown in Fig. 3.10, sidelobes can be reduced by increasing the mainlobe width. In contrast, if the mainlobes of multiple components overlap, the IF error becomes large. Therefore, an appropriate mainlobe width must be selected to reduce both effects. The Rényi entropy has also been used as a metric to select the variance of the Gaussian window in STFT and FSST [51, 94]. Since the mainlobe widths of the designed windows will also affect the Rényi entropy of the FSST, we evaluated the Rényi entropy of the FSST for various mainlobe widths.

To further evaluate the FSST of an artificial signal, we used the Earth mover's distance [95]. The Earth mover's distance is an index that measures between two distributions and has been used to evaluate the synchrosqueezing-based method [21, 22, 96]. This evaluation consists of averaging the 1D Earth mover's distance between the estimated T-F representation and the ideal representation at each time index n. If the ideal representation is known, the Earth mover's distance is a more valid metric for estimating the T-F representation than the Rényi entropy.

3.7.1 FSST of artificial signals

The proposed windows were evaluated with the FSST of a real-valued artificial signal, which contained a sinusoid, a linear chirp, and a quadratic chirp. The window length was set to $N = 2^6$ because the calculation of the Rényi entropy of FSST was numerically unstable when $N = 2^7$.

The Rényi entropies of the FSST with different mainlobe widths ω_{MW} are plotted in Fig. 3.17. Table 3.1 shows the minimum values of the Rényi entropies in Fig. 3.17 and the corresponding bandwidths. Compared to the other four windows, the widely used truncated Gaussian window shows poor performance. The Slepian window and the proposed Chebyshev window show approximately equivalent performance. The proposed Slepian window achieves the best performance among the five windows. Then, the Earth mover's distance of FSST with different mainlobe widths ω_{MW} are shown in Fig. 3.18, and their minimum values are shown in Table 3.2. Although the mainlobe widths that obtain the minima are different, the FSST spectrogram using the proposed Slepian has the smallest distance. These results indicate that the proposed Slepian window provides the most concentrated FSST spectrogram of the five windows.

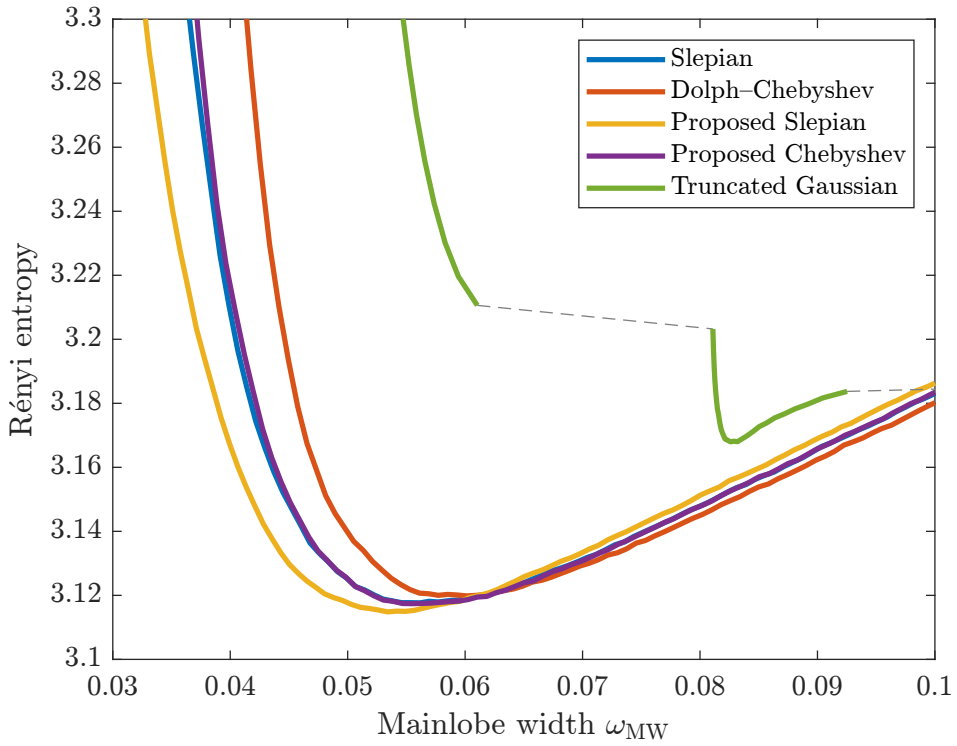

Figure 3.17: Rényi entropies of FSST spectrograms of a synthesized signal.

Fig. 3.19 illustrates the T-F representations obtained by the STFT and FSST in the case of Table 3.1. Focusing on the enlargement in the red box (the right column of Fig. 3.19), it can be confirmed that the component outside the center frequency is reduced using the proposed Slepian window. The FSST spectrogram using the proposed Chebyshev window is almost as sharp as that using the Slepian window.

Table 3.1: Minimum values of the Rényi entropies in Fig. 3.17 and the corresponding bandwidths.

Window function	Mainlobe width	Rényi entropy
Slepian	0.0562	3.1177
Dolph–Chebyshev	0.0601	3.1198
Proposed Slepian	0.0534	3.1148
Proposed Chebyshev	0.0554	3.1174
Truncated Gaussian	0.0826	3.1680

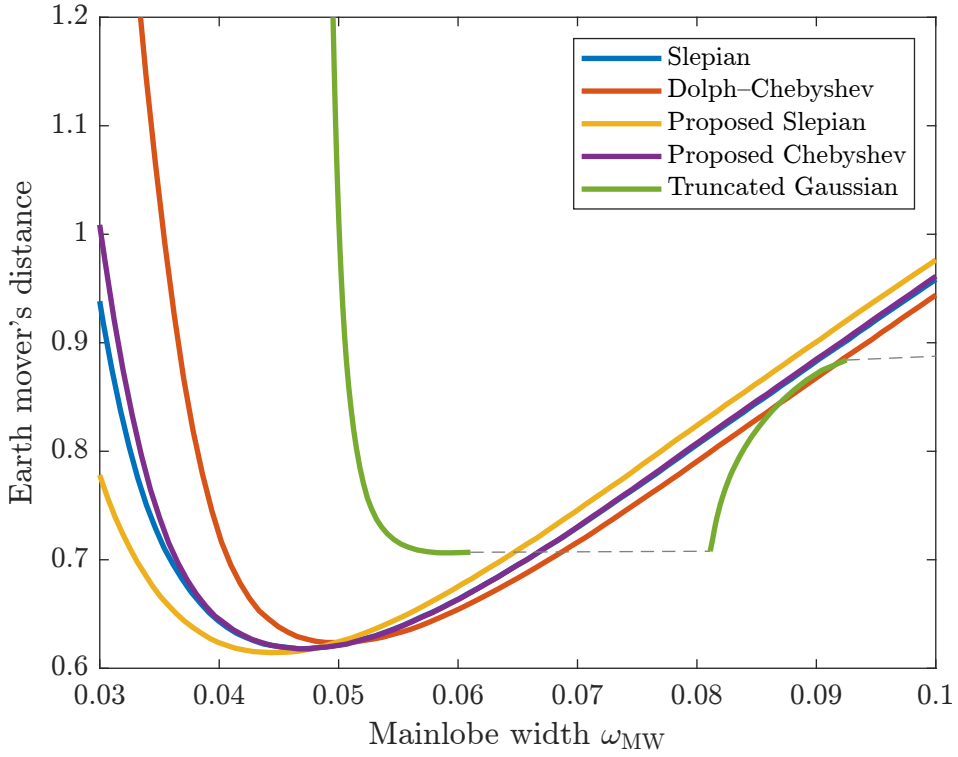

Figure 3.18: Earth mover's distance of FSST spectrograms of a synthesized signal.

Table 3.2: Minimum values of the Earth mover's distance in Fig. 3.18 and the corresponding bandwidths.

Window function	Mainlobe width	Earth mover's distance
Slepian	0.0468	0.6178
Dolph–Chebyshev	0.0505	0.6235
Proposed Slepian	0.0443	0.6144
Proposed Chebyshev	0.0467	0.6178
Truncated Gaussian	0.0594	0.7064

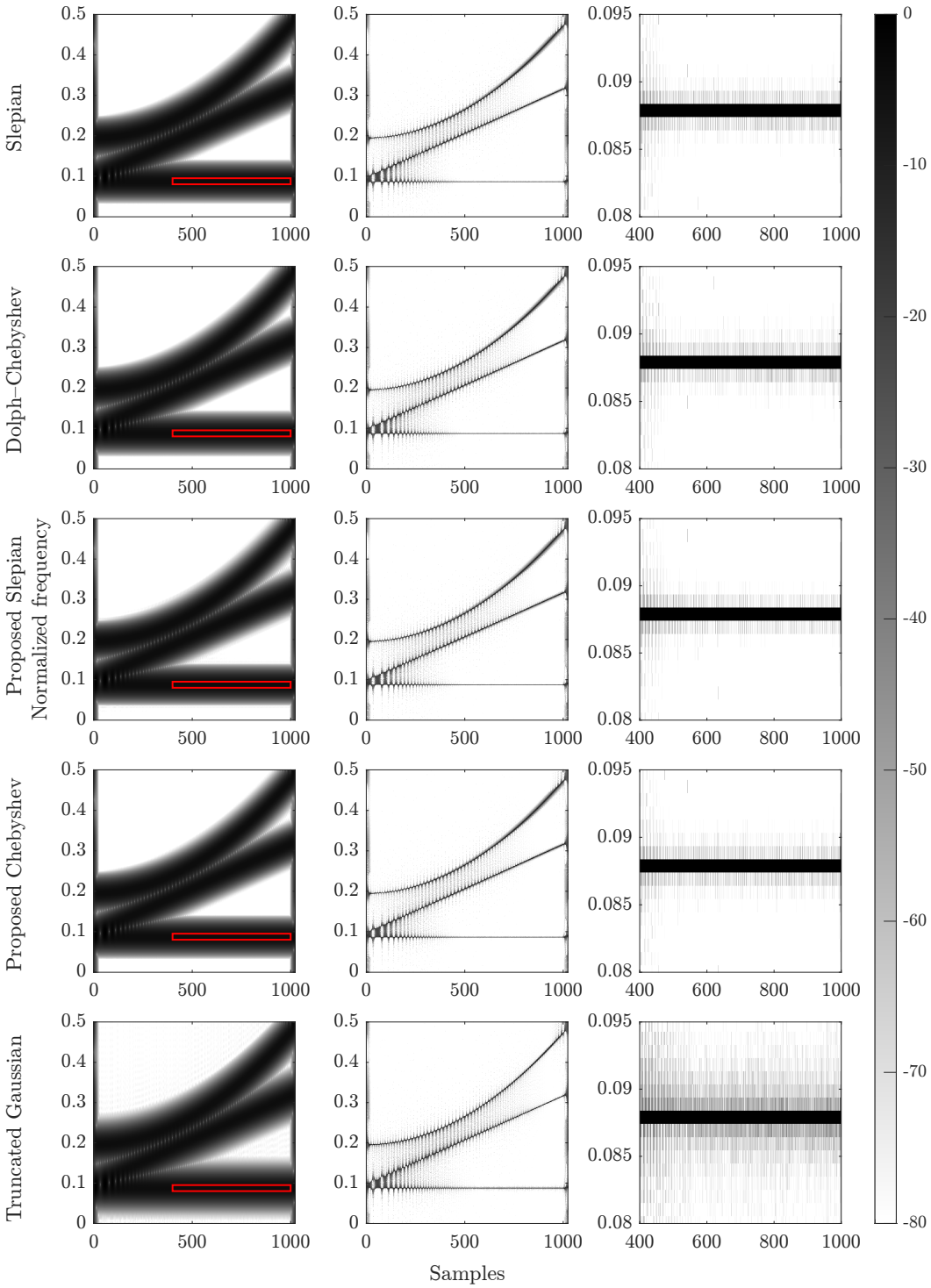

Figure 3.19: FSST spectrograms of an artificial signal. Each column shows (from left to right) the spectrograms, FSST spectrograms, and enlargements of the FSST spectrograms in the red box. Each row shows (from top to bottom) the results for the Slepian, Dolph–Chebyshev, proposed Slepian, proposed Chebyshev, and truncated Gaussian windows.

3.7.2 FSST of speech signal

We considered a speech signal with a sampling frequency of 7,418 Hz. The window length was set to $N = 2^7$. Since the ideal T-F representation is unknown, the Earth mover's distance cannot be used for evaluation, and only the Rényi entropy is used in this experiment.

Fig. 3.20 and Table 3.3 show Rényi entropies of SST with different bandwidths and their minimum values, respectively. As in the case of the synthesized signal, FSST with the proposed Slepian window achieves the best performance in terms of the Rényi entropy. Fig. 3.21 depicts the T-F representations obtained by the STFT and FSST in the case of Table 3.3. Although the appearances of the spectrograms are similar, the proposed Slepian window provides the sharpest FSST spectrogram of the four windows. The results confirmed that the proposed Slepian window provides the sharpest FSST spectrogram of the four windows for a real speech signal as well.

The proposed Slepian window showed the best performance throughout the experiments in this section. Therefore, the proposed Slepian window can be useful for FSST applications. However, the proposed Chebyshev window may be preferable depending on the signal of interest because the sidelobes of the Chebyshev window derivative near the mainlobe are smaller than those of the Slepian window derivative.

Table 3.3: Minimum values of the Rényi entropies in Fig. 3.20 and the corresponding bandwidths.

Window function	Mainlobe width	Rényi entropy
Slepian	0.0215	5.0835
Dolph–Chebyshev	0.0239	5.1175
Proposed Slepian	0.0208	5.0794
Proposed Chebyshev	0.0221	5.0899
Truncated Gaussian	0.0284	5.2594

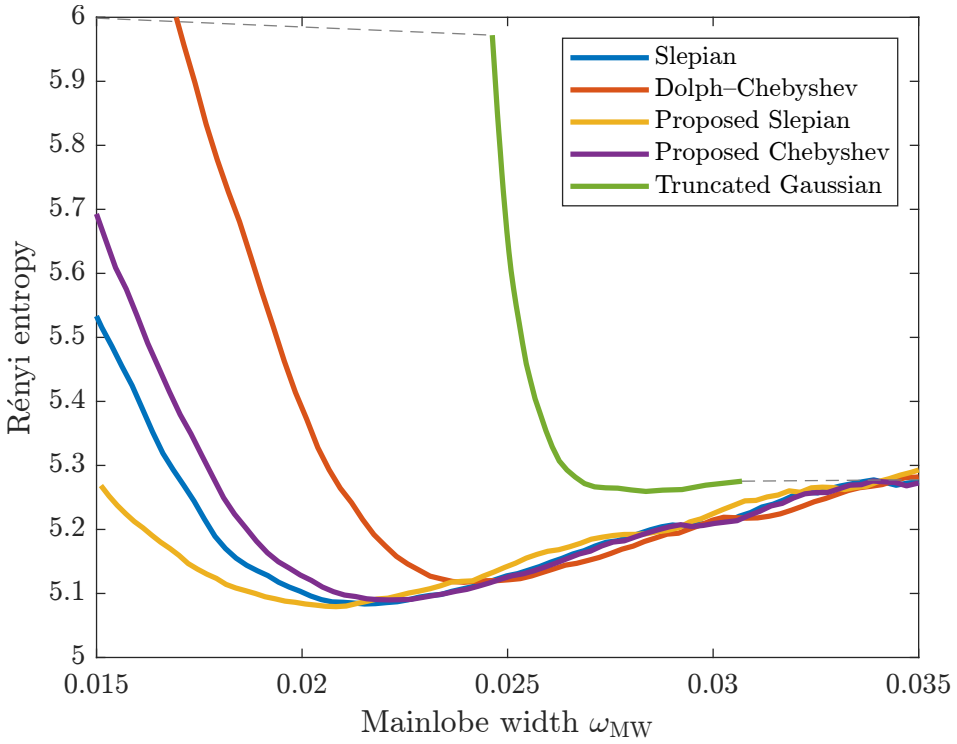

Figure 3.20: Rényi entropies of FSST spectrograms of a speech signal.

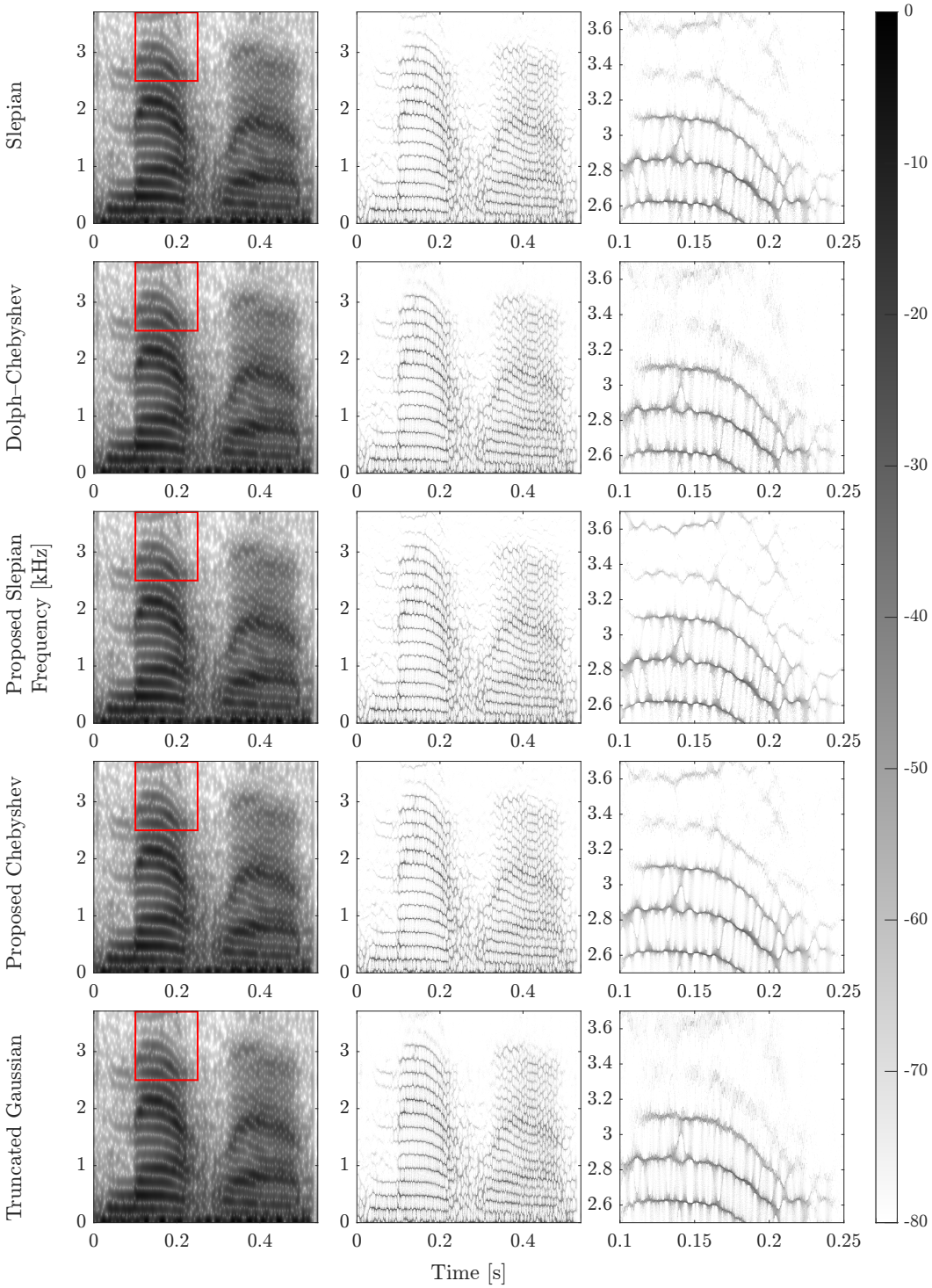

Figure 3.21: FSST spectrograms of a speech signal. Each column shows (from left to right) the spectrograms, FSST spectrograms, and enlargements of the FSST spectrograms in the red box. Each row shows (from top to bottom) the results for the Slepian, Dolph–Chebyshev, proposed Slepian, proposed Chebyshev, and truncated Gaussian windows.

3.8 Conclusion

In this chapter, we proposed a framework to design a window function for the IF computation. The proposed method first designs the window derivative to minimize the sidelobes and then estimates the window function from the designed window derivative. We designed two minimum-sidelobe derivatives to minimize the SE and HSL, referred to as the Slepian and Chebyshev window derivatives, respectively. The integral constant that appears when integrating the designed window derivative is estimated to minimize the truncation effect. In the IF computation of a sinusoid, the proposed windows reduce the error compared to certain other windows. In addition, the proposed Slepian window showed the best performance in the FSST among the several windows.

Some phase-aware techniques use not only the IF but also the group delay and higher-order derivatives of the phase [20–22]. Moreover, phase derivatives also play an important role in wavelet synchrosqueezing transform and other filterbank-based methods. Therefore, future work will include generalizing the proposed method to these elements and investigating the method's computational efficiency.

Chapter 4

Sparse time-frequency representation via atomic norm minimization

4.1 Introduction

In this chapter, an estimation method of a sparse T-F representation using atomic norm minimization is proposed. The T-F representation obtained by the STFT or the DGT is spread due to the windowing of the analyzed signal. This spread may affect the performance of T-F domain analysis and processing.

To achieve a well-localized T-F representation, many approaches have been proposed [5, 7, 14, 18, 25–32]. Reassignment and synchrosqueezing methods aim to relocate the spread components into the original positions using phase derivative information [5, 7, 14, 18]. Their performances are affected by the mixing of components due to windowing [33]. Sparsity-aware methods are powerful tools that are robust against such mixing of components and noises [25–32]. Sparsity-aware methods aim to find a sparse solution of an underdetermined system. However, the typical formulation based on ℓ_1-norm minimization involves discretizing a continuous parameter onto a grid. It may degrade the performance due to a model mismatch between the signal and the predefined grid [97].

Recently, sparse optimization using the atomic norm has been studied [98–101] and applied to many applications such as line spectrum estimation [102, 103], direction of arrival estimation [104–106], and target localization in radar [107].

The atomic norm does not require the discretization of continuous parameters. Thus, introducing the atomic norm into a sparse T-F representation should obtain a better-localized T-F representation.

In this chapter, we propose an estimation method of a sparse T-F representation. In the proposed method, the estimation problem is formulated as atomic norm minimization under the condition that the analyzed time-domain signal can be reconstructed. Numerical experiments confirm that the proposed method provides a sparser T-F representation than the conventional methods.

4.2 Line spectrum estimation using atomic norm

$\mathcal{G}(\mathbf{g}, a, M)$ in (2.30) contains windowed sinusoids whose frequency is discretized onto the grid $\{m/M\}_{m=0,\dots,M-1}$. Eq. (2.37) may provide a poor result when the signal $\mathbf{f}$ has a component whose frequency is not included in the grid. To avoid the effects of grid mismatch, a method using the atomic norm has been studied as a gridless sparse optimization method [98–107]. Here, line spectrum estimation is applied only to the nth windowed signal. A windowed signal at time index n is denoted as $\mathbf{f}_n = \mathbf{W}_n\mathbf{f}$, where $\mathbf{W}_n \in \mathbb{R}^{L\times L}$ is a diagonal matrix whose diagonal elements are given by $\mathbf{W}_n[l,l] = \mathbf{g}[l-an]$. We assume that the nth windowed signal $\mathbf{f}_n$ can be expressed as a sum of complex sinusoids,

$$\mathbf{f}_n = \mathbf{W}_n \sum_k c_{n,k}\mathbf{a}_{n,k}, \qquad \mathbf{a}_{n,k} \in \mathcal{A}, \tag{4.1}$$

where $\mathcal{A}$ is a collection of complex sinusoids

$$\mathcal{A} = \left\{ \mathbf{a} \in \mathbb{C}^L \,\middle|\, \mathbf{a}[l] = \mathrm{e}^{\mathrm{i}2\pi\omega l}, \omega \in [0,1) \right\}. \tag{4.2}$$

The atomic norm is used to express the nth windowed signal $\mathbf{f}_n$ with a few coefficient $c_{n,k}$. Let us denote $\mathbf{x}_n = \sum_k c_{n,k}\mathbf{a}_{n,k}$, then the atomic norm of $\mathbf{x}_n$ associated with a set of atoms $\mathcal{A}$ is given by [99]

$$\begin{aligned} \|\mathbf{x}_n\|_{\mathcal{A}} &= \inf\left\{ \nu_n \geq 0 \,|\, \mathbf{x}_n \in \nu_n \operatorname{conv}(\mathcal{A}) \right\}, \\ &= \inf\left\{ \sum_k |c_{n,k}| \,\middle|\, \mathbf{x}_n = \sum_k c_{n,k}\mathbf{a}_{n,k},\ \ \mathbf{a}_{n,k} \in \mathcal{A} \right\}, \end{aligned} \tag{4.3}$$

where $\operatorname{conv}(\mathcal{A})$ is the convex hull of $\mathcal{A}$. It corresponds to the infimum of the ℓ_1-norm of coefficients when $\mathbf{x}_n$ is represented by a linear combination of elements in

$\mathcal{A}$. That is, the atomic norm can be interpreted as an extension of the ℓ_1-norm to the continuous parameter $\omega \in [0, 1)$. The atomic norm in Eq. (4.3) is characterized by the following optimization problem [100]:

$$\begin{aligned} \|\mathbf{x}_n\|_{\mathcal{A}} = \min_{\mathbf{u}_n, \nu_n} \quad & \frac{1}{2L} \operatorname{Tr}\left(T(\mathbf{u}_n)\right) + \frac{1}{2}\nu_n \\ \text{subject to} \quad & \begin{bmatrix} T(\mathbf{u}_n) & \mathbf{x}_n \\ \mathbf{x}_n^* & \nu_n \end{bmatrix} \succeq 0, \end{aligned} \tag{4.4}$$

where $T : \mathbb{C}^L \to \mathbb{C}^{L \times L}$ is the Hermitian Toeplitz operator:

$$T(\mathbf{u}) = \begin{bmatrix} \mathbf{u}[0] & \overline{\mathbf{u}[1]} & \cdots & \overline{\mathbf{u}[L-1]} \\ \mathbf{u}[1] & \mathbf{u}[0] & \ddots & \vdots \\ \vdots & \ddots & \ddots & \overline{\mathbf{u}[1]} \\ \mathbf{u}[L-1] & \cdots & \mathbf{u}[1] & \mathbf{u}[0] \end{bmatrix}. \tag{4.5}$$

If the Hermitian Toeplitz matrix $T(\mathbf{u}_n)$ is positive semidefinite and singular, it can be uniquely decomposed as [108]

$$T(\mathbf{u}_n) = \sum_{k=0}^{K-1} |c_{n,k}|\, \mathbf{a}_{n,k}\mathbf{a}_{n,k}^*, \qquad \mathbf{a}_{n,k} \in \mathcal{A}, \tag{4.6}$$

where K corresponds to the rank of $T(\mathbf{u}_n)$. $\mathbf{a}_{n,k}$ in Eq. (4.6) can be obtained by Prony's method [109], the matrix pencil method [110], or other linear prediction methods [111]. Then, coefficients $c_{n,k}$ can be obtained by solving the linear equation

$$[\mathbf{a}_{n,0}, \ldots, \mathbf{a}_{n,K-1}]\, \mathbf{c}_n = \mathbf{x}_n, \tag{4.7}$$

where $\mathbf{c}_n = [c_{n,0}, \ldots, c_{n,K-1}]^{\mathrm{T}}$. Therefore, estimating sinusoids in the windowed signal $\mathbf{f}_n$ using the atomic norm is formulated as

$$\underset{\mathbf{x}_n}{\text{minimize}} \quad \|\mathbf{x}_n\|_{\mathcal{A}} \quad \text{subject to} \quad \mathbf{f}_n = \mathbf{W}_n\mathbf{x}_n. \tag{4.8}$$

Substituting Eq. (4.4) into Eq. (4.8), Eq. (4.8) can be rewritten as the following semidefinite programming:

$$\begin{aligned} \underset{\mathbf{x}_n, \mathbf{u}_n, \nu_n}{\text{minimize}} \quad & \frac{1}{2L} \operatorname{Tr}(T(\mathbf{u}_n)) + \frac{1}{2}\nu_n \\ \text{subject to} \quad & \begin{bmatrix} T(\mathbf{u}_n) & \mathbf{x}_n \\ \mathbf{x}_n^* & \nu_n \end{bmatrix} \succeq 0, \quad \mathbf{f}_n = \mathbf{W}_n\mathbf{x}_n. \end{aligned} \tag{4.9}$$

While the eigenvalue decomposition of the $(L+1) \times (L+1)$ matrix needs to be iterated when solving Eq. (4.9), it can be reduced to the eigenvalue decomposition of a $(J+1) \times (J+1)$ matrix if the window function is supported on $[0, J-1]$.

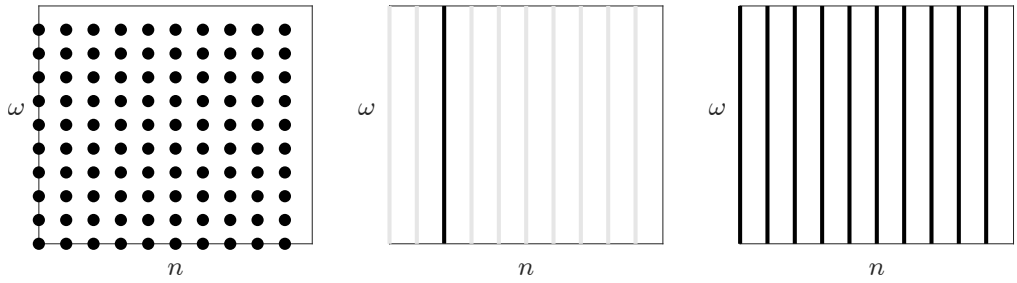

Figure 4.1: Estimating T-F representations using (a) ℓ_1-norm, (b) the line spectrum estimation by atomic norm, and (c) the proposed method.

4.3 Proposed method

The method of estimating the T-F representation based on the ℓ_1-norm suffers from degradation of performance due to discretization onto the grid, as shown in Fig. 4.1 (a). By contrast, by performing the line spectrum estimation using the atomic norm for each time index, a sparse T-F representation can be obtained without a grid in the frequency direction. However, Eq. (4.9) estimates sinusoids at each time index n independently as Fig. 4.1 (b), which does not efficiently take advantage of the sparsity of the T-F representation (see Fig. 4.2).

In this chapter, we propose an estimation method for a sparse T-F representation using the atomic norm to avoid the effect of grid mismatch. In our formulation, the atomic norm is minimized under the constraint of the perfect reconstruction of the entire signal. It can be interpreted as an extension of Eq. (2.37) with an infinite number of frequency channels, shown in Fig. 4.1 (c).

4.3.1 Sparse T-F representation using atomic norm

Reconstruction of the nth windowed signal $\mathbf{f}_n = \mathbf{W}_n\mathbf{f}_n$ is considered in Eq. (4.9). In contrast, reconstructing the entire signal $\mathbf{f}$ by windowing and summing $\mathbf{x}_n$ is considered in the proposed method, written as

$$\mathbf{A}_\mathbf{g}\mathbf{x} = \sum_{n=0}^{N-1} \mathbf{W}_n\mathbf{x}_n, \tag{4.10}$$

where $\mathbf{x} = \left[\mathbf{x}_0^{\mathrm{T}}, \mathbf{x}_1^{\mathrm{T}}, \ldots, \mathbf{x}_{N-1}^{\mathrm{T}}\right]^{\mathrm{T}}$. A sum of atomic norms for time index n is chosen as a cost function because it corresponds to the grid-less version of the cost function

in Eq. (2.37),

$$\sum_{n=0}^{N-1} \|\mathbf{x}_n\|_{\mathcal{A}} = \inf \left\{ \sum_{n,k} |c_{n,k}| \,\middle|\, \mathbf{x}_n = \sum_k c_{n,k}\mathbf{a}_{n,k},\ \ \mathbf{a}_{n,k} \in \mathcal{A} \right\}.$$

Thus, the estimation of a sparse T-F representation is formulated as

$$\underset{\mathbf{x}}{\text{minimize}} \quad \sum_{n=0}^{N-1} \|\mathbf{x}_n\|_{\mathcal{A}} \quad \text{subject to} \quad \mathbf{f} = \mathbf{A}_{\mathbf{g}}\mathbf{x}. \tag{4.11}$$

Since $\mathbf{x}_n = \sum_k c_{n,k}\mathbf{a}_{n,k}$, the nth windowed signal $\mathbf{W}_n\mathbf{x}_n = \sum_k c_{n,k}(\mathbf{W}_n\mathbf{a}_{n,k})$ corresponds to the vertical sum of elements in Fig. 4.1. The signal $\mathbf{f}$ can be reconstructed by $\sum_{n=0}^{N-1} \mathbf{W}_n\mathbf{x}_n = \mathbf{A}_{\mathbf{g}}\mathbf{x}$. After solving Eq. (4.11), the coefficients $c_{n,k}$ can be calculated as the line spectrum estimation using the atomic norm (in Sec. 4.2).

4.3.2 Algorithm for solving Eq. (4.11)

To solve Eq. (4.11), we firstly reformulate it as a semidefinite programming. Substituting Eq. (4.4) into Eq. (4.11), it can be rewritten as

$$\begin{aligned} \underset{\mathbf{x},\mathbf{u},\boldsymbol{\nu}}{\text{minimize}} \quad & \sum_{n=0}^{N-1} \frac{1}{2L}\mathrm{Tr}(T(\mathbf{u}_n)) + \frac{1}{2}\nu_n \\ \text{subject to} \quad & \begin{bmatrix} T(\mathbf{u}_n) & \mathbf{x}_n \\ \mathbf{x}_n^* & \nu_n \end{bmatrix} \succeq 0, \ \text{for } n = 0, \ldots, N-1 \\ & \mathbf{f} = \mathbf{A}_{\mathbf{g}}\mathbf{x}, \end{aligned} \tag{4.12}$$

where $\mathbf{u} = \left[\mathbf{u}_0^{\mathrm{T}}, \mathbf{u}_1^{\mathrm{T}}, \ldots, \mathbf{u}_{N-1}^{\mathrm{T}}\right]^{\mathrm{T}}$ and $\boldsymbol{\nu} = [\nu_0, \nu_1, \ldots, \nu_{N-1}]^{\mathrm{T}}$. We adopt the alternating direction method of multipliers (ADMM) [112] to solve Eq. (4.12). To apply ADMM to the proposed method, we introduce auxiliary variables $\mathbf{Z}_n \in \mathbb{C}^{(L+1)\times(L+1)}$ for $n = 0, \ldots, N-1$ and a set corresponding to the reconstruction constraint $C = \{\mathbf{x} | \mathbf{A}_{\mathbf{g}}\mathbf{x} = \mathbf{f}\}$. Then, Eq. (4.12) is reformulated as

$$\begin{aligned} \underset{\substack{\mathbf{x}\in C,\mathbf{u},\boldsymbol{\nu} \\ \mathbf{Z}_n \succeq 0}}{\text{minimize}} \quad & \sum_{n=0}^{N-1} \frac{1}{2L}\mathrm{Tr}(T(\mathbf{u}_n)) + \frac{1}{2}\nu_n \\ \text{subject to} \quad & \begin{bmatrix} T(\mathbf{u}_n) & \mathbf{x}_n \\ \mathbf{x}_n^* & \nu_n \end{bmatrix} = \mathbf{Z}_n, \ \text{for } n = 0, \ldots, N-1. \end{aligned}$$

The augmented Lagrangian associated with this problem is given by

$$\begin{aligned}\mathcal{L}(\mathbf{x},\mathbf{u},\boldsymbol{\nu},\mathbf{Z}_n,\boldsymbol{\Lambda}_n) = &\sum_{n=0}^{N-1}\frac{1}{2L}\mathrm{Tr}(T(\mathbf{u}_n))+\frac{1}{2}\nu_n \\ &+\left\langle \boldsymbol{\Lambda}_n, \begin{bmatrix} T(\mathbf{u}_n) & \mathbf{x}_n \\ \mathbf{x}_n^* & \nu_n \end{bmatrix} - \mathbf{Z}_n \right\rangle_{\mathrm{F}} \\ &+\frac{\rho}{2}\left\| \begin{bmatrix} T(\mathbf{u}_n) & \mathbf{x}_n \\ \mathbf{x}_n^* & \nu_n \end{bmatrix} - \mathbf{Z}_n \right\|_{\mathrm{F}}^2, \end{aligned} \tag{4.13}$$

where $\boldsymbol{\Lambda}_n \in \mathbb{C}^{(L+1)\times(L+1)}$ for $n = 0,\ldots,N-1$ are dual variables, and $\rho > 0$ is the augmented Lagrangian parameter. Then, ADMM consists of the following iterations:

$$(\mathbf{x}^{(i+1)},\mathbf{u}^{(i+1)},\boldsymbol{\nu}^{(i+1)}) = \underset{\mathbf{x}\in C,\mathbf{u},\boldsymbol{\nu}}{\mathrm{argmin}}\ \mathcal{L}(\mathbf{x},\mathbf{u},\boldsymbol{\nu},\mathbf{Z}_n^{(i)},\boldsymbol{\Lambda}_n^{(i)}), \tag{4.14}$$

$$\mathbf{Z}_n^{(i+1)} = \underset{\mathbf{Z}_n\succeq 0}{\mathrm{argmin}}\ \mathcal{L}(\mathbf{x}^{(i+1)},\mathbf{u}^{(i+1)},\boldsymbol{\nu}^{(i+1)},\mathbf{Z}_n,\boldsymbol{\Lambda}_n^{(i)}), \tag{4.15}$$

$$\boldsymbol{\Lambda}_n^{(i+1)} = \boldsymbol{\Lambda}_n^{(i)} + \rho\left(\begin{bmatrix} T(\mathbf{u}_n^{(i+1)}) & \mathbf{x}_n^{(i+1)} \\ (\mathbf{x}_n^{(i+1)})^* & \nu_n^{(i+1)} \end{bmatrix} - \mathbf{Z}_n^{(i+1)} \right). \tag{4.16}$$

Introducing

$$\mathbf{Z}_n = \begin{bmatrix} \mathbf{Z}_{\mathrm{T}n} & \mathbf{z}_{\mathrm{x}n} \\ \mathbf{z}_{\mathrm{x}n}^* & z_{\nu n} \end{bmatrix}, \quad \boldsymbol{\Lambda}_n = \begin{bmatrix} \boldsymbol{\Lambda}_{\mathrm{T}n} & \boldsymbol{\lambda}_{\mathrm{x}n} \\ \boldsymbol{\lambda}_{\mathrm{x}n}^* & \lambda_{\nu n} \end{bmatrix},$$

$\mathbf{z}_{\mathrm{x}} = \left[\mathbf{z}_{\mathrm{x}0}^{\mathrm{T}},\mathbf{z}_{\mathrm{x}1}^{\mathrm{T}},\ldots,\mathbf{z}_{\mathrm{x}N-1}^{\mathrm{T}}\right]^{\mathrm{T}}$, and $\boldsymbol{\lambda}_{\mathrm{x}} = \left[\boldsymbol{\lambda}_{\mathrm{x}0}^{\mathrm{T}},\boldsymbol{\lambda}_{\mathrm{x}1}^{\mathrm{T}},\ldots,\boldsymbol{\lambda}_{\mathrm{x}N-1}^{\mathrm{T}}\right]^{\mathrm{T}}$, Eq. (4.14) can be separately solved for $\mathbf{x}$, $\mathbf{u}$ and $\boldsymbol{\nu}$.

The proposed algorithm is summarized in Algorithm 1. The update for $\mathbf{x}$ is written as the projection onto the set $C = \{\mathbf{x}|\mathbf{A}_{\mathbf{g}}\mathbf{x} = \mathbf{f}\}$, denoted by $P_C(\cdot)$. It is given by

$$P_C(\mathbf{v}) = \mathbf{v} - \mathbf{A}_{\mathbf{g}}^*(\mathbf{A}_{\mathbf{g}}\mathbf{A}_{\mathbf{g}}^*)^{-1}(\mathbf{A}_{\mathbf{g}}\mathbf{v} - \mathbf{f}). \tag{4.17}$$

The computation of matrix inversion in Eq. (4.17) can be avoided by the canonical dual window $\tilde{\mathbf{h}}$ as

$$\mathbf{A}_{\mathbf{g}}^*\left(\mathbf{A}_{\mathbf{g}}\mathbf{A}_{\mathbf{g}}^*\right)^{-1} = \mathbf{A}_{\tilde{\mathbf{h}}}^*. \tag{4.18}$$

The updates for $\mathbf{u}$, $\boldsymbol{\nu}$, $\mathbf{Z}_n$, and $\boldsymbol{\Lambda}_n$ can be computed in parallel for each n. $T^\dagger : \mathbb{C}^{L\times L} \to \mathbb{C}^L$ in the update for $\mathbf{u}$ represents the pseudo-inverse operator of T,

$$T^\dagger(\mathbf{X})[n] = \frac{1}{2(L-n)}\sum_{k=0}^{L-n-1}\left(\mathbf{X}[k,k+n] + \overline{\mathbf{X}[k+n,k]}\right).$$

Algorithm 1 ADMM for solving Eq. (4.11)

Input: $\mathbf{A}$, $\mathbf{f}$, ρ

Output: $\mathbf{x}$, $\mathbf{u}$, $\boldsymbol{\nu}$

Initialize $\mathbf{Z}_n$ and $\mathbf{\Lambda}_n$ for $n = 0, \ldots, N-1$

for $i = 0, 1, \cdots$ **do**

$\mathbf{x} \leftarrow P_C\left(\mathbf{z}_{\mathrm{x}} - \frac{1}{\rho}\boldsymbol{\lambda}_{\mathrm{x}}\right)$

for $n = 1, \cdots, N$ **do**

$\mathbf{u}_n \leftarrow T^{\dagger}\left(\mathbf{Z}_{\mathrm{T}n} - \frac{1}{\rho}\left(\mathbf{\Lambda}_{\mathrm{T}n} - \frac{1}{2}\mathbf{I}_L\right)\right)$

$\nu_n \leftarrow z_{\nu n} - \frac{1}{\rho}\left(\lambda_{\nu n} - \frac{1}{2}\right)$

$$\mathbf{Z}_n \leftarrow P_{\mathbb{S}_+}\left(\begin{bmatrix} T(\mathbf{u}_n) & \mathbf{x}_n \\ \mathbf{x}_n^* & \nu_n \end{bmatrix} + \frac{1}{\rho}\mathbf{\Lambda}_n\right)$$

$$\mathbf{\Lambda}_n \leftarrow \mathbf{\Lambda}_n + \rho\left(\begin{bmatrix} T(\mathbf{u}_n) & \mathbf{x}_n \\ \mathbf{x}_n^* & \nu_n \end{bmatrix} - \mathbf{Z}_n\right)$$

end for

end for

The update for $\mathbf{Z}_n$ can be calculated by a projection onto the positive semidefinite cone $\mathbb{S}_+$, which is implemented by setting the negative eigenvalues to 0. Since Eq. (4.12) is a convex optimization problem, Algorithm 1 can obtain the global optimal solution regardless of the initialization for $\mathbf{Z}_n$ and $\mathbf{\Lambda}_n$.

4.4 Numerical experiments

To evaluate its performance, the proposed method was first applied to an artificial signal containing a sinusoid, a linear chirp, and a quadratic chirp. The proposed method was compared to DGT with the canonical dual window, the reassignment method [7], the ℓ_1-norm minimization [29], and the window-wise atomic norm minimization. The Slepian window [36] was chosen as a window $\mathbf{g}$, and its length and bandwidth were set to 2^7 and 0.04. The time-shifting width and the number of frequency channels were set to $a = 2^4$ and $M = 2^{10}$. $\mathbf{Z}_n$ and $\mathbf{\Lambda}_n$ in Algorithm 1 were initialized as zero matrices. Prony's method was used to estimate $\mathbf{a}_{n,k}$ from the solution of Eq. (4.12).

The estimated T-F representations are shown in Fig. 4.2. The T-F representation obtained using the ℓ_1-norm was better-localized than DGT and the reassign-

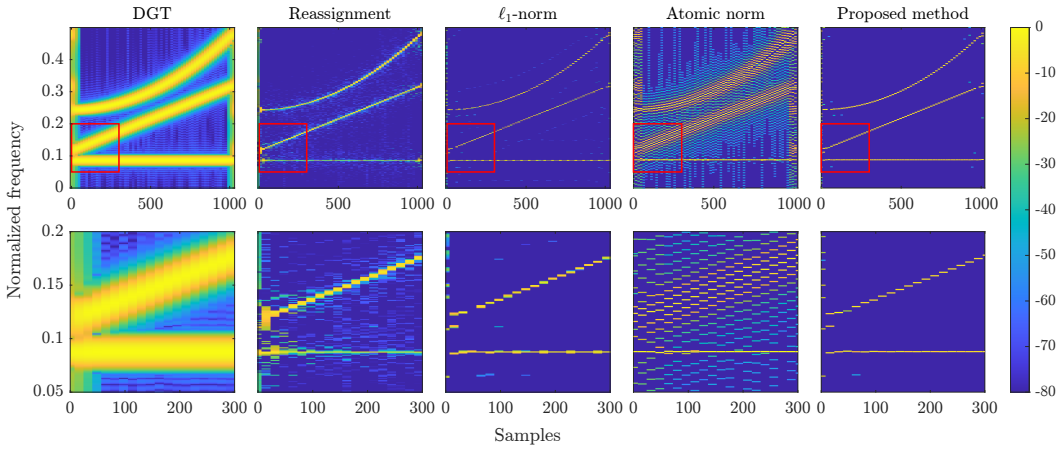

Figure 4.2: T-F representations of an artificial signal. Each column shows (from left to right) the T-F representations obtained by DGT, the reassignment method, the ℓ_1-norm minimization, the window-wise atomic norm minimization, and the proposed method, respectively. The bottom row illustrates these enlargements in the red box.

ment method, but had multiple non-zero coefficients in each time index to represent a sinusoid. The window-wise atomic norm minimization estimated a sparse representation corresponding to the sinusoid. However, it cannot express the chirps sparsely. The T-F representation obtained by the proposed method was the most-localized among these T-F representations.

To evaluate the sparseness of these T-F representations, the squared absolute value of coefficients is plotted in Fig. 4.3. It can be seen that the proposed method can represent the signal using the fewest coefficients. These results indicate that the proposed method provides a sparse representation using the atomic norm while considering the relationship among each time index.

Finally, the proposed method was applied to a speech signal. The settings associated with the Gabor system and the proposed algorithm were the same as in the previous experiment. The estimated T-F representations are shown in Fig. 4.4. It can be seen that the T-F representation obtained by the proposed method was the sparsest among them. The result suggests that the proposed method performs well for a real audio signal.

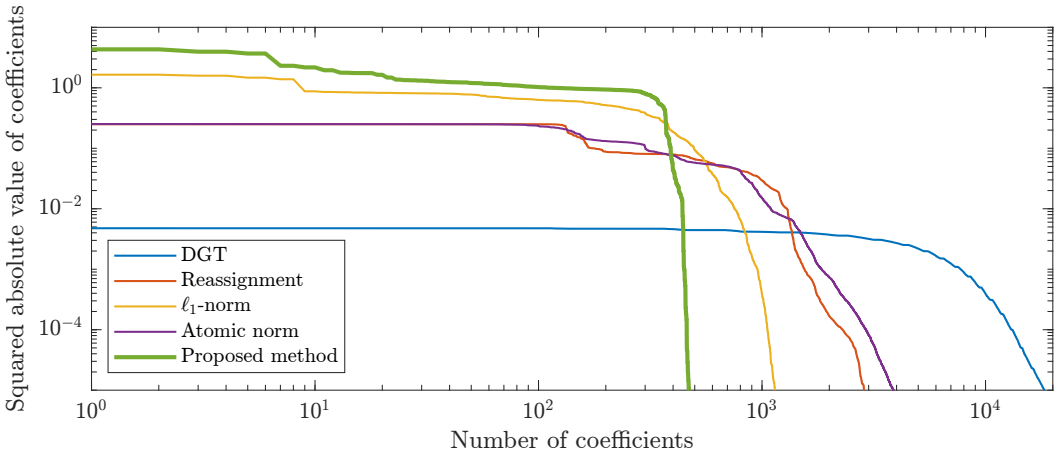

Figure 4.3: Squared absolute value of coefficients in decending order.

4.5 Conclusion

In this chapter, we proposed a method for estimating a sparse T-F representation via atomic norm minimization. The proposed method estimates a sparse T-F representation without discretizing frequency using the atomic norm. The experimental results show that the proposed method can estimate a sparser T-F representation than existing methods.

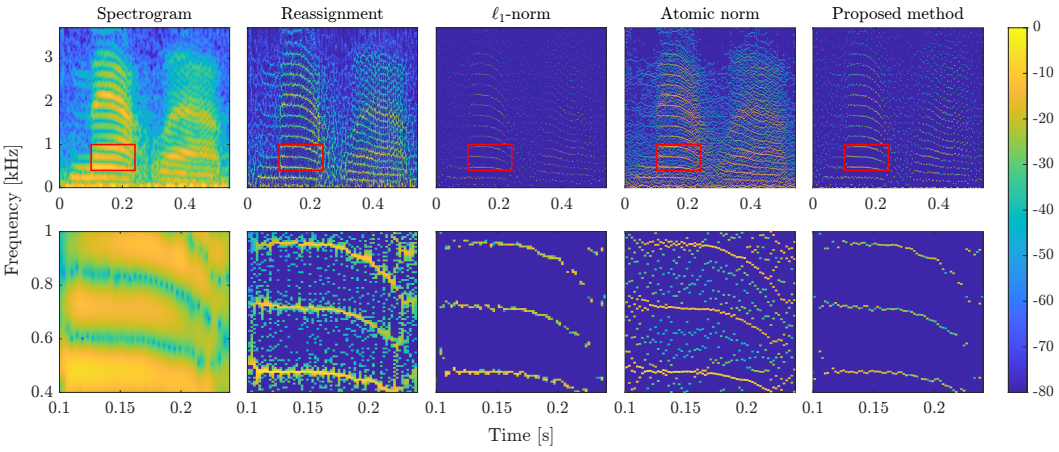

Figure 4.4: T-F representations of a speech signal. Each row and column represents the same as Fig. 4.2.

Chapter 5

Introducing nonconvex regularization to gridless sparse method

5.1 Introduction

Sparse optimization is a technique for estimating a signal from measurements by assuming that the signal is sparse in certain domains [113, 114]. This approach has been widely employed in signal processing [115, 116]. Although the direct formulation for finding sparse representation involves minimizing the number of nonzero values, called the ℓ_0 norm, this problem is usually an intractable combinatorial optimization problem. A typical formulation of sparse optimization uses the ℓ_1 norm instead of the ℓ_0 norm as a sparsity-inducing regularization term. The ℓ_1 norm corresponds to a convex relaxation of the ℓ_0 norm, and efficient algorithms are available for solving this problem. However, even though the actual parameters take continuous values in applications such as line spectrum estimation and direction of arrival estimation, this formulation involves discretizing a continuous parameter onto a grid. As a result, the performance is often degraded due to model mismatches between the signal and the predefined grid [97]. To address such grid mismatch issues, some researchers have proposed off-grid methods based on some techniques including Taylor series approximation [117–119], linear interpolation [120], adaptive grid estimation [121, 122], and clustering [123]. These models perform better than fixed grid-based approaches, but they are often nonconvex and challenging to solve.

In addition, their performances depend on the initial grids.

Recently, a gridless sparse optimization method via atomic norm minimization (ANM) has been developed [98–101]. In ANM, an infinite-dimensional sparse optimization problem is transformed into a finite-dimensional semidefinite programming (SDP) problem. It uses the Vandermonde structure, which enables continuous parameters to be handled without discretization. ANM-based sparse optimization has been applied in line spectrum estimation [102, 103], direction of arrival estimation [104–106], target localization in radar [107], and time-frequency analysis [124] and has demonstrated its effectiveness.

Although ANM can perform estimation without grid mismatches, the resulting estimator is biased due to decreasing the amplitudes of coefficients, as in the ℓ_1 norm-based method. In addition, the atomic norm suffers from a resolution limit due to the convex relaxation [125, 126]. For cases in which some prior information about the parameters is available, bias reduction methods that introduce weights to the atomic norm have been proposed [127, 128]. In [127], prior information is introduced as weights by dividing continuous parameters into several blocks. [128] proposed an efficient algorithm for ANM weighted by a prior probability distribution. For situations where no prior information is available, reweighting methods have been proposed to enhance sparsity [129, 130]. These reweighting methods originate from the reweighting method in the on-grid ℓ_1 norm-based method [131].

Other than the reweighting methods, nonconvex regularizations have been proposed as alternatives to the ℓ_0 and ℓ_1 norms [132, 133]. These methods have been designed to reduce the bias while relaxing the ℓ_0 norm. Examples include the ℓ_q $(0 < q < 1)$ norm [134], the logarithmic penalty [135], the smoothly clipped absolute deviation (SCAD) [136], the minimax concave penalty (MCP) [137], and the capped ℓ_1 norm [138]. These nonconvex regularizations have performed better than the ℓ_1 norm. Introducing such nonconvex regularization into the gridless sparse approach is expected to improve its performance, as well as the success of reweighting-based ANM.

In this chapter, we propose a framework for incorporating such nonconvex regularizations into the gridless sparse optimization method. Our method is based on an SDP formulation of ANM. The proposed method uses proximal algorithms and can handle various existing regularizations with the same algorithm. Numerical experiments confirmed that the proposed method provides performance better than

conventional ANM and reweighted ANM (RAM).

The rest of this chapter is organized as follows. Sec. 5.2 introduces the concept of sparse optimization via the atomic norm. Sec. 5.3 explains a nonconvex regularization approach that promotes sparsity and the proximal algorithms that can be used to solve it. Then, Sec. 5.4 explains the proposed method, which combines ANM and nonconvex regularization. Sec. 5.5 provides numerical experiments to confirm the effectiveness of the proposed method, and the conclusion of this chapter is presented in Sec. 5.6.

5.2 Gridless sparse method with atomic norm

This section describes the signal model and the theoretical background of the gridless sparse method executed via ANM.

5.2.1 Signal model

Let us consider the problem of estimating a signal $\mathbf{S} \in \mathbb{C}^{N\times M}$ from an observation $\mathbf{y} \in \mathbb{C}^{L}$ represented as

$$\mathbf{y} = \mathcal{B}(\mathbf{S}) + \mathbf{n}, \tag{5.1}$$

where $\mathcal{B} : \mathbb{C}^{N\times M} \to \mathbb{C}^{L}$ is a known linear mapping and $\mathbf{n} \in \mathbb{C}^{L}$ represents additive Gaussian noise. The signal $\mathbf{S}$ is assumed to be expressed as the sum of a few complex sinusoids, which is given by

$$\mathbf{S} = \sum_{k=0}^{K-1} \mathbf{a}_k \mathbf{s}_k^{\mathrm{H}}, \qquad \mathbf{a}_k \in \mathcal{A}, \tag{5.2}$$

where $\mathcal{A}$ is a collection of complex sinusoids

$$\mathcal{A} = \left\{ \mathbf{a} \in \mathbb{C}^{N} \,\middle|\, \mathbf{a}[l] = \mathrm{e}^{\mathrm{i}2\pi\omega l}, \omega \in [0,1) \right\}, \tag{5.3}$$

and $\mathbf{s}_k \in \mathbb{C}^{M}$ is the nth coefficient vector of the sinusoids. The ANM-based sparse method aims to estimate the frequencies ω_k and the coefficients $\mathbf{s}_k$ without a priori knowledge of the number of sinusoids K. ANM is based on the assumption that a signal $\mathbf{X}$ can be expressed as a sum of a few atoms in the set $\mathcal{A}$.

5.2.2 Atomic ℓ_0 norm

A direct metric for the number of sinusoids K is the atomic ℓ_0 norm, which is defined as

$$\|\mathbf{X}\|_{\mathcal{A},0} = \inf_{\mathbf{a}_k,\mathbf{s}_k} \left\{ K \,\middle|\, \mathbf{X} = \sum_{k=0}^{K-1} \mathbf{a}_k \mathbf{s}_k^{\mathrm{H}},\ \ \mathbf{a}_k \in \mathcal{A} \right\}. \tag{5.4}$$

The atomic ℓ_0 norm in (5.4) equals the optimal value of the following optimization problem [139]:

$$\begin{aligned} \underset{\mathbf{Z},\mathbf{u},\mathbf{Q}}{\text{minimize}} \quad & \operatorname{rank}(\mathbf{Z}) \\ \text{subject to} \quad & \mathbf{Z} = \begin{bmatrix} T(\mathbf{u}) & \mathbf{X} \\ \mathbf{X}^{\mathrm{H}} & \mathbf{Q} \end{bmatrix} \succeq 0, \end{aligned} \tag{5.5}$$

where $T : \mathbb{C}^N \to \mathbb{C}^{N\times N}$ is the Hermitian Toeplitz operator:

$$T(\mathbf{u}) = \begin{bmatrix} \mathbf{u}[0] & \overline{\mathbf{u}[1]} & \cdots & \overline{\mathbf{u}[N-1]} \\ \mathbf{u}[1] & \mathbf{u}[0] & \ddots & \vdots \\ \vdots & \ddots & \ddots & \overline{\mathbf{u}[1]} \\ \mathbf{u}[N-1] & \cdots & \mathbf{u}[1] & \mathbf{u}[0] \end{bmatrix}. \tag{5.6}$$

Denoting the optimal solutions of (5.5) by $\mathbf{u}^\star$, $\mathbf{Q}^\star$, and $\mathbf{Z}^\star$, it is known that $\operatorname{rank}(\mathbf{Z}^\star)$ equals $\operatorname{rank}(T(\mathbf{u}^\star))$. Utilizing the ℓ_0 norm $\|\cdot\|_{\mathcal{A},0}$, a problem of estimating the signal $\mathbf{S}$ from an observation $\mathbf{y}$ is formulated as

$$\begin{aligned} \underset{\mathbf{X}}{\text{minimize}} \quad & \|\mathbf{X}\|_{\mathcal{A},0} \\ \text{subject to} \quad & \|\mathcal{B}(\mathbf{X}) - \mathbf{y}\|_2 \le \eta, \end{aligned} \tag{5.7}$$

where $\eta \ge 0$ is a data fidelity parameter. $\eta = 0$ corresponds to the case where the observation $\mathbf{y}$ contains no noise. Substituting (5.5) into (5.7), the equation can be rewritten as

$$\begin{aligned} \underset{\mathbf{Z},\mathbf{u},\mathbf{X},\mathbf{Q}}{\text{minimize}} \quad & \operatorname{rank}(\mathbf{Z}) \\ \text{subject to} \quad & \mathbf{Z} = \begin{bmatrix} T(\mathbf{u}) & \mathbf{X} \\ \mathbf{X}^{\mathrm{H}} & \mathbf{Q} \end{bmatrix} \succeq 0 \\ & \|\mathcal{B}(\mathbf{X}) - \mathbf{y}\|_2 \le \eta. \end{aligned} \tag{5.8}$$

From the property of the Schur complement [140], the solution $T(\mathbf{u}^\star)$ satisfies $T(\mathbf{u}^\star) \succeq 0$. Thus, the Vandermonde decomposition can be applied to $T(\mathbf{u}^\star)$. Any

positive semidefinite Hermitian Toeplitz matrix can be decomposed as [108]

$$T(\mathbf{u}^\star) = \sum_{k=0}^{K^\star-1} p_k^\star \mathbf{a}_k^\star \mathbf{a}_k^{\star\mathrm{H}}, \qquad \mathbf{a}_k^\star \in \mathcal{A}, \tag{5.9}$$

where $K^\star$ corresponds to $\mathrm{rank}(T(\mathbf{u}^\star)) = \mathrm{rank}(\mathbf{Z}^\star)$. If $K^\star < N$, its decomposition is uniquely determined. In practical computation scenarios, $\mathbf{a}_k^\star$ in (5.9) can be obtained by Prony's method [109], the matrix pencil method [110], or other linear prediction methods [111]. Then, the coefficients $\mathbf{p}^\star = \left[p_0^\star, p_1^\star, \ldots, p_{K^\star-1}^\star\right]$ can be obtained by solving the following linear equation:

$$\mathbf{A}^\star \mathbf{p}^\star = \mathbf{u}^\star, \tag{5.10}$$

where $\mathbf{A}^\star = \left[\mathbf{a}_0^\star, \mathbf{a}_1^\star, \ldots, \mathbf{a}_{K^\star-1}^\star\right]$.

5.2.3 Convex relaxation of the atomic ℓ_0 norm

Once the optimal solution to (5.8) is obtained, the frequencies ω_k and coefficients $\mathbf{s}_k$ can be calculated through the Vandermonde decomposition. However, the atomic ℓ_0 norm is an intractable combinatorial optimization problem. Instead, the atomic (ℓ_1) norm [99],

$$\|\mathbf{X}\|_{\mathcal{A}} = \inf_{\mathbf{a}_k, \mathbf{s}_k} \left\{ \sum_k \|\mathbf{s}_k\|_2 \,\middle|\, \mathbf{X} = \sum_k \mathbf{a}_k \mathbf{s}_k^{\mathrm{H}}, \;\; \mathbf{a}_k \in \mathcal{A} \right\}, \tag{5.11}$$

is used as convex relaxation of the atomic ℓ_0 norm. The sparse estimation problem formulated in terms of the atomic norm is expressed as

$$\begin{aligned} \underset{\mathbf{X}}{\text{minimize}} \quad & \|\mathbf{X}\|_{\mathcal{A}} \\ \text{subject to} \quad & \|\mathcal{B}(\mathbf{X}) - \mathbf{y}\|_2 \leq \eta. \end{aligned} \tag{5.12}$$

The atomic norm can also be characterized by the solution to the following optimization problem:

$$\begin{aligned} \underset{\mathbf{Z},\mathbf{u},\mathbf{Q}}{\text{minimize}} \quad & \frac{1}{2\sqrt{N}} \mathrm{Tr}\left(\mathbf{Z}\right) \\ \text{subject to} \quad & \mathbf{Z} = \begin{bmatrix} T(\mathbf{u}) & \mathbf{X} \\ \mathbf{X}^{\mathrm{H}} & \mathbf{Q} \end{bmatrix} \succeq 0, \end{aligned} \tag{5.13}$$

Hence, (5.12) is reformulated as an SDP problem:

$$\begin{aligned} \underset{\mathbf{Z},\mathbf{u},\mathbf{X},\mathbf{Q}}{\text{minimize}} \quad & \frac{1}{2\sqrt{N}}\operatorname{Tr}(\mathbf{Z}) \\ \text{subject to} \quad & \mathbf{Z} = \begin{bmatrix} T(\mathbf{u}) & \mathbf{X} \\ \mathbf{X}^{\mathrm{H}} & \mathbf{Q} \end{bmatrix} \succeq 0 \\ & \|\mathcal{B}(\mathbf{X}) - \mathbf{y}\|_2 \leq \eta. \end{aligned} \tag{5.14}$$

5.3 Proximal operator for sparse and low-rank estimation

This section provides the prior knowledge required for the proposed method.

5.3.1 Proximal operator of nonconvex sparse regularizations

Sparsity-inducing penalties, including the ℓ_1 norm, are nonsmooth functions. Proximal algorithms can efficiently solve optimization problems involving such regularizations [141, 142]. The main operation in a proximal algorithm involves evaluating the proximal operator of a regularizer iteratively.

Let R be a proper and lower semicontinuous function. The proximal operator of R is defined by

$$\operatorname{prox}_{\lambda R(\cdot)}(\mathbf{x}) = \underset{\mathbf{y}}{\operatorname{argmin}} \; R(\mathbf{y}) + \frac{1}{2\lambda}\|\mathbf{x}-\mathbf{y}\|_2^2, \tag{5.15}$$

where $\lambda > 0$. If $R(\cdot)$ is fully separable, i.e., $R(\mathbf{x}) = \sum_l r(\mathbf{x}[l])$, then $\operatorname{prox}_{\lambda R(\cdot)}(\mathbf{x})$ can be evaluated in an elementwise manner as

$$\begin{aligned} \left(\operatorname{prox}_{\lambda R(\cdot)}(\mathbf{x})\right)[l] &= \underset{y}{\operatorname{argmin}} \; r(y) + \frac{1}{2\lambda}|\mathbf{x}[l]-y|^2 \\ &= \operatorname{prox}_{\lambda r(\cdot)}(\mathbf{x}[l]). \end{aligned} \tag{5.16}$$

Table 5.1 and Fig. 5.1 show examples of nonconvex penalties for inducing soarsity and their proximal operators. All of these penalties except for the ℓ_1 norm are nonconvex. The estimator using the ℓ_1 norm is biased due to decreasing the amplitudes of the coefficients when the true parameters are relatively large. To reduce the bias while maintaining numerical stability, these nonconvex penalties have been designed to fill in the gap between the ℓ_0 and ℓ_1 norms. Specifically, these penalties behave like the ℓ_1 norm when the amplitudes of the parameters are small and the

ℓ_0 norm when the amplitudes are large. The proximal operator of the ℓ_q norm $(0 < q < 1)$ does not have a closed-form expression except in the two special cases where $q = 1/2$ [143] and $q = 2/3$ [144]. An iterative algorithm has been proposed to evaluate the proximal operator of the ℓ_q norm for arbitrary q $(0 < q < 1)$ [145].

5.3.2 Proximal operator for low-rank matrix positive semidefinite matrix

Estimation of low-rank positive semidefinite matrices is formulated as minimizing the sum of low-rank inducing regularization $\mathcal{R}(\mathbf{Z})$ and the indicator function of a positive semidefinite cone $\iota_{\mathbb{S}_+^N}(\mathbf{Z})$. Inducing low-rankness of a Hermitian matrix $\mathbf{Z}$ can be achieved by regularizing the sparsity of its eigenvalues. The positive semidefiniteness of a Hermitian matrix corresponds to the nonnegativity of its eigenvalues. Hence, the proximal operators of $\mathcal{R}(\mathbf{Z})$ and $\iota_{\mathbb{S}_+^N}(\mathbf{Z})$ are calculated as [141]

$$\begin{aligned} \mathrm{prox}_{\lambda\mathcal{R}(\cdot)}(\mathbf{X}) &= \mathbf{V}_{\mathbf{X}} \,\mathrm{diag}(\mathrm{prox}_{\lambda R(\cdot)}(\boldsymbol{\sigma}_{\mathbf{X}}))\mathbf{V}_{\mathbf{X}}^{\mathrm{H}} \qquad (5.17) \\ \mathrm{prox}_{\iota_{\mathbb{S}_+^N}}(\mathbf{X}) &= \mathbf{V}_{\mathbf{X}} \,\mathrm{diag}(\mathrm{prox}_{\iota_{\mathbb{R}_+}}(\boldsymbol{\sigma}_{\mathbf{X}}))\mathbf{V}_{\mathbf{X}}^{\mathrm{H}} \\ &= \mathbf{V}_{\mathbf{X}} \,\mathrm{diag}((\boldsymbol{\sigma}_{\mathbf{X}})_+)\mathbf{V}_{\mathbf{X}}^{\mathrm{H}}, \qquad (5.18) \end{aligned}$$

respectively, where $\boldsymbol{\sigma}_{\mathbf{Z}} \in \mathbb{R}^N$ is the eigenvalues of $\mathbf{Z}$, $\mathbf{V}_{\mathbf{X}}$ is a matrix whose columns are eigenvectors of $\mathbf{X}$, and $(\cdot)_+ = \max(\cdot, 0)$. In addition, since both $R(\cdot)$ and $\iota_{\cdot\geq 0}(\cdot)$ are separable, the proximal operator of $R_+(\cdot) = R(\cdot) + \iota_{\cdot\geq 0}(\cdot)$ can be evaluated as

$$\begin{aligned} \left(\mathrm{prox}_{\lambda R_+(\cdot)}(\mathbf{x})\right)[l] &= \underset{y\in\mathbb{R}_+}{\mathrm{argmin}} \; r(y) + \frac{1}{2\lambda}\left|\mathbf{x}[l] - y\right|^2 \\ &= \left(\mathrm{prox}_{\lambda r(\cdot)}(\mathbf{x}[l])\right)_+ . \qquad (5.19) \end{aligned}$$

5.4 Proposed method

The SDP formulations of the atomic ℓ_0 norm and atomic norm are explained in Sec. 5.2. The SDP formulation of ANM (5.13) can be viewed as the minimization of the trace norm of a semidefinite matrix. The trace norm induces low rankness in the semidefinite matrix since the trace norm corresponds to the convex relaxation of the rank of the matrix. Similar to that of the ℓ_1 norm, the estimator of the trace norm is biased to reduce the amplitudes of the eigenvalues.

Table 5.1: Popular penalties for inducing sparsity and their proximal operators.

Penalty	$\lambda r(x)$	$\mathrm{prox}_{\lambda r(\cdot)}(x)$
ℓ_1 norm	$\lambda\|x\|$	$\mathrm{sign}(x)\max(\|x\|-\lambda,0)$
ℓ_0 norm	$\lambda\|x\|_0$	$\begin{cases} 0 & \text{if } \|x\|<\sqrt{2\lambda} \\ \{0,x\} & \text{if } \|x\|=\sqrt{2\lambda} \\ x & \text{if } \|x\|>\sqrt{2\lambda} \end{cases}$
ℓ_q norm $(0<q<1)$	$\lambda\|x\|^q$	$\begin{cases} 0 & \text{if } \|x\|<\tau \\ \{0,\mathrm{sign}(x)\beta\} & \text{if } \|x\|=\tau \\ \mathrm{sign}(x)y & \text{if } \|x\|>\tau \end{cases}$ where $\beta=(2\lambda(1-q))^{\frac{1}{2-q}}$, $\tau=\beta+\lambda q\beta^{q-1}$, and $y\in[\beta,\|x\|]$ such that $\lambda q y^{q-1}+y=\|x\|$
Logarithmic [135]	$\frac{\lambda}{a}\log(1+a\|x\|)$	$\begin{cases} 0 & \text{if } \|x\|<\lambda \\ \left[\frac{\|x\|}{2}-\frac{1}{2a}+\sqrt{\left(\frac{\|x\|}{2}+\frac{1}{2a}\right)^2-\frac{\lambda}{a}}\right]\mathrm{sign}(x) & \text{if } \|x\|\geq\lambda \end{cases}$
SCAD [136]	$\begin{cases} \lambda\|x\| & \text{if } \|x\|<\lambda \\ \frac{2\mu\lambda\|x\|-x^2-\lambda^2}{2(\mu-1)} & \text{if } \lambda\leq\|x\|<\mu\lambda \\ \frac{\mu+1}{2}\lambda^2 & \text{if } \|x\|\geq\mu\lambda \end{cases}$	$\begin{cases} \mathrm{sign}(x)(\|x\|-\lambda)_+ & \text{if } \|x\|\leq 2\lambda \\ \frac{(\mu-1)x-\mathrm{sign}(x)\mu\lambda}{\mu-2} & \text{if } 2\lambda<\|x\|\leq\mu\lambda \\ x & \text{if } \|x\|>\mu\lambda \end{cases}$
MCP [137]	$\lambda\int_0^{\|x\|}\max\left(1-\frac{t}{\gamma\lambda},0\right)\mathrm{d}t$	$\begin{cases} 0 & \text{if } \|x\|\leq\lambda \\ \frac{\mathrm{sign}(x)(\|x\|-\lambda)}{1-1/\gamma} & \text{if } \lambda<\|x\|\leq\gamma\lambda \\ x & \text{if } \|x\|>\gamma\lambda \end{cases}$
Capped ℓ_1	$\lambda\min(\|x\|,\theta)$	$\begin{cases} \mathrm{sign}(x)\max(\|x\|-\lambda) & \text{if } \|x\|<\theta+\frac{\lambda}{2} \\ \mathrm{sign}(x)\left(\theta\pm\frac{\lambda}{2}\right) & \text{if } \|x\|=\theta+\frac{\lambda}{2} \\ x & \text{if } \|x\|>\theta+\frac{\lambda}{2} \end{cases}$

In this chapter, we propose a method to reduce the bias by replacing the trace norm of (5.13) with rank-reducing nonconvex penalties. The relationship among the atomic ℓ_0 norm, the atomic norm, and the proposed method is illustrated in Fig. 5.2. The proposed method bridges the atomic ℓ_0 norm and the atomic norm through the nonconvex regularizations as shown in Table 5.1.

5.4.1 Formulation

Since the trace norm $\|\cdot\|_*$ of $\mathbf{Z}\in\mathbb{S}_+^{N+M}$ equals

$$\sum_{k=0}^{N+M-1}|\boldsymbol{\sigma}_{\mathbf{Z}}[k]|=\sum_{k=0}^{N+M-1}\boldsymbol{\sigma}_{\mathbf{Z}}[k]=\mathrm{Tr}(\mathbf{Z}), \tag{5.20}$$

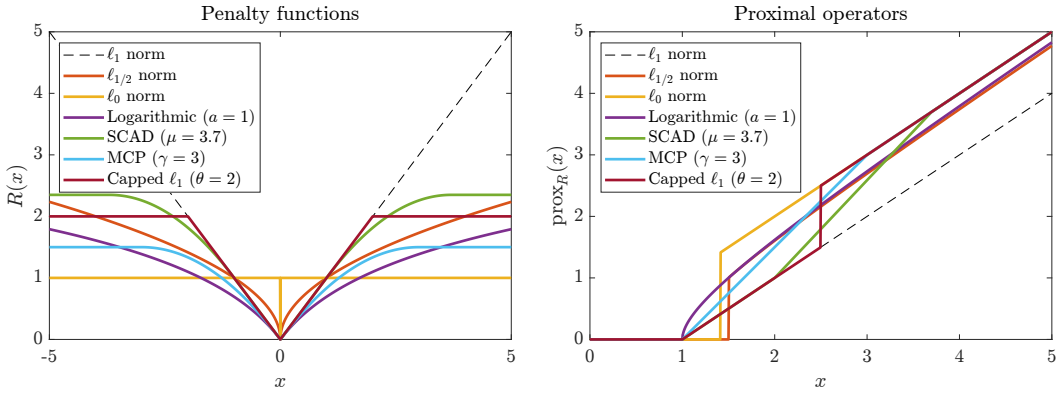

Figure 5.1: Nonconvex penalties for inducing sparsity and their proximal operators ($\lambda = 1$).

(5.13) can be rewritten as

$$\begin{aligned} \|\mathbf{X}\|_{\mathcal{A}} = \min_{\mathbf{Z},\mathbf{u},\mathbf{Q}} \quad & \frac{1}{2\sqrt{N}} \|\mathbf{Z}\|_* \\ \text{subject to} \quad & \mathbf{Z} = \begin{bmatrix} T(\mathbf{u}) & \mathbf{X} \\ \mathbf{X}^{\mathrm{H}} & \mathbf{Q} \end{bmatrix} \succeq 0. \end{aligned} \tag{5.21}$$

The trace norm $\|\cdot\|_*$ corresponds to convex relaxation of the rank of the matrix. According to (5.21), the SDP formulation of the atomic norm can be considered a convex relaxation of the SDP representation of the atomic ℓ_0 norm.

Replacing the trace norm in the SDP formulation of the atomic norm with a rank-reducing penalty $\mathcal{R}(\cdot)$ is expected to improve the performance of the gridless sparse estimation. Therefore, we propose a regularization

$$\begin{aligned} \mathscr{R}(\mathbf{X}) = \min_{\mathbf{Z},\mathbf{u},\mathbf{Q}} \quad & \lambda \mathcal{R}\left(\mathbf{Z}\right) \\ \text{subject to} \quad & \mathbf{Z} = \begin{bmatrix} T(\mathbf{u}) & \mathbf{X} \\ \mathbf{X}^{\mathrm{H}} & \mathbf{Q} \end{bmatrix} \succeq 0. \end{aligned} \tag{5.22}$$

When $\mathcal{R}(\cdot)$ is the trace norm $\|\cdot\|_*$ or the rank $\mathrm{rank}(\cdot)$, the regularization $\mathscr{R}(\cdot)$ is equal to the atomic norm $\|\cdot\|_{\mathcal{A}}$ or the atomic ℓ_0 norm $\|\cdot\|_{\mathcal{A},0}$, respectively. The proposed regularization $\mathscr{R}(\cdot)$ corresponds to the nonconvex relaxation of the SDP formulation of the atomic ℓ_0 norm (5.5). However, the relationship between the direct nonconvex relaxation of the atomic ℓ_0 norm,

$$\|\mathbf{X}\|_{\mathcal{A},r} = \inf_{\mathbf{a}_k,\mathbf{s}_k} \left\{ \sum_k r(\|\mathbf{s}_k\|_2) \,\middle|\, \mathbf{X} = \sum_k \mathbf{a}_k \mathbf{s}_k^{\mathrm{H}}, \;\; \mathbf{a}_k \in \mathcal{A} \right\},$$

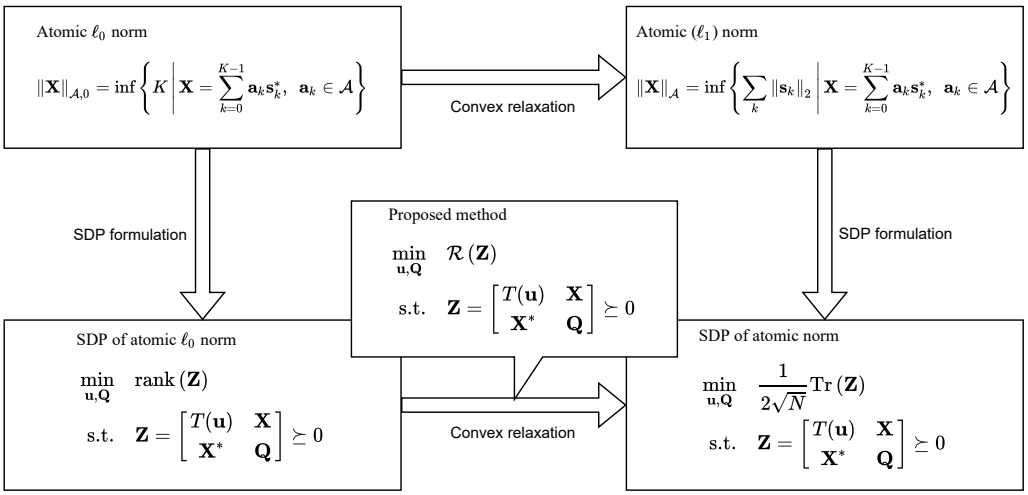

Figure 5.2: Relationship among the atomic ℓ_0 norm, the atomic norm, and the proposed method.

and the proposed regularization is not clear.

Utilizing the regularization $\mathscr{R}(\cdot)$, the problem of estimating a signal $\mathbf{S}$ from an observation $\mathbf{y}$ is formulated as

$$\underset{\mathbf{X}}{\text{minimize}} \quad \mathscr{R}(\mathbf{X}) \quad \text{subject to} \quad \|\mathcal{B}(\mathbf{X}) - \mathbf{y}\|_2 \leq \eta, \tag{5.23}$$

which can also be reformulated as the following SDP problem:

$$\begin{aligned} \underset{\mathbf{Z},\mathbf{u},\mathbf{X},\mathbf{Q}}{\text{minimize}} \quad & \lambda \mathcal{R}(\mathbf{Z}) \\ \text{subject to} \quad & \mathbf{Z} = \begin{bmatrix} T(\mathbf{u}) & \mathbf{X} \\ \mathbf{X}^{\mathrm{H}} & \mathbf{Q} \end{bmatrix} \succeq 0 \\ & \|\mathcal{B}(\mathbf{X}) - \mathbf{y}\|_2 \leq \eta. \end{aligned} \tag{5.24}$$

A standard SDP solver is available for solving the conventional ANM problem. In contrast, since (5.24) is nonconvex optimization problem, the solver cannot be used to solve (5.24).

5.4.2 Algrithm via proximal ADMM

In this chapter, we adopt the proximal alternating direction method of multipliers (ADMM) to solve (5.24) [146]. For applying the proximal ADMM to the proposed

method, (5.24) is reformulated as

$$\begin{aligned} \underset{\mathbf{Z},\mathbf{u},\mathbf{X},\mathbf{Q}}{\text{minimize}} \quad & \lambda\mathcal{R}_+(\mathbf{Z}) + \iota_C(\mathbf{X}) \\ \text{subject to} \quad & \mathbf{Z} = \begin{bmatrix} T(\mathbf{u}) & \mathbf{X} \\ \mathbf{X}^* & \mathbf{Q} \end{bmatrix}, \end{aligned} \tag{5.25}$$

when $\eta = 0$, and

$$\begin{aligned} \underset{\mathbf{Z},\mathbf{u},\mathbf{X},\mathbf{Q}}{\text{minimize}} \quad & \lambda\mathcal{R}_+(\mathbf{Z}) + \frac{1}{2}\left\|\mathcal{B}(\mathbf{X}) - \mathbf{y}\right\|_2^2 \\ \text{subject to} \quad & \mathbf{Z} = \begin{bmatrix} T(\mathbf{u}) & \mathbf{X} \\ \mathbf{X}^* & \mathbf{Q} \end{bmatrix}, \end{aligned} \tag{5.26}$$

when $\eta > 0$, where

$$\mathcal{R}_+(\mathbf{Z}) = \mathcal{R}(\mathbf{Z}) + \iota_{\mathbb{S}_+^{N+M}}(\mathbf{Z}), \tag{5.27}$$

and $C = \{\mathbf{X} | \mathcal{B}(\mathbf{X}) = \mathbf{y}\}$. Although λ in (5.25) does not change the solution to the optimization problem, An estimation obtained by the algorithm depends on λ. Additionally, λ in (5.26) controlls the data fidelity instead of η.

The augmented Lagrangian associated with (5.25) is given by

$$\begin{aligned} &\mathcal{L}_\rho(\mathbf{Z}, \mathbf{X}, \mathbf{u}, \mathbf{Q}, \mathbf{\Lambda}) \\ &= \lambda\mathcal{R}_+(\mathbf{Z}) + \iota_C(\mathbf{X}) \\ &+ \left\langle \mathbf{\Lambda}, \mathbf{Z} - \begin{bmatrix} T(\mathbf{u}) & \mathbf{X} \\ \mathbf{X}^{\mathrm{H}} & \mathbf{Q} \end{bmatrix} \right\rangle_{\mathrm{F}} + \frac{\rho}{2}\left\| \mathbf{Z} - \begin{bmatrix} T(\mathbf{u}) & \mathbf{X} \\ \mathbf{X}^{\mathrm{H}} & \mathbf{Q} \end{bmatrix} \right\|_{\mathrm{F}}^2, \end{aligned}$$

where $\mathbf{\Lambda} \in \mathbb{C}^{(L+M)\times(L+M)}$ is a dual variable, and $\rho > 0$ is the augmented Lagrangian parameter. Then, the proximal ADMM for solving (5.25) consists of the following iterations:

$$\begin{aligned} \mathbf{Z}^{(i+1)} &= \underset{\mathbf{Z}}{\operatorname{argmin}} \ \mathcal{L}_\rho(\mathbf{Z}, \mathbf{X}^{(i)}, \mathbf{u}^{(i)}, \mathbf{Q}^{(i)}, \mathbf{\Lambda}^{(i)}) \\ &\qquad + \frac{\beta}{2}\left\|\mathbf{Z} - \mathbf{Z}^{(i)}\right\|_{\mathrm{F}}^2 \\ (\mathbf{X}^{(i+1)}, \mathbf{u}^{(i+1)}, \mathbf{Q}^{(i+1)}) &= \underset{\mathbf{X},\mathbf{u},\mathbf{Q}}{\operatorname{argmin}} \ \mathcal{L}_\rho(\mathbf{Z}^{(i+1)}, \mathbf{X}, \mathbf{u}, \mathbf{Q}, \mathbf{\Lambda}^{(i)}) \\ &\qquad + \frac{\beta}{2}\left\| \begin{bmatrix} T(\mathbf{u}) & \mathbf{X} \\ \mathbf{X}^{\mathrm{H}} & \mathbf{Q} \end{bmatrix} - \begin{bmatrix} T(\mathbf{u}^{(i)}) & \mathbf{X}^{(i)} \\ (\mathbf{X}^{(i)})^{\mathrm{H}} & \mathbf{Q}^{(i)} \end{bmatrix} \right\|_{\mathrm{F}}^2 \\ \mathbf{\Lambda}^{(i+1)} &= \mathbf{\Lambda}^{(i)} + \rho\left(\mathbf{Z}^{(i+1)} - \begin{bmatrix} T(\mathbf{u}^{(i+1)}) & \mathbf{X}^{(i+1)} \\ (\mathbf{X}^{(i+1)})^{\mathrm{H}} & \mathbf{Q}^{(i+1)} \end{bmatrix} \right), \end{aligned}$$

where $\beta > 0$. The update for $\mathbf{Z}$ can be rewritten as

$$\begin{aligned}
\mathbf{Z}^{(i+1)} &= \underset{\mathbf{Z}}{\operatorname{argmin}}\ \lambda\mathcal{R}_+(\mathbf{Z}) + \left\langle \mathbf{\Lambda}^{(i)}, \mathbf{Z} \right\rangle_{\mathrm{F}} \\
&\quad + \frac{\rho}{2}\left\| \mathbf{Z} - \begin{bmatrix} T(\mathbf{u}^{(i)}) & \mathbf{X}^{(i)} \\ (\mathbf{X}^{(i)})^{\mathrm{H}} & \mathbf{Q}^{(i)} \end{bmatrix} \right\|_{\mathrm{F}}^2 + \frac{\beta}{2}\left\| \mathbf{Z} - \mathbf{Z}^{(i)} \right\|_{\mathrm{F}}^2 \\
&= \underset{\mathbf{Z}}{\operatorname{argmin}}\ \lambda\mathcal{R}_+(\mathbf{Z}) \\
&\quad + \frac{\rho+\beta}{2}\left\| \mathbf{Z} - \frac{1}{\rho+\beta}\left(\rho \begin{bmatrix} T(\mathbf{u}^{(i)}) & \mathbf{X}^{(i)} \\ (\mathbf{X}^{(i)})^{\mathrm{H}} & \mathbf{Q}^{(i)} \end{bmatrix} - \mathbf{\Lambda}^{(i)} + \beta\mathbf{Z}^{(i)} \right) \right\|_{\mathrm{F}}^2 \\
&= \operatorname{prox}_{\frac{\lambda}{\rho+\beta}\mathcal{R}_+(\cdot)}\left(\frac{1}{\rho+\beta}\left(\rho \begin{bmatrix} T(\mathbf{u}^{(i)}) & \mathbf{X}^{(i)} \\ (\mathbf{X}^{(i)})^{\mathrm{H}} & \mathbf{Q}^{(i)} \end{bmatrix} - \mathbf{\Lambda}^{(i)} + \beta\mathbf{Z}^{(i)} \right) \right). \qquad (5.28)
\end{aligned}$$

Introducing

$$\mathbf{Z}^{(i)} = \begin{bmatrix} \mathbf{Z}_{\mathrm{T}}^{(i)} & \mathbf{Z}_{\mathrm{X}}^{(i)} \\ (\mathbf{Z}_{\mathrm{X}}^{(i)})^{\mathrm{H}} & \mathbf{Z}_{\mathrm{Q}}^{(i)} \end{bmatrix}, \quad \mathbf{\Lambda}^{(i)} = \begin{bmatrix} \mathbf{\Lambda}_{\mathrm{T}}^{(i)} & \mathbf{\Lambda}_{\mathrm{X}}^{(i)} \\ (\mathbf{\Lambda}_{\mathrm{X}}^{(i)})^{\mathrm{H}} & \mathbf{\Lambda}_{\mathrm{Q}}^{(i)} \end{bmatrix},$$

$\mathbf{X}$, $\mathbf{u}$ and $\mathbf{Q}$ can be updated separately as

$$\begin{aligned}
\mathbf{X}^{(i+1)} &= \underset{\mathbf{X}}{\operatorname{argmin}}\ \iota_C(\mathbf{X}) - 2\left\langle \mathbf{\Lambda}_{\mathrm{X}}^{(i)}, \mathbf{X} \right\rangle_{\mathrm{F}} \\
&\quad + \rho\left\| \mathbf{Z}_{\mathrm{X}}^{(i+1)} - \mathbf{X} \right\|_{\mathrm{F}}^2 + \beta\left\| \mathbf{X} - \mathbf{X}^{(i)} \right\|_{\mathrm{F}}^2 \\
&= P_C\left(\frac{1}{\rho+\beta}\left(\rho\mathbf{Z}_{\mathrm{X}}^{(i+1)} + \mathbf{\Lambda}_{\mathrm{X}}^{(i)} + \beta\mathbf{X}^{(i)} \right) \right), \qquad (5.29)
\end{aligned}$$

$$\begin{aligned}
\mathbf{u}^{(i+1)} &= \underset{\mathbf{u}}{\operatorname{argmin}}\ -\left\langle \mathbf{\Lambda}_{\mathrm{T}}^{(i)}, T(\mathbf{u}) \right\rangle_{\mathrm{F}} + \frac{\rho}{2}\left\| \mathbf{Z}_{\mathrm{T}}^{(i+1)} - T(\mathbf{u}) \right\|_{\mathrm{F}}^2 \\
&\quad + \frac{\beta}{2}\left\| T(\mathbf{u}) - T(\mathbf{u}^{(i)}) \right\|_{\mathrm{F}}^2 \\
&= T^{\dagger}\left(\frac{1}{\rho+\beta}\left(\rho\mathbf{Z}_{\mathrm{T}}^{(i+1)} + \mathbf{\Lambda}_{\mathrm{T}}^{(i)} + \beta T(\mathbf{u}^{(i)}) \right) \right), \qquad (5.30)
\end{aligned}$$

$$\begin{aligned}
\mathbf{Q}^{(i+1)} &= \underset{\mathbf{Q}}{\operatorname{argmin}}\ -\left\langle \mathbf{\Lambda}_{\mathrm{Q}}^{(i)}, \mathbf{Q} \right\rangle_{\mathrm{F}} + \frac{\rho}{2}\left\| \mathbf{Z}_{\mathrm{Q}}^{(i+1)} - \mathbf{Q} \right\|_{\mathrm{F}}^2 \\
&\quad + \frac{\beta}{2}\left\| \mathbf{Q} - \mathbf{Q}^{(i)} \right\|_{\mathrm{F}}^2 \\
&= \frac{1}{\rho+\beta}\left(\rho\mathbf{Z}_{\mathrm{Q}}^{(i+1)} + \mathbf{\Lambda}_{\mathrm{Q}}^{(i)} + \beta\mathbf{Q}^{(i)} \right), \qquad (5.31)
\end{aligned}$$

respectively. The projection $P_C(\cdot)$ in the update for $\mathbf{X}$ is given by

$$P_C(\mathbf{X}) = \mathbf{X} - \mathcal{B}^*(\mathcal{B}\mathcal{B}^*)^{-1}(\mathcal{B}(\mathbf{X}) - \mathbf{y}). \qquad (5.32)$$

The linear operator $\mathcal{B}(\cdot)$ can be rewritten as the product of a matrix $\mathbf{B} \in \mathbb{C}^{L \times NM}$ and the vectorization of $\mathbf{X}$

$$\mathcal{B}(\mathbf{X}) = \mathbf{B}\mathrm{vec}(\mathbf{X}). \tag{5.33}$$

Then, its adjoint $\mathcal{B}^*$ and $(\mathcal{B}\mathcal{B}^*)^{-1}$ are calculated as

$$\mathcal{B}^*(\mathbf{v}) = \mathrm{mat}_{N,M}\left(\mathbf{B}^{\mathrm{H}}(\mathbf{v})\right) \tag{5.34}$$

$$(\mathcal{B}\mathcal{B}^*)^{-1}(\mathbf{v}) = (\mathbf{B}\mathbf{B}^{\mathrm{H}})^{-1}\mathbf{v}, \tag{5.35}$$

where $\mathrm{mat}_{N,M} : \mathbb{C}^{NM} \to \mathbb{C}^{N \times M}$ is a matricization operator. $T^\dagger : \mathbb{C}^{L \times L} \to \mathbb{C}^L$ in the update for $\mathbf{u}$ represents the pseudo-inverse operator of T,

$$T^\dagger(\mathbf{X})[n] = \frac{1}{2(L-n)} \sum_{k=0}^{L-n-1} \left(\mathbf{X}[k, k+n] + \overline{\mathbf{X}[k+n, k]}\right).$$

Algorithm 2 summarizes the algorithm for solving (5.25).

The augmented Lagrangian associated with (5.26) is similar to that of (5.25). Hence, The algorithm for solving (5.26) differs only in the update of $\mathbf{X}$, which is given by

$$\begin{aligned}
\mathbf{X}^{(i+1)} &= \underset{\mathbf{X}}{\mathrm{argmin}}\ \frac{1}{2}\left\|\mathcal{B}(\mathbf{X}) - \mathbf{y}\right\|_2^2 - 2\left\langle \mathbf{\Lambda}_{\mathrm{X}}^{(i)}, \mathbf{X}\right\rangle_{\mathrm{F}} \\
&\quad + \rho\left\|\mathbf{Z}_{\mathrm{X}}^{(i+1)} - \mathbf{X}\right\|_{\mathrm{F}}^2 + \beta\left\|\mathbf{X} - \mathbf{X}^{(i)}\right\|_{\mathrm{F}}^2 \\
&= (\mathcal{B}^*\mathcal{B} + 2(\rho+\beta)\mathcal{I})^{-1} \\
&\quad \cdot \left(\mathcal{B}^*(\mathbf{y}) + 2\left(\rho\mathbf{Z}_{\mathrm{X}}^{(i+1)} + \mathbf{\Lambda}_{\mathrm{X}}^{(i)} + \beta\mathbf{X}^{(i)}\right)\right),
\end{aligned} \tag{5.36}$$

where $\mathcal{I}$ is the identity operator. As in (5.34) and (5.35), $(\mathcal{B}^*\mathcal{B} + 2(\rho+\beta)\mathcal{I})^{-1}$ is computed as

$$\begin{aligned}
&(\mathcal{B}^*\mathcal{B} + 2(\rho+\beta)\mathcal{I})^{-1}(\mathbf{X}) \\
&= \mathrm{mat}_{N,M}\left(\left(\mathbf{B}^{\mathrm{H}}\mathbf{B} + 2(\rho+\beta)\mathbf{I}_{NM}\right)^{-1}\mathrm{vec}(\mathbf{X})\right).
\end{aligned} \tag{5.37}$$

The algorithm for solving (5.26) is summarized in Algorithm 3.

The algorithms are applicable if the proximal operator of $\mathcal{R}_+(\cdot)$ can be computed. Convergence of the proximal ADMM in nonconvex optimization problems has been discussed [147], but convergence of this problem is not guaranteed. For weakly convex $\mathcal{R}_+(\mathbf{Z})$, the convergence of the proximal augmented Lagrangian

Algorithm 2 Proximal ADMM for solving (5.25)

Input: $\mathcal{B}$, $\mathbf{y}$, and ϵ, ρ

Output: $\mathbf{X}$, $\mathbf{u}$, $\mathbf{Q}$

Initialize $\mathbf{Z}^{(0)}$, $\mathbf{X}^{(0)}$, $\mathbf{u}^{(0)}$, $\mathbf{Q}^{(0)}$, and $\mathbf{\Lambda}^{(0)}$

for $i = 0, 1, \cdots$ **do**

Calculate $\mathbf{Z}^{(i+1)}$ by (5.28)

Calculate $\mathbf{X}^{(i+1)}$ by (5.29)

Calculate $\mathbf{u}^{(i+1)}$ by (5.30)

Calculate $\mathbf{Q}^{(i+1)}$ by (5.31)

$$\mathbf{\Lambda}^{(i+1)} \leftarrow \mathbf{\Lambda}^{(i)} + \rho \left(\mathbf{Z}^{(i+1)} - \begin{bmatrix} T(\mathbf{u}^{(i+1)}) & \mathbf{X}^{(i+1)} \\ (\mathbf{X}^{(i+1)})^{\mathrm{H}} & \mathbf{Q}^{(i+1)} \end{bmatrix} \right)$$

end for

method is guaranteed [148] [1]. The discussion of the convergence of the proposed method is a future issue.

5.5 Numerical experiments

This section compares the proposed method with ANM and RAM [130]. RAM consists of repeatedly solving the following SDP problem:

$$\begin{aligned} \underset{\mathbf{X},\mathbf{u},\mathbf{Q}}{\text{minimize}} \quad & \mathrm{Tr}\left(\mathbf{W}^{(i)} T(\mathbf{u})\right) + \mathrm{Tr}(\mathbf{Q}) \\ \text{subject to} \quad & \mathbf{Z} = \begin{bmatrix} T(\mathbf{u}) & \mathbf{X} \\ \mathbf{X}^{\mathrm{H}} & \mathbf{Q} \end{bmatrix} \succeq 0 \\ & \|\mathcal{B}(\mathbf{X}) - \mathbf{y}\|_2 \leq \eta, \end{aligned} \tag{5.38}$$

where $\mathbf{W}^{(i)} = \left(T(\mathbf{u}^{(i)}) + \varepsilon \mathbf{I}_N\right)^{-1}$ is a weight matrix constructed by the solution of the previous iteration $\mathbf{u}^{(i)}$. The first iteration was computed with $\mathbf{u}^{(0)} = \mathbf{0}_N$ and $\varepsilon = 1$. Then, ε was halved at the beginning of the next iteration until $\varepsilon = \frac{1}{2^{10}}$ or $\varepsilon < \frac{L\eta^2}{10\|\mathbf{y}\|_2^2}$. We used CVX [149] with SDP3 [150] to solve (5.38).

The nonconvex penalties shown in Table 5.1 and Fig. 5.1 were used in the proposed method. Each parameter of the nonconvex penalties was set to $a = 1/\lambda$

[1] Precisely, [148] proposes the Moreau envelope augmented Lagrangian (MEAL) method and discusses its convergence. The MEAL method is identical to the proximal augmented Lagrangian method with a certain parameter.

Algorithm 3 Proximal ADMM for solving (5.26)

Input: $\mathcal{B}$, $\mathbf{y}$, ρ, and β

Output: $\mathbf{X}$, $\mathbf{u}$, $\mathbf{Q}$

Initialize $\mathbf{Z}^{(0)}$, $\mathbf{X}^{(0)}$, $\mathbf{u}^{(0)}$, $\mathbf{Q}^{(0)}$, and $\mathbf{\Lambda}^{(0)}$

for $i = 0, 1, \cdots$ **do**

Calculate $\mathbf{Z}^{(i+1)}$ by (5.28)

Calculate $\mathbf{X}^{(i+1)}$ by (5.36)

Calculate $\mathbf{u}^{(i+1)}$ by (5.30)

Calculate $\mathbf{Q}^{(i+1)}$ by (5.31)

$$\mathbf{\Lambda}^{(i+1)} \leftarrow \mathbf{\Lambda}^{(i)} + \rho \left(\mathbf{Z}^{(i+1)} - \begin{bmatrix} T(\mathbf{u}^{(i+1)}) & \mathbf{X}^{(i+1)} \\ (\mathbf{X}^{(i+1)})^{\mathrm{H}} & \mathbf{Q}^{(i+1)} \end{bmatrix} \right)$$

end for

for the logarithmic penalty, $\mu = 3.7$ for SCAD, $\gamma = 3$ for MCP, and $\theta = 2\lambda$ for the capped ℓ_1 norm. The parameters in Algorithm 2 and 3 were set to $\rho = 1$ and $\beta = 1$. Since the optimization problems (5.25) and (5.26) in the proposed method are non-convex, there is a dependency on the intial value. We set the initial values to the zero matrices or the solution of ANM.

5.5.1 Sparsity separation and phase transition

We conducted an experiment to evaluate the separation performance of sinusoids in noise-free situations. The proposed method was applied to estimate the line spectrum of the sum of K sinusoids whose frequencies were separated by at least Δ_ω. Let $\mathcal{B}$ in (5.1) be a sampling operator that samples L elements from a signal of length N. The sampling index is generated uniformly at random. If the frequency is correctly estimated, the entire signal can be reconstructed. If the error between the unsampled signal and the estimated signal $\|\mathbf{S} - \mathbf{X}\|_{\mathrm{F}}^2$ is less than 10^{-6}, this estimation result is treated as a success. The success rate of each combination (K, Δ_ω) was calculated during ten runs. The size of the signal was set to $N = 64$, $M = 1$, and $L = 30$.

Fig. 5.3 shows the success rates of ANM, RAM, and the proposed method with the different regularizations when the initial values were set to the zero matrices. The proposed method with the $\ell_{1/2}$ norm and the logarithmic penalty succeeded in more conditions than the ANM. However, from Fig. 5.3, the proposed method

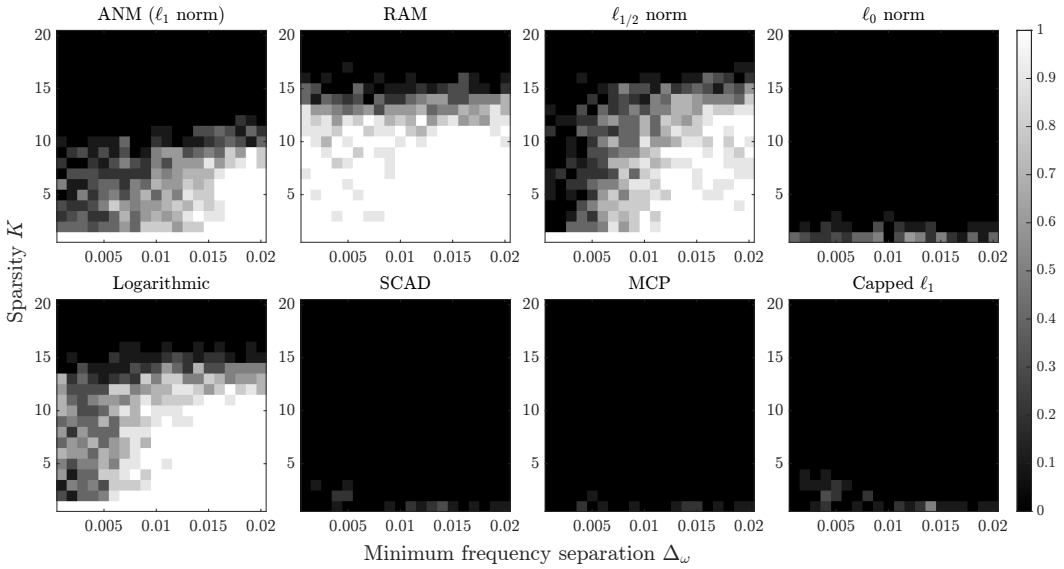

Figure 5.3: Success rates of ANM, RAM, and the proposed method with the different regularizations when the initial values were set to the zero matrices.

with the other regularizations rarely estimated correctly. Fig. 5.4 shows the success rates when the initial values were set to the estimates by the ANM. Note that the ANM shows the same result in Fig. 5.3. Furthermore, the first iteration of the RAM matches the ANM, so the RAM also shows the equivalent performance in Fig. 5.3. By setting the initial value to the estimations by the ANM, the proposed method performed better than when the initial values were set to the zero matrices and the ANM. The results suggest that incorporating nonconvex regularizations into the gridless sparse method improves the performance of line spectrum estimation. On the other hand, compared to the RAM, the performance issues remain when minimum frequency separation Δ_ω is small.

5.5.2 Line spectrum denoising

Next, the proposed method was used to estimate the sum of the sinusoids $\mathbf{S}$ obtained from a noisy measurement $\mathbf{y}$. The observation signal was generated by adding complex Gaussian noise with a mean of 0 and a variance of σ^2. This variance of σ^2 was chosen so that the signal-to-noise ratio (SNR)

$$\mathrm{SNR} = \frac{\|\mathbf{S}\|_{\mathrm{F}}^2}{MN\sigma^2} \tag{5.39}$$

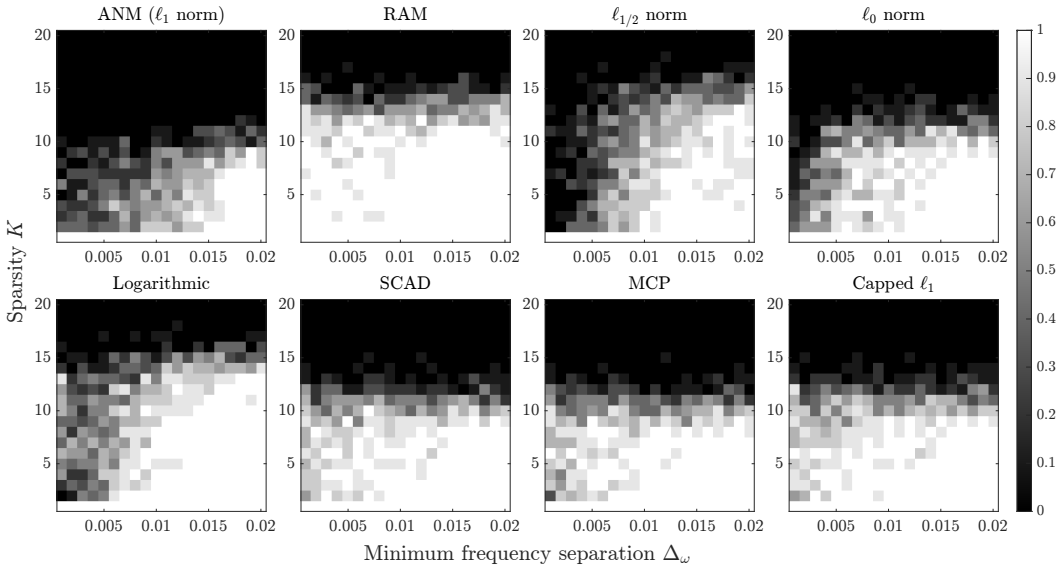

Figure 5.4: Success rates of ANM, RAM, and the proposed method with the different regularizations when the initial values were set to the estimations by ANM.

equals $0, 5, \ldots, 40$ dB. The estimation was performed 100 times for each SNR, and the average of error $\|\mathbf{S} - \mathbf{X}\|_{\mathrm{F}}^2$ was evaluated. The number of signal samples is $N = L = 24$, $M = 1$, and $\mathcal{B}(\cdot)$ is the identity operator $\mathcal{I}$. That is, $\mathbf{X}$ was estimated by solving

$$\begin{aligned} \underset{\mathbf{Z},\mathbf{u},\mathbf{X},\mathbf{Q}}{\text{minimize}} \quad & \lambda \mathcal{R}(\mathbf{Z}) + \frac{1}{2} \|\mathbf{X} - \mathbf{y}\|_{\mathrm{F}}^2 \\ \text{subject to} \quad & \mathbf{Z} = \begin{bmatrix} T(\mathbf{u}) & \mathbf{X} \\ \mathbf{X}^{\mathrm{H}} & \mathbf{Q} \end{bmatrix} \succeq 0. \end{aligned} \tag{5.40}$$

The regularization parameter λ was chosen to minimize the resulting averaged error.

The estimation results obtained at SNR $= 20$ when the initial values were set to the zero matrices are shown in Fig. 5.5. ($\bigcirc$) and ($\times$) represent the ground truth and the estimated parameters, respectively. In the results of the ANM, the amplitude tends to be smaller than the ground truth because of the effect of including bias. The RAM and the proposed method maintain the original amplitude. Furthermore, the estimates using the proposed method have less variation than the estimate of the RAM.

Then, Fig. 5.6 plots the averaged errors induced at each SNR. For SNR greater than 20 dB, the proposed methods outperformed the ANM and RAM, indicating that the original signal can be estimated with high accuracy by suppressing the

bias. On the other hand, for SNR below 15 dB, the proposed methods using some regularizations had worse performance than ANM. This is thought to be due to the sensitivity to noise caused by nonconvexity. The nonconvex regularizations, except for the ℓ_0 norm, include a parameter that adjusts for nonconvexity. Hence, choosing the parameter appropriately could improve the performance of the proposed method. In fact, when the parameter θ in the capped ℓ_1 is chosen large enough, the estimation results of the capped ℓ_1 equals those of the ℓ_1 norm. The results suggest that the proposed method can improve the performance of gridless sparse optimization on noisy data.

The estimation results and average error at SNR = 20 when the initial values are estimated by ANM are illustrated in Figs. 5.7 and 5.8. Compared to Sec. 5.5.1, the proposed method works well in both cases of initial values.

5.6 Conclusion

In this chapter, we propose a framework to introduce nonconvex regularizations for gridless sparse estimation. The proposed method reduces bias by replacing the trace in the sparse estimation method (which uses the atomic norm technique) with a nonconvex function that induces low rankness in the given matrix. The proposed method outperformed the conventional ANM approach in a frequency separation experiment. Furthermore, the proposed method performed better than ANM and RAM in a line spectrum denoising at SNR greater than 20 dB.

The proposed method employs the proximal ADMM to solve the optimization problem, but its convergence is not guaranteed. Future work will include a discussion about the convergence of the algorithm. Future work will also include conducting a theoretical analysis to clarify the relationships between the coefficients of sinusoids and the proposed cost functions.

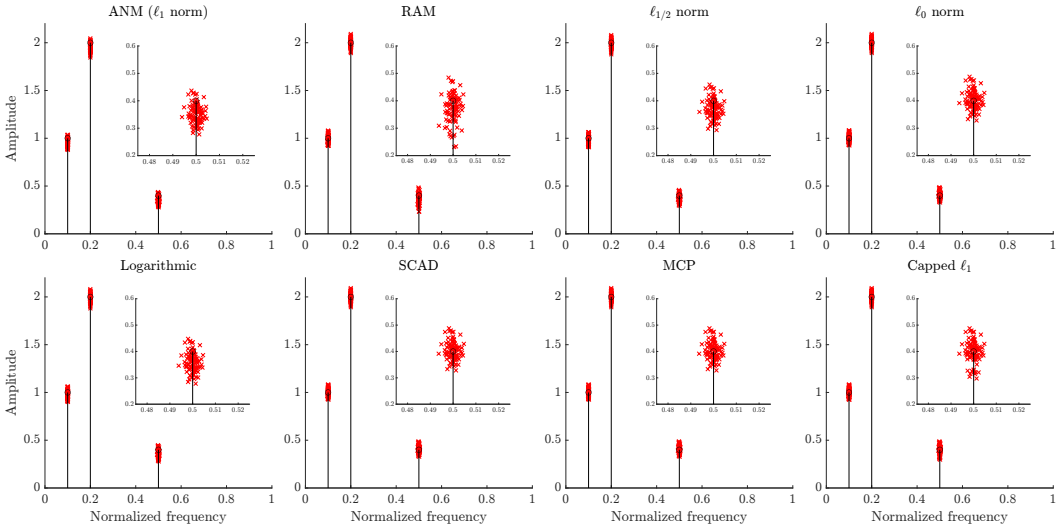

Figure 5.5: Results of line spectrum denoising for SNR = 20 when the initial values were set to the zero matrices. (◯) and (×) represent the ground truth and the estimated parameters, respectively. The upper right subfigure in each figure is an enlargement of the part in which the frequency $\omega = 0.5$.

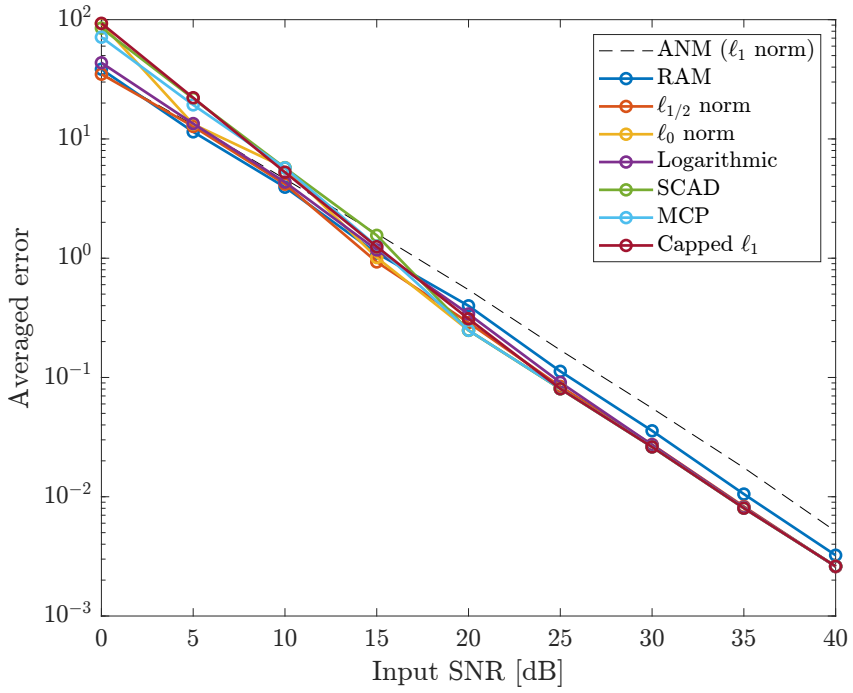

Figure 5.6: Avaraged errors of line spectrum denoising when the initial values were set to the zero matrices.

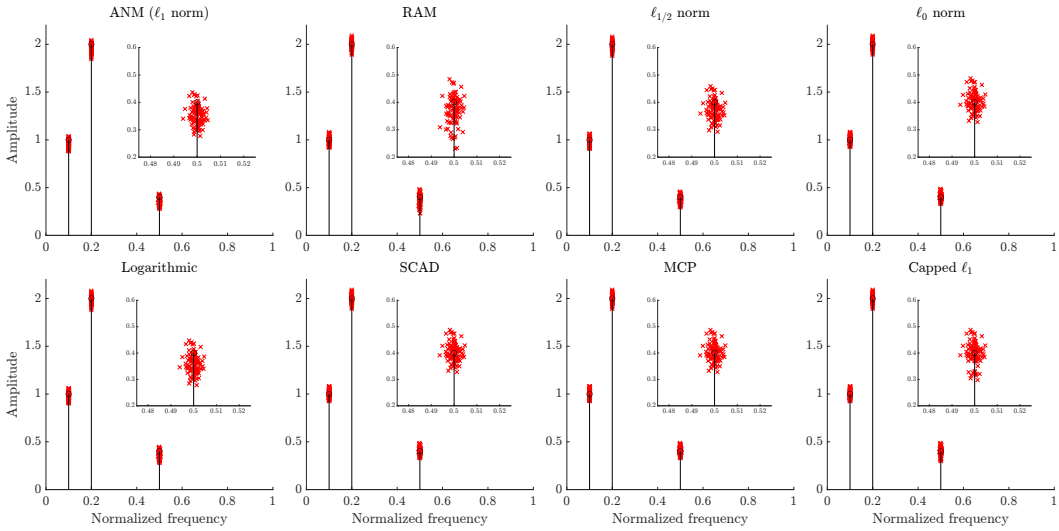

Figure 5.7: Results of line spectrum denoising for SNR = 20 when the initial values were set to the estimates by the ANM. (◯) and (×) represent the ground truth and the estimated parameters, respectively. The upper right subfigure in each figure is an enlargement of the part in which the frequency $\omega = 0.5$.

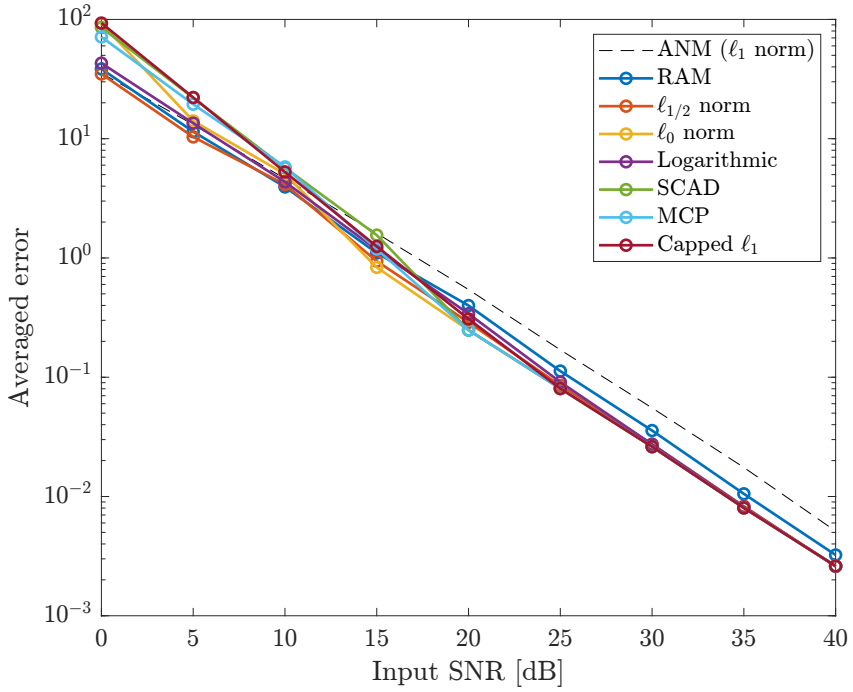

Figure 5.8: Avaraged errors of line spectrum denoising when the initial values were set to the estimates by the ANM.

Chapter 6

Conclusions

6.1 Summary

T-F representations are essential in nonstationary signal analysis and processing. The short-time Fourier transform (STFT) is widely used to convert a signal into the T-F domain owing to its simplicity and well-understood structure. The resolution of a T-F representation obtained by the STFT is limited in Heisenberg's uncertainty principle. In this thesis, we explored methods for estimating sparse T-F representations based on STFT and proposed the methods for improving them.

In Chapter 3, we proposed a framework to design a window function for computing the IF of STFT. Computing the IF requires the STFT with a window and STFT with its derivative, i.e., the IF computation depends on both the window function and its derivative. To design a window suitable for computing the IF, we formulated the window design problem as a sidelobe minimization problem of the corresponding derivative. Two windows were designed considering the sidelobe energy or the highest sidelobe level as cost functions to minimize the sidelobes of their derivatives. The SST using the proposed window provided a sharper T-F representation compared with those produced using ordinary bandwidth-adjustable windows.

In Chapter 4, we proposed a method of estimating a sparse T-F representation using the atomic norm. The atomic norm enables sparse optimization without the discretization of continuous parameters. Introducing the atomic norm avoids the effects of grid mismatch and allows to estimate sparser T-F representation than the method using the ℓ_1 norm.

In Chapter 5, we proposed a framework for introducing non-convex penalties into the gridless sparse optimization method. The proposed method aims to reduce the bias that reduces the amplitude of the estimation. It was realized by replacing the trace in sparse estimation using the atomic norm with a nonconvex function that induces a lower-ranked matrix. Numerical experiments confirmed that the proposed method performs better than the conventional ANM-based method.

6.2 Contribution to intermedia studies

Nonstationary signals are commonly difficult to understand visually, and they should be converted into some representations. Converting these signals into a T-F domain is effective to grasp the signal intuitively. Classical T-F representations have limited resolutions and sparsity, which may disturb understanding of the signal. The sparse T-F representation provides higher resolution and aids in further understanding of the signal. Furthermore, sparse T-F representation is also useful in signal processing. Our study has proposed the methods to obtain better sparse T-F representations. This contributes to people grasping the signal intuitively.

6.3 Future remarks

Although the methods to estimate sparse T-F representations including the reassignment method, SST, and sparsity-aware methods have been proposed, most audio signal processings are still performed using the STFT. In this thesis proposed the methods improving the T-F representations. These works will contribute to the expanded applications of sparse T-F representations and more accurate signal processing using them. There are several remaining works to improve their performance and expand their applications.

6.3.1 T-F representation via nonconvex gridless sparse method

In Chapter 4, we proposed a method of estimating a sparse T-F representation using the atomic norm. The gridless sparse method with nonconvex regularization was proposed in Chapter 5. These methods can be combined and expected to obtain a sparser representation. Recall that the estimation of a sparse T-F representation

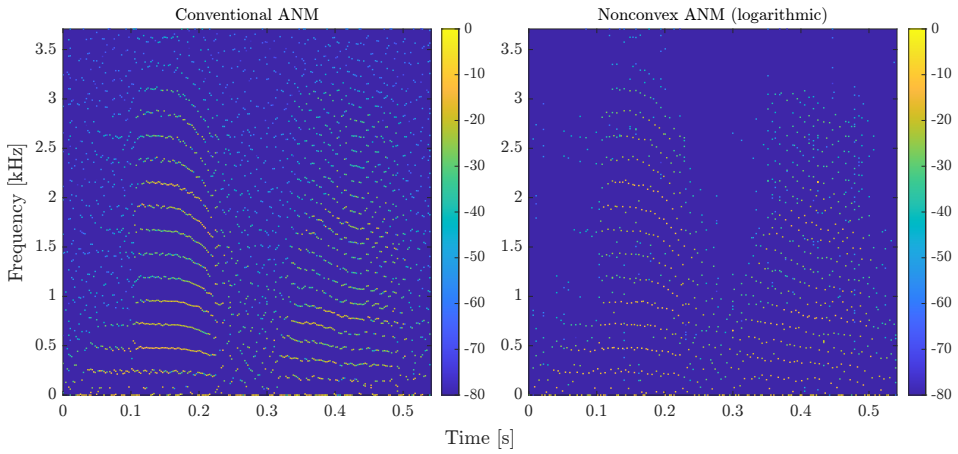

Figure 6.1: T-F representations of a speech signal using gridless sparse methods. The left and right sides show the T-F representations obtained using the atomic norm (in Chapter 4) and the nonconvex gridless sparse method, respectively.

using ANM is formulated as

$$\underset{\mathbf{x}}{\text{minimize}} \quad \sum_{n=0}^{N-1} \|\mathbf{x}_n\|_{\mathcal{A}} \quad \text{subject to} \quad \mathbf{f} = \mathbf{A}_{\mathbf{g}}\mathbf{x}. \tag{6.1}$$

By replacing $\|\cdot\|_{\mathcal{A}}$ in (6.1) with nonconvex regularization $\mathscr{R}(\cdot)$ in (5.22), estimating a sparse T-F representation via nonconvex gridress sparse method is formulated as

$$\underset{\mathbf{x}}{\text{minimize}} \quad \sum_{n=0}^{N-1} \mathscr{R}\left(\mathbf{x}_n\right) \quad \text{subject to} \quad \mathbf{f} = \mathbf{A}_{\mathbf{g}}\mathbf{x}. \tag{6.2}$$

The two methods (4.11) and (6.2) were applied to a speech signal for comparing these methods. The Slepian window [36] was chosen as a window $\mathbf{g}$, and its length and bandwidth were set to 2^7 and 0.04. The time-shifting width and the number of frequency channels were set to $a = 2^4$ and $M = 2^{10}$. The logarithmic penalty was used as a non-convex regularization $\mathscr{R}(\cdot)$.

The estimated T-F representations are displayed in Fig. 6.1. The T-F representation estimated using the nonconvex regularization appears to be more sparse. However, it has discontinuities in the harmonics, which would disturb signal analysis. Sparse optimization that takes into account the relationship in the time direction should be considered to solve this problem.

6.3.2 Generalization for two-dimensional signals

In this thesis, we have discussed STFT for one-dimensional signals such as acoustic signals. STFT is employed in the analysis and processing of two-dimensional signals as well as one-dimensional signals [151–153]. This section shows the application of sparse T-F representation to fringe restoration as an example for two-dimensional signals.

Optical interferometry is a method for measuring the phase difference using light interference with a reference light. The observed fringe pattern can be expressed as

$$f(\mathbf{x}) = a(\mathbf{x}) + b(\mathbf{x})\cos(\phi(\mathbf{x})) + n(\mathbf{x}), \tag{6.3}$$

where $\mathbf{x} = [x_1, x_2]^\top$ is the position vector, a is the background intensity, b is the amplitude of the fringe, ϕ is the phase of the fringe and n is additive noise.

Two-dimensional STFT, often called the windowed Fourier transform (WFT), has been used to extract the fringe $b(\mathbf{x})\cos(\phi(\mathbf{x}))$ from the observation $f(\mathbf{x})$. Applying the WFT to the observation image, the components of the fringe will be localized. On the other hand, the noise components are spread over the entire domain. Thus, the fringe can be restored by truncating values below a preset threshold. This process is referred to as the windowed Fourier filtering (WFF). As in the case of one-dimensional STFT, the representation of the WFT is spread depending on the choice of a window function, which affects the performance of the WFF. SST for two-dimensional signals has been studied in several literatures. Introducing the SST in the WFF can localize the components of fringe and will improve the restoration performance.

We simulated the data measured from interferometry and compared with WFF and SST-based methods in the fringe restoration. A window function of WFT and SST is designed as

$$g(\mathbf{x}) = w(x_1)w(x_2) \tag{6.4}$$

where w is a window with length $N = 32$ and bandwidth $W = 0.1$ designed to minimize the sidelobe energy of its derivative shown in Chapter 3.

The simulated data and the frequency representation obtained by WFT and SST are shown in Fig. 6.2. From Fig. 6.2, the frequency representation obtained by the SST is more localized than that obtained by the WFT. The fringe patterns restored by the WFF and filtering with the SST, are shown in Fig. 6.3. The fringe

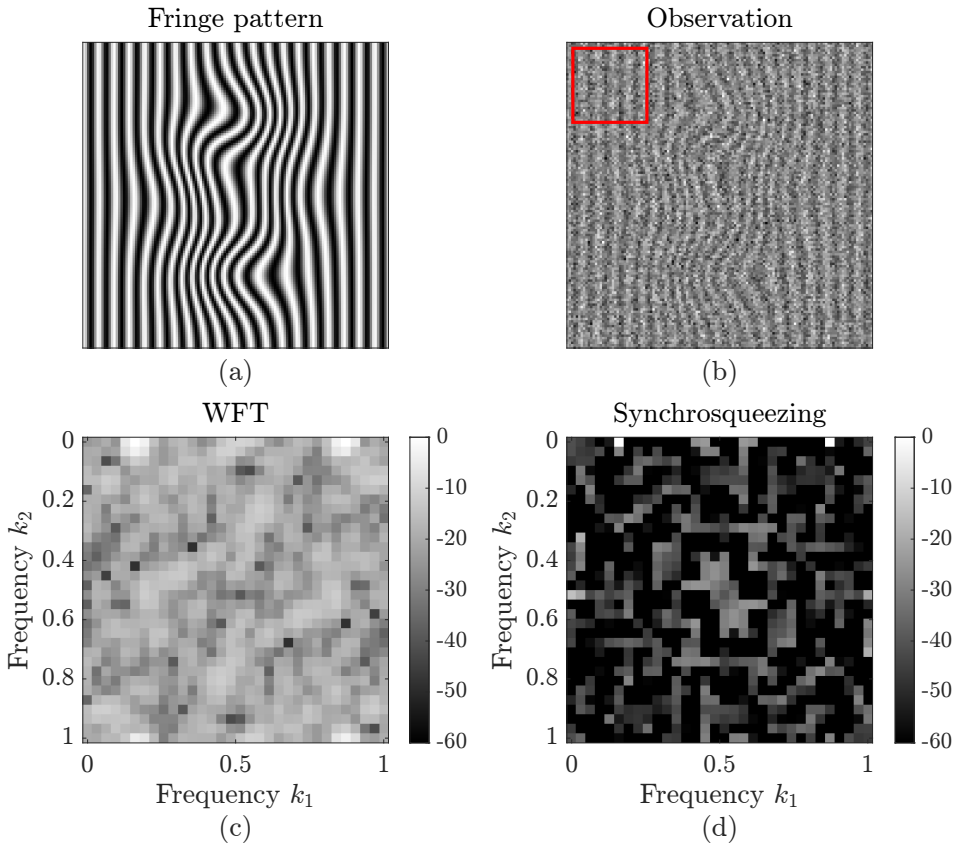

Figure 6.2: Two-dimensional SST of the noisy fringe data. (a) is the fringe pattern, (b) is the observed data generated by adding Gaussian noise to the fringe, (c) is the frequency representation obtained by windowing the red frame in (b), and (d) is the frequency representation obtained by SST of (c).

restored by the WFF contains oscillations that are not included in the true fringe (Fig. 6.2(a)). On the other hand, the fringe patterns restored using the SST contains no such oscillations. The result suggests that filtering with the SST can removed the high-frequency noise that could not be removed by the WFF since the SST localize the components of the fringe.

6.3.3 Application of frequencies obtained with continuous values

The observed data in signal processing is generally a discrete signal, and its T-F representation obtained by STFT is discretized in the frequency direction as well. Thus, signal processing methods with STFT are designed to handle such discretized T-F representations. By contrast, the sparse T-F representation obtained by reassignment, SST, and the sparse optimization using ANM have a continuous

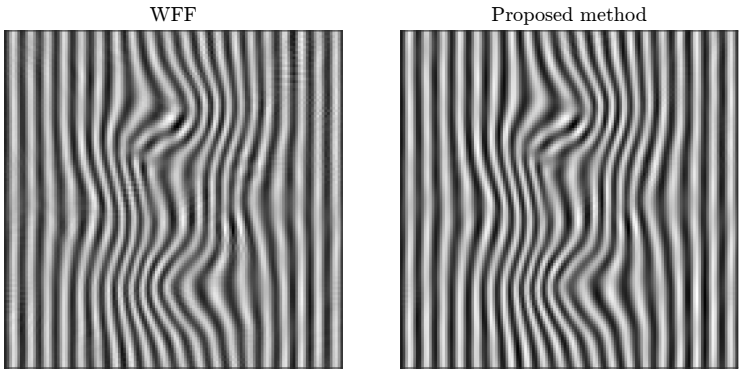

Figure 6.3: Result of fringe restorations. The left and right sides illustrate the fringes restored by the WFF and filtering with the SST, respectively.

frequency parameter. It is not easy to apply the processing methods used in STFT to these T-F representations. For ease to handle, the T-F representation of SST are often rounded to the same grid as STFT. This does not take advantage of all the information that is originally contained. The consideration of the signal processing methods using continuous frequency parameter may establish better signal processing methods.

Appendix A

Trigonometric representation of the frequency response of the window derivative

This appendix explains the derivation of (3.43) and the relation between $\boldsymbol{\alpha}$ and $\mathbf{z}$ based on [86]. The frequency response of the antisymmetric window $\mathbf{z}$ can be represented as a sine series. When N is odd,

$$(\mathcal{F}\mathbf{z})(\omega) = \mathrm{e}^{\mathrm{i}\left(\frac{\pi}{2}-(N-1)\pi\omega\right)} \sum_{k=1}^{K} \mathbf{a}[k] \sin(2\pi\omega k), \tag{A.1}$$

where $K = (N-1)/2$ and $\mathbf{a}[k] = 2\mathbf{z}[K-k]$ for $k = 1, \ldots, K$, and $\mathbf{z}[K] = 0$. When N is even,

$$(\mathcal{F}\mathbf{z})(\omega) = \mathrm{e}^{\mathrm{i}\left(\frac{\pi}{2}-(N-1)\pi\omega\right)} \sum_{k=1}^{K} \mathbf{a}[k] \sin\left(2\pi\omega\left(k-\frac{1}{2}\right)\right) \tag{A.2}$$

where $K = N/2$ and $\mathbf{a}[k] = 2\mathbf{z}[K-k]$ for $k = 1, \ldots, K$. (A.1) and (A.2) can be rewritten as a cosine series using the trigonometric identities,

$$\cos\theta_1 \cos\theta_2 = \frac{1}{2}\left[\cos(\theta_1+\theta_2) + \cos(\theta_1-\theta_2)\right], \tag{A.3}$$

$$\cos\theta_1 \sin\theta_2 = \frac{1}{2}\left[\sin(\theta_1+\theta_2) - \sin(\theta_1-\theta_2)\right]. \tag{A.4}$$

When window length N is odd,

$$\sum_{k=1}^{K} \mathbf{a}[k] \sin(2\pi\omega k) = \sin(2\pi\omega) \sum_{k=0}^{K-1} \boldsymbol{\alpha}[k] \cos(2\pi\omega k), \tag{A.5}$$

where

$$\mathbf{a}[k] = \begin{cases} \boldsymbol{\alpha}[0] - \frac{1}{2}\boldsymbol{\alpha}[2] & \text{if } k = 1 \\ \frac{1}{2}(\boldsymbol{\alpha}[k-1] - \boldsymbol{\alpha}[k+1]) & \text{if } k = 2, 3, \ldots, K-2 \\ \frac{1}{2}\boldsymbol{\alpha}[k-1] & \text{if } k = K-1, K \end{cases} . \tag{A.6}$$

When window length N is even,

$$\sum_{k=1}^{K} \mathbf{a}[k] \sin\left(2\pi\omega\left(k - \frac{1}{2}\right)\right) = \sin(\pi\omega) \sum_{k=0}^{K-1} \boldsymbol{\alpha}[k] \cos(2\pi\omega k), \tag{A.7}$$

where

$$\mathbf{a}[k] = \begin{cases} \boldsymbol{\alpha}[0] - \frac{1}{2}\boldsymbol{\alpha}[1] & \text{if } k = 1 \\ \frac{1}{2}(\boldsymbol{\alpha}[k-1] - \boldsymbol{\alpha}[k]) & \text{if } k = 2, 3, \ldots, K-1 \\ \frac{1}{2}\boldsymbol{\alpha}[k-1] & \text{if } k = K \end{cases} . \tag{A.8}$$

Hence, using (3.44) and (3.45), $(\mathcal{F}\mathbf{z})(\omega)$ in the odd and even cases can be uniformly rewritten as

$$(\mathcal{F}\mathbf{z})(\omega) = \mathrm{e}^{\mathrm{i}\left(\frac{\pi}{2} - (N-1)\pi\omega\right)} Q(\omega) P(\omega, \boldsymbol{\alpha}), \tag{A.9}$$

which equals (3.43).

After obtaining the solution $\boldsymbol{\alpha}^\star$ by solving (3.48), the window derivative $\mathbf{z}$ is calculated using (A.6) and

$$\mathbf{z}[n] = \begin{cases} \frac{1}{2}\mathbf{a}[K-n] & \text{for } n = 0, \ldots, K-1 \\ 0 & \text{for } n = K \\ -\frac{1}{2}\mathbf{a}[n-K] & \text{for } n = K+1, \ldots, N-1 \end{cases} \tag{A.10}$$

when N is odd, and (A.8) and

$$\mathbf{z}[n] = \begin{cases} \frac{1}{2}\mathbf{a}[K-n] & \text{for } n = 0, \ldots, K-1 \\ -\frac{1}{2}\mathbf{a}[n-K] & \text{for } n = K, \ldots, N-1 \end{cases} \tag{A.11}$$

when N is even.

Appendix B

Derivation of (3.74)

Let $\rho \in \mathbb{R}_+$ and $\varphi \in (-\pi, \pi]$ such that

$$r\mathrm{e}^{\mathrm{i}\Delta\phi}\frac{(\mathcal{F}\mathbf{g})(m/L - \xi_{\mathbf{s}})}{(\mathcal{F}\mathbf{g})(m/L - \xi_{\boldsymbol{i}})} =: \rho\mathrm{e}^{\mathrm{i}\varphi}. \tag{B.1}$$

Then, (3.72) can be rewritten as

$$e = \Re\left\{\frac{(\mathrm{IF}^{\boldsymbol{i}}_{\mathbf{g}}[n,m] - \mathrm{IF}^{\mathbf{s}}_{\mathbf{g}}[n,m])(\rho\mathrm{e}^{-\mathrm{i}\varphi}+1)}{(\rho\mathrm{e}^{\mathrm{i}\varphi}+1)(\rho\mathrm{e}^{-\mathrm{i}\varphi}+1)}\right\}. \tag{B.2}$$

Since $\mathrm{IF}^{\mathbf{s}}_{\mathbf{g}}$ and $\mathrm{IF}^{\boldsymbol{i}}_{\mathbf{g}}$ are real,

$$e = \frac{(\mathrm{IF}^{\boldsymbol{i}}_{\mathbf{g}}[n,m] - \mathrm{IF}^{\mathbf{s}}_{\mathbf{g}}[n,m])(\rho\cos\varphi+1)}{\rho^2+2\rho\cos\varphi+1}. \tag{B.3}$$

When e becomes extrema, φ must satisfy

$$\frac{\partial e}{\partial\varphi} = \frac{(\mathrm{IF}^{\boldsymbol{i}}_{\mathbf{g}}[n,m] - \mathrm{IF}^{\mathbf{s}}_{\mathbf{g}}[n,m])\rho(1-\rho^2)\sin\varphi}{(\rho^2+2\rho\cos\varphi+1)^2} = 0. \tag{B.4}$$

Because we assumed that $|(\mathcal{F}\mathbf{g})(m/L - \xi_{\mathbf{s}})/(\mathcal{F}\mathbf{g})(m/L - \xi_{\boldsymbol{i}}))| > 1/r$, ρ must be greater than 1. Thus, the denominator of (B.4) is not zero, and $\frac{\partial e}{\partial\varphi}$ becomes zero if $\varphi = 0$ or $\varphi = \pi$. Substituting $\varphi = 0$ or $\varphi = \pi$ for (B.2),

$$\begin{aligned} e_0 &= \frac{(\mathrm{IF}^{\boldsymbol{i}}_{\mathbf{g}}[n,m] - \mathrm{IF}^{\mathbf{s}}_{\mathbf{g}}[n,m])(\rho+1)}{\rho^2+2\rho+1} \\ &= \frac{\mathrm{IF}^{\boldsymbol{i}}_{\mathbf{g}}[n,m] - \mathrm{IF}^{\mathbf{s}}_{\mathbf{g}}[n,m]}{1+\rho}, \end{aligned} \tag{B.5}$$

$$\begin{aligned} e_\pi &= \frac{(\mathrm{IF}^{\boldsymbol{i}}_{\mathbf{g}}[n,m] - \mathrm{IF}^{\mathbf{s}}_{\mathbf{g}}[n,m])(-\rho+1)}{\rho^2-2\rho+1} \\ &= \frac{\mathrm{IF}^{\boldsymbol{i}}_{\mathbf{g}}[n,m] - \mathrm{IF}^{\mathbf{s}}_{\mathbf{g}}[n,m]}{1-\rho}. \end{aligned} \tag{B.6}$$

The difference between the squares of these is

$$\begin{aligned} e_0^2 - e_\pi^2 &= (\mathrm{IF}_{\mathbf{g}}^{\boldsymbol{i}}[n,m] - \mathrm{IF}_{\mathbf{g}}^{\mathbf{s}}[n,m])^2 \left\{ \frac{1}{(1+\rho)^2} + \frac{1}{(1-\rho)^2} \right\} \\ &= (\mathrm{IF}_{\mathbf{g}}^{\boldsymbol{i}}[n,m] - \mathrm{IF}_{\mathbf{g}}^{\mathbf{s}}[n,m])^2 \frac{-4\rho}{(1-\rho^2)^2} < 0. \end{aligned}$$

Hence, the worst-case error is e_π, and e_π is equal to (3.74) by substituting $\rho = r\,|(\mathcal{F}\mathbf{g})(m/L - \xi_{\mathbf{s}})/(\mathcal{F}\mathbf{g})(m/L - \xi_{\boldsymbol{i}})|$.

Bibliography

[1] D. Gabor, "Theory of communication. Part 1: The analysis of information," *J. Inst. Electr. Eng.*, vol. 93, no. 26, pp. 429–441, 1946.

[2] H. G. Feichtinger and T. Strohmer, Eds., *Gabor analysis and algorithms.* Boston, MA, USA: Birkhäuser Boston, 1998.

[3] K. Gröchenig, *Foundations of time-frequency analysis.* Boston, MA, USA: Birkhäuser Boston, 2001.

[4] B. Boashash, *Time-frequency signal analysis and processing: a comprehensive reference*, 2nd ed. London, UK: Academic Press, 2016.

[5] F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H.-T. Wu, "Time-frequency reassignment and synchrosqueezing: An overview," *IEEE Signal Process. Mag.*, vol. 30, no. 6, pp. 32–41, Nov. 2013.

[6] K. Kodera, C. De Villedary, and R. Gendrin, "A new method for the numerical analysis of non-stationary signals," *Phys. Earth Planet. Inter.*, vol. 12, no. 2–3, pp. 142–150, Aug. 1976.

[7] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1068–1089, May 1995.

[8] S. Fenet, R. Badeau, and G. Richard, "Reassigned time-frequency representations of discrete time signals and application to the constant-Q transform," *Signal Process.*, vol. 132, pp. 170–176, Mar. 2017.

[9] N. Holighaus, Z. Průša, and P. L. Søndergaard, "Reassignment and synchrosqueezing for general time-frequency filter banks, subsampling and processing," *Signal Process.*, vol. 125, pp. 1–8, Aug. 2016.

[10] S. A. Fulop and K. Fitz, "Algorithms for computing the time-corrected instantaneous frequency (reassigned) spectrogram, with applications," *J. Acoust. Soc. Am.*, vol. 119, no. 1, pp. 360–371, Jan. 2006.

[11] H.-T. Wu, "Current state of nonlinear-type time–frequency analysis and applications to high-frequency biomedical signals," *Curr. Opin. Syst. Biol.*, vol. 23, pp. 8–21, Oct. 2020.

[12] Z. Peng, F. Chu, and Y. He, "Vibration signal analysis and feature extraction based on reassigned wavelet scalogram," *J. Sound Vib.*, vol. 253, no. 5, pp. 1087–1100, Jun. 2002.

[13] I. Daubechies and S. Maes, "A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models," in *Wavelets in medicine and biology*, A. Aldroubi and M. A. Unser, Eds. Boca Raton, FL, USA: CRC Press, 1996, pp. 527–546.

[14] I. Daubechies, J. Lu, and H.-T. Wu, "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 2, pp. 243–261, 2011.

[15] G. Thakur and H.-T. Wu, "Synchrosqueezing-based recovery of instantaneous frequency from nonuniform samples," *SIAM J. Math. Anal.*, vol. 43, no. 5, pp. 2078–2095, Jan. 2011.

[16] T. Oberlin, S. Meignen, and V. Perrier, "The Fourier-based synchrosqueezing transform," in *Proc. IEEE Int. Conf. Acoust. Speech. Signal. Process. (ICASSP)*, May 2014, pp. 315–319.

[17] Z.-l. Huang, J. Zhang, T.-H. Zhao, and Y. Sun, "Synchrosqueezing S-transform and its application in seismic spectral decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 2, pp. 817–825, Feb. 2016.

[18] S. Meignen, T. Oberlin, and S. McLaughlin, "A new algorithm for multicomponent signals analysis based on synchrosqueezing: with an application to signal sampling and denoising," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 5787–5798, Nov. 2012.

[19] G. Thakur, E. Brevdo, N. S. Fučkar, and H.-T. Wu, "The Synchrosqueezing algorithm for time-varying spectral analysis: Robustness properties and new paleoclimate applications," *Signal Process.*, vol. 93, no. 5, pp. 1079–1094, 2013.

[20] T. Oberlin, S. Meignen, and V. Perrier, "Second-order synchrosqueezing transform or invertible reassignment? Towards ideal time-frequency representations," *IEEE Trans. Signal Process.*, vol. 63, no. 5, pp. 1335–1344, Mar. 2015.

[21] D.-H. Pham and S. Meignen, "High-order synchrosqueezing transform for multicomponent signals Analysis—with an application to gravitational-wave signal," *IEEE Trans. Signal Process.*, vol. 65, no. 12, pp. 3168–3178, Jun. 2017.

[22] R. Behera, S. Meignen, and T. Oberlin, "Theoretical analysis of the second-order synchrosqueezing transform," *Appl. Comput. Harmon. Anal.*, vol. 45, no. 2, pp. 379–404, 2018.

[23] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[24] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, Jan. 1998.

[25] G. E. Pfander and H. Rauhut, "Sparsity in time-frequency representations," *J. Fourier Anal. Appl.*, vol. 16, pp. 233–260, 2010.

[26] A. Gholami, "Sparse time-frequency decomposition and some applications," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 6, pp. 3598–3604, Jun. 2013.

[27] M. D. Plumbley, T. Blumensath, L. Daudet, R. Gribonval, and M. E. Davies, "Sparse representations in audio and music: From coding to source separation," *Proc. IEEE*, vol. 98, no. 6, pp. 995–1005, 2010.

[28] I. Bayram, "Mixed norms with overlapping groups as signal priors," in *Proc. IEEE Int. Conf. Acoust. Speech. Signal. Process. (ICASSP)*, no. 4, May 2011, pp. 4036–4039.

[29] P. Balazs, M. Doerfler, M. Kowalski, and B. Torresani, "Adapted and adaptive linear time-frequency representations: A synthesis point of view," *IEEE Signal Process. Mag.*, vol. 30, no. 6, pp. 20–31, 2013.

[30] P.-Y. Chen and I. W. Selesnick, "Group-sparse signal denoising: Non-convex regularization, convex optimization," *IEEE Trans. Signal Process.*, vol. 62, no. 13, pp. 3464–3478, Jul. 2014.

[31] E. Sejdić, I. Orović, and S. Stanković, "Compressive sensing meets time-frequency: An overview of recent advances in time-frequency processing of sparse signals," *Digital Signal Process.*, vol. 77, pp. 22–35, 2018.

[32] M. Kowalski, A. Meynard, and H. T. Wu, "Convex optimization approach to signals with fast varying instantaneous frequency," *Appl. Comput. Harmon. Anal.*, vol. 44, no. 1, pp. 89–122, 2018.

[33] T. Kusano, K. Yatabe, and Y. Oikawa, "Maximally energy-concentrated differential window for phase-aware signal processing using instantaneous frequency," in *Proc. IEEE Int. Conf. Acoust. Speech. Signal. Process. (ICASSP)*, May 2020, pp. 5825–5829.

[34] T. Kusano, K. Yatabe, and Y. Oikawa, "Window functions with minimum-sidelobe derivatives for computing instantaneous frequency," *IEEE Access*, vol. 10, pp. 32 075–32 092, 2022.

[35] C. Dolph, "A current distribution for broadside arrays which optimizes the relationship between beam width and side-lobe level," *Proc. IRE*, vol. 34, no. 6, pp. 335–348, Jun. 1946.

[36] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty-V: The discrete case," *Bell Syst. Tech. J.*, vol. 57, no. 5, pp. 1371–1430, May 1978.

[37] F. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proc. IEEE*, vol. 66, no. 1, pp. 51–83, 1978.

[38] J. Kaiser and R. Schafer, "On the use of the I_0-sinh window for spectrum analysis," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 28, no. 1, pp. 105–107, Feb. 1980.

[39] T. Saramäki, "A class of window functions with nearly minimum sidelobe energy for designing FIR filters," in *Proc. IEEE Int. Symp. Circuits. Syst. (ISCAS)*, vol. 1, 1989, pp. 359–362.

[40] S. W. Bergen and A. Antoniou, "Generation of ultraspherical window functions," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, 2002, pp. 1–4.

[41] S. W. Bergen and A. Antoniou, "Design of ultraspherical window functions with prescribed spectral characteristics," *EURASIP J. Adv. Signal Process.*, vol. 2004, no. 13, pp. 2053–2065, Dec. 2004.

[42] K. Avci and A. Nacaroglu, "Cosh window family and its application to FIR filter design," *AEU - Int. J. Electron. Commun.*, vol. 63, no. 11, pp. 907–916, Nov. 2009.

[43] A. Nuttall, "Some windows with very good sidelobe behavior," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 29, no. 1, pp. 84–91, Feb. 1981.

[44] H. Kawahara, K.-I. Sakakibara, M. Morise, H. Banno, T. Toda, and T. Irino, "A new cosine series antialiasing function and its application to aliasing-free glottal source models for speech and singing synthesis," in *Proc. Annu. Conf. Int. Speech Commun. Assoc. (Interspeech)*, Aug. 2017, pp. 1358–1362.

[45] Z. Cvetkovic, "On discrete short-time Fourier analysis," *IEEE Trans. Signal Process.*, vol. 48, no. 9, pp. 2628–2640, 2000.

[46] M. R. Wilbur, T. N. Davidson, and J. P. Reilly, "Efficient design of oversampled NPR GDFT filterbanks," *IEEE Trans. Signal Process.*, vol. 52, no. 7, pp. 1947–1962, 2004.

[47] T. Werther, Y. C. Eldar, and N. N. Subbanna, "Dual Gabor frames: Theory and computational aspects," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4147–4158, 2005.

[48] J. Jiang, S. Ouyang, and F. Zhou, "Design of NPR DFT-modulated filter banks via iterative updating algorithm," *Circuits Syst. Signal Process.*, vol. 32, no. 3, pp. 1351–1362, 2013.

[49] N. Perraudin, N. Holighaus, P. L. Søndergaard, and P. Balazs, "Designing Gabor windows using convex optimization," *Appl. Math. Comput.*, vol. 330, pp. 266–287, Aug. 2018.

[50] T. Kusano, Y. Masuyama, K. Yatabe, and Y. Oikawa, "Designing nearly tight window for improving time-frequency masking," in *Proc. Int. Congr. Acoust. (ICA)*, Sep. 2019, pp. 2885–2892.

[51] S. Meignen, M. Colominas, and D.-H. Pham, "On the use of Rényi entropy for optimal window size computation in the short-time Fourier transform," in *Proc. IEEE Int. Conf. Acoust. Speech. Signal. Process. (ICASSP)*, May 2020, pp. 5830–5834.

[52] E. Wigner, "On the quantum correction for thermodynamic equilibrium," *Phys. Rev.*, vol. 40, no. 5, pp. 749–759, Jun. 1932.

[53] J. Ville, "Théorie et applications de la notion de signal analytique," *Cables Transm.*, vol. 2A, pp. 61–74, 1948.

[54] L. Cohen, *Time-frequency analysis*. Englewood Cliffs, NJ, USA: Prentice Hall, 1995.

[55] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA, USA: SIAM, Jan. 1992.

[56] R. Stockwell, L. Mansinha, and R. Lowe, "Localization of the complex spectrum: the S transform," *IEEE Trans. Signal Process.*, vol. 44, no. 4, pp. 998–1001, Apr. 1996.

[57] J. C. Brown, "Calculation of a constant Q spectral transform," *J. Acoust. Soc. Am.*, vol. 89, no. 1, pp. 425–434, Jan. 1991.

[58] H. Bolcskei, F. Hlawatsch, and H. Feichtinger, "Frame-theoretic analysis of oversampled filter banks," *IEEE Trans. Signal Process.*, vol. 46, no. 12, pp. 3256–3268, Dec. 1998.

[59] G. A. Velasco, N. Holighaus, M. Dörfler, and T. Grill, "Constructing an invertible constant-Q transform with nonstationary Gabor frames," in *Proc. Int. Conf. Digit. Audio Eff. (DAFx-11)*, 2011, pp. 93–99.

[60] P. Balazs, M. Dörfler, F. Jaillet, N. Holighaus, and G. Velasco, "Theory, implementation and applications of nonstationary Gabor frames," *J. Comput. Appl. Math.*, vol. 236, no. 6, pp. 1481–1496, Oct. 2011.

[61] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. Lond. Ser. Math. Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.

[62] D. He, H. Cao, S. Wang, and X. Chen, "Time-reassigned synchrosqueezing transform: The algorithm and its applications in mechanical signal processing," *Mechanical Systems and Signal Processing*, vol. 117, pp. 255–279, Feb. 2019.

[63] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals," *Proc. IEEE*, vol. 80, no. 4, pp. 520–538, Apr. 1992.

[64] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal. II. Algorithms and applications," *Proc. IEEE*, vol. 80, no. 4, pp. 540–568, Apr. 1992.

[65] N. Perraudin, N. Holighaus, P. Majdak, and P. Balazs, "Inpainting of long audio segments with similarity graphs," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 26, no. 6, pp. 1083–1094, Jun. 2018.

[66] K. Yatabe, Y. Masuyama, T. Kusano, and Y. Oikawa, "Representation of complex spectrogram via phase conversion," *Acoust. Sci. Technol.*, vol. 40, no. 3, pp. 170–177, 2019.

[67] P. Flandrin, F. Auger, and E. Chassande-Mottin, "Time-Frequency Reassignment: From Principles to Algorithms," in *Applications in Time-Frequency Signal Processing*, 1st ed., A. Papandreou-Suppappola, Ed. Boca Raton, FL, USA: CRC Press, Oct. 2018, pp. 179–204.

[68] J. L. Flanagan and R. M. Golden, "Phase vocoder," *Bell Syst. Tech. J.*, vol. 45, no. 9, pp. 1493–1509, Nov. 1966.

[69] M. Dolson, "The phase vocoder: A tutorial," *Comput. Music J.*, vol. 10, no. 4, pp. 14–27, 1986.

[70] S. Abeysekera and K. Padhi, "An investigation of window effects on the frequency estimation using the phase vocoder," *IEEE Trans. Audio Speech Lang. Process.*, vol. 14, no. 4, pp. 1432–1439, Jul. 2006.

[71] M. Betser, P. Collen, G. Richard, and B. David, "Estimation of frequency for AM/FM models using the phase vocoder framework," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 505–517, Feb. 2008.

[72] A. Hiruma, K. Yatabe, and Y. Oikawa, "Separating stereo audio mixture having no phase difference by convex clustering and disjointness map," in *Proc. Int. Workshop Acoust. Signal Enhance. (IWAENC)*, Sep. 2018, pp. 266–270.

[73] A. Hiruma, K. Yatabe, and Y. Oikawa, "Detection of clean time-frequency bins based on phase derivative of multichannel signals," in *Proc. Int. Congr. Acoust. (ICA)*, Sep. 2019, pp. 2797–2804.

[74] K. Yatabe and Y. Oikawa, "Phase corrected total variation for audio signals," in *Proc. IEEE Int. Conf. Acoust. Speech. Signal. Process. (ICASSP)*, Apr. 2018, pp. 656–660.

[75] Y. Masuyama, K. Yatabe, and Y. Oikawa, "Phase-aware harmonic/percussive source separation via convex optimization," in *Proc. IEEE Int. Conf. Acoust. Speech. Signal. Process. (ICASSP)*, May 2019, pp. 985–989.

[76] Y. Masuyama, K. Yatabe, and Y. Oikawa, "Low-rankness of complex-valued spectrogram and its application to phase-aware audio processing," in *Proc. IEEE Int. Conf. Acoust. Speech. Signal. Process. (ICASSP)*, May 2019, pp. 855–859.

[77] A. P. Stark and K. K. Paliwal, "Speech analysis using instantaneous frequency deviation," in *Proc. Annu. Conf. Int. Speech Commun. Assoc. (Interspeech)*, 2008, pp. 2602–2605.

[78] M. Grimaldi and F. Cummins, "Speaker identification using instantaneous frequencies," *IEEE Trans. Audio Speech Lang. Process.*, vol. 16, no. 6, pp. 1097–1111, Aug. 2008.

[79] A. Gaich and P. Mowlaee, "On speech quality estimation of phase-aware single-channel speech enhancement," in *Proc. IEEE Int. Conf. Acoust. Speech. Signal. Process. (ICASSP)*, Apr. 2015, pp. 216–220.

[80] Z. Průša and P. L. Søndergaard, "Real-time audio visualization with reassigned non-uniform filter banks," in *Proc. Int. Conf. Digit. Audio Eff. (DAFx-16)*, Sep. 2016, pp. 3–8.

[81] C. K. Chui, Y.-T. Lin, and H.-T. Wu, "Real-time dynamics acquisition from irregular samples — With application to anesthesia evaluation," *Anal. Appl.*, vol. 14, no. 04, pp. 537–590, 2016.

[82] P. L. Søndergaard, "Efficient algorithms for the discrete gabor transform with a long FIR window," *J. Fourier Anal. Appl.*, vol. 18, pp. 456–470, 2012.

[83] P. Bloomfield, *Fourier Analysis of Time Series: An Introduction*, 2nd ed. New York, NY, USA: John Wiley & Sons, 2000.

[84] G. H. Golub, "Some modified matrix eigenvalue problems," *SIAM Rev.*, vol. 15, no. 2, pp. 318–334, Apr. 1973.

[85] J. R. Weaver, "Centrosymmetric (cross-symmetric) matrices, their basic properties, eigenvalues, and eigenvectors," *Am. Math. Mon.*, vol. 92, no. 10, pp. 711–717, 1985.

[86] J. McClellan and T. Parks, "A unified approach to the design of optimum FIR linear-phase digital filters," *IEEE Trans. Circuit Theory*, vol. 20, no. 6, pp. 697–701, Nov. 1973.

[87] D. Gruenbacher and D. Hummels, "A simple algorithm for generating discrete prolate spheroidal sequences," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 3276–3278, 1994.

[88] T. Verma, S. Bilbao, and T. Meng, "The digital prolate spheroidal window," in *Proc. IEEE Int. Conf. Acoust. Speech. Signal. Process. (ICASSP)*, vol. 3, 1996, pp. 1351–1354.

[89] X. Lai, "Constrained Chebyshev design of FIR filters," *IEEE Trans. Circuits Syst. II Express Briefs*, vol. 51, no. 3, pp. 143–146, Mar. 2004.

[90] J. McClellan, T. Parks, and L. Rabiner, "A computer program for designing optimum FIR linear phase digital filters," *IEEE Trans. Audio Electroacoustics*, vol. 21, no. 6, pp. 506–526, Dec. 1973.

[91] X. Lai, "Chebyshev design of a class of FIR filters with frequency equation constraints," *Circuits Syst. Signal Process.*, vol. 21, no. 2, pp. 181–193, Mar. 2002.

[92] M. J. D. Powell, *Approximation theory and methods.* New York, NY, USA: Cambridge University Press, 1981.

[93] R. Baraniuk, P. Flandrin, A. Janssen, and O. Michel, "Measuring time-frequency information content using the Rényi entropies," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1391–1409, May 2001.

[94] S. Meignen, D.-H. Pham, and S. McLaughlin, "On demodulation, ridge detection, and synchrosqueezing for multicomponent signals," *IEEE Trans. Signal Process.*, vol. 65, no. 8, pp. 2093–2103, Apr. 2017.

[95] O. Pele and M. Werman, "Fast and robust Earth mover's distances," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Sep. 2009, pp. 460–467.

[96] I. Daubechies, Y. Wang, and H.-T. Wu, "ConceFT: concentration of frequency and time via a multitapered synchrosqueezed transform," *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.*, vol. 374, no. 2065, p. 20150193, Apr. 2016.

[97] Y. Chi, L. L. Scharf, A. Pezeshki, and A. R. Calderbank, "Sensitivity to basis mismatch in compressed sensing," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2182–2195, 2011.

[98] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Commun. Pure Appl. Math.*, vol. 67, no. 6, pp. 906–956, 2014.

[99] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The convex geometry of linear inverse problems," *Found. Comput. Math.*, vol. 12, no. 6, pp. 805–849, Dec. 2012.

[100] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7465–7490, 2013.

[101] Y. Chi and M. Ferreira Da Costa, "Harnessing sparsity over the continuum: Atomic norm minimization for superresolution," *IEEE Signal Process. Mag.*, vol. 37, no. 2, pp. 39–57, 2020.

[102] Z. Yang and L. Xie, "On gridless sparse methods for line spectral estimation from complete and incomplete data," *IEEE Trans. Signal Process.*, vol. 63, no. 12, pp. 3139–3153, 2015.

[103] Y. Li and Y. Chi, "Off-the-grid line spectrum denoising and estimation with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 64, no. 5, pp. 1257–1269, Mar. 2016.

[104] A. Xenaki and P. Gerstoft, "Grid-free compressive beamforming," *J. Acoust. Soc. Am.*, vol. 137, no. 4, pp. 1923–1935, Apr. 2015.

[105] Y. Yang, Z. Chu, Z. Xu, and G. Ping, "Two-dimensional grid-free compressive beamforming," *J. Acoust. Soc. Am.*, vol. 142, no. 2, pp. 618–629, 2017.

[106] K. Mahata and M. M. Hyder, "Grid-less T.V minimization for DOA estimation," *Signal Process.*, vol. 132, pp. 155–164, Mar. 2017.

[107] R. Heckel, V. I. Morgenshtern, and M. Soltanolkotabi, "Super-resolution radar," *Inf. Inference*, vol. 5, no. 1, pp. 22–75, Mar. 2016.

[108] C. Caratheodory and L. Fejér, "über den zusammenhang der extremen von harmonischen funktionen mit ihren koeffizienten und über den picard-landau'schen satz," *Rendiconti Circolo Mat. Palermo*, vol. 32, no. 1, pp. 218–239, Dec. 1911.

[109] G. Riche de Prony, "Essai expérimental et analytique: sur les lois de la dilatabilité de fluides élastiques et sur celles de la force expansive de la vapeur de l'eau et de la vapeur de l'alcool à différentes températures," *J. L'École Polytech.*, vol. 1, no. 22, pp. 24–76, 1795.

[110] Y. Hua and T. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 38, no. 5, pp. 814–824, May 1990.

[111] D. Tufts and R. Kumaresan, "Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood," *Proc. IEEE*, vol. 70, no. 9, pp. 975–989, 1982.

[112] S. Boyd, N. Parikh, and E. Chu, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2010.

[113] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[114] Y. C. Eldar and G. Kutyniok, *Compressed sensing: theory and applications*. New York, NY, USA: Cambridge University Press, 2012.

[115] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Rev.*, vol. 51, no. 1, pp. 34–81, Feb. 2009.

[116] M. Zibulevsky and M. Elad, "L1-L2 optimization in signal and image processing," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 76–88, May 2010.

[117] H. Zhu, G. Leus, and G. B. Giannakis, "Sparsity-cognizant total least-squares for perturbed compressive sampling," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2002–2016, May 2011.

[118] C. Ekanadham, D. Tranchina, and E. P. Simoncelli, "Recovery of sparse translation-invariant signals with continuous basis pursuit," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4735–4744, Oct. 2011.

[119] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse Bayesian inference," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 38–43, Jan. 2013.

[120] X. Wu, W.-P. Zhu, and J. Yan, "Direction of arrival estimation for off-grid signals based on sparse Bayesian learning," *IEEE Sens. J.*, vol. 16, no. 7, pp. 2004–2016, Apr. 2016.

[121] L. Hu, Z. Shi, J. Zhou, and Q. Fu, "Compressed sensing of complex sinusoids: An approach based on dictionary refinement," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3809–3822, Jul. 2012.

[122] J. Fang, F. Wang, Y. Shen, H. Li, and R. S. Blum, "Super-resolution compressed sensing for line spectral estimation: An iterative reweighted approach," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4649–4662, Sep. 2016.

[123] T. Tachikawa, K. Yatabe, and Y. Oikawa, "3D sound source localization based on coherence-adjusted monopole dictionary and modified convex clustering," *Appl. Acoust.*, vol. 139, no. May, pp. 267–281, 2018.

[124] T. Kusano, K. Yatabe, and Y. Oikawa, "Sparse time-frequency representation via atomic norm minimization," in *Proc. IEEE Int. Conf. Acoust. Speech. Signal. Process. (ICASSP)*, Jun. 2021, pp. 5075–5079.

[125] G. Tang, "Resolution limits for atomic decompositions via Markov-Bernstein type inequalities," in *Proc. Int. Conf. Sampl. Theory. Appl. (SampTA)*, Mar. 2015, pp. 548–552.

[126] C. Fernandez-Granda, "Super-resolution of point sources via convex programming," *Inf. Inference*, vol. 5, no. 3, pp. 251–303, 2016.

[127] K. V. Mishra, M. Cho, A. Kruger, and W. Xu, "Spectral super-resolution with prior knowledge," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5342–5357, 2015.

[128] Z. Yang and L. Xie, "Fast convex optimization method for frequency estimation with prior knowledge in all dimensions," *Signal Process.*, vol. 142, pp. 271–280, Jan. 2018.

[129] M. Cho, K. V. Mishra, J.-F. Cai, and W. Xu, "Block iterative reweighted algorithms for super-resolution of spectrally sparse signals," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2319–2313, Dec. 2015.

[130] Z. Yang and L. Xie, "Enhancing sparsity and resolution via reweighted atomic norm minimization," *IEEE Trans. Signal Process.*, vol. 64, no. 4, pp. 995–1006, 2016.

[131] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *J. Fourier Anal. Appl.*, vol. 14, no. 5-6, pp. 877–905, 2008.

[132] R. Chartrand and W. Yin, "Nonconvex Sparse Regularization and Splitting Algorithms," in *Splitting Methods in Communication, Imaging, Science, and Engineering*. Cham, Switzerland: Springer International Publishing, 2016, pp. 237–249.

[133] F. Wen, L. Chu, P. Liu, and R. C. Qiu, "A survey on nonconvex regularization-based sparse and low-rank recovery in signal processing, statistics, and machine learning," *IEEE Access*, vol. 6, pp. 69 883–69 906, 2018.

[134] G. Marjanovic and V. Solo, "On l_q optimization and matrix completion," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 5714–5724, Nov. 2012.

[135] D. Malioutov and A. Aravkin, "Iterative log thresholding," in *Proc. IEEE Int. Conf. Acoust. Speech. Signal. Process. (ICASSP)*, May 2014, pp. 7198–7202.

[136] J. Fan and R. Li, "Variable selection via nonconcave penalized likelihood and its oracle properties," *J. Am. Stat. Assoc.*, vol. 96, no. 456, pp. 1348–1360, Dec. 2001.

[137] C.-H. Zhang, "Nearly unbiased variable selection under minimax concave penalty," *Ann. Stat.*, vol. 38, no. 2, pp. 894–942, Apr. 2010.

[138] T. Zhang, "Analysis of multi-stage convex relaxation for sparse regularization," *J. Mach. Learn. Res.*, vol. 11, pp. 1081–1107, Mar. 2010.

[139] Z. Yang and L. Xie, "Exact joint sparse frequency recovery via optimization methods," *IEEE Trans. Signal Process.*, vol. 64, no. 19, pp. 5145–5157, Oct. 2016.

[140] F. Zhang, *The Schur complement and its applications*. New York, NY, USA: Springer Science & Business Media, 2006.

[141] N. Parikh and S. Boyd, "Proximal algorithms," *Found. Trends Optim.*, vol. 1, no. 3, pp. 127–239, 2014.

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

[143] Z. Xu, X. Chang, F. Xu, and H. Zhang, "$l_{1/2}$ regularization: A thresholding representation theory and a fast solver," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 7, pp. 1013–1027, Jul. 2012.

[144] W. Cao, J. Sun, and Z. Xu, "Fast image deconvolution using closed-form thresholding formulas of regularization," *J. Vis. Commun. Image Represent.*, vol. 24, no. 1, pp. 31–41, Jan. 2013.

[145] W. Zuo, D. Meng, L. Zhang, X. Feng, and D. Zhang, "A generalized iterated shrinkage algorithm for non-convex sparse coding," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2013, pp. 217–224.

[146] R. T. Rockafellar, "Augmented Lagrangians and applications of the proximal point algorithm in convex programming," *Math. Oper. Res.*, vol. 1, no. 2, pp. 97–116, 1976.

[147] R. I. Boţ and D.-K. Nguyen, "The proximal alternating direction method of multipliers in the nonconvex setting: Convergence analysis and rates," *Math. Oper. Res.*, vol. 45, no. 2, pp. 682–712, Mar. 2020.

[148] J. Zeng, W. Yin, and D.-X. Zhou, "Moreau envelope augmented Lagrangian method for nonconvex optimization with linear constraints," *J. Sci. Comput.*, vol. 91, no. 2, p. 61, Apr. 2022.

[149] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," Mar. 2014.

[150] K. C. Toh, M. J. Todd, and R. H. Tütüncü, "SDPT3 — A Matlab software package for semidefinite programming, Version 1.3," *Optim. Methods Softw.*, vol. 11, no. 1–4, pp. 545–581, 1999.

[151] T. Reed and H. Wechsler, "Segmentation of textured images and Gestalt organization using spatial/spatial-frequency representations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 1, pp. 1–12, Jan. 1990.

[152] S. Chikkerur, A. N. Cartwright, and V. Govindaraju, "Fingerprint enhancement using STFT analysis," *Pattern Recognit.*, vol. 40, no. 1, pp. 198–211, Jan. 2007.

[153] Q. Kemao, "Applications of windowed Fourier fringe analysis in optical measurement: A review," *Opt. Lasers Eng.*, vol. 66, pp. 67–73, Mar. 2015.

Publications

Journal

1. ◯T. Kusano, K. Yatabe, and Y. Oikawa, "Window functions with minimum-sidelobe derivatives for computing instantaneous frequency," IEEE Access, vol. 10, pp. 32075–32092, Mar. 2022.

2. Y. Masuyama, T. Kusano, K. Yatabe, and Y. Oikawa, "Modal decomposition of musical instrument sounds via optimization-based non-linear filtering," Acoustical Science and Technology, vol. 40, no. 3, pp. 186–197, May 2019.

3. T. Kusano, K. Yatabe, and Y. Oikawa, "Localization of marine seismic vibrator based on hyperbolic Radon transform," Acoustical Science and Technology, vol. 39, no. 3, pp. 215–225, May. 2018.

Review

1. K. Yatabe, Y. Masuyama, T. Kusano, and Y. Oikawa, "Representation of complex spectrogram via phase conversion," Acoustical Science and Technology, vol. 40, no. 3, pp. 170–177, May 2019.

2. 矢田部浩平，升山義紀，草野翼，及川靖広，"位相変換による複素スペクトログラムの表現，" 日本音響学会誌，vol. 75，no. 3，pp. 147–155，Mar. 2019.

International conference (peer-reviewed)

1. ◯T. Kusano, K. Yatabe, and Y. Oikawa, "Sparse time-frequency representation via atomic norm minimization," 46th IEEE International Conference on

Acoustics, Speech, and Signal Processing (ICASSP 2021), Toronto, Canada (fully virtual), pp. 5060–5064, June 2021.

2. ◯T. Kusano, K. Yatabe, and Y. Oikawa, "Maximally energy-concentrated differential window for phase-aware signal processing using instantaneous frequency," 45th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2020), Barcelona, Spain (fully virtual), pp. 5825–5829, May 2020.

3. T. Kusano, K. Yatabe, and Y. Oikawa, "Envelope estimation by tangentially constrained spline," 43th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2018), Calgary, Canada, pp. 4374–4378, Apr. 2018.

4. Y. Masuyama, T. Kusano, K. Yatabe, and Y. Oikawa, "Modal decomposition of musical instrument sound via alternating direction method of multipliers," 43th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2018), Calgary, Canada, pp. 631–635, Apr. 2018.

International conference

1. T. Kusano, Y. Masuyama, K. Yatabe, and Y. Oikawa, "Designing nearly tight window for improving time-frequency masking," 23rd International Congress on Acoustics (ICA 2019), Aachen, Germany, pp. 2885–2892, Sep. 2019.

2. Y. Kohase, T. Kusano, K. Yatabe, and Y. Oikawa, " Fourier spectral method for acoustic simulation with domain enclosed by curved boundary," 23rd International Congress on Acoustics (ICA 2019), Aachen, Germany, pp. 4504–4510, Sep. 2019.

3. T. Kusano, Y. Matsumoto, Y. Kataoka, W. Teraoka, Y. Oikawa, K. Yoshida, and Y. Kitazumi, "Three-dimensional sound source localization using microphone array and its display with mixed reality technology," 48th International Congress and Exposition on Noise Control Engineering (Inter-Noise 2019), Madrid, Spain, p. 1776, June 2019.

Domestic Conference

1. ◯草野翼，矢田部浩平，及川靖広，“2 次元 synchrosqueezing 変換を用いた干渉縞復元，” 日本音響学会秋季研究発表会講演論文集，pp. 173–174，Sep. 2022.

2. 後藤昌彦，草野翼，潘明宇，井上敦登，寺岡航，及川靖広，佐藤考浩，岩根康之，小林真人，“光学透過型ヘッドマウントディスプレイを用いた音圧分布可視化手法に関する研究 －その 5 MVDR を用いた三次元音源位置推定の検討－，” 日本音響学会秋季研究発表会講演論文集，pp. 827–828，Sep. 2022.

3. 野澤遙, 草野翼, 及川靖広, “偏光高速度干渉計を用いた口笛演奏音収録への最小分散無歪応答法の適用” 日本音響学会秋季研究発表会講演論文集，pp. 225–226，Sep. 2022.

4. ◯草野翼，矢田部浩平，及川靖広，“アトミックノルムによるスパース最適化の非凸最適化への拡張，” 日本音響学会春季研究発表会講演論文集，pp. 203–204，Mar. 2022.

5. 草野翼，矢田部浩平，及川靖広，“アトミックノルム最小化によるスパース時間周波数表現を利用したノイズ除去，” 日本音響学会秋季研究発表会講演論文集，pp. 205–206，Sep. 2021.

6. 草野翼，矢田部浩平，及川靖広，“アトミックノルム最小化によるスパース時間周波数表現，” 日本音響学会春季研究発表会講演論文集，pp. 269–270，Mar. 2021.

7. 草野翼，矢田部浩平，及川靖広，“瞬時周波数計算のための微分窓のサイドローブ最小化，” 日本音響学会秋季研究発表会講演論文集，pp. 179–180，Sep. 2020.

8. 草野翼，矢田部浩平，及川靖広，“瞬時周波数計算のためのエネルギー局在微分窓，” 日本音響学会春季研究発表会講演論文集，pp. 247–248，Mar. 2020.

9. 草野翼，矢田部浩平，及川靖広，“短時間 Fourier 変換瞬時周波数推定における窓関数の影響，” 日本音響学会秋季研究発表会講演論文集，pp. 203–204，Sep. 2019.

10. 小長谷悠，草野翼，矢田部浩平，及川靖広，“スペクトル法の曲面形状への拡張，” 日本音響学会秋季研究発表会講演論文集，pp. 505–506，Sep. 2019.

11. 矢田部浩平，升山義紀，草野翼，及川靖広，“MATLAB 瞬時周波数 Toolbox，”日本音響学会秋季研究発表会講演論文集，pp. 255–258，Sep. 2019.

12. 草野翼，升山義紀，矢田部浩平，及川靖広，“時間周波数マスキング性能を向上させる窓関数，”日本音響学会春季研究発表会講演論文集，pp. 173–174，Mar. 2019.

13. 草野翼，升山義紀，矢田部浩平，及川靖広，“音響信号処理に対する逆短時間 Fourier 変換の合成窓関数の影響，”日本音響学会秋季研究発表会講演論文集，pp. 167–168，Sep. 2018.

14. 草野翼，矢田部浩平，及川靖広，“接スプラインによる包絡の新たな定義とその推定，”日本音響学会春季研究発表会講演論文集，pp. 501–504，Mar. 2018.

15. 升山義紀，草野翼，矢田部浩平，及川靖広，宮城雄介，大石耕史，高橋健，“交互方向乗数法を用いたモード分解による楽器音の解析，”日本音響学会春季研究発表会講演論文集，pp. 793–794，Mar. 2018.

16. 升山義紀，草野翼，矢田部浩平，及川靖広，宮城雄介，大石耕史，“データ忠実性を制約とした最適化による楽器音のモード分解，”日本音響学会音楽音響研究会資料，vol. MA2017-36，pp. 109–114, Oct. 2017.

17. 草野翼，矢田部浩平，及川靖広，“4 次 C2 級スプライン関数を用いた制約付き最適化による包絡推定，”日本音響学会秋季研究発表会講演論文集，pp. 471–472，Sep. 2017.

18. 升山義紀，草野翼，矢田部浩平，及川靖広，宮城雄介，大石耕史，“制約付き最適化を用いた楽器音のモード分解，”日本音響学会秋季研究発表会講演論文集，pp. 1085–1086，Sep. 2017.

19. 草野翼，矢田部浩平，及川靖広，小笹弘晃，田中浩一郎，佐藤文男，“海底資源探査のためのスパース最適化を用いたパルス圧縮法，”日本音響学会春季研究発表会講演論文集，pp. 1321–1322, Mar. 2017.

20. 草野翼，矢田部浩平，及川靖広，小笹弘晃，田中浩一郎，佐藤文男，“双曲型ラドン変換を用いたマリンセイスミックバイブレータの位置推定，”日本音響学会秋季研究発表会講演論文集，pp. 1085–1086，Sep. 2016.

Other

- 及川靖広，矢田部浩平，草野翼，升山義紀，宮城雄介，大石耕史，“モード分解装置、モード分解方法、プログラム,” 特開 2019-159018 (特願 2018-43193), 2018 年 3 月 9 日出願.
- 小笹弘晃，田中浩一郎，佐藤文男，及川靖広，矢田部浩平，草野翼，“震源位置推定方法及び震源位置推定システム，” 特開 2018-31606 (特願 2016-162296), 2016 年 8 月 22 日出願.

Award

1. 日本音響学会 第 21 回学生優秀発表賞, 2020.