Theoretical Analysis and Perceptual Evaluations on Nonlinear Devices in Virtual Bass System

Nay Oo

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2012

ACKNOWLEDGEMENTS

This work has been performed at the Digital Signal Processing laboratory (DSP Lab) of the school of Electrical and Electronic Engineering in Nanyang Technological University (NTU). I would like to express my gratitude to a number of people who have contributed directly or indirectly in numerous ways to support and encourage me throughout the period of my candidature.

First and foremost, I would like to express my greatest gratitude to my dissertation supervisor, Professor Dr. Gan Woon Seng, for his invaluable guidance, continual supports and encouragements throughout the course of this work. He initiated me to the world of research, taught me audio signal processing techniques, real-time fixed-point digital signal processors programming skills, reviewed my works, gave lots of constructive criticism, and provided a productive research atmosphere. He also inspired me to discover new ideas and further research developments.

In addition, I would like to thank Mr. Chua Song Heng, Director, Innovation Site Singapore, and Mr. Gozali Goh Kong San, Function Manager on Sound, of Philips Electronics (S) Pte Ltd for giving me opportunities to transfer research findings and technological know-how to Philips staff, and to develop a commercial grade implementation of NTU DSP Lab's Virtual-Bass technology for use in Philips televisions.

I am also grateful to Mr. Yeo Sung Kheng and Mr. Ong Say Cheng, for their support and help in logistic, administrative matters as well as social activities. I would like to show my gratitude to all my friends from DSP Lab—present and past—Ee

Leng (Joseph), Quy, Dr. Ji Peifeng, Dahyanto, Wang Liang, Reuben Johannes, Sam, Ji Wei, Liao Lei, Wang Tongwei, Seth, Dileep, Kushal, Kaushik, Rishabh, Jin Feng and Mu Hao. They have made the lab an interesting place for doing research and studying, and also organizing many activities during my Ph.D. candidature in DSP lab.

Over the years, I had also very fruitful interactions with the final year project students in the DSP lab at School of Electrical and Electronic Engineering in NTU. I would like to say special thanks to Lim Wee Tong and Jason Teo Zu-Tian. I am indebted to my friend, Marla Brunker for proofreading the dissertation.

Lastly but not least, I am indebted to my parents, my wife, and all family members for their encouragements and supports throughout the course of my Ph.D.

Nay Oo

August, 2011

Singapore

Table of Contents

SUMMARY

With the development of small consumer audio-enabled devices, such as mobile phones, tablet computers, light-weight laptop computers, and flat-panel televisions, the demand to reproduce good bass sounds with inherently small loudspeakers has never been greater. Enhancement of bass sounds by simply amplifying low frequency audio components does not work well, and may even overload and damage small loudspeakers. Thus, due to the limitations of miniature loudspeakers, physical bass enhancement is impossible without using bulky subwoofers.

However, an alternative way to enhance bass is by tricking our human auditory system into perceiving fuller bass sounds. This is called the psychoacoustic bass enhancement system, or virtual bass system (VBS). By making use of the phenomenon known as the "missing fundamental," by which our auditory system reconstructs the fundamental frequency (virtual pitch) from harmonics even if the fundamental frequency is completely missing, we can trick our ears to hear low frequencies, which cannot be efficiently reproduced by the small loudspeakers, from their higher harmonics, which are carefully injected into the original audio signal stream and situated in the mid-frequency range where loudspeakers can reproduce them well. Thus, physical bass frequencies that cannot be reproduced by small loudspeakers can be replaced by the *perception* of psychoacoustic bass frequencies that is synthesized in the human auditory system, thereby extending the experience of low-frequency bandwidth for listeners.

One technique to inject harmonics in the mid-range is processing signals nonlinearly, using nonlinear devices (NLDs). While injecting harmonics in the

original audio signal stream can be viewed as introducing distortions, such distortions can be either good or bad. A good and carefully introduced distortion can cause warmer and fuller audio effects due to the missing fundamental effect creating the sense of richer harmonics. Some examples of good, intentionally introduced audio distortions are those produced by *FuzzBox* in electric guitar amplifiers, and by vacuum tube synthesizers.

Whereas NLDs will produce both harmonic (desirable) and intermodulation components (undesirable); the latter have no harmonic structure and therefore contribute only to the perception of signal distortion. However, due to the auditory masking phenomena of our human auditory system, we cannot hear all such physical audio components; the extent to which these intermodulation distortion components cause perceivable artifacts is a good question for study. Moreover, there are infinite possibilities for NLDs (even if we restrict ourselves to static nonlinear systems only). The kinds of NLD static curves that can create the perception of deep bass with minimal undesirable bad distortions offer another possibility for research.

To address these research questions, this dissertation performed a single-tone harmonic analysis of a selected and derived pool of NLDs and a multitone intermodulation distortion analysis of such NLDs, and conducted two listening test experiments, on bass intensity and distortion quality, respectively; relating the subjective results with objective results. To aid in objective analysis, exact distortion components computation algorithms, including harmonic analysis and synthesis on polynomial (HASP) and the Yeary-Oo-Gan algorithm (YOGA), were developed from first principles and used in harmonic and intermodulation distortion analysis of

NLDs. The perceptual nonlinear distortion model *Rnonlin* was used to correlate subjective listening tests results. Derivatives of NLD analyses were performed to ascertain optimal NLDs for VBS, and for psychoacoustic enhancement of bass for miniature loudspeakers.

List of Figures

List of Tables

List of Abbreviations and Acronyms

ANOVA	Analysis of Variance
ACF	Autocorrelation Function
ANC	Anchor
ANSI	American National Standards Institute
AS	Audio Sequence/Stimulus
ASA	Auditory Scene Analysis
ASQRT	Absolute Square Root
ATSR	Arc-Tangent Square Root
B	Both Even and Odd
CD	Compact Disc
CLP	Clipper
CUBE	Cube
dBFS	Decibel Full Scale
DC	Direct Current
df	Degree of Freedom
DFT	Discrete Fourier Transform
DS	Distortion Score
DUT	Device Under Test
E	Even
ERB	Equivalent Rectangular Bandwidth
et al.	et allii
etc.	et cetera
EXP1	Exponential 1
EXP2	Exponential 2

EXPs	Exponentials
FEXP1	Fuzz Exponential 1
FEXP2	Fuzz Exponential 2
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FREXP1	Full-range Exponential 1
FREXP2	Full-range Exponential 2
FWR	Full-Wave Rectifier
GPU	Graphics Processing Unit
GUI	Graphical User Interface
HASP	Harmonic Analysis and Synthesis on Polynomials
HAT	Harmonic Addition Theorem
HPF	High-Pass Filter
HREXP1	Half-range Exponential 1
HREXP2	Half-range Exponential 2
HWR	Half-Wave Rectifier
i.e.	id est (that is)
IMD	Intermodulation Distortion
LCB	Lower Confidence Bound
LPF	Low-Pass Filter
LTI	Linear Time-Invariant
MAF	Minimum Audible Field
MPIPCT	Modified PIPCT
MS	Mean Score
MUSHRA	Multiple Stimuli with Hidden Reference and Anchor
NLD	Nonlinear Device

NSIG	Normalized Sigmoid
NTANH	Normalized Hyperbolic Tangent
O	Odd
OLA	OverLapAdd
OME	Outer and Middle Ear
OS	Opinion Score
PIPCT	Precise Intermodulation Product Computing Technique
Q.E.D.	quod erat demonstrandum
REF	Reference
RRMSE	Relative Root-Mean-Square Error
SPL	Sound Pressure Level
SQAM	Sound Quality Assessment Material
SQL	Square Law
SQS	Square Sign
SR	SortRows
TC	Test Condition
THD	Total Harmonic Distortion
UCB	Upper Confidence Bound
USB	Universal Serial Bus
VB	Virtual Bass
VBS	Virtual Bass System
YOGA	Yeary Oo Gan Algorithm

CHAPTER 1

INTRODUCTION

1.1 Motivation

Modern consumer electronics, such as flat-screen televisions (TVs) and portable handheld devices, require smaller and flatter loudspeakers, often at the expense of audio performance. Human hearing perception generally ranges from 20–20000 Hz [Moo03] [Zwi61], while audio bass frequency generally ranges from 20–250 Hz [Ows99]. However, due to the form-factor limitation, the physical bass reproduction in small loudspeakers can be extremely poor [NH07] [OG09a] [LA04], causing an unavoidably poor bass response [LA04, p.54]. Enhancing the bass by simply amplifying low frequency audio components is usually undesirable because of limited excitation in small loudspeakers, and may overload or damage the loudspeakers [GKT01].

However, an arguably better way to enhance bass would be to trick our human auditory system into perceiving bass that is not physically present [BC99] [SD99] [PPA00] [GKT01] [LA02] [TSY10]. This kind of system is called the psychoacoustic bass enhancement system [Kar02] [OG09b]. In this dissertation, we call this system the virtual bass system (VBS). A psychoacoustic phenomenon, known as the "missing fundamental" (also called residue pitch or virtual pitch [Moo03]), states that our auditory system reconstructs the fundamental frequency from the harmonics in

the absence of low frequencies [Moo03] [POF05]. In other words, we can trick our ears to hear low frequencies from the higher harmonics in the absence of low frequencies. In that sense, physical bass frequencies, which cannot be reproduced by the loudspeakers, can be reproduced virtually by the psychoacoustic bass frequencies that are synthesized in the human auditory systems. Thus, we can virtually extend the low frequency bandwidth with VBS [LA04] [GKT01] [LA02] [BC99] [Kar02] [OG09b]. Some of the successful deployments of VBS include wave-field synthesis [PRL10], audio crosstalk cancellation [TGT00], virtual surround [TG99], game stations [GKT01], flat TV sets [BBH07], active noise control headsets [GK04], VoIP speech intelligibility [THH08], piezoelectric loudspeakers [LOG09] [CAH11] [KGE02], wide-area psychoacoustic correction for problematic room modes [HH10a], and acoustic echo cancellation [RCP10] [CRP11].

The virtual bass system may be implemented in either time or frequency domains. For the time-domain approach, a nonlinear device (NLD) [LA02] [OG09b] [PRL10] [BBH07] [LOG09] [CAH11] [OGH11] [OGL10] [OG08a] [OG08b] is commonly used to create harmonics to enhance bass perception. In other words, NLD is used to distort the original clean signal and produces harmonics that can extend the human auditory an octave lower, or perceive the missing fundamental. This is a blind approach [LA04], which passes a pre-filtered signal directly to the NLD to synthesize harmonics related to the missing fundamental. The blind approach using NLD can create intermodulation distortion in addition to the desired harmonic distortion (see pp. 666 of [LSP08]). Intermodulation components are not harmonically related to the original frequency components in the clean signal, and are hypothesized to be purely

distortion artifacts. However, our human auditory system also works nonlinearly [Moo03], and there are complex psychoacoustic phenomena taking place when audio signals are being processed by the auditory system. For example, the psychoacoustic masking effect [Moo03] [FZ07] [GL03a] [GL03b] [Voi06] plays an important role in determining which part of the audio signal can be perceived. Not all the tonal components are perceived by our human auditory system. In other words, a tone may be physically present, but psychoacoustically absent. Moreover, the advantage of the blind approach is its lower computation complexity, which is much more suited for real-time signal processing.

There is, however, a non-blind approach [LA04], where a pitch detector is employed to determine the fundamental frequency within the original signal from which a related harmonic series can be synthesized. There are some drawbacks to the non-blind approach. First, the pitch detector can find up to two multi-pitches at the same time, using an adaptive notch filter approach [LSP08]. Second, the pitch detector approach results in increased computational complexity, compared to the NLD-based approaches. For example, due to the ease of implementation and lighter computational load, an NLD-based approach has been successfully implemented in a fixed-point real-time DSP platform (demonstrated in [OG09b]). Third, the pitch detector needs highly accurate pitch-tracking capability to avoid faulty pitch detection and audible (annoying or—for some—unbearable) artifacts. (For problems in pitch detection, see [CRR76] and [RCR76].). Fourth, a harmonic generator is required to synthesize harmonics that do not cause severe artifacts and timbre differences. The

main advantage of the non-blind approach is that no intermodulation component is generated.

Under the frequency-domain approach, a phase vocoder is used to shift harmonics from the bass frequency region into the mid-frequency region [BL06]. Recently, Hill and Hawksford proposed a hybrid VBS approach (combining phase vocoder and NLD); their subjective listening test indicated good performance from this hybrid approach [HH10b]. They also showed that some types of audio tracks (drums and "oldies") are more suited to the NLD-based approach.

In this dissertation, we shall focus on the NLD-based blind approach, in which static (memoryless) nonlinear systems are used in the VBS implementation. Nonlinear devices produce both harmonics and intermodulation components. However, due to the auditory masking phenomena, intermodulation distortion may or may not be perceived as artifacts. More in-depth, objective and subjective studies on how intermodulation affects bass perception have been carried out in this dissertation. Moreover, there are many possible types of NLDs, even if we restrict ourselves to static nonlinear systems only; we wish to know what kind of static nonlinear curves can create deeper bass perception without sacrificing quality.

The major objectives of this dissertation are

- To conduct *objective* evaluations of NLDs; and
- To conduct *subjective* evaluations of NLDs.

The above objectives are sub-divided into the following goals. For the objective evaluations, we intend

- To perform single-tone harmonic analyses of a selected and derived pool of NLDs;
- To perform the multitone intermodulation distortion analyses of such NLDs; and
- To correlate the subjective data with perceptually motivated models.

For the subjective evaluations, the tasks are

- To conduct a first subjective listening experiment, to test the bass intensity generated by such NLDs; and
- To conduct a second subjective listening experiment to test the distortion artifacts caused by NLDs.

Thus, the central objective of this research is greater understanding of the distortion and bass perception effects of various types of NLDs and to relate the results of two listening tests with the objective measures, such as single tone, multitone and predictive nonlinear distortion perceptual models. While fulfilling the targeted objectives, we will also find suitable NLDs, as well as new mathematical tools to perform tonal analyses.

1.2 Contributions of the Dissertation

This dissertation focuses on the analytical and perceptual evaluations of NLDs for VBS. Its major contributions are highlighted as follows:

Harmonic Analysis and Synthesis on Polynomials (HASP): The Schaefer-Suen equation of nonlinear harmonic analysis and synthesis for power series nonlinearity [Sch70] [Sue70] is modified in closed-form equations to accommodate the

polynomial nonlinearities. The essence of this work is the development of *Triple Duality Theorem*, relating harmonic analysis, synthesis, and symmetries of the nonlinearities [OGC11]. Insights gained from the theoretical work are translated into the HASP algorithm. The HASP is used in harmonic analysis of NLDs as well as harmonic synthesis with a modulator to enhance low frequency performance of a directional loudspeaker [LOG09]. Fast Fourier transform (FFT) has well-known inherent limitations in spectrum analysis, such as sampling aliasing, spectrum leakage, and the need to use windows. However, all the mathematical analysis tools developed in this dissertation do not have the above-mentioned limitations of FFT. Due to the fast computational speed of FFT, it is mainly used for measurement systems, but may not be a suitable tool for analysis and synthesis, where individual tonal components can be exactly analyzed and synthesized.

Yeary-Oo-Gan Algorithm (YOGA): The intermodulation analysis is a well-researched topic in signal processing, communication, circuit design, and measurement. To develop the mathematical tool for the intermodulation analysis of NLDs, we build on the recent works of Yeary [Yea08]. In his work, the intermodulation distortion components are efficiently computed at the output of polynomial nonlinearities by two lemmas. However, there are two main limitations in his original technique. We solve both limitations and cut the computational load by half. The resultant algorithm is named YOGA and will be used in our intermodulation distortion analysis of selected NLDs in this dissertation.

Harmonic Addition Theorem (HAT): In literature, *Phasor Addition Theorem* is used to add multiple tonal components having the same frequency but different

amplitudes and phases [MSY03]. The trick is to convert the tones into phasors, add them in phasors and convert them back into trigonometric forms. However, the closed-form formula to perform such tricks is lagging in the literature. To fill this gap, we derive the closed-form and exact formula; we also provide a mathematical proof for this closed-form formula and convert the formula into an algorithm called HAT.

Design and Statistical Analysis of Distortion and Bass Listening Experiments: The statistical analysis on analysis of variance (ANOVA) and confidence intervals were carried out. Listening test results are correlated with perceptual model, *Rnonlin* [TMZ04]. The final contribution of this dissertation is the NLD curvature derivative analysis and categorization of the NLDs into the following groups: 1) Good; 2) Bass-Killers; 3) Not Recommended; and 4) Highly Distorted [OGH11].

1.3 Organization of the Dissertation

This dissertation is organized as shown in Figure 1.1.

In Chapter 2, research on the missing fundamental effect is reviewed, and blind and non-blind approaches on virtual bass enhancement and different NLDs used in previous studies are presented and reviewed.

In Chapter 3, *Triple Duality Theorem* is formulated with detailed and rigorous mathematical proofs. This work result in the HASP algorithm and its pseudocode is presented. HASP is used to perform both harmonic analysis and harmonic synthesis.

In Chapter 4, the intermodulation analysis of NLDs is performed. The limitations of precise intermodulation product computing technique (PIPCT) are

solved. In addition, the exponential computational growth of PIPCT is reduced by half, by using a proposed algorithm known as YOGA.

In Chapter 5, we design listening tests to evaluate distortion and bass intensity derived from different NLDs. Statistical analyses are also carried out, with a detailed correlation study between the experimental data and perceptual nonlinear distortion model, *Rnonlin*. In addition, a curvature derivative analysis is performed to classify NLDs by different degrees of bass enhancement and distortion perception.

Finally, Chapter 6 concludes the dissertation and discusses some future works related to our research topics and contributions.

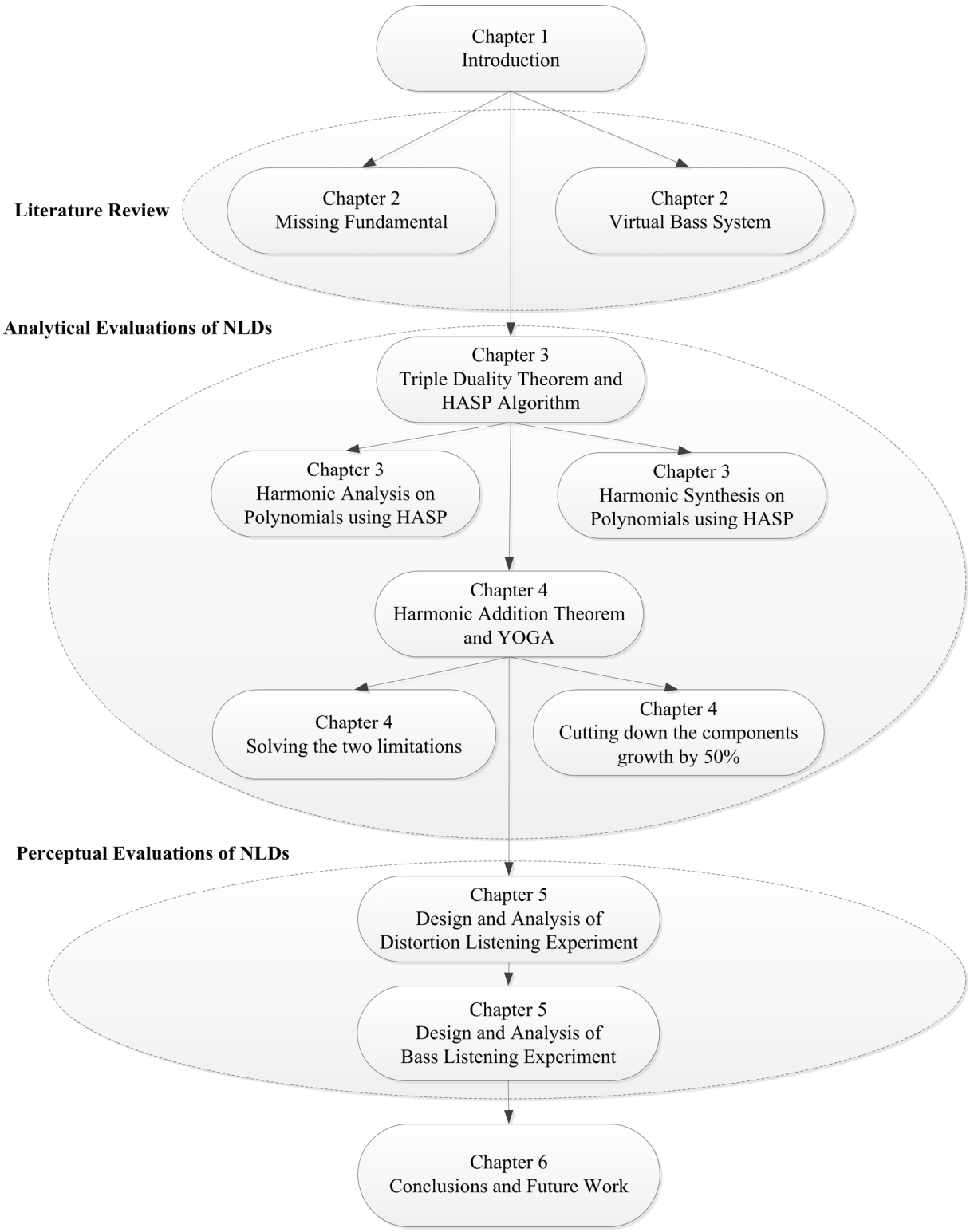

Figure 1.1: Links of dissertation chapters.

CHAPTER 2

THEORY OF MISSING FUNDAMENTAL AND VIRTUAL BASS SYSTEM

2.1 Introduction

This chapter presents the missing fundamental theory, which is used in the development of the virtual bass system (VBS). The missing fundamental theory originates from pitch perception research in psychoacoustics. The theory was later applied into audio engineering field to solve audio low-frequency (bass) reproduction problem of small- or medium-sized loudspeakers and low-quality headphones.

The organization of this chapter is as follows: Section 2.2 reviews the history and relevant findings on the theory of missing fundamentals in pitch perception. A review of the literature on VBS with regard to the theory of missing fundamentals is provided in Section 2.3, with a special emphasis on the nonlinear device (NLD)-based approach, which is the main thrust of this dissertation.

2.2 Pitch and the Theory of Missing Fundamentals

Pitch was defined as "attribute of auditory sensation in terms of which sounds may be ordered on a musical scale" in the 1960 American National Standard Institute (ANSI) standard [ANSI60] [Moo03, p. 3] [Ter00]. In the 1994 edition of ANSI standard [ANSI94], pitch is described as "that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high. Pitch depends mainly on the frequency content of the sound stimulus, but it also depends on the sound pressure and the waveform of the stimulus." In [Moo03], Moore explains the concept of pitch as ". . . a subjective attribute of a stimulus, and as such cannot be measured directly. However, for a sinusoid, the pitch is closely related to the frequency; the higher the frequency, the higher the pitch. For a more complex sound, the pitch is often investigated by asking the subject to adjust a sinusoid so that it has the same pitch as the complex sound. The frequency of the sinusoid is then taken as a measure of the pitch of the complex sound."

Other pitch-related terms used in this dissertation include:

- Fundamental frequency (F_0) is the rate at which a periodic wave repeats itself [Che10]; pitch is the perceptual correlate of F_0.
- Period of a wave is related to pitch; wave period is the reciprocal of F_0, so the longer the period, the lower the pitch. The range of F_0 for musical pitch is within about 30–5000 Hz. [Che10].
- Tone loosely refers to the perceived musical qualities of a sound. Less subjectively, however, pure tones and complex tones differ in that a *pure* tone is a sinusoidal waveform containing a single component, whereas a *complex* tone is a

periodic waveform with multiple components that are harmonically related to (i.e., are integer multiples of) F_0. According to Moore, "a sound that evokes a pitch is often called a 'tone,' especially when the pitch has a clear musical quality" [Moo03].

Figure 2.1 illustrates some pitch-evoking stimuli. Figure 2.1 (a) shows a pure tone stimulus of 100 Hz. A pure tone evokes pitch as well as complex tones; the pitch of a pure tone is similar to the pitch of complex tones of the same period. For example, except for Figure 2.1 (d), all stimuli have the same period in Figure 2.1 and the evoked pitch is hypothesized to be same, or very similar. Moreover, for complex tones, the fundamental frequency need not be present to evoke pitch. This important finding is known as the theory of the missing fundamental.

Figure 2.1 illustrates the missing fundamental concept. The waveforms in Figure 2.1 (b) and (c) have the same periodicity, but panel (c) has the fundamental frequency (F_0) while panel (b) does not. However, they both evoke the same pitch as that of the pure tone in panel (a). Panel (d) shows a complex tone with its partial frequencies, spaced over twice the length of the F_0, which is equal to the largest common divisor of the partial frequencies. Therefore, the pitch evoked by the complex tone in panel (d) is twice as high—an octave—than those of panels (a), (b) and (c). Notice the relationship between period and pitch. For higher pitch, the period of the stimulus is narrower and vice versa. Figure 2.1 (e) shows a complex tone with six partials in alternating sine/cosine phases, whereas panel (f) shows the same in cosine phase. Notice that, although spectral envelopes look the same in panels (e) and (f), their temporal envelopes look different.

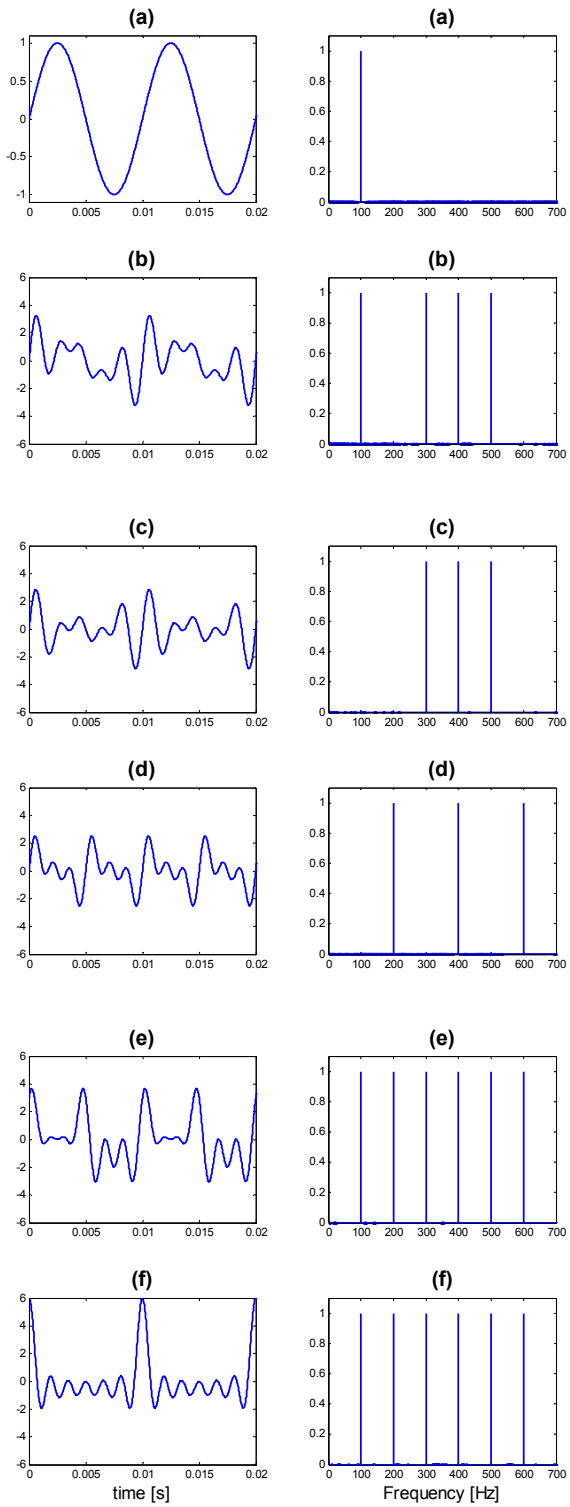

Figure 2.1: Time (left) and spectral (right) plots of various stimuli that evoke pitch (adapted and modified from [Che10]).

Fourier's theorem, which has had great impact on mathematics, physics, and engineering, states that any periodic wave can be constructed as the linear addition (superposition) of sinusoids with frequencies that are integer multiples of F_0 [Fou09a] [Fou09b].

In 1834, Georg Simon Ohm extended Fourier's theorem to the auditory domain by stating that the human ear is able to separate musical sound into pure harmonic tones, which can be expanded by means of Fourier's theorem [Lin66] [Tur77], [Ohm73]. This hypothesis is also stated as "a pitch corresponding to a certain frequency can only be heard if the acoustic wave contains power at that frequency" [HA09], and is known as Ohm's *acoustical* law [HA09].

August Seebeck was unconvinced [Ohm43] [Ohm44] [Ohm73] [See41] [See43] [See44a] [See44b]; he countered Ohm on this hypothesis with a series of experiments using synthesized stimuli where the fundamental was weak or absent [See41] [See43] [See44a] [See44b]. Seebeck raised the question of why harmonic series are dominated by their fundamental pitch [Plo67]; he concluded that "harmonics must reinforce the fundamental tone by direct physical action, and the presence of F_0 is not essential for hearing a pitch corresponding to this frequency." For this insight, Seeback should be regarded as the father of *missing fundamental* theory.

Hermann von Helmholtz revived the Ohm–Seebeck controversy, which had seemed to have ended in 1849 with Seebeck's death (at only forty-four years of age) [Tur77]. In 1856, Helmholtz publicly announced his acceptance of Ohm's hypothesis and recognized the hypothesis as Ohm's accoustical law [Tur77] [Hel54] [Hel56] [Hel95]. Alain de Cheveigné reasoned that Helmoltz' acceptance might be due to the

audibility of upper harmonics and desire of Helmoltz to fit together his auditory theory with the physical theory of Fourier [Che04]. Such was Helmoltz's professional prestige that his favoring of Ohm's acoustical law led to its general acceptance for about a century [POF05].

In 1924, Fletcher published a study concluding that pitch is the same even when F_0 and a large number of harmonics from a complex tone are eliminated, and hypothesized that the missing fundamental phenomenon is due to the nonlinearity of ear [Fle24]. In 1939–1940, Schouten, using an optical siren (an acoustical instrument for producing musical tones) investigated the missing fundamental phenomenon again and confirmed that the presence of F_0 was not a requirement for pitch perception [Sch38] [Sch39] [Sch40a] [Sch40b] [Sch40c]. Follow-up work by Licklider supported the missing fundamental theory [Lic51]. Licklider masked F_0 with noise and reported that perceived low pitch was unaffected [Lic54]. These experimental studies contradict Ohm's accoustical law, according to which F_0 must be present for a stimulus to sound low-pitched. Although Ohm's hypothesis of the ear as a frequency analyzer was accepted by Helmholtz, modern psychoacoustic experiments repeatedly confirm the perception of missing fundamentals. Plomp investigated further and concluded that the frequency components can be "heard out" only if their frequency separation exceeds the critical bandwidth [Plo64], [Plo68]. Once the psychoacoustic perception of missing fundamentals had been confirmed, researchers could establish models for existence region, dominance region, and pitch perception—for which reviews of the literature are herein presented.

Existance region

Among pure tones, frequencies above 4000–5000 Hz do not create a melodic pitch [POF05]. The human ear is also limited in its ability to perceive pitch for harmonic complex tones. Much of the initial research into the existence region of harmonic complex pitch was performed by Ritsma [Rit62] [Rit63]. Using amplitude-modulated sinusoids containing three sinusoidal components, he showed that residue pitch exists only for the existence region, which has both upper and lower limits [POF05]. His findings showed that there is no pitch for components with frequencies above 6000 Hz, which is therefore considered to be the upper limit of the existence region [POF05]. The lower limit depends on a sound's F_0 and harmonic spectral region. In Pressnitzer *et al.* [DPK01], the lower limit ranges from 35 Hz with a lower harmonic cutoff frequency of 200 Hz to around 300 Hz with a lower harmonic cutoff frequency of 3200 Hz [POF05]. The studies conducted by Moore and Rosen [MR79] and by Kaernbach and Bering [KB01] show that pitch produced by unresolved harmonics, though weaker than that produced by resolved harmonics, can still produce a perception of musical pitch. *Resolvability* refers to the ability of cochlear filtering to isolate individual partials of a complex sound [Che10].

Dominance region

Investigations of the dominance region raise the question of which harmonics play vital roles in pitch perception. Ritsma's study concluded that the frequency band that includes the third, fourth and fifth harmonics tends to dominate the perception of pitch for fundamental frequencies in the range of 100–400 Hz, for stimuli of at least

50 dB and above, and amplitude exceeding an minimum absolute level of about 10 dB above threshold [Rit67].

On the other hand, Plomp said that pitch is determined by the fourth and higher harmonics for F_0 up to about 350 Hz; and by the third and higher harmonics for F_0 up to about 700 Hz. His finding indicates that pitch is not only dependent on harmonic number but also on frequency [Plo67]. Pattern and Wightman, however, found that the dominance region was a function of the absolute as well as the relative frequency of the components [PW76].

Moore *et al.* in 1984 concluded that "the harmonics, which dominate in determining the pitch of a complex tone are those which are most discriminable. For complex tones with equal-amplitude harmonics these are usually the lowest harmonics (1-5). However, high harmonics can be dominant if they are the highest components in a complex, or if their level is incremented relative to other components" [MGS84]. Apart from these, Moore *et al.* in 1985 [MGP85] reported, "For complex tones with equal amplitude harmonics, and with fundamental frequencies of 100, 200, or 400 Hz, the dominant harmonics always lie within the first six harmonics. However, there are considerable individual differences in which of the lower harmonics are dominant, and in the distribution of dominance across harmonics. For some subjects the first, second, and third harmonics are more dominant than the fourth and fifth harmonics, contrary to the classical concepts of the dominance region."

Dai found that harmonics closest to 600 Hz tend to dominate; his results contradicted previous findings that the dominant harmonics were of fixed harmonic ranks regardless of their frequencies [Dai00].

Thus, although researchers could not agree on a unified theory for the dominance region, some agreements could be made among them. In summary, the dominant harmonics are generally from first to fifth, while the dominant harmonic number decreases with increasing F_0. For the very low F_0 (e.g., 50 Hz), harmonic numbers higher than fifth may be dominant [POF05].

Pitch perception

Another important research area is pitch models. Cheveigné's broad definition of a model is: a thing that represents another thing in some way that is useful [POF05]. For the pitch perception model, he argued that "one pitch model may predict behavioral data quantitatively, while another is easier to explain, and a third fits physiology more closely [POF05]. Criteria of quality are not one-dimensional, so models cannot always be ordered from best to worst."

Understanding of how humans perceive pitch, and how scientists might best model pitch perception has been a matter of intense debate for many years and is not yet resolved. However, most of that controversy is beyond the purview of this dissertation. Instead, the theory of missing fundamental is herein assumed to be valid; our main focus is how best it can be applied to solve the audio low-frequency bandwidth limitation problem in audio engineering perspective. However, for the sake of completeness, a brief review of pitch models found in literature on the missing fundamental theory is offered here.

The pioneer in this field is the *place* hypothesis by Helmholtz [Hel54]. According to Helmholtz, pitch is determined from the position of maximum excitation along the basilar membrane, within the cochlea, in the inner ear. However, the place hypothesis cannot explain the fact that a pure tone and complex tone (a sum of multiple pure tones with harmonic frequencies) can evoke the same pitch. Therefore, it fails to explain the theory of missing fundamental.

Modern pitch models can be generally classified into three major groups: temporal models, pattern recognition models, and autocorrelation models.

Temporal models include the works of de Boer [Boe56] and Schouten [Sch62] [POF05]. According to the temporal model, pitch is extracted from the periodicity of the acoustic waveform. De Boer [Boe56] suggested that the effective cue is the spacing between peaks of the temporal fine structure of the stimulus waveform; Schouten *et al.* [Sch62] reported that zero-crossings worked as well [POF05]. Wightman [Wig73] argued that such fine-structure theories are phase-sensitive, and mentioned that "since the relative phases of the spectral components of a waveform determines its temporal fine structure, changes in these phase relations are expected to affect pitch." In addition, virtual pitches (i.e., missing fundamentals) were still heard even when the harmonics could not interact because they were presented sequentially or to different ears. Due to the above-mentioned two problems, the debate has now shifted to pattern recognition and autocorrelation models [Che10].

The best known pattern recognition models are of Goldstein [Gol73], Wightman [Wig73], and Terhardt [Ter74]. In pattern recognition models, the partials of a periodic sound form a pattern of frequencies. Two parts are involved in pattern

recognition models: one produces the *pattern* and the other looks for a match within *templates* [POF05]. De Boer describes the pattern recognition in his Ph.D. dissertation [Boe56] and later promoted by Goldstein [Gol73], Wightman [Wig73], and Terhardt [Ter74]. The drawback is that the pattern recognition models work only when partials are resolved. However, pitch can also be heard using stimuli for which there are no resolved partials. Pattern recognition models cannot explain such a pitch.

The pioneer work in autocorrelation model was done by Licklider [Lic51]. In autocorrelation models, each sample of the waveform is used as an event. Each is compared to every other sample and the inter-event interval that gives the best match on average indicates the period. If the samples are equal, their products are large, and the *autocorrelation function* (ACF) has peaks at the period and its multiples. The peak is the cue to the pitch. Licklider's duplex theory of pitch perception states that the auditory system employs both frequency analysis (in the cochlea) and autocorrelation analysis (in the central auditory system). The latter is therefore an analysis of trains of nerve impulses that have been transformed by the cochlea from acoustic stimuli. In other words, in Lickerlider's model, the ACF is calculated within the auditory nervous system, for each channel of the auditory filter bank [Che04]. Licklider's theory was later simplified and implemented computationally by Meddis and Hewitt [MH91a]. Meddis and Hewitt's model was categorized as a special case of the four-stage pitch extraction general model by Meddis and O'Mard [MO97]. There exists the *multiple mechanism hypothesis,* which states that pitch is processed by different mechanisms: *pattern recognition* for resolved and *autocorrelation* for

unresolved. However, current understanding of how we perceive pitch is still incomplete [Che10].

2.3 Application of the Missing Fundamental Theory in the Low-Frequency Audio Bandwidth Limitation Problem

Due to physical limitation of audio reproduction devices, small loudspeakers and low quality headphones cannot reproduce low frequency components in audio signals. This problem is commonly known as the low-frequency audio bandwidth limitation problem. For loudspeaker acoustics, this bass reproduction problem can be summarized by two loudspeaker modeling equations as follows [LA04]:

$$\eta \propto \left(\frac{S}{m} \right)^2, \tag{2.1}$$

$$f_0 = \frac{1}{2\pi} \sqrt{\frac{k_t}{m}}, \tag{2.2}$$

where η is the energy conversion efficiency that is defined as a ratio between time-averaged acoustically radiated power and time-averaged electrical power, and f_0 is the resonance frequency of the loudspeaker. In (2.1) and (2.2), m is the mass, S is the loudspeaker cone area, and k_t is the total compliance that combines suspension and cabinet influence. For small loudspeakers, cone area and mass are small, and total compliance is high. Therefore, the resonance frequency is high, while energy conversion efficiency is low. To further reduce the resonance frequency of loudspeaker, the total compliance must be decreased. Otherwise, the mass must be

increased, requiring larger cabinet volume. Therefore, small loudspeakers physically limit bass reproduction.

Apart from this physical limitation, Figure 2.2 shows that more amplification of sound level is required to achieve the same loudness perception for low-frequency signals compared to mid-frequency signals. In order to generate higher sound pressure level (SPL), larger cone area and mass are needed in addition to higher power requirements. In fact, very small loudspeakers cannot reproduce bass frequencies at all [BL05], and their output sound quality is unpleasing [Bec02].

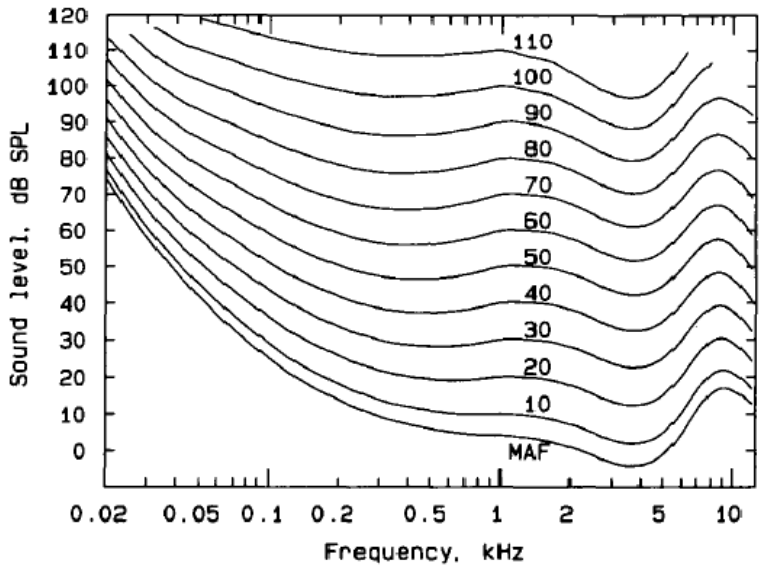

Figure 2.2: Equal loudness contours depicting the variation in loudness with frequency. MAF indicates Minimum Audible Field. Source from [Moo03].

There are two ways to overcome this bandwidth limitation problem: physical bandwidth extension and psychoacoustic bandwidth extension. Figure 2.3 illustrates these concepts. The thin arrows indicate the shifting of energy when extending low-frequency bandwidth. The wide arrows point to the direction of audio bandwidth

extension (audio bandwidth extension direction is to the left, i.e., the low frequency bandwidth extension).

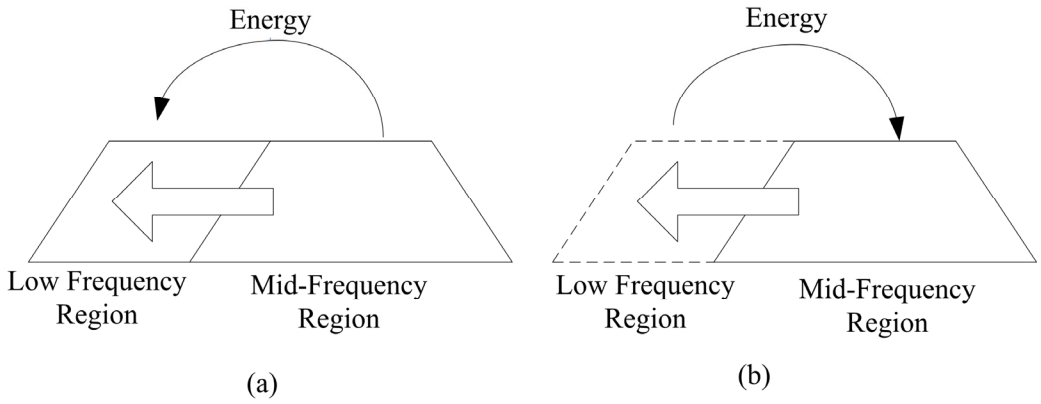

Figure 2.3: (a) Physical audio bandwidth extension. (b) Psychoacoustic bandwidth extension. Planes indicate audio bandwidth; dotted line extension indicates the bandwidth is to be extended virtually. (Adapted and modified from [LA04]).

In the direct approach, physical bandwidth extension boosts low frequency components by adding signal components in the lower frequency region, as shown in Figure 2.3 (a). Nevertheless, the loudspeaker must be able to reproduce the added audio low frequency components for this approach to be effective.

On the other hand, the psychoacoustic bandwidth extension technique (Figure 2.3 (b)) injects useful signal components in mid-frequency range, where majority of loudspeakers have relatively flat responses. By applying the "missing fundamental" as described in Section 2.1, the bandwidth extension takes place psychoacoustically. Even though physical bass frequencies may be missing, listeners can still perceive bass enhancement in this latter mode. The following sections introduce the techniques behind the psychoacoustic signal processing approach to enhance bass perception.

Psychoacoustic signal processing techniques are employed to enhance bass virtually [LA04] [GKT01] [LA02] [BC99] [Kar02] [OG09b]. These techniques are also known as virtual bass (VB) processing and can be categorized mainly as blind and non-blind signal processing. The non-blind VB signal processing algorithms extract characteristics from incoming signals; the blind VB signal processing algorithms do not exploit incoming signal characteristics (such as pitch information), but passively modify the signal with certain transformations or rules. Mono-channel processing of the blind/non-blind virtual bass system (VBS) is diagrammed in Figure 2.4; the stereo channel processing diagram is shown in Figure 2.5.

As shown in both Figures 2.4 and 2.5, bass frequencies are filtered by a low-pass filter (LPF); its output is fed into a signal processing module that can be either blind or non-blind. Audio high-frequency components are filtered by a high-pass filter (HPF) while LPF extracts audio low-frequency components, which lie in the bass frequency range (60–250 Hz) or sub-bass range (16–60 Hz). Generally, these frequency ranges cannot be reproduced by small loudspeakers due to their physical limitations. Figure 2.5 depicts the basic stereo VBS structure. Since bass is considered non-directional due to the larger wavelength (see Table 2.1), the incoming left and right stereo channels are combined as a middle channel. This mono channel is then low-pass filtered and fed into a signal processing module. The processed signal is added back into the high-passed filtered left and right channels to create a stereo virtual bass effect.

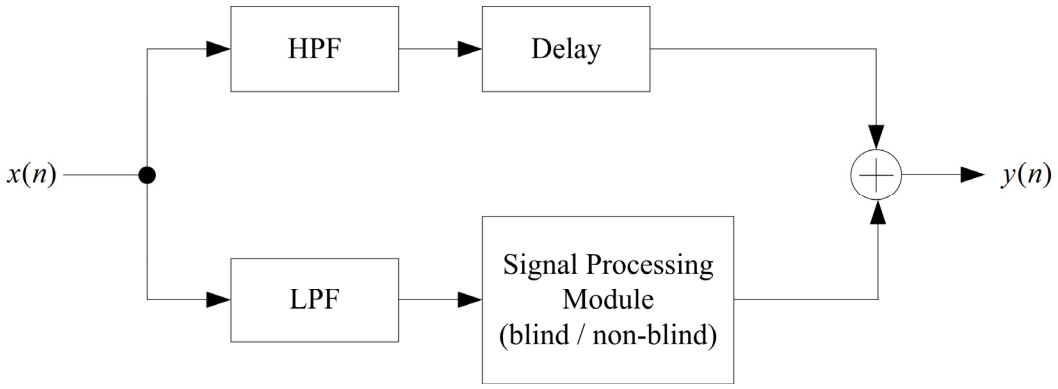

Figure 2.4: The blind/non-blind VBS in simplified blocks (mono channel).

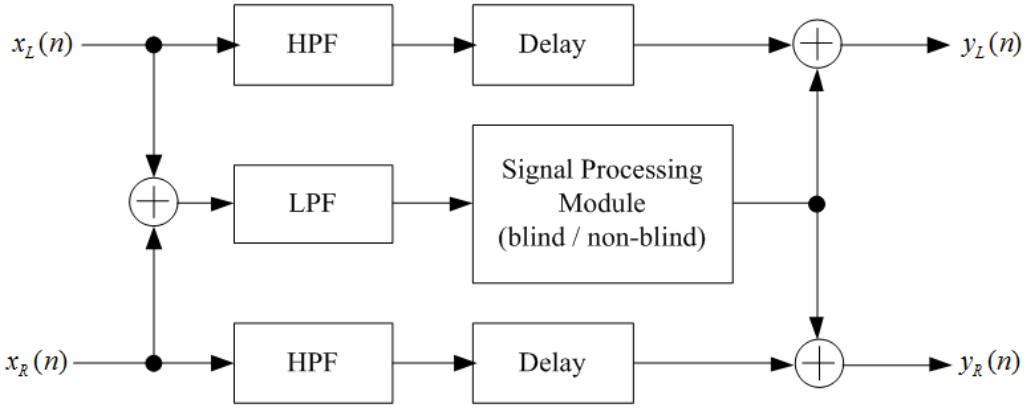

Figure 2.5: The blind/non-blind VBS in simplified blocks (stereo channels).

Table 2.1 shows the classification of audible frequency ranges and their wavelengths. Note that the bass and sub-bass regions have wavelengths that are long and behave as if omnidirectional. The higher frequency ranges represent the presence and brilliance characteristics of audio signal while the frequencies have shorter wavelength that behave as more directional sound waves.

Table 2.1: Audio frequency ranges and wavelengths (speed of sound = 343 m/s at 20° C) [Ows99].

Range	Name	Wavelength (m)
16 – 60 Hz	Sub-Bass	21.4 – 5.7 m
60 – 250 Hz	Bass	5.7 – 1.372 m
250 – 2 kHz	Low Mids	1.372 – 0.172 m
2 kHz – 4 kHz	High Mids	0.172 – 0.086 m
4 kHz – 6 kHz	Presence	0.086 – 0.0572 m
6 kHz – 16 kHz	Brilliance	0.0572 – 0.0214 m

2.3.1 VBS non-blind signal processing module

The details of VBS using non-blind signal processing module are further illustrated in Figure 2.6. The signal processing module consists of two sub-modules, namely pitch detector and harmonic generator. The pitch detector detects the pitch (or fundamental frequency F_0), and multiple pitches of the audio signal. The detected pitches are fed into the harmonic generator, which generates harmonics that are integral multiples of the detected pitches. The objective of the non-blind signal processing module is to create the missing fundamental effect so that virtual bass can be perceived by the listener.

In Figure 2.7, F_1 is the first harmonic and it is in the same frequency bin as F_0, but magnitude may not be the same; ΔF is the gap between subsequent harmonics and is equal to the frequency bin of F_0. The main reason for using a pitch detector is to avoid audio intermodulation distortion (IMD) products, which are generally present in blind VBS. (Blind VBS will be described in more detail in Section 2.3.2.)

Musical signals generally consist of different instrumentations and vocals, which make the pitch detection a challenging task. That is why multiple pitch

detection plays a key role in the success of non-blind VBS. In the literature, multiple pitch detection is a branch of auditory scene analysis (ASA) [Bre94] [Kla06]. Multiple pitch detection can be implemented through different methods, which vary in their ability to distinguish different aspects of multiple pitches in real time.

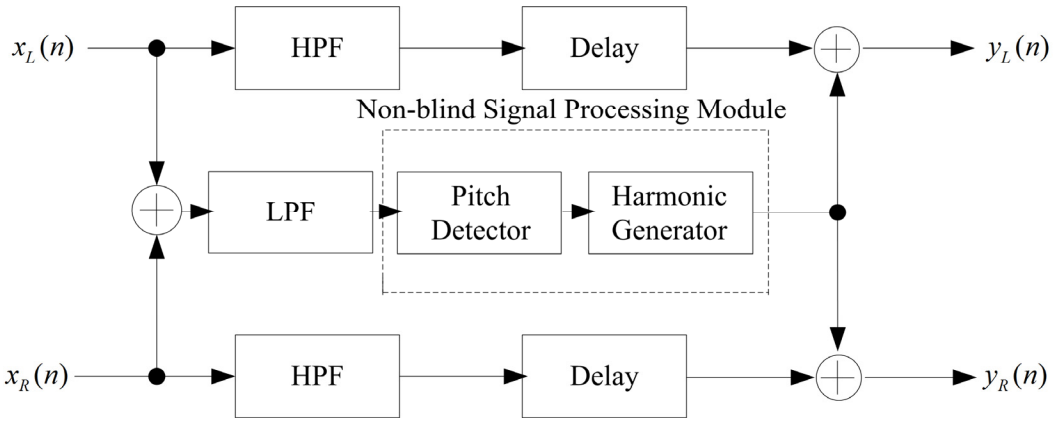

Figure 2.6: VBS non-blind signal processing making use of pitch detector and harmonic generator.

The generation of harmonics by the harmonic generator itself is another challenging task and open to future research, because the perception of music is mainly based on three pillars—pitch, timbre and loudness. The harmonic generator's main task is to create the virtual pitch effect by generating harmonics. However, harmonic amplitudes play a major role in timbre perception [RP82] [MR79]. If harmonics are generated without consideration of perceptual quality, severe distortion can be perceived and the naturalness of the original audio signal can be destroyed. The recent work of Lee *et al.* on this topic did not cover how to generate harmonics for psychoacoustic bass enhancement without sacrificing audio quality [LSP08], contending that blind VBS generates IMD products that are perceived as pure

distortion in the ears. However, the masking effect also takes place in the human auditory system when music is being played. The generated IMDs may not be heard even though they exist, due to the masking effect. Therefore, a tone may be physically present, but not perceived.

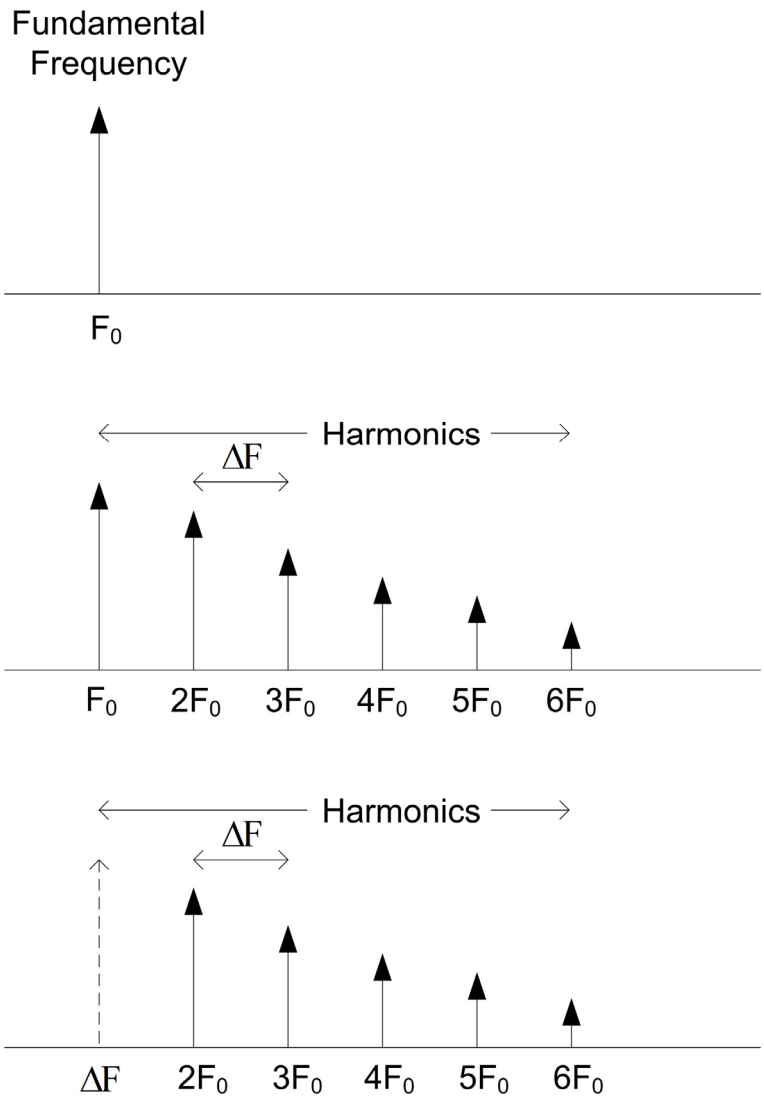

Figure 2.7: The missing fundamental phenomenon. F_0 is perceived by the listener even though it may be absent in the harmonics. This is also called virtual pitch.

2.3.2 VBS blind signal processing module

VBS blind signal processing techniques can be categorized into two approaches: the frequency domain approach (diagrammed in Figure 2.8) and the time domain approach (diagrammed in Figure 2.9).

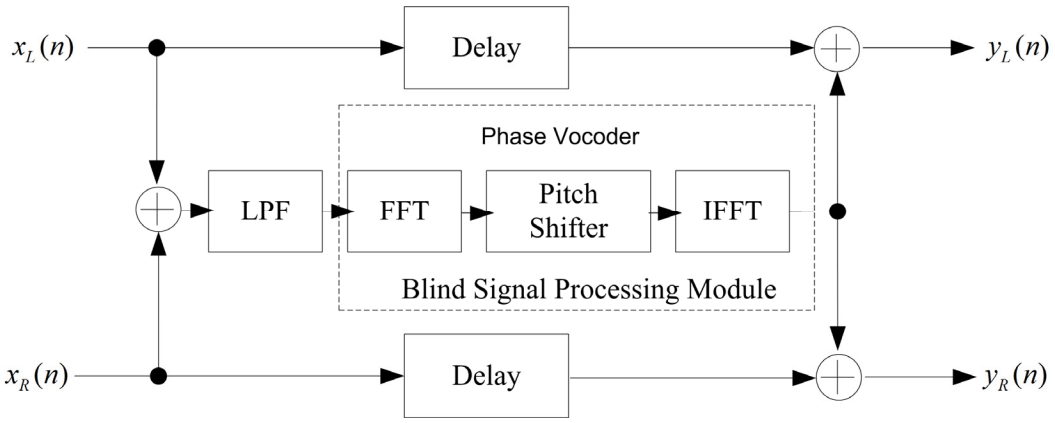

Figure 2.8: The frequency domain approach using a *phase vocoder*.

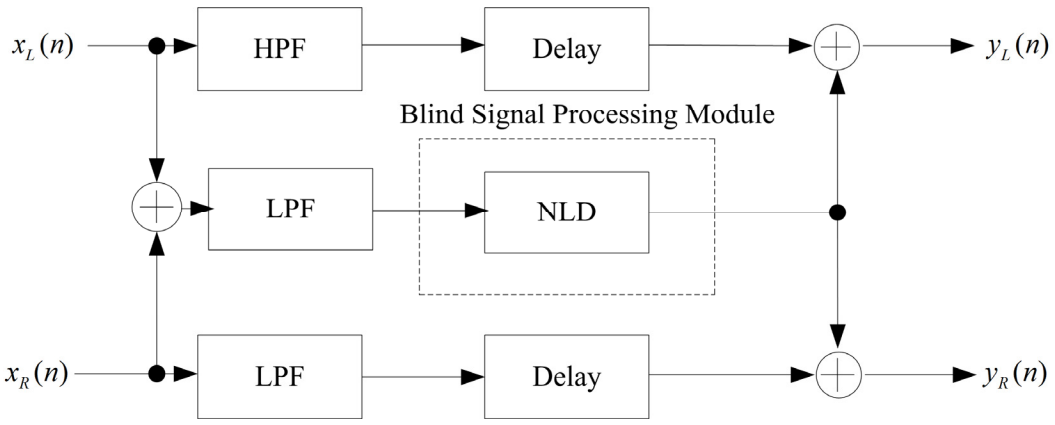

Figure 2.9: The time domain NLD-based VBS blind signal processing framework. .

The VBS frequency domain approach using phase-vocoder [LD99] was proposed by Bai *et al.* [BL06]. Low frequencies, which might not be reproduced by the small loudspeakers, are low-pass-filtered and pitch-shifted to the mid-range audio frequency region. Pitch shifting and frequency shifting differ in that pitch shifting multiplies a constant to every frequency components that is to be shifted, whereas frequency shifting adds a constant to every frequency components. The advantage of pitch shifting over frequency shifting is that phase coherence is preserved, according to Bai *et al.* [BL06]. Since NLD is not used in the frequency domain approach, there is no IMD. Bai *et al.* argue that IMD, generated by NLD, may be perceived as distortion; therefore, their approach does not have this problem. However, their method has two limitations. The first one is that the signals are processed frame-by-frame because FFT is used. If phase propagation is not properly dealt with, the phasiness artifact will be audible [BL06]. The second is that as pitch is shifted to mid-range and signal amplitudes are adjusted to match the loudness, the naturalness of the original signal may be lost.

The NLD-based approach is diagrammed in Figure 2.9. The system makes use of nonlinearity to generate harmonics, which in turn create virtual bass effects, while NLD is a nonlinearity or memoryless nonlinear system. The main difference between linear and nonlinear systems results from the superposition principle. All linear systems hold the superposition property, whereas nonlinear systems do not.

The superposition principle can be further generalized as follows [Sch80].

$$y_n(t) = T\{x_n(t)\}, \text{ for } n = 1, 2, 3, \ldots, N, \tag{2.3}$$

$$x(t) = \sum_{n=1}^{N} \alpha_n x_n(t), \tag{2.4}$$

where α_n are arbitrary constants (real or complex) and N is the number of inputs. Then, the superposition principle is

$$y(t) = T\{x(t)\} = T\left\{\sum_{n=1}^{N} \alpha_n x_n(t)\right\} = \sum_{n=1}^{N} \alpha_n T\{x_n(t)\} = \sum_{n=1}^{N} \alpha_n y_n(t). \tag{2.5}$$

Equation (2.5) implies that the linear system's response to the linear combination of inputs is equivalent to the linear combination of responses to each individual input. The implication of this superposition principle is the development of the convolution integral. This convolution integral means that if the response $h(t)$ for a specific input $\partial(t)$, unit impulse, is known, then the response to any input that is a linear combination of translations of $\partial(t)$ is available. That is,

$$h(t) = T\{\partial(t)\}, \tag{2.6}$$

$$y(t) = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau. \tag{2.7}$$

The transfer function, or frequency response of the linear time-invariant (LTI) system can be derived from (2.7). Another major difference between linear and nonlinear systems is that linear systems can never introduce new frequency components at the output, whereas nonlinear systems always do. This property can be described mathematically in the following derivations. Using (2.7),

$$x(t) = e^{j\omega t}, \tag{2.8}$$

$$y(t) = \int_{-\infty}^{\infty} h(\tau)e^{j\omega(t-\tau)}d\tau = e^{j\omega t}\int_{-\infty}^{\infty} h(\tau)e^{-j\omega\tau}d\tau = H(j\omega)e^{j\omega t}, \tag{2.9}$$

where $H(j\omega)$ is the Fourier transform of the unit impulse response, $h(t)$. The Fourier transform and its inverse are

$$H(j\omega) = \int_{-\infty}^{\infty} h(\tau) e^{-j\omega\tau} d\tau, \tag{2.10}$$

$$h(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(j\omega) e^{j\omega\tau} d\omega. \tag{2.11}$$

If the input is

$$x(t) = \sum_n \alpha_n e^{j\omega_n t}, \tag{2.12}$$

using (2.5) and (2.9),

$$y(t) = \sum_n \alpha_n H(j\omega_n) e^{j\omega_n t}. \tag{2.13}$$

Then, if the input signal is a cosine tone,

$$x(t) = A\cos\omega_0 t = \frac{A}{2} e^{j\omega_0 t} + \frac{A}{2} e^{-j\omega_0 t} = \frac{A}{2} x_1(t) + \frac{A}{2} x_2(t), \tag{2.14}$$

where $x_1(t) = e^{j\omega_0 t}$ and $x_2(t) = x_1^*(t) = e^{-j\omega_0 t}$, then

$$y_1(t) = T\{x_1(t)\} = H(j\omega_0) e^{j\omega_0 t}, \tag{2.15}$$

$$y_2(t) = T\{x_2(t)\} = H(-j\omega_0) e^{-j\omega_0 t} = H^*(j\omega_0) e^{-j\omega_0 t} = y_1^*(t), \tag{2.16}$$

$$y(t) = T\left\{\frac{A}{2} x_1(t) + \frac{A}{2} x_2(t)\right\} = \frac{A}{2} y_1(t) + \frac{A}{2} y_2(t) = \frac{A}{2} y_1(t) + \frac{A}{2} y_1^*(t)$$

$$= \frac{A}{2} H(j\omega_0) e^{j\omega_0 t} + \frac{A}{2} H^*(j\omega_0) e^{-j\omega_0 t}$$

$$= \frac{A}{2} |H(j\omega_0)| e^{j\theta(\omega_0)} e^{j\omega_0 t} + \frac{A}{2} |H(j\omega_0)| e^{-j\theta(\omega_0) t} e^{-j\omega_0 t}$$

$$= \frac{A}{2} |H(j\omega_0)| e^{j[\omega_0 t + \theta(\omega_0)]} + \frac{A}{2} |H(j\omega_0)| e^{-j[\omega_0 t + \theta(\omega_0)]}$$

$$= \frac{A}{2} |H(j\omega_0)| \left[e^{j[\omega_0 t + \theta(\omega_0)]} + e^{-j[\omega_0 t + \theta(\omega_0)]} \right]$$

$$= A\left|H(j\omega_0)\right|\cos[\omega_0 t + \theta(\omega_0)] \tag{2.17}$$

Equation (2.17) expresses a special characteristic of a LTI system when LTI system is fed by a single tone. The only changes are the single tone's amplitude and its phase. No additional frequency component is generated at the output. In contrast, in nonlinear systems, for which the superposition principle does not hold true, there are harmonics at the output when the system is fed by a single tone.

To show how a memoryless NLD generates harmonics when fed by a single tone, let the NLD be

$$y(t) = [x(t)]^n, \tag{2.18}$$

where $x(t) = A\cos\omega_0 t$ is the input signal, and $y(t)$ is the output of nth order monomial NLD. From (2.18),

$$y(t) = \left[A\cos\omega_0 t\right]^n = A^n\left[\cos\omega_0 t\right]^n. \tag{2.19}$$

Using the following trigonometric identities,

$$\cos^n\theta \equiv \begin{cases} \dfrac{2}{2^n}\displaystyle\sum_{k=0}^{\frac{n-1}{2}}\binom{n}{k}\cos\big((n-2k)\theta\big) & \text{if } n \text{ is odd,} \\ \dfrac{1}{2^2}\binom{n}{n/2} + \dfrac{2}{2^n}\displaystyle\sum_{k=0}^{\frac{n}{2}-1}\binom{n}{k}\cos\big((n-2k)\theta\big) & \text{if } n \text{ is even.} \end{cases} \tag{2.20}$$

Substituting $\theta = \omega_0 t$ and by using (2.19), (2.20) becomes

$$y(t) = A^n\left[\cos\omega_0 t\right]^n \equiv \begin{cases} \dfrac{2A^n}{2^n}\displaystyle\sum_{k=0}^{\frac{n-1}{2}}\binom{n}{k}\cos\big((n-2k)\omega_0 t\big) & \text{if } n \text{ is odd,} \\ \dfrac{A^n}{2^2}\binom{n}{n/2} + \dfrac{2A^n}{2^n}\displaystyle\sum_{k=0}^{\frac{n}{2}-1}\binom{n}{k}\cos\big((n-2k)\omega_0 t\big) & \text{if } n \text{ is even.} \end{cases} \tag{2.21}$$

By using (2.21), the harmonic generation pattern of monomial NLD can be seen, when the NLD is fed by a single cosine tone. Table 2.2—which can be derived from (2.18) and (2.21) — illustrates this harmonic generation pattern up to the fifth order of monomial NLD.

Table 2.2: Harmonic generation pattern of monomial NLD of order *n*, fed by a single tone.

n	$y(t)=[x(t)]^n, x(t)=A\cos\omega_0 t, \omega_0=2\pi f_0 t$	DC	ω_0	$2\omega_0$	$3\omega_0$	$4\omega_0$	$5\omega_0$
2	$A^2\left[\frac{1}{2}+\frac{1}{2}\cos(2\omega_0 t)\right]$	$\frac{A^2}{2}$		$\frac{A^2}{2}$			
3	$A^3\left[\frac{3}{4}\cos(\omega_0 t)+\frac{1}{4}\cos(3\omega_0 t)\right]$		$\frac{3A^3}{4}$		$\frac{A^3}{4}$		
4	$A^4\left[\frac{3}{8}+\frac{4}{8}\cos(2\omega_0 t)+\frac{1}{8}\cos(4\omega_0 t)\right]$	$\frac{3A^4}{8}$		$\frac{4A^4}{8}$		$\frac{A^4}{8}$	
5	$A^5\left[\frac{10}{16}\cos(\omega_0 t)+\frac{5}{16}\cos(3\omega_0 t)+\frac{1}{16}\cos(5\omega_0 t)\right]$		$\frac{10A^5}{16}$		$\frac{5A^5}{16}$		$\frac{A^5}{16}$

From Table 2.2, the following are observed:

1. A monomial NLD is a nonlinear system without memory, where the superposition principle does not hold true. Harmonics are generated when the system is fed by a single tone.
2. The odd-ordered monomial NLD generates only odd harmonics, whereas the even-ordered monomial NLD generates only even harmonics and DC.
3. The highest harmonic's number is equal to the order of monomial NLD—e.g., the second-order monomial NLD generates up to the second harmonics, the third-order monomial NLD generates up to the third harmonics, etc.

The above three observations were reported in [OG08a], The mathematical proofs of the relationships between harmonic number and polynomial order and

symmetries of nonlinear function will be presented in Chapter 3 as the Triple Duality Theorem. Moreover, the theoretical developments of tonal analysis algorithms are presented in Chapter 3 and 4 in detail.

If a nonlinear system is time-invariant and uses memory, the relationship between the input and the output can be expressed as

$$y = h_0 + \int_{-\infty}^{\infty} h_1(\tau_1)x(t-\tau_1)d\tau_1 + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(\tau_1,\tau_2)x(t-\tau_1)x(t-\tau_2)d\tau_1 d\tau_2$$

$$+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_3(\tau_1,\tau_2,\tau_3)x(t-\tau_1)x(t-\tau_2)x(t-\tau_3)d\tau_1 d\tau_2 d\tau_3 + \ldots$$

$$+\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} h_n(\tau_1,\tau_2,\ldots,\tau_n)x(t-\tau_1)x(t-\tau_2)\ldots x(t-\tau_n)d\tau_1\ldots d\tau_n \qquad (2.22)$$

where $h_n(\tau_1,\ldots,\tau_n) = 0$ for any $\tau_i < 0, i = 1,2,\ldots,n$. The $h_n(\tau_1,\ldots,\tau_n)$ are multi-dimensional functions and are called the *Volterra kernels*. To simplify the complex expression of (2.22), let $H_n[x(t)]$ be an n^{th}-order Volterra operator that is defined as

$$H_n[x(t)] = \int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} h_n(\tau_1,\ldots,\tau_n)x(t-\tau_1)\ldots x(t-\tau_n)d\tau_1\ldots d\tau_n \qquad (2.23)$$

Hence, (2.22) can be expressed as

$$y(t) = h_0 + H_1[x(t)] + H_2[x(t)] + \ldots + H_n[x(t)], \qquad (2.24)$$

and can be visualized as in Figure 2.10.

However, a non-linear system can be approximated as a polynomial function if it is memory-less. This polynomial function is also named as a polynomial NLD [OG08b] [LOG09]. The polynomial NLD or function can be expressed as

$$y(t) = \hat{h}_0 + \hat{h}_1 x(t) + \hat{h}_2[x(t)]^2 + \ldots + \hat{h}_Q[x(t)]^Q, \qquad (2.25)$$

where $x(t)$ is the input signal, $y(t)$ is the output signal, and $\hat{h}_i, i = 0,\ldots,Q$ are the polynomial coefficients. Due to the multi-dimensional nature, the nonlinear systems

with memory as expressed in (2.22) are harder to analyze and are not used in VBS. So far, NLDs in VBS are memoryless and can be approximated as in (2.25), and visualized as in Figure 2.11.

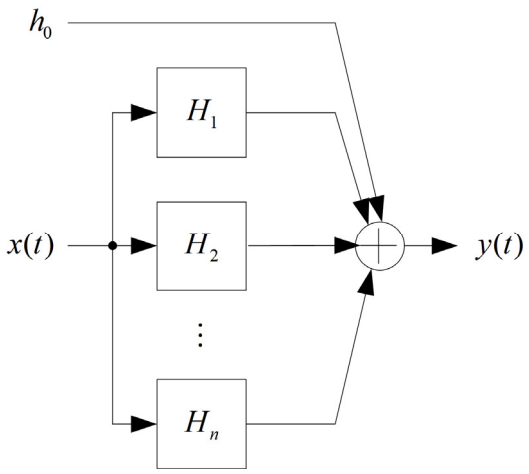

Figure 2.10: Nonlinear time-invariant system with memory as expressed in the equation (2.24).

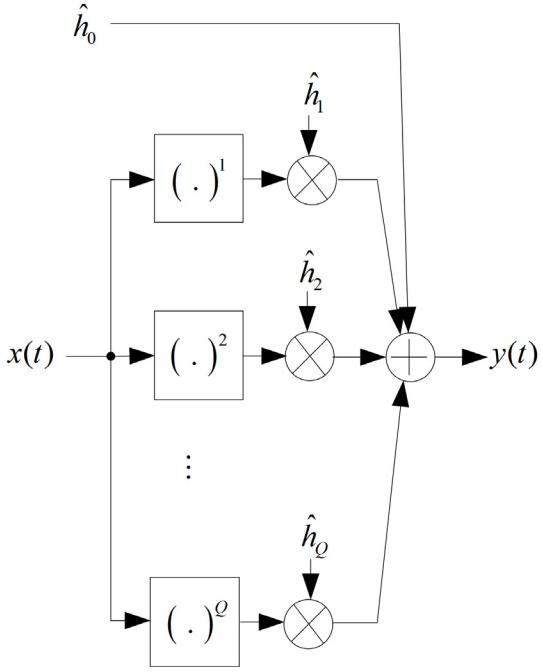

Figure 2.11: Polynomial memoryless NLD, as expressed in the equation (2.25).

2.4 Conclusions

In this chapter, we have reviewed the theory of missing fundamental, including the Ohm-Seeback-Helmholtz controversy, existence region and dominance region of the pitch perception, and categories of pitch models. Finally, the application of missing fundamental theory to VBS research has been analyzed. The literature detailing VBS research has also been reevaluated with special emphasis on NLD-based blind signal processing technique. In the next chapter, harmonic analysis and synthesis on NLD will be assessed together with the development of the Triple Duality Theorem that links to the relationship between symmetries of nonlinear functions, and harmonic analysis and synthesis.

CHAPTER 3

HARMONIC ANALYSIS AND SYNTHESIS IN NONLINEAR DEVICES

3.1 *Introduction*

In Chapter 2, the theory of missing fundamental was previewed, and the signal processing techniques to apply missing fundamental in VBS were introduced. Thus, to design sound systems or audio signal processing systems, such as VBS in this dissertation, audio researchers must work with various synthesis tools, and sometimes need to create new tools. To analyze sound, they have an array of choices, including traditional discrete Fourier transform (DFT), spectrograms, and wavelets etc. However, every tool has its own strengths and weaknesses. Because sound systems, such as audio amplifiers and loudspeakers, are nonlinear [Cor11] [Kli06], objective tonal analysis metrics such as total harmonic distortion (THD) and intermodulation distortion (IMD) are commonly used to analyze or quantify any distorted sounds they produce [Cab99]. In contrast, to create enjoyable nonlinear audio effects [Bar98] ("Fuzzbox," e.g. [Zol02] [Bod84]), or to overcome audio bandwidth limitations [LA02], such as VBS [OG10] [OGL10], nonlinearities can be carefully inserted into the signal processing chain.

A sound synthesis technique heavily used in computer music synthesis (CMS) [DJ97] and sound design [Far10] is waveshaping [Bur79] [Arf79] [Sch70] [Sue70], which uses Chebyshev polynomials of the first kind, to synthesize sounds with user-specified harmonic spectra.

This chapter presents the mathematical and computational works on harmonic analysis and synthesis of NLDs. The duality relationships between harmonic sound analysis, synthesis, and nonlinear function symmetries are derived in Section 3.2. An algorithm to perform harmonic analysis and synthesis is proposed with case studies. Generalized harmonic analysis using the Taylor series approximation technique on arc-tangent square root (ATSR) and a family of parameterized exponentials (EXPs) NLDs are presented in Section 3.3. In addition to harmonic analysis, harmonic synthesis with modulation technique is investigated with a subjective listening test for the bass enhancement of a parametric loudspeaker in Section 3.4. Finally, Section 3.5 summarizes and concludes this chapter.

3.2 Triple Duality Theorem

In [Far10], Farnell stated that any sound can be generated from first principles of sound analysis and synthesis. Inspired by his slogan, the objectives of this section are

1. To show mathematically how nonlinear sound analysis and synthesis are dual in nature;
2. To relate analysis–synthesis to symmetries of nonlinear functions; and
3. To transform design equations into algorithms for audio researchers.

3.2.1 Power series and polynomial nonlinearities

Let a power series be defined as

$$y = f(x) = \sum_{n=0}^{\infty} h_n x^n, \tag{3.1}$$

where $h_n \in \mathbb{R}$ represents the power series coefficient of the nth term, x, and y denote input and output of the power series, respectively, and $f(x)$ represents a power series function. The symbol $\mathbb{R}$ denotes the set of all real numbers. Because the upper bound of summation in (3.1) is the positive infinity, the degree of the power series is unbounded. Let us define a polynomial nonlinearity,

$$\hat{y} = \hat{f}(x) = \sum_{n=0}^{Q} \hat{h}_n x^n, \tag{3.2}$$

under the assumption that $\lim_{Q\to\infty} \hat{y} = y$ where $\hat{h}_n$, Q, and $\hat{y}$ denote the coefficient of the nth term, the highest degree, and the output of the polynomial nonlinearity, respectively. Like $f(x)$ in (3.1), $\hat{f}(x)$ denotes a polynomial function.

3.2.2 Harmonic analysis and synthesis of power series

Let us define the input single tone with arbitrary amplitude $A \in \mathbb{R}$ and zero initial phase as

$$x(t) = A\cos(\omega t), \tag{3.3}$$

where ω is the angular frequency in radian per second and t is the time in seconds. Since (3.1) and (3.2) are memoryless (static or time-independent) systems, we omit the time parameter t throughout this chapter.

When the single tone (3.3) is applied to the input of (3.1), the output is a Fourier series:

$$g(t) = \frac{c_0}{2} + \sum_{k=1}^{\infty} c_k \cos(k\omega t), \tag{3.4}$$

where $c_k \in \mathbb{R}$ are Fourier coefficients if $k \in \mathbb{Z}^+$, and harmonic amplitudes if $k \in \mathbb{N}$. The symbols $\mathbb{Z}^+$ and $\mathbb{N}$ denote the set of all positive integers including zero and the set of all natural numbers, respectively. Note that in (3.4), $c_0/2$ is the DC component.

Two analysis and synthesis problems are defined as follows:

(i) Given (3.3) and (3.1), (3.4) can be expressed as a function of A and h_n such that $c_k = \Im_a(A, h_n)$ for $k \in \mathbb{Z}^+$, where $\Im_a(.)$ is an analysis function that generates c_k from h_n and A.

(ii) Given (3.3) and (3.4), (3.1) can be expressed as a function of A and c_k such that $h_n = \Im_s(A, c_k)$ for $n \in \mathbb{Z}^+$, where $\Im_s(.)$ is a synthesis function that generates h_n from c_k and A.

Note that $\Im_a(.)$ and $\Im_s(.)$ are harmonic analysis and synthesis equations of the power series, respectively. Let $n=k+2j$ for analysis and $k=n+2j$ for synthesis. Schaefer derived $\Im_a(.)$ and $\Im_s(.)$ for the special case of $A=1$ using Chebyshev polynomials (see (15) and (18) of [Sch70]). Rearranging the terms in the original equations of Schaefer to seek some interesting patterns that are to be used in the next section, Schaefer harmonic analysis and synthesis equations can be rewritten as

$$c_k = \Im_a(A=1, h_n) = \frac{1}{2^{k-1}} \sum_{j=0}^{\infty} \frac{h_n}{2^{2j}} \binom{k+2j}{j}, \tag{3.5}$$

for $n = k + 2j$ and $k = 0, 1, 2, \ldots,$ and

$$h_n = \Im_s(A = 1, c_k) = 2^{n-1} \sum_{j=0}^{\infty} (-1)^j \frac{n+2j}{n+j} \binom{n+j}{j} c_k, \tag{3.6}$$

for $k = n + 2j$ and $n = 0, 1, 2, \ldots,$ where

$$\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha-\beta)!} \quad \text{for } 0 \le \beta \le \alpha, \tag{3.7}$$

are binomial coefficients. To overcome the unit input amplitude constraint, Suen used Pascal's triangle and some trigonometric identities (see (9) and (10) of [Sue70] for the original equations); we rearranged the terms of his equations, to extract the pattern in the next section. The resultant analysis and synthesis equations are given as:

$$c_k = \Im_a(A, h_n) = \frac{1}{2^{k-1}} \sum_{j=0}^{\infty} A^{k+2j} \frac{h_n}{2^{2j}} \binom{k+2j}{j}, \tag{3.8}$$

for $n = k + 2j$ and $k = 0, 1, 2, \ldots,$ and

$$h_n = \Im_s(A, c_k) = \frac{2^{n-1}}{A^n} \sum_{j=0}^{\infty} (-1)^j \frac{n+2j}{n+j} \binom{n+j}{j} c_k, \tag{3.9}$$

for $k = n + 2j$ and $n = 0, 1, 2, \ldots.$

Comparing (3.5) and (3.6) with (3.8) and (3.9), we note that the differences are A^{k+2j} inside the summation for analysis, and $1/A^n$ for synthesis. Therefore, (3.8) and (3.9) are solutions of the analysis and synthesis problems in a power series.

3.2.3 Harmonic analysis and synthesis of polynomial nonlinearities

Equations (3.8) and (3.9) are analysis and synthesis equations for the power series in (3.1), but these are not closed-form equations for polynomial nonlinearities in (3.2).

In this section, we derive the closed-form design equations for polynomial nonlinearities.

Let us define a truncated Fourier series as

$$\hat{g}(t) = \frac{\hat{c}_0}{2} + \sum_{k=1}^{P} \hat{c}_k \cos(k\omega t), \tag{3.10}$$

where P is the upper bound of the summation or the highest harmonic number, $\hat{c}_k (k = 1, 2, ..., P)$ are truncated Fourier series coefficients or harmonic amplitudes, and $\hat{c}_0/2$ is the DC component. Suppose that the total number of harmonics at the output of polynomial nonlinearity in (3.2) is finite, and can be represented by the truncated Fourier series in (3.10). As in Section 3.2.2, two problems are defined as follows:

(i) Given (3.3) and (3.2), (3.10) can be expressed as a function of A and $\hat{h}_n$ such that $\hat{c}_k = \hat{\Im}_a(A, \hat{h}_n)$ for $k \in \mathbb{Z}^+$, where $\hat{\Im}_a(.)$ is an analysis function that generates $\hat{c}_k$ from $\hat{h}_n$ and A.

(ii) Given (3.3) and (3.10), (3.2) can be expressed as a function of A and $\hat{c}_k$ such that $\hat{h}_n = \hat{\Im}_s(A, \hat{c}_k)$ for $n \in \mathbb{Z}^+$, where $\hat{\Im}_s(.)$ is a synthesis function that generates $\hat{h}_n$ from $\hat{c}_k$ and A.

Note that $\hat{\Im}_a(.)$ $\hat{\Im}_s(.)$ are harmonic analysis and synthesis equations of polynomial nonlinearities, respectively. Unlike $\Im_a(.)$ and $\Im_s(.)$ in (3.8) and (3.9), which are infinite series, $\hat{\Im}_a(.)$ and $\hat{\Im}_s(.)$ are closed-form equations that we derive, based on some integer number patterns in harmonic index k, power series index n and summation index j, as shown in (3.8) and (3.9).

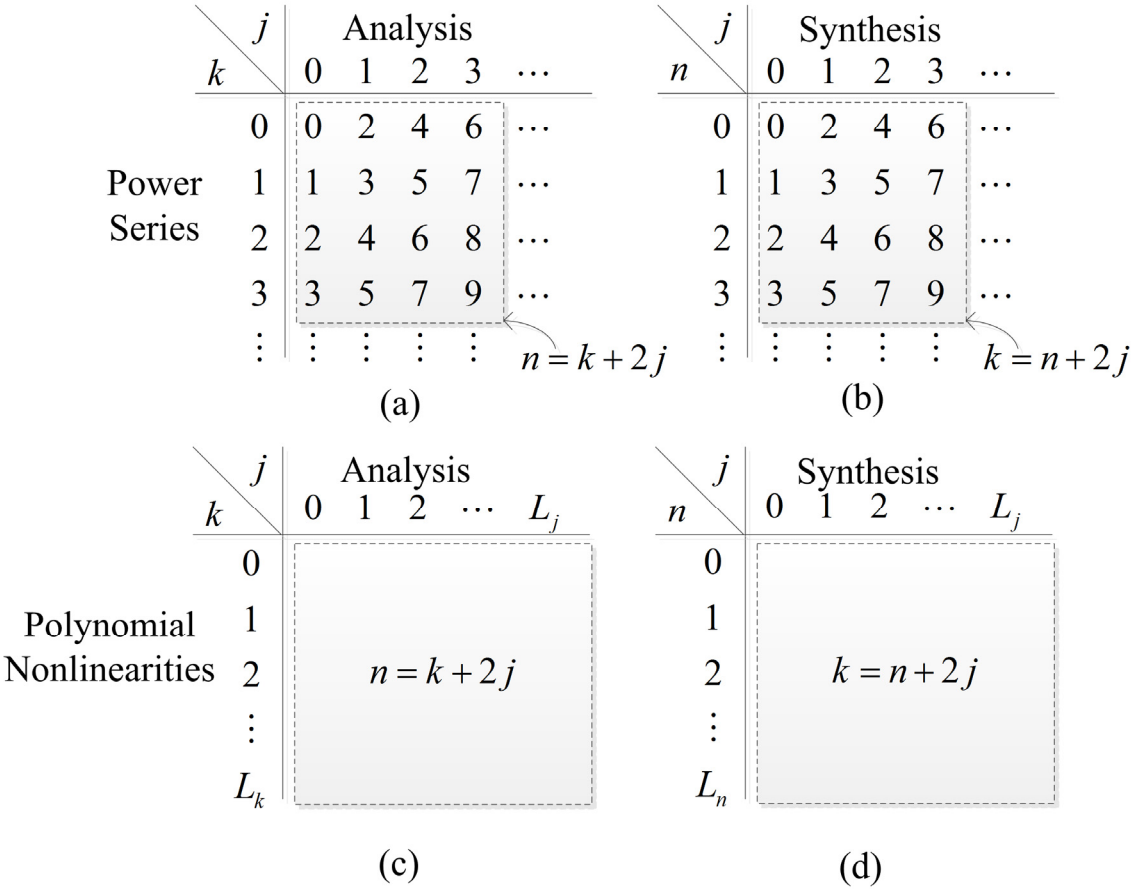

Figure 3.1: The mathematical relationship between indices *k*, *n,* and *j*. (a) Analysis table of power series (b) Synthesis table of power series (c) Analysis table of polynomial nonlinearities (d) Synthesis table of polynomial nonlinearities.

From (3.8) and (3.9), the index number relationships are $n=k+2j$ for analysis and $k=n+2j$ for synthesis. The differences between power series and polynomial nonlinearities are infinite index upper bounds in Figure 3.1 (a) and (b). The finite index upper bounds, such as L_k, L_n, and L_j are shown in Figure 3.1 (c) and (d) represent finite index upper bounds of harmonic number, polynomial coefficient, and summation, respectively. Therefore, by obtaining such upper bounds, we provide a lemma and its proof as solutions for the above defined problems.

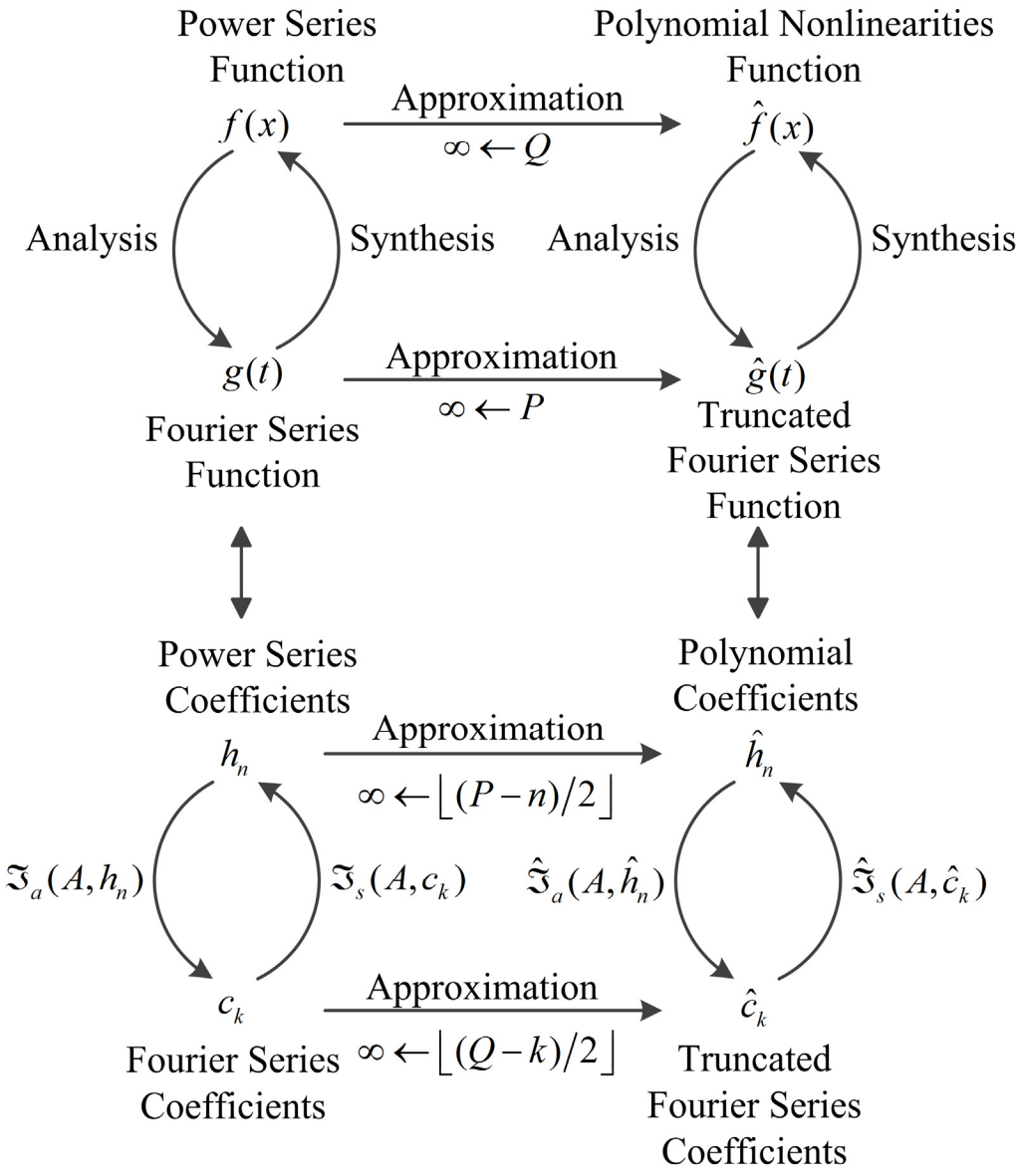

Figure 3.2: Analysis and synthesis link between power series, polynomial Nonlinearities, Fourier series, and truncated Fourier series.

Lemma 1:

For analysis, $L_k = \max(k) = Q \in \mathbb{Z}^+$ and $L_j = \lfloor (L_k - k)/2 \rfloor = \lfloor (Q-k)/2 \rfloor$ for $k \in \{0,1,2,\ldots,Q\}$. $L_n = \max(n) = P \in \mathbb{Z}^+$, $L_j = \lfloor (L_n - n)/2 \rfloor = \lfloor (P-n)/2 \rfloor$ for synthesis. The closed-form design equations, $\hat{c}_k = \hat{\Im}_a(A, \hat{h}_n)$ and $\hat{h}_n = \hat{\Im}_s(A, \hat{c}_k)$, are

$$\hat{c}_k = \hat{\Im}_a(A, \hat{h}_n) = \frac{1}{2^{k-1}} \sum_{j=0}^{L_j = \lfloor (Q-k)/2 \rfloor} \frac{A^{k+2j} \hat{h}_n}{2^{2j}} \binom{k+2j}{j}, \tag{3.11}$$

for $n = k + 2j$ and $k = 0, 1, 2, \ldots, (L_k = Q)$,

$$\hat{h}_n = \hat{\Im}_s(A, \hat{c}_k) = \frac{2^{n-1}}{A^n} \sum_{j=0}^{L_j = \lfloor (P-n)/2 \rfloor} (-1)^j \frac{n+2j}{n+j} \binom{n+j}{j} \hat{c}_k, \tag{3.12}$$

for $k = n + 2j$ and $n = 0, 1, 2, \ldots, (L_n = P)$.

Proof:

Denote sup(.) as the *supremum* or the least upper bound and inf(.) as the *infimum* or the greatest lower bound. Suppose $\hat{h}_n$, for $n = k + 2j$ exists in (3.11) and $\hat{c}_k$, for $k = n + 2j$ exists in (3.12). For (3.11), L_k denotes the highest harmonic number. $L_k = \sup\{k + \inf(2j)\} \le Q$.

Because $j \in \mathbb{Z}^+$ and $\inf(2j) = 2j\big|_{j=0} = 0$, $L_k = \sup(k) \le Q$. Also, $k, Q \in \mathbb{Z}^+, L_k = Q$. For (3.11), L_j denotes the index upper bound. $L_j = \sup(j)$ and $k + 2\sup(j) \le Q$ where $k, j \in \mathbb{Z}^+$. $L_j \le (Q-k)/2$ implies $L_j = \lfloor (Q-k)/2 \rfloor$, where $\lfloor x \rfloor$ denotes the largest integer not greater than x. For (3.12), $L_n = \sup\{n + \inf(2j)\} \le P$ denotes the highest degree of the polynomial. Because

$j \in \mathbb{Z}^+$ and $\inf(2j) = 2j\big|_{j=0} = 0$, $L_n = \sup(n) \le P$. Because $n, P \in \mathbb{Z}^+$, $L_n = P$. For (3.12), L_j denotes the index upper bound. $L_j = \sup(j)$ and $n + 2\sup(j) \le P$ where $n, j \in \mathbb{Z}^+$. Thus, $L_j = \lfloor (P-n)/2 \rfloor$. *Q. E. D.*

The analysis and synthesis duality relationships are illustrated in Figure 3.2, where $f(x)$, $\hat{f}(x)$, $g(t)$, and $\hat{g}(t)$ are expressed in (3.1), (3.2), (3.4), and (3.10), and $\Im_a(A, h_n)$, $\Im_s(A, c_k)$, $\hat{\Im}_a(A, \hat{h}_n)$, and $\hat{\Im}_s(A, \hat{c}_k)$ are shown in (3.8), (3.9), (3.11), and (3.12), respectively.

3.2.4 Triple Duality Theorem for sound analysis and synthesis

In this section, the mathematical relationships between nonlinear function symmetries and harmonic generation patterns are discussed.

For $\forall j \in \mathbb{Z}^+$, let $n = k + 2j$, $n \ge 2j$ for analysis equations in (3.8) and (3.11), and $k = n + 2j$, $k \ge 2j$ for synthesis equations in (3.9) and (3.12). Let $\mathrm{E}^+ = \{0, 2, 4, \ldots\}$ be a set of all positive even numbers and $\mathrm{O}^+ = \{1, 3, 5, \ldots\}$ be a set of all positive odd numbers with $\mathrm{E}^+ \bigcup \mathrm{O}^+ = \mathbb{Z}^+$ and $\mathrm{E}^+ \bigcap \mathrm{O}^+ = \varnothing$, where $\bigcup$ and $\bigcap$ are union and intersection set operators, respectively, and $\varnothing$ is the symbol for the empty set.

Lemma 2: The following statements are true:

1) $\forall n \in \mathrm{E}^+ \leftrightarrow \forall k \in \mathrm{E}^+$,
2) $\forall n \in \mathrm{O}^+ \leftrightarrow \forall k \in \mathrm{O}^+$,
3) $\forall n \in \mathbb{Z}^+ \leftrightarrow \forall k \in \mathbb{Z}^+$,

where the symbols $\leftrightarrow$, and $\forall$ denote "if and only if (iff)," and "for all," respectively.

Proof:

Let $A \in \{n \in \mathrm{E}^+ \rightarrow true, n \notin \mathrm{E}^+ \rightarrow false\}$, and $B \in \{k \in \mathrm{E}^+ \rightarrow true, k \notin \mathrm{E}^+ \rightarrow false\}$ be two conditions that can either be true or false, and $n, k \in \mathbb{Z}^+$. Because

$$\left(\mathrm{E}^+ \bigcup \mathrm{O}^+ = \mathbb{Z}^+\right) \wedge \left(\mathrm{E}^+ \bigcap \mathrm{O}^+ = \varnothing\right), \tag{3.13}$$

where $\wedge$ denotes the logical conjunction (AND) operator. Set $\neg(A) = \neg(n \in \mathrm{E}^+) = n \notin \mathrm{E}^+ = n \in \mathrm{O}^+$, and $\neg(B) = \neg(k \in \mathrm{E}^+) = k \notin \mathrm{E}^+ = k \in \mathrm{O}^+$, where $\neg$ denotes negation. Let $\rightarrow$ denotes implication. To prove $A \rightarrow B$, assume that $A \rightarrow \neg(B)$. By (3.11), $k = n - 2j$, $n \geq 2j$. Setting $n = 2 \in \mathrm{E}^+$, and $j = 0 \in \mathbb{Z}^+$, $k = 2 - 2(0) = 2 \in \mathrm{E}^+$, which contradicts the assumption. Because $A \not\rightarrow \neg(B)$, $A \rightarrow B$.

To prove $B \rightarrow A$, assume that $B \rightarrow \neg(A)$. By (3.12), $n = k - 2j$, $k \geq 2j$. Setting $k = 2 \in \mathrm{E}^+$, and $j = 0 \in \mathbb{Z}^+$, $n = 2 - 2(0) = 2 \in \mathrm{E}^+$, which contradicts the assumption. Because $B \not\rightarrow \neg(A), B \rightarrow A$. Because $A \rightarrow B$ and $B \rightarrow A$, $A \leftrightarrow B$. Furthermore, due to (3.13), $n \in \mathrm{E}^+ \notin \mathrm{O}^+$ and $k \in \mathrm{E}^+ \notin \mathrm{O}^+$, which implies that $\forall n \in \mathrm{E}^+ \leftrightarrow \forall k \in \mathrm{E}^+$. Suppose $C \in \{n \in \mathrm{O}^+ \rightarrow true, n \notin \mathrm{O}^+ \rightarrow false\}$, and $D \in \{k \in \mathrm{O}^+ \rightarrow true, k \notin \mathrm{O}^+ \rightarrow false\}$ be two conditions that can either be true or false, and $n, k \in \mathbb{Z}^+$. By (3.13), $\neg(C) = \neg(n \in \mathrm{O}^+) = n \notin \mathrm{O}^+ = n \in \mathrm{E}^+$, and $\neg(D) = \neg(k \in \mathrm{O}^+) = k \notin \mathrm{O}^+ = k \in \mathrm{E}^+$.

To prove $C \to D$, assume that $C \to \neg(D)$. Set $n = 1 \in \mathrm{O}^+$ and $j = 0 \in \mathbb{Z}^+$. Because $k = n - 2j$ for $n \geq 2j$ from (3.11), $k = 1 - 2(0) = 1 \in \mathrm{O}^+$ that contradicts the assumption. Because $C \not\to \neg(D)$, $C \to D$. To prove $D \to C$, assume that $D \to \neg(C)$. Set $k = 1 \in \mathrm{O}^+$ and $j = 0 \in \mathbb{Z}^+$. Because $n = k - 2j$ for $k \geq 2j$ from (3.12), $n = 1 - 2(0) = 1 \in \mathrm{O}^+$ that contradicts the assumption. Because $D \not\to C$, $D \to C$. Because $C \to D$ and $D \to C$, $C \leftrightarrow D$. Furthermore, due to (3.13), $n \in \mathrm{O}^+ \notin \mathrm{E}^+$ and $k \in \mathrm{O}^+ \notin \mathrm{E}^+$, which implies that $\forall n \in \mathrm{O}^+ \leftrightarrow \forall k \in \mathrm{O}^+$.

To prove $\forall n \in \mathbb{Z}^+ \leftrightarrow \forall k \in \mathbb{Z}^+$, let $\{n_1, n_2\} \subset \mathbb{Z}^+$, $n_1 \in \mathrm{E}^+$, and $n_2 \in \mathrm{O}^+$, whereas $\{k_1, k_2\} \subset \mathbb{Z}^+, k_1 \in \mathrm{E}^+$ and $k_2 \in \mathrm{O}^+$. Because $n_1 \in \mathrm{E}^+ \leftrightarrow k_1 \in \mathrm{E}^+$ and $n_2 \in \mathrm{O}^+ \leftrightarrow k_2 \in \mathrm{O}^+$, $n_1, n_2 \in \mathrm{E}^+ \bigcup \mathrm{O}^+ = \mathbb{Z}^+ \leftrightarrow k_1, k_2 \in \mathrm{E}^+ \bigcup \mathrm{O}^+ = \mathbb{Z}^+$.

Because $\mathrm{E}^+ \bigcap \mathrm{O}^+ = \varnothing$, $\forall n \in \mathbb{Z}^+ \leftrightarrow \forall k \in \mathbb{Z}^+$. *Q. E. D.*

The mathematical results of Lemma 2 can be interpreted as follows: For analysis, if system functions $f(x)$ or $\hat{f}(x)$ in (3.1) and (3.2) contain indices n that are all even (or all odd) numbers, the signal functions $g(t)$ or $\hat{g}(t)$ in (3.4) and (3.10) will contain indices k that are all even (or odd) numbers, respectively. In other words, power series or polynomials with even (or odd) orders can generate only even (or odd) harmonics. Note that the DC component is a constant term $c_0/2$ in (3.4) or $\hat{c}_0/2$ in (3.10) with $k = 0$ which is an even number. For the synthesis, if the signal functions contain indices k that are all even (or all odd) numbers, the system functions contain indices n that are all even (or all odd) numbers, respectively. The last statement of Lemma 2 means that for analysis, if the system functions contain

both even and odd numbers, the generated harmonics contain both even and odd numbers and vice versa for synthesis.

Following this line of thought, we link the analysis and synthesis dualities with function symmetries. If $f(x)$ is even symmetric, $f(x) = f(-x)$ and if $f(x)$ is odd symmetric, $f(x) = -f(-x)$ If $f(x)$ is non-symmetric, $f(x) \neq f(-x) \neq -f(-x)$. Lemma 3 and its proof show the links between function symmetries and order n of $f(x)$ or $\hat{f}(x)$. Because convergence condition $\lim_{n\to\infty} \hat{f}(x) = f(x)$ is assumed, the general statement of Lemma 3 and its proof is based on $f(x)$ but the result is also applicable to $\hat{f}(x)$.

Lemma 3: The following statements are true.

1) $f(x) = f(-x) \leftrightarrow \forall n \in \mathrm{E}^+$,

2) $f(x) = -f(-x) \leftrightarrow \forall n \in \mathrm{O}^+$,

3) $f(x) \neq f(-x) \neq -f(-x) \leftrightarrow \forall n \in \mathbb{Z}^+$.

Proof: Let $f(x) = \sum_{n\in\mathbb{Z}^+} h_n x^n$.

$$f(-x) = \sum_n h_n (-x)^n = \sum_n h_n (x)^n = f(x) \leftrightarrow \forall n \in \mathrm{E}^+.$$

$$f(-x) = \sum_n h_n (-x)^n = -\sum_n h_n (x)^n = -f(x) \leftrightarrow \forall n \in \mathrm{O}^+.$$

$$f(-x) = \sum_n h_n (-x)^n \neq \sum_n h_n (x)^n = f(x) \neq -\sum_n h_n (x)^n = -f(x)$$

$$\leftrightarrow \neg\left(\forall n \in \mathrm{E}^+\right) \wedge \neg\left(\forall n \in \mathrm{O}^+\right) = \left(\exists n \in \mathrm{E}^+\right) \wedge \left(\exists n \in \mathrm{O}^+\right) = \forall n \in \mathbb{Z}^+. \quad \textit{Q.E.D.}$$

The logic symbol $\exists$ denotes "there exists." Based on these lemmas, we state and prove the Triple Duality Theorem, which shows the relationships between function symmetries, order of nonlinearities, and harmonic numbers.

Triple Duality Theorem:

1) $f(x) = f(-x) \leftrightarrow \forall n \in \mathrm{E}^+ \leftrightarrow \forall k \in \mathrm{E}^+,$

2) $f(x) = -f(-x) \leftrightarrow \forall n \in \mathrm{O}^+ \leftrightarrow \forall k \in \mathrm{O}^+,$

3) $f(x) \neq f(-x) \neq -f(-x) \leftrightarrow \forall n \in \mathbb{Z}^+ \leftrightarrow \forall k \in \mathbb{Z}^+.$

Proof: Using Lemma 2 and 3,

$\left(\forall n \in \mathrm{E}^+ \leftrightarrow \forall k \in \mathrm{E}^+\right) \wedge \left(f(x) = f(-x) \leftrightarrow \forall n \in \mathrm{E}^+\right)$

$\rightarrow f(x) = f(-x) \leftrightarrow \forall n \in \mathrm{E}^+ \leftrightarrow \forall k \in \mathrm{E}^+.$

$\left(\forall n \in \mathrm{O}^+ \leftrightarrow \forall k \in \mathrm{O}^+\right) \wedge \left(f(x) = -f(-x) \leftrightarrow \forall n \in \mathrm{O}^+\right)$

$\rightarrow f(x) = -f(-x) \leftrightarrow \forall n \in \mathrm{O}^+ \leftrightarrow \forall k \in \mathrm{O}^+.$

$\left(\forall n \in \mathbb{Z}^+ \leftrightarrow \forall k \in \mathbb{Z}^+\right) \wedge \left(f(x) \neq f(-x) \neq -f(-x) \leftrightarrow \forall n \in \mathbb{Z}^+\right)$

$\rightarrow f(x) \neq f(-x) \neq -f(-x) \leftrightarrow \forall n \in \mathbb{Z}^+ \leftrightarrow \forall n \in \mathbb{Z}^+.$

Q.E.D.

Corollary: Let Λ denote any positive integer including zero. $\exists n = \Lambda \in \mathbb{Z}^+ \leftrightarrow \exists k = \Lambda \in \mathbb{Z}^+.$

Proof: By Lemma 1, $k = n - 2j,$ for $j = 0,1,\ldots, L_j.$ Thus, $\exists n = \Lambda \rightarrow \exists k = \Lambda - 2\min(j) = \Lambda - 2(0) = \Lambda.$ By Lemma 1, $n = k - 2j$ for $j = 0,1,\ldots, L_j.$ Thus, $\exists k = \Lambda \rightarrow \exists n = \Lambda - 2\min(j) = \Lambda - 2(0) = \Lambda.$ Because $\exists n = \Lambda \rightarrow \exists k = \Lambda$ and $\exists k = \Lambda \rightarrow \exists n = \Lambda,$ $\exists n = \Lambda \leftrightarrow \exists k = \Lambda$ for $\Lambda \in \mathbb{Z}^+.$ *Q.E.D.*

The Triple Duality Theorem describes the relationships between function symmetries, harmonic patterns, and order of nonlinearity. The corollary states that if

an integer number is a nonlinearity order, this particular integer number is the harmonic number and vice versa.

3.2.5 Harmonic Analysis and Synthesis on Polynomial (HASP)

Using design equations in (3.11) and (3.12), we propose the HASP algorithm. HASP consists of only two routines *HASP_Analysis* and *HASP_Synthesis*, which are presented in Table 3.1 and Table 3.2, respectively.

Table 3.1: Analysis routine of HASP.

function $C =$ HASP_Analysis (A, H)

Input: Amplitude of input single tone A and an array H of length $Q+1$ containing polynomial coefficients such that $H[1]=\hat{h}_0$, $H[2]=\hat{h}_1$, …, $H[Q+1]=\hat{h}_Q$.

Output: An array of C length $Q+1$ containing amplitudes of harmonics such that $C[1]=\hat{c}_0$, $C[2]=\hat{c}_1$, …, $C[Q+1]=\hat{c}_Q$.

for $k \leftarrow 0$ **to** Q **step** 1 **do**

  $sum \leftarrow 0$;

  for $j \leftarrow 0$ **to** $floor((Q-k)/2)$ **step** 1 **do**

    $$sum \leftarrow sum + A^{k+2\times j} \times \frac{H[k+2\times j-1]}{2^{2\times j}} \times \binom{k+2\times j}{j};$$

  od;

  $$C[k+1] \leftarrow \frac{1}{2^{k-1}} \times sum;$$

od;

Table 3.2: Synthesis routine of HASP.

function $H =$ HASP_Synthesis(A , C)

Input: Amplitude of input single tone A and an array C of length $P+1$ containing truncated Fourier series coefficients such that $C[1] = \hat{c}_0$, $C[2] = \hat{c}_1$, … , $C[P+1] = \hat{c}_P$.

Output: An array H of length $P+1$ containing polynomial coefficients such that $H[1] = \hat{h}_0$, $H[2] = \hat{h}_1$, … , $H[P+1] = \hat{h}_P$.

$sum \leftarrow 0$;

for $j \leftarrow 1$ **to** $floor(P/2)$ **step** 1 **do**

$sum \leftarrow sum + (-1)^j \times C[2 \times j + 1]$;

od;

comment : computations of $\hat{h}_n$, $n = 1, \ldots, P$;

for $n \leftarrow 1$ **to** P **step** 1 **do**

$sum \leftarrow 0$;

for $j \leftarrow 0$ **to** $floor((P-n)/2)$ **step** 1 **do**

$$sum \leftarrow sum + (-1)^j \times \frac{(n+2\times j)}{(n+j)} \times \binom{n+j}{j} \times C[n+2\times j+1];$$

od;

$$H[n+1] \leftarrow \frac{2^{n-1}}{A^n} \times sum;$$

od;

3.2.6 Case studies using HASP

$f(x)$ and $\hat{f}(x)$ are the power series and polynomial nonlinearity, defined in (3.1) and (3.2), respectively. They can be related as follows:

$$f(x) = \sum_{n=0}^{\infty} h_n x^n \approx \hat{f}(x) = \sum_{n=0}^{Q} \hat{h}_n x^n, \tag{3.14}$$

with the convergence assumption

$$\lim_{Q \to \infty} \hat{f}(x) = f(x), \tag{3.15}$$

where $h_0, h_1, \ldots$, and $\hat{h}_0, \hat{h}_1, \ldots$ are series and polynomial coefficients, respectively. Q is the highest order of polynomial. Given the values of $f(x)$ at $M+1$ distinct values of x, and for $m = 0,1,\ldots,M$, $\hat{f}(x_m) = f(x_m)$. That is

$$\sum_{n=0}^{Q} \hat{h}_n x_m^2 = \hat{h}_0 + \hat{h}_1 x_m + \hat{h}_2 x_m^2 + \cdots + \hat{h}_Q x_m^Q = f(x_m), \tag{3.16}$$

for $0 \le m \le Q \le M$. Equation (3.16) can also be written as

$$\underbrace{\begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^Q \\ 1 & x_1 & x_1^2 & \cdots & x_1^Q \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_M & x_M^2 & \cdots & x_M^Q \end{bmatrix}}_{\mathbf{V} \in \mathbb{R}^{(M+1)\times(Q+1)}} \underbrace{\begin{bmatrix} \hat{h}_0 \\ \hat{h}_1 \\ \vdots \\ \hat{h}_M \end{bmatrix}}_{\mathbf{h} \in \mathbb{R}^{(M+1)\times 1}} = \underbrace{\begin{bmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_M) \end{bmatrix}}_{\mathbf{f} \in \mathbb{R}^{(M+1)\times 1}}, \tag{3.17}$$

where $\mathbf{V}$ is the *Vandermonde* matrix. Because $\mathbf{V}$ is singular [Phi03], the approximating polynomial vector $\mathbf{h}$ can be obtained as

$$\hat{\mathbf{h}} = \mathbf{V}^{-1}\mathbf{f}. \tag{3.18}$$

By (3.18), memoryless nonlinear functions can be polynomial-approximated and *HASP_Analysis* can be used to compute the individual harmonic contents at the

output of polynomial nonlinearities. On the other hand, *HASP_Synthesis* can be applied to construct the polynomial nonlinearities instantly when the information on the harmonic content is pre-defined.

Table 3.3 shows numerical results as case studies or examples. The first column of the table are acronyms of four nonlinear functions that are odd-symmetric. They are normalized hyperbolic tangent (NTANH), inverted NTANH, negated NTANH, and inverted and negated NTANH functions. The functions are plotted in Figure 3.3. The second column shows the closed-form formulae of the functions. The third and fourth columns consist of approximated polynomial coefficients and truncated Fourier coefficients for the case of unity input amplitude. This example shows that given the third column, using *HASP_Analysis*, we can obtain the fourth column, or given the fourth column, using *HASP_synthesis,* we can obtain the third column data. Note that triple duality theorem predicts that odd symmetric nonlinearities can generate only odd harmonics, and maximum harmonic number is equivalent to the largest polynomial order. In these examples, the maximum harmonic number is 3, and the largest polynomial order is third order.

Table 3.3: Harmonic analysis and synthesis results of a family of normalized hyperbolic tangent functions. Note: $\alpha = (e^2+1)/(e^2-1)$.

Acronym	$f(x)$	$\hat{h}_0$	$\hat{h}_1$	$\hat{h}_2$	$\hat{h}_3$	$\hat{c}_0$	$\hat{c}_1$	$\hat{c}_2$	$\hat{c}_3$
NTANH	$\alpha\frac{(e^x-e^{-x})}{(e^x+e^{-x})}$	0	1.2863	0	-0.2977	0	1.0630	0	-0.0744
INTANH	$\frac{1}{2}\ln\left(\frac{1+x/\alpha}{1-x/\alpha}\right)$	0	0.7372	0	0.2438	0	0.9200	0	0.0609
N2TANH	$\alpha\frac{(e^{-x}-e^{x})}{(e^{-x}+e^{x})}$	0	-1.2863	0	0.2977	0	-1.0630	0	0.0744
IN2TANH	$\frac{1}{2}\ln\left(\frac{1-x/\alpha}{1+x/\alpha}\right)$	0	-0.7372	0	-0.2438	0	-0.9200	0	-0.0609

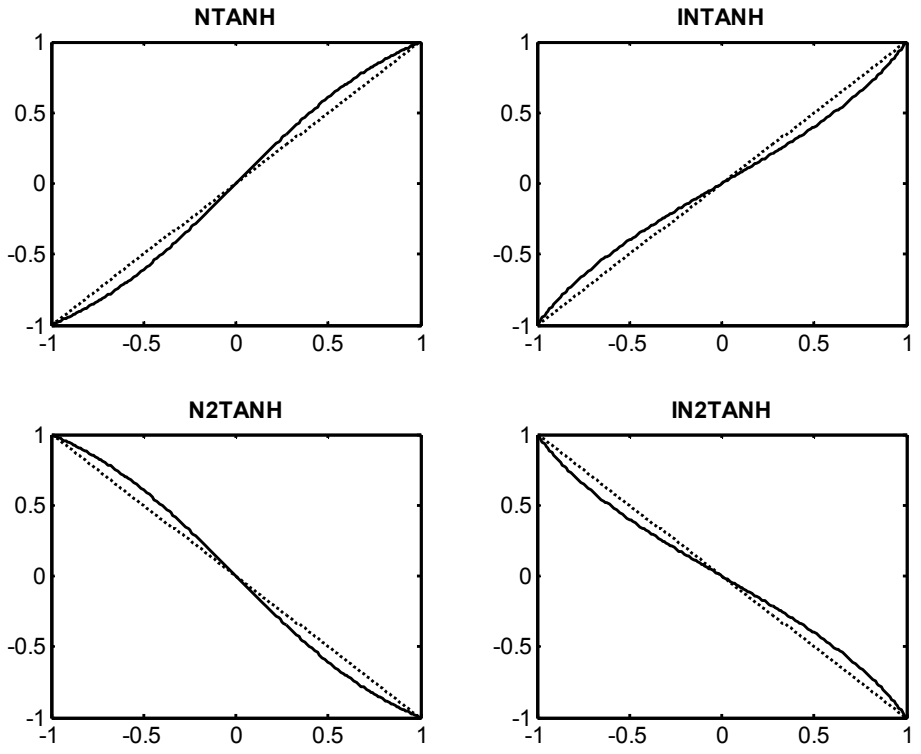

Figure 3.3: In-out plots of four normalized hyperbolic tangent functions. The dotted diagonal lines in upper left and right subplots represent $y = x$ and the dotted diagonal lines in lower left and right subplots are $y = -x$. The four solid curves represent the four nonlinearities presented in Table 3.3, respectively.

3.3 Generalized Harmonic Analysis

Generalized harmonic analysis is a term used here to derive generalized closed-form harmonic analysis formulae from the given NLD nonlinear functions with the help of Taylor's series. The objective is to study the effect of the parameterizations of the nonlinear functions on the harmonic generation patterns and to relate the findings with the single-tone objective metric, THD.

3.3.1 Generalized harmonic analysis of ATSR NLD

In [OGL10], we proposed an NLD that can generate both even and odd harmonics that are essential for bass perception. The proposed NLD is termed arc-tangent square root (ATSR) and is defined as

$$f_{\text{ATSR}}(x) \equiv \alpha \tan^{-1}(\beta x) + \psi\sqrt{1-(\zeta x)^2} - \psi, \tag{3.19}$$

where $\{\alpha, \beta, \psi, \zeta\}$ are the four parameters that control the harmonics generation pattern and intensity. The mathematical relation between these four parameters and harmonic generation will be considered next.

By using Taylor's series [CC04], (3.19) can be expressed as follows:

$$f_{\text{ATSR}}(x) = \underbrace{\alpha\beta\sum_{n=0}^{\infty}\left(\frac{(-1)^n}{2n+1}\beta^{2n}\right)x^{2n+1}}_{f_1(x)} + \underbrace{\psi\sum_{n=0}^{\infty}\frac{(-\zeta)^{2n}(2n)!}{(1-2n)n!^2\,4^n}x^{2n}}_{f_2(x)} - \psi. \tag{3.20}$$

If we limit the highest order of polynomials to Q, with the highest even order as Q_e, and the highest odd order as Q_o, we can state this relationship as

$$\{Q_e, Q_o\} = \begin{cases} Q_e = Q, \quad Q_o = Q_e - 1, & \text{if } \text{mod}(Q,2) = 0, \\ Q_o = Q, \quad Q_e = Q_o - 1, & \text{if } \text{mod}(Q,2) = 1, \end{cases} \tag{3.21}$$

where $\text{mod}(Q,2)$ is the modulus operator finding the remainder of division of Q by 2. An estimated polynomial approximation of the ATSR is written as

$$\hat{f}(x) = \underbrace{\alpha\beta\sum_{n=0}^{\frac{Q_o-1}{2}}\left(\frac{(-1)^n}{2n+1}\beta^{2n}\right)x^{2n+1}}_{\hat{f}_1(x)} + \underbrace{\psi\sum_{n=0}^{\frac{Q_e}{2}}\frac{(-\zeta)^{2n}(2n)!}{(1-2n)n!^2\,4^n}x^{2n}}_{\hat{f}_2(x)} - \psi. \tag{3.22}$$

The input-output transfer characteristic of (3.22) is shown in Figure 3.4. Notice that $\hat{f}_1(x)$ and $\hat{f}_2(x)$ are the odd-symmetric and even-symmetric functions that

generate odd and even harmonics, respectively. These odd and even harmonic blocks are combined in parallel, and a constant ψ is subtracted from their combined result. The parameters α and β are used to control the intensity and harmonic patterns of the odd harmonics, while ψ and ζ are used to control the even harmonics.

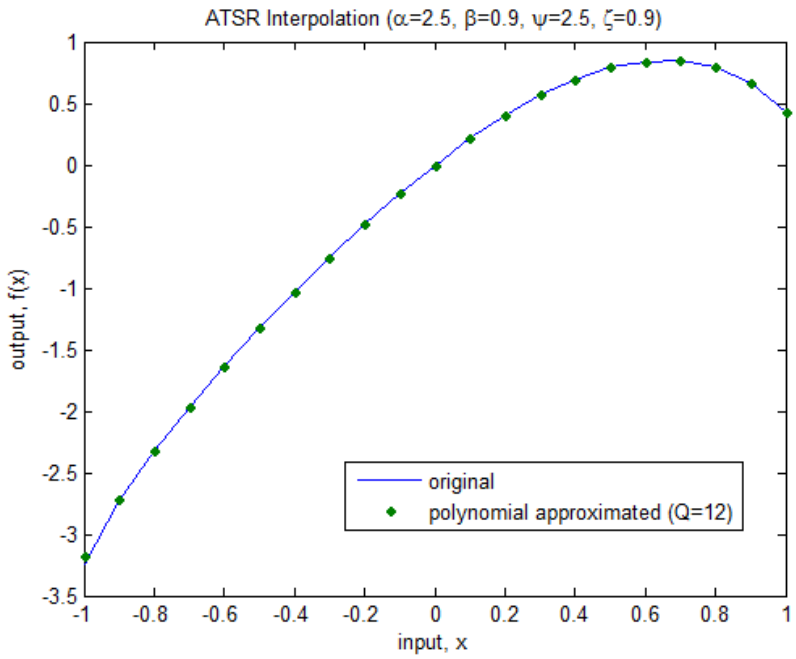

Figure 3.4: The input-output plot of the original ATSR NLD and polynomial approximated ATSR NLD. The parameters are set as $\alpha = 2.5$, $\beta = 0.9$, $\psi = 2.5$, and $\zeta = 0.9$. The polynomial approximation order is $Q = 12$, $Q_e = 12$, and $Q_o = 11$.

The output signal of the ATSR, excited by the sinusoidal signal in (3.3), can be formulated by [OGL10]

$$\hat{y}(t) = \frac{\hat{c}_0}{2} + \sum_{k=1}^{Q} \hat{c}_k \cos(k\omega t), k \in \{1,2,3,\ldots,Q\}$$

$$= \underbrace{\sum_{l=1}^{Q_o} \hat{d}_l \cos(l\omega t)}_{\hat{y}_1(t)} + \underbrace{\frac{\hat{e}_0}{2} + \sum_{m=2}^{Q_e} \hat{e}_m \cos(m\omega t)}_{\hat{y}_2(t)} - \psi, \qquad (3.23)$$

$$\text{for } l \in \{1,3,\ldots,Q_0\}, m \in \{2,4,\ldots,Q_e\}.$$

The generalized closed-form formulae for both even and odd harmonics' amplitudes can be formulated as [OGL10]

$$\hat{d}_l = \frac{\alpha\beta}{2^{l-1}} \sum_{j=0}^{\frac{Q_o - l}{2}} \frac{A^{l+2j}(-1)^{\frac{l+2j-1}{2}} \beta^{l+2j-1}}{2^{2j}(l+2j)} \binom{l+2j}{j}, \qquad \text{for } l \in \{1,3,\ldots,Q_o\}, \qquad (3.24)$$

and

$$\hat{e}_m = \frac{\psi}{2^{m-1}} \sum_{j=0}^{\frac{Q_e - m}{2}} \frac{(-A\zeta)^{m+2j}(m+2j)!}{(1-m-2j)\left(\frac{m+2j}{2}\right)!^2\, 2^{m+4j}} \binom{m+2j}{j}, \qquad \text{for } m \in \{0,2,\ldots,Q_e\}. \quad (3.25)$$

Equations (3.24) and (3.25) generalize the single-tone harmonic analysis of the ATSR function. It is noted that in (3.24), odd harmonics' amplitudes, $\hat{d}_l$ are influenced by the two parameters of ATSR function, α, β, and the input single tone amplitude, A. Likewise in (3.25), even harmonics' amplitudes, $\hat{e}_m$ are influenced by ψ, ζ, and A.

3.3.2 Generalized harmonic analysis of parameterized EXP NLDs

The exponential NLDs are used in psychoacoustic bass enhancement system (also called a virtual bass system) [Kar02] [OG10] [PRL10]. In this subsection, we analyze the harmonic generation pattern of exponential NLDs, and show that THD may not be a preferred metric to measure the nonlinear distortion caused by NLDs.

Let us first define two basis exponential functions as

$$f_+(x,e) = e^x, \quad f_-(x,e) = e^{-x}, \qquad (3.26)$$

where x is single tone input signal defined in (3.3). We then derive some closed-form formulae for parameterized exponential NLDs from (3.26). Let us define exponential

1 (EXP1) and exponential 2 (EXP2). First, a tunable exponential parameter b is substituted in the place of e in (3.26):

$$f_{EXP1}(x,b) = b^x, \tag{3.27}$$

$$f_{EXP2}(x,b) = b^{-x}. \tag{3.28}$$

By Taylor's series, from (3.27),

$$f_{EXP1}(x,b) = b^x = e^{x\ln b} \approx \hat{f}_{EXP1}(x,b) = \sum_{k=0}^{Q} \frac{(\ln b)^k}{k!} x^k, \tag{3.29}$$

and from (3.28),

$$f_{EXP2}(x,b) = b^{-x} = e^{-x\ln b} \approx \hat{f}_{EXP2}(x,b) = \sum_{k=0}^{Q} \frac{(-\ln b)^k}{k!} x^k. \tag{3.30}$$

The transfer characteristics of (3.29) and (3.30) are plotted in Figure 3.5 for b = 1.1 to 10 with a step-size of 1.1. The closed-form harmonic analysis equations for (3.29) and (3.30) can be obtained as follows:

$$\hat{c}_k\big|_{EXP1} = 2\left(\frac{A\ln b}{2}\right)^k \sum_{j=0}^{\left\lfloor \frac{Q-k}{2} \right\rfloor} \frac{\left(A\ln b/2\right)^{2j}}{j!\Gamma(k+j+1)}, \tag{3.31}$$

and

$$\hat{c}_k\big|_{EXP2} = 2\left(\frac{-A\ln b}{2}\right)^k \sum_{j=0}^{\left\lfloor \frac{Q-k}{2} \right\rfloor} \frac{\left(-A\ln b/2\right)^{2j}}{j!\Gamma(k+j+1)}, \tag{3.32}$$

where $\Gamma(.)$ is a gamma function [Sch71] [Leb72], A is the amplitude of input tone in (3.3), b is the base parameter of the exponential NLD, and $\hat{c}_k$ are the truncated Fourier series coefficients that can be written as

$$\hat{f}\left[A\cos(\omega t)\right] = \frac{\hat{c}_0}{2} + \sum_{k=1}^{Q} \hat{c}_k \cos(k\omega t), \tag{3.33}$$

where Q is the highest order of polynomial NLD and the largest harmonic number.

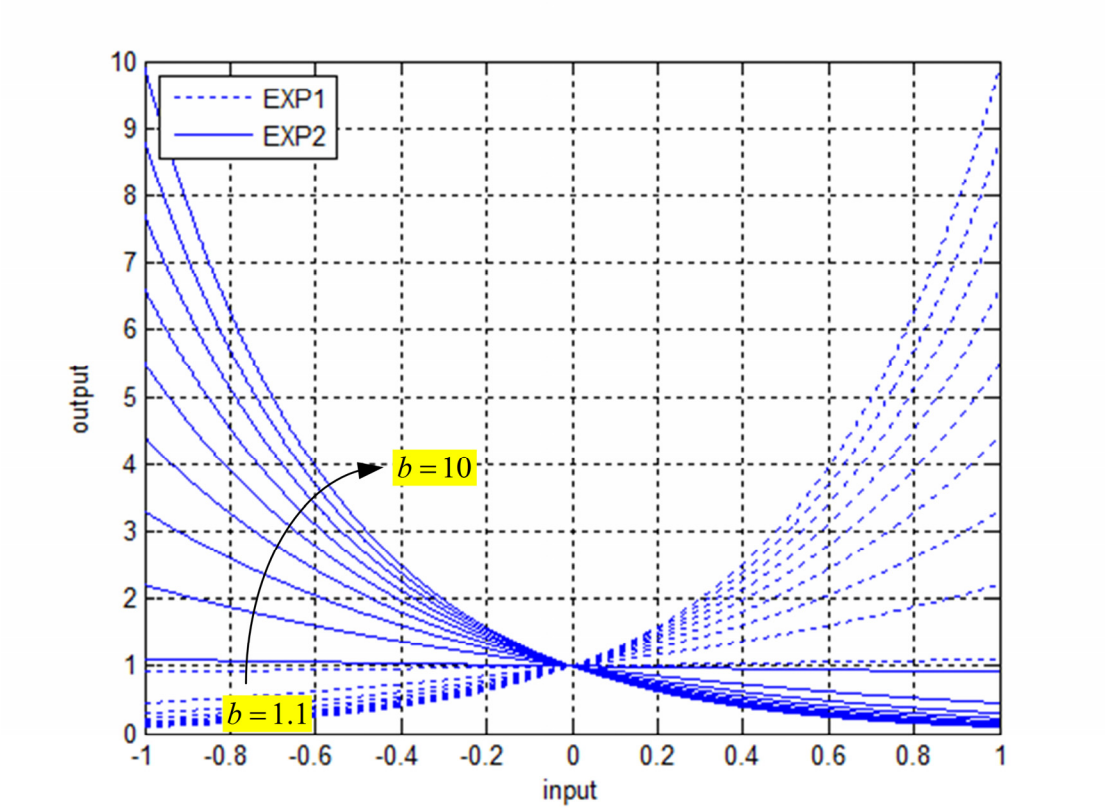

Figure 3.5: Nonlinear transfer characteristics of EXP1 and EXP2. The parameter *b* varies from 1.1 to 10 with a step-size of 1.1.

To derive half-range exponential 1 (HREXP1), half-range exponential 2 (HREXP2), full-range exponential 1 (FREXP1), and full-range exponential 2 (FREXP2), we borrow the idea from data mining applications [HK06]: the essence of the idea is to transform the output interval from $[y_{\min}, y_{\max}]$ to $[\tilde{y}_{\min}, \tilde{y}_{\max}]$ for the input interval $[x_{\min}, x_{\max}]$. The followings are set:

$$
\begin{aligned}
& y = f(x); && \tilde{y} = \tilde{f}(x); \\
& y_{\max} = \max[f(x)]; && \tilde{y}_{\max} = \max[\tilde{f}(x)]; \\
& y_{\min} = \min[f(x)]; && \tilde{y}_{\min} = \min[\tilde{f}(x)],
\end{aligned} \tag{3.34}
$$

where $\tilde{f}(x)$ denotes the normalized nonlinear function. By (3.34), the min-max normalization formula can be written as

$$
\tilde{y} = \frac{y - y_{\min}}{y_{\max} - y_{\min}} (\tilde{y}_{\max} - \tilde{y}_{\min}) + \tilde{y}_{\min}, \tag{3.35}
$$

where y is the original data value to be normalized. $y_{\max}$ and $y_{\min}$ are maximum and minimum of the original data values, whereas $\tilde{y}$ is the normalized data value. $\tilde{y}_{\max}$ and $\tilde{y}_{\min}$ are maximum and minimum of the normalized data values, respectively.

Table 3.4 shows the numerical values settings to be substituted into (3.35) for full-range and half-range EXP1 and EXP2 nonlinearities and the resultant equations.

Table 3.4: Numerical values settings.

NLD	$\tilde{y}_{\max}$	$\tilde{y}_{\min}$	$y_{\max}$	$y_{\min}$	y	$\tilde{y}$
HREXP1	+1	0	b	b^{-1}	b^x	$\frac{b^x - b^{-1}}{b - b^{-1}}$
HREXP2	+1	0	b^{-1}	b	b^{-x}	$\frac{b^2 - b^{1-x}}{b^2 - 1}$
FREXP1	+1	-1	b	b^{-1}	b^x	$2\left(\frac{b^x - b^{-1}}{b - b^{-1}}\right) - 1$
FREXP2	+1	-1	b^{-1}	b	b^{-x}	$2\left(\frac{b^2 - b^{1-x}}{b^2 - 1}\right) - 1$

We define

$$
\Phi_1(b) = \frac{b}{b^2 - 1}; \quad \Phi_2(b) = \frac{-1}{b^2 - 1}, \tag{3.36}
$$

and use (3.36) and Table 3.4 to present the relationships between the half-range and full-range and half-range exponential NLDs and (3.27) and (3.28) as follows:

$$f_{HREXP1}(x,b) = \Phi_1(b) f_{EXP1}(x,b) + \Phi_2(b), \tag{3.37}$$

$$f_{HREXP2}(x,b) = b\left[\Phi_2(b) f_{EXP2}(x,b) + \Phi_1(b)\right], \tag{3.38}$$

$$f_{FREXP1}(x,b) = 2 f_{HREXP1}(x;b) - 1, \tag{3.39}$$

and

$$f_{FREXP2}(x,b) = 2 f_{HREXP2}(x,b) - 1. \tag{3.40}$$

The nonlinear transfer characteristics of (3.37), (3.38), (3.39), and (3.40) are plotted in Figure 3.6 for $b = 1.1$ to 10 with a step-size of 1.1.

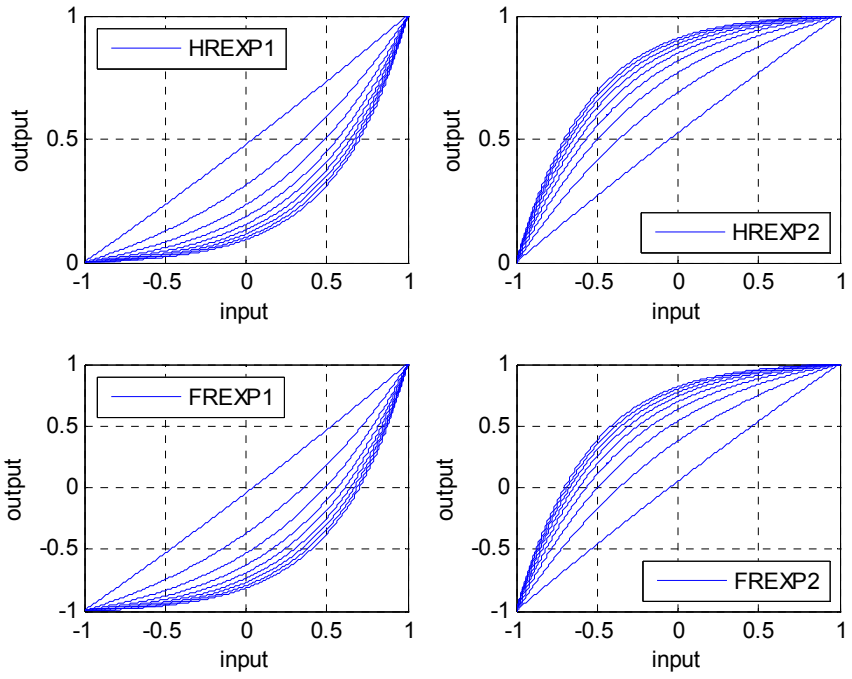

Figure 3.6: Nonlinear transfer characteristics of HREXP1, HREXP2, FREXP1, and FREXP2. The parameter *b* varies from 1.1 to 10 with a step-size of 1.1. The increase in parameter *b* increase nonlinear curvatures.

The following closed-form harmonic analysis equations for the half-range and full-range exponential NLDs are derived:

$$\hat{c}_0\big|_{HREXP1} = 2\left[\Phi_1(b)\frac{\hat{c}_0\big|_{EXP1}}{2} + \Phi_2(b)\right]; \quad \hat{c}_{k>0}\big|_{HREXP1} = \Phi_1(b)\,\hat{c}_{k>0}\big|_{EXP1}. \tag{3.41}$$

$$\hat{c}_0\big|_{HREXP2} = 2\left\{b\left[\Phi_2(b)\frac{\hat{c}_0\big|_{EXP2}}{2} + \Phi_1(b)\right]\right\}; \quad \hat{c}_{k>0}\big|_{HREXP2} = b\Phi_2(b)\,\hat{c}_{k>0}\big|_{EXP2}. \tag{3.42}$$

$$\hat{c}_0\big|_{FREXP1} = 2\left(\hat{c}_0\big|_{HREXP1} - 1\right); \quad \hat{c}_{k>0}\big|_{FREXP1} = 2\,\hat{c}_{k>0}\big|_{HREXP1}. \tag{3.43}$$

$$\hat{c}_0\big|_{FREXP2} = 2\left(\hat{c}_0\big|_{HREXP2} - 1\right); \quad \hat{c}_{k>0}\big|_{FREXP2} = 2\,\hat{c}_{k>0}\big|_{HREXP2}. \tag{3.44}$$

The THD metric is defined by [Shm05]

$$THD = \frac{\sqrt{\sum_{n=2}^{\infty} H_n^2}}{H_1}, \tag{3.45}$$

where H_n for $n = 2,3,\ldots$ are the amplitudes of harmonics. Note that $H_n = |\hat{c}_n|$ because $\hat{c}_n$ can be either positive or negative, representing the same phase or 180 degrees out of phase with respect to the input sinusoidal signal in (3.3). Next, we compare the THD scores of six derived exponential NLDs for $A = 1$ and $b = \exp(1)$. We observe that, although exponential NLDs have different closed-form expressions, they all turn out to yield the same THD scores, which is $24.34\,\%$ in Table 3.5. The numerical results are listed in Table 3.5 for ease of comparison.

The three-dimensional THD plots for the six exponential NLDs are also shown in Figure 3.7. Two parameters A and b vary from -60 dBFS to 0 dBFS and from 1.1 to 10, respectively. The dBFS stands for decibel relative to digital full scale. As predicted by Table 3.5, all the plots are exactly equal implying that, even though

nonlinear transfer function plots are different, six exponential NLDs result in the same THD scores regardless of input parameter and base parameter.

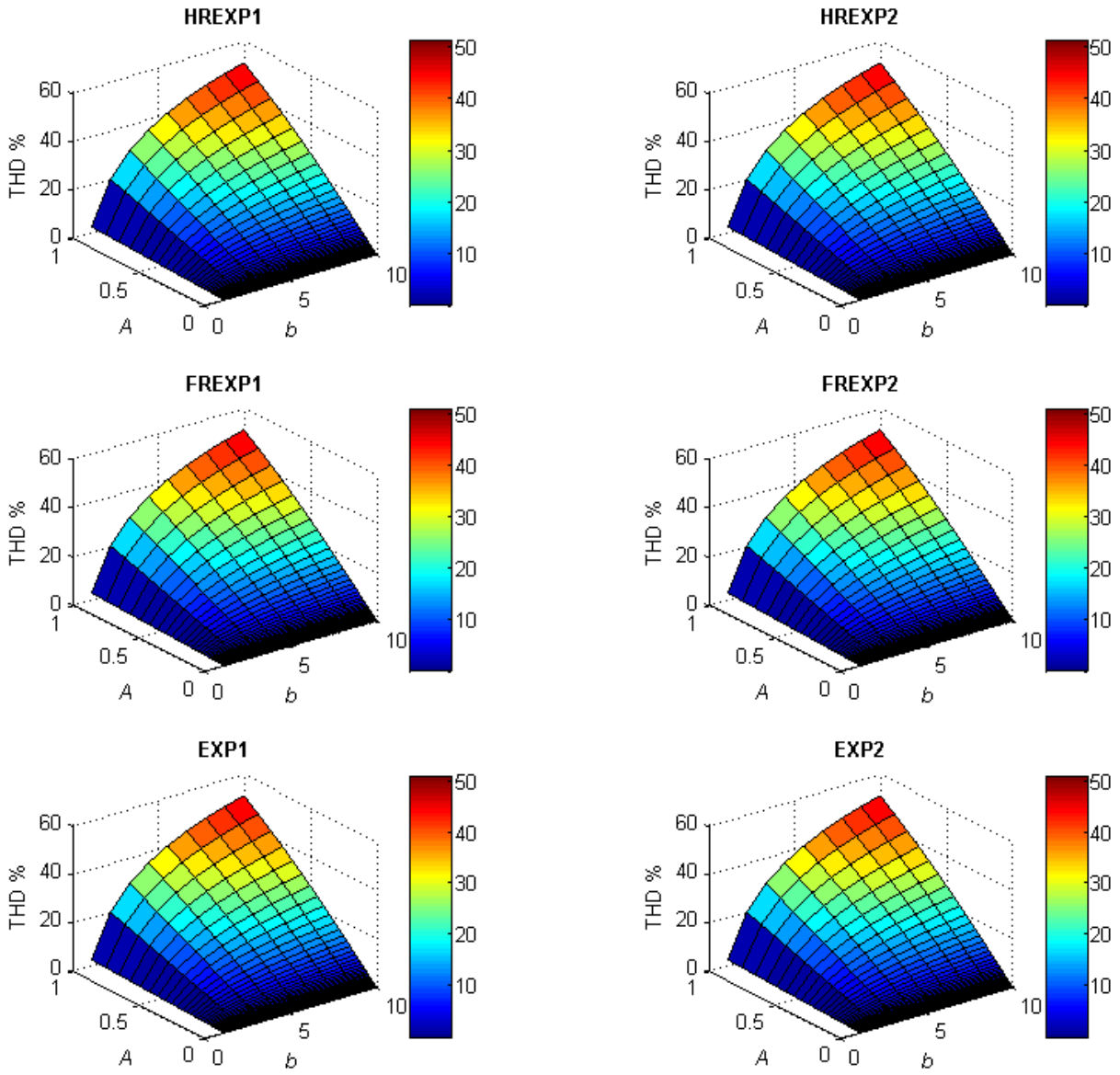

Figure 3.7: Three dimensional THD plots of six exponential NLDs.

Table 3.5: The truncated Fourier series coefficients (up to 5^{th} harmonic) of six exponential NLDs and their respective THD scores $[A = 1, b = \exp(1)]$.

	EXP1	EXP2	HREXP1	HREXP2	FREXP1	FREXP2
DC	1.2661	1.2661	0.3821	0.6179	-0.2357	0.2357
$\hat{c}_1$	1.1303	-1.1303	0.4809	0.4809	0.9618	0.9618
$\hat{c}_2$	0.2715	0.2715	0.1155	-0.1155	0.2310	-0.2310
$\hat{c}_3$	0.0443	-0.0443	0.0189	0.0189	0.0377	0.0377
$\hat{c}_4$	0.0055	0.0055	0.0023	-0.0023	0.0047	-0.0047
$\hat{c}_5$	0.0005	-0.0005	0.0002	0.0002	0.0005	0.0005
THD	24.34%	24.34%	24.34%	24.34%	24.34%	24.34%

3.3.3 Synthesis of polynomial NLD and harmonic shifting technique

In the last sections, we showed that static memoryless nonlinearities can be approximated using polynomials. Hence, with a set of appropriate coefficients, a polynomial-based NLD can be used to represent different NLDs. Here in this section, we use *HASP_Synthesis*, as outlined in Table 3.2, to control the amplitude of the generated harmonics.

However, as some harmonics (depending on the fundamental frequency) may lie in the non-reproducible frequency band of a loudspeaker (as shown in Figure 3.8), we attempt to counter this problem by splitting the low frequencies (0–770 Hz) into 3 filter bands and introduce a harmonic shifting technique to shift the affected subbands to the reproducible frequency band.

A polynomial NLD is described in (3.2). It is a memoryless nonlinear system—i.e., the current output sample depends only on the current input sample; and superposition principle does not hold true. This system is shown in Figure 3.9.

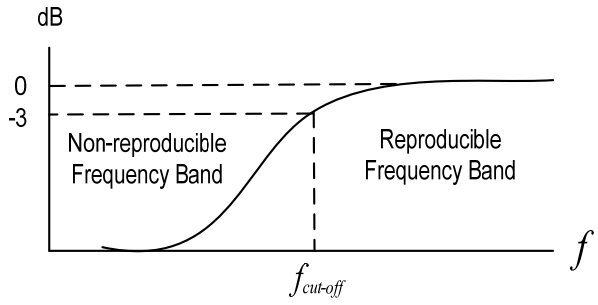

Figure 3.8. Typical frequency response of a loudspeaker.

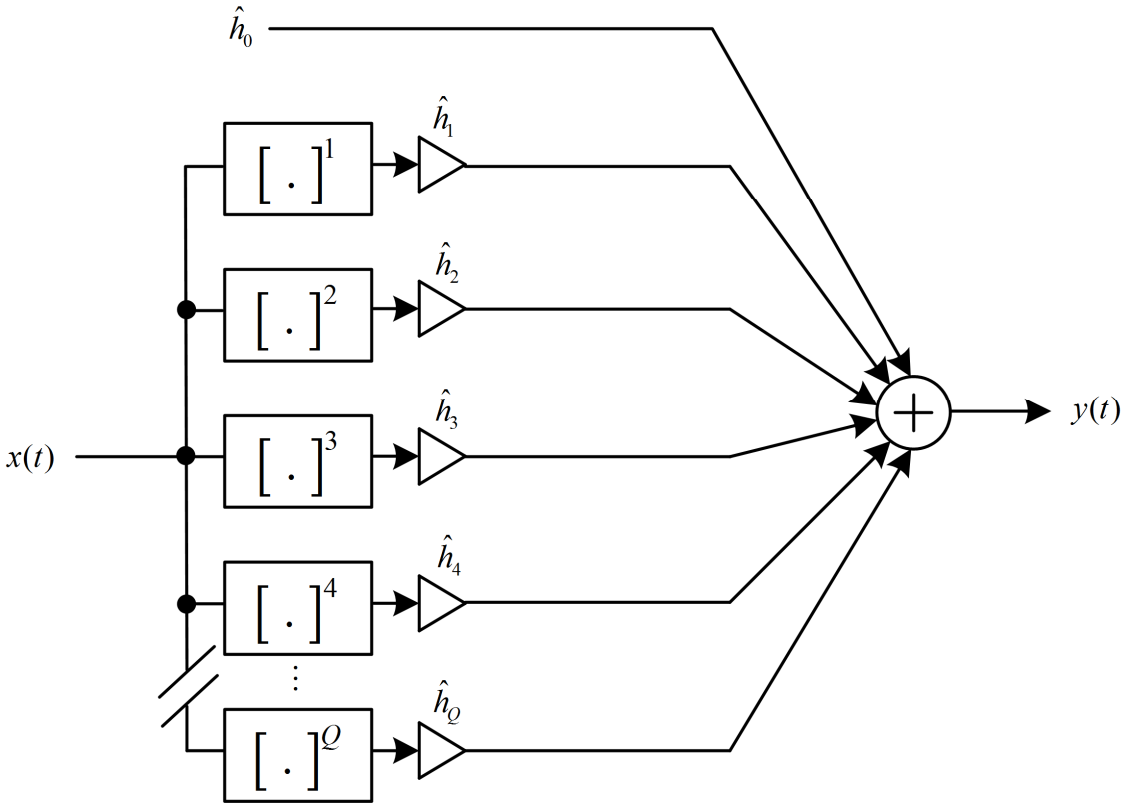

Figure 3.9. Block diagram of a polynomial NLD.

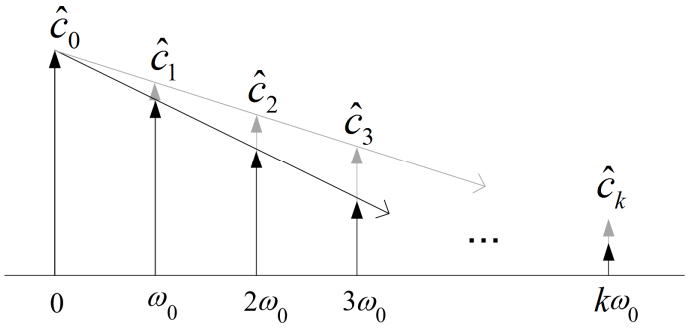

Figure 3.10. Magnitudes of generated harmonics and DC-term of NLD in frequency domain with different decay rates.

The generated harmonics (by a single tone) including the DC term and the fundamental frequency can be described by (3.10). In the frequency domain, they can be viewed as shown in Figure 3.10.

The process of NLD analysis is defined as finding the amplitudes of the harmonics, given the polynomials' coefficients. Conversely, NLD synthesis obtains the polynomial coefficients from the amplitudes of the harmonics. Based on psychoacoustic research findings [Dai00] [MGP85], the first five harmonics are critical for bass perception. In order to generate up to the fifth order harmonics, the highest order of polynomials should be five, i.e. $\hat{h}_n, n \in \{1,2,3,4,5\}$, are to be synthesized, given $\hat{c}_k, k \in \{1,2,3,4,5\}$. Individual harmonic's magnitude can be set and the polynomial-based NLD's five coefficients can be synthesized easily using *HASP_Synthesis*.

However, a problem associated with the complexity and harmonic wastage may arise. In the case of a 5th-order polynomial-based NLD and a single tone with fundamental frequency of 100 Hz from band 1 (see Figure 3.11), the synthesized harmonics of up to 5th order are 100, 200, 300, 400 and 500 Hz. In the case of a

highly-directional speaker that uses ultrasonic transducers with cut-off frequency of 700 Hz, none of the harmonics can be reproduced. Higher harmonics may need to be reproduced, but with reduced impact on the bass perception. These ultrasonic transducers cannot reproduce low frequencies due to their physical limitation (i.e., not having enough air volume or large diaphragm to vibrate to emit low frequencies).

Hence, our next proposed technique on harmonic shifting attempts to address this issue by shifting the synthesized harmonics in the non-reproducible frequency band to the reproducible frequency band.

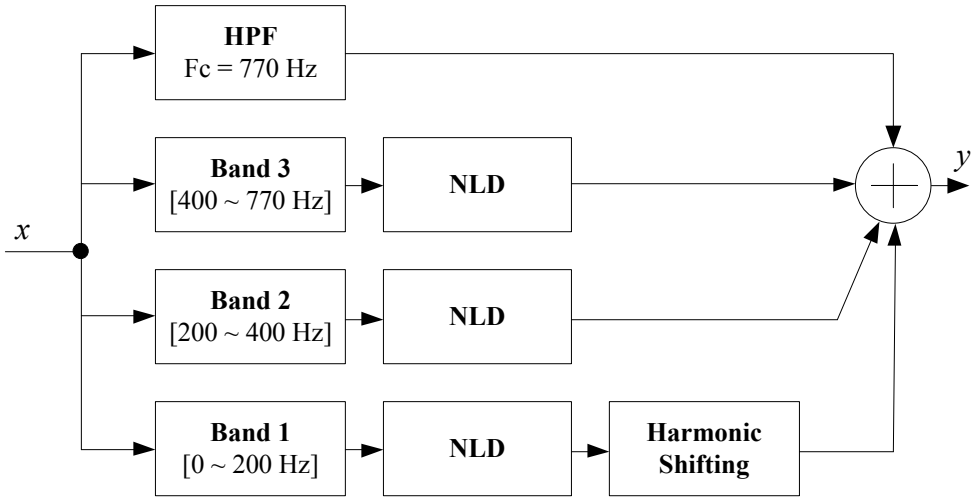

Figure 3.11. Block diagram of a simplified VBS.

3.3.4 Harmonic shifting using double sideband suppressed carrier modulation scheme

Double sideband suppressed carrier (DSB-SC) modulation technique can be used to shift specific low-frequency subbands that lie in the non-reproducible frequency band of the speaker to the resonating frequency of the speaker. The carrier frequency can be suitably chosen to lie in the reproducible frequency band (i.e., above the

fundamental resonant frequency) of the speaker that faithfully reproduces the harmonics. The frequency spectrum produced by amplitude modulation is symmetrically spaced above and below the carrier frequency. One of the reasons for using DSB-SC modulation technique over single-side band (SSB) modulation is decreased complexity. Furthermore, sub-harmonics (lower side band) can also be produced that may enhance the bass perception [HM07].

Assuming the case of a single tone, f_0 with unity amplitude, the DSB-SC equation can be expressed as

$$\begin{aligned} y(t) &= \underbrace{V_m \cos(\omega_m t)}_{f_0} \times \underbrace{V_c \cos(\omega_c t)}_{f_c} \times 2 \\ &= \frac{V_m V_c}{2} \left\{ \cos\left[(\omega_c + \omega_m) t \right] + \cos\left[(\omega_c - \omega_m) t \right] \right\} \times 2 \\ &= V_m V_c \left\{ \cos\left[(\omega_c + \omega_m) t \right] + \cos\left[(\omega_c - \omega_m) t \right] \right\}, \end{aligned} \tag{3.46}$$

where $y(t)$ is the modulated spectrum, V_m is the amplitude of the single tone, V_c is the amplitude of unmodulated carrier, $\omega_m = 2\pi f_m$ is the single tone angular frequency in radians/s and $\omega_c = 2\pi f_c$ is the unmodulated carrier angular frequency in radians/s [HM07]. Note that (3.46) has been multiplied by 2 in order to get unity output.

3.3.5 Case study: combining polynomial synthesis and harmonic shifting

This section illustrates an example in combining polynomial synthesis and harmonic shifting. A single tone, $f_0 = 100$ Hz, which falls within band 1 (see Figure 3.11) is fed into a 5th-order polynomial-based NLD with a cutoff frequency f_c of 700 Hz.

Coefficients for the polynomial-based NLD are computed based on the desired harmonics' magnitudes. The following will present a guideline for selection of the harmonics' magnitude. $\hat{c}_k$ can be set based on 2 known parameters: desired decay rate and magnitude of the first harmonic. Setting the magnitude of the first harmonic $\hat{c}_1 = 0.9$ and a linear decay rate, based on the difference between 2 harmonics, Δ of 0.1 dBFS, the other 4 magnitudes can be computed based on (3.47) below:

$$\Delta = \hat{c}_k - \hat{c}_{k+1}, \tag{3.47}$$

where $\hat{c}_k, k \in \{1,2,3,4,5\}$. Table 3.6 shows the computed magnitudes.

Table 3.6: Computed magnitudes of harmonics.

$\hat{c}_1$	0.9
$\hat{c}_2$	0.8
$\hat{c}_3$	0.7
$\hat{c}_4$	0.6
$\hat{c}_5$	0.5

After applying the harmonic shifting technique, the synthesized harmonics (100, 200, 300, 400, and 500) Hz are now shifted to (800, 900, 1000, 1100 and 1200) Hz with additional harmonics (200, 300, 400, 500 and 600) Hz being generated as a result of the DSB-SC modulation. Assuming the case of a highly directional speaker with cut-off frequency of 700 Hz, all the frequencies in the upper side band (above the carrier), lie in the reproducible region, while most frequencies in the lower side band (below the carrier) are greatly attenuated. However, these attenuated frequencies may still be perceivable to a certain extent. The magnitude of these shifted harmonics

(dependent on the frequency band) and the f_c (dependent on the speaker's cut-off frequency) must be carefully selected so as to produce desirable results.

Equation (3.48) below can be used to shape the harmonic spectrum in 2 steps if necessary.

$$\breve{c}_k = \begin{cases} G.\hat{c}_k{}^{\alpha} & \alpha > 1, \\ G.\hat{c}_k{}^{\alpha} & 0 < \alpha < 1, \\ G.\hat{c}_k & \alpha = 1, \\ G & \alpha = 0. \end{cases} \tag{3.48}$$

The first step is to select a suitable decay rate by adjusting α. $\alpha > 1$ results in a steeper decay curve, while $\alpha < 1$ results in a more gradual decay curve. Once a suitable α is selected, the next step is to select a suitable gain, G where $G > 1$ causes amplification and $0 \le G < 1$ causes attenuation. The result is a new set of harmonics, $\breve{c}_k, k \in \{1,2,3,4,5\}$. We can set $\alpha = 1$ and $G = 1$ to retain the initial set of $\hat{c}_k$ since it may not be necessary to further modify the spectrum for bands that do not require the process of harmonic shifting.

Figure 3.11 illustrates an example of a simplified VBS where frequencies below 770 Hz are split into filter bands and each is processed by a NLD. As most of the synthesized harmonics from band 1 fall into non-reproducible frequency band, DSB-SC modulation technique is used to shift frequencies in band 1 to the reproducible frequency band before mixing. This combined technique can further enhance the bass perception.

Figure 3.12 shows spectrogram plots of (a) original input signal, (b) processed signal without modulation, and (c) processed signal with modulation. Figures 3.12 (b) and (c) both clearly show increased energies across the frequencies (y-axis), illustrating the effects of the synthesized harmonics. On the other hand, Figure 3.12 (c) shows increased energies from 700 to 1100 Hz and a reduction of energies from 0 to 200 Hz, signifying the effects of the shifted harmonics.

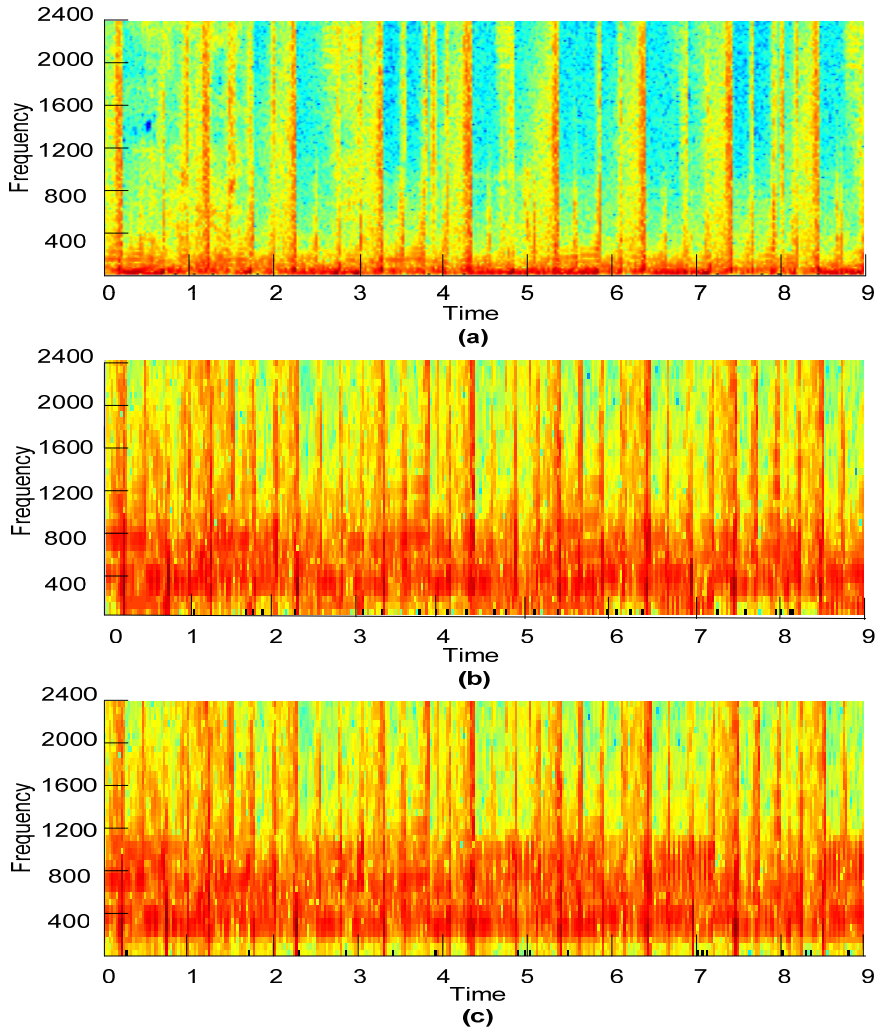

Figure 3.12: Spectrogram plots with (a) orginal input signal, (b) processed signal without modulation, and (c) processed signal with modulation. [Color Keys: Red = Strong energy, Yellow = weak energy, Blue = weaker or lack of energy]

3.3.6 Subjective evaluation and results

Subjective testing was carried out to examine the effectiveness of the proposed method over the unprocessed sound track. The bass enhancement algorithm was implemented using Analog Device's BlackFin 533 Ez-Lite Kit [Ana11]. Appropriate values for the harmonics' magnitudes were selected. In order to demonstrate the problem of high cut-off, a highly directional speaker [Ame07] with cut-off of approximately 1000 Hz was used in the test. Twenty-four subjects, 20–40 years of age, participated in the tests. Arrangement of the speaker, position of subject and the conditions of listening room followed the ITU standard ITU-R BS.1116 [ITU97]. Multi-stimuli with hidden reference and anchor (MUSHRA) of ITU-R BS.1534-1 [ITU01] were used as the test procedure. Rating scales were based on ITU-T P.1534 continuous scale and ITU-R 5-point continuous impairment scale [ITU90]. Three stimuli with appropriate amounts of bass energies were used for our subjective tests. They were:

- Stimulus 1: Puff Daddy – I'll Be Missing You
- Stimulus 2: Britney Spears – I Love Rock 'N' Roll
- Stimulus 3: Empire Brass Quintet – Hopper Dance

Two models are evaluated in the listening test. Model 1 (M1) uses the polynomial-based NLD without DSB-SC modulation and Model 2 (M2) uses the same polynomial-based NLD but with DSB-SC modulation technique. Data are compiled and plotted in a bar diagram. Figure 3.13 below contains the compiled scores for both models on two attributes, bass quantity and impairment. Scores for

anchor (high-passed signal) and reference signal are not included here. Scores were computed by taking the mean of the scores from 24 subjects.

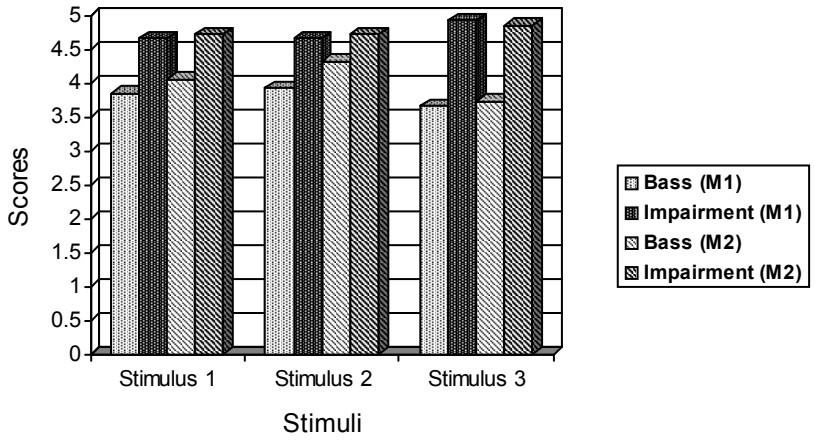

Figure 3.13: Listening test results.

It is clear from Figure 3.13 that both models showed significant enhancement of bass perception, especially on stimulus 2, with M1 and M2 having scores of 3.9 and 4.1, respectively. It can also be observed that impairment was generally imperceptible with a score of 4.5 and higher for each stimulus. Feedback from subjects mentioned the better quality in terms of naturalness and sensation of the processed sounds. Results also show that M2 employing DSB-SC modulation outperformed M1, with at least half a unit difference in all stimuli and a maximum of 1.2 in stimulus 2.

3.4 Conclusions

Triple Duality Theorem has been stated and proved in this chapter. Detailed mathematical analysis on the relationship between harmonic analysis, synthesis and symmetries of nonlinear function has been proved. The mathematical findings are

converted into an algorithm, known as HASP. HASP has been listed and applied to the cases of harmonic analysis and synthesis. The generalized harmonic analysis on ATSR and EXP nonlinearities are presented. The harmonic analysis on EXP NLDs shows that THD metric may not be suitable to measure the degree of nonlinearities. The harmonic synthesis idea is combined with DSB-SC modulation technique to make parametric array loudspeakers sound bassy. The subjective listening test results support the feasibility of the idea.

In general, this chapter has shown the duality relationship between sound analysis and synthesis. We have also proposed a set of closed-form design equations for sound analysis and synthesis, and the HASP algorithm. The significance of this work lies in its duality relationship and simplicity. For duality, we uncover the relationships between analysis-synthesis, and nonlinear functions' symmetries, whereas, for simplicity, we propose computational efficient HASP routines that are easy to implement for sound synthesis and nonlinear harmonic distortion analysis.

CHAPTER 4

MULTITONE ANALYSIS ON NONLINEAR DEVICES

4.1 Introduction

When a multitone signal is applied to an NLD, both harmonics and intermodulation components are generated. There are different techniques to compute the harmonics and intermodulation components at the output of polynomial nonlinearities without using discrete Fourier transform (DFT), which has well-known limitations if used to compute output components, such as frequency resolution, spectral leakage, and sampling aliasing [OSB99].

The precise intermodulation product computing technique (PIPCT) has been proposed to compute intermodulation products at the output of polynomial nonlinear equations [Yea08], for which PIPCT is used to compute intermodulation components precisely and efficiently, when NLD is fed by a multitone signal.

However, there are two limitations in PIPCT: (a) input multitone components must have unitary or constant amplitude; and (b) each phase of the input multitone components must be preset to zero (i.e., the phase values of the intermodulation components cannot be computed). Given these limitations, the following research problems are defined:

1. Reduction of the computational load of PIPCT up to 50%;

2. Overcoming the constant input amplitudes limitation; and

3. Elimination of the zero-input phase limitation.

This chapter is organized as follows: First, an overview on PIPCT is introduced in Section 4.2. In Section 4.3, mathematical analysis on the two lemmas from [Yea08] is carried out. With the insight gained from this analysis, a technique to reduce computational load is proposed, with solutions for the two limitations in Section 4.4. In Section 4.5, the harmonic addition theorem is proposed and two original proofs are offered for it. The harmonic addition theorem is used to add overlapping frequency components having different amplitudes and phases. A detailed description of the proposed algorithm is also described in this section. Section 4.6 shows the simulation results and point out FFT pitfalls on intermodulation distortion spectral analysis. Finally, Section 4.7 summarizes and concludes this chapter.

4.2 Overview of PIPCT

The two lemmas of PIPCT are based on the monomial nonlinearity. The monomial nonlinearity is defined as

$$y_q(t) = \left[x(t)\right]^q, \tag{4.1}$$

where $x(t)$ is the input, $y_q(t)$ is the output, and q is the order of monomial nonlinearity. Let us consider the input as a multitone signal with its components having arbitrary amplitudes (A_k), frequencies (ω_k), and phases (φ_k) as follows:

$$x(t) = \sum_{k=0}^{N-1} A_k \cos(\omega_k t + \varphi_k), \tag{4.2}$$

where k is the index number and N is the number of input multitone components. Let us define the output from the monomial nonlinearity as

$$y_q(t) = \sum_{k=0}^{L-1} B_k \cos(\tilde{\omega}_k t + \psi_k), \tag{4.3}$$

where B_k, $\tilde{\omega}_k$, and ψ_k are amplitudes, frequencies, and phases of the output components, respectively. In (4.3), k is the index number and L is the number of output components. Note that some of the output components may overlap, i.e., share the same frequencies. Let us write a polynomial nonlinearity as

$$\begin{aligned} y(t) &= \hat{f}\left[x(t)\right] \\ &= \hat{h}_0 + \hat{h}_1 x(t) + \hat{h}_2\left[x(t)\right]^2 + \cdots + h_{Q-1}\left[x(t)\right]^{Q-1} + h_Q\left[x(t)\right]^Q \\ &= \sum_{q=0}^{Q} \hat{h}_q\left[x(t)\right]^q \\ &= \sum_{q=0}^{Q} \hat{h}_q y_q(t), \end{aligned} \tag{4.4}$$

where $\hat{f}(.)$ is the polynomial-approximated nonlinear function, $\hat{h}_q$ is the polynomial coefficients, Q is the highest order of polynomial nonlinearity, and q is the index of the monomial nonlinearity. The polynomial nonlinearity is formed by adding monomial nonlinearities, each of which is multiplied by a polynomial coefficient.

The input signal to the monomial nonlinearity for the case of PIPCT is

$$x(t) = A\sum_{k=0}^{N-1} \cos(\omega_k t), \tag{4.5}$$

where $A_k = A$ and $\varphi_k = 0$ in (4.2). That is, the amplitudes of the input multitone are a constant and initial phases are all set to zero.

Let us form a column vector **f** with N number of input multitone frequencies, $\omega_0, \omega_1, \ldots, \omega_{N-1}$ from (4.2) or (4.5) as follows:

$$\mathbf{f} = \begin{bmatrix} \omega_0 & \omega_1 & \ldots & \omega_{N-1} \end{bmatrix}^T \in \mathbb{R}^N, \tag{4.6}$$

where $\mathbb{R}^N$ denotes all column vectors with N real components, while $[.]^T$ is the transposed operation. Since the polynomial multiplication is equivalent to the convolution [Yea08], the following process can be restated:

$$A(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_{N-1} z^{N-1}, \tag{4.7}$$

$$B(z) = b_0 + b_1 z + b_2 z^2 + \cdots + b_{N-1} z^{N-1}, \tag{4.8}$$

and

$$C(z) = A(z) \times B(z) = c_0 + c_1 z + c_2 z^2 + \cdots + c_{2(N-1)} z^{2(N-1)}. \tag{4.9}$$

The column vectors with polynomial coefficients of (4.7), (4.8), and (4.9) are

$$\mathbf{a} = \begin{bmatrix} a_0 & a_1 & \ldots & a_{N-2} & a_{N-1} \end{bmatrix}^T \in \mathbb{R}^N, \tag{4.10}$$

$$\mathbf{b} = \begin{bmatrix} b_0 & b_1 & \cdots & b_{N-2} & b_{N-1} \end{bmatrix}^T \in \mathbb{R}^N, \tag{4.11}$$

and

$$\mathbf{c} = \begin{bmatrix} c_0 & c_1 & \cdots & c_{2N-1} & c_{2(N-1)} \end{bmatrix}^T \in \mathbb{R}^{2N-1}. \tag{4.12}$$

Notice that the vector **c** can be formed from the vectors **a** and **b** by convolution as follows:

$$\mathbf{c} \leftarrow \text{conv}(\mathbf{a}, \mathbf{b}) = \mathbf{a} \otimes \mathbf{b}, \tag{4.13}$$

where $\otimes$ is the convolution operator and $\leftarrow$ is the symbolic assignment operator. The vector elements of $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ can be related by

$$c_i = \sum_{j=0}^{N-1} a_j b_{i-j}, \tag{4.14}$$

for $0 \le i \le 2(N-1)$ and $0 \le i-j \le N-1$.

Lemma 1 for PIPCT:

A new column vector **d**, which lists the sum and difference $a_j + b_{i-j}$ and $a_j - b_{i-j}$ correspondingly to each produce term $a_j b_{i-j}$ in (4.14) can be constructed. The vector **d** is defined as

$$\mathbf{d} = \begin{bmatrix} d_0 & d_1 & \dots & d_{L_\mathbf{d}-2} & d_{L_\mathbf{d}-1} \end{bmatrix}^T \in \mathbb{R}^{L_\mathbf{d}}, \tag{4.15}$$

where each element of **d** is formed by symbolic splitting as

$$\mathbf{d} \leftarrow \text{split}(\mathbf{c}), \tag{4.16}$$

where each element of **c** by convolution in (4.14) is split into two elements of **d** as

$$a_j b_{i-j} \mapsto \left[\left(a_j + b_{i-j} \right), \left(a_j - b_{i-j} \right) \right] \mapsto \left(d_k, d_{k+1} \right), \tag{4.17}$$

for $0 \le i \le 2(N-1)$, $0 \le i-j \le N-1$, and $0 \le k \le (L_\mathbf{d} - 1)$. The vector length of **d** is $L_\mathbf{d} = 2N^2$.

Lemma 2 for PIPCT:

Referring to the monomial nonlinearity in (4.1) and the output signal in (4.3), the scaling factor of $(1/2)^{q-1}$ is associated with each spectral component $[\cos(\tilde{\omega}_0 t), \dots, \cos(\tilde{\omega}_{L-1} t)]$. That is each $B_k = (1/2)^{q-1}$ in (4.3) for $k = 0, 1, \dots, L-1$.

Due to the two limitations of PIPCT, as described in Section 4.1, ψ_n in (4.3), the phases of the output components cannot be computed. Lemma 2 is only valid for the special case of $A_k = 1$ in (4.2) for the monomial nonlinearity in (4.1).

The two limitations are solved in the next sections with the detailed descriptions of algorithmic development from the monomial nonlinearity in (4.1) to the polynomial nonlinearity in (4.4). Moreover, the insights gained from the analysis of Lemma 1 help us to further develop the computational reduced technique to compute the output components—amplitudes, frequencies, and phases.

4.3 Analysis of Lemma 1 of PIPCT

The trigonometric interpretation of Lemma 1 is provided as follows: For $q = 2$ in (4.1) and $N = 2$ in (4.5), the trigonometric expansion (and their respective polynomial multiplication and expansion) is given by

$$y(t) = \overbrace{[x(t)]^2}^{C(z)} = [\overbrace{\cos(\omega_0 t) + \cos(\omega_1 t)}^{A(z)=a_0+a_1 z}] \times [\overbrace{\cos(\omega_0 t) + \cos(\omega_1 t)}^{B(z)=b_0+b_1 z}]$$
$$= \overbrace{\cos(\omega_0 t)\cos(\omega_0 t)}^{a_0 b_0} + \overbrace{\cos(\omega_0 t)\cos(\omega_1 t)}^{a_0 b_1 z} + \overbrace{\cos(\omega_1 t)\cos(\omega_0 t)}^{a_1 b_0 z} + \overbrace{\cos(\omega_1 t)\cos(\omega_1 t)}^{a_1 b_1 z^2}, \tag{4.18}$$

where the equivalent polynomial multiplication in (4.9) is

$$\begin{aligned} C(z) &= A(z)B(z) \\ &= (a_0 + a_1 z) \times (b_0 + b_1 z) \\ &= a_0 b_0 + a_0 b_1 z + a_1 b_0 z + a_1 b_1 z^2 \\ &= \overbrace{a_0 b_0}^{c_0} + \overbrace{\left(a_0 b_1 + a_1 b_0\right)}^{c_1} z + \overbrace{a_1 b_1}^{c_2} z^2 \\ &= c_0 + c_1 z + c_2 z^2. \end{aligned} \tag{4.19}$$

For $q = 2$ in (4.1) and $N = 3$ in (4.5),

$$\begin{aligned} y(t) &= \overbrace{[x(t)]^2}^{C(z)} = [\overbrace{\cos(\omega_0 t) + \cos(\omega_1 t) + \cos(\omega_2 t)}^{A(z)=a_0+a_1 z+a_2 z^2}] \times [\overbrace{\cos(\omega_0 t) + \cos(\omega_1 t) + \cos(\omega_2 t)}^{B(z)=b_0+b_1 z+b_2 z^2}] \\ &= \overbrace{\cos(\omega_0 t)\cos(\omega_0 t)}^{a_0 b_0} + \overbrace{\cos(\omega_0 t)\cos(\omega_1 t)}^{a_0 b_1 z} + \overbrace{\cos(\omega_1 t)\cos(\omega_0 t)}^{a_1 b_0 z} + \overbrace{\cos(\omega_0 t)\cos(\omega_2 t)}^{a_0 b_2 z^2} \\ &\quad + \overbrace{\cos(\omega_1 t)\cos(\omega_1 t)}^{a_1 b_1 z^2} + \overbrace{\cos(\omega_2 t)\cos(\omega_0 t)}^{a_2 b_0 z^2} + \overbrace{\cos(\omega_1 t)\cos(\omega_2 t)}^{a_1 b_2 z^3} \\ &\quad + \overbrace{\cos(\omega_2 t)\cos(\omega_1 t)}^{a_2 b_1 z^3} + \overbrace{\cos(\omega_2 t)\cos(\omega_2 t)}^{a_2 b_2 z^4}, \end{aligned} \tag{4.20}$$

where

$$\begin{aligned} C(z) &= A(z)B(z) \\ &= (a_0 + a_1 z + a_2 z^2 + a_3 z^3)(b_0 + b_1 z + b_2 z^2 + b_3 z^3) \\ &= \overbrace{a_0 b_0}^{c_0} + \overbrace{(a_0 b_1 + a_1 b_0)}^{c_1} z + \overbrace{(a_0 b_2 + a_1 b_1 + a_2 b_0)}^{c_2} z^2 + \overbrace{(a_0 b_3 + a_1 b_2 + a_2 b_1 + a_3 b_0)}^{c_3} z^3 \\ &= c_0 + c_1 z + c_2 z^2 + c_3 z^3. \end{aligned} \tag{4.21}$$

Notice that the polynomial coefficients of $C(z)$ in (4.9) are related to the polynomial coefficients of $A(z)$ in (4.7) and those of $B(z)$ in (4.11) by the convolution operation in (4.14). The algebraic expansions of polynomial coefficients of $C(z)$ are in triangular pattern, as shown in Table 4.1. The first two sub-tables represents $N = 2$ and $N = 3$ in (4.19) and (4.21), respectively, while the last sub-table shows the generalized triangular pattern for $N = n + 1$, where n is a positive integer greater than or equal to one.

Table 4.1: Triangular pattern of the polynomial coefficients of $C(z)$ for the cases of $N = 2,3$ and the general case of $N = n + 1$. Note $q = 2$ for all the cases.

$N = 2$	$c_0 = a_0 b_0$
	$c_1 = a_0 b_1 + a_1 b_0$
	$c_2 = a_1 b_1$
$N = 3$	$c_0 = a_0 b_0$
	$c_1 = a_0 b_1 + a_1 b_0$
	$c_2 = a_0 b_2 + a_1 b_1 + a_2 b_0$
	$c_3 = a_1 b_2 + a_2 b_1$
	$c_4 = a_2 b_2$
$N = n+1$	$c_0 = a_0 b_0$
	$c_1 = a_0 b_1 + a_1 b_0$
	$c_2 = a_0 b_2 + a_1 b_1 + a_2 b_0$
	$\vdots$
	$c_n = a_0 b_n + a_1 b_{n-1} + \cdots + a_{n-1} b_1 + a_n b_0$
	$\vdots$
	$c_{2n-2} = a_{n-2} b_n + a_{n-1} b_{n-1} + a_n b_{n-2}$
	$c_{2n-1} = a_{n-1} b_n + a_n b_{n-1}$
	$c_{2n} = a_n b_n$

The number of components generated in **d** by PIPCT [Yea08] is derived by recursion as follows: Starting from the second monomial [i.e., $q = 2$ in (4.1)] , set **a** ← **f** and **b** ← **f**, where **a**, **b**, and **f** are as defined as shown in (4.10), (4.11), and (4.6) respectively. Based on Lemma 1, the output frequency components are available in **d** (which is also equivalent to the vector length $L_{\mathbf{d}_q}^{\text{PIPCT}}$, where q denotes the nonlinearity order of the monomial nonlinearity) as defined in (4.15) by the convolution operation in (4.14) and splitting operation in (4.17). The total number of generated frequency components for the second-order case is given as

$$L_{\mathbf{d}_{q=2}}^{\text{PIPCT}} = 2N^2 = \frac{(2N)^2}{2}, \tag{4.22}$$

where N is the vector length of **a** and **b**, respectively. It should also be noted that N is also the number of input components in (4.2). Equation (4.22) is the vector length of **d** by the Lemma 1 for the second order monomial nonlinearity. For the third order monomial [i.e., $q = 3$ in (4.1)] , the computed **d** in the second order is reapplied, and the two input vectors are reset as follows: **a** ← **d** and **b**← **f**. Again, a new vector **d** is generated by applying the Lemma 1. Hence, the total number of generated frequency components for the third order monomial case is

$$L_{\mathbf{d}_{q=3}}^{\text{PIPCT}} = 2NL_{\mathbf{d}_{q=2}}^{\text{PIPCT}} = \frac{(2N)^3}{2}. \tag{4.23}$$

Therefore, the total number of generated frequency components in **d** can be formulated recursively for any monomial nonlinearity order q that is a positive integer and greater than one as

$$L_{\mathbf{d}_q}^{\text{PIPCT}} = 2NL_{\mathbf{d}_{q-1}}^{\text{PIPCT}} = \frac{(2N)^q}{2}. \tag{4.24}$$

Equation (4.24) shows a computational growth of the Lemma 1. It can be concluded that the number of frequency components grows exponentially with q. In the next section, a technique will be proposed to decrease the computational growth by exploiting triangular symmetric pattern of polynomial multiplication, while obtaining the same final results as the original technique.

4.4 Proposed Technique for Monomial NLDs

In this section, theoretical analysis on reduction of exponential computational growth is presented, the upper and lower bounds for computational cost savings are derived, and the solutions for the two limitations of PIPCT are proposed.

4.4.1 Reduction of exponential computation growth

The proposed technique is identified as *modified* PIPCT (MPIPCT) and compared with PIPCT described in the last section. Since Lemma 1 is based on the recursive computation, the number of generated components in **d** is first reduced for the case of q = 2, and a generalized formula is derived, stating how much the size of **d** is reduced if the triangular symmetric pattern is exploited. This reduction leads to significant computational savings, especially for high-order monomials.

The essence of this idea is based on the triangular symmetric pattern of polynomial coefficients expansion in **c**, as illustrated in Table 4.2 and Figure 4.1. The reduction is achieved based on the following observations: As an example, note that

$$\begin{aligned}\overbrace{\cos(\omega_0 t)\cos(\omega_1 t)}^{a_0 b_1} &= 1/2\overbrace{\cos((\omega_0+\omega_1)t)}^{a_0+b_1} + 1/2\overbrace{\cos((\omega_0-\omega_1)t)}^{a_0-b_1} \\ \overbrace{\cos(\omega_1 t)\cos(\omega_0 t)}^{a_1 b_0} &= 1/2\overbrace{\cos((\omega_1+\omega_0)t)}^{a_1+b_0} + 1/2\overbrace{\cos((\omega_1-\omega_0)t)}^{a_1-b_0}.\end{aligned} \tag{4.25}$$

By (4.25) and since $\cos(-\phi) = \cos\phi$,

$$a_0 b_1 = a_1 b_0, \tag{4.26}$$

Therefore, $a_x b_y$ and $a_y b_x$ are the same and can be reduced to $2a_x b_y$ such that the repeating terms can be reduced and recorded using *gain multiplier* $\in \{1,2\}$. When the two terms in any coefficient of $C(z)$ are repeating, they can be reduced to one as

$$\left(a_0 b_1, a_1 b_0\right) \rightarrow 2 \times \left(a_0 b_1\right), \tag{4.27}$$

where *gain multiplier = 2*. Based on this idea of reducing repeating terms by introducing the *gain multiplier*, we study the triangular symmetric pattern of polynomial coefficient expansion of $C(z)$, to derive the generalized equations revealing how much components growth can be reduced by such reduction from the exploitation of triangular symmetry (See Figure 4.1). For the purpose of mathematical analysis, let us define two more vectors as follows:

$$\mathbf{c}' = \begin{bmatrix} c'_0 & c'_1 & \dots & c'_{R-1} \end{bmatrix}^T \in \mathbb{R}^R, \tag{4.28}$$

and

$$\mathbf{e} = \begin{bmatrix} e_0 & e_1 & \dots & e_{R-1} \end{bmatrix}^T \in \mathbb{R}^R, \tag{4.29}$$

where $\mathbf{c}'$ holds the coefficients of the reduced triangular symmetric pattern and $\mathbf{e}$ holds the repeating terms (See Table 4.2 and Figure 4.1). Let us define

$$\mathbf{g} = \begin{bmatrix} g_0 & g_1 & \dots & g_{S-2} & g_{S-1} \end{bmatrix}^T \in \{1,2\}^S, \tag{4.30}$$

to store the *gain multipliers*, where S is the length of $\mathbf{g}$.

Figure 4.1 is the pictorial representation of triangular symmetric pattern and reduced triangular symmetric patterns presented in Table 4.1. From Table 4.2, the repeating terms for each coefficient of $C(z)$ can be formulated as

$$e_k = \begin{cases} \lceil k/2 \rceil, & \text{for } k = 0,1,\ldots,N-1, \\ e_{2n-k}, & \text{for } k = N, N+1,\ldots,2(N-1), \end{cases} \tag{4.31}$$

$$\text{for } k = 0,1,\ldots 2n, \text{ where } n = N-1,$$

where $\lceil x \rceil$ is the ceiling function: $\min\{m \mid m \geq x, \text{ integer } m\}$. Notice that n = $N - 1$ is the coefficient index for the symmetry of the folding row of the triangular symmetric pattern and reduced triangular symmetric pattern.

Table 4.2: Triangular symmetric pattern, reduced triangular symmetric pattern, and repeating terms for the cases of N = 2, 3 and the general case of $N = n+1$ for the second order monomial nonlinearity.

N	Elements of **c**	Elements of **c'**	Elements of **e**	Elements of **g**
2	$c_0 = a_0 b_0$	$c'_0 = a_0 b_0$	$e_0 = \lceil 0/2 \rceil = 0$	$g_0 = 1$
	$c_1 = \overbrace{a_0 b_1}^{(1)} + \overbrace{a_1 b_0}^{(1)}$	$c'_1 = \overbrace{a_0 b_1}^{(1)}$	$e_1 = \lceil 1/2 \rceil = 1$	$g_1 = 2$
	$c_2 = a_1 b_1$	$c'_2 = a_1 b_1$	$e_2 = e_{2-2} = e_0 = 0$	$g_2 = 1$
3	$c_0 = a_0 b_0$	$c'_0 = a_0 b_0$	$e_0 = \lceil 0/2 \rceil = 0$	$g_0 = 1$
	$c_1 = \overbrace{a_0 b_1}^{(1)} + \overbrace{a_1 b_0}^{(1)}$	$c'_1 = \overbrace{a_0 b_1}^{(1)}$	$e_1 = \lceil 1/2 \rceil = 1$	$g_1 = 2$
	$c_2 = \overbrace{a_0 b_2}^{(1)} + a_1 b_1 + \overbrace{a_2 b_0}^{(1)}$	$c'_2 = \overbrace{a_0 b_2}^{(1)} + a_1 b_1$	$e_2 = \lceil 2/2 \rceil = 1$	$g_2 = 2, g_3 = 1$
	$c_3 = \overbrace{a_1 b_2}^{(1)} + \overbrace{a_2 b_1}^{(1)}$	$c'_3 = \overbrace{a_1 b_2}^{(1)}$	$e_3 = e_{4-3} = e_1 = 1$	$g_4 = 2$
	$c_4 = a_2 b_2$	$c'_4 = a_2 b_2$	$e_4 = e_{4-4} = e_0 = 0$	$g_5 = 1$
$n+1$	$c_0 = a_0 b_0$	$c'_0 = a_0 b_0$	$e_0 = \lceil 0/2 \rceil = 0$	$g_0 = 1$
	$c_1 = \overbrace{a_0 b_1}^{(1)} + \overbrace{a_1 b_0}^{(1)}$	$c'_1 = \overbrace{a_0 b_1}^{(1)}$	$e_1 = \lceil 1/2 \rceil = 1$	$g_1 = 2$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$c_n = \overbrace{a_0 b_n}^{(1)} + \overbrace{a_1 b_{n-1}}^{(2)} + \cdots + \overbrace{a_{n-1} b_1}^{(2)} + \overbrace{a_n b_0}^{(1)}$	$c'_n = \overbrace{a_0 b_n}^{(1)} + \overbrace{a_1 b_{n-1}}^{(2)} + \cdots$	$e_n = \lceil n/2 \rceil$	$g_n = 2, g_{n+1} = 2, \ldots$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$c_{2n-1} = \overbrace{a_{n-1} b_n}^{(1)} + \overbrace{a_n b_{n-1}}^{(1)}$	$c'_{2n-1} = \overbrace{a_{n-1} b_n}^{(1)}$	$e_{2n-1} = e_{2n-(2n-1)}$	$g_{S-2} = 2$
	$c_{2n} = a_n b_n$	$c'_{2n} = a_n b_n$	$e_{2n} = e_{2n-2n}$	$g_{S-1} = 1$

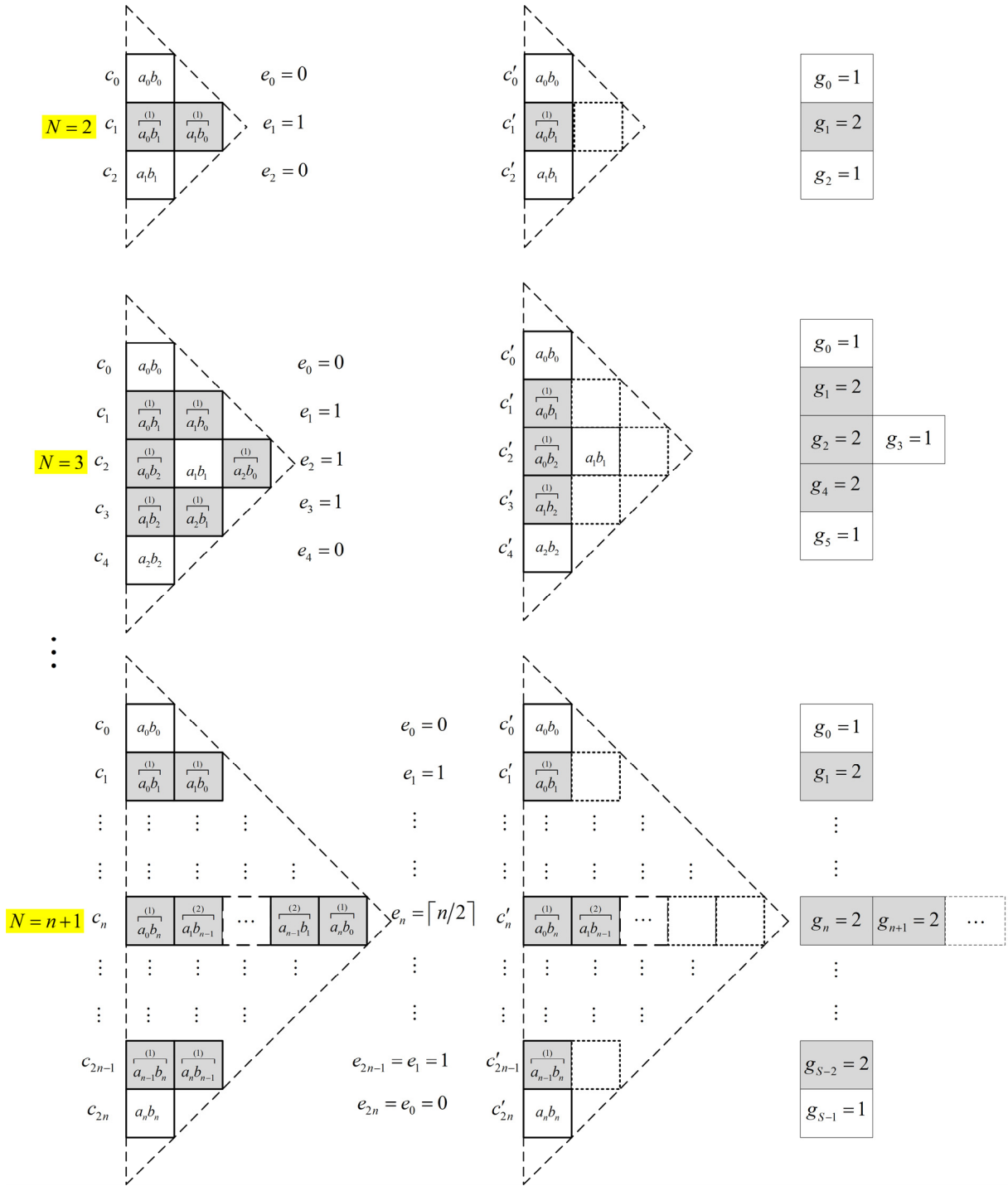

Figure 4.1: Triangular symmetric pattern, reduced triangular symmetric pattern of polynomial expansions, and *gain multipliers* in Table 4.2.

To derive the generalized formula describing the number of terms that can be discarded without losing the information while reducing the computational load, let $\Psi(N)_{q=2}$ be the total number of repeating terms, parameterized by N for the case of the second-order monomial nonlinearity. The pattern of the repeating terms can also be observed from Table 4.3. The total number of repeating terms for second-order monomial nonlinearity is

$$\Psi(N)_{q=2} = \sum_{k=0}^{2(N-1)} e_k = \begin{cases} 3\left(\frac{N-1}{2}\right) + 4\sum_{k=1}^{\frac{N-3}{2}} k, & \text{if } N \text{ is odd,} \\ \frac{N}{2} + 4\sum_{k=1}^{\frac{N}{2}-1} k, & \text{if } N \text{ is even.} \end{cases} \tag{4.32}$$

Table 4.3: Pattern of repeating terms in **c** for $N = 2,3,\ldots,11$.

N	$\Psi(N)_{q=2}$	Pattern
2	1	$1+2\times(0)$
3	3	$1+2\times(0+1)$
4	6	$2+2\times(0+1+1)$
5	10	$2+2\times(0+1+1+2)$
6	15	$3+2\times(0+1+1+2+2)$
7	21	$3+2\times(0+1+1+2+2+3)$
8	28	$4+2\times(0+1+1+2+2+3+3)$
9	36	$4+2\times(0+1+1+2+2+3+3+4)$
10	45	$5+2\times(0+1+1+2+2+3+3+4+4)$
11	55	$5+2\times(0+1+1+2+2+3+3+4+4+5)$

Equation (4.32) can be further simplified by using Gauss summation formula [Eve69] and the closed-form formula for triangular numbers [CG96]. Let us denote

$$T(n) = 1+2+3+\cdots+(n-1)+n = \sum_{k=1}^{n} k = \frac{n(n+1)}{2}, \tag{4.33}$$

where *T(n)* is both the triangular number and the sum of the first *n* natural numbers. By applying (4.33) to (4.32) for both cases of odd and even N (substituting $n=(N-3)/2$ and $n=N/2-1$ into (4.33) first and then back into (4.32)), (4.32) becomes

$$\Psi(N)_{q=2}=\frac{N(N-1)}{2},\forall N\geq 2. \tag{4.34}$$

Therefore, the total number of repeating terms in $\mathbf{d}$ for $q=2$ can be expressed as

$$\Phi(N)_{q=2}=2\Psi(N)_{q=2}=N(N-1). \tag{4.35}$$

By incorporating **g**, such repeating terms can be discarded from the computation, the total number of essential components left in **d** for $q=2$,

$$L_{\mathbf{d}_{q=2}}^{\text{MPIPCT}}=L_{\mathbf{d}_{q=2}}^{\text{PIPCT}}-\Phi(N)_{q=2}=2\left[N^2-\Psi(N)_{q=2}\right]. \tag{4.36}$$

Substituting (4.22) and (4.35) into (4.36), the simpler expression is obtained as

$$L_{\mathbf{d}_{q=2}}^{\text{MPIPCT}}=N(N+1). \tag{4.37}$$

Here, we proceed to derive the expression for any arbitrary positive integer *q*. We note that for *q* > 2, the *distributive rule*, i.e., $a\times(b+c)=a\times b+a\times c$ can be applied to extend the result for $q\geq 2$.

Notice that PIPCT is a recursive computational technique. Nevertheless, at the initialization stage of the algorithm, the following vector assignments are performed.

$$\mathbf{a}\leftarrow\mathbf{f}, \tag{4.38}$$

and

$$\mathbf{b}\leftarrow\mathbf{f}. \tag{4.39}$$

At the end of every recursion, except for the very first one, the following vector assignments are performed:

$$\mathbf{a} \leftarrow \mathbf{d}, \tag{4.40}$$

and

$$\mathbf{b} \leftarrow \mathbf{f}, \tag{4.41}$$

where **d** in (4.15) is the resultant vector from one recursion. For $q > 2$, the distributive rule can be applied as follows:

$$\begin{aligned} C(z) = D(z) \times B(z) &= \overbrace{\left(d_0 + d_1 z + \cdots + d_s z^s\right)}^{D(z)}\left(b_0 + b_1 z + \cdots + b_{N-1} z^{N-1}\right) \\ &= b_0 D(z) + b_1 D(z) z + \cdots + b_{N-1} D(z) z^{N-1}. \end{aligned} \tag{4.42}$$

Two explicit examples for $N = 2$ and $q = 2, 3$ are described as follows. From (4.18)

$$\begin{aligned} y_{q=2}(t) = [x(t)]^2 = &\frac{1}{2}\overbrace{\cos[(\omega_0 + \omega_0)t]}^{d_0 = a_0 + b_0,\ \ g_0 = 1} + \frac{1}{2}\overbrace{\cos[(\omega_0 - \omega_0)t]}^{d_1 = a_0 - b_0,\ \ g_1 = 1} + \frac{1}{2} \times \overbrace{2\cos[(\omega_0 + \omega_1)t]}^{d_2 = a_0 + b_1,\ \ g_2 = 2} \\ &+ \frac{1}{2} \times \overbrace{2\cos[(\omega_0 - \omega_1)t]}^{d_3 = a_0 - b_1,\ \ g_3 = 2} + \frac{1}{2}\overbrace{\cos[(\omega_1 + \omega_1)t]}^{d_4 = a_1 + b_1,\ \ g_4 = 1} + \frac{1}{2}\overbrace{\cos[(\omega_1 - \omega_1)t]}^{d_5 = a_1 - b_1,\ \ g_5 = 1}. \end{aligned} \tag{4.43}$$

By applying the distributive rule,

$$y_{q=3}(t) = [x(t)]^3 = y_{q=2}(t)[x(t)]^2$$

$$= \left(\frac{1}{2}\right)^2 \overbrace{\cos[(\omega_0 + \omega_0 + \omega_0)t]}^{d_0 = a_0 + b_0 + b_0,\ \ g_0 = 1} + \left(\frac{1}{2}\right)^2 \overbrace{\cos[(\omega_0 + \omega_0 - \omega_0)t]}^{d_1 = a_0 + b_0 - b_0,\ \ g_1 = 1} + \left(\frac{1}{2}\right)^2 \overbrace{\cos[(\omega_0 - \omega_0 + \omega_0)t]}^{d_2 = a_0 - b_0 + b_0,\ \ g_2 = 1}$$

$$+ \left(\frac{1}{2}\right)^2 \overbrace{\cos[(\omega_0 - \omega_0 - \omega_0)t]}^{d_3 = a_0 - b_0 - b_0,\ \ g_3 = 1} + \left(\frac{1}{2}\right)^2 2\overbrace{\cos[(\omega_0 + \omega_1 + \omega_0)t]}^{d_4 = a_0 + b_1 + b_0,\ \ g_4 = 2} + \left(\frac{1}{2}\right)^2 2\overbrace{\cos[(\omega_0 + \omega_1 - \omega_0)t]}^{d_5 = a_0 + b_1 - b_0,\ \ g_5 = 2}$$

$$+ \left(\frac{1}{2}\right)^2 2\overbrace{\cos[(\omega_0 - \omega_1 + \omega_0)t]}^{d_6 = a_0 - b_1 + b_0,\ \ g_6 = 2} + \left(\frac{1}{2}\right)^2 2\overbrace{\cos[(\omega_0 - \omega_1 - \omega_0)t]}^{d_7 = a_0 - b_1 - b_0,\ \ g_7 = 2} + \left(\frac{1}{2}\right)^2 \overbrace{\cos[(\omega_1 + \omega_1 + \omega_0)t]}^{d_8 = a_1 + b_1 + b_0,\ \ g_8 = 1}$$

$$+ \left(\frac{1}{2}\right)^2 \overbrace{\cos[(\omega_1 + \omega_1 - \omega_0)t]}^{d_9 = a_1 + b_1 - b_0,\ \ g_9 = 1} + \left(\frac{1}{2}\right)^2 \overbrace{\cos[(\omega_1 - \omega_1 + \omega_0)t]}^{d_{10} = a_1 - b_1 + b_0,\ \ g_{10} = 1} + \left(\frac{1}{2}\right)^2 \overbrace{\cos[(\omega_1 - \omega_1 - \omega_0)t]}^{d_{11} = a_1 - b_1 - b_0,\ \ g_{11} = 1}$$

$$+\left(\frac{1}{2}\right)^2 \overbrace{\cos[(\omega_0+\omega_0+\omega_1)t]}^{d_{12}=a_0+b_0+b_1,\quad g_{12}=1} + \left(\frac{1}{2}\right)^2 \overbrace{\cos[(\omega_0+\omega_0-\omega_1)t]}^{d_{13}=a_0+b_0-b_1,\quad g_{13}=1} + \left(\frac{1}{2}\right)^2 \overbrace{\cos[(\omega_0-\omega_0+\omega_1)t]}^{d_{14}=a_0-b_0+b_1,\quad g_{14}=1}$$

$$+\left(\frac{1}{2}\right)^2 \overbrace{\cos[(\omega_0-\omega_0-\omega_1)t]}^{d_{15}=a_0-b_0-b_1,\quad g_{15}=1} + \left(\frac{1}{2}\right)^2 \overbrace{2\cos[(\omega_0+\omega_1+\omega_1)t]}^{d_{16}=a_0+b_1+b_1,\quad g_{16}=2} + \left(\frac{1}{2}\right)^2 \overbrace{2\cos[(\omega_0+\omega_1-\omega_1)t]}^{d_{17}=a_0+b_1-b_1,\quad g_{17}=2}$$

$$+\left(\frac{1}{2}\right)^2 \overbrace{2\cos[(\omega_0-\omega_1+\omega_1)t]}^{d_{18}=a_0-b_1+b_1,\quad g_{18}=2} + \left(\frac{1}{2}\right)^2 \overbrace{2\cos[(\omega_0-\omega_1-\omega_1)t]}^{d_{19}=a_0-b_1-b_1,\quad g_{19}=2} + \left(\frac{1}{2}\right)^2 \overbrace{\cos[(\omega_1+\omega_1+\omega_1)t]}^{d_{20}=a_1+b_1+b_1,\quad g_{20}=1}$$

$$+\left(\frac{1}{2}\right)^2 \overbrace{\cos[(\omega_1+\omega_1-\omega_1)t]}^{d_{21}=a_1+b_1-b_1,\quad g_{21}=1} + \left(\frac{1}{2}\right)^2 \overbrace{\cos[(\omega_1-\omega_1+\omega_1)t]}^{d_{22}=a_1-b_1+b_1,\quad g_{22}=1} + \left(\frac{1}{2}\right)^2 \overbrace{\cos[(\omega_1-\omega_1-\omega_1)t]}^{d_{23}=a_1-b_1-b_1,\quad g_{23}=1}.$$

(4.44)

To show the pattern more clearly, polynomial coefficients of $D(z)$, the *gain mulitpliers* vector **g**, the respective amplitudes and frequencies of components [B_n and $\tilde{\omega}_n$ in (4.3)] are listed in Table 4.4 for (4.43) and (4.44).

Table 4.4: Symbolic representations for the cases of $q = 2,3$ and $N = 2$.

$q=2$				$q=3$			
$D(z)$	**g**	B_n	$\tilde{\omega}_n$	$D(z)$	**g**	B_n	$\tilde{\omega}_n$
$d_0 = a_0 + b_0$	1	$B_0 = 1/2 \times 1$	$\tilde{\omega}_0 = \omega_0 + \omega_0$	$d_0 = a_0 + b_0 + b_0$	1	$B_0 = (1/2)^2 \times 1$	$\tilde{\omega}_0 = \omega_0 + \omega_0 + \omega_0$
$d_1 = a_0 - b_0$	1	$B_1 = 1/2 \times 1$	$\tilde{\omega}_1 = \omega_0 - \omega_0$	$d_1 = a_0 + b_0 - b_0$	1	$B_1 = (1/2)^2 \times 1$	$\tilde{\omega}_1 = \omega_0 + \omega_0 - \omega_0$
$d_2 = a_0 + b_1$	2	$B_3 = 1/2 \times 2$	$\tilde{\omega}_2 = \omega_0 + \omega_1$	$d_2 = a_0 - b_0 + b_0$	1	$B_2 = (1/2)^2 \times 1$	$\tilde{\omega}_2 = \omega_0 - \omega_0 + \omega_0$
$d_3 = a_0 - b_1$	2	$B_4 = 1/2 \times 2$	$\tilde{\omega}_3 = \omega_0 - \omega_1$	$d_3 = a_0 - b_0 - b_0$	1	$B_3 = (1/2)^2 \times 1$	$\tilde{\omega}_3 = \omega_0 - \omega_0 - \omega_0$
$d_4 = a_1 + b_1$	1	$B_5 = 1/2 \times 1$	$\tilde{\omega}_4 = \omega_1 + \omega_1$	$d_4 = a_0 + b_1 + b_0$	2	$B_4 = (1/2)^2 \times 2$	$\tilde{\omega}_4 = \omega_0 + \omega_1 + \omega_0$
$d_5 = a_1 - b_1$	1	$B_6 = 1/2 \times 1$	$\tilde{\omega}_5 = \omega_1 - \omega_1$	$d_5 = a_0 + b_1 - b_0$	2	$B_5 = (1/2)^2 \times 2$	$\tilde{\omega}_5 = \omega_0 + \omega_1 - \omega_0$
				$d_6 = a_0 - b_1 + b_0$	2	$B_6 = (1/2)^2 \times 2$	$\tilde{\omega}_6 = \omega_0 - \omega_1 + \omega_0$
				$d_7 = a_0 - b_1 - b_0$	2	$B_7 = (1/2)^2 \times 2$	$\tilde{\omega}_7 = \omega_0 - \omega_1 - \omega_0$
				$d_8 = a_1 + b_1 + b_0$	1	$B_8 = (1/2)^2 \times 1$	$\tilde{\omega}_8 = \omega_1 + \omega_1 + \omega_0$
				$d_9 = a_1 + b_1 - b_0$	1	$B_9 = (1/2)^2 \times 1$	$\tilde{\omega}_9 = \omega_1 + \omega_1 - \omega_0$
				$d_{10} = a_1 - b_1 + b_0$	1	$B_{10} = (1/2)^2 \times 1$	$\tilde{\omega}_{10} = \omega_1 - \omega_1 + \omega_0$
				$d_{11} = a_1 - b_1 - b_0$	1	$B_{11} = (1/2)^2 \times 1$	$\tilde{\omega}_{11} = \omega_1 - \omega_1 - \omega_0$
				$d_{12} = a_0 + b_0 + b_1$	1	$B_{12} = (1/2)^2 \times 1$	$\tilde{\omega}_{12} = \omega_0 + \omega_0 + \omega_1$
				$d_{13} = a_0 + b_0 - b_1$	1	$B_{13} = (1/2)^2 \times 1$	$\tilde{\omega}_{13} = \omega_0 + \omega_0 - \omega_1$
				$d_{14} = a_0 - b_0 + b_1$	1	$B_{14} = (1/2)^2 \times 1$	$\tilde{\omega}_{14} = \omega_0 - \omega_0 + \omega_1$
				$d_{15} = a_0 - b_0 - b_1$	1	$B_{15} = (1/2)^2 \times 1$	$\tilde{\omega}_{15} = \omega_0 - \omega_0 - \omega_1$
				$d_{16} = a_0 + b_1 + b_1$	2	$B_{16} = (1/2)^2 \times 2$	$\tilde{\omega}_{16} = \omega_0 + \omega_1 + \omega_1$
				$d_{17} = a_0 + b_1 - b_1$	2	$B_{17} = (1/2)^2 \times 2$	$\tilde{\omega}_{17} = \omega_0 + \omega_1 - \omega_1$
				$d_{18} = a_0 - b_1 + b_1$	2	$B_{18} = (1/2)^2 \times 2$	$\tilde{\omega}_{18} = \omega_0 - \omega_1 + \omega_1$
				$d_{19} = a_0 - b_1 - b_1$	2	$B_{19} = (1/2)^2 \times 2$	$\tilde{\omega}_{19} = \omega_0 - \omega_1 - \omega_1$
				$d_{20} = a_1 + b_1 + b_1$	1	$B_{20} = (1/2)^2 \times 1$	$\tilde{\omega}_{20} = \omega_1 + \omega_1 + \omega_1$
				$d_{21} = a_1 + b_1 - b_1$	1	$B_{21} = (1/2)^2 \times 1$	$\tilde{\omega}_{21} = \omega_1 + \omega_1 - \omega_1$
				$d_{22} = a_1 - b_1 + b_1$	1	$B_{22} = (1/2)^2 \times 1$	$\tilde{\omega}_{22} = \omega_1 - \omega_1 + \omega_1$
				$d_{23} = a_1 - b_1 - b_1$	1	$B_{23} = (1/2)^2 \times 1$	$\tilde{\omega}_{23} = \omega_1 - \omega_1 - \omega_1$

The pattern in Table 4.4 repeats for any positive integer number $q > 2$ and $N \geq 2$. When the splitting rule in (4.26) is applied, the total number of repeating terms in **d** for $q > 2$ is

$$\Phi(N)_{q>2} = 2N\Phi(N)_{q-1}. \tag{4.45}$$

By (4.35) and recursion,

$$\Phi(N)_{q>2} = \left(2N\right)^{q-2} \Phi(N)_{q=2}. \tag{4.46}$$

Hence, the number of components generated in **d** for $q > 2$ after removing the repeating terms is

$$L_{\mathbf{d}_{q>2}}^{\text{MPIPCT}} = L_{\mathbf{d}_{q>2}}^{\text{PIPCT}} - \Phi(N)_{q>2}. \tag{4.47}$$

By (4.24), (4.34), (4.35), (4.36), (4.46), and (4.47), the total number of generated components in **d** is

$$\begin{aligned} L_{\mathbf{d}_q}^{\text{MPIPCT}} &= \begin{cases} 2\left[N^2 - \Psi(N)_{q=2}\right], & \text{for } q = 2, \\ L_{\mathbf{d}_q}^{\text{PIPCT}}\left[1 - \dfrac{\Psi(N)_{q=2}}{N^2}\right], & \text{for } q > 2, \end{cases} \\ &= \begin{cases} \left(N+1\right)N, & \text{for } q = 2, \\ \left(N+1\right)L_{\mathbf{d}_{q-1}}^{\text{PIPCT}}, & \text{for } q > 2. \end{cases} \end{aligned} \tag{4.48}$$

Notice that (4.24) and (4.48) are analytical formulae describing the number of resultant components in **d** for the monomial nonlinearity in (4.1) by PIPCT and MPIPCT. For visual comparison, the number of components in **d** for both techniques is plotted in Figure 4.2 for $q \in \{2,3,4,5\}$ and Figure 4.3 for $q \in \{6,7,8,9\}$ and for $N = 2,3,\ldots,20$. Although the number of generated frequency components in **d** using MPIPCT is always fewer than those generated by PIPCT, the final result of intermodulation products' frequencies, and amplitudes are exactly the same.

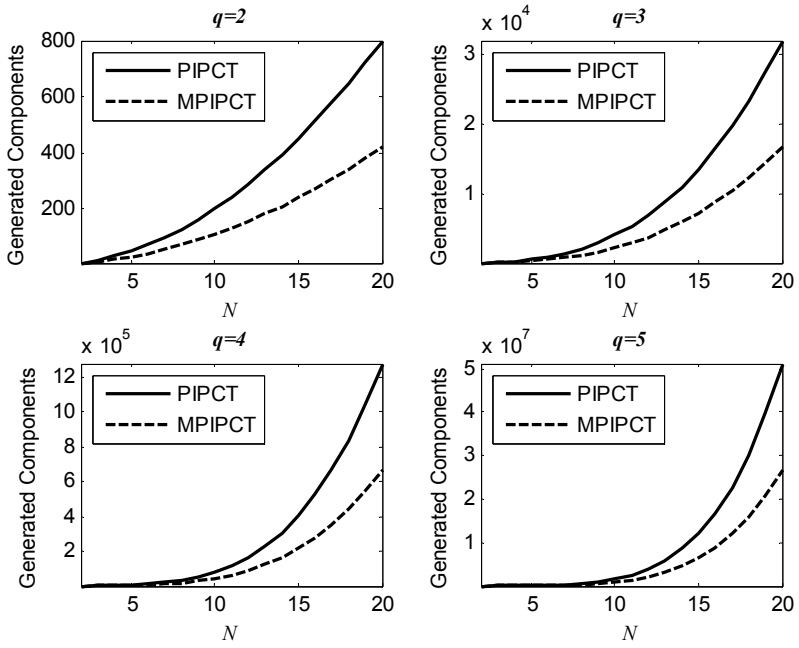

Figure 4.2: Visual comparison of PIPCT and MPIPCT for $q \in \{2,3,4,5\}$ and $N = 2,3,\ldots,20$.

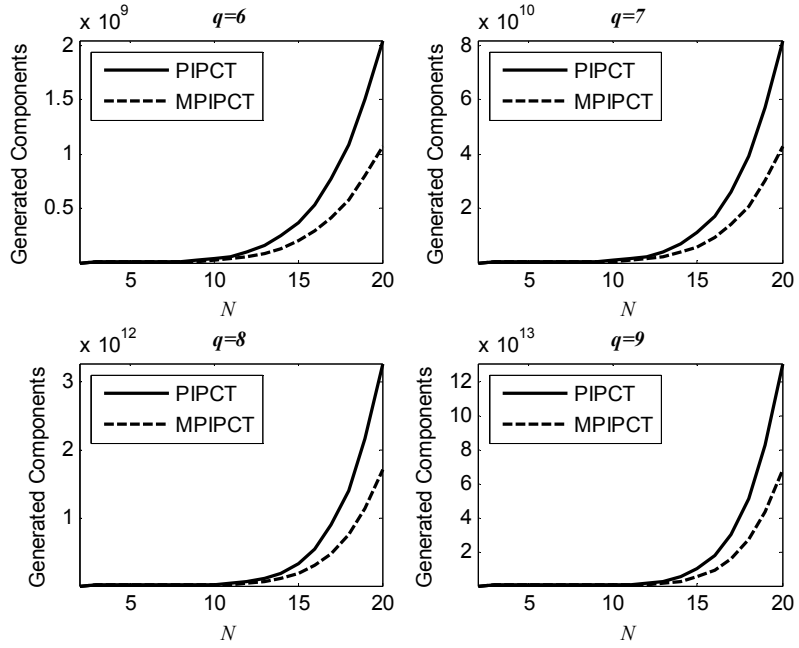

Figure 4.3: Visual comparison of PIPCT and MPIPCT for $q \in \{6,7,8,9\}$ and $N = 2,3,\ldots,20$.

By incorporating *gain multipliers*, g_k, the Lemma 2 can be modified as

$$\text{Amplitudes of Ouput Components} = B_k = \left(\frac{1}{2}\right)^{q-1} g_k, \quad k = 0, 2, ..., L_{\mathbf{d}_q}^{\text{MPIPCT}} - 1, \tag{4.49}$$

where $L_{\mathbf{d}_q}^{\text{MPIPCT}}$, which is formulated in (4.48), is the total number of generated components for the qth order monomial using the proposed technique, MPIPCT.

4.4.2 Upper and lower bounds for computational cost savings

This section presents the derivations of the theoretical upper and lower bounds on how much we can save the number of component growth in **d**, while retaining the same computational results of PIPCT and reducing the computational complexity.

To quantify the number of components we can save in **d**, a metric termed *% Saving*, which direct affects computational time and complexity of the algorithm development, is defined as follows:

$$\%\ \text{Saving} = \frac{L_{\mathbf{d}_q}^{\text{PIPCT}} - L_{\mathbf{d}_q}^{\text{MPIPCT}}}{L_{\mathbf{d}_q}^{\text{PIPCT}}} \times 100\%, \tag{4.50}$$

where $L_{\mathbf{d}_q}^{\text{PIPCT}}$ and $L_{\mathbf{d}_q}^{\text{MPIPCT}}$ are the number of components in **d** of PIPCT and MPIPCT, respectively, for any positive integer number of monomial nonlinearity, $q \geq 2$.

Theorem 4.1: By applying the triangular symmetric pattern, and *gain multipliers*, the frequency component growth in **d** can be saved from 25% for $N = 2$, which is the lower bound, to 50% for $N \to +\infty$, which is the theoretical upper bound, regardless of the monomial nonlinearity order, q.

Proof:

For the lower bound (i.e., two input tones, $N = 2$),

the second-order monomial, $q = 2$: by (4.24) and (4.35),

$$\begin{aligned}\left(\%\ \text{Saving}\right)_{q=2} &= \frac{\Phi(N)_{q=2}}{L^{\text{PIPCT}}_{\mathbf{d}_{q=2}}} \times 100\% \\ &= \frac{N(N-1)}{(2N)^2/2} \times 100\% \\ &= \left(\frac{1}{2} - \frac{1}{2N}\right) \times 100\% = 25\%.\end{aligned} \tag{4.51}$$

orders larger than the second order, $q > 2$: by (4.24), (4.35), and (4.46)

$$\begin{aligned}(\%\ \text{Saving})_{q>2} &= \frac{\Phi(N)_{q>2}}{L^{\text{PIPCT}}_{\mathbf{d}_{q>2}}} \times 100\% \\ &= \frac{(2N)^{q-2}\Phi(N)_{q=2}}{(2N)^q/2} \times 100\% \\ &= \left(\frac{1}{2} - \frac{1}{2N}\right) \times 100\% = 25\%.\end{aligned} \tag{4.52}$$

For the theoretical upper bound $(N \to +\infty)$

$$\lim_{N\to+\infty}\left(\%\ \text{Saving}\right)_{q\geq 2} = \lim_{N\to+\infty}\left\{\left(\frac{1}{2} - \frac{1}{2N}\right) \times 100\%\right\} = 50\%. \tag{4.53}$$

Hence, component growth in **d** can be reduced from 25% as the lower bound to a theoretical maximum of 50% as the upper bound, by exploiting the triangular symmetric pattern of polynomials for q = 2, and by using *gain multiplier*, $g_k, k = 0,1,\ldots, L^{\text{MPIPCT}}_{\mathbf{d}_q} - 1,$ to record the number of repeating terms. Since component

growth in **d** is reduced, computational speed becomes faster without any compromises.

4.4.3 Solutions for the two limitations of PIPCT

To solve the two limitations of PIPCT described in Section 4.1—the need to use constant input amplitudes, and the absence of input and output phase information—we start by observing the following patterns for $N = 2$ and $q = 2$,

$$\begin{aligned} y_{q=2}(t) &= \left[\sum_{k=0}^{1} A_k \cos(\omega_k t + \varphi_k)\right]^2 \\ &= B_0 \cos(\tilde{\omega}_0 t + \psi_0) + \ldots + B_5(\tilde{\omega}_5 t + \psi_5) \\ &= \sum_{n=0}^{5} B_n \cos(\tilde{\omega}_n t + \psi_n), \end{aligned} \tag{4.54}$$

and for $N = 2$ and $q = 3$,

$$\begin{aligned} y_{q=3}(t) &= \left[\sum_{k=0}^{1} A_k \cos(\omega_k t + \varphi_k)\right]^3 \\ &= B_0 \cos(\tilde{\omega}_0 t + \psi_0) + \ldots + B_{23}(\tilde{\omega}_{23} t + \psi_{23}) \\ &= \sum_{n=0}^{23} B_n \cos(\tilde{\omega}_n t + \psi_n), \end{aligned} \tag{4.55}$$

by incorporating the amplitude and phase information (A_k and φ_k) of input multitone signal in (4.2). B_n, $\tilde{\omega}_n$, and ψ_n of (4.54) and (4.55) are listed in Table 4.5 and 4.6 to reveal the emerged patterns vividly.

Table 4.5: Amplitudes, frequencies, and phases computation with *gain multiplier* **g** for the case of $N = 2$ and $q = 2$.

$D(z)$	**g**	Amplitude B_n	Frequency $\tilde{\omega}_n$	Phase ψ_n
$d_0 = a_0 + b_0$	1	$B_0 = 1/2 \times 1 \times A_0 \times A_0$	$\tilde{\omega}_0 = \omega_0 + \omega_0$	$\psi_0 = \varphi_0 + \varphi_0$
$d_1 = a_0 - b_0$	1	$B_1 = 1/2 \times 1 \times A_0 \times A_0$	$\tilde{\omega}_1 = \omega_0 - \omega_0$	$\psi_1 = \varphi_0 - \varphi_0$
$d_2 = a_0 + b_1$	2	$B_2 = 1/2 \times 2 \times A_0 \times A_1$	$\tilde{\omega}_2 = \omega_0 + \omega_1$	$\psi_2 = \varphi_0 + \varphi_1$
$d_3 = a_0 - b_1$	2	$B_3 = 1/2 \times 2 \times A_0 \times A_1$	$\tilde{\omega}_3 = \omega_0 - \omega_1$	$\psi_3 = \varphi_0 - \varphi_1$
$d_4 = a_1 + b_1$	1	$B_4 = 1/2 \times 1 \times A_1 \times A_1$	$\tilde{\omega}_4 = \omega_1 + \omega_1$	$\psi_4 = \varphi_1 + \varphi_1$
$d_5 = a_1 - b_1$	1	$B_5 = 1/2 \times 1 \times A_1 \times A_1$	$\tilde{\omega}_5 = \omega_1 - \omega_1$	$\psi_5 = \varphi_1 - \varphi_1$

From Tables 4.5 and 4.6, with arbitrary amplitudes of input multitone, (4.49) can be further modified to solve the constant input amplitude limitation of PIPCT as follows:

$$\text{Amplitudes of Ouput Components} = B_k = \left(\frac{1}{2}\right)^{q-1} g_k \prod_n A_n, \quad \text{for } k = 0,1,\ldots, L_{\mathbf{d}_q}^{\text{MPIPCT}} - 1, \tag{4.56}$$

where $\prod_n A_n$ is the product of the respective input amplitudes (having the same index number as that of frequencies ω_n; see Tables 4.5 and 4.6) of the output frequency $\tilde{\omega}_k$, g_k is the *gain multiplier* of the output component, and B_k is the amplitude of the output component, which is indexed by *k*, respectively. As shown in Tables 4.5 and 4.6, the phases of the output components, ψ_k, can be computed exactly in the same way as those of frequencies, $\tilde{\omega}_k$. For example, from Table 4.5: $\tilde{\omega}_1 = \omega_0 - \omega_0$ and $\psi_1 = \varphi_0 - \varphi_0$. From Table 4.6: $\tilde{\omega}_1 = \omega_0 + \omega_0 - \omega_0$ and $\psi_1 = \varphi_0 + \varphi_0 - \varphi_0$.

Table 4.6: Amplitudes, frequencies, and phases computation with gain multiplier **g** for the case of $N = 2$ and $q = 3$.

$D(z)$	**g**	Amplitude B_n	Frequency $\tilde{\omega}_n$	Phase ψ_n
$d_0 = a_0 + b_0 + b_0$	1	$B_0 = (1/2)^2 \times 1 \times A_0 \times A_0 \times A_0$	$\tilde{\omega}_0 = \omega_0 + \omega_0 + \omega_0$	$\psi_0 = \varphi_0 + \varphi_0 + \varphi_0$
$d_1 = a_0 + b_0 - b_0$	1	$B_1 = (1/2)^2 \times 1 \times A_0 \times A_0 \times A_0$	$\tilde{\omega}_1 = \omega_0 + \omega_0 - \omega_0$	$\psi_1 = \varphi_0 + \varphi_0 - \varphi_0$
$d_2 = a_0 - b_0 + b_0$	1	$B_2 = (1/2)^2 \times 1 \times A_0 \times A_0 \times A_0$	$\tilde{\omega}_2 = \omega_0 - \omega_0 + \omega_0$	$\psi_2 = \varphi_0 - \varphi_0 + \varphi_0$
$d_3 = a_0 - b_0 - b_0$	1	$B_3 = (1/2)^2 \times 1 \times A_0 \times A_0 \times A_0$	$\tilde{\omega}_3 = \omega_0 - \omega_0 - \omega_0$	$\psi_3 = \varphi_0 - \varphi_0 - \varphi_0$
$d_4 = a_0 + b_1 + b_0$	2	$B_4 = (1/2)^2 \times 2 \times A_0 \times A_1 \times A_0$	$\tilde{\omega}_4 = \omega_0 + \omega_1 + \omega_0$	$\psi_4 = \varphi_0 + \varphi_1 + \varphi_0$
$d_5 = a_0 + b_1 - b_0$	2	$B_5 = (1/2)^2 \times 2 \times A_0 \times A_1 \times A_0$	$\tilde{\omega}_5 = \omega_0 + \omega_1 - \omega_0$	$\psi_5 = \varphi_0 + \varphi_1 - \varphi_0$
$d_6 = a_0 - b_1 + b_0$	2	$B_6 = (1/2)^2 \times 2 \times A_0 \times A_1 \times A_0$	$\tilde{\omega}_6 = \omega_0 - \omega_1 + \omega_0$	$\psi_6 = \varphi_0 - \varphi_1 + \varphi_0$
$d_7 = a_0 - b_1 - b_0$	2	$B_7 = (1/2)^2 \times 2 \times A_0 \times A_1 \times A_0$	$\tilde{\omega}_7 = \omega_0 - \omega_1 - \omega_0$	$\psi_7 = \varphi_0 - \varphi_1 - \varphi_0$
$d_8 = a_1 + b_1 + b_0$	1	$B_8 = (1/2)^2 \times 1 \times A_1 \times A_1 \times A_0$	$\tilde{\omega}_8 = \omega_1 + \omega_1 + \omega_0$	$\psi_8 = \varphi_1 + \varphi_1 + \varphi_0$
$d_9 = a_1 + b_1 - b_0$	1	$B_9 = (1/2)^2 \times 1 \times A_1 \times A_1 \times A_0$	$\tilde{\omega}_9 = \omega_1 + \omega_1 - \omega_0$	$\psi_9 = \varphi_1 + \varphi_1 - \varphi_0$
$d_{10} = a_1 - b_1 + b_0$	1	$B_{10} = (1/2)^2 \times 1 \times A_1 \times A_1 \times A_0$	$\tilde{\omega}_{10} = \omega_1 - \omega_1 + \omega_0$	$\psi_{10} = \varphi_1 - \varphi_1 + \varphi_0$
$d_{11} = a_1 - b_1 - b_0$	1	$B_{11} = (1/2)^2 \times 1 \times A_1 \times A_1 \times A_0$	$\tilde{\omega}_{11} = \omega_1 - \omega_1 - \omega_0$	$\psi_{11} = \varphi_1 - \varphi_1 - \varphi_0$
$d_{12} = a_0 + b_0 + b_1$	1	$B_{12} = (1/2)^2 \times 1 \times A_0 \times A_0 \times A_1$	$\tilde{\omega}_{12} = \omega_0 + \omega_0 + \omega_1$	$\psi_{12} = \varphi_0 + \varphi_0 + \varphi_1$
$d_{13} = a_0 + b_0 - b_1$	1	$B_{13} = (1/2)^2 \times 1 \times A_0 \times A_0 \times A_1$	$\tilde{\omega}_{13} = \omega_0 + \omega_0 - \omega_1$	$\psi_{13} = \varphi_0 + \varphi_0 - \varphi_1$
$d_{14} = a_0 - b_0 + b_1$	1	$B_{14} = (1/2)^2 \times 1 \times A_0 \times A_0 \times A_1$	$\tilde{\omega}_{14} = \omega_0 - \omega_0 + \omega_1$	$\psi_{14} = \varphi_0 - \varphi_0 + \varphi_1$
$d_{15} = a_0 - b_0 - b_1$	1	$B_{15} = (1/2)^2 \times 1 \times A_0 \times A_0 \times A_1$	$\tilde{\omega}_{15} = \omega_0 - \omega_0 - \omega_1$	$\psi_{15} = \varphi_0 - \varphi_0 - \varphi_1$
$d_{16} = a_0 + b_1 + b_1$	2	$B_{16} = (1/2)^2 \times 2 \times A_0 \times A_1 \times A_1$	$\tilde{\omega}_{16} = \omega_0 + \omega_1 + \omega_1$	$\psi_{16} = \varphi_0 + \varphi_1 + \varphi_1$
$d_{17} = a_0 + b_1 - b_1$	2	$B_{17} = (1/2)^2 \times 2 \times A_0 \times A_1 \times A_1$	$\tilde{\omega}_{17} = \omega_0 + \omega_1 - \omega_1$	$\psi_{17} = \varphi_0 + \varphi_1 - \varphi_1$
$d_{18} = a_0 - b_1 + b_1$	2	$B_{18} = (1/2)^2 \times 2 \times A_0 \times A_1 \times A_1$	$\tilde{\omega}_{18} = \omega_0 - \omega_1 + \omega_1$	$\psi_{18} = \varphi_0 - \varphi_1 + \varphi_1$
$d_{19} = a_0 - b_1 - b_1$	2	$B_{19} = (1/2)^2 \times 2 \times A_0 \times A_1 \times A_1$	$\tilde{\omega}_{19} = \omega_0 - \omega_1 - \omega_1$	$\psi_{19} = \varphi_0 - \varphi_1 - \varphi_1$
$d_{20} = a_1 + b_1 + b_1$	1	$B_{20} = (1/2)^2 \times 1 \times A_1 \times A_1 \times A_1$	$\tilde{\omega}_{20} = \omega_1 + \omega_1 + \omega_1$	$\psi_{20} = \varphi_1 + \varphi_1 + \varphi_1$
$d_{21} = a_1 + b_1 - b_1$	1	$B_{21} = (1/2)^2 \times 1 \times A_1 \times A_1 \times A_1$	$\tilde{\omega}_{21} = \omega_1 + \omega_1 - \omega_1$	$\psi_{21} = \varphi_1 + \varphi_1 - \varphi_1$
$d_{22} = a_1 - b_1 + b_1$	1	$B_{22} = (1/2)^2 \times 1 \times A_1 \times A_1 \times A_1$	$\tilde{\omega}_{22} = \omega_1 - \omega_1 + \omega_1$	$\psi_{22} = \varphi_1 - \varphi_1 + \varphi_1$
$d_{23} = a_1 - b_1 - b_1$	1	$B_{23} = (1/2)^2 \times 1 \times A_1 \times A_1 \times A_1$	$\tilde{\omega}_{23} = \omega_1 - \omega_1 - \omega_1$	$\psi_{23} = \varphi_1 - \varphi_1 - \varphi_1$

4.4.4 Modified PIPCT

Based on the derivations in the previous section, we propose the computational steps of MPIPCT. Equation (4.1) defines the monomial nonlinearity with input $x(t)$ in (4.2) and output $y_q(t)$ in (4.3). In matrix form, let us define

$$\mathbf{X} = \begin{bmatrix} \mathbf{A} & \boldsymbol{\omega} & \boldsymbol{\varphi} \end{bmatrix} \in \mathbb{R}^{N\times 3}, \tag{4.57}$$

$$\mathbf{Y}_q = \begin{bmatrix} \mathbf{B}_q & \tilde{\boldsymbol{\omega}}_q & \boldsymbol{\psi}_q \end{bmatrix} \in \mathbb{R}^{L_{\mathbf{d}_q}^{\mathrm{MPIPCT}}\times 3}, \tag{4.58}$$

for the input and output matrices for MPIPCT, where

$$\mathbf{A} = \begin{bmatrix} A_0 & A_1 & \dots & A_{N-1} \end{bmatrix}^T \in \mathbb{R}^{N\times 1}, \tag{4.59}$$

$$\boldsymbol{\omega} = \begin{bmatrix} \omega_0 & \omega_1 & \dots & \omega_{N-1} \end{bmatrix}^T \in \mathbb{R}^{N\times 1}, \tag{4.60}$$

$$\boldsymbol{\varphi} = \begin{bmatrix} \varphi_0 & \varphi_1 & \dots & \varphi_{N-1} \end{bmatrix}^T \in \mathbb{R}^{N\times 1}, \tag{4.61}$$

$$\mathbf{B}_q = \begin{bmatrix} B_0 & B_1 & \dots & B_{L_{\mathbf{d}_q}^{\mathrm{MPIPCT}}-1} \end{bmatrix}^T \in \mathbb{R}^{L_{\mathbf{d}_q}^{\mathrm{MPIPCT}}\times 1}, \tag{4.62}$$

$$\tilde{\boldsymbol{\omega}}_q = \begin{bmatrix} \tilde{\omega}_0 & \tilde{\omega}_1 & \dots & \tilde{\omega}_{L_{\mathbf{d}_q}^{\mathrm{MPIPCT}}-1} \end{bmatrix}^T \in \mathbb{R}^{L_{\mathbf{d}_q}^{\mathrm{MPIPCT}}\times 1}, \tag{4.63}$$

$$\boldsymbol{\psi}_q = \begin{bmatrix} \psi_0 & \psi_1 & \dots & \psi_{L_{\mathbf{d}_q}^{\mathrm{MPIPCT}}-1} \end{bmatrix}^T \in \mathbb{R}^{L_{\mathbf{d}_q}^{\mathrm{MPIPCT}}\times 1}, \tag{4.64}$$

are amplitude, frequency, and phase vectors of the input and output signal in (4.2) and (4.3), respectively, where $L_{\mathbf{d}_q}^{\mathrm{MPIPCT}}$ is given by (4.48). The MPIPCT algorithm takes in the input matrix $\mathbf{X}$ in (4.57) and the scalar integer number $q \geq 2$, which indicates the order of monomial nonlinearity. In functional form, this can be represented as:

$$\mathbf{Y}_q \leftarrow \mathrm{MPIPCT}(\mathbf{X}, q), \tag{4.65}$$

The simplified pseudo-code for MPIPCT is presented in Table 4.7.

Table 4.7: Pseudocode of MPIPCT algorithm.

function $[\mathbf{Y}_q] = \text{MPIPCT}(\mathbf{X}, q)$

/* frequency */
$\mathbf{a} \leftarrow \boldsymbol{\omega}$
$\mathbf{b} \leftarrow \boldsymbol{\omega}$
$\mathbf{c} \leftarrow \text{conv}(\mathbf{a}, \mathbf{b})$
$\mathbf{d} \leftarrow \text{split}(\mathbf{c})$
$\mathbf{g} \leftarrow \text{compute\_gain}(\mathbf{d})$
$\mathbf{d} \leftarrow \text{reduce}(\mathbf{d})$
$\tilde{\boldsymbol{\omega}}_{q=2} \leftarrow \mathbf{d}$
/* phase */
$\mathbf{a} \leftarrow \boldsymbol{\varphi}$
$\mathbf{b} \leftarrow \boldsymbol{\varphi}$
$\mathbf{c} \leftarrow \text{conv}(\mathbf{a}, \mathbf{b})$
$\mathbf{d} \leftarrow \text{split}(\mathbf{c})$
$\mathbf{d} \leftarrow \text{reduce}(\mathbf{d})$
$\boldsymbol{\psi}_{q=2} \leftarrow \mathbf{d}$
/* amplitude */
$\mathbf{B}_q \leftarrow \prod_n A_n$
for $i \leftarrow 3$ **to** q **step** 1 **do**
/* frequency */
$\mathbf{a} \leftarrow \tilde{\boldsymbol{\omega}}_{q=i-1}$
$\mathbf{b} \leftarrow \boldsymbol{\omega}$
$\mathbf{d} \leftarrow \text{distribute}(\mathbf{a}, \mathbf{b})$
$\tilde{\boldsymbol{\omega}}_{q=i} \leftarrow \mathbf{d}$
/* phase */
$\mathbf{a} \leftarrow \boldsymbol{\psi}_{q=i-1}$
$\mathbf{b} \leftarrow \boldsymbol{\varphi}$
$\mathbf{d} \leftarrow \text{distribute}(\mathbf{a}, \mathbf{b})$
$\boldsymbol{\psi}_{q=i} \leftarrow \mathbf{d}$
/* amplitude */
$\mathbf{B}_{q=i} \leftarrow \prod_n A_n$
$\mathbf{g} \leftarrow \text{update}(\mathbf{g})$
od
/* amplitude */
$\mathbf{B}_q \leftarrow (1/2)^{q-1} \times \mathbf{g} \circ \mathbf{B}_q$ /* element-wise multiplication */
/* frequency and phase computation */
for $i \leftarrow 1$ *to length(* $\mathbf{B}_q$)
if $\tilde{\boldsymbol{\omega}}_q[i] < 0$
$\boldsymbol{\psi}_q[i] \leftarrow -\boldsymbol{\psi}_q[i]; \quad \tilde{\boldsymbol{\omega}}_q[i] \leftarrow -\tilde{\boldsymbol{\omega}}_q[i]$
od
$\mathbf{Y}_q \leftarrow \left[\mathbf{B}_q \quad \tilde{\boldsymbol{\omega}}_q \quad \boldsymbol{\psi}_q\right]$

4.5 Proposed Technique for Polynomial NLD

In the previous section, the proposed technique for monomial nonlinearity for intermodulation distortion analysis was presented. In this section, we extend the proposed technique in the case of polynomial nonlinearity. We first derive the closed-form expression to add components having the same frequencies but arbitrary amplitudes and phases. This expression is termed *Harmonic Addition Theorem* (HAT), which is also known in literature as *Phasor Addition Theorem* [MSY03]. However, the derived expression in this chapter is new and in closed form. Based on this derived expression, a pseudo-code sub-routine is presented. The second half of this section presents the detailed steps to develop the technique used in cases of monomial nonlinearity for polynomial nonlinearity, using HAT.

4.5.1 Harmonic Addition Theorem (HAT)

Theorem 4.2:

$$\sum_{m=1}^{M} B_m \cos(\widehat{\omega} t + \psi_m) = \widehat{B} \cos(\widehat{\omega} t + \widehat{\psi}), \tag{4.66}$$

where

$$\widehat{B} = \sqrt{\sum_{m=1}^{M} B_m^2 + 2\sum_{m=1}^{M-1} \sum_{n=m+1}^{M} B_m B_n \cos(\psi_m - \psi_n)}, \tag{4.67}$$

$$\widehat{\psi} = \operatorname{atan2}\left(\frac{\sum_{m=1}^{M} B_m \sin \psi_m}{\sum_{m=1}^{M} B_m \cos \psi_m} \right), -\pi < \widehat{\psi} \le \pi, \tag{4.68}$$

and B_m and ψ_m are amplitudes and phase of the components having the same frequency $\hat{\omega}$. M is the total number of components to be added. $\hat{B}$ and $\hat{\psi}$ are the resultant amplitude and phase, respectively.

Proof:

Let us define a sub-problem to formulate a generalized formula for the addition of complex numbers in polar form without needing to convert them into complex rectangular form. Let $x_e(t)$ be a complex exponential signal function, given by

$$x_e(t) = \sum_{m=1}^{M} B_m \exp(j\psi_m) = \sum_{m=1}^{M} B_m \angle \psi_m, \tag{4.69}$$

where B_m is the amplitude and ψ_m is the phase of the mth complex number in polar form or in phasor notation. M is the total number of complex numbers to add, the objective is to convert (4.69) into

$$x_e(t) = \hat{B} \exp(j\hat{\psi}) = \hat{B} \angle \hat{\psi}, \tag{4.70}$$

where $\hat{B}$ is the amplitude and $\hat{\psi}$ is the phase. In other words, the problem is to formulate $\hat{B}$ and $\hat{\psi}$ of (4.70) in terms of B_m and ψ_m for $m = 1, 2, \ldots, M$ of (4.69).

To solve this sub-problem, starting from $M = 2$, we note that

$$\begin{aligned} \hat{B} &= \sqrt{(B_1 \cos\psi_1 + B_2 \cos\psi_2)^2 + (B_1 \sin\psi_1 + B_2 \sin\psi_2)^2} \\ &= \sqrt{\gamma_2} = \sqrt{B_1^2 + B_2^2 + 2B_1 B_2 \cos(\psi_1 - \psi_2)},. \end{aligned} \tag{4.71}$$

and

$$\hat{\psi} = \text{atan2}\left(\sum_{m=1}^{2} B_m \sin\psi_m, \sum_{m=1}^{2} B_m \cos\psi_m\right), -\pi < \hat{\psi} \le \pi. \tag{4.72}$$

For the computation of $\hat{\psi}$, *atan2* function [MM89] is used to exactly locate the angle in any of the four quadrants in the complex plane. Notice that the ordinary *atan* function range is, however, $-\pi/2 < \hat{\psi} \le \pi/2$ in contrast to the *atan2* function range of $-\pi < \hat{\psi} \le \pi$. Table 4.8 displays γ_M for $M = 2,3,4,5,6$. The symbolic expansion patterns in Table 4.8 can be generalized as follows:

$$\hat{B}^2 = \gamma_M = \left(\sum_{m=1}^{M} B_m \cos\psi_m \right)^2 + \left(\sum_{m=1}^{M} B_m \sin\psi_m \right)^2 = \xi_M + 2\lambda_M, \tag{4.73}$$

where

$$\xi_M = \sum_{m=1}^{M} B_m^2, \tag{4.74}$$

and

$$\lambda_M = \sum_{m=1}^{M-1} \sum_{n=m+1}^{M} B_m B_n \cos(\psi_m - \psi_n). \tag{4.75}$$

For both cases of positive and negative angles in (4.70), let us define

$$x_{e^-}(t) = \sum_{m=1}^{M} B_m \exp(-j\psi_m) = \hat{B}\exp(-j\hat{\psi}), \tag{4.76}$$

and

$$x_{e^+}(t) = \sum_{m=1}^{M} B_m \exp(j\psi_m) = \hat{B}\exp(j\hat{\psi}). \tag{4.77}$$

Using (4.76) and (4.77), and Euler's formula,

$$\sum_{m=1}^{M} B_m \cos(\hat{\omega}t + \psi_m) = \frac{1}{2}\exp(j\hat{\omega}t) \underbrace{\sum_{m=1}^{M} B_m \exp(j\psi_m)}_{x_{e^+}(t)} + \frac{1}{2}\exp(-j\hat{\omega}t) \underbrace{\sum_{m=1}^{M} B_m \exp(-j\psi_m)}_{x_{e^-}(t)}$$

$$= \frac{1}{2}\exp(j\hat{\omega}t) \underbrace{B_m \exp(j\psi_m)}_{x_{e^+}(t)} + \frac{1}{2}\exp(-j\hat{\omega}t) \underbrace{B_m \exp(-j\psi_m)}_{x_{e^-}(t)}$$

$$= \frac{\widehat{B}}{2}\left\{\exp\left[j(\widehat{\omega}t + \widehat{\psi})\right] + \exp\left[-j(\widehat{\omega}t + \widehat{\psi})\right]\right\}$$

$$= \widehat{B}\cos(\widehat{\omega}t + \widehat{\psi}). \qquad Q.E.D.$$

Table 4.8: Algebraic and trigonometric expansions of γ_M for the cases of $M = 2,3,4,5,6$.

M	γ_M
2	$= \overbrace{B_1^2 + B_2^2}^{\xi_2} + 2\overbrace{B_1 B_2 \cos(\psi_1 - \psi_2)}^{\lambda_2}$
3	$= \overbrace{B_1^2 + B_2^2 + B_3^2}^{\xi_3} + 2[$ $\left.\begin{array}{l} B_1B_2\cos(\psi_1-\psi_2) + B_1B_3\cos(\psi_1-\psi_3) + \\ B_2B_3\cos(\psi_2-\psi_3)] \end{array}\right\}\lambda_3$
4	$= \overbrace{B_1^2 + B_2^2 + B_3^2 + B_4^2}^{\xi_4} + 2[$ $\left.\begin{array}{l} B_1B_2\cos(\psi_1-\psi_2) + B_1B_3\cos(\psi_1-\psi_3) + B_1B_4\cos(\psi_1-\psi_4) + \\ B_2B_3\cos(\psi_2-\psi_3) + B_2B_4\cos(\psi_2-\psi_4) + \\ B_3B_4\cos(\psi_3-\psi_4)] \end{array}\right\}\lambda_4$
5	$= \overbrace{B_1^2 + B_2^2 + B_3^2 + B_4^2 + B_5^2}^{\xi_5} + 2[$ $\left.\begin{array}{l} B_1B_2\cos(\psi_1-\psi_2) + B_1B_3\cos(\psi_1-\psi_3) + B_1B_4\cos(\psi_1-\psi_4) + B_1B_5\cos(\psi_1-\psi_5) + \\ B_2B_3\cos(\psi_2-\psi_3) + B_2B_4\cos(\psi_2-\psi_4) + B_2B_6\cos(\psi_2-\psi_5) + \\ B_3B_4\cos(\psi_3-\psi_4) + B_3B_5\cos(\psi_3-\psi_5) + \\ B_4B_5\cos(\psi_4-\psi_5)] \end{array}\right\}\lambda_5$
6	$= \overbrace{B_1^2 + B_2^2 + B_3^2 + B_4^2 + B_5^2 + B_6^2}^{\xi_6} + 2[$ $\left.\begin{array}{l} B_1B_2\cos(\psi_1-\psi_2) + B_1B_3\cos(\psi_1-\psi_3) + B_1B_4\cos(\psi_1-\psi_4) + B_1B_5\cos(\psi_1-\psi_5) + B_1B_6\cos(\psi_1-\psi_6) \\ B_2B_3\cos(\psi_2-\psi_3) + B_2B_4\cos(\psi_2-\psi_4) + B_2B_5\cos(\psi_2-\psi_5) + B_2B_6\cos(\psi_2-\psi_6) \\ B_3B_4\cos(\psi_3-\psi_4) + B_3B_5\cos(\psi_3-\psi_5) + B_3B_6\cos(\psi_3-\psi_6) \\ B_4B_5\cos(\psi_4-\psi_5) + B_4B_5\cos(\psi_4-\psi_5) \\ B_5B_6\cos(\psi_5-\psi_6)] \end{array}\right\}\lambda_6$

Table 4.9: Pseudocode of HAT algorithm.

function $[\hat{B},\hat{\psi}]=$ HAT(B,ψ)

Input: Two arrays B and ψ of length M containing the amplitudes and phases of components in (4.66), respectively, such that $B[1]=B_1,B[2]=B_2,\ldots,B[M]=B_M$; $\psi[1]=\psi_1,\psi[2]=\psi_2,\ldots,\psi[M]=\psi_M$.

Output: The resultant amplitude and phase, $\hat{B}$ and $\hat{\psi}$ in (4.66).

$M \leftarrow$ length(B);
$sum1 \leftarrow 0$; $sum2 \leftarrow 0$; $sum3 \leftarrow 0$; $sum4 \leftarrow 0$;
for $m \leftarrow 1$ **to** M **step** 1 **do**
  $sum1 \leftarrow sum1 + B[m] \times B[m]$;
od;
for $m \leftarrow 1$ **to** $M-1$ **step** 1 **do**
  for $n \leftarrow m+1$ **to** M **step** 1 **do**
    $sum2 \leftarrow sum2 + B[m] \times B[n] \times \cos(\psi[m]-\psi[n])$;
  od;
od;
$\hat{B} \leftarrow sqrt(sum1 + 2 \times sum2)$;
for $m \leftarrow 1$ **to** M **step** 1 **do**
  $sum3 \leftarrow sum3 + B[m] \times \sin(\psi[i])$;
  $sum4 \leftarrow sum4 + A[i] \times \cos(\psi[i])$;
od;
$\hat{\psi} \leftarrow$ atan2($sum3$,$sum4$);

Alternatively, by using (4.76) and (4.77), and Fourier transformation, denoted by $\overset{\mathbb{F}}{\Leftrightarrow}$,

$$\sum_{m=1}^{M} B_m \cos(\hat{\omega}t+\psi_m) \overset{\mathbb{F}}{\Leftrightarrow} \pi\delta(\omega-\hat{\omega}) \underbrace{\sum_{m=1}^{M} B_m \exp(j\psi_m)}_{x_{e^+}(t)} + \pi\delta(\omega+\omega_0) \underbrace{\sum_{m=1}^{M} B_m \exp(-j\psi_m)}_{x_{e^-}(t)}$$

$$= \pi\delta(\omega-\hat{\omega}) \underbrace{\hat{B}\exp(j\hat{\psi})}_{x_{e^+}(t)} + \pi\delta(\omega+\hat{\omega}) \underbrace{\hat{B}\exp(-j\hat{\psi})}_{x_{e^-}(t)}$$

$$= \hat{B}\left[\pi\delta(\omega-\hat{\omega})\exp(j\hat{\psi}) + \pi\delta(\omega+\hat{\omega})\exp(-j\hat{\psi})\right]$$

$$\overset{\mathbb{F}}{\Leftrightarrow} \hat{B}\cos(\hat{\omega}t+\hat{\psi}). \qquad Q.E.D.$$

We next convert HAT into the algorithm using (4.67) and (4.68). The pseudocode is listed in Table 4.9.

4.5.2 Yeary Oo Gan Algorithm (YOGA)

In this subsection, the detailed algorithmic development of the proposed technique from the case of monomial nonlinearity to polynomial nonlinearity is presented. The polynomial nonlinearity equation is presented in (4.4) and is illustrated in Figure 4.4. The pseudo-code of YOGA is shown in Table 4.10. The block diagram with symbols and signal-flow is presented in Figure 4.4 for the theoretical development.

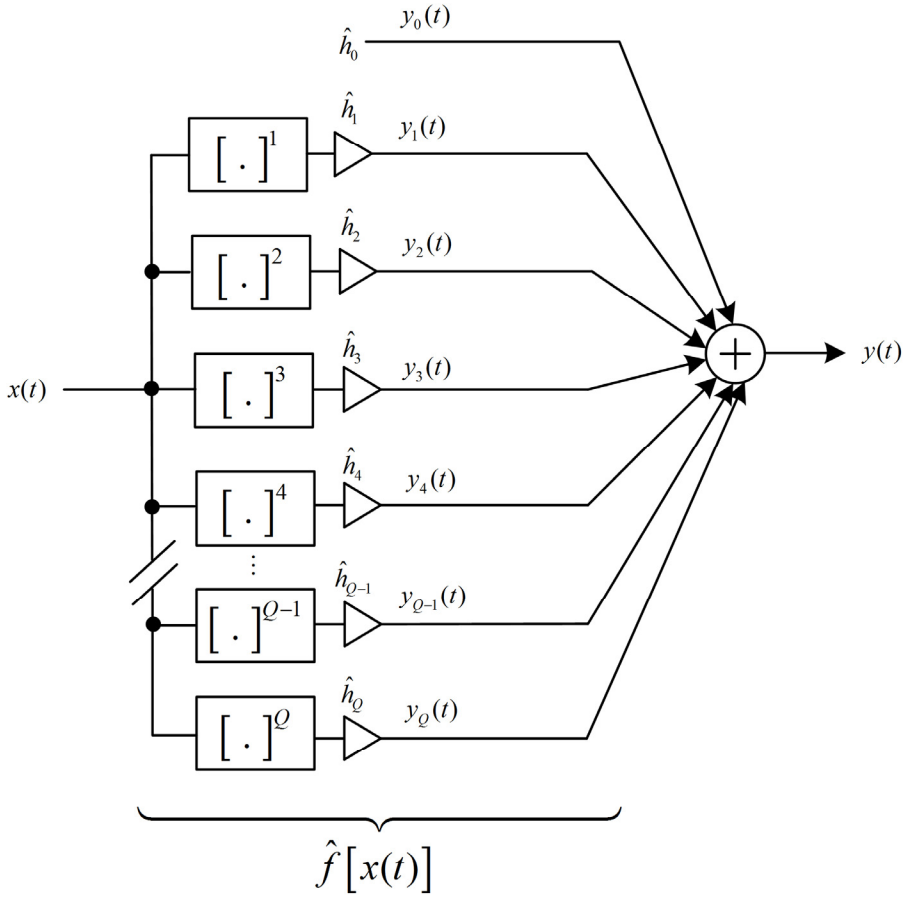

Figure 4.4: Block diagram representation of the polynomial nonlinearity in (4.4).

To develop the technique further, let us denote the output signal function as

$$y(t) = \widehat{DC} + \sum_{k=1}^{P} \hat{B}_k \cos(\hat{\omega}_k t + \hat{\psi}_k), \tag{4.78}$$

where $\widehat{DC}$ is the DC component, and $\hat{B}_k$, $\hat{\omega}_k$, and $\hat{\psi}_k$ are the amplitude, frequency, and phase, respectively, of the *k*th component. From (4.78), the total number of non-overlapping (not having the same frequency) components is $P + 1$. The objective is to obtain $\widehat{DC}$, $\hat{B}_k$, $\hat{\omega}_k$, and $\hat{\psi}_k$, $k = 1, \ldots, P$, given (4.2) and (4.4).

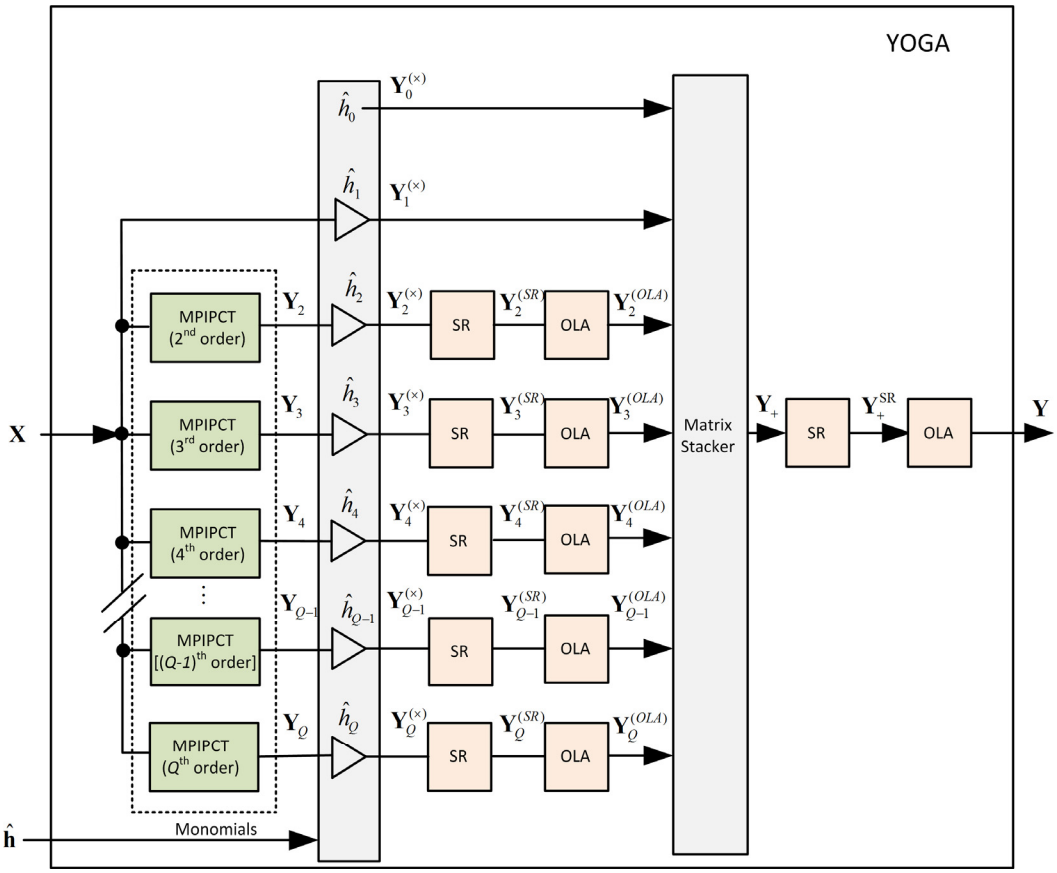

Figure 4.5: Block diagram, mathematical symbols and signal flows of YOGA. SR: *SortRow*; OLA: *OverLapAdd*.

As shown in Figure 4.4 (and algorithmic configuration in Figure 4.5), a polynomial nonlinearity consists of monomial nonlinearities connecting in parallel,

each of which is multiplied by a gain (polynomial coefficient) $\hat{h}_q$, where q is the order of polynomial in this case. For the algorithmic development, referring to Figure 4.5, we define the following matrices:

$$\hat{\mathbf{h}} = \begin{bmatrix} \hat{h}_0 & \hat{h}_1 & \cdots & \hat{h}_{Q-1} & \hat{h}_Q \end{bmatrix}^T \in \mathbb{R}^{(Q+1)\times 1}, \tag{4.79}$$

which contains the polynomial coefficients $\{\hat{h}_0, \hat{h}_1, \ldots, \hat{h}_{Q-1}, \hat{h}_Q\}$.

$$\mathbf{Y}_0^{(\times)} = \begin{bmatrix} \hat{h}_0 & 0 & 0 \end{bmatrix}^T \in \mathbb{R}^{1\times 3}, \text{ for } q = 1, \tag{4.80}$$

$$\mathbf{Y}_1^{(\times)} = \begin{bmatrix} \hat{h}_1\mathbf{A} & \boldsymbol{\omega} & \boldsymbol{\varphi} \end{bmatrix}^T \in \mathbb{R}^{N\times 3}, \text{ for } q = 2, \tag{4.81}$$

and

$$\mathbf{Y}_q^{(\times)} = \begin{bmatrix} \hat{h}_q\mathbf{B}_q & \tilde{\boldsymbol{\omega}}_q & \boldsymbol{\psi}_q \end{bmatrix}^T \in \mathbb{R}^{L_{\mathbf{d}_q}^{\text{MPIPCT}}\times 3}, \text{ for } q = 2, 3, \ldots, Q, \tag{4.82}$$

where $\{\mathbf{A}, \boldsymbol{\omega}, \boldsymbol{\varphi}\}$ are amplitude, frequency, and phase vectors of input multitone as defined in (4.59), (4.60), and (4.61), respectively. $\{\mathbf{B}_q, \tilde{\boldsymbol{\omega}}_q, \boldsymbol{\psi}_q\}$ are amplitude, frequency and phase vectors of output components from the qth order monomial nonlinearity as defined in (4.62), (4.63), and (4.64) respectively. N is the total number of input components and $L_{\mathbf{d}_q}^{\text{MPIPCT}}$ in (4.48) is the total number of output components from the qth order monomial nonlinearity. At each and every branch of the qth order monomial, the two operations *SortRows* (SR) and *OverLapAdd* (OLA) are performed to further reduce the computational growth of components, which in turn leads to saving the computational load. The SR operation can be mathematically described as follows.

$$\mathbf{Y}_q^{(\text{SR})} \xleftarrow{\text{SR}} \text{SortRows}\left(\mathbf{Y}_q^{(\times)}, \tilde{\boldsymbol{\omega}}_q, \downarrow\right), \text{ for } q = 2, 3, \ldots, Q, \tag{4.83}$$

where SortRows(.) sorts the matrix $\mathbf{Y}_q^{(\times)}$ as specified by the column, whereby the symbol (↓) denotes the ascending order. Let us define

$$\mathbf{Y}_q^{(\mathrm{SR})} = \begin{bmatrix} \mathbf{DC}^{(\mathrm{SR})} & \mathbf{0}^{(\mathrm{SR})} & \mathbf{0}^{(\mathrm{SR})} \\ \mathbf{B}_1^{(\mathrm{SR})} & \widehat{\boldsymbol{\omega}}_1^{(\mathrm{SR})} & \boldsymbol{\psi}_1^{(\mathrm{SR})} \\ \vdots & \vdots & \vdots \\ \mathbf{B}_{\Gamma_q}^{(\mathrm{SR})} & \widehat{\boldsymbol{\omega}}_{\Gamma_q}^{(\mathrm{SR})} & \boldsymbol{\psi}_{\Gamma_q}^{(\mathrm{SR})} \end{bmatrix} \in \mathbb{R}^{L_{\mathbf{d}_q}^{\mathrm{MPIPCT}} \times 3}, \text{ for } q = 2,3,\ldots,Q, \tag{4.84}$$

as the output matrix from the SR module at the monomial nonlinearity branches for the second order and above, where

$$\mathbf{DC}^{(\mathrm{SR})} = \begin{bmatrix} DC_1^{(\mathrm{SR})} & \ldots & DC_M^{(\mathrm{SR})} \end{bmatrix}^T, \tag{4.85}$$

represents DC component column vector. Let

$$\mathbf{0}^{(\mathrm{SR})} = \begin{bmatrix} 0_1 & \ldots & 0_M \end{bmatrix}^T \tag{4.86}$$

denote the zero column vector for frequencies and phases of DC components.

$$\mathbf{B}_k^{(\mathrm{SR})} = \begin{bmatrix} B_1^{(\mathrm{SR})} & \ldots & B_M^{(\mathrm{SR})} \end{bmatrix}^T, \text{ for } k = 1,\ldots,\Gamma_q, \tag{4.87}$$

$$\widehat{\boldsymbol{\omega}}_k^{(\mathrm{SR})} = \begin{bmatrix} \widehat{\omega}_k & \ldots & \widehat{\omega}_k \end{bmatrix}^T, \text{ for } k = 1,\ldots,\Gamma_q, \tag{4.88}$$

$$\boldsymbol{\psi}_k^{(\mathrm{SR})} = \begin{bmatrix} \psi_1^{(\mathrm{SR})} & \ldots & \psi_M^{(\mathrm{SR})} \end{bmatrix}^T, \text{ for } k = 1,\ldots,\Gamma_q, \tag{4.89}$$

represent amplitude, frequency, and phase column vectors, whose components are grouped together if their respective frequencies are the same. We denote M for the the total number equal-frequency components in the kth column vector and Γ_q denotes the total number of $\mathbf{B}_k^{(\mathrm{SR})}$ or $\widehat{\boldsymbol{\omega}}_k^{(\mathrm{SR})}$ or $\boldsymbol{\psi}_k^{(\mathrm{SR})}$ in $\mathbf{Y}_q^{(\mathrm{SR})}$.

The OLA operation can be described by

$$\mathbf{Y}_q^{(\mathrm{OLA})} \xleftarrow{\mathrm{OLA}} \mathrm{OverLapAdd}\left(\mathbf{Y}_q^{(\mathrm{SR})}\right), \tag{4.90}$$

where the input matrix $\mathbf{Y}_q^{(\mathrm{SR})}$ is given in (4.84) and the output matrix can be formulated by

$$\mathbf{Y}_q^{(\mathrm{OLA})} = \begin{bmatrix} DC^{(OLA)} \leftarrow \sum_{m=1}^{M} DC_m^{(\mathrm{SR})} & 0 & 0 \\ B_1^{(\mathrm{OLA})} \leftarrow \mathrm{HAT}\left(\mathbf{B}_1^{(\mathrm{SR})}, \boldsymbol{\psi}_1^{(\mathrm{SR})}\right) & \widehat{\omega}_1 & \psi_1^{(\mathrm{OLA})} \leftarrow \mathrm{HAT}\left(\mathbf{B}_1^{(\mathrm{SR})}, \boldsymbol{\psi}_1^{(\mathrm{SR})}\right) \\ \vdots & \vdots & \vdots \\ B_{\Gamma_q}^{(\mathrm{OLA})} \leftarrow \mathrm{HAT}\left(\mathbf{B}_{\Gamma_q}^{(\mathrm{SR})}, \boldsymbol{\psi}_{\Gamma_q}^{(\mathrm{SR})}\right) & \widehat{\omega}_{\Gamma_q} & \psi_{\Gamma_q}^{(\mathrm{OLA})} \leftarrow \mathrm{HAT}\left(\mathbf{B}_{\Gamma_q}^{(\mathrm{SR})}, \boldsymbol{\psi}_{\Gamma_q}^{(\mathrm{SR})}\right) \end{bmatrix},$$

$$= \begin{bmatrix} DC^{(OLA)} & 0 & 0 \\ B_1^{(\mathrm{OLA})} & \widehat{\omega}_1 & \psi_1^{(\mathrm{OLA})} \\ \vdots & \vdots & \vdots \\ B_{\Gamma_q}^{(\mathrm{OLA})} & \widehat{\omega}_{\Gamma_q} & \psi_{\Gamma_q}^{(\mathrm{OLA})} \end{bmatrix} \in \mathbb{R}^{\left(1+\Gamma_q\right)\times 3}$$

$$= \begin{bmatrix} DC^{(\mathrm{OLA})} & 0 & 0 \\ \mathbf{B}_q^{(\mathrm{OLA})} & \widehat{\boldsymbol{\omega}}_q^{(\mathrm{OLA})} & \boldsymbol{\psi}_q^{(\mathrm{OLA})} \end{bmatrix} \in \mathbb{R}^{\left(1+\Gamma_q\right)\times 3}, \text{ for } q = 2,3,\ldots,Q, \tag{4.91}$$

where

$$\mathbf{B}_q^{(\mathrm{OLA})} = \begin{bmatrix} B_1^{(\mathrm{OLA})} & \cdots & B_{\Gamma_q}^{(\mathrm{OLA})} \end{bmatrix}^T \in \mathbb{R}^{\Gamma_q \times 3}, \text{ for } q = 2,3,\ldots,Q, \tag{4.92}$$

$$\widehat{\boldsymbol{\omega}}_q^{(\mathrm{OLA})} = \begin{bmatrix} \widehat{\omega}_1 & \cdots & \widehat{\omega}_{\Gamma_q} \end{bmatrix}^T \in \mathbb{R}^{\Gamma_q \times 3}, \text{ for } q = 2,3,\ldots,Q, \tag{4.93}$$

$$\boldsymbol{\psi}_q^{(\mathrm{OLA})} = \begin{bmatrix} \psi_1^{(\mathrm{OLA})} & \cdots & \psi_{\Gamma_q}^{(\mathrm{OLA})} \end{bmatrix}^T \in \mathbb{R}^{\Gamma_q \times 3}, \text{ for } q = 2,3,\ldots,Q. \tag{4.94}$$

The OLA operation, as its name implies, adds the amplitudes and phases of the overlapping frequency components. HAT algorithm, as presented in Section 4.4.1, is used to add the components having the same frequency but different amplitudes and phases.

All the output matrices from the monomial nonlinearity branches are stacked as follows:

$$\mathbf{Y}_{+} \leftarrow \text{MatrixStacker}\left(\mathbf{Y}_0^{(\times)}, \mathbf{Y}_1^{(\times)}, \left\{\mathbf{Y}_q^{(\text{OLA})}\right\}_{q=2}^{Q}\right), \tag{4.95}$$

where

$$\mathbf{Y}_{+} = \begin{bmatrix} \mathbf{Y}_0^{(\times)} \\ \cdots\cdots \\ \mathbf{Y}_1^{(\times)} \\ \cdots\cdots \\ \mathbf{Y}_2^{(\text{OLA})} \\ \cdots\cdots \\ \vdots \\ \cdots\cdots \\ \mathbf{Y}_Q^{(\text{OLA})} \end{bmatrix} = \begin{bmatrix} \hat{h}_0 & 0 & 0 \\ \hat{h}_1\mathbf{A} & \boldsymbol{\omega} & \boldsymbol{\varphi} \\ \mathbf{B}_{q=2}^{(\text{OLA})} & \widehat{\boldsymbol{\omega}}_{q=2}^{(\text{OLA})} & \boldsymbol{\psi}_{q=2}^{(\text{OLA})} \\ \vdots & \vdots & \vdots \\ \mathbf{B}_{q=Q}^{(\text{OLA})} & \widehat{\boldsymbol{\omega}}_{q=Q}^{(\text{OLA})} & \boldsymbol{\psi}_{q=Q}^{(\text{OLA})} \end{bmatrix} = \begin{bmatrix} \mathbf{B}_{+} & \boldsymbol{\omega}_{+} & \boldsymbol{\psi}_{+} \end{bmatrix} \in \mathbb{R}^{\Gamma_{\text{poly}} \times 3}, \tag{4.96}$$

and

$$\Gamma_{\text{poly}} = 1 + N + \sum\nolimits_{q=2}^{Q} (1 + \Gamma_q), \tag{4.97}$$

$$\mathbf{B}_{+} = \begin{bmatrix} \hat{h}_0 & \hat{h}_1\mathbf{A} & \mathbf{B}_{q=2}^{(\text{OLA})} & \cdots & \mathbf{B}_{q=Q}^{(\text{OLA})} \end{bmatrix}^T, \tag{4.98}$$

$$\boldsymbol{\omega}_{+} = \begin{bmatrix} 0 & \boldsymbol{\omega} & \widehat{\boldsymbol{\omega}}_{q=2}^{(\text{OLA})} & \cdots & \widehat{\boldsymbol{\omega}}_{q=Q}^{(\text{OLA})} \end{bmatrix}^T, \tag{4.99}$$

$$\boldsymbol{\psi}_{+} = \begin{bmatrix} 0 & \boldsymbol{\varphi} & \boldsymbol{\psi}_{q=2}^{(\text{OLA})} & \cdots & \boldsymbol{\psi}_{q=Q}^{(\text{OLA})} \end{bmatrix}^T. \tag{4.100}$$

Referring to Figure 4.5, similar to the monomial branches, the stacked matrix $\mathbf{Y}_{+}$ goes through RS and OLA operations. The followings are mathematical descriptions of these operations.

$$\mathbf{Y}_{+}^{(\text{SR})} \xleftarrow{\text{SR}} \text{sortrows}\left(\mathbf{Y}_{+}, \boldsymbol{\omega}_{+}, \downarrow\right), \text{ for } q = 2,3,\ldots,Q, \tag{4.101}$$

where

$$\mathbf{Y}_{+}^{(\mathrm{SR})}=\begin{bmatrix}\widehat{\mathbf{DC}}_{+}^{(\mathrm{SR})} & \mathbf{0}_{+}^{(\mathrm{SR})} & \mathbf{0}_{+}^{(\mathrm{SR})}\\ \widehat{\mathbf{B}}_{+,1}^{(\mathrm{SR})} & \widehat{\boldsymbol{\omega}}_{+,1}^{(\mathrm{SR})} & \widehat{\boldsymbol{\psi}}_{+,1}^{(\mathrm{SR})}\\ \vdots & \vdots & \vdots\\ \widehat{\mathbf{B}}_{+,P}^{(\mathrm{SR})} & \widehat{\boldsymbol{\omega}}_{+,P}^{(\mathrm{SR})} & \widehat{\boldsymbol{\psi}}_{+,P}^{(\mathrm{SR})}\end{bmatrix}\in\mathbb{R}^{\Gamma_{\mathrm{poly}}\times 3},\ \text{for } q=2,3,\ldots,Q, \tag{4.102}$$

where

$$\widehat{\mathbf{DC}}_{+}^{(\mathrm{SR})}=\begin{bmatrix}\widehat{DC}_{+,1}^{(\mathrm{SR})} & \ldots & \widehat{DC}_{+,M}^{(\mathrm{SR})}\end{bmatrix}^{T}, \tag{4.103}$$

$$\mathbf{0}_{+}^{(\mathrm{SR})}=\begin{bmatrix}0_1 & \ldots & 0_M\end{bmatrix}^{T} \tag{4.104}$$

$$\widehat{\mathbf{B}}_{+,k}^{(\mathrm{SR})}=\begin{bmatrix}\widehat{B}_{+,1}^{(\mathrm{SR})} & \ldots & \widehat{B}_{+,M}^{(\mathrm{SR})}\end{bmatrix}^{T},\ \text{for } k=1,\ldots,P, \tag{4.105}$$

$$\widehat{\boldsymbol{\omega}}_{+,k}^{(\mathrm{SR})}=\begin{bmatrix}\widehat{\omega}_{+,k} & \ldots & \widehat{\omega}_{+,k}\end{bmatrix}^{T},\ \text{for } k=1,\ldots,P, \tag{4.106}$$

$$\widehat{\boldsymbol{\psi}}_{+,k}^{(\mathrm{SR})}=\begin{bmatrix}\widehat{\psi}_{+,1}^{(\mathrm{SR})} & \ldots & \widehat{\psi}_{+,M}^{(\mathrm{SR})}\end{bmatrix}^{T},\ \text{for } k=1,\ldots,P, \tag{4.107}$$

are DC, zero, amplitude, frequency, and phase column vectors, respectively. P is the number of output components excluding DC [See (4.78)]. Notice that components with the same frequency [$\widehat{\omega}_{+,k}$ in (4.106)] has been grouped together. Next, OLA operation is carried out to add the components having the same frequency as follows:

$$\mathbf{Y}\xleftarrow{\text{OLA}}\text{OverLapAdd}\left(\mathbf{Y}_{+}^{(\mathrm{SR})}\right), \tag{4.108}$$

where the final output matrix is

$$\mathbf{Y} = \begin{bmatrix} \widehat{DC} \leftarrow \sum_{m=1}^{M} \widehat{DC}_{+,m}^{(\mathrm{SR})} & 0 & 0 \\ \widehat{B}_1 \leftarrow \mathrm{HAT}\left(\widehat{\mathbf{B}}_{+,1}^{(\mathrm{SR})}, \widehat{\boldsymbol{\psi}}_{+,1}^{(\mathrm{SR})}\right) & \widehat{\omega}_1 & \widehat{\psi}_1 \leftarrow \mathrm{HAT}\left(\widehat{\mathbf{B}}_{+,1}^{(\mathrm{SR})}, \widehat{\boldsymbol{\psi}}_{+,1}^{(\mathrm{SR})}\right) \\ \vdots & \vdots & \vdots \\ \widehat{B}_P \leftarrow \mathrm{HAT}\left(\widehat{\mathbf{B}}_{+,P}^{(\mathrm{SR})}, \widehat{\boldsymbol{\psi}}_{+,P}^{(\mathrm{SR})}\right) & \widehat{\omega}_P & \widehat{\psi}_P \leftarrow \mathrm{HAT}\left(\widehat{\mathbf{B}}_{+,P}^{(\mathrm{SR})}, \widehat{\boldsymbol{\psi}}_{+,P}^{(\mathrm{SR})}\right) \end{bmatrix}$$

$$= \begin{bmatrix} \widehat{DC} & 0 & 0 \\ \widehat{B}_1 & \widehat{\omega}_1 & \widehat{\psi}_1 \\ \vdots & \vdots & \vdots \\ \widehat{B}_P & \widehat{\omega}_P & \widehat{\psi}_P \end{bmatrix} \in \mathbb{R}^{(1+P)\times 3} \tag{4.109}$$

$$= \begin{bmatrix} \widehat{DC} & 0 & 0 \\ \widehat{\mathbf{B}} & \widehat{\boldsymbol{\omega}} & \widehat{\boldsymbol{\psi}} \end{bmatrix} \in \mathbb{R}^{(1+P)\times 3},$$

and

$$\widehat{\mathbf{B}} = \begin{bmatrix} \widehat{B}_1 & \widehat{B}_2 & \dots & \widehat{B}_P \end{bmatrix} \in \mathbb{R}^{P\times 1}, \tag{4.110}$$

$$\widehat{\boldsymbol{\omega}} = \begin{bmatrix} \widehat{\omega}_1 & \widehat{\omega}_2 & \dots & \widehat{\omega}_P \end{bmatrix} \in \mathbb{R}^{P\times 1}, \tag{4.111}$$

$$\widehat{\boldsymbol{\psi}} = \begin{bmatrix} \widehat{\psi}_1 & \widehat{\psi}_2 & \dots & \widehat{\psi}_P \end{bmatrix} \in \mathbb{R}^{P\times 1}, \tag{4.112}$$

are amplitude, frequency, and phase vectors of the output multitone components whose elements $\{\widehat{B}_k, \widehat{\omega}_k, \widehat{\psi}_k\}$, which we want to compute for harmonic and intermodulation analysis, for $k = 1, 2, \dots, P$ from (4.78). Based on these detailed analysis, the pseudocode for YOGA is presented in Table 4.10.

Table 4.10: Pseudocode of YOGA.

function $[\mathbf{Y}] = \text{YOGA}(\mathbf{X}, \hat{\mathbf{h}})$

/* step 1: monomial nonlinearity computations */

$\mathbf{Y}_q = \left[\mathbf{B}_q \quad \tilde{\boldsymbol{\omega}}_q \quad \boldsymbol{\psi}_q\right] \leftarrow \text{MPIPCT}(\mathbf{X}, q)$, for $q = 2,3,\ldots,Q$;

/* step 2: $q = 1$ and $q = 2$ */

$\mathbf{Y}_0^{(\times)} = \left[\hat{h}_0 \quad 0 \quad 0\right]^T \in \mathbb{R}^{1\times 3}$, for $q = 1$;

$\mathbf{Y}_1^{(\times)} = \left[\hat{h}_1\mathbf{A} \quad \boldsymbol{\omega} \quad \boldsymbol{\varphi}\right]^T \in \mathbb{R}^{N\times 3}$, for $q = 2$;

/* step 3: amplitude multiplication with polynomial coefficients */

$\mathbf{Y}_q^{(\times)} = \left[\hat{h}_q\mathbf{B}_q \quad \tilde{\boldsymbol{\omega}}_q \quad \boldsymbol{\psi}_q\right]^T \in \mathbb{R}^{L_q^{(\text{ETAM\_MT})}\times 3}$, for $q = 2,3,\ldots,Q$;

/* step 4: SR operation for monomial nonlinearity branches*/

$\mathbf{Y}_q^{(\text{SR})} \xleftarrow{\text{SR}} \text{SortRows}\left(\mathbf{Y}_q^{(\times)}, \tilde{\boldsymbol{\omega}}_q, \downarrow\right)$, for $q = 2,3,\ldots,Q$;

/* step 5: OLA operation for monomial nonlinearity branches */

$\mathbf{Y}_q^{(\text{OLA})} \xleftarrow{\text{OLA}} \text{OverLapAdd}\left(\mathbf{Y}_q^{(\text{SR})}\right)$, for $q = 2,3,\ldots,Q$;

/* step 6: matrix stacking operation */

$\mathbf{Y}_+ \leftarrow \text{MatrixStacker}\left(\mathbf{Y}_0^{(\times)}, \mathbf{Y}_1^{(\times)}, \left\{\mathbf{Y}_q^{(\text{OLA})}\right\}_{q=2}^{Q}\right)$,

/* step 7: SR operation for polynomial nonlinearity */

$\mathbf{Y}_+^{(\text{SR})} \xleftarrow{\text{SR}} \text{sortrows}\left(\mathbf{Y}_+, \boldsymbol{\omega}_+, \downarrow\right)$, for $q = 2,3,\ldots,Q$;

/* step 8: OLA operation for polynomial nonlinearity */

$\mathbf{Y} \xleftarrow{\text{OLA}} \text{OverLapAdd}\left(\mathbf{Y}_+^{(\text{SR})}\right)$;

4.6 Simulations and Comparisons with FFT Outputs

Fast Fourier transform (FFT) is a fast computational algorithm of DFT. FFT has four fundamental limitations (pitfalls), namely frequency resolution, aliasing, spectral leakage, and the picket fence effect [OSB99] [Sun01]. FFT is a batch-processing technique that processes signals in the digital domain. Signals are assumed by the FFT analyzer to be periodic. The frequency resolution problem is due to $\Delta f = f_s / L$, where Δf is the frequency resolution, f_s is the sampling frequency, and L is the FFT buffer size. A large buffer size is needed to analyze high-frequency signals in fine-

frequency resolution since f_s must be large in order to avoid aliasing problems. Aliasing problems occur when nonlinearity-introduced components or original signal components exceed the Nyquist frequency (folding frequency; $f_s/2$). These components are folded back to the Nyquist interval, and FFT cannot distinguish between the original spectral line and aliased spectral lines. Spectral leakage problems occur when the signal is not captured in integral number of periods of the waveform, and due to non-periodic truncation in time, discontinuity in the sampled signal occurs. This discontinuity results in energy leakage in spectral domain, causing line spectrum to combine with adjacent lines. FFT samples the frequency spectrum and its inability to distinguish between two frequency sample points causes the picket fence effect, i.e., the computation of the spectrum is limited to integer multiples of Δf.

To illustrate these FFT pitfalls in spectral analysis, and advantages of using the proposed technique that does not use FFT, we simulate the square law nonlinear static device, $y(t) = 0.5x^2(t)$, with a multitone input signal having two components (i.e., $N = 2$). We used four cases to compare FFT results with our proposed technique. The numerical data for both input and output obtained from our proposed technique are listed in Table 4.11. A simple FFT with rectangular windowing is used to capture both input and output signals of such four cases and presented in Figure 4.6. MATLAB was used to compute both our technique and FFT. The sampling frequency f_s is 48 kHz. The input signal is $(1-1/fs)$ seconds long from time zero. Therefore, FFT size is 48,000 and the frequency resolution is 1 Hz. The input multitone signal has two components whose amplitudes, frequencies, and phases are

listed in Table 4.11 for the four cases. These input signals' FFT results are plotted in the left column of Figure 4.6. For each case, input multitone is synthesized as the data in Table 4.11, and output signal is then fed into FFT and plotted the results in the right column of Figure 4.6. Note that the input and output spectra are zoomed into the span of 0–100 Hz for easy comparison.

Table 4.11: Amplitudes, frequencies, and phases for the four examples.

Case	Input Multitone Components			Output Multitone Components		
	Amplitude	Frequency (Hz)	Phase (Deg)	Amplitude	Frequency (Hz)	Phase (Deg)
I	1.000	20.000	36.000	0.045000	10.000	9.000
	0.120	30.000	63.000	0.385800	20.000	36.000
				0.090648	30.000	63.000
				0.005400	40.000	90.000
				0.125000	60.000	108.000
				0.045000	70.000	135.000
				0.005400	80.000	162.000
				0.000216	90.000	189.000
II	1.000	20.240	36.000	0.045000	19.180	9.000
	0.120	21.300	63.000	0.385800	20.240	36.000
				0.090648	21.300	63.000
				0.005400	22.360	90.000
				0.125000	60.720	108.000
				0.045000	61.780	135.000
				0.005400	62.840	162.000
				0.000216	63.900	189.000
III	1.000	20.000	36.000	0.385800	20.000	36.000
	0.120	47910.000	63.000	0.125000	60.000	108.000
				0.045000	47870.000	-9.000
				0.090648	47910.000	63.000
				0.045000	47950.000	135.000
				0.005400	95800.000	90.000
				0.005400	95840.000	162.000
				0.000216	143730.000	189.000
IV	1.000	20.000	36.000	0.385800	20.000	36.000
	0.120	90.000	63.000	0.045000	50.000	-9.000
				0.125000	60.000	108.000
				0.090648	90.000	63.000
				0.045000	130.000	135.000
				0.005400	160.000	90.000
				0.005400	200.000	162.000
				0.000216	270.000	189.000

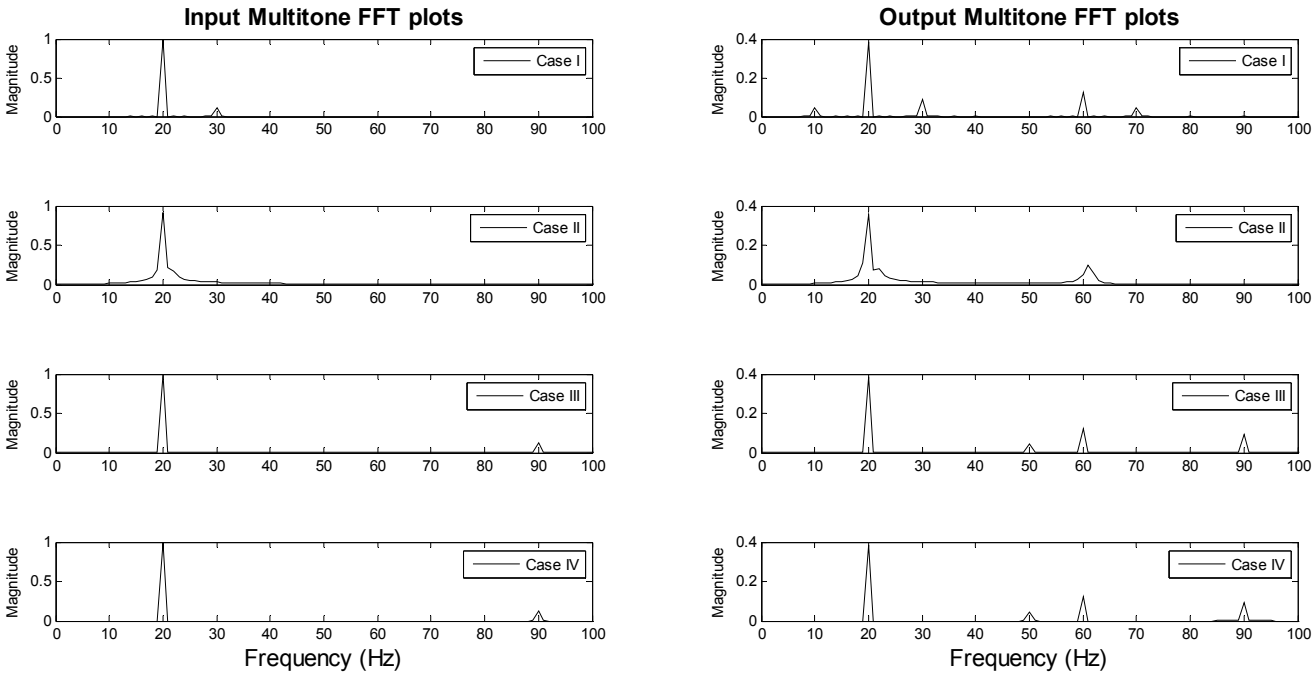

Figure 4.6: Frequency-zoomed (0–100Hz) FFT-captured signals' spectral plots for the four example cases. Each row represents one of the four cases under this study. Spectral plots in left column are input to the square law NLD, and those in right column are output from the NLD. These spectral plots are to be compared with numerical results obtained using our technique, as listed in Table 4.11.

For verification, the output signals from the square law device and synthesized signals using results from Table 4.11 are plotted in continuous time domain, showed side-by-side in Figure 4.7. The two columns are exactly the same. Hence, the proposed technique can reconstruct the output signal because the output components' amplitudes, frequencies, and phases can be precisely computed. The rest of this section contains the study of the four cases.

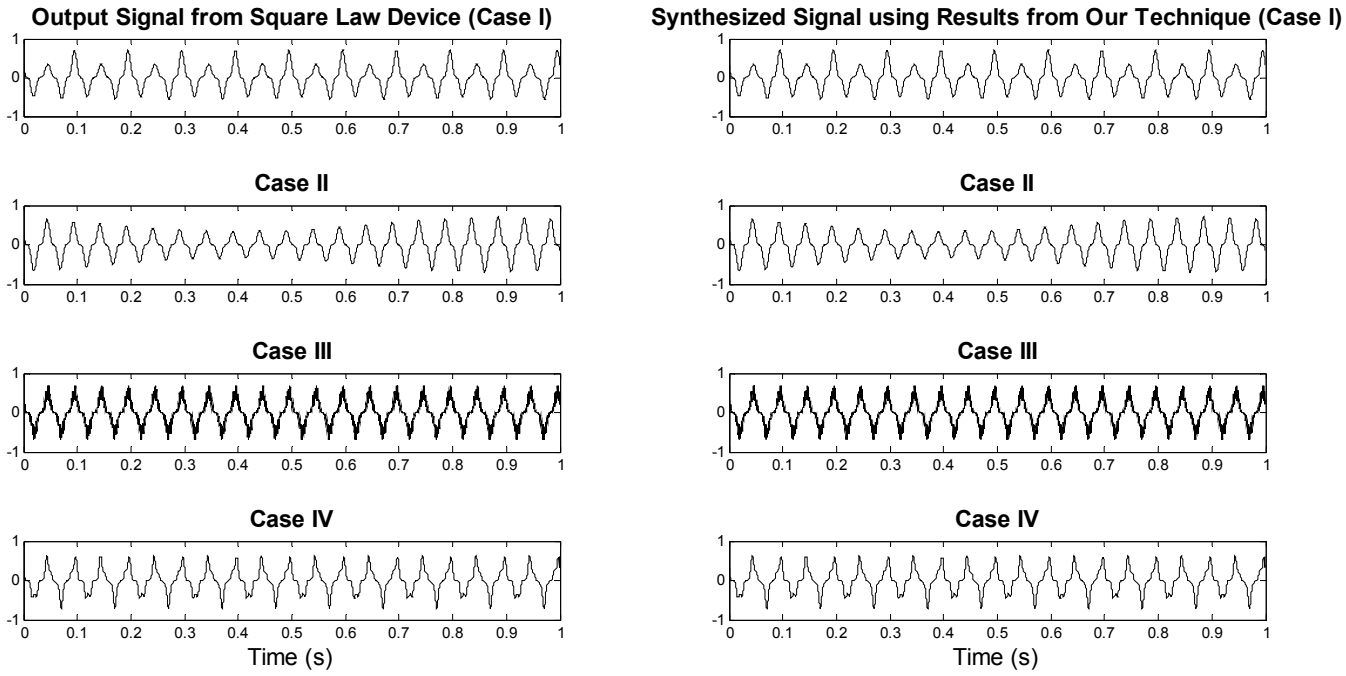

Figure 4.7: Output signals in continuous time domain from the square law device are plotted in left column, and synthesized signals using results from our technique in Table 4.11 are plotted in right column for the four cases. This figure also verifies our technique and shows that the output signal in continuous time can be easily synthesized using the results obtained from our technique.

In Case I, at 40 Hz, FFT cannot display the output spectrum line at amplitude 0.0054 which is smaller than the frequencies at 10, 20, 30, 60, and 70 Hz (See Table 4.11 and Figure 4.6). Note that our technique can precisely compute amplitudes, frequencies, and phases of the output multitone components. For this particular case, no intermodulation components, but all harmonically related output tones are generated because the two components of the input multitone signal are harmonically related to each other. Since the input components' frequencies are in integer multiples, and FFT does not have any problem distinguishing between frequency components for this case, there is no resolution problem for this case.

Case II demonstrates the frequency resolution problem and the picket fence effect of FFT. From Table 4.11, two input frequencies are very close to each other, and are not in integer numbers. Since FFT resolution is set to 1 Hz, it cannot resolve both input and output components well. In second row, right column of Figure 4.6, at the output, only two major peaks are observed around 20 and 61 Hz. The frequencies 19.180, 20.240, 21.300, and 22.360 Hz cannot be clearly distinguished among one another, and appear as one peak around 20 Hz. The frequencies 60.720, 61.780, 62.840, and 63.900 Hz also appear as a single peak around 61 Hz. Exact phase computation of each components using FFT method in this case is thus impossible.

Case III demonstrates the aliasing problem. The second component of 47,910 Hz is far more than Nyquist frequency of 24000 Hz. Therefore, folding effect occurs in both input and output FFT spectrum. The output component observed at 90 Hz is an aliased component due to 47,910 Hz (i.e., 48,000 Hz – 47,910 Hz = 90 Hz). Nonlinearity generates more components at the output; some of them are even higher frequencies. There are multiple aliasing (folding effect) occurrences in the Nyquist interval. Hence, the actual computed components from Table 4.11, 20, 60, 47870, 47910, 47950, 95800, 95840, and 143730 Hz appear at 20, 60, 130, 90, 50, 200, 160, and 270 Hz, respectively, on the FFT spectrum as a mixture of aliased and non-aliased components. From 0–100 Hz, only four components of 20, 50, 60, and 90 Hz are observed as shown in the third row, right column of Figure 4.6.

Case IV shows that on the FFT spectrum, aliased and non-aliased components look exactly the same although their time domain plots look different (See the fourth row of Figures 4.6 and 4.7). The output frequencies set of 20, 50, 60, 90, 130, 160,

200, and 270 Hz turn out to be the same as the aliased and non-aliased frequencies of Case III. Hence, on the FFT spectrum, frequency lines of Case IV look exactly the same as those of Case III, although their time domain plots show otherwise. In other words, this case reveals FFT aliasing problem from another point of view.

Hence, phase computation using FFT is not straightforward and approximate, and is prone to error under the conditions of aliasing, and other pitfalls mentioned above, whereas the proposed technique can precisely compute individual amplitudes, frequencies, and phases without any variations in error.

4.7 Conclusions

In this chapter, we proposed a technique to reduce the computational steps of the Lemma 1 of PIPCT by using triangular symmetric pattern and *gain multiplier*. With this technique, a large number of components in **d** do not need to be computed, thus significantly reducing computational time. The generalized formulae were derived from the first principle to compute the number of components generated by two techniques, PIPCT and MPIPCT. These techniques were compared, with step-reduction concepts depicted in equations, tables, and plots. A theorem was stated and proved that MPIPCT can reduce exponential component growth by 25–50%. We further showed how the two limitations of PIPCT can be solved by observing a mathematical pattern and modifying Lemma 2.

The Harmonic Addition Theorem was proposed to extend the computational algorithm for monomial nonlinearities to cases of polynomial nonlinearity. The proposed algorithm, YOGA, allows exact and efficient computation of any order of

polynomial nonlinearity with any number or combinations of amplitude, frequency and phase of input multitone. Simulation examples were given for easy comparison with FFT methods, and to highlight the pitfalls of using FFT in the four cases shown in Section 4.6.

In the next chapter, YOGA from this chapter and HASP algorithms (developed in last chapter) will be used to assess the traditional objective metrics such as THD and TIMD and compared with the perceptually motivated nonlinear distortion metrics and computational models, such as GedLee and *Rnonlin*.

CHAPTER 5

PERCEPTUAL EVALUATIONS OF NLDS

5.1 Introduction

The core objective of this chapter is to identify the most appropriate type of NLDs in the context of a VBS to achieve the best bass-intensity perception with minimum perceived impairment arising principally from intermodulation distortion. The techniques and algorithms developed in Chapter 3 and 4 are also applied in this chapter to compute the objective metric computations in an attempt to find a suitable metric to predict nonlinear distortion prediction.

The main thesis of VBS is that, when listening to a harmonic series with missing fundamental, the human auditory system tends to perceive the fundamental even though it is physically absent. Some simple examples can illuminate this process. Consider the truncated harmonic series {50, 100, 150, 200, 250, 300} Hz; it is evident that even if the series excludes 50 Hz, then it can easily be inferred, as the relationships are integer-based. On the other hand, if the series only has even multiples of 50 Hz, i.e. {100, 200, 300} Hz, then this is actually a harmonic series with a fundamental of 100 Hz, so the original fundamental of 50 Hz cannot be uniquely inferred, implying weaker likelihood of perception. However, for an odd multiple series, i.e. {50, 150, 250} Hz, then it is clear that 50 Hz can be inferred. Consequently, it could be argued that an NLD that produces mainly odd-order

harmonics should be preferable to one which only produces even-order harmonics, due to the way the fundamental is encoded uniquely into the sequence. Also, if the aim is to minimize the perception of distortion, then an NLD that yields only a small number of low-order harmonics—and preferably odd-order dominant terms—could prove attractive. Indeed, an odd-order sequence could use as few as two harmonics to uniquely define the fundamental, e.g. {150, 250}. For an NLD that delivers both even and odd terms simultaneously, a question arises as to whether the perceptual roles of even and odd harmonics are equal, or whether there is bias towards the odd terms in the sequence. However, there is a further complication relating to the perception of distortion, as the NLD must operate on audio signals normally possessing spread audio spectra, possibly without dominant bass frequencies (notes). Consequently, the nonlinearity will produce both harmonic (*desirable*) and intermodulation components (*undesirable*, even though filtering is used in the synthesis process to mitigate their generation). As the latter has no harmonic structure, it contributes only to the perception of signal distortion. Therefore, each NLD will offer a distinct balance of distortion and bass synthesis which is inevitably signal dependent; this leads to complications in formal subjective evaluation, especially as some imperfections may prove more acceptable (possibly even perceived as an enhancement) to a listener than others [Bar98]. This is why non-perceptually motivated metrics fail to establish the true performance of a VBS. As a result, intermodulation distortion analysis alone cannot reveal the true perceptually relevant artifacts caused by such nonlinear distortion, implying subjectively motivated measurements that are sensitive to the psychoacoustics of VBS must be performed to arrive at any conclusive comments on

perceived distortion. [CVA01a] [CVA01b] [VTC04]. As such, the selection of an optimum NLD methodology that minimizes the physical distortion components while maximizing the perception of bass extension can be intelligently performed.

Traditional listening tests are notoriously time-consuming, expensive and suffer numerous consistency problems; consequently, perceptually motivated models that predict distortion scores are desirable. In [GL03a] [GL03b], Lee and Geddes proposed the *GedLee(G_m)* metric, which incorporates an objective model of human masking [GL03a]. However, this failed to accurately predict nonlinear distortion in VBS [OG10]. A weakness of the *GedLee* metric for VBS is its application to the entire auditory spectrum; because the low-frequency distortion generated by the NLD-based VBS covers typically less than five octaves, its contribution is almost hidden in the context of the full-band model. To address this limitation, Tan and Moore *et al.* proposed two nonlinear distortion prediction models—*Distortion Score* (DS) [TMZ03] and *Rnonlin* [TMZ04]. However, the input signal to the DS model must be restricted to a multi-tone signal, whereas the input signal to the *Rnonlin* model can be either music or speech signals. In this chapter, the *Rnonlin* model is studied in terms of its applicability to VBS in the context of an audio system with a highly restricted low-frequency performance.

In this chapter the following three research questions are addressed:

1. From a pre-selected pool of NLDs, which introduce severe and inappropriate distortion artifacts?
2. From the subset of best-performing NLDs, which achieve good bass perception?

3. Can the *Rnonlin* perceptual model be applied successfully to NLD-based VBS distortion analysis?

Two experiments were performed. The first experiment focused on Questions 1 and 3, where perceptual contributions of nonlinear distortion were tested and correlated with the *Rnonlin* scores. The second experiment was designed to address Question 2.

This chapter is organized as follows: Section 5.2 describes the pre-selected set of NLDs, while Section 5.3 compares and contrasts the objective metrics. Section 5.4 introduces the *Rnonlin* model for VBS. Section 5.5 and 5.6 then present the statistical data analysis and results for distortion analysis and perception of bass enhancement derived from Experiments 1 and 2, respectively. From these subjective studies, NLDs are then categorized and analyzed further in Section 5.7 in terms of their input-output characteristics and derivatives, in order to arrive at common features that correlate well with good bass enhancement. Finally, Section 5.8 discusses approaches taken to assess bass enhancement, and draws conclusions as to the efficacy of using the *Rnonlin* model as a means of NLD performance sifting for VB applications.

5.2 NLD Functional Descriptors

Oda [Oda97] was first to propose a full-wave rectifier as an NLD in VBS. Subsequently, Feremans *et al.* [FS98] employed an analog half-wave rectifier, and Unemura [Une98] from Matsushita Electric again considered full-wave rectification as the harmonic generator within television (TV) audio circuits, to overcome the bandwidth limitation problem of modest loudspeakers. Larsen and Aarts then developed a generic NLD framework [LA02] [LA04], which has been adopted within

many systems [PRL10] [AMJ06] [OG08a] [OG08b] [OG10] [OGL10], including the present study. The generic framework is reproduced in Figure 5.1(a). However, in order to evaluate certain NLD-specific subjective attributes, the modified framework depicted in Figure 5.1(b) is used.

Thirteen NLDs were selected to exhibit a range of input-output symmetries and harmonic generation; their respective closed-form functions are listed in Table 5.1 with corresponding nonlinear transfer characteristics plotted in Figure 5.2. All selected NLDs investigated in this chapter are zero-memory (or static) systems, implying that a given output sample depends only on the corresponding input sample, with no convolution dependency. Generation of even, odd or both even-odd harmonics depends on the symmetry of the nonlinear transfer characteristics. The NLDs possessing either even- or odd-order transfer characteristics are therefore termed even-symmetric and odd-symmetric NLDs, respectively, whereas NLDs with non-symmetric transfer characteristics are designated non-symmetric. As shown in Figure 5.2, all the NLD input–output transfer characteristics pass through the origin; hence for zero input, the DC output offset is zero. Also, for the input signal normalized between ±1 range, the output range is bounded between ±1, with the exception of the "EXP2" and "ATSR" NLDs.

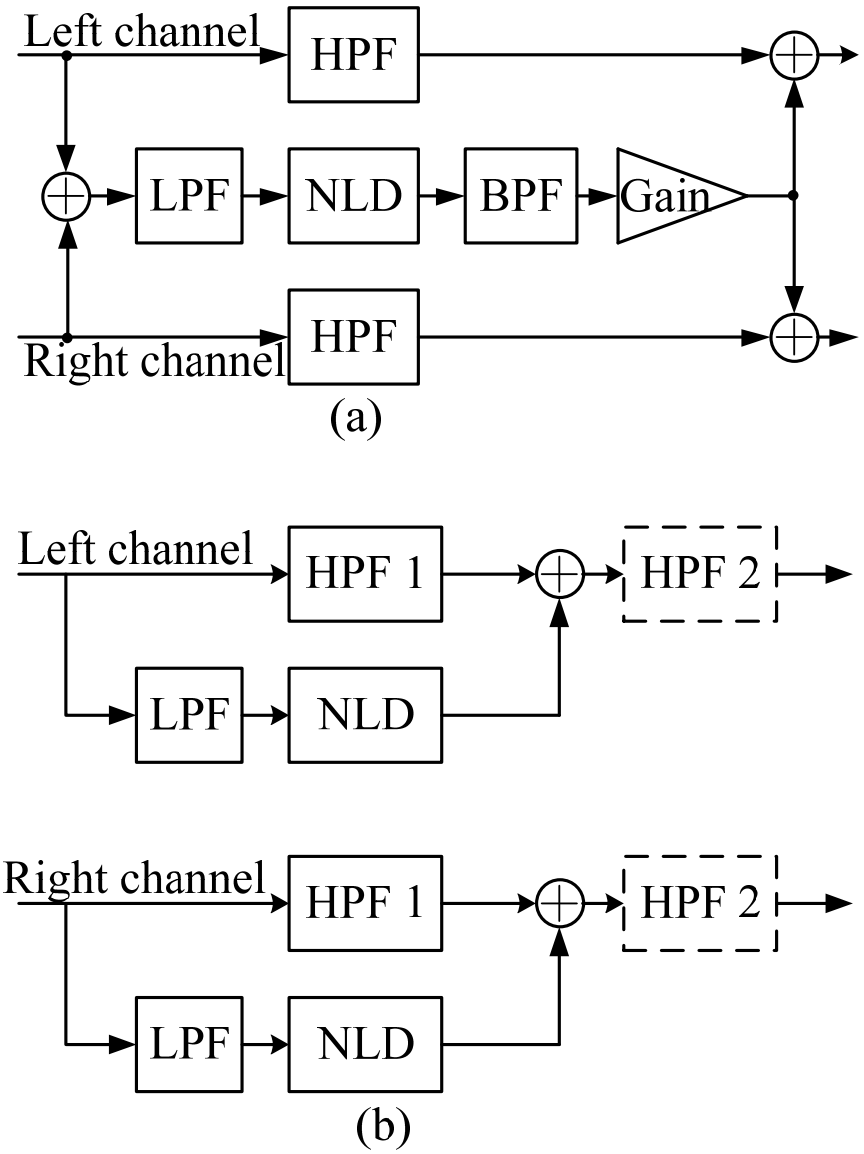

Figure 5.1: (a) The NLD-based signal processing framework proposed by Larsen and Aarts [LA02] [LA04]. LPF: low-pass filter; HPF: high-pass filter. (b) Block diagram of our VBS configuration used in Experiment 1 and 2. Left and right channels are processed separately; HPF2 in the dotted box is included only in Experiment 2, the bass listening test.

Table 5.1: The closed-form nonlinear equations, types of symmetry and harmonics generation of the thirteen NLDs used in the Experiment 1.

TC	Acronym	Name	$f(x)$	$f'(x)$	[1]S	[2]H
1	ASQRT	Absolute Square Root	$\sqrt{\lvert x\rvert}$	$\operatorname{sgn}(x)\Big/\left(2\sqrt{\lvert x\rvert}\right), x\neq 0$	[3]E	[4]E
2	FWR	Full-wave Rectifier	$\lvert x\rvert$	$\operatorname{sgn}(x), x\neq 0$	●	●
3	SQL	Square Law	x^2	$2x$	●	●
4	HWR	Half-wave Rectifier	$0.5\left(x+\lvert x\rvert\right)$	$0.5\left(x+\operatorname{sgn}(x)\right), x\neq 0$	[5]N	[6]f_0
5	CLP	Clipper	$\begin{cases}0.5\operatorname{sgn}(x), & \text{if } \lvert x\rvert>0.5\\ x & \text{otherwise}\end{cases}$	$\begin{cases}0 & \text{if } \lvert x\rvert>0.5\\ 1 & \text{otherwise}\end{cases}$	[7]O	[8]O
6	CUBE	Cube	x^3	$3x^2$	●	●
7	FEXP1	Fuzz Exponential 1	$\operatorname{sgn}(x)\dfrac{1-e^{-\lvert x\rvert}}{1-e^{-1}}$	$2\delta(x)\left(\dfrac{1-e^{-\lvert x\rvert}}{1-e^{-1}}\right)+\dfrac{\operatorname{sgn}^2(x)e^{-\lvert x\rvert}}{1-e^{-1}}$	●	●
8	FEXP2	Fuzz Exponential 2	$\operatorname{sgn}(-x)\dfrac{1-e^{\lvert x\rvert}}{e-1}$	$-2\delta(x)\left(\dfrac{1-e^{\lvert x\rvert}}{e-1}\right)+\dfrac{\operatorname{sgn}^2(x)e^{\lvert x\rvert}}{e-1}$	●	●
9	NSIG	Normalized Sigmoid	$\dfrac{(e^x-1)(e+1)}{(e^x+1)(e-1)}$	$\dfrac{e^x(e+1)}{(e^x+1)(e-1)}-\dfrac{(e^x-1)(e+1)e^x}{(e^x+1)^2(e-1)}$	●	●
10	NTANH	Normalized Hyperbolic Tangent	$\dfrac{(e^x-e^{-x})(e+e^{-1})}{(e^x+e^{-x})(e-e^{-1})}$	$\dfrac{e+e^{-1}}{e-e^{-1}}-\dfrac{(e^x-e^{-x})^2(e+e^{-1})}{(e-e^{-1})(e^x+e^{-x})^2}$	●	●
11	SQS	Square Sign	$x^2\operatorname{sgn}(x)$	$2x\operatorname{sgn}(x)+2x^2\delta(x)$	●	●
12	ATSR	Arc-tangent Square Root	$2.5\tan^{-1}(0.9x)+2.5\sqrt{1-(0.9x)^2}-2.5$	$\dfrac{2.25}{1+0.81x^2}-\dfrac{2.025x}{\sqrt{1-0.81x^2}}$	[5]N	[9]B
13	EXP2	Exponential 2	$\dfrac{e-e^{1-x}}{e-1}$	$\dfrac{e^{1-x}}{e-1}$	●	●

[1]Function Symmetry.
[2]Harmonics.
[3]Even Symmetry.
[4]Even Harmonics.
[5]Nonsymmetry.
[6]Fundamental frequency (first harmonic) and even harmonics.
[7]Odd Symmetry
[8]Odd Harmonics
[9]Both even and odd Harmonics.

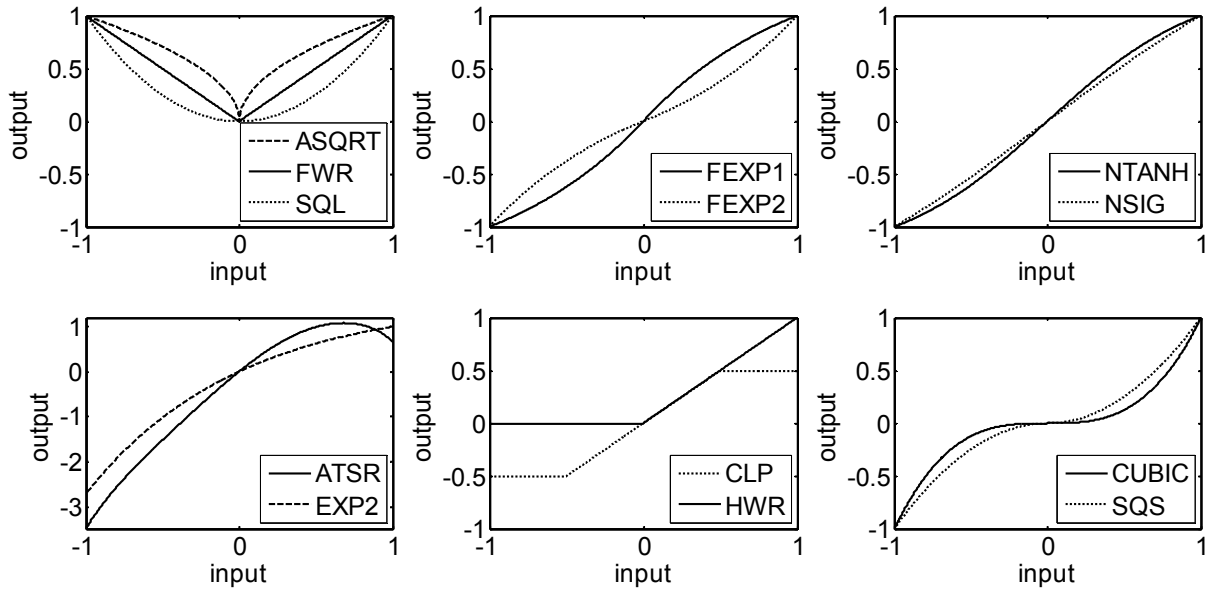

Figure 5.2: Zero-memory nonlinear transfer function plots of the thirteen NLDs.

5.3 Objective Metrics Calculations

Seven NLDs out of thirteen NLDs are polynomial approximated to the sixth order in a least-square sense. According to the triple duality theorem, presented in Chapter 3, the highest harmonic number generated by the sixth order polynomial is six. Moreover, as reviewed in Chapter 2, the dominant harmonics for the missing fundamental effect are within the first six harmonics. Therefore, the sixth order polynomial is sufficient to represent the NLD that can create virtual pitch perception.

The objective of this section is to show that selected objective metrics fail to reach the conclusion of which NLDs are suitable for VBS, and that the scores given by them to predict the nonlinear distortion contradict one another.

First let us define the system and metrics. The sixth order polynomial equation is given by

$$\hat{f}(x) = \hat{h}_0 + \hat{h}_1 x + \hat{h}_2 x^2 + \hat{h}_3 x^3 + \hat{h}_4 x^4 + \hat{h}_5 x^5 + \hat{h}_6 x^6, \tag{5.1}$$

where $\hat{h}_i, i \in \{0,1,2,\ldots,6\}$ are polynomial coefficients, $\hat{f}(x)$ is the polynomial nonlinear function or NLD, and x is the input. To approximate the nonlinear transfer function curves, the *polyfit* MATLAB function is applied. The *polyfit* function uses the Vandermonde matrix, as discussed in Chapter 3. The coefficients for the seven polynomials are listed in Table 5.2.

To assess the fitness of polynomial approximation, the relative root-mean-square error (RRMSE) metric is defined as follows [Abu10]:

$$\text{RRMSE} = \frac{\sqrt{\sum_{i=1}^{NP}\left[\hat{f}(x_i) - f(x_i)\right]^2}}{\sqrt{\sum_{i=1}^{NP}\left[f(x_i)\right]^2}}, \tag{5.2}$$

where $f(x_i)$ represents the exact value of NLD function in Table 5.1, $\hat{f}(x_i)$ represents the approximate value obtained from (5.1), and *NP* represents the total number of points used in the calculation. The results of RRMSE for the seven NLDs are also shown in Table 5.2.

Three categories of objective metrics will be used to calculate the scores. The first kind is single-tone metric, THD, which equation is defined in (3.45) in Chapter 3. The *HASP_Analysis* algorithm from Chapter 3 is used to compute the THD scores; results are listed in Table 5.2. The second kind is the multitone metrics. Two multitone metrics such as $\text{THD}_{\text{MT}}$and IMD_{MT} are used, where IMD and MT stand for intermodulation distortion and multitone, respectively. The THD_{MT} and IMD_{MT} metrics are calculated as follows [CVA01a]:

$$\mathrm{THD}_{\mathrm{MT}} = \sqrt{\frac{\sum_{k=1}^{L} H_k^2}{\sum_{k=1}^{L} H_k^2 + \sum_{i=1}^{N} MT_i^2}}, \tag{5.3}$$

$$\mathrm{TIMD}_{\mathrm{MT}} = \sqrt{\frac{\sum_{k=1}^{M} IM_k^2}{\sum_{k=1}^{M} IM_k^2 + \sum_{i=1}^{N} MT_i^2}}, \tag{5.4}$$

where $\sum_{i=1}^{N} MT_i^2$ represents the power of all input multitone signal, $\sum_{k=1}^{M} IM_k^2$ represents the power of all intermodulation products, $\sum_{k=1}^{L} H_k^2$ is the power of all harmonic products, N is the number of input primary tones, M is the number of intermodulation products, and L is the number of harmonics.

Czerwinski et *al.* recommended the use of logarithmically equally spaced multitone signals because they provide more accurate information about nonlinear systems, compared to harmonic, two-tone intermodulation distortion measurements [CVA01a] [CVA01b]. Logarithmically equally spaced multitone signal are also preferable to other test signals in this case because the statistical properties of a multitone signal resemble those of a noise or a musical signal. The detailed derivation of the generation of multitone signal is not available in [CVA01a] and [CVA02b].

Here, we derive a generalized formula to generate logarithmically equally spaced multitone signals. For the multitone frequencies $\omega_k, k \in \{0,1,\ldots,N-1\}$, assume that $\omega_0 < \omega_1 < \cdots < \omega_{N-1}$ and $\omega_i \neq \omega_j$ for $i \neq j$; $i, j \in \{0,1,\ldots,N-1\}$. The objective is to obtain a generalize formula for the linear frequency multiplication factor β such that

$$\omega_{k+1} = \beta\omega_k, \text{ for } k = 0,1,\ldots,N-2, \tag{5.5}$$

while the following logarithmically equally spaced condition is guaranteed:

$$\begin{aligned}\alpha &= \log_{10}(\omega_1) - \log_{10}(\omega_0) = \cdots = \log_{10}\left(\omega_{N-1}\right) - \log_{10}\left(\omega_{N-2}\right) \\ &= \log_{10}(\omega_{k+1}) - \log(\omega_k), \text{ for } k = 0,1,2,\ldots,N-2,\end{aligned} \tag{5.6}$$

This condition can be expressed as follows:

$$\alpha = \frac{\log_{10}\left(\omega_{N-1}\right) - \log_{10}\left(\omega_0\right)}{\left(N-1\right)} = \log_{10}\left(\omega_{N-1}/\omega_0\right)/\left(N-1\right). \tag{5.7}$$

Since

$$\beta = 10^{\alpha} = \log_{10}^{-1}\{\alpha\}, \tag{5.8}$$

substituting (5.7) into (5.8), the following generalized formula is obtained:

$$\begin{aligned}\beta &= \log_{10}^{-1}\left\{\log_{10}\left(\omega_{N-1}/\omega_0\right)/\left(N-1\right)\right\} \\ &= \sqrt[N-1]{\log_{10}^{-1}\left\{\log_{10}\left(\omega_{N-1}/\omega_{1-1}\right)\right\}} \\ &= \sqrt[N-1]{\frac{\omega_{N-1}}{\omega_0}}.\end{aligned} \tag{5.9}$$

By recursion, the logarithmically equally spaced multitone frequencies can be easily computed as follows:

$$\omega_n \leftarrow \beta^n \omega_0, \text{ for } n = 0,1,\ldots,N-1, \tag{5.10}$$

where β is computed using (5.9), and ω_0 is the lowest frequency. For example, multitones with N={5, 10, 15, 20} Hz partials from the lowest frequency $\omega_0 = 1000$ Hz to the highest frequency $\omega_{N-1} = 10000$ Hz are shown in Figure 5.3.

The respective βs are obtained as {1.7783, 1.2915, 1.1788, 1.1288}. As for the computations of THD_{MT} and $TIMD_{MT}$ metrics, a multitone with five partials is used with the unitary amplitudes, frequencies {1000.000, 1778.279, 3162.278, 5623.413, 10000.00} Hz and phases {2.690427, 3.351363, 4.750792, 4.362870, 4.953241}

radians. The computational algorithm *YOGA* from Chapter 4 is used to compute THD_{MT} and $TIMD_{MT}$ scores, and the results are also shown in Table 5.2.

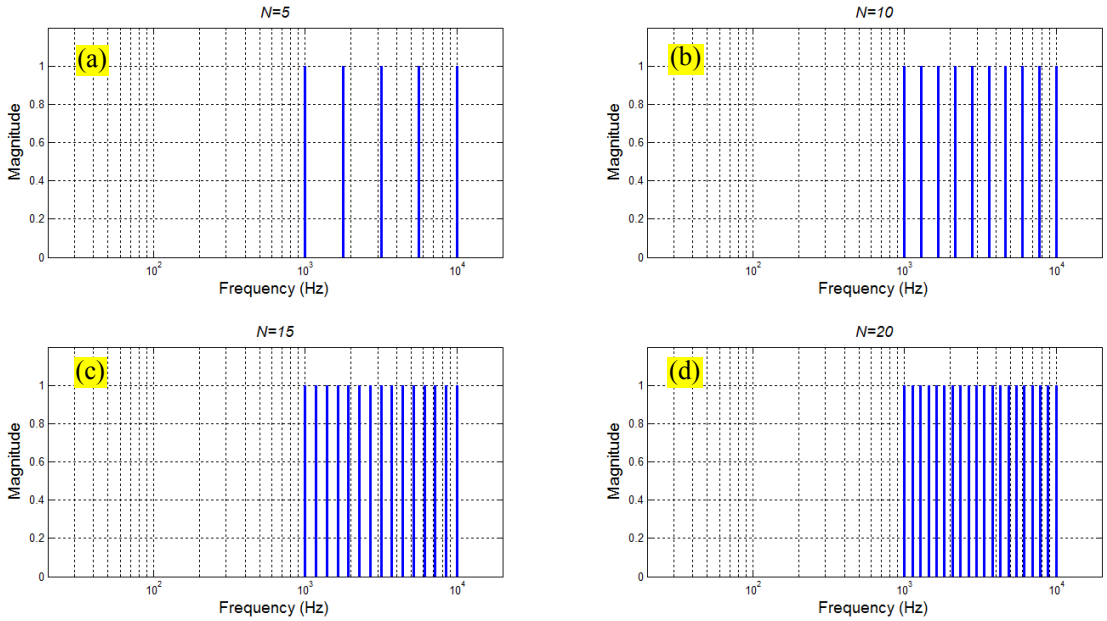

Figure 5.3: Logarithmically equal spaced multitone signals. (a) five (b) ten (c) fifteen, and (d) twenty components.

The third kind of objective metric used in this Section is the GedLee metric (G_m). The equation is given by

$$G_m = \sqrt{\int_{-1}^{1} \left(\cos \frac{\pi}{2} x \right)^2 \left(\frac{d^2 \hat{f}(x)}{dx^2} \right)^2 dx}, \tag{5.11}$$

where x is the input, $\hat{f}(x)$ is the polynomial nonlinear equation in (5.1). The main difference between G_m and {THD, THD_{MT}, $TIMD_{MT}$} is that G_m is the property of the system, not of a signal sent through the system [GL03a] [GL03b]. When designing the G_m metric, Geddes and Lee proposed the three distortion principles:

1. High-frequency distortion products are masked to a greater extent than those that are lower in the frequency.
2. Distortion products that lie closer to the masker are less likely to be perceived than those that lie farther away.
3. Distortion products are likely to be more perceptible at lower signal levels than at higher signal levels.

The G_m has the following features:

1. It is more sensitive to higher order nonlinearities than the lower order nonlinearities.
2. It is weighted towards greater values for nonlinearities at lower signal levels.
3. It is immune to changes in offset and gain (first-order slope).
4. It is completely independent of the actual input signal that passes through the system.
5. It is a frequency-independent measure for the frequency-independent (memoryless) nonlinear systems.

The second derivative of the nonlinear transfer function in (5.11) makes G_m more sensitive to the higher order nonlinearities. The second derivative is squared to alleviate the sign problem. To address the second feature above, a cosine-squared function, which is unity for small values of the signal, and zero for large ones, is multiplied to the square of the second derivative of the nonlinear transfer function. The resultant function is integrated from -1 to 1, giving a number of which the square root yields a single number: G_m [GL03a]. Therefore, the G_m metric is derived from the general theory of nonlinear systems and the auditory masking mechanism.

Table 5.2: The polynomial coefficients, MMSE, THD, THD_{MT}, $TIMD_{MT}$, and G_m scores of the seven NLDs.

	FEXP1	FEXP2	NSIG	NTANH	SQS	ATSR	EXP2
$\hat{h}_0$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0015	0.0000
$\hat{h}_1$	1.4215	0.6410	1.0819	1.3100	0.2051	2.7335	1.5820
$\hat{h}_2$	0.0000	0.0000	0.0000	0.0000	0.0000	-1.0596	-0.7910
$\hat{h}_3$	-0.7413	0.4665	-0.0896	-0.4080	1.2304	-0.9365	0.2634
$\hat{h}_4$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0132	0.0659
$\hat{h}_5$	0.3318	-0.1120	0.0077	0.0991	-0.4511	0.2907	0.0137
$\hat{h}_6$	0.0000	0.0000	0.0000	0.0000	0.0000	-0.3555	-0.0023
RRMSE	0.0059	0.0025	0.0000	0.0004	0.0088	0.0010	0.0000
THD	7.8516 %	8.8973 %	1.9613 %	6.6871 %	**19.9775 %**	**31.9922 %**	24.3424 %
THD_{MT}	49.3573 %	37.4723 %	**69.6498 %**	**50.8743 %**	**45.5609 %**	**35.2551 %**	57.2327 %
$TIMD_{MT}$	**63.5630 %**	**67.4929 %**	**23.3829 %**	**62.7880 %**	65.1644 %	67.3945 %	55.8381 %
G_m	**0.9742**	**0.7737**	0.1771	0.6743	1.7628	3.0520	1.7971

Referring to Table 5.2, interpretation of the scores is that the higher the value is, the higher perceived distortion will be.

However, the G_m of FEXP1 is 0.9742 and that of FEXP2 is 0.7737; whereas the $TIMD_{MT}$ of FEXP1 is 63.5630 %, and that of FEXP2 is 67.4929 %. The G_m of FEXP1 is higher than that of FEXP2, but $TIMD_{MT}$ of FEXP1 is lower than that of FEXP2. Thus, the two metrics give contradictory predictions of degrees of nonlinear distortion. Another example of contradicting metrics is the THD_{MT} and $TIMD_{MT}$ of NSIG and NTANH. As shown in Table 5.2, the THD_{MT} of NSIG (69.6498%) is higher than the THD_{MT} of NTANH (50.8743%). On the other hand, the $TIMD_{MT}$ of NSIG (23.3829%) is lower than the $TIMD_{MT}$ of NTANH (62.7880%). A third example of contradicting metrics from Table 5.2 is the THD and THD_{MT}. The THD of SQS (19.9775%) is lower than the THD of ATSR (31.9922%), whereas the THD_{MT} of SQS (45.5609%) is higher than the THD_{MT} of ATSR (35.2551%).

Based on the above three examples, the four metrics contradict one another. In the next Section, *Rnonlin* nonlinear distortion prediction model is verified with a subjective listening test.

5.4 Rnonlin Psychoacoustic Nonlinear Distortion Model

The *Rnonlin* [TMZ04] is a computational model for perceptually motivated nonlinear distortion prediction that can be used with both music and speech excitations [Voi06] [SH07]. Critically, it has been reported in [TMZ04] to achieve enhanced correlation over the DS model [TMZ03] using distortion scores derived from listening tests. The *Rnonlin* model is applied here to an NLD-based VBS evaluation, where the following paragraphs describe key implementation blocks.

Figure 5.4 shows the block diagram of the *Rnonlin* model (shaded box), as proposed in [SH07]. In this VB study, the *Rnonlin* model is based on the published steps given by Tan *et al.* [TMZ04] and Santis *et al.* [SH07]. Both papers recommended time alignment of input and output signals to compensate for the delay caused by the system under test. Instead of performing offline time alignment as proposed by previous researchers, zero-delay filters [OSB99] [Gus96] are employed as shown in Figure 5.1(b), to time-align the clean signal with the VBS-processed signal. The zero-delay filter was implemented using the MATLAB *filtfilt* function.

Referring to Figure 5.4, the *Rnonlin* model takes in both the clean signal and the processed signal and produces a score, known as the *Rnonlin* score. Both the clean signal and the processed signal pass through the outer and middle ear (OME) transfer function, which is based on the ANSI/ASA S3.4-2007 standard [ANSI07]. The 40 immediate frequency values, together with their respective decibel-magnitudes of the

middle-ear transfer function are obtained from Table 3 of the standard [ANSI07]. In addition, the function *gene_Middle_Ear_Transfer_function* in [Gen10] is used to implement the middle-ear transfer function, plotted in Figure 5.5 (a).

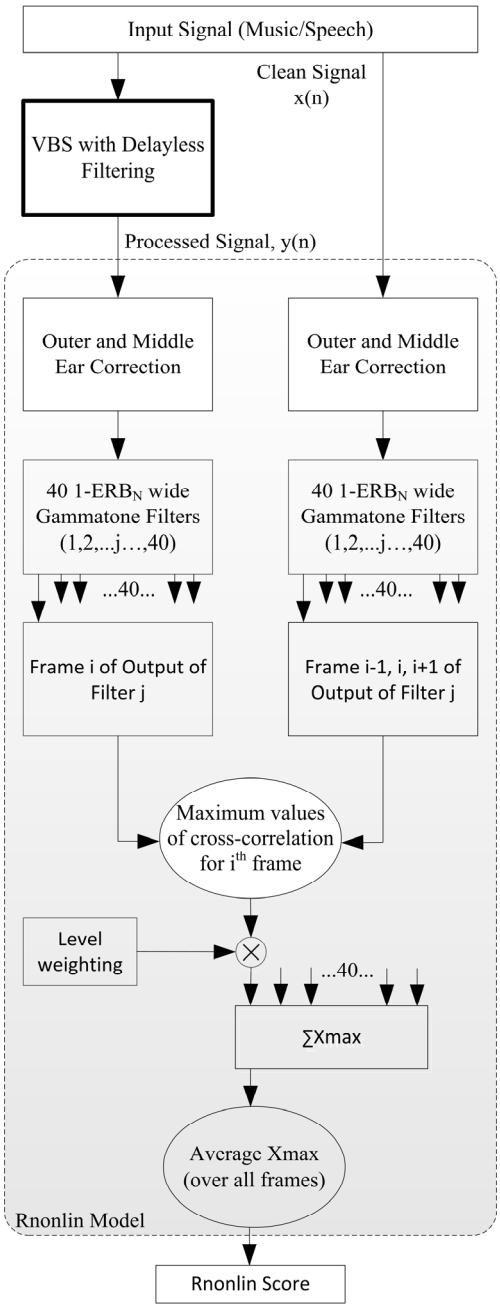

Figure 5.4: Block diagram of *Rnonlin* calculation algorithm (Redrawn from [SH07]).

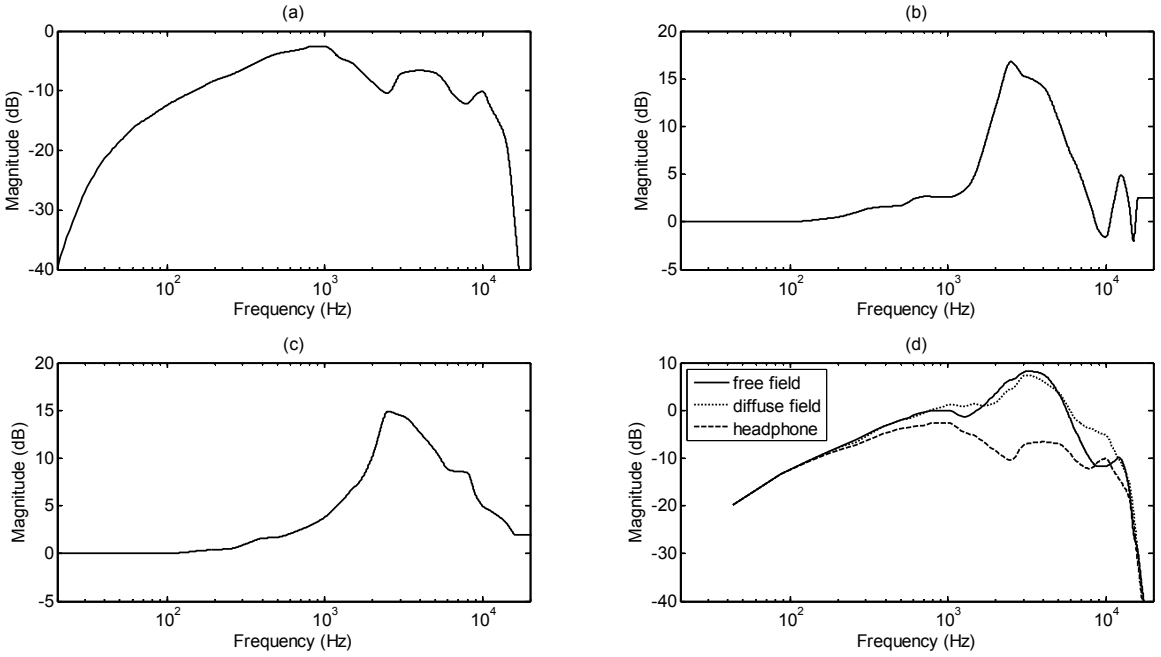

Figure 5.5: (a) middle-ear transfer function (b) free field to ear drum transfer function (c) diffuse field to ear drum transfer function (d) outer and middle-ear transfer functions for the three listening conditions: free field, diffuse field, and headphone.

The OME transfer function filter is a cascade structure of the outer-ear transfer function filter and the middle-ear transfer function filter. For the outer-ear transfer function, there are three listening conditions [ANSI07] to be considered: listening in free field, listening in diffuse field, and listening through headphones. For cases of listening in free field and diffuse field, numerical values described in Table 1 and 2 of the ANSI standard [ANSI07] are used, respectively. The resultant transfer functions are plotted in Figure 5.5 (b) and (c). In the case of listening through headphones, the outer-ear transfer function can be removed, leaving only middle-ear transfer function, with the assumption that the frequency response of the headphone is flat [ANSI07]. Based on interpolated values, the OME transfer function is designed

using a 4097-coefficient finite-impulse response (FIR) filter for the *fir2* function in MATLAB [Gen10].

To implement 40 equivalent rectangular-bandwidth (ERB) filter banks [Moo03] in the *Rnonlin* model, Auditory Toolbox [Sta98] functions, *MakeERBFilters* and *ERBFilter-Bank,* are used. Filter coefficients for the frequencies range from 35 Hz to 24 kHz are obtained using the *MakeERBFilters* function; actual filtering for both clean and processed signals are performed using the *ERBFilterBank* function. As shown in Figure 5.4, both clean and processed signals pass through the OME and 40 ERB filter banks.

At the output of each ERB filter, both clean and processed signals are split into 30 ms frames. Let L denotes the number of samples in a frame, which is equal to 1323 samples for a 30 ms frame sampling at 44.1 kHz. The cross-correlation between the clean signal and the processed signal at the *i*th frame at the *j*th filter is computed by concatenating the *i*-1, *i* and *i*+1 frames for the time lags of η between ± 10 ms or ± 441 unit samples [Voi06]. This cross correlation computation step can be expressed concisely as,

$$r_{x,y} = \frac{\mathrm{cov}\{x(n), y(n-\eta)\}}{\sqrt{\mathrm{var}\{x(n)\} \times \mathrm{var}\{y(n-\eta)\}}}, \tag{5.12}$$

or in full form as

$$r_{x,y}(i;j;\eta) = \frac{\sum_{n=(i-1)L+1+\eta}^{iL+\eta} x(n;j)\,y(n-\eta;j)}{\sqrt{\left(\sum_{n=(i-1)L+1+\eta}^{iL+\eta} x(n;j)^2\right)\left(\sum_{n=(i-1)L+1+\eta}^{iL+\eta} y(n-\eta;j)^2\right)}}, \tag{5.13}$$

for $-441 \le \eta \le +441$ [TMZ04] [SH07]. In (5.12), cov{.} and var{.} denote covariance and variance, respectively. For each *i*th frame at the *j*th filter, the maximum correlated value is computed as

$$X_{max}(i;j) = \max_{-441 \le \eta \le +441} \left[r_{x,y}(i;j;\eta) \right], \tag{5.14}$$

for $i = 1, 2, \ldots, \lfloor \text{total number of samples}/L \rfloor$ and $j = 1,\ 2, \ldots,\ 40$. Tan *et al.* [TMZ04] assumed that the value of X_{max} provides a measure of the processed signal that is affected by the frequency components not present in the original signal. Referring to Figure 5.4, the remaining steps are weight computation and application of weights to $X_{\max}(i;j)$. The weighting is implemented in the model based on the assumption that the low-level output at the ERB filter banks corresponds to small levels of perceived distortion [TMZ04]. The individual levels for each frame and each filter are computed as

$$Level(i;j) = 10 \times \log_{10} \left[\frac{1}{L} \sum_{n=(i-1)L+1}^{iL} y(n;j) y(n;j) \right], \tag{5.15}$$

and for each *i*th frame, the maximum value of the *j*th filter is taken as

$$max\_weight(i) = \max \left[Level(i;j) \right], \tag{5.16}$$

for $j = 1, 2, ..., 40$. By using $max\_weight(i)$, the weights, which are applied to $X_{\max}(i;j)$, are computed as

if $Level(i;j) > max\_weight(i) - 40$

$weighting(i,j) = max\_weight(i);$

elseif $Level(i,j) > max\_weight(i) - 80$ **and** $Level(i,j) < max\_weight(i) - 40$

$weighting(i,j) = \left(Level(i,j) - \left(max\_weight(i) - 40 \right) \right) / 40;$

else $weighting(i,j) = 0.$ (5.17)

The pseudocode (5.17) shows that the scaling is performed in such a way that if the level is within 40 dB from the maximum value, the weight is assigned to the maximum weight. If the level is lower than 40 dB from the maximum value but higher than 80 dB from the maximum value, the weight is scaled linearly with level. If the level is lower than 80 dB from the maximum value, the weight is assigned to zero. The objective is to ensure that low-level outputs are assigned weight equal to zero, very high level outputs are assigned the maximum weight and the remaining intermediate levels are assigned weights proportionally to the level.

The final *Rnonlin* score is obtained by the following steps. First, the normalized weights,

$$normalized\_weight(i,j) = \frac{weighting(i,j)}{\sum_{j=1}^{40} weighting(i,j)} \tag{5.18}$$

are applied to $X_{\max}(i;j)$ as,

$$X_{max\_weighted}(i,j) = X_{max}(i,j) \times normalized\_weight(i,j). \tag{5.19}$$

$X_{max\_weighted}(i,j)$ are then summed across all filters and averaged across the frames to yield the final *Rnonlin* score,

$$Rnonlin = \underset{i}{Mean}\left\{\sum_{j} X_{max\_weighted}(i,j)\right\}. \tag{5.20}$$

Rnonlin ranges from 0 to 1, where 0 represents the worst perceptual distortion score and 1 represents no distortion; the *Rnonlin* score therefore decreases with increases in distortion.

5.5 Experiment 1: Distortion Listening Test

The objective of this section is to evaluate the perceived quality of different NLDs and to investigate the ability to predict the quality of perceived distortion using the *Rnonlin* model. The design and analysis of the distortion listening test are presented while statistical analysis techniques such as ANOVA, nonlinear regression and correlation analysis are used to analyze the listening test data.

5.5.1 Audio materials

Six audio files from the Sound Quality Assessment Material (SQAM) CD [EBU08] were selected for Experiment 1. The selected audio files together with their respective durations in seconds are listed in Table 5.3. The term Audio Sequence/Stimulus (AS) is used to denote the audio files. A speech signal AS1 was included in the listening test to investigate whether the NLDs under test have any special influences on speech signal as compared to musical soundtracks.

Table 5.3: Audio materials selected for the Experiment 1.

AS	Content	SQAM Tracks	Seconds	Genres
1	Female speech in English	49	4	Speech
2	Pop music Eddie Rabbit	70	14	Pop
3	Bass-clarinet	17	8	Clarinet
4	Piano Schubert	60	12	Piano
5	Vocal Bass	47	11	Vocal
6	Soprano Mozart	61	4	Classical

5.5.2 Listeners

A total of 40 pre-screened and paid listeners (10 females and 30 males, average age: 25 years [range: 20–35 years]) were selected to carry out the subjective listening test for this experiment. All the listeners were briefed on the details of the listening test.

5.5.3 Processing of Audio Materials

All *.wav* files were processed offline and based on the VBS block diagram shown in Figure 5.1 (b). The input signal is first separated into low- and high-frequency bands using the low-pass filter (LPF) and the high-pass filter 1 (HPF1), respectively. The Park-McClellan filter design algorithm [OSB99] was used to design both HPF and LPF FIR digital filters, each with a cutoff frequency of 280 Hz. The magnitude responses for both HPF and LPF FIR filters are plotted in Figure 5.6 with pass-band ripple and stop-band attenuation set to 1 dB and 60 dB, respectively.

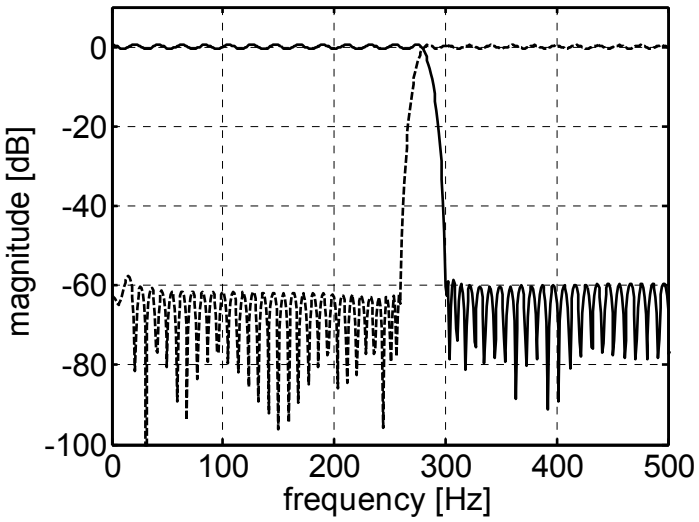

Figure 5.6: Magnitude responses of LPF and HPF1 FIR filters.

As discussed in Section 5.3, zero-phase digital filtering is performed using the MATLAB *filtfilt* function instead of the normal *filter* function. Note that high-pass filter 2 (HPF2) is not used in Experiment 1 in order to evaluate the full-range perceived distortion. There are 84 audio files consisting of one reference signal, thirteen NLD-processed signals for six audio stimuli (AS) where all *.wav* files are normalized to -1 dB full scale (dBFS) before inputting into the VBS.

5.5.4 Listening test configuration

Listening tests were conducted in a quiet room located at NTU where a personal computer ran the listening software. A Sennheiser HD600 headphone set was connected to a Creative Extigy USB-external sound card and a comfortable listening level established which then was rigorously maintained throughout all experiments. The graphical user interface (GUI) program (i.e. listening software) shown in Figure 5.7 was developed in Microsoft Visual Basic .NET 2008.

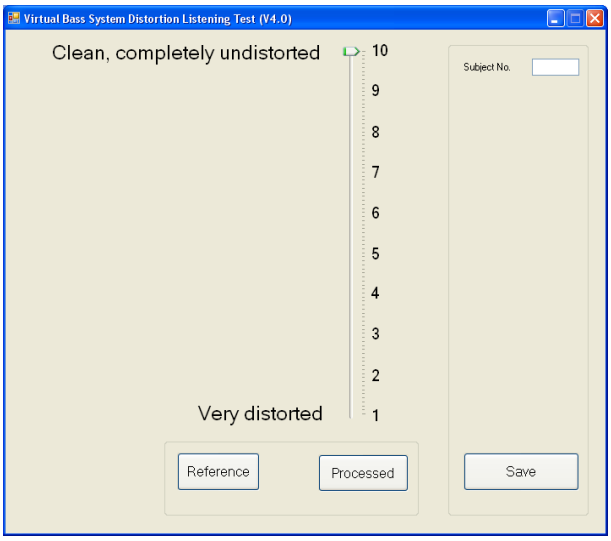

Figure 5.7: Graphical user interface used during Experiment 1.

Each NLD represents one test condition (TC) in Experiment 1. There are thirteen TCs, each of which was tested on the six AS. TC/AS randomization was performed by the GUI program during startup for every listener to facilitate a robust double-blind testing procedure and to reduce the recency effect (i.e., bias toward recent events) [ZRB08]. A 10-points scale with 0.1 resolution was employed to grade the sound quality of each of the processed audio tracks. The highest unit 10 represents a "clean, completely undistorted" audio signal while the lowest unit 1 represents a "highly distorted" audio signal. This scale was used to compare the obtained distortion scores with the distortion scores predicted by the *Rnonlin* model, whose nonlinear regression results are shown in Section 5.5.6.

5.5.5 ANOVA analysis

The resulting mean scores for the six AS, with 95% confidence intervals, are sorted in descending order and listed in Table 5.4. From Table 5.2 and 5.4, notice that G_m metric fails to predict correctly for the case of FEXP1 and FEXP2 for AS 1, 2 and 4. According to G_m, FEXP1 is more distorted than FEXP2, but the mean values of AS 1, 2, and 4 indicate otherwise.

To compare the mean values using global hypothesis testing, one-way analyses of variance (ANOVA) were performed. In ANOVA, there are independent variables (factors) and dependent variables (levels). One-way ANOVA is used to test whether two or more levels of a factor have the same mean, as it allows inferences about whether samples are drawn from a population having the same mean. For the first ANOVA test, the factor is the mean score of a TC for a fixed AS with 13 levels,

each representing a TC. Because six AS were used in the listening test, this ANOVA test is repeated six times for different AS. On the other hand, for the second ANOVA test, the factor is the mean score of an AS for a fixed TC with 6 levels, again each representing an AS. Again, 13 ANOVA tests were carried out for 13 different TCs.

Table 5.4: Mean scores (MSs), upper confidence bounds (UCBs) and lower confidence bound (LCBs) of six audio sequences from Experiment 1 data. Mean scores are sorted in descending order.

AS	TC	MS	LCB	UCB	AS	TC	MS	LCB	UCB
1	NTANH	7.66	6.97	8.34	4	NSIG	9.52	9.31	9.74
	CLP	7.43	6.79	8.07		CLP	9.50	9.23	9.76
	NSIG	7.40	6.72	8.09		FEXP1	9.46	9.19	9.73
	FEXP1	7.38	6.75	8.01		NTANH	9.33	9.07	9.60
	FEXP2	6.81	6.10	7.51		FEXP2	9.29	8.99	9.59
	EXP2	6.46	5.74	7.17		EXP2	9.06	8.73	9.39
	HWR	5.11	4.44	5.78		HWR	8.98	8.66	9.30
	SQS	5.05	4.39	5.71		SQS	7.95	7.31	8.60
	ATSR	4.96	4.30	5.61		CUBIC	7.70	7.06	8.33
	CUBIC	4.62	3.88	5.35		SQL	7.58	6.92	8.24
	SQL	4.34	3.70	4.98		ATSR	7.17	6.56	7.78
	FWR	2.99	2.40	3.59		FWR	5.22	4.45	5.99
	ASQRT	1.62	1.32	1.93		ASQRT	1.40	1.15	1.64
2	NSIG	9.16	8.60	9.72	5	CLP	9.40	9.02	9.78
	CLP	9.14	8.81	9.47		FEXP2	9.40	9.20	9.61
	FEXP1	9.09	8.75	9.42		NSIG	9.23	8.69	9.77
	NTANH	8.98	8.56	9.41		EXP2	9.14	8.64	9.65
	EXP2	8.92	8.42	9.41		FEXP1	9.13	8.64	9.62
	FEXP2	8.59	8.10	9.08		HWR	9.04	8.58	9.49
	HWR	7.54	6.83	8.26		NTANH	9.04	8.52	9.57
	ATSR	7.17	6.44	7.91		ATSR	7.95	7.34	8.57
	SQS	5.47	4.70	6.24		SQS	6.93	6.33	7.53
	FWR	5.10	4.39	5.81		CUBIC	6.51	5.79	7.23
	SQL	5.02	4.25	5.78		SQL	6.33	5.64	7.02
	CUBIC	5.01	4.26	5.76		FWR	5.38	4.58	6.18
	ASQRT	1.69	1.37	2.02		ASQRT	1.71	1.31	2.12
3	CLP	9.25	8.83	9.67	6	CLP	9.62	9.44	9.81
	NSIG	9.19	8.74	9.63		HWR	9.53	9.26	9.80
	FEXP2	9.12	8.74	9.49		EXP2	9.50	9.23	9.76
	FEXP1	8.97	8.62	9.33		NTANH	9.42	9.02	9.81
	NTANH	8.97	8.48	9.46		FEXP2	9.41	9.11	9.72
	EXP2	8.75	8.27	9.23		FEXP1	9.39	9.04	9.74
	HWR	7.47	6.76	8.18		NSIG	9.39	9.14	9.65
	ATSR	6.62	5.88	7.36		ATSR	8.97	8.59	9.34
	SQS	5.99	5.28	6.70		SQS	8.25	7.62	8.87
	CUBIC	5.63	4.91	6.36		FWR	8.18	7.53	8.83
	SQL	5.63	4.92	6.35		CUBIC	8.16	7.60	8.72
	FWR	5.54	4.78	6.30		SQL	7.85	7.18	8.53
	ASQRT	1.66	1.34	1.99		ASQRT	3.12	2.49	3.76

For each factor, the global hypotheses to test are defined as follows. For a fixed AS (i.e. for each $AS1, AS2, \ldots, AS6$),

$$\begin{aligned} H_0 &: \mu_{TC1} = \mu_{TC2} = \ldots = \mu_{TC13} \\ &\quad \text{(i.e., means of all levels are equal),} \\ H_a &: \mu_{TCi} \neq \mu_{TCj} \text{ for } i \neq j \in \{1,2,\ldots,13\} \\ &\quad \text{(i.e., at least two means are not equal).} \end{aligned} \tag{5.21}$$

For a fixed TC (i.e. for each $TC1, TC2, \ldots, TC13$),

$$\begin{aligned} H_0 &: \mu_{AS1} = \mu_{AS2} = \ldots = \mu_{AS6} \\ &\quad \text{(i.e., means of all levels are equal),} \\ H_a &: \mu_{ASi} \neq \mu_{ASj} \text{ for } i \neq j \in \{1,2,\ldots,6\} \\ &\quad \text{(i.e., not all the means are equal).} \end{aligned} \tag{5.22}$$

In (5.21) and (5.22), H_0 is the null-hypothesis and H_a is the alternative-hypothesis, where $\mu_{TC1}, \mu_{TC2}, \ldots, \mu_{TC13}$ denote mean-opinion scores of the 13 TCs for a fixed AS and $\mu_{AS1}, \mu_{AS2}, \ldots, \mu_{AS6}$ denote mean-opinion scores of the 6 AS for a fixed TC. In (5.21) the condition whether all TCs (each representing a respective NLD) generate the same mean is tested. In addition, the condition whether all AS generate the same mean is tested in (5.22). In other words, (5.21) tests the equivalence of the mean values of different NLDs on distortion perception, while (5.22) tests for the condition whether at any fixed TC or NLD, the choice of any audio stimulus can generate the same mean value.

Table 5.5 shows the generic data format for a single-factor ANOVA test which holds true for both ANOVA tests (5.21) and (5.22), where the symbols are defined as follows: L, number of levels, namely $L = 13$ is used for the fixed-AS and $L = 6$ is

used for the fixed-TC ANOVA tests; n, number of subjects, $n = 40$ is applied to both fixed-AS and fixed-TC ANOVA tests; $N = L \times n$, the total number of observations; $y_{i,j}$, listening test distortion scores where i ranges from 1 to L and j from 1 to n; $\bar{y}_i$, mean scores for level i and $\bar{\bar{y}}$, grand mean.

Table 5.5: Data format for a single-factor ANOVA [Mon01].

Level	Observations				Averages
1	y_{11}	y_{12}	$\cdots$	y_{1n}	$\bar{y}_1 = \sum_{j=1}^{n} y_{1j} \big/ n$
2	y_{21}	y_{22}	$\cdots$	y_{2n}	$\bar{y}_2 = \sum_{j=1}^{n} y_{2j} \big/ n$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
L	y_{L1}	y_{L2}	$\cdots$	y_{Ln}	$\bar{y}_L = \sum_{j=1}^{n} y_{Lj} \big/ n$
					$\bar{\bar{y}} = \sum_{i=1}^{L} \bar{y}_i \big/ L$

Table 5.6 summarizes a single-factor ANOVA table. The following symbols are defined. Let SS_B, SS_W and SS_T be *Between Group* sum-of-squares, *Within Group* sum-of-squares and *Total* sum-of-squares, respectively. Note that *df* denotes the degree of freedom that is equal to $L-1$ and $N-L$ for *Between Group* and *Within Group*, respectively. These degrees of freedom are used to calculate mean squares from sum-of-squares. Hence, MS_B and MS_W denote the treatments, mean squares and error mean squares, which can also be interpreted as *explained variance* and *unexplained variance*, respectively, for the *F-Ratio* calculation as follows:

$$\begin{aligned} F_0 &= \frac{\text{Explained Variance}}{\text{Unexplained Variance}} \\ &= \frac{\text{Between groups (levels) variance}}{\text{Within groups (levels) variance}} \\ &= \frac{MS_B}{MS_W}. \end{aligned} \tag{5.23}$$

Table 5.6: ANOVA table for a single-factor model [Mon01].

Source Variation	Sum of Squares	*df*	Mean Square	*F*-Ratio
Between Groups	$SS_B = n\sum_{i=1}^{L}(\bar{y}_i - \bar{\bar{y}})^2$	$L-1$	$MS_B = SS_B/(L-1)$	$F_0 = MS_B/MS_W$
Within Groups	$SS_W = \sum_{i=1}^{L}\sum_{j=1}^{n}(y_{ij} - \bar{y}_i)^2$	$N-L$	$MS_W = SS_W/(N-L)$	
Total	$SS_T = SS_B + SS_W$	$N-1$		

The *F-Ratio* is compared against the *F-critical* value which is defined as

$$F_{Crit.} = F_{\alpha,L-1,N-L}, \tag{5.24}$$

where α is the significance level (for 95% confidence level, $\alpha = 0.05$); $L-1$ and $N-L$ are numerator and denominator degree of freedoms, respectively. $F_{crit.}$ can be obtained from the standard statistical tables or software packages. The decision rule is stated as

$$\begin{aligned} &F_0 > F_{crit.} \Rightarrow \text{reject } H_0, \\ &F_0 \le F_{crit.} \Rightarrow \text{do not reject } H_0. \end{aligned} \tag{5.25}$$

An alternative interpretation of the same information uses the *P*-value that is defined as

$$P = prob(F_{crit.} > F_0), \tag{5.26}$$

where $prob(.)$ is the probability. Since α is chosen as 0.05, which can also be considered as the probability upper bound, the decision rule is formulated as

$$\begin{aligned} &P < (\alpha = 0.05) \Rightarrow \text{reject } H_0, \\ &P \ge (\alpha = 0.05) \Rightarrow \text{do not reject } H_0. \end{aligned} \tag{5.27}$$

In Table 5.7, six F-tests are performed for the six AS, for each hypothesis in (5.21). From the standard F statistic table, $F_{crit.} = F_{0.05,12,507} \approx 1.75$. All the six F-Ratios are

much greater than $F_{crit.}$, so H_0s can be rejected for all six AS. Alternatively, the same conclusion can be reached—that the means for all 13 TCs are not the same at the 95% confidence level for all six AS—by noting that the F-tests are all less than 0.05.

Table 5.7: Fixed-AS ANOVA Table using Experiment 1 data.

AS	Source	Sum of Squares	*df*	Mean Square	*F*-Ratio	*P*-Value
AS1	Between Groups	1690.08	12	140.84	34.28	0.0000
	Within Groups	2083.03	507	4.10853		
	Total	3773.1	519			
AS2	Between Groups	2643.5	12	220.292	63.87	0.0000
	Within Groups	1748.64	507	3.449		
	Total	4392.15	519			
AS3	Between Groups	2425.63	12	202.136	61.09	0.0000
	Within Groups	1677.46	507	3.30861		
	Total	4103.1	519			
AS4	Between Groups	2569.2	12	214.1	100.17	0.0000
	Within Groups	1083.61	507	2.13729		
	Total	3652.81	519			
AS5	Between Groups	2438.17	12	203.18	67.55	0.0000
	Within Groups	1525.02	507	3.00793		
	Total	3963.19	519			
AS6	Between Groups	1456.52	12	121.377	58.57	0.0000
	Within Groups	1050.67	507	2.07234		
	Total	2507.19	519			

In Table 5.8, 13 F-tests are performed for the 13 TCs, each testing hypothesis (5.22). All F-tests show that all *P*-values are less than 0.05. Thus, the alternate hypothesis is accepted in preference of the null hypothesis, which indicates that there is a statistically significant difference between the means of the 6 AS at the 95% confidence level for all 13 TCs. The same conclusion can be obtained by examining the F-values and noting that $F_{crit.} = F_{0.05,5,234} \approx 2.21$. Hence, from the ANOVA analysis tests, it can be concluded that not all NLDs cause the same degree of the distortion perception, and that distortion perception is dependent on the audio stimuli used in the listening test experiment.

Table 5.8: Fixed-TC ANOVA Table using Experiment 1 data.

TC	Source	Sum of Squares	*df*	Mean Square	*F*-Ratio	*P*-Value
ASQRT	Between Groups	77.9837	5	15.5967	10.28	0.0000
	Within Groups	354.984	234	1.51703		
	Total	432.968	239			
ATSR	Between Groups	361.423	5	72.2846	18.33	0.0000
	Within Groups	922.593	234	3.94271		
	Total	1284.02	239			
CLP	Between Groups	132.696	5	26.5391	17.17	0.0000
	Within Groups	361.769	234	1.54602		
	Total	494.465	239			
CUBIC	Between Groups	415.812	5	83.1623	17.87	0.0000
	Within Groups	1089.05	234	4.65407		
	Total	1504.86	239			
EXP2	Between Groups	240.63	5	48.1261	20.79	0.0000
	Within Groups	541.612	234	2.31458		
	Total	782.242	239			
FEXP1	Between Groups	118.353	5	23.6707	13.55	0.0000
	Within Groups	408.662	234	1.74642		
	Total	527.016	239			
FEXP2	Between Groups	203.409	5	40.6818	22.70	0.0000
	Within Groups	419.339	234	1.79205		
	Total	622.748	239			
FWR	Between Groups	546.706	5	109.341	21.68	0.0000
	Within Groups	1180.24	234	5.04377		
	Total	1726.95	239			
HWR	Between Groups	527.427	5	105.485	34.87	0.0000
	Within Groups	707.889	234	3.02516		
	Total	1235.32	239			
NSIG	Between Groups	123.353	5	24.6705	10.98	0.0000
	Within Groups	525.687	234	2.24653		
	Total	649.04	239			
NTANH	Between Groups	81.2107	5	16.2421	7.16	0.0000
	Within Groups	530.659	234	2.26777		
	Total	611.87	239			
SQL	Between Groups	391.902	5	78.3805	16.67	0.0000
	Within Groups	1100.26	234	4.70196		
	Total	1492.16	239			
SQS	Between Groups	347.715	5	69.543	15.87	0.0000
	Within Groups	1025.69	234	4.38327		
	Total	1373.4	239			

5.5.6 Nonlinear regression for *Rnonlin*

In this section, the statistical nonlinear regression model fitting and validation work are presented. The goal is to verify that the *Rnonlin* model can be used to predict perceptual distortion created by different zero-memory NLDs. From the ANOVA

result it was concluded that the mean scores of each TC of the six AS are not equal in the statistical sense. Hence, an individual nonlinear regression analysis is required for each and every AS.

Figure 5.8 shows the conceptual block diagram illustrating the data analysis sequences involved in nonlinear regression statistical model building and correlation analysis for the *Rnonlin* perceptual model. Note that this work involves two stages of cascade modeling, namely perceptual distortion score computing and statistical nonlinear regression. The *Rnonlin* model, labeled as Block (a) in Figure 5.8, takes in both clean and processed signals for each TC. Since the listening tests were conducted using stereo headphones, *Rnonlin* scores are computed separately for left and right channels, each using the process depicted in Figure 5.8. The average *Rnonlin* score for both channels is used for each TC.

Block (b) in Figure 5.8 represents the listening test data preparation step. Opinion scores (OS) from the listening test were split into 60 % training data and 40% validation data. The training data set is used to build the nonlinear regression model and the validation data set is used to verify whether statistical modeling has over-fit the data. The model that is built using the training data set might also capture statistical noise and systematic components of the model. Thus, to assess this risk of over-fitting, the validation data set is used to access the model fitness in terms of correlation analysis.

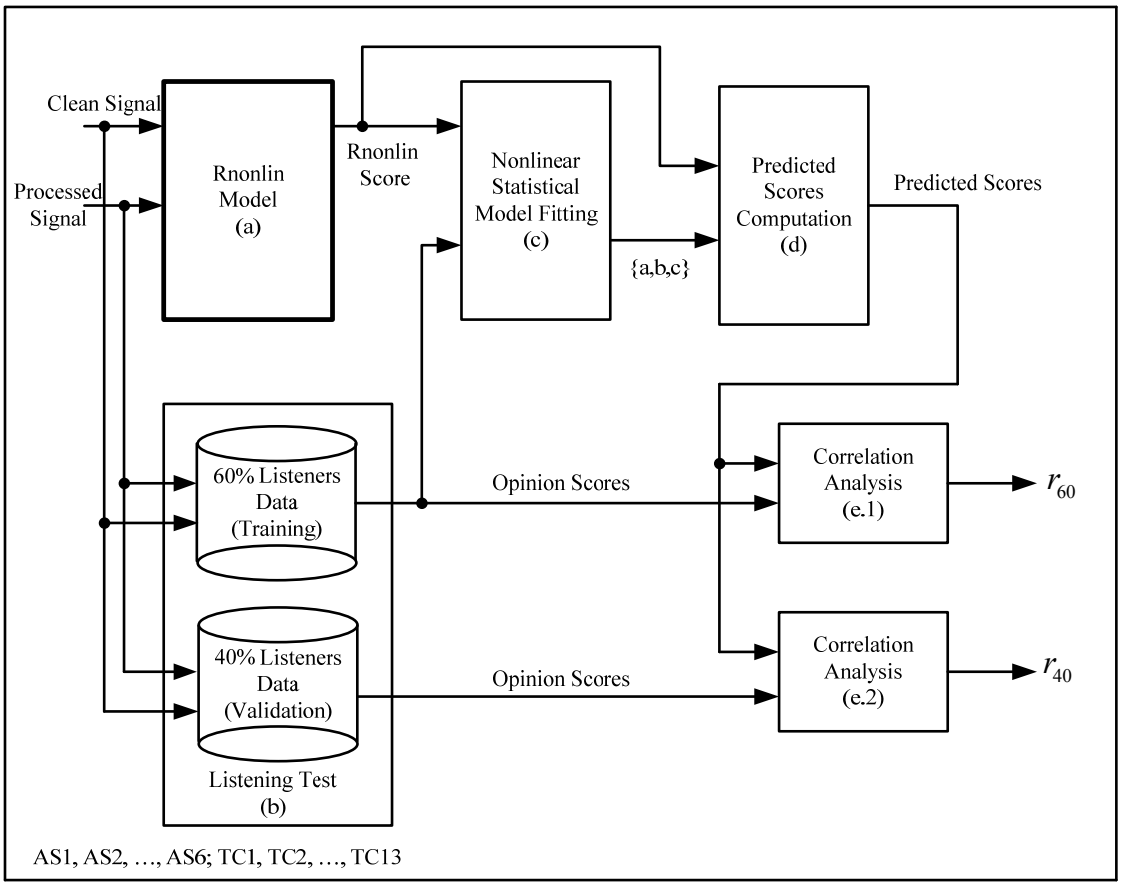

Figure 5.8: A block diagram describing the statistical data analysis steps from the *Rnonlin* model to nonlinear regression model curve-fitting to correlation analysis. The data flow in this diagram is repeated for six times for the six AS. The accompanying data preparation steps are illustrated in Figure 5.9.

For every statistical analysis the data preparation step is crucial. Figure 5.9 shows the data preparation steps for a single AS, a process repeated over all AS. The uppermost table in Figure 5.9 represents the $OS_{i,j}$ data set obtained from the listening test, where i represents the subjects' index from 1 to 40 and j represents the TC from 1 to 13. Therefore the $OS_{i,j}$ table consists of 40 × 13 entries.

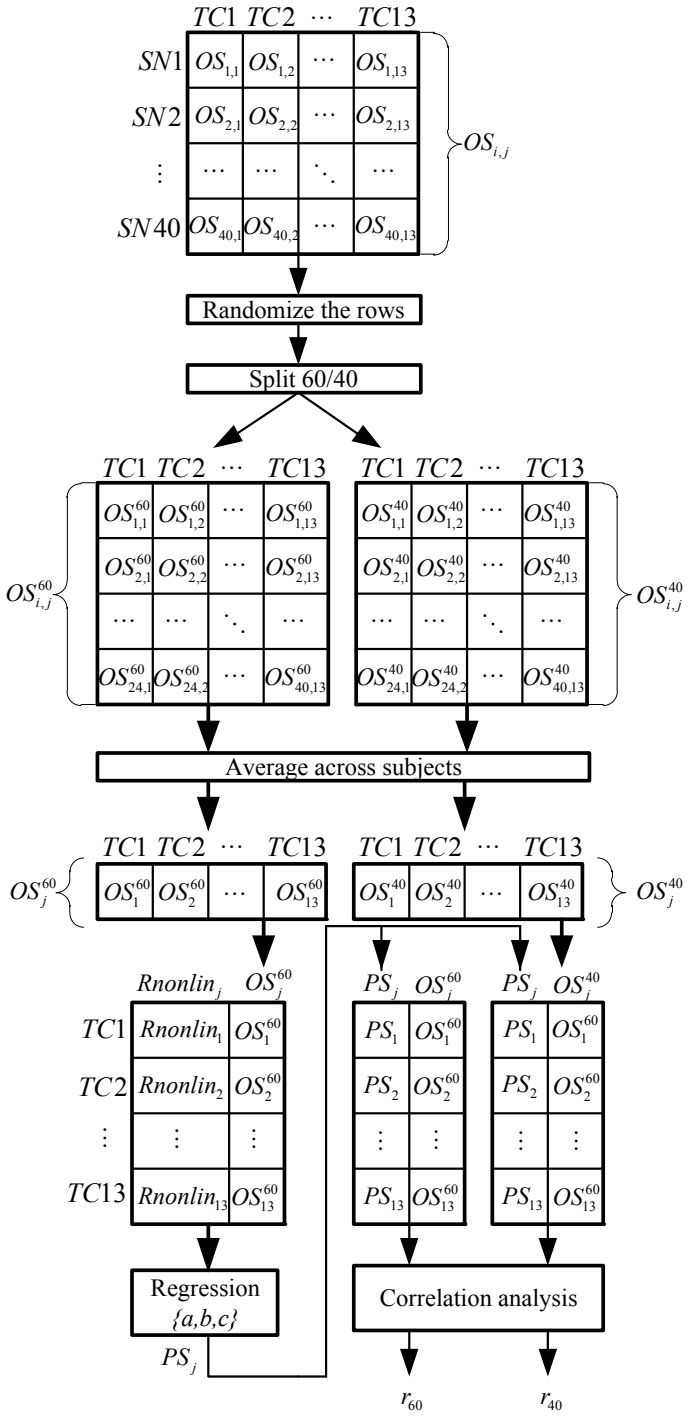

Figure 5.9: Data preparation steps and data structures used in Rnonlin regression model analysis and correlation analysis and model validation process.

Next to ensure observations recorded in rows are independent, the rows in the AS data set are shuffled and split into 60 % training data set, $OS_{i,j}^{60}$, and 40 % validation data set, $OS_{i,j}^{40}$. The OS data records in both training data set and validation data set are then averaged across all subjects and the resultant values stored in two data sets namely, OS_j^{60}, and OS_j^{40} for $j = 1,\ldots,13$. The averaged scores for the training data set, OS_j^{60}, is augmented with the *Rnonlin* data set, $Rnonlin_j$, as shown in the bottom left block of Figure 5.9, to form the final data preparation step. Note that $Rnonlin_j$ denotes the *Rnonlin* scores for the 13 TCs obtained from the *Rnonlin* model as shown in the Block (a) of Figure 5.8. For the nonlinear regression, the dependent variable is OS_j^{60} and the independent variable is $Rnonlin_j$. Figure 5.8 shows the nonlinear regression model curve-fitting block labeled as Block (c) where the inputs are $Rnonlin_j$ as computed by the *Rnonlin* model and the opinion scores, OS_j^{60}.

The scatter plots and their respective superimposed regression curves for the six AS are plotted in Figure 5.10. The nonlinear empirical equation [TMZ04] used in fitting the regression model is expressed as,

$$Score = \frac{a + b \times Rnonlin^c}{a + Rnonlin^c}, \tag{5.28}$$

where a, b, and c are the regression coefficients, *Rnonlin* is the number computed by the *Rnonlin* model and *Score* is the averaged score obtained from the experiment. Note that *Rnonlin* ranges from 0 to 1 and *Score* ranges from 1 to 10. The Levenberg-Marquardt nonlinear least square algorithm [SW03] was chosen to

compute the estimated regression coefficients, $\{a, b, c\}$. Table 5.9 presents the nonlinear regression coefficients, asymptotic 95% confidence intervals of the coefficients, and goodness-of-fit measures, such as R^2 and R^2_{adj} for the six different AS. The R^2 statistics indicate how much the fitted model explains the percentage variability in *Score*. The adjusted R^2 statistics, R^2_{adj}, which are more suitable for comparing models with different numbers of independent variables, also indicates the percentage variability in *Score*. In Table 5.9, R^2 and R^2_{adj} are presented in percentage. Higher percentage values of R^2 and R^2_{adj} results in better fitting of the model to the experimental data.

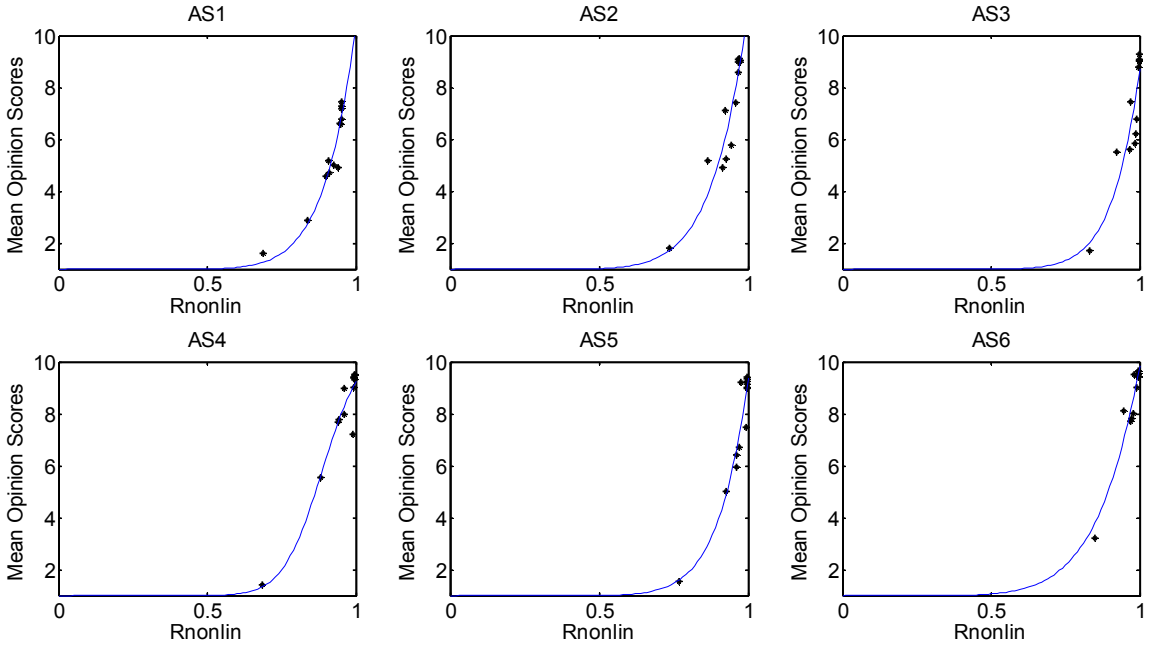

Figure 5.10: Mean subjective opinion scores from Experiment 1 plotted against the computed distortion scores, *Rnonlin* for the six AS.

Table 5.9: Nonlinear regression coefficients, confidence interval and goodness of fit results for the six AS.

AS	Parameter	Estimate	95% Confidence Interval		R^2	R^2_{adj}
			Lower	Upper		
AS1	*a*	60355.8	55868.9	64842.8	91.5%	89.8 %
	b	567758.	567758.	567758.		
	c	9.47863	9.02134	9.93592		
AS2	*a*	59352.6	54561.5	64143.7	89.2%	87.0%
	b	599961.	599961.	599961.		
	c	8.57878	7.57847	9.57909		
AS3	*a*	47294.0	42468.6	52119.4	80.9%	77.1%
	b	364875.	364875.	364875.		
	c	10.8835	10.0142	11.7528		
AS4	*a*	0.253971	-0.464548	0.972491	93.0%	91.6%
	b	11.3465	4.95164	17.7413		
	c	12.5454	-2.20465	27.2954		
AS5	*a*	159.941	-66486.4	66806.3	91.2%	89.5%
	b	1352.27	-552739.	555444.		
	c	9.98696	-8.00839	27.9823		
AS6	*a*	-70711.6	-73915.8	-67507.3	91.6%	89.9%
	b	-626481.	-626481.	-626481.		
	c	7.03893	6.3548	7.72305		

After the model has been built and the estimated regression coefficients $\{a,b,c\}$ obtained, the next step is to compute the prediction scores, PS_j, for $j=1,2,\ldots,13$ using the regression equation in (5.28). This process is represented by Block (d) in Figure 5.8. Inputs to Block (d) in Figure 5.8 are $Rnonlin_j$ and the model coefficients $\{a,b,c\}$. The output from Block (d) is PS_j. The scatter plots of mean-opinion scores versus predicted scores using both training data and validation data sets are plotted in Figures 10 and 11, respectively.

The final step of the statistical analysis of the *Rnonlin* model and verification is the correlation analysis which is illustrated as a Blocks (e.1) and (e.2) in Figure 5.8. There are two correlation analyses: (1) between the model-predicted scores PS_j and

the opinion scores OS_j^{60} from training data set and (2) between the model-predicted scores PS_j and the opinion scores OS_j^{40} from the validation data set. The correlation coefficient between PS_j and OS_j^{60} is denoted as r_{60} and that between PS_j and OS_j^{40} is denoted as r_{40}. Note that OS_j^{60} is used to build the regression model while PS_j is obtained from the model using $Rnonlin_j$. OS_j^{40} is a validation data set and is not used in the model building process. The purpose of this analysis is to assess model accuracy by correlating the OSs with the predicted scores using both training and validation data sets. If the correlation scores using the validation data set does not drop significantly, it can safely be claimed that the model does not over-fit the data and can be used for predicting perceived distortion without running the subjective listening test.

The correlation coefficients r_{60} and r_{40} are computed as follows:

$$r_{60} = \frac{Cov\left(PS_j, OS_j^{60}\right)}{\sqrt{Var\left(PS_j\right)Var\left(OS_j^{60}\right)}}, \tag{5.29}$$

$$r_{40} = \frac{Cov\left(PS_j, OS_j^{40}\right)}{\sqrt{Var\left(PS_j\right)Var\left(OS_j^{40}\right)}}, \tag{5.30}$$

where $Cov(X,Y)$ denotes the covariance between the variable X and Y and $Var(Z)$ denotes the variance of the variable Z. r_{60} and r_{40} are also known as product-moment correlation coefficients or Pearson correlation coefficients [BCB03] [HWJ03]. The values of these correlation coefficients always lie between ±1. A value of correlation coefficient close to +1 indicates a strong positive linear relationship,

while a value close to –1 indicates a strong negative linear relationship resulting in inverse proportionality between variables. A value of correlation coefficient closes to 0 represents neither linear nor nonlinear relationship between variables. The r_{60} and r_{40} values for the six AS are listed in the second column of Table 5.10.

Table 5.10: Correlation analysis results for the training data set and the validation data set.

Training Data Set (60%) (Sub-table A)					
AS	r_{60}	t statistics	*P*-value	95% LCL	95% UCL
1	0.9577	11.0333	0.0000	0.8610	0.9876
2	0.9448	9.5644	0.0000	0.8214	0.9837
3	0.8997	6.8375	0.0000	0.6916	0.9699
4	0.9646	12.1278	0.0000	0.8827	0.9896
5	0.9551	10.6949	0.0000	0.8531	0.9868
6	0.9574	10.9910	0.0000	0.8600	0.9875
Validation Data Set (40%) (Sub-table B)					
AS	r_{40}	t statistics	*P*-value	95% LCL	95% UCL
1	0.9303	8.4090	0.0000	0.7781	0.9793
2	0.8929	6.5764	0.0000	0.6730	0.9678
3	0.8621	5.6437	0.0002	0.5927	0.9580
4	0.9592	11.2581	0.0000	0.8659	0.9880
5	0.9816	17.0280	0.0000	0.9377	0.9946
6	0.9459	9.6715	0.0000	0.8248	0.9840
Correlation Difference Between Training Data Set and Validation Data Set (Sub-table C)					
AS	$r_{60} - r_{40}$	z statistics	*P*-value	95% LCL	95% UCL
1	0.0274	0.5739	0.5660	-0.5510	0.8121
2	0.0519	0.7716	0.4403	-0.4865	0.8401
3	0.0376	0.3783	0.7052	-0.6090	0.7801
4	0.0053	0.1601	0.8728	-0.6668	0.7389
5	-0.0264	-1.0090	0.3130	-0.8687	0.4014
6	0.0114	0.2722	0.7855	-0.6380	0.7609

After obtaining the correlation coefficients, the next statistical procedure is to test whether there is any linear relationship between the variables in the population as a whole. That is, the sample correlation r estimates the population correlation ρ. This type of statistical test is known as inferential statistics and uses hypothesis testing of the correlation technique. For the two-tail test, hypotheses are: $H_0 : \rho = 0$

versus $H_a : \rho \neq 0$, where H_0 is the null hypothesis and H_a is the alternative hypothesis. Test statistics is given by

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}, \tag{5.31}$$

where r is the correlation coefficient representing r_{60} or r_{40} and n is the number of data points (n = 13, in this case). A basic assumption is that test statistics are distributed as a t-random variable with $n-2$ degrees of freedom. Thus, for a test with significance level $\alpha = 0.05$, the critical value is $t_{\alpha/2,n-2} = t_{0.025,11} = 2.2$ from the standard t table. The decision rule is given by

$$\begin{aligned} |t| > t_{\alpha/2,n-2} &\Rightarrow \text{reject } H_0, \\ |t| \leq t_{\alpha/2,n-2} &\Rightarrow \text{do not reject } H_0. \end{aligned} \tag{5.32}$$

Notice that the *t*-values are listed in the third column of Table 5.10, which are all greater than the critical value of 2.2; hence, it can be concluded that the null hypotheses are all rejected and the linear relationships between the predicted scores and the opinion scores are strong. Alternatively, *P*-values show that they are all less than the significance level of 0.05 which also points to the strong linear relationship between the predicted scores and the OS.

The subsequent statistical test for the correlation analysis is the confidence interval computation. Although the hypothesis test indicates that there is a linear relationship, it gives no indication of the confidence interval. However, this additional information can be obtained from a confidence interval for the population coefficient.

To calculate this confidence interval, r must be transformed into a normal distribution using the Fisher z-transformation,

$$z = \frac{1}{2}\ln\frac{1+r}{1-r} = \arctan(r), \tag{5.33}$$

where r is the sample correlation coefficient representing r_{60} or r_{40} for each AS. The standard error of z is approximately $1/\sqrt{n-3}$. Thus the 95% confidence interval for the true population value for the transformed correlation coefficient z is given by

$$z \pm 1.96 \times 1/\sqrt{n-3}. \tag{5.34}$$

Since z has normal distribution, 1.96 standard deviations from the mean results in a 95% confidence level. However, the inverse of the Fisher transformation must be used for the lower and upper limits of the confidence interval to obtain the 95% confidence interval for the correlation coefficient. That is, the inverse Fisher transformation is given as

$$r = \frac{\exp(2z)-1}{\exp(2z)+1}, \tag{5.35}$$

where z is the transformed coefficient and r the correlation coefficient that can be used to construct a confidence interval for the population correlation parameter ρ. The upper and lower limits of the 95% confidence interval for the training data set and the validation data set are listed in the fifth and sixth columns in Sub-tables A and B of Table 5.10.

Sub-table C in Table 5.10 shows the statistical test results of the correlation coefficients' difference between training and validation data sets. The second column of this section shows the difference between the correlation coefficients, i.e.

$r_{60-40} = r_{60} - r_{40}$. To determine whether a difference is significance, the hypothesis test is repeated. Since the hypothesis test is inferential statistics, the null hypothesis is formulated such that there is no difference between the correlation coefficients for the two populations, i.e. training and validation. The null hypothesis will therefore be tested against the non-directional alternative hypothesis. The hypotheses are: $H_0 : \rho_{60} = \rho_{40}$ versus $H_a : \rho_{60} \neq \rho_{40}$. Note that ρ_{60} and ρ_{40} are population correlation coefficient parameters corresponding to sample statistics r_{60} and r_{40}. To test this hypothesis, the first step is to convert the sample correlation coefficients into z using the Fisher transformation as follows: $z_{60} = \arctan(r_{60})$ and $z_{40} = \arctan(r_{40})$. The z statistics in the third column of the Sub-table C of the Table 5.10 is computed as

$$z_{60-40} = \frac{z_{60} - z_{40}}{\sqrt{1/(n_{60} - 3) + 1/(n_{40} - 3)}}, \tag{5.36}$$

where n_{60} and n_{40} are the number of observations in training and validation data sets and equal 13. For $\alpha = 0.05$ significance level the decision rule is governed by

$$\begin{aligned} &|z_{60-40}| > z_{\alpha/2} \text{ or } |z_{60-40}| > 1.96 \Rightarrow \text{reject } H_0, \\ &|z_{60-40}| \leq z_{\alpha/2} \text{ or } |z_{60-40}| \leq 1.96 \Rightarrow \text{do not reject } H_0. \end{aligned} \tag{5.37}$$

Notice that all the absolute values of the z statistics in the third column of the sub-table C are less than 1.96, hence the null hypotheses for all six AS are rejected, concluding that there is no difference between the correlation coefficients of the training and validation data sets. Similar conclusion can be obtained from the *P*-values analysis. Therefore, the obtained statistical models listed in Table 5.10 do not over-fit the data.

The last statistics test is the 95% confidence interval for the difference in the correlation coefficients. This is carried out by the following procedures. z_{60} and z_{40} are computed using the Fisher transformation, the 95% confidence interval in z is

$$\left(z_{60}-z_{40}\right)\pm 1.96\sqrt{1/\left(n_{60}-3\right)+1/\left(n_{40}-3\right)}, \tag{5.38}$$

where $\sqrt{1/\left(n_{60}-3\right)+1/\left(n_{40}-3\right)}$ is the standard error. Using the inverse Fisher transformation described in (5.35), the upper and lower limits of the 95% confidence interval for the correlation coefficient differences are computed. The results are listed in the fifth and sixth columns of Section C of Table 5.10.

From this correlation analysis, it can be concluded that there is a strong linear relationship between the predicted scores obtained from the *Rnonlin* model and those scores obtained from the listening test, as verified by using separate validation data sets. The differences in correlation coefficients between training and validation data sets are found to be statistically insignificant. Hence it may be concluded that the *Rnonlin* model can be used to predict the degree of perceived distortion produced by different NLDs, and that the estimated nonlinear regression coefficients can be derived so as to predict distortion scores for any zero-memory NLD under any of the auditory stimuli reported in this chapter. These findings are carried forward into the VBS experimental domain to assess NLD suitability.

5.6 Experiment 2: Bass-Intensity Listening Test

Experiment 2 is carried out to assess the bass intensity of the NLDs selected from the results obtained in Experiment 1. The results of Experiment 1 have shown that not all NLDs have the same perceived distortion, and the *Rnonlin* model can be used to predict the distortion scores. Based on the distortion scores obtained in Experiment 1, five NLDs (ASQRT, FWR, SQL, CUBE and SQL) are found to have poor distortion scores and were discarded in Experiment 2. The remaining eight NLDs are used in Experiment 2 to test their reproduced bass intensity. The following subsections present the detailed experiment design and the results obtained from this experiment.

5.6.1 Audio materials

Table 5.11 presents the three audio files used in Experiment 2, each of which contains significant bass frequency content. A speech signal was not included in this experiment, as the main objective is to evaluate only the bass-intensity enhancement of the selected NLDs. The genre of the music signals are country, hip-hop and pop with special emphasis on instrumental drums and synthesized bass.

Table 5.11: Audio materials selected for the Experiment 2.

AS	Title	Band/Single	Seconds	Genres
1	Hotel California	Eagles	5	Country
2	Superman	Eminem	6	Hip-hop
3	Always	Atlantic Starr	8	Pop

5.6.2 Processing of audio materials

A total of 11 TCs were generated based on the *Multiple Stimuli with Hidden Reference and Anchor* (MUSHRA) methodology [ITU01]. These TCs consist of eight

NLD-processed soundtracks, one equalized (+12 dB gain for frequency < 250 Hz) soundtrack, one original reference signal, and one anchor (500 Hz high-pass filtered) signal. These TCs were performed over the three AS listed in Table 5.11. In this experiment, all processed audio files are high-pass filtered by HPF2 (see Figure 5.1(b)) at 100 Hz to evaluate the virtual-bass effect generated by different NLDs.

5.6.3 Software and Hardware Experimental Configuration

Figure 5.11 shows the GUI used in Experiment 2. Instead of the conventional scale of 0 to 100 commonly found in MUSHRA software [Vin05], the scale ranged from –3 to +3, where 0 indicates no bass enhancement when compared to the reference signal (REF). The negative score represents "lesser bass" and the positive score represents "more bass" intensity. Listeners could compare the performance of each TC with the reference and the anchor signals together with other TCs to give an overall bass-intensity rating. Two-stage double-blind randomization was performed with the MUSHRA software while the hardware configuration is the same as described in Experiment 1.

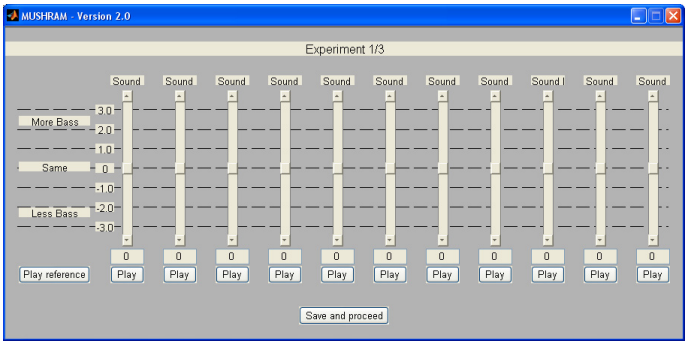

Figure 5.11: Graphical user interface used during Experiment 2 [Vin05].

5.6.4 ANOVA analysis

A total of fifty subjects participated in Experiment 2. The obtained mean scores for the three AS with 95% confidence intervals are sorted in descending order and listed in Table 5.12. From Table 5.12, ATSR scores the highest, followed by EXP2, FEXP1 and NTANH. All listeners consistently identified the anchor (ANC), which is rated as the lowest among all TCs. Note that among the TCs: NSIG, EQ and REF scores are close to zero, while CLP, HWR and FEXP2 had negative bass-intensity scores.

Table 5.12: Mean scores (MS), upper confidence bounds (UCBs) and lower confidence bound (LCBs) of six audio sequences from Experiment 2 data. MSs are sorted in descending order.

AS	TC	MS	LCB	UCB	AS	TC	MS	LCB	UCB
1	ATSR	2.68	2.54	2.82	3	ATSR	2.73	2.58	2.88
	EXP2	1.81	1.62	2.00		EXP2	1.77	1.57	1.97
	FEXP1	1.18	1.03	1.33		FEXP1	1.20	1.02	1.38
	NTANH	0.70	0.54	0.86		NTANH	0.80	0.63	0.97
	NSIG	0.24	0.03	0.45		NSIG	0.30	0.09	0.51
	EQ	0.13	-0.07	0.33		EQ	0.26	0.09	0.43
	CLP	-0.05	-0.24	0.14		REF	0.10	-0.07	0.27
	REF	-0.16	-0.33	0.01		CLP	-0.21	-0.38	-0.04
	HWR	-0.19	-0.64	0.26		FEXP2	-1.14	-1.30	-0.98
	FEXP2	-0.83	-1.05	-0.61		HWR	-1.63	-1.78	-1.48
	ANC	-2.88	-2.96	-2.80		ANC	-2.91	-3.00	-2.82
2	ATSR	2.77	2.65	2.89					
	EXP2	1.83	1.64	2.02					
	FEXP1	1.18	0.98	1.38					
	NTANH	0.99	0.79	1.19					
	EQ	0.42	0.26	0.58					
	NSIG	0.37	0.17	0.57					
	REF	0.04	-0.12	0.20					
	CLP	-0.38	-0.57	-0.19					
	FEXP2	-0.92	-1.11	-0.73					
	HWR	-1.34	-1.60	-1.08					
	ANC	-2.93	-2.99	-2.87					

Two hypotheses were tested with two single-factor ANOVA tests, such as the fixed-AS and the fixed-TC. For a fixed AS (i.e., for each $AS1, AS2, AS3$),

$$
\begin{aligned}
H_0 &: \mu_{TC1} = \mu_{TC2} = \ldots = \mu_{TC11} \quad \text{(i.e. means of all levels are equal),} \\
H_a &: \mu_{TCi} \neq \mu_{TCj} \text{ for } i \neq j \in \{1,2,\ldots,11\} \quad \text{(i.e. at least two means are not equal).}
\end{aligned} \tag{5.39}
$$

For a fixed TC (i.e., for each $TC1, TC2, \ldots, TC11$),

$$
\begin{aligned}
H_0 &: \mu_{AS1} = \mu_{AS2} = \mu_{AS3} \quad \text{(i.e. means of all levels are equal),} \\
H_a &: \mu_{ASi} \neq \mu_{ASj} \text{ for } i \neq j \in \{1,2,3\} \quad \text{(i.e. not all the means are equal).}
\end{aligned} \tag{5.40}
$$

The fixed-AS ANOVA performance listed in Table 5.13 confirms that there is a statistically significant difference between the means of the 11 variables (TCs in this case) at the 95% confidence level, since the *P*-values of the *F*-tests are less than 0.05 for all three AS.

Table 5.13: Fixed-AS ANOVA Table using Experiment 2 data.

AS	Source	Sum of Squares	*df*	Mean Square	*F*-Ratio	*P*-Value
AS1	Between groups	1041.7	10	104.17	182.53	0.0000
	Within groups	307.605	539	0.570696		
	Total	1349.31	549			
AS2	Between groups	1235.29	10	123.529	298.27	0.0000
	Within groups	223.225	539	0.414147		
	Total	1458.52	549			
AS3	Between groups	1257.77	10	125.777	363.22	0.0000
	Within groups	186.645	539	0.34628		
	Total	1444.42	549			

Table 5.14 shows the fixed-TC ANOVA model used to test the validity of whether the use of different AS cause any statistically significant differences in means between the three variables (AS in this case) observed for each TC. The *P*-values for CLP and HWR are 0.0413 and 0, respectively, with the remainder greater than 0.05. Hence, contrary to the findings of Experiment 1, different AS do not influence bass perception, even though the music tracks contained significant bass frequency content.

Table 5.14: Fixed-TC ANOVA Table using Experiment 2 data.

TC	Source	Sum of Squares	*df*	Mean Square	*F*-Ratio	*P*-Value
REF	Between groups	1.85333	2	0.926667	2.61	0.0767
	Within groups	52.14	147	0.354694		
	Total	53.9933	149			
ANCHOR	Between groups	0.0633333	2	0.0316667	0.40	0.6709
	Within groups	11.63	147	0.0791156		
	Total	11.6933	149			
EQ	Between groups	2.11	2	1.055	2.76	0.0666
	Within groups	56.205	147	0.382347		
	Total	58.315	149			
ATSR	Between groups	0.203333	2	0.101667	0.44	0.6459
	Within groups	34.09	147	0.231905		
	Total	34.2933	149			
CLP	Between groups	2.72333	2	1.36167	3.26	0.0413
	Within groups	61.45	147	0.418027		
	Total	64.1733	149			
EXP2	Between groups	0.0933333	2	0.0466667	0.10	0.9018
	Within groups	66.355	147	0.451395		
	Total	66.4483	149			
FEXP1	Between groups	0.0133333	2	0.00666667	0.02	0.9830
	Within groups	57.26	147	0.389524		
	Total	57.2733	149			
FEXP2	Between groups	2.54333	2	1.27167	2.88	0.0596
	Within groups	65.005	147	0.442211		
	Total	67.5483	149			
HWR	Between groups	58.0033	2	29.0017	24.08	0.0000
	Within groups	177.07	147	1.20456		
	Total	235.073	149			
NSIG	Between groups	0.423333	2	0.211667	0.40	0.6684
	Within groups	77.025	147	0.52398		
	Total	77.4483	149			
NTANH	Between groups	2.17	2	1.085	2.69	0.0711
	Within groups	59.245	147	0.403027		
	Total	61.415	149			

5.7 Selecting the best NLD for VBS

In the context of VBS, distortion can be classified as *desirable* or *undesirable*. A *desirable* distortion can create the sensation of a warmer, richer and pleasing sound quality by subjectively extending the low-frequency bandwidth, while *undesirable* distortion induces perceptually unpleasant and discordant artifacts. This distortion classification and the subsequent choice of NLD relates directly to harmonic and intermodulation distortion generation as discussed in Section 0 and to the research

Questions 1 and 2. Based on results from the two subjective listening experiments, NLDs can now be classified into four sub-classes of *Good*, *Bass-Killer*, *Not Recommended* and *Highly Distorted*, as shown in Tables 5.14 and 5.15. The *Good* NLDs are chosen based on their good bass enhancement with minimum *undesirable* distortion, while the *Bass-Killer* NLDs suppress bass perception. The *Not Recommended* NLDs produce modest bass enhancement together with *undesirable* distortion while the *Highly Distorted* NLDs just produce a highly distorted bass effect.

For the selected NLD transfer characteristics, Table 5.15 presents their second derivatives, limits of first derivatives and harmonic generation; the nonlinear functions and their first derivatives have already been listed in Table 5.1. $f'(x)$ and $f''(x)$ denote the first derivative and second derivative of the nonlinear function $f(x)$, respectively. $x = (0, +1]$ denotes that x is a real number larger than 0 but less than or equal to 1. $\lim_{x \to 0, x>0} f'(x)$ and $\lim_{x \to 0, x<0} f'(x)$ denote the right-hand limit and the left-hand limit of the first-derivative function $f'(x)$, respectively. Some of the nonlinear functions are discontinuous at the origin for the first-derivative function, so both left-hand and right-hand limits are considered for these nonlinear functions. In Figure 5.12 the first-derivative curves of the thirteen NLDs are shown, grouped under the four classes.

In addition there are three types of symmetries {even (E), odd (O) and both even and odd (B)} generation listed in Table 5.15. Table 5.16 summarizes the main characteristics of the four classes of NLDs in terms of trends in their second derivatives and harmonic generation.

Table 5.15: Four classes of NLDs from 13 NLDS under study.

Class	NLD	$f''(x)$	$\lim_{x\to 0} f'(x)$	Harmonic
Good	EXP2	$\frac{-e^{1-x}}{e-1}$	$f'(0)$	B
	ATSR	$\frac{-3.6450x}{(1+0.81x^2)^2} - \frac{1.64025x^2}{(1-0.81x^2)^{3/2}} - \frac{2.025}{\sqrt{1-0.81x^2}}$	$f'(0)$	B
	FEXP1	$\frac{-e^{-\lvert x\rvert}}{1-e^{-1}}, x \neq 0$	$+\infty$	O
	NSIG	$\frac{e^x(e+1)}{(e^x+1)(e-1)} - \frac{2e^{2x}(e+1)}{(e^x+1)^2(e-1)} + \frac{2e^{2x}(e^x-1)(e+1)}{(e^x+1)^3(e-1)} - \frac{e^x(e^x-1)(e+1)e^x}{(e^x+1)^2(e-1)}$	$f'(0)$	O
	NTANH	$\frac{-2(e^x-e^{-x})(e+e^{-1})}{(e-e^{-1})(e^x+e^{-x})} + \frac{2(e^x-e^{-x})^3(e+e^{-1})}{(e-e^{-1})(e^x+e^{-x})^3}$	$f'(0)$	O
Bass-Killer	FEXP2	$\frac{e^{\lvert x\rvert}}{e-1}, x \neq 0$	$-\infty$	O
	SQS	2	$+\infty$	O
	SQL	2	$f'(0)$	E
	CUBE	$6x$	$f'(0)$	O
Not Rec-commended	HWR	$0, x \neq 0$	$\lim_{x\to 0,x>0} f'(x) = +1$ $\lim_{x\to 0,x<0} f'(x) = 0$	E+f_0
	FWR	$0, x \neq 0$	$\lim_{x\to 0,x>0} f'(x) = +1$ $\lim_{x\to 0,x<0} f'(x) = -1$	E
	CLP	0	$f'(0)$	O
Highly Distorted	ASQRT	$\frac{-\mathrm{sgn}^2(x)}{4\lvert x\rvert^{3/2}}, x \neq 0$	$\lim_{x\to 0,x>0} f'(x) = +\infty$ $\lim_{x\to 0,x<0} f'(x) = -\infty$	E

In line with the discussion in Section 5.1 on VBS requirements, Table 5.16 reveals that even-only harmonic NLDs are inappropriate for virtual bass enhancement, whereas NLDs that generate both even and odd harmonics or odd-only harmonics yield good bass enhancement. However, considering just the types of harmonics generated is insufficient to differentiate between *Good* and *Bad* NLDs, as intermodulation distortion must be considered. Interestingly, all *Good* NLDs also exhibit a second derivative in their transfer characteristic of less than zero over the interval $x = (0, +1]$. This trend can clearly be seen in the top row of Figure 5.12, which shows negative second derivatives for all *Good* NLDs.

Table 5.16: Four classes of NLDs and the second-derivative pattern and harmonic generation pattern. E: even; O: odd; B: both even and odd.

		Harmonics		
Class	Derivative/Limit Pattern	E	O	B
Good	$f''(x) < 0, x \in (0, +1]$		●	●
Bass-Killers	$f''(x) > 0, x \in (0, +1]$	●	●	
Not Recommended	$f''(x) = 0, x \in (0, +1]$	●	●	
Highly Distorted	$\lim_{x \to 0, x > 0} f'(x) = +\infty, \lim_{x \to 0, x < 0} f'(x) = -\infty$	●	●	

Bass-Killers, Not-Recommended and *Highly-Distorted* NLDs are all considered *Bad*. From Table 5.16 all three groups contain either even or odd symmetries but critically, Figure 5.12 shows that for *Bass-Killer* NLDs, the second derivatives are greater than zero, the *Not-Recommended* NLDs have second derivatives exactly zero over the interval $x = (0, +1]$, while *Highly-Distorted* NLDs have discontinuous first derivatives where left- and right-hand limits are infinite with opposite signs (i.e. $\pm\infty$).

Therefore by examining the harmonics generated by NLDs and taking their second derivatives as presented in Table 5.16, a good performance set of NLDs can be identified for virtual bass enhancement. In addition, when coupled with the *Rnonlin* model the perceived distortion generated from different types of zero-memory NLDs can be quantified, using disparate genres of musical soundtracks.

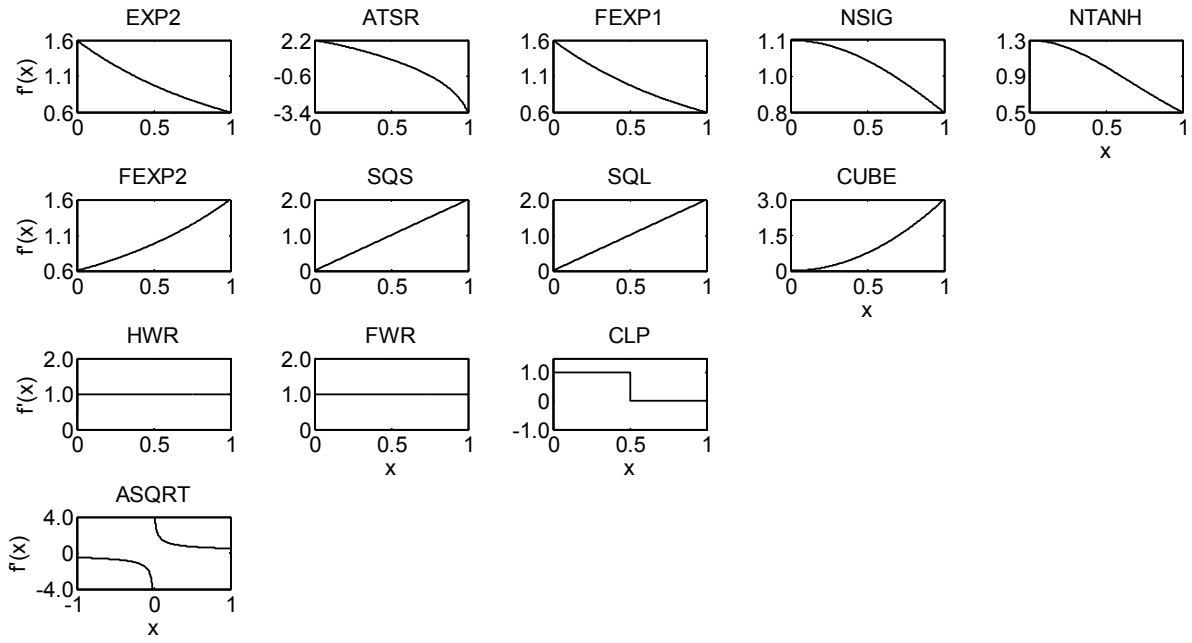

Figure 5.12: The first-derivative function plots of thirteen NLDs.

5.8 Conclusions

The three research questions posed concerning the usage and selection of NLDs in VBS in terms of bass-intensity perception and distortion perception have been addressed. The objective metrics were defined and computed across seven NLDs. The results were compared and found to contradict one another in their predictions of nonlinear distortion. On the other hand, a statistical correlation analysis confirmed that the *Rnonlin* model can be used to predict perceived distortion generated by NLD-specific VBS reliably. The ANOVA test showed that the amount of perceived distortion is highly dependent on the selected AS, whereas the ANOVA test applied to bass intensity derived from using different NLDs is not greatly dependent on the

selected AS. Therefore, as long as there is bass content in the AS, a given NLD would be expected to produce a consistent bass enhancement performance.

From the research findings discussed in the previous section, even-only harmonic NLDs are inappropriate for virtual bass enhancement, whereas NLDs that generate both even and odd harmonics or odd-only harmonics are considered suitable candidates for good bass enhancement. However, there is still danger of falling within the category of "bass killer" when the second derivative of the NLD curve is greater than zero on the input interval 0 to 1. Therefore, to be considered a *Good* NLD for virtual bass system, we conclude an NLD should not be even–symmetric, and that its second derivative should be less than zero in the input interval 0–1.

In addition, some general performance trends have been identified, derived from inspection of the input-output transfer characteristics of a set of zero-memory NLDs. These observations correlate well with the results obtained for the set of *Good* NLDs that return a good bass performance with minimal perceived distortion. These trends should therefore serve as a useful guide in selecting suitable NLDs for VBS.

CHAPTER 6

CONCLUSIONS AND FUTURE WORKS

6.1 Conclusions

The background, motivations, and objectives of this dissertation on NLD-based VBS for psychoacoustic bass enhancement were introduced in Chapter 1. The main conclusions of this dissertation are presented as follows:

- Analytical and computational techniques, HASP and YOGA, were developed to evaluate nonlinear distortion caused by the NLDs, and to evaluate traditional distortion metrics such as THD, THD_{MT}, and $TIMD_{MT}$.
    - The GedLee metric was included for cross-comparisons among these objective distortion metrics.
    - The *Rnonlin* psychoacoustic nonlinear distortion perceptual model was used, as a more perception-focused metric was needed; it was cross-validated with subjective listening test data and found to be suitable for VBS's nonlinear distortion modeling.
- Two main listening tests—for audio quality and bass intensity—were carried out.

- Statistical analyses—ANOVA, nonlinear regression, and correlation analysis—were used to answer research questions; their derivative analyses showed a general trend to aid selection of both generic and specific NLDs for better VBS.

Chapter 2 introduced and reviewed the missing fundamental theory from pitch perception research in psychoacoustics. Starting from the Ohm-Seebeck controversy, subtopics in pitch perception research, including existence regions, dominance regions, and modeling for temporality, pattern recognition, and autocorrelation pitch were discussed. A review of psychoacoustic pitch perception research was linked to audio engineering research into the bass reproduction problem of small- or medium-sized loudspeakers and low-quality headphones. Blind and non-blind signal processing modules for VBS were introduced with special emphasis on NLD-based signal processing technique.

Chapter 3 focused on three areas: development of the triple duality theorem, generalized harmonic analysis, and synthesis of polynomials with harmonic shifting technique for VBS with directional loudspeaker application. The original mathematical proofs of the triple duality theorem were presented; a computational harmonic analysis and synthesis algorithm, HASP, was derived from the theorem. A generalized harmonic analysis was performed on ATSR and EXP NLDs to study the effect of nonlinearity's parameters. The THD metric was also tested and found to be unreliable in assessing nonlinearity. A real-time synthesizer and harmonic shifting technique was implemented, and a subjective listening test results showed the feasibility of the algorithm.

Chapter 4 deals with a computational algorithm development for multitone analysis of NLDs. YOGA was developed as a general tool for computing intermodulation distortion products and harmonics at the output of polynomial NLD. The two limitations of PIPCT were solved; MPIPCT was proposed, and a theoretical analysis showed that MPIPCT can reduce exponential component growth by 25–50%. Thus, better computational efficiency (in terms of memory usage and speed) was obtainable by using MPIPCT rather than PIPCT. The algorithm YOGA extended MPIPCT from monomial NLD to polynomial NLD and allowed exact and efficient computation of any order of nonlinearity with any number or combination of amplitude, frequency and phase of input multitone. In addition to the algorithmic development of YOGA, harmonic addition theorem (HAT) was proved with two original mathematical proofs. The theorem was again transformed into the HAT algorithm, which is a key component in YOGA. The pitfalls of FFT in nonlinear distortion analysis were discussed, and the advantages of using YOGA over FFT were highlighted. A generalized formula to generate logarithmically equally spaced multitone signals was also derived and presented in Chapter 4.

Chapters 3 and 4 developed mathematical analysis and computational techniques to assess nonlinear distortion using single-tone and multitone signals. To achieve the objectives, many new mathematical properties of signals and systems were discovered, proved, and presented.

Three research goals were addressed in Chapter 5—to (a) find out which types of NLDs are suitable for VBS; (b) create deeper bass while maintaining pleasant audio quality; and (c) ascertain whether *Rnonlin* nonlinear distortion model can be

used to assess nonlinear distortion caused by NLDs. Statistical tests confirmed that *Rnonlin* could be used for future nonlinear distortion predictions of NLDs in VBS. Moreover, the derivative analysis in terms of input-output characteristics of NLDs showed common features that correlate well with good bass enhancement.

6.2 Future Works

This section presents several investigations that are worthy to be pursued for further research in psychoacoustic bass enhancement.

All the NLDs in this dissertation are zero-memory (memory-less) NLDs, and due to the complexity of the analysis, non-zero memory NLDs were not considered to be used in VBS. However, previously proposed technique such as integrator [LA02] [LA04] and recursive harmonic generation technique with a feedback loop [SD99] can be considered non-zero memory NLDs. An examination of the theoretical link between harmonic/intermodulation generation patterns of such non-zero memory NLDs and perceptual nonlinear distortion scores prediction by *Rnonlin* can provide a useful research topic.

Findings in Chapter 5 confirm that *Rnonlin* model can predict the nonlinear distortion prediction scores for VBS. The goal of VBS is to enhance the bass while maintaining the lesser distortion artifacts. However, to predict the bass intensity, a model is not yet available. A good and reliable model that can predict bass intensity scores without having to run tedious subjective tests can be valuable for VBS research, if this model does exist or can be constructed.

Next, a computational aspect of YOGA can be further explored. As discussed in Chapter 4, YOGA is a generic algorithm that can take in any order of polynomial order and any number of multitone components. For the large input size and very high order polynomials, however, YOGA takes a long time to simulate due to the components growth in the memory even though 50% theoretical reduction of the components has been achieved. To speed up simulation time, the graphics processing unit (GPU) technology can be used to parallelize the execution of the algorithm on many processor cores [Mcc08] [OHL08] [SK11]. However, memory access time is the main bottleneck in parallelism [HW11], the tradeoff between the memory usage and computational speed gain for porting YOGA over GPU platform may be an interesting research area to pursue.

AUTHOR'S PUBLICATIONS

Journals

N. Oo, *et al.*, "Perceptually-Motivated Objective Grading of Nonlinear Processing in Virtual-Bass System," *Journal of Audio Engineering Society*, Vol. 59, No. 11, pp. 804-824, Nov. 2011.

N. Oo, "Exact Tonal Analysis on Polynomials (ETAP) for Computational Harmonic Analysis," *Bahria University Journal of Information & Communication Technology (BUJICT)*, Vol. 4, No. 1, pp. 22-23, Aug. 2011.

Conference Papers

N. Oo and W. S. Gan, "Harmonic and intermodulation analysis of nonlinear devices used in virtual bass systems," presented at the AES 124th Convention, Amsterdam, the Netherlands, 2008.

N. Oo and W. S. Gan, "Harmonic analysis of nonlinear devices for virtual bass system," presented at the IEEE International Conference on Audio, Language and Image Processing (ICALIP), Shanghai, China, 2008.

W. T. Lim, N. Oo and W. S. Gan, "Synthesis of polynomial-based nonlinear device and harmonic shifting technique for virtual bass system," presented at the IEEE International Symposium on Circuits and Systems (ISCAS), Taipei, Taiwan, 2009.

N. Oo and W. S. Gan, "Loudspeaker frequency response measurements and bass reproduction," presented at the First International Conference on Science and Engineering (ICSE), Yangon, Myanmar, 2009.

N. Oo and W. S. Gan, “Psychoacoustic Bass Enhancement System”, presented at the IEEE international conference on Acoustics, Speech and Signal Processing (ICASSP) Show & Tell Demonstration, Taipei, Taiwan, 2009. Online: http://www.icassp09.com/ShowAndTell_04.asp

N. Oo and W. S. Gan, “Analytical and perceptual evaluation of nonlinear devices for virtual bass system,” presented at the AES 128th Convention, London, UK, 2010.

N. Oo, W. T. Lim and W. S. Gan, “Generalized harmonic analysis of arc-tangent square root (ATSR) nonlinear device for virtual bass system,” presented at the IEEE International conference on Acoustics, Speech and Signal Processing (ICASSP), Texas, USA, 2010.

N. Oo and W. S. Gan, “Analysis on precise intermodulation product computing technique and reduction of computational load,” presented at the 5th International Symposium on Communications, Control, and Signal Processing (ISCCSP), Rome, Italy, 2012.

N. Oo and W. S. Gan, “An exact harmonic computing technique for polynomial nonlinearities,” presented at the AES 132nd Convention, Budapest, Hungary, 2012.

BIBLIOGRAPHY

[Ame07] American Technology Corporation, *Hyper Sonic Sound H460 Product Sheet*, September 2007.

[Aar04] R. M. Aarts, "Low-complexity tracking and estimation of frequency and amplitude of sinusoids," *Digital Signal Processing*, vol. 14, no. 4, pp. 372–378, July 2004.

[Abu10] M. T. Abuelma'atti, "Harmonic and intermodulation performance of hearing aids under large sound pressure levels," *Applied Acoustics*, vol. 71, pp. 269–275, 2010.

[AMJ06] M. Arora, H.-G. Moon, and S. C. Jang, "Low Complexity Virtual Bass Enhancement Algorithm for Portable Multimedia Device," presented at the 29th AES International Conference; Seoul, Korea, September 2006.

[Ana11] Analog Devices, *ADSP-BF533: High Performance General Purpose Blackfin Processor*, 2011. Available at: http://www.analog.com/en/processors-dsp/blackfin/BF533-HARDWARE/processors/product.html

[ANSI07] *American National Standard Procedure for the Computation of Steady Sound*, ANSI/ASA Standard S3.4-2007, Acoustical Society of America, May 2007.

[ANSI60] *American Standard Acoustical Terminology*, ANSI Standard S1.1-1960.

[ANSI94] *American Standard Acoustical Terminology,* ANSI Standard S1.1-1994 (R2004).

[Arf79] D. Arfib, "Digital Synthesis of Complex Spectra by Means of Multiplication of Nonlinear Distorted Sine Waves," *J. Audio Eng. Soc.*, vol. 27, pp. 757–768, October 1979.

[Bar98] E. Barbour, "The Cool Sound of Tubes [Vacuum Tube Musical Applications]," *IEEE Spectrum*, vol. 35, pp. 24–35, August 1998.

[BBH07] H. Behrends, W. Bradinal, and C. Heinsberger, “Loudspeaker Systems for Flat Television Sets,” presented at the 123rd AES Convention, New York, USA, 2007.

[BC99] D. Ben-Tzur and M. Colloms, “The effect of MaxxBass psychoacoustic bass enhancement on loudspeaker design,” presented at the 106th Convention of the Audio Engineering Society, Munich, Germany, preprint 4892, May 1999.

[BCB03] V. Bewick, L. Cheek and J. Ball, “Statistical Review 7: Correlation and Regression,” *Crit. Care*, vol. 7, pp. 451–459, December 2003.Online: http://ccforum. com/ content/7/6/451

[Bec02] S. Bech, “Requirement for low-frequency sound reproduction, Part I: The audibility of changes in passband amplitude ripple and lower system cutoff frequency and slope,” *J. Audio Eng. Soc.*, vol. 50, pp. 564–580, Jul.–August 2002.

[BL05] M. R. Bai and J. Liao, “Acoustic analysis and design of miniature loudspeakers for mobile phone,” *J. Audio Eng. Soc.*, vol. 53, pp. 1061–1076, November 2005.

[BL06] M. R. Bai and W. C. Lin, “Synthesis and implementation of Virtual Bass System with a phase-vocoder approach,” *J. Audio Eng. Soc.*, vol. 54, no. 11, pp. 1077–1091, November 2006.

[Bod84] H. Bode, “History of Electronic Sound Modification,” J. Audio Eng. Soc., vol. 32, pp. 730–739, October 1984.

[Boe56] E. de Boer, “On the ‘residue’ in hearing,” Ph.D. dissertation, Faculty of Mathematics and Physics, Univ. of Amsterdam, the Netherlands, 1956.

[Bre94] A. S. Bregman, “Role of harmonic relations and F0,” in *Auditory Scene Analysis: The Perceptual Organization of Sound*, MIT Press, 1994, ch. 6, pp. 559–571.

[Bur79] M. Le Burn, “Digital Waveshaping Synthesis,” *J. Audio Eng. Soc.*, vol. 27, pp. 250–266, April 1979.

[Cab99] R. C. Cabot, “Fundamentals of modern audio measurement,” *J. Audio Eng. Soc.*, vol. 47, pp. 738–762, September 1999.

[CAH11] L. K. Chiu, D. V. Anderson, and B. Hoomes, “Audio output enhancement algorithms for piezoelectric loudspeakers,” in *Proc. DSP workshop and IEEE Signal Process. Educ. Workshop (DSP/SPE)*, Arizona, 2011, pp. 317–320.

[CC04] O. Christensen, K. L. Christensen, *Approximation Theory from Taylor Polynomials to Wavelets*, Boston: Birkhäuser, (2004).

[CG96] J. H. Conway, and R. K. Guy, “Triangular Numbers” in *The Book of Numbers*, Spinger-Verlag: 1996, pp. 33–38.

[Che04] A. de Cheveigné, “Pitch perception models – a historical review,” *Proc. the 18th Int. Conf. on Acoust.*, Kyoto, 2004.

[Che10] A. de Cheveigné, “Pitch perception,” in *Oxford Handbook of Auditory Science–Auditory Perception*, C. J. Plack, Eds., Oxford: Oxford University Press, 2010, pp. 71–104.

[CKM05] J. H. Chang, N. S. Kim and S. K. Mitra, “Pitch estimation of speech signal based on adaptive lattice notch filter,” *Signal Processing*, vol. 85, no. 3, pp. 637–641, March 2005.

[Cor11] B. Cordell, *Designing Audio Power Amplifiers*, McGraw-Hill, 2011.

[CR07] C. D. Capua and E. Romeo, “A smart THD meter performing an original uncertainty evaluation procedure,” *IEEE Trans. Instrum. Meas.*, vol. 56, no. 4, pp. 1257–1264, August 2007.

[CRP11] S. Cecchi, L. Romoli, P. Peretti, and F. Piazza, “A combined psychoacoustic approach for stereo acoustic echo cancellation,” *IEEE Trans. on Audio, Speech, and Lang. Process.*, vol. 19, pp. 1530–1539, August 2011.

[CRR76] M. J. Cheng, L. R. Rabiner, A. E. Rosenberg, and C. A. McGonegal, “Some comparisons among several pitch detection algorithms,” in *Proc. IEEE Int. Conf. on Audio, Speech, and Signal Processing (ICASSP)*, Philadelphia, 1976, pp. 332–335.

[CVA01a] E. Czerwinski, A. Voishvillo, S. Alexandrov, and A. Terekhov, “Multitone Testing of Sound System Components—Some Results and Conclusions, Part 1: History and Theory,” *J. Audio Eng. Soc.*, vol. 49, pp. 1011–1048, November 2001.

[CVA01b] E. Czerwinski, A. Voishvillo, S. Alexandrov and A. Terekhov, “Multitone Testing of Sound System Components—Some Results and Conclusions, Part 2: Modeling and Application,” *J. Audio Eng. Soc.*, vol. 49, pp. 1181–1192, December 2001.

[Dai00] H. Dai, “On the relative influence of individual harmonics on pitch judgment,” *J. Acoust. Soc. Am.*, vol. 107, pp. 953–959, 2000.

[DJ97] C. Dodge, and T. A. Jerse, *Computer Music: Synthesis, Composition, and Performance*, 2nd. ed., Schirmer, 1997.

[DPK01] D. Pressnitzer, R. D. Patterson, and K. Krumbholz, “The lower limit of melodic pitch,” *J. Acoust. Soc. Am.*, vol. 109, pp. 2074–2084, 2001.

[EBU08] EBU-TECH3253, *Sound Quality Assessment Material Recordings for Subjective Tests*, EBU, Geneva, 2008.

[Eve69] H. W. Eves, *Carl Friedrich Gauss* in “In Mathematical Circles, A Selection of Mathematical Stories and Anecdotes: Quadrant III and IV,” Prindle Weber & Schmidt, 1969, pp. 112–117.

[Dun04] J. Dunn, *Measurement Techniques for Digital Audio, Audio Precision Application Note #5*, Oregon: Audio Precision, 2004.

[Far10] A. Farnell, *Designing Sound*, MIT Press, 2010.

[Fle24] H. Fletcher, “The physical criterion for determining the pitch of a musical tone,” *Phys. Rev.*, vol. 23, pp. 427–437, March 1924.

[Fou09a] J. B. J. Fourier, *Théorie Analytique de la Chaleur*, 2009, reprint of 1822 ed., Cambridge Uni. Press.

[Fou09b] J. B. J. Fourier, *The Analytical Theory of Heat*, 2009, reprint of 1878 ed., trans. A. Freeman, Cambridge Uni. Press.

[FS98] E. E. Feremans and F. D. Smet, "Method and Device for Processing Signals," US Patent 5,828,755, October 1998.

[FZ07] H. Fastl and E. Zwicker, *Psychoacoustics Facts and Models*, Springer, 2007.

[Gen10] Genesis Acoustics, *Loudness Online—Free Loudness Toolbox*, 2010. Available at: http://www.genesis-acoustics.com/loudness_online-32.html.

[GK04] W. S. Gan and S. M. Kuo, "Integration of virtual bass reproduction in active noise control headsets," in *Proc. 7th Int. Conf. on Signal Process. (ICSP)*, Beijing, 2004, pp. 368–371.

[GKT01] W. S. Gan, S. M. Kuo, and C. W. Toh, "Virtual bass for home entertainment, multimedia PC, game station and portable audio systems," *IEEE Trans. Consum. Electron.*, vol. 47, pp. 787–794, November 2001.

[GL03a] E. R. Geddes and L. W. Lee, "Auditory perception of nonlinear distortion—theory," presented at the 115th AES Convention, New York, USA, 2003.

[GL03b] L. W. Lee and E. R. Geddes, "Auditory perception of nonlinear distortion," presented at the 115th AES Convention, New York, USA, 2003.

[Gol73] J. L. Goldstein, "An optimum processor theory for the central information of the pitch of complex tones," *J. Acoust. Soc. Am.*, vol. 54, pp. 1496–1516, 1973.

[Got04] M. Goto, "A real-time music scene-description system: predominant-F0 estimation for detecting melody and bass lines in real-world audio signals," *Speech Communication*, vol. 43, no. 4, pp. 311–329, September 2004.

[Gus96] F. Gustafsson, "Determining the Initial State in Forward-Backward Filtering," *IEEE Trans. Signal Process.*, vol. 44, pp.

988–992, April 1996.

[HA09] D. V. M. Howard and J. A. S. Angus, *Acoustics and Psychoacoustics*, 4th ed., Burlington, MA: Focal Press, 2009.

[Hel54] H. L. F. Helmholtz, *On the sensations of tones as a physiological basis for the theory of music,* the Dover ed., an unaltered republication of the 2nd ed. (1885) of [Hel95], Trans. A. J. Ellis, New York: Dover Publications, 1954.

[Hel56] H. Helmholtz, "Ueber Combinationstöne," *Annalen der Physik*, vol. 175, pp. 497–540, 1856.

[Hel95] H. L. F. Helmholtz, *On the sensations of tones as a physiological basis for the theory of music*, 3rd ed., Trans. A. J. Ellis, London: Longmans, Green, and Co., 1895. Available at: http://www.archive.org/details/onsensationsofto00helmrich

[HG71] A. J. M. Houtsma and J. L. Goldstein, "The central origin of the pitch of complex tones: evidence from musical interval recognition," *J. Acoust. Soc. Am.*, vol. 51, pp. 520–529, November 1971.

[HH10a] A. J. Hill and M. O. J. Hawksford, "Wide-area psychoacoustic correction for problematic room modes using non-linear bass synthesis," presented at the 129th AES Convention, San Francisco, CA, USA, 2010.

[HH10b] A. J. Hill, and M. O. J. Hawksford, "A hybrid virtual bass system for optimized steady-state and transient performance," in *Proc. IEEE 2nd Comp. Sci. and Electron. Eng. Conf. (CEEC)*, Essex, 2010, pp. 1–6.

[HK06] J. Han, and M. Kamber, *Data Mining Concepts and Techniques*, 2nd ed., Morgan Kaufmann Publishers, 2006.

[HM07] S. Haykin and M. Moher, *Introduction to Analog and Digital Communications*, Wiley, 2007.

[HVK99] M. Hanmandlu, O. P. Verma, N. K. Kumar, and M. Kulkarni, "A novel optimal fuzzy system for color image enhancement using

bacterial foraging," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 8, pp. 2867–2878, August 1999.

[HW11] G. Hager and G. Wellein, *Introduction to High Performance Computing for Scientifists and Engineers*, CRC Press, 2011.

[HWJ03] D. E. Hinkle, W. Wiersma and S. G. Jurs, *Applied Statistics for the Behavioral Sciences*, MA: Houghton Mifflin, 2003.

[ITU01] ITU-R BS.1534-1, *Method for the Subjective Assessment of Intermediate Sound Quality (MUSHRA)*, International Telecommunications Union, Geneva, Switzerland, 2001.

[ITU90] ITU-R, *Recommendation BS.562-3, Subjective assessment of sound quality*, International Telecommunications Union Radiocommunication Assembly, 1990.

[ITU97] ITU-R, *Recommendation BS.1116-1, Methods for subjective assessment of small impairments in audio systems including multi-channel sound systems,* International Telecommunications Union Radiocommunication Assembly, 1997.

[Kar02] F. A. Karnapi, "Psychoacoustic signal processing techniques in bass perception enhancement," M.Eng. thesis, School of Electr. and Electron. Eng., Nanyang Tech. Univ., Singapore, 2002.

[KB01] C. Kaernbach and C. Bering, "Exploring the temporal mechanism involved in the pitch of resolved harmonics," *J. Acoust. Soc. Am.*, vol. 110, pp. 1039–1048, 2001.

[KGE02] K. A. Karnapi, W. S. Gan, and M. H. Er, "Method to enhance low frequency perception from a parametric array loudspeaker," presented at the AES 112th Convention, Munich, Germany, 2002.

[Kla01] A. I. Klayman, "Low-frequency audio enhancement," U. S. Patent 6285767, September 4, 2001.

[Kla03] A. Klapuri, "Multiple fundamental frequency estimation based on harmonicity and spectral smoothness," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 6, pp. 804–815, November 2003.

[Kla06] A. Klapuri, "Auditory model-based methods for multiple

fundamental frequency estimation," in *Signal Processing Methods for Music Transcription*, pp. 229–264; Springer, 2006.

[Kla08] A. Klapuri, "Multipitch analysis of polyphonic music and speech signals using an auditory model," *IEEE Trans. Audio, Speech and Lang. Process.*, vol. 16, no. 2, pp. 255–266, February 2008.

[Kli06] W. Klippel, "Tutorial: Louspeaker Nonlinearities—Causes, Parameters, Symptoms," *J. Audio Eng. Soc*., vol. 54, pp. 907–939, October 2006.

[Kuh90] W. B. Kuhn, "A real-time pitch recognition algorithm for music application," *Computer Music Journal*, vol. 14, no. 3, pp.60–71, Fall 1990.

[LA02] E. Larsen and R. M. Aarts, "Reproducing low-pitched signals through small loudspeakers," *J. Audio Eng. Soc.*, vol. 50, pp. 147–164, March 2002.

[LA04] E. Larsen and R. M. Aarts, *Audio Bandwidth Extension Application of Psychoacoustics, Signal Processing and Loudspeaker Design*, Chichester, West Sussex: John Wiley & Sons Ltd, 2004.

[Lan90] J. E. Lane, "Pitch detection using a tunable IIR filter," *Computer Music Journal*, vol. 14, no. 3, pp. 46–59, Autumn 1990.

[LD99] J. Laroche and M. Dolson, "New phase-vocoder techniques for real-time pitch shifting, chorusing, harmonizing, and other exotic audio modifications," *J. Audio Eng. Soc.*, vol. 47, no. 11, pp. 928–936, November 1999.

[Leb72] N. N. Lebedev, *Special Functions and Their Applications*, translated from the Russian and edited by R. A. Silverman, Dover, 1972.

[Lic51] J. C. R. Licklider, "A duplex theory of pitch perception," *Experientia*, vol. 7, pp. 128–134, 1951.

[Lic54] J. C. R. Licklider, "Periodicity pitch and place pitch," *J. Acoust. Soc. Am.*, vol. 26, p. 945, September 1954.

[Lic59] J. C. R. Licklider, "Three auditory theories," in *Psychology: A Study of a Science.* Vol 1, S. Koch, Ed., New York: McGraw-Hill, pp. 41–144, 1959.

[Lin66] R. B. Lindsay, "The story of acoustics," *J. Acoust. Soc. Am.*, vol. 39, pp. 629–644, April 1966.

[LOG09] W. T. Lim, N. Oo, and W. S. Gan, "Synthesis of polynomial-based nonlinear device and harmonic shifting technique for virtual bass system," in *Proc. IEEE Int. Symp. Circ. and Syst. (ISCAS)*, Taipei, 2009, pp. 1871–1874.

[LSP08] J. Lee, E. Song, Y. Park, and D. Youn, "Effective bass enhancement using second-order adaptive notch filter," *IEEE Trans. Multimedia*, vol. 54, pp. 663–668, May 2008.

[Mcc08] M. D. McCool, "Signal processing and general-purpose computing on GPUs," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 109–114, 2008.

[MGP85] B. C. J. Moore, B. R. Glasberg, and R. W. Peters, "Relative dominance of individual partials in determining the pitch of complex tones," *J. Acoust. Soc. Am.*, vol. 77, pp. 1853–1860, 1985.

[MGS84] B. C. J. Moore, B. R. Glasberg, and M. J. Shailer, "Frequency and intensity difference limens for harmonics within complex tones," *J. Acoust. Soc. Am.*, vol. 75, pp. 550–561, February 1984.

[MH91a] R. Meddis and M. J. Hewitt, "Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification," *J. Acoust. Soc. Am.*, vol. 89, pp. 2866–2882, June 1991.

[MH91b] R. Meddis and M. J. Hewitt, "Virtual pitch and phase sensitivity of a computer model of the auditory periphery. II: Phase sensitivity," *J. Acoust. Soc. Am.*, vol. 89, pp. 2883–2894, June 1991.

[MM89] S. Moritsugu and M. Matsumoto, "A Note on the Numerical Evaluation of Arctangent Function," *ACM SIGSAM Bulletin*, vol.

23, pp. 8–12, July 1989.

[MO97] R. Meddis and L. O'Mard, "A unitary model of pitch perception," *J. Acoust. Soc. Am.*, vol. 102, no. 3, pp. 1811–1820, September 1997.

[Mon01] D. C. Montgomery, *Design and Analysis of Experiments*, 5th ed. John Wiley & Sons, 2001.

[Moo03] B. C. J. Moore, *An Introduction to the Psychology of Hearing*, 5th ed., Academic Press, 2003.

[MR79] B. C. J. Moore and S. M. Rosen, "Tune recognition with reduced pitch and interval information," *Q. J. Exp. Psychol.*, 31, pp. 229–240, 1979.

[MSY03] J. H. McClellan, R. W. Schafer, and M. A. Yoder, *Signal Processing First*, Prentice Hall, 2003.

[NH07] P. Newell and K. Holland, *Loudspeakers for Music Recording and Reproduction*. Oxford, UK: Focal Press, 2007.

[Oda97] M. Oda, "Low Frequency Audio Conversion Circuit," US Patent 5,668,886, September 1997.

[OG08a] N. Oo and W. S. Gan, "Harmonic analysis of nonlinear devices for virtual bass system," in *Proc. IEEE Int. Conf. on Audio, Lang., and Image Process. (ICALIP)*, Shanghai, pp. 279–284, 2008.

[OG08b] N. Oo and W. S. Gan, "Harmonic and intermodulation analysis of nonlinear devices used in virtual bass systems," presented at the 124th AES Convention, Amsterdam, The Netherlands, 2008.

[OG09a] N. Oo and W. S. Gan, "Loudspeaker Frequency Response Measurements and Bass Reproduction," *Proc. First International Conference on Science and Engineering*, Yangon, 2009.

[OG09b] N. Oo and W. S. Gan, "Psychoacoustic bass enhancement system," Presented at IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP2009). Available at: http://www.icassp09.com/ShowAndTell_04.asp

[OG10] N. Oo and W. S. Gan, "Analytical and perceptual evaluation of

nonlinear devices for virtual bass system," presented at the 128th AES Convention, London, UK, 2010.

[OGC11] N. Oo, W. S. Gan, and Y. K. Chong, "Design equations for nonlinear sound analysis and synthesis," *J. Audio Eng. Soc.*, submitted for publication.

[OGH11] N. Oo, W. S. Gan, and M. O. Hawksford, "Perceptually-motivated objective grading of nonlinear processing in virtual bass systems," *J. Audio Eng. Soc.*, to be published.

[OGL10] N. Oo, W. S. Gan, and W. T. Lim, "Generalized harmonic analysis of arc-tangent square root (ATSR) nonlinear device for virtual bass system," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Process. (ICASSP)*, Dallas, pp. 301–304, 2010.

[OHL08] J. D. Owens, M. Houston, D. Luebke, S. Green, J. E. Stone, and J. C. Phillips, "GPU Computing: Graphics Processing Units—powerful, programmable, and highly parallel—are increasingly targeting general-purpose computing applications," *Pro. IEEE*, vol. 96, no. 5, pp. 879–899, May 2008.

[Ohm43] G. S. Ohm, "Ueber die Definition des Tones, nebst daran geknüpfter Theorie der Sirene und ähnlicher tonbildender Vorrichtungen," *Annalen der Physik*, vol. 135, pp. 513–565, 1843.

[Ohm44] G. S. Ohm, "Noch ein Paar Worte über die Definition des Tones," *Annalen der Physik*, vol. 138, pp. 1–18, 1844.

[Ohm73] G. S. Ohm, "On the definition of a tone with associated theory of the siren and similar sound producing device," (1843) trans. R. B. Lindsay in *Acoustics: Historical and Philosophical Development*, Stroudsburg, PA: Dowden, Hutchinson and Ross, pp. 243–247, 1973.

[OSB99] A. V. Oppenheim and R. W. Schafer with J. R. Buck, *Discrete-time Signal Processing*, 2nd ed., Prentice Hall, New Jersey, 1999.

[Ows99] B. Owsinski, "Element Three: Frequency Range—Equalizing," in *The Missing Engineer's Handbook*, Auburn Hills, MI: Artistpro,

pp. 26–27, 1999.

[Phi03] G. M. Phillips, *Interpolation and Approximation by Polynomials*, NY: Springer-Verlag, 2003.

[Plo64] R. Plomp, "The ear as a frequency analyzer," *J. Acoust. Soc. Am.*, vol. 36, pp. 1628–1636, 1964.

[Plo67] R. Plomp, "Pitch of complex tones," *J. Acoust. Soc. Am.*, vol. 41, pp. 1967–1533, 1967.

[Plo68] R. Plomp, "The ear as a frequency analyzer. II," *J. Acoust. Soc. Am.*, vol. 43, pp. 764–767, 1968.

[POF05] C. J. Plack, A. J. Oxenham, R. R. Fay, and A. N. Popper, Eds., *Pitch: Neural Coding and Perception*, Springer, 2005.

[PPA00] G. F. M. de Poortere, C. M. Polisset, and R. M. Aarts, "Ultra bass," U. S. Patent 6134330, October 17, 2000.

[PRL10] B. Pueo, G. Ramos, and J. J. Lopez, "Strategies for bass enhancement in multiactuator panels for wave field synthesis," *Applied Acoustics*, vol. 71, pp. 722–730, August 2010.

[PW76] R. S. Turner, "The Ohm–Seebeck dispute, Hermann von Helmholtz, and the origins of physiological acoustics," *The British J. for the Hist. of Sci*., vol. 10, pp. 1–24, March 1977.

[RCP10] L. Romoli, S. Cecchi, L. Palestini, P. Peretti, and F. Piazza, "A novel approach to channel decorrelation for stereo acoustic echo cancellation based on missing fundamental theory," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Process. (ICASSP)*, Dallas, pp. 329–332, 2010.

[RCR76] L. R. Rabiner, M. J. Cheng, A. E. Rosenberg, and C. A. McGonegal, "A comparative performance study of several pitch detection algorithms," *IEEE Trans. Audio, Speech, and Signal Process.*, vol. 5, pp. 399–418, October 1976.

[Rit62] R. J. Ritsma, "Existence region of the tonal residue. I," *J. Acoust. Soc. Am.*, vol. 34, pp. 1224–1229, 1962.

[Rit63] R. J. Ritsma, "Existence region of the tonal residue. II," *J. Acoust.*

Soc. Am., vol. 35, pp. 1241–1245, 1963.

[Rit67] R. J. Ritsma, "Frequencies dominant in the perception of the pitch of complex sounds," *J. Acoust. Soc. Am,* vol. 42, pp. 191–198, 1967.

[RP82] R. A. Rasch and R. Plomp, "The perception of musical tones," in *The Psychology of Music*, Florida: AP, 1982, ch. 1, pp. 1–21.

[RS04] P. Rao and S. Shandilya, "On the detection of melodic pitch in a percussive background," *J. Audio Eng. Soc.*, vol. 52, no. 4, pp. 378–391, April 2004.

[Sch38] J. F. Schouten, "The perception of subjective tones," *Proc. Kon. Ned. Akad. Wetenschap*., vol. 41, pp. 1086–1093, 1938.

[Sch39] J. F. Schouten, "Synthetic sound," *Philips Tech. Rev.*, vol. 4, pp. 167–173, June 1939.

[Sch40a] J. F. Schouten, "The perception of pitch," *Philips Tech. Rev.*, vol. 5, pp. 286–294, October 1940.

[Sch40b] J. F. Schouten, "The residue a new component in subjective sound analysis," *Proc. Kon. Ned. Akad. Wetenschap*., vol. 43, pp. 356–365, 1940.

[Sch40c] J. F. Schouten, "The residue and the mechanism of hearing," *Proc. Kon. Ned. Akad. Wetenschap*., vol. 43, pp. 990–999, 1940.

[Sch62] J. F. Schouten, R. J. Ritsma, and B. L. Cardozo, "Pitch of the residue," *J. Acoust. Soc. Am.*, vol. 34, pp. 1418–1424, 1962.

[Sch70] R. A. Schaefer, "Electronic musical tone production by nonlinear waveshaping," *J. Audio Eng. Soc.*, vol. 18, pp. 413–417, August 1970.

[Sch71] R. A. Schaefer, "Production of Harmonics and Distortion in P-N Junctions," *J. Audio Eng. Soc.*, vol. 19, pp. 759–768, October 1971.

[Sch80] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*, New York: Wiley, 1980.

[SD99] M. Shashoua and D. Glotter, "Method and system for enhancing

quality of sound signal," U. S. Patent 5930373, July 27, 1999.

[See41] A. Seebeck, "Beobachtungen über einige Bedingungen der Enstehung von Tönen," *Annalen der Physik*, vol. 129, pp. 417–436, 1841.

[See43] A. Seebeck, "Ueber die Sirene," *Annalen der Physik*, vol. 136, pp. 449–481, 1843.

[See44a] A. Seebeck, "Ueber die definition des tones," *Annalen der Physik*, vol. 139, pp. 353–368, 1844.

[See44b] A. Seebeck, "Ueber die Erzeugung von Tönen durch getrennte Eindrücke, mit Beziehung auf die Definition des Tones," *Annalen der Physik*, vol. 139, pp. 368–380, 1844.

[SH07] E. M. de Santis and S. Henin, "Perception & Thresholds of Nonlinear Distortion using Complex Signals," Masters dissertation, Department of Electronic Systems, Aalborg University, 2007.

[Shm05] D. Shmilovitz, "On the definition of total harmonic distortion and its effect on measurement interpretation," *IEEE Trans. Power Del.*, vol. 20, no. 1, pp. 526–528, January 2005.

[SK11] J. Sanders and E. Kandrot, *CUDA by Example: An introduction to general-purpose GPU programming*, Addison-Wesley, 2011.

[Sta98] M. Slaney, "Auditory Toolbox: Version 2," *Interval Research Corporation 1988-010*, 1998. Available at http://cobweb.ecn.purdue.edu/~malcolm/interval/1998-010/

[Sue70] C. Y. Suen, "Derivation of harmonic equations in nonlinear circuits," *J. Audio Eng. Soc.*, vol. 18, pp. 675–676, December 1970.

[Sun01] D. Sundararajan, "Aliasing and other effects," in *The Discrete Fourier Transform: Theory, Algorithms and Applications*, World Scientific, 2001, pp. 225–248

[SW03] G. A. F. Seber and C. J. Wild, "Nonlinear Regression," New Jersey: John Wiley & Sons, 2003.

[Ter00] E. Terhardt. (2000, February 12). *Definition of Pitch* [Online]. Available at: http://www.mmk.ei.tum.de/persons/ter/top/defpitch.html

[Ter74] E. Terhardt, "Pitch, consonance, and harmony," *J. Acoust. Soc. Am.*, vol. 55, pp. 1961–1969, May 1974.

[TG99] C. W. Toh and W. S. Gan, "A Real-time Virtual Surround Sound System with Bass Enhancement," presented at the 107th AES Convention, New York, USA, 1999.

[TGT00] S. E. Tan, W. S. Gan, C. W. Toh, and J. Yang, "Application of virtual bass in audio crosstalk cancellation," *Electronic Letters*, vol. 36, pp. 1500–1501, August 2000.

[THH08] R. Turnbull, P. Hughes, and S. Hoare, "Audio enhancement for portable device based speech applications," presented at the AES 124th Convention, Amsterdam, the Netherlands, 2008.

[TJG10] E. L. Tan, P. Ji, and W. S. Gan, "On preprocessing techniques for bandlimited parametric loudspeakers," *Applied Acoustics*, vol. 71, pp. 486–492, 2010.

[TK00] T. Tolonen and M. Karjalainen, "A computationally efficient multipitch analysis model," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 6, pp. 708–716, November 2000.

[TMZ03] C. T. Tan, B. C. J. Moore and N. Zacharov, "The Effect of Nonlinear Distortion on the Perceived Quality of Music and Speech Signals," *J. Audio Eng. Soc.*, vol. 51, pp. 1012–1031, November 2003.

[TMZ04] C. T. Tan, B. C. J. Moore, N. Zacharov and V. V. Mattila, "Predicting the Perceived Quality of Nonlinearly Distorted Music and Speech Signals," *J. Audio Eng. Soc.*, vol. 52, pp. 699–711, July/August 2004.

[TSY10] R. Tsutsui, I. Setiawan, and Y. Iwata, "Audio bandwidth expansion," U. S. Patent 7676043B1, March 9, 2010.

[Tur77] R. S. Turner, "The Ohm–Seebeck dispute, Hermann von

Helmholtz, and the origins of physiological acoustics," *The British J. for the Hist. of Sci.*, vol. 10, pp. 1–24, March 1977.

[Une98] T. Unemura, "Audio Circuit," US Patent 5771296, June 1998.

[VBR10] E. Vincent, N. Bertin, R. Badeau, "Adaptive harmonic spectral decomposition for multiple pitch estimation," *IEEE Trans. Audio, Speech, and Lang. Proess.*, vol. 18, pp. 528–537, 2010

[Vin05] E. Vincent, "A MATLAB interface for MUSHRA listening tests," 2005. Available at http://www.elec.qmul.ac. uk//people/emmanuelv/mushram/

[Voi06] A. Voishvillo, "Assessment of nonlinearity in transducers and sound systems—from THD to perceptual models," presented at the 121st AES Convention, San Francisco, CA, USA, 2006.

[VTC04] A. Voishvillo, A. Terekhov, E. Czerwinski and S. Alexandrov, "Graphing, Interpretation and Comparison of Results of Loudspeaker Nonlinear Distortion Measurements," *J. Audio Eng. Soc.*, vol. 52, pp. 332–357, April 2004.

[VZD00] S. Vassiliadis, M. Zhang, and J. G. Delgado-Frias, "Elementary function generators for neural-network emulators," *IEEE Trans. Neural Netw.*, vol. 11, no. 6, pp. 1438–1449, November 2000.

[Wei11] E. W. Weisstein (2011). *Harmonic Addition Theorem*, from: MathWorld—A Wolfram web resource. [Online]. Available at: http://mathworld.wolfram.com/HarmonicAdditionTheorem.html

[Wig73] F. L. Wightman, "The pattern-transformation model of pitch," *J. Acoust. Soc. Am.,* vol. 54, pp. 407–416, 1973.

[Wit93] R. A. Witte, Hewlett-Packard Company, *Spectrum & Network Measurements*, New Jersey: Prentice Hall, 1993.

[XP08] J. W. Xu and J. C. Principe, "A pitch detector based on generalized correlation function," *IEEE Trans. Audio, Speech and Lang. Process.*, vol. 16, no. 8, pp. 1420–1432, November 2008.

[Yea08] M. B. Yeary, "An efficient intermodulation product computing technique for broadband active transmit systems," *IEEE Trans.*

Instrum. Meas., vol. 57, pp. 438–443, February 2008.

[Zol02] U. Zölzer (Ed.), *DAFX – Digital Audio Effects*, West Sussex, UK: John Wiley & Sons, 2002.

[ZRB08] S. Zieliński, F. Rumsey and S. Bech, "On Some Biases Encountered in Modern Audio Quality Listening Tests—A Review," *J. Audio Eng. Soc.*, vol. 56, pp. 427–451, June 2008.

[Zwi61] E. Zwicker, "Subdivision of the Audible Frequency Range into Critical Bands (Frequenzgruppen)," *J. Acoust. Soc. Am.*, vol. 33, no. 2, p. 248, 1961.