Novel Digital Alias-Free Signal Processing Approaches to FIR Filtering Estimation

Darawsheh, H.

UNIVERSITY OF WESTMINSTER

Novel Digital Alias-Free Signal Processing Approaches to FIR Filtering Estimation

Hikmat Y. Darawsheh

School of Computer Science and Engineering

University of Westminster

A thesis submitted in partial fulfilment of the requirements

of the University of Westminster for the degree of

Doctor of Philosophy

London, June 2021.

Declaration

I declare that all the material contained in this thesis is my own work and has not been submitted for any other award. To the best of my knowledge, none of the material of this thesis has been previously published or created by another person except where due reference is made in the text.

Hikmat Y. Darawsheh

Abstract

This thesis aims at developing a new methodology of filtering continuous-time bandlimited signals and piecewise-continuous signals from their discrete-time samples. Unlike the existing state-of-the-art filters, my filters are not adversely affected by aliasing, allowing the designers to flexibly select the sampling rates of the processed signal to reach the required accuracy of signal filtering rather than meeting stiff and often demanding constraints imposed by the classical theory of digital signal processing (DSP). The impact of this thesis is cost reduction of alias-free sampling, filtering and other digital processing blocks, particularly when the processed signals have sparse and unknown spectral support.

Novel approaches are proposed which can mitigate the negative effects of aliasing, thanks to the use of nonuniform random/pseudorandom sampling and processing algorithms. As such, the proposed approaches belong to the family of digital alias-free signal processing (DASP). Namely, three main approaches are considered: total random (ToRa), stratified (StSa) and antithetical stratified (AnSt) random sampling techniques.

First, I introduce a finite impulse response (FIR) filter estimator for each of the three considered techniques. In addition, a generalised estimator that encompasses the three filter estimators is also proposed. Then, statistical properties of all estimators are investigated to assess their quality. Properties such as expected value, bias, variance, convergence rate, and consistency are all inspected and unveiled. Moreover, closed-form mathematical expression is devised for the variance of each single estimator.

Furthermore, quality assessment of the proposed estimators is examined in two main cases related to the smoothness status of the filter convolution's integrand function, $g(t,\tau) := x(\tau)h(t-\tau)$, and its first two derivatives. The first main case is continuous and differentiable functions $g(t,\tau)$, $g'(t,\tau)$, and $g''(t,\tau)$. Whereas in the second main case, I cover all possible instances where some/all of such functions are piecewise-continuous and involving a finite number of bounded discontinuities.

Primarily obtained results prove that all considered filter estimators are unbiassed and consistent. Hence, variances of the estimators converge to zero after certain number

of sample points. However, the convergence rate depends on the selected estimator and which case of smoothness is being considered.

In the first case (i.e. continuous $g(t,\tau)$ and its derivatives), ToRa, StSa and AnSt filter estimators converge uniformly at rates of N^{-1}, N^{-3}, and N^{-5} respectively, where $2N$ is the total number of sample points. More interestingly, in the second main case, the convergence rates of StSa and AnSt estimators are maintained even if there are some discontinuities in the first-order derivative (FOD) with respect to τ of $g(t,\tau)$ (for StSa estimator) or in the second-order derivative (SOD) with respect to τ of $g(t,\tau)$ (for AnSt). Whereas these rates drop to N^{-2} and N^{-4} (for StSa and AnSt, respectively) if the zero-order derivative (ZOD) (for StSa) and FOD (for AnSt) are piecewise-continuous. Finally, if the ZOD of $g(t,\tau)$ is piecewise-continuous, then the uniform convergence rate of the AnSt estimator further drops to N^{-2}.

For practical reasons, I also introduce the utilisation of the three estimators in a special situation where the input signal is pseudorandomly sampled from otherwise uniform and dense grid. An FIR filter model with an oversampled finite-duration impulse response, timely aligned with the grid, is proposed and meant to be stored in a lookup table of the implemented filter's memory to save processing time. Then, a synchronised convolution sum operation is conducted to estimate the filter output.

Finally, a new unequally spaced Lagrange interpolation-based rule is proposed. The so-called composite 3-nonuniform-sample (C3NS) rule is employed to estimate area under the curve (AUC) of an integrand function rather than the simple Rectangular rule. I then carry out comparisons for the convergence rates of different estimators based on the two interpolation rules. The proposed C3NS estimator outperforms other Rectangular rule estimators on the expense of higher computational complexity. Of course, this extra cost could only be justifiable for some specific applications where more accurate estimation is required.

Acknowledgements

All praise be to Allah, Lord of the worlds, Who taught man what he did not know. Indeed, without the will of Allah, I would not be where I am now. Peace and blessing of Allah be upon His Messenger, Prophet Muhammad, who had always been urging us to research for, learn, and spread all useful sorts of knowledge and sciences.

I would like to express deepest gratitude to my supervisor Dr Andrzej Tarczynski, who offered me the opportunity to conduct this interesting research. His constant mentoring, support and collaboration have guided me to gain grateful knowledge and develop my research skills. Despite the tough times of the Coronavirus pandemic, he devoted a lot of his time to provide me with friendship, encouragement and advice. Thank you, Dr Tarczynski, for everything, I really appreciate your help.

I am very grateful to the examination panel, Prof Miroslaw Pawlak and Dr Adem Coskun, for their careful revisions, constructive comments, and valuable suggestions which helped improve the final presentation of this thesis. Moreover, I would also like to thank Prof. Izzet Kale for the fruitful discussions and encouragement.

This work could not have been achieved without the generous support from University of Westminster, the BC's HESPAL Scholarship and Al-Quds Open University (Palestine). This cooperation has made my research journey stable and less worrying. Many thanks to all of you for making this research possible.

Extreme appreciation goes to my family, relatives, and friends for their unconditional love. My dear Dad, I will never forget the care and support you immersed me with in whole of my life. My beloved wife, Suheir, I will never forget your patience, endurance, and looking after the boys during my study. Your love is a treasure that I will cherish forever. Mohammad, Ameer, Abdullah, Waleed and Hamza, you have always relieved me and topped up my heart with joy and happiness. A'alia, Ziyyad, Moayyad, Hussam, Mohannad, and all other brothers and sisters, your kindness, encouragement and cooperation were incredible. You are really wonderful.

Finally, my late Mum, your genuine prayers always accompanied me day and night and never abandoned my heart and mind, may Allah bestow his mercy and forgiveness upon your soul. I ask Allah to grant you the highest rank of the Paradise, Jannatul Firdaws. Aameen. This thesis is dedicated to you and my Dad.

In memory of my affectionate Mum, Sadiqah Awwad,

and

In loyalty to my dear Dad, Yousef Darawsheh.

Contents

List of Acronyms

ADC	Analog to Digital Converter.
AnSt	Antithetical Stratified (random sampling technique).
BPF	Band Pass Filter.
CT	Continuous-Time/Computed Tomography (depending on the context).
DASP	Digital Alias-free Signal Processing.
DFT	Discrete Fourier Transform.
DT	Discrete-Time
DSP	Digital Signal Processing.
FIR	Finite Impulse Response.
FOD	First-Order Derivative.
FT	Fourier Transform.
HPF	High Pass Filter.
HySt	Hybrid-Stratified (random sampling technique).
IIR	Infinite Impulse Response.
LPF	Low Pass Filter.
LTI	Linear Time-Invariant.
MRI	Magnetic Resonance Imaging.
NUS	Nonuniform Sampling.
PNS	Periodic Nonuniform Sampling.
SDR	Software-Defined Radio
SECOEX	Sequential Component Extraction.
SNR	Signal-to-Noise Ratio.

SOD	Second-Order Derivative.
SSF	Spectral Support Function
StSa	Stratified Sampling (random sampling technique).
ToRa	Total Random (random sampling technique).
TE	Total Error
ZOD	Zero-Order Derivative.

Notation

$[t-T, t)$	The time interval of the observation window
T	The length of the observation window
N	Half the number of random sample points = number of strata in AnSt
$x(t)$	The analog input signal
$h(t)$	The analog impulse response of a CT LTI filter
$f(t,\tau)$	The integrand function of the filtering convolution $= x(\tau)h(t-\tau)$
Δ_j	The length of the j-th stratum ($= \Delta$ for equidistant strata)
S_{j-1}	The time-instant of the beginning of the j-th stratum, with $S_0 = t - T$
S_j	The time-instant of the end of the j-th stratum, with $S_{2N} = t$ (for StSa) or $S_N = t$ (for AnSt)
C_j	The time-instant of the centre of the j-th stratum $= (S_{j-1} + S_j)/2$
τ_{D_j}	The time-instant in the j-th stratum at which there is a discontinuity in zero-, first-, and/or second-order derivative(s) of $f(t,\tau)$
A_j	The interval of the whole j-th stratum, $[S_{j-1}, S_j)$
$A_{j,L}$	The left subinterval of the j-th stratum in which there is a jump discontinuity, $[S_{j-1}, \tau_{D_j})$
$A_{j,R}$	The right subinterval of the j-th stratum in which there is a jump discontinuity, $[\tau_{D_j}, S_j)$
D_j	The time length of $A_{j,L}$, i.e. $D_j = \tau_{D_j} - S_{j-1}$
$\Delta_j - D_j$	The time length of $A_{j,\mathrm{R}}$, i.e. $\Delta_j - D_j = S_j - \tau_{D_j}$
τ_j	The time-instant of the random sample point in the j-th stratum
τ_j^a	The antithetical counterpart of τ_j, i.e. $\tau_j^a = 2C_j - \tau_j$, (for AnSt only)
K_j	The ratio D_j/Δ_j, i.e. $0 \leq K_j \leq 1$.

List of Figures

List of Tables

CHAPTER 1

Introduction

1.1. Overview

An inherent feature of all digital signal processing (DSP) algorithms is that the processed signals are sampled. This means that the values of these signals are only known at discrete time instants, whereas in-between there is no information collected. Many signal processing applications rely on full knowledge of the processed signals. Therefore, a common requirement is that the collected samples of some processed signal should somehow provide the complete required information about it. This is a challenging requirement that in practice can rarely be fulfilled.

A classical approach defined by Shannon is to use uniform sampling of the processed signal not less than a specific rate known as the *Nyquist rate* $= F_{s\_Nyq}$ [1], [2]. For a bandlimited signal, the Nyquist rate is two times the highest frequency present in it. If this is satisfied, then it can be uniquely represented by interpolating an *infinite* number of equally spaced signal samples and the *sinc* kernel (basis).

However, even in this case there are practical limitations. It is impossible to collect an infinite number of samples over a limited period of time. So, the infinite summation of the interpolation has to be truncated to a finite summation, meaning that we indirectly assume the signal is zero outside the time window of observation. But it is well-known that signals with finite support in the time domain have infinite bandwidth. Consequently, the assumption that they could be sampled at a finite rate is not entirely true.

In practice, signals are oversampled, then after truncating them to a finite window, reconstruction errors can be upper-bounded and dealt with. While Shannon sampling theorem states a sufficient condition for signal reconstruction; it is by no means necessary.

An interesting observation was made by Landau who formulated a necessary condition [3]. He proved that if the processed signal is a multiband signal and has single-sided spectrum sub-bandwidths $B_1, B_2, B_3, \dots, B_M$, then the minimum sampling rate that allows the perfect reconstruction of the signal is equivalent to the sum of all positive and negative sub-bandwidths, i.e. $F_{s\_min\_L} = \sum_{k=1}^{M} 2B_k$, provided that the number and locations of such sub-bandwidths are a priori known.

Unfortunately, this theorem does not state what sampling scheme should be used, or how to reconstruct the signal from the samples if they are correctly collected. Uniform sampling often fails in this case, because of the wide spectrum support of the signal and the distortion effects of aliasing phenomenon. Hence, suitable nonuniform random sampling schemes could be the best sampling approach to consider in this case.

Some special cases that use Landau rate, rather than Nyquist-Shannon's, are solved, and some are widely used in practice. So-called bandpass sampling allows users to sample high-frequency signals, e.g. multiband and communication signals, at the rates being many times slower than those recommended by Nyquist [4]. Multiband signals with known spectrum support can be tackled with by the use of periodic nonuniform sampling (PNS) (or multicoset sampling) patterns [5]–[7], where the minimum Landau sampling rate is approached. Additional challenges come when the spectrum support function of the processed signal is not fully known. In this case, blind spectrum multiband sub-Nyquist sampling and reconstruction techniques, such as the emerging compressed sensing (CS) framework, are exploited [8]–[10]. Though, such techniques are not completely "blind", as they assume that the number of bands and their widths are known in advance. It has been shown that the minimum achievable rate for blind spectrum-based sampling and perfect signal reconstruction is *twice* the minimum Landau sampling rate [9], i.e. $F_{s\_min\_BSS} = 2F_{s\_min\_L}$.

However, all the methods mentioned above, except for the last one in the previous paragraph (i.e. CS and blind spectrum techniques), use regular sampling schemes, which means that the classes of processed signals have to be heavily constrained in order to avoid aliasing. In this thesis, I understand *aliasing as a scenario when two different continuous-time signals, belong to the class of the processed signal, have identical sampled (discrete-time) versions*. Of course, this understanding takes into

account those cases where aliasing may partially occur in specific spectral component(s) of, for example, a composite signal. In this case, the discrete-time sequence obtained using insufficient uniform sampling rate (i.e. slightly less than the required Nyquist rate) is interpolated back to a new continuous-time signal which is really different from the original processed one, albeit this difference is sometimes insignificant, especially if the aliased component is a single sinusoid with negligible amplitude compared to the rest of the signal.

Digital alias-free signal processing (DASP) is an innovative technology that utilises random/pseudorandom sampling and/or quantising techniques and subsequent suitable algorithms to digitise and process analog signals without the bounds and restrictions imposed by the classical DSP theory and practice. Definitely, this novel technology, if properly understood and well-implemented, exhibits a considerable gain over conventional techniques, especially in mitigating the destructive effects of aliasing phenomenon and expanding the bandwidth of processed analog signals [11, p. 54].

This significant advantage of DASP in suppressing aliasing while tackling with signals beyond the Nyquist limit, despite being considered unachievable from DSP perspective, is a direct result of correctly employing randomisation in signal digitising and processing. In effect, proper randomised sampling and/or quantising leads, under certain conditions, to the fact that different analog signals are converted to different sequences of digital data. Meaning, the mapping between continuous-time domain and random discrete-time domain is unique [8], [12, pp. 101–103].

However, a new problem emerges resulting from the fact that the random discrete-time signal is not known, and it can be observed usually through a single realization. Therefore, the question arises about the quality of signal processing obtained from such single realization. The problem is, then, shifted from avoiding aliasing to statistical estimation and accuracy of the estimation.

DASP is heavily used to solving problems related to spectrum estimation, spatial signal processing, instrumentation, and some others [11, pp. 55–60]. But there is no much research done so far in DASP environment about signal filtering, or in general,

about the response of continuous-time linear time-invariant (CT LTI) systems. This thesis addresses such scarcity and explores different aspects that exist in this area.

1.2. Objectives and Scope

The thesis aims at developing a new methodology of filtering continuous-time bandlimited signals and piecewise-continuous signals (= having a finite number of bounded discontinuities) from their discrete-time samples. Such filters are used in many areas including communication systems, dynamic spectrum access, data collections, audio processing and satellite navigation. Unlike the existing, state-of-the-art filters, my filters are not adversely affected by aliasing, letting the designers flexibly select the sampling rates of the processed signal to reach the required accuracy of signal filtering rather than meeting stiff and often demanding constraints imposed by the classical theory of digital signal processing. The impact of this research is cost reduction of digital alias-free filtering and introducing applicable nonuniform random/pseudorandom approaches to filtering estimation where the traditional DSP solutions fall short for economical or technical reasons, particularly when the processed signals have sparse and unknown presence in the frequency domain. Said this, DASP should not be deemed as a replacement for the classical DSP, rather, DASP complements it and provide viable and promising alternatives.

The theory and practice of DASP technology is evolving rapidly, with new random sampling and quantisation techniques, algorithms, designs and implementations continuously emerge in research and application environments. However, the scope of this thesis does not cover all areas and applications under the nonuniform sampling and quantisation. For example, reconstruction of randomly sampled signals, compressed sensing, periodic nonuniform sampling and random quantisation techniques are not deeply discussed. Though, brief demonstrations of such applications and schemes are included in the literature review chapter (Chapter 2) for comparison purposes and comprehensiveness.

Now, regarding the thesis's title, although it includes the words "FIR Filtering", the scope of the thesis is neither about the various design characteristics of analog and digital filters, nor about the practical implementation (i.e. building hardware testbeds

on FPGA kits or other electronic devices). Numerous publications exist in the research literature, as well as many design software, about this well-established topic despite the fact that the vast majority of it is related to uniform sampling. Real-life applications involving filtering (analog or digital) are uncountable. Therefore, I am not going to discuss the detailed features, types, pros and cons, and other aspects related to filter design as a device.

The main goal associated with the term "FIR Filtering Estimation" in the title is to indicate that I am going to study the FIR filtering process, i.e. the convolution operation, as a system response to an input signal with specific features (randomly sampled, continuous or piecewise continuous, bandlimited or timelimited, etc). Consequently, it focuses on how to carry out the randomised filtering to estimate the output signal and investigate its various statistical properties.

Of course, I will design and use several FIR filters to validate my analytical findings through numerical and simulation examples. However, the design of a specific filter for a given numerical example is not exclusive or of highly significant importance. Meaning that several alternatives for the same filter could be selected, as long as they have equivalent bandwidth, centre frequency, passband and stopband(s) attenuations, and transition width. Other features could be relaxed with some acceptable tolerance depending on the application.

As is well-known, FIR filters are stable, causal and LTI systems with linear phase response, which makes them a good choice for my investigation. Actually, I need to explore the effects of employing random sampling techniques on the filter output, and how this would mitigate aliasing errors, albeit there will be statistical errors as a result of sampling irregularities. Therefore, the filter should have linear phase response, be stable and LTI so that it does not introduce extra errors that may affect my investigations and results.

Furthermore, the scope of the thesis includes the statistical impact of the presence of a limited number of bounded discontinuities in the argument product function of the filter convolution operation and/or its derivatives. Fundamentally, I introduce the use of stratification-based filter estimators of randomly sampled signals and analyse the variance and the convergence rate of the estimators. This does not involve how such

discontinuities could be detected in real-time applications or what suggested techniques are there to eliminate them. I abandoned this as a future research work to expand upon this thesis, although there is already some research addressing functions discontinuities, especially in audio signal processing [13], but from different perspective.

1.3. Potential Application Scenarios

Many wideband applications could make use of pseudorandomised filtering based on DASP alias-free sampling and processing techniques. While conventional DSP filters are usually utilised in such applications; they often entail expensive specialised hardware/software implementation devices that require high computational complexity. Moreover, there are some specific application areas or situations where classical digital filtering based on uniform sampling is not feasible. Hence, using nonuniform filtering alternatives bridge this gap, and present lower complexity solutions.

To show the strengths of DASP based filtering, and why I carry out this research in the first place, I provide below a non-exhaustive list of possible applications that can tolerate the statistical errors of randomised filtering in return for reducing the cost of their implementation.

• **Spectrum Management.** Regulatory bodies in each country, who are responsible for governing the telecommunications systems and licensing frequency bands to the end users, monitor the electromagnetic spectrum (or specific bands of it) for unauthorised access or harmful interference caused to the licensed users. When it is required to scan an especially empty or sparse wideband of the spectrum to detect potential unlicensed users, or even unknown signals, then this would be expensive using classical DSP uniform sampling and filtering techniques. Indeed, the *completely blind* sampling of a sparse and wide frequency range of the spectrum, in the conventional DSP way, requires high-bandwidth analog-to-digital converters (ADCs) or a complete system of lower bandwidth interleaved ADCs. The sampling rate of the applied high-bandwidth ADC or the total compound sampling rate of the interleaved

ADCs system should match the Nyquist rate in this case, which makes this approach economically unaffordable.

By using suitable randomised sampling and filtering DASP approaches, it is possible to monitor broad clusters of empty or sparse spectrum without abiding by the Nyquist limit on the sampling rate.

• **Software Defined Radios (SDRs).** These multi-standard transceiver devices feature special importance for DASP sampling and filtering approaches, because of the general nature of such devices regarding the wide frequency bands, waveforms and techniques they deal with. The embedded ADC can limit the variety of applications of the SDR if uniform sampling is the only considered technique. Many SDR designs often rely on reconfigurable ADCs and reconfigurable analog filters through switched capacitor technology or other techniques [14]–[18]. Despite these sophisticated hardware-based reconfigurability properties, these devices are still confined to the traditional DSP constraints for each operational wireless standard currently selected by the SDR. Coupling such devices with alias-free sampling and filtering capabilities would emphasise their strength in terms of adding more communications standards and widening the operational bandwidth.

• **Instrumentation.** Spectrum analysers', particularly, and many other digital instrumentation devices' performances can be significantly enhanced with randomised signal processing. Mitigating spectral aliases makes such devices operate beyond the limits imposed by classical theory of DSP. Some implementations of voltmeters, oscilloscopes and spectrum analysers have already been manufactured by commercial companies, like Hewlett Packard (*hp*), many decades ago [11, p. 52]. Further improvements and faster performance can always be achieved with statistical randomisation-based techniques.

• **Cognitive Radio.** In this emerging area of communications, licensed (*primary*) users are those who have been assigned specific frequency bands by the regulatory bodies. Temporally or spatially based spectrum *white holes* normally appear in the licensed bands when there is inactivity by the allocated primary user. Such holes could be exploited by potential unlicensed (*secondary*) users under certain conditions that guarantee no harmful interference will be caused to the primary users. This criterion

allows better utilisation of the naturally limited, yet crowded, resource; the electromagnetic spectrum.

In order for a cognitive radio principle to work efficiently, the secondary users need to continuously search a relatively wide spectrum band for possible white spaces. Similar to the first item above, Spectrum Management, uniform sampling-based approaches that may have to deal with such circumstances, i.e. sampling, digitising and processing high-frequency multiband signals, are costly and sometimes technically unviable. However, this kind of applications could be conveniently tackled with by applying nonuniform pseudorandom sampling schemes. Not only because of the beyond-Nyquist-limit capabilities of such randomised schemes, but also there is no need to actually detect the very detailed characteristics of the primary signals that might exist in the frequency band of interest, as the only required condition for the secondary users is to conduct a hypotheses test identifying whether or not there is a hole in this specific spectrum band. Traditional dynamic spectrum access techniques, e.g. an energy detector [19], [20], are usually used to perform this sort of hypotheses testing, which limits the width of the band to be searched at a given time and space to half the uniform sampling rate. Nevertheless, it is more efficient to use randomised/pseudorandomised techniques in this case, as broader bands could be considered while using average random sampling rate approaching the minimum Landau rate, or even less for specific types of estimations (e.g. first and second moments) and random processes (ergodic, for instance).

- **Wireless sensor networks (WSNs).** Imagine a large collection of wireless sensors spread across a geographical region to monitor specific natural (physical, chemical, optical, …) properties or conditions [21]. Each sensor collects and transmits data to either a locally centralised sink station or a globally centralised centre using a preassigned narrow bandwidth and centre frequency. All sensors are assumed to be event-triggered transmitters. Therefore, it is unknown which sensor would send data and when. This application area is also best utilised using nonuniform sampling schemes, like the three considered random approaches; ToRa, StSa and AnSt.

Other applications would include ad hoc networks, radar, and applications intended to detect signals with sparse and unknown spectral occupancy.

All mentioned scenarios above use deliberate randomised/pseudorandomised sampling techniques to extend the bandwidth of detectable signals beyond the traditional DSP limits using sub-Nyquist sampling rates. Nonetheless, there are many other application areas where unintentional nonuniform sampling is the only affordable digitising scheme. This includes, but not limited to, nuclear magnetic resonance (NMR) which is used to extract some features of proteins and nucleic acids [22], magnetic resonance imaging (MRI), astronomy, missing data from otherwise uniformly sampled signals (random skip sampling), sampling with jitter, hardware imperfections [11, p. 163].

1.4. Original Contribution to Knowledge

In Chapter 7, I present a detailed list of the novel contributions of this Thesis. Though, to get a brief insight, I list below the main contributions.

- Randomly/Pseudorandomly sampled continuous-time wideband signals and piecewise-smooth signals can, under certain conditions, be filtered using DASP-based filtering approaches without the negative effects of aliasing even if the average sampling frequency is less than the Nyquist rate. Thanks to the sparse spectrum occupancy features of the signal of interest, estimating the filter output is practically feasible using cost-effective random sampling techniques.

- Introducing ToRa-, StSa- and AnSt-based consistent filter estimators and identifying their statistical advantages and limitations.

- Devising mathematical expressions for the variances of the estimators and revealing their uniform convergence rates in different cases regarding the smoothness of the integrand/summand function, and its derivatives, of the filter convolution operation. My results prove that filter estimators' uniform convergence rates are adversely affected by the presence of discontinuities in the convolution integrand/summand function.

- Proposing new nonuniform interpolation technique, C3NS, based on Lagrange second-degree polynomial, to numerically integrate the area under the curve of the

filtering convolution summand function. Compared with the conventional Rectangular rule, the proposed method shows performance improvement.

1.5. List of Related Publications

Here are the journal and conference proceedings publications associated with this research.

- **H. Y. Darawsheh** and A. Tarczynski, "*Antithetical stratified sampling estimator for filtering signals with discontinuities*," Signal Processing, vol. 181, p. 107910, Apr. 2021.

- A. Tarczynski and **H. Y. Darawsheh**, "*DASP Implementation of Continuous-Time, Finite-Impulse-Response Systems*," in '2020 28th European Signal Processing Conference (EUSIPCO), Amsterdam, the Netherlands, Jan. 2021, pp. 2244–2248.

- **H. Y. Darawsheh** and A. Tarczynski, "*FIR Filtering of Discontinuous Signals: A Random-Stratified Sampling Approach*," in ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 2020, pp. 5800–5804.

- **H. Y. Darawsheh** and A. Tarczynski, "*Comparison Between Uniform and Nonuniform Interpolation Techniques for Digital Alias-free FIR Filtering*," in International Conference on Digital Image & Signal Processing (DISP'19), Oxford, United Kingdom, May 2019, pp. 1–5.

- **H. Y. Darawsheh** and A. Tarczynski, "*Filtering Nonuniformly Sampled Grid-Based Signals*," in 2018 4th International Conference on Frontiers of Signal Processing (ICFSP), Sep. 2018, pp. 56–60.

1.6. Thesis Layout

The rest of this thesis comprises six chapters. Chapter 2 reviews the research literature in the areas of traditional and randomised signal processing to formulate a background for this thesis. The core of this thesis is divided into two main parts, Part I: Continuous-Time Integrand/Summand Functions of the Filter Convolution Operation; and Part II:

Non-Smooth Integrand Functions. In the first part, which comprises two chapters, I introduce filter estimators, in the case of continuous-time input signal/filter impulse response, for the three considered DASP random sampling and processing approaches and assess their statistical qualities. Namely, in Chapter 3, I present ToRa-, StSa- and AnSt-based filter estimators. Then, I verify their unbiasedness and consistency, and devise mathematical expressions for the variance and uniform convergence rate for each estimator. While in Chapter 4, I consider a dense and uniform grid on which pseudorandomisation of the sampling process is taking place. The three estimators, mentioned above, are considered again, this time on practical backgrounds. Moreover, I propose a new nonuniform on-grid Lagrange-based interpolation technique to calculate the convolution sum besides the well-known Rectangular rule. A comparison between the different approaches is also provided.

In the second part, where discontinuities in the integrand/summand functions of the convolution operation and/or their derivatives are considered, there are two chapters, too. Chapter 5 demonstrates the impact of such discontinuities on the statistical properties of StSa-based estimator, while Chapter 6 considers the AnSt-based estimator and its new and different statistical features in the same underlying conditions regarding the non-smoothness of the integrand function and its derivatives.

Finally, conclusion remarks and future research opportunities are listed in Chapter 7. Appendices are presented at the end of the thesis. They include proofs of some of the proposed theorems throughout the text. References are also provided at the end.

CHAPTER 2

Processing Uniformly and Randomly Sampled Signals

2.1. Overview

Many physical properties and phenomena we observe everyday are analog in nature. Weather temperature, airplane location, speech volume, light intensity, magnetic field strength, car speed, and battery charge are all examples of continuous-time properties. If we would like to *measure* their absolute or relative values in order to build an electrical/electronic monitoring and/or controlling system, we somehow need to *sense* such properties in the first place. Fortunately, many types of sensors and transducers can easily be manufactured to do this very job. They can map/transform those physical properties to measurable quantities, such as electrical voltage or current, which are basically considered analog signals. Therefore, we have to design and implement either a completely analog system or a mixed (i.e. analog-and-digital) system to achieve the requested task.

In completely analog systems, signals are generated and processed in continuous-time domain. Whereas in digital (mixed) systems, we need to convert the analog signals to sequences of digital data before further processing, since modern computers, processors, microcontrollers, memory storage and other digital encoding/decoding blocks cannot deal with analog signals directly.

The conversion procedure (or digitization), which is normally carried out by an analog-to-digital converter (ADC), comprises three main stages: sampling; quantisation; and encoding. The digital data is then forwarded to subsequent hardware/software units for further processing, storage or transmission. In this thesis, I only pay attention to sampling and, to less extent, quantisation, whereas no discussion will be provided for encoding, as it is a pure digital representation of the

quantised data and therefore has nothing to do with sampling randomisation or aliasing mitigation.

Although ADCs are incredible and indispensable systems in todays' digital life, unfortunately, they have some drawbacks. Indeed, they are power-hungry and relatively time-consuming devices. Meaning, for remote or mobile applications with no direct or continuous access to main power sources, ADCs must be carefully selected to save batteries from rapid discharge. Additionally, being time-consuming devices, traditional ADCs have a limited operating bandwidth and, hence, their maximum uniform conversion rates cannot be exceeded. On the one hand, for a specific bandlimited analog signal with highest frequency component at f_m, for example, a considered ADC's uniform sampling rate must be equal to $2f_m$, at least, to avoid undesired aliasing effects [2]. On the other hand, even if the ADC is technically able to oversample the same signal at rates largely higher than $2f_m$, this means, again, more power consumption and more processing time. Hence, there is a trade-off between the required/affordable sampling rate of a given ADC and its power consumption/computational complexity.

Nonuniform sampling [23]–[25], including random and pseudorandom forms [12], emerged several decades ago to mitigate negative effects of aliasing while reducing power consumption and/or computational complexity. Classical restrictions imposed on the sampling rate (Nyquist rate) can be relaxed by using suitable random sampling techniques and subsequent processing algorithms. Nothing for free! This relaxation comes on the expense of other statistical errors due to this irregularity in sampling. With this stochastic approach, filter outputs, Fourier transforms, and other processing blocks become random and have to be statistically estimated. Now there is a new trade-off between the quality of estimates and the power consumption/computational complexity of the systems/devices that adopt and manifest sampling randomisation.

Next section shows a suggested taxonomy of sampling-related terminology normally used in the research literature, despite the fact that there is no official standard for such classification. Rather, some terminology in this area are used interchangeably, ambiguously or even inconsistently.

Afterwards, I demonstrate how uniform sampling and classical DSP work in order to better comprehend and *taste* the impact of using random sampling in DASP applications. By providing comparisons between the two approaches, we know exactly what the opportunities and challenges each approach has. Consequently, most of this chapter will be devoted to review the research literature on different aspects, techniques and algorithms of both approaches.

In Section 2.3, several uniform sampling-based areas, directly related to this research work, are revisited. This includes Shannon sampling theorem, aliasing, Fourier transform (FT), signal filtering, Lagrange interpolation and signal reconstruction. While in the nonuniform sampling section, Section 2.4, I demonstrate, in addition to the aforementioned topics, several random sampling and quantisation techniques along with their impact on the aliasing problem. In Section 2.5, I provide a quick comparison between common time-based randomisation schemes. Finally, Section 2.6. explores the literature on processing unsmooth signals, that is signals with limited number of discontinuities in their function-representations and/or their derivatives.

2.2. Sampling Terminology

Fig. 1 shows a suggested taxonomy of sampling schemes according to the time spacing between the sample points. As illustrated, there are different terminology for the same sampling scheme used in the research literature.

Throughout this thesis, I use *uniform sampling* term to denote equally spaced time instants of the samples, whereas the term *nonuniform sampling* is used to represent any irregularly spaced sampling schemes. Thus, nonuniform sampling may refer to random, pseudorandom or even hybrid (uniform + random/pseudorandom) sampling. The context will implicitly or explicitly assign which one is meant at a particular point.

Furthermore, by traditional/classical/conventional terms I mean uniform-based approaches. Accordingly, traditional DSP refers to the concepts, techniques and algorithms of acquiring, processing, transmission, detection or reconstruction of signals with equally spaced sample points. Whereas DASP approach will be used for processing nonuniformly sampled signals with the ability of completely eliminating the undesirable effects of aliasing or reducing it considerably.

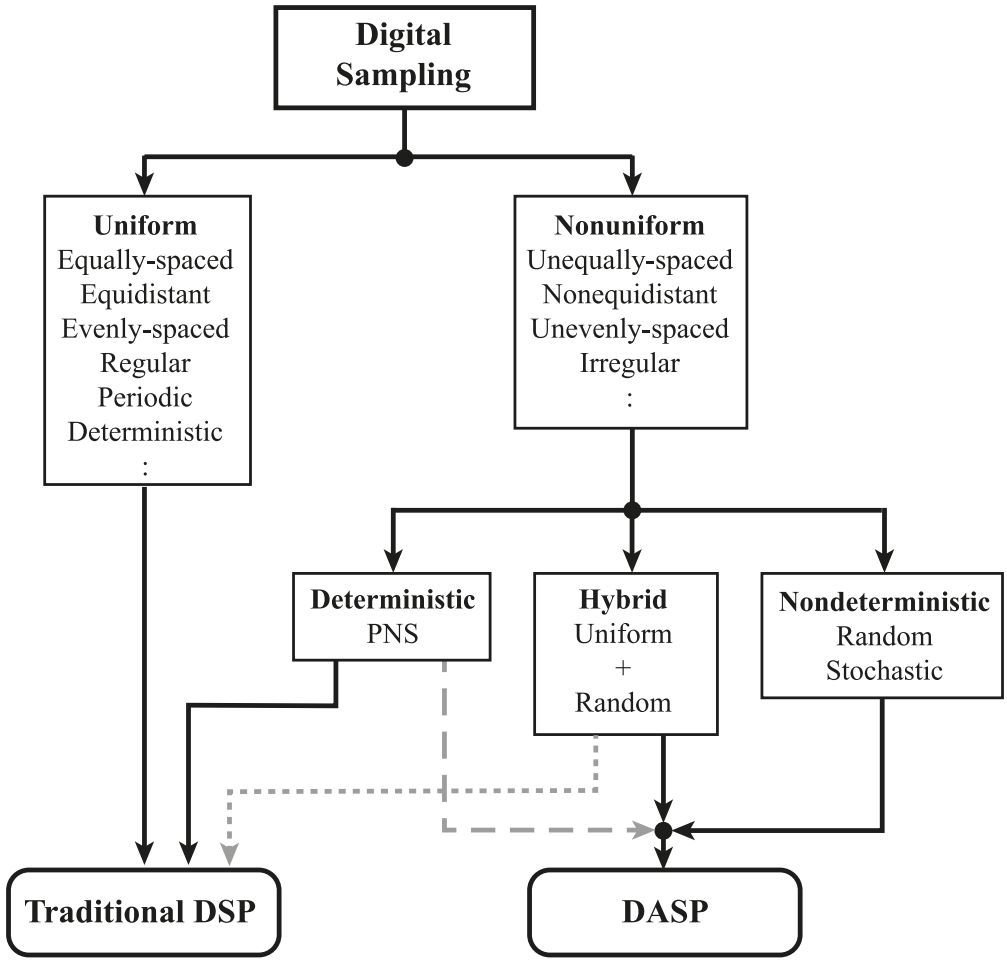


Fig. 1. Sampling taxonomy, with different terminology found in the literature. Some sampling schemes have features from both DSP and DASP depending on the details of the applied sampling pattern, and the dashed lines mean weaker relation, usually.

2.3. Uniform Sampling

In uniform sampling schemes, continuous-time analog signals are sampled periodically at a constant rate known as sampling frequency, F_s. Therefore, the sampling period is $T_s = 1/F_s$. The sampling process is usually represented by multiplying the continuous-time analog signal with a train of equally spaced versions of the Dirac delta function, $\{\delta(t - kT_s)\}_{k\in\mathbb{Z}}$, as shown in Fig. 2. The result is a sequence of equidistant data, $\{x(kT_s)\}$, which is also known as uniform discrete-time signal (or just the discrete signal). In this section, I revisit several topics in the research literature involving (or based on) uniform sampling.

2.3.1. Sampling Theory

Whittaker–Kotelnikov–Shannon (WKS) sampling theorem, [1], [26], [27], stipulates that the maximum time spacing, T_s, between two consecutive samples of a uniformly sampled bandlimited baseband signal, having $2f_m$ double-sided bandwidth, should be less than $1/2f_m$ for the signal to be perfectly reconstructed from such samples. Thus, the minimum sampling frequency, $F_s = 1/T_s$, that would achieve the above reconstruction condition is called the Nyquist rate. Indeed, this condition is sufficient but not necessary. Though, alias replicas of spectral components exceeding half the Nyquist rate appear in the bandwidth of the signal, $[-f_m, f_m]$, if the utilised scheme of sampling is uniform and the sampling rate is less than the Nyquist's.

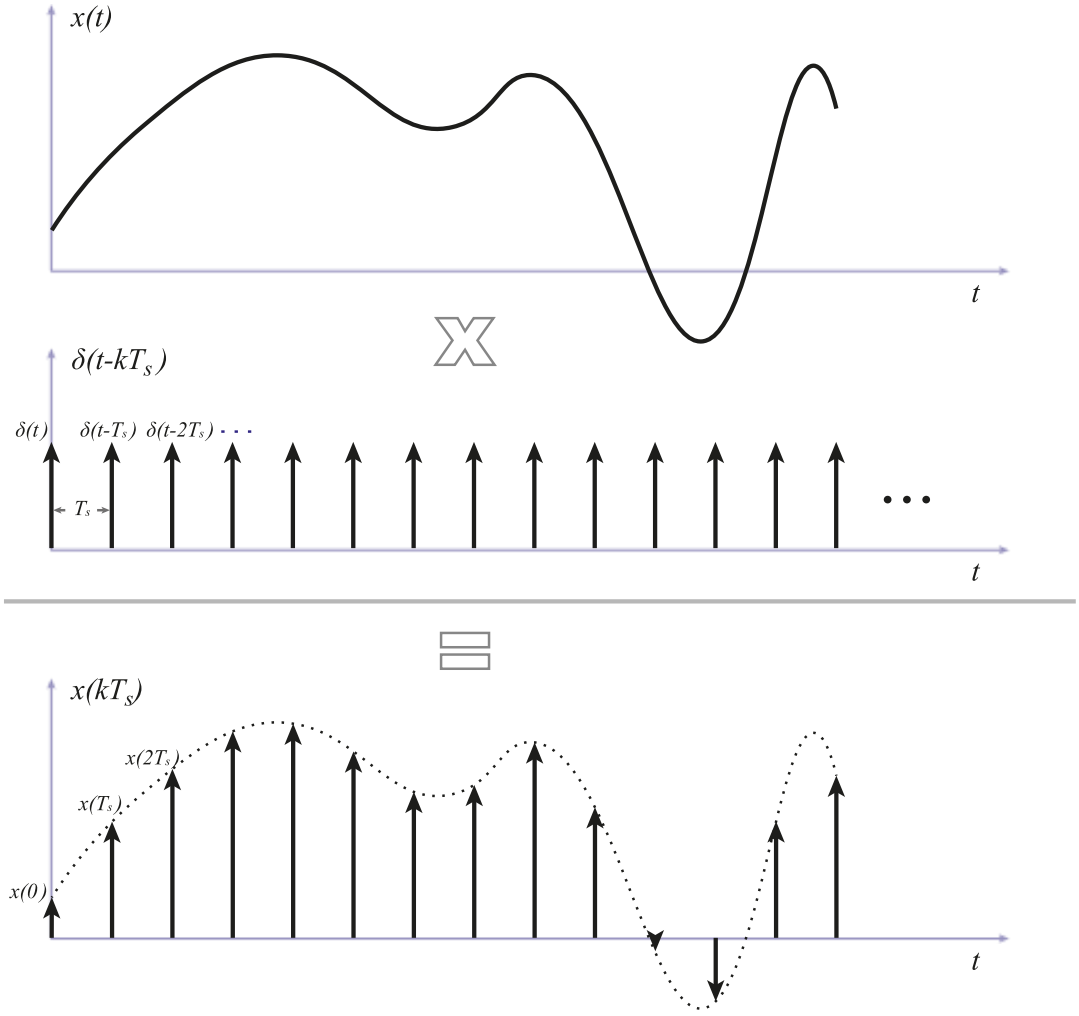


Fig. 2. Uniform sampling example, where the continuous-time analog signal, $x(t)$, is sampled by being multiplied with a train of equally spaced and shifted versions of the Dirac delta function $\{\delta(t - kT_s)\}_{k\in\mathbb{Z}}$. The resulting sequence of data, $\{x(kT_s)\}$, is the discrete-time sampled signal.

For the case of bandlimited bandpass signal with *known* bandwidth, B, and central frequency, f_c, and having lower and upper frequencies, $f_L = f_c - B/2$ and $f_U = f_c + B/2$, respectively (i.e. $B = f_U - f_L$ and $f_U > f_L$), then the cost of sampling the signal uniformly with a frequency rate more than $2f_U$, which is equivalent to the Nyquist rate for sampling a baseband signal with a maximum frequency $= f_U$, is very high and unaffordable. Therefore, another cost-effective approach of regularly sampling bandpass signals has to be established. Fortunately, such approach was also considered in the research literature, making use of the constructive effects (advantages) of aliasing phenomenon, where the lowest spectral replica of the bandpass signal is sampled instead. In [28], it is shown that the minimum uniform sampling rate (Nyquist rate) of the bandpass signal is $\frac{2f_U}{n}$, where n is the maximum integer number of the quotient $\frac{f_u}{B}$, i.e. the number of bandwidths where the upper frequency of the multiband signal spectrum is located away from the 0Hz origin (DC component). Therefore, the theoretical Nyquist rate is equal to $2B$ only if the bandpass signal is exactly shifted by a whole integer number, otherwise, it could reach up to $4B$. Moreover, Vaughan et al in [4] showed that the required uniform sampling frequency, F_s, to ensure no overlapping of positive and negative components is $\frac{2f_U}{n} \leq F_s \leq \frac{2f_L}{n-1}$.

Furthermore, in the case of sparse multiband signals, applying the classical bandpass sampling approach, with Nyquist rate $= 2B$ or $2f_U/n$, is not cost-effective, since the actual spectrum of this signal is comparatively wide with low spectrum occupancy rate. Hence, Landau suggested a theoretically lower sampling rate than the Nyquist rate for a given multiband signal [3]. For example, if the multiband signal consists of sub-bandwidths $(B_1, B_2, B_3, \dots, B_M)$ as shown in Fig. 3, then Landau minimum sampling rate is equal to $2(B_1 + B_2 + \cdots + B_M)$. But in this case, there is no guarantee that aliasing will not occur using uniform sampling. However, aliasing could be avoided (or reduced dramatically) by simply using proper random sampling techniques and subsequent processing units with average sampling rate approaching the minimum Landau rate. Remark, however, that this is only applicable if the spectral support function is fully known, otherwise the minimum achievable sampling rate is double the Landau rate [9].

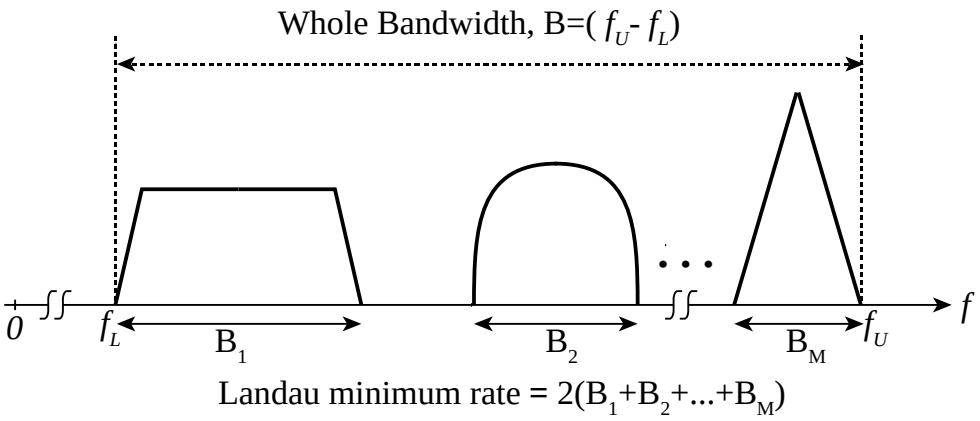


Fig. 3. Landau minimum sampling rate of multiband signal, with sub-bandwidths $B_1, B_2, B_3, \ldots, B_M$, is equal to $F_{s\_\min\ \_L} = 2(B_1 + B_2 + B_3 + \ldots + B_M)$. If only f_L and f_U are known but not the spectral support of the signal, then the Nyquist rate $= 2f_U/n$. If f_L is exactly positioned at the DC component (0Hz), then the Nyquist rate is equal to $2B = 2f_U$. If f_L is also unknown, then the Nyquist rate $= 2f_U$, i.e. the highest of all.

2.3.2. Signal Reconstruction

In traditional DSP, where continuous time signals are uniformly sampled and digitized for further stages of processing, one of the main concerns is how to reconstruct the signal faultlessly without loss of information. Fundamentally, this can be done using the *sinc* interpolation method [1], [29]. A bandlimited baseband signal $x(t)$ can be perfectly reconstructed from its equally spaced sequence of samples, $\{x(kT_s)\}_{k\in\mathbb{Z}}$, using (2.1), if the uniform sampling frequency conforms to the Nyquist rate,

$$x(t) = \sum_{k=-\infty}^{\infty} x(kT_s)\ sinc\left(\frac{t}{T_s} - k\right), \tag{2.1}$$

where $sinc(t) = \sin(\pi t)/\pi t$.

Since the FT of the *sinc* function is the *rect* boxcar function in the frequency domain, and because the above formula is a convolution in the time domain, therefore, it is equivalent to a multiplication in the frequency domain, meaning that the discrete sequence of signal samples is being filtered with an ideal lowpass filter.

Furthermore, signal reconstruction can also be carried out using noisy samples from bandlimited or non-bandlimited signals [30]. However, the area of sampling, filtering

and reconstruction of signals in the presence of noise is out of the scope of this paper and will be examined in a future research work, as indicated in Chapter 7.

2.3.3. Fourier Transform

The philosophy behind Fourier series is that any periodic signal, $x_p(t)$, which is piecewise continuous on the interval $[0, T_p]$ with a fundamental frequency, $f_p = 1/T_p$, can be decomposed into an infinite number of uniform sinusoidal signals having frequencies integer multiples of its fundamental frequency [31], that is

$$x_p(t) = a_0 + \sum_{k=1}^{\infty} a_k \cos(k\omega_p t) + b_k \sin(k\omega_p t), \tag{2.2}$$

where $a_0 = \frac{1}{T_p}\int_0^{T_p} x_p(t)dt$ is the DC component (average) of the signal, $\omega_p = 2\pi f_p$ and a_k and b_k are given by

$$a_k = \frac{2}{T_p}\int_0^{T_p} x_p(t)\cos(k\omega_p t)\,dt\,, \qquad k \neq 0, \tag{2.3}$$

$$b_k = \frac{2}{T_p}\int_0^{T_p} x_p(t)\sin(k\omega_p t)\,dt. \tag{2.4}$$

While Fourier series represents periodic signals using linear combination of orthogonal sinusoids, FT is usually used to analyse aperiodic signals in frequency domain. It can also be used, mathematically, with any signal type, but for periodic signals it reduces to the exponential form of Fourier series. For instance, the periodic function, $x_p(t)$, can be synthesized from its FT coefficients, $\{X_p(f_k)\}$, by using this series $x_p(t) = \sum_{k=-\infty}^{\infty} X_p(f_k)e^{j2\pi f_k t}$, where $j = \sqrt{-1}$. In general, the FT and the inverse FT of any function $x(t)$ are calculated using (2.5) and (2.6), respectively,

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi f t}dt. \tag{2.5}$$

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{\mathrm{j}2\pi f t} df. \tag{2.6}$$

Advantages of the FT are numerous. Converting signals from time domain to frequency domain, or vice versa, helps in understanding the nature of the signal and its properties. FT is heavily used in communication systems and channel estimation. Filtering of signals cannot be properly carried out without knowing minimal spectral occupancy information of the signal to be filtered, i.e. the bandwidth for a baseband signal and the bandwidth and centre frequency for a bandpass signal. Thus, it is very important to compute (or estimate) the FT of the processed signal in a large collection of applications in both environments; DSP and DASP.

2.3.4. Filtering of Digital Signals

Digital filtering is a well-established area in traditional DSP. For a given N-sample size sequence (i.e. a truncated discrete-time input signal), $\{x(kT_s)\}_{k=1}^{N}$, with $T_s = T/N$ uniform sampling period within the interval $[0, T]$, and a filter impulse response $h(t)$, then the filter output, $y(t)$, can be computed by using the discrete convolution operation

$$y(t) = T_s \sum_{k=0}^{N-1} x(kT_s)\, h(t - kT_s). \tag{2.7}$$

In real-life applications, the output signal is observed in discrete form as well. Thus, $y(t)$ is normally calculated at nT_s-apart time instants, with n is an integer. It follows from (2.7) that

$$y(nT_s) = T_s \sum_{k=0}^{N-1} x(kT_s)\, h(nT_s - kT_s). \tag{2.8}$$

Normalising by the sampling period T_s, (2.8) can be reinterpreted as

$$y(n) = \sum_{k=0}^{N-1} x(k)\, h(n - k). \tag{2.9}$$

Later in the next section, we will see that (2.8) and (2.9) can be tweaked a little bit to cover the more general case of unequally spaced sample points.

A close look at (2.8), one can intuitively see that the value of the discrete output signal at a given time instant is actually a numerical integration problem that uses the *Rectangular* (or *midpoint*) rule, having a T_s width, to accumulatively calculate the area formed by the product function inside the convolution sum. Many other approaches to estimate the same area are discussed in the research literature [32]. One approach of my interest is to use Lagrange polynomial interpolation [33], since it usually exhibits faster convergence rates than the Rectangular rule.

2.3.5. Uniform Interpolation and Composite Simpson's 1/3 Rule

Interpolation is a mathematical tool to find a curve (e.g. polynomial or other) that best fits for a set of equally or nonequally spaced points. It is commonly used in numerical analysis to integrate/sum a function that is only evaluated at a discrete sequence of values [34], [35]. The result then represents an estimate of the area under the curve (AUC) of the integrand/summand function. The error of estimation depends on many factors, such as the characteristics of the function itself and its derivatives, the number and locations of available evaluation points, and the used interpolation method.

As mentioned above, the interpolated points (or nodes) could be equally spaced or not, but in general, interpolating evenly distributed points is more accurate. Sometimes, the uniform data points are not available, hence, nonuniform interpolation techniques have to be considered. Nevertheless, this sub-section focuses only on interpolation of equally spaced points using Lagrange polynomials, while interpolation of nonuniformly spaced samples will be addressed in the next section.

Several uniform interpolation techniques are widely used in numerical integration area. I demonstrate here one of them related to the family of closed Newton–Cotes interpolation formulas [36]. Namely, composite Simpson's 1/3 rule, which will be studied to show how it is related to my proposed nonuniform C3NS rule, to be introduced in Chapter 4.

A function, $f(x)$, can be estimated according to the regular Simpson's 1/3 interpolation rule if there are available three discrete sample points of $f(x)$ acquired at x_0, x_1 and x_2.

To this end, a unique second-order polynomial, $P(x) = a_0 + a_1 x + a_2 x^2$ where $\{a_0, a_1, a_2\}$ are the polynomial coefficients, which exactly passes through the function values $f(x_0)$, $f(x_1)$, and $f(x_2)$, is formulated. This polynomial is often referred to as the Lagrange polynomial and is illustrated in Fig. 4.

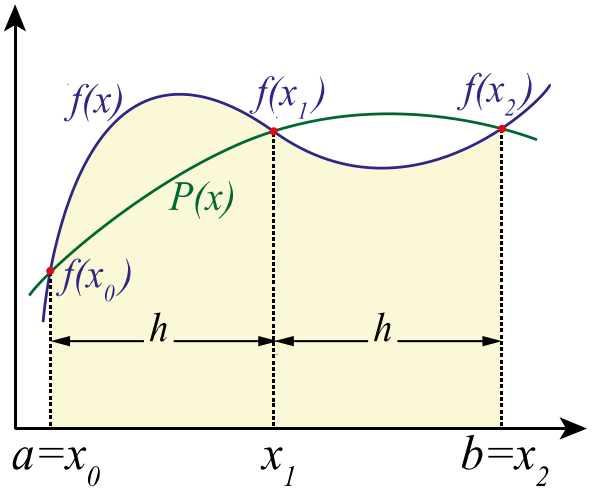


Fig. 4. The second-order parabolic function $P(x)$ is an approximate interpolation curve for $f(x)$. $P(x)$ exactly passes through the three equally spaced samples of $f(x)$ taken at x_0, x_1 and x_2. Simpson's 1/3 rule suggests that the area under the curve of $P(x)$ within the interval $[a, b]$ is an estimate of that of $f(x)$ within the same interval.

With the parabolic equation becomes known, it is easy now to estimate the AUC of $f(x)$ in a finite interval, $[a, b]$, by means of the definite integral of $P(x)$ from a to b. In [32], it has been shown that this integral is equal to

$$\int_a^b f(x)dx \approx \frac{h}{3}\big(f(x_0) + 4f(x_1) + f(x_2)\big), \tag{2.10}$$

where $x_0 = a,\ x_1 = \frac{b+a}{2},\ x_2 = b$ are the three samples' x-axis coordinates of $f(x)$, and $h = \frac{b-a}{2}$ is the uniform sampling (evaluation) step or the segment width [37].

The error of estimation in Simpson's 1/3 rule is bounded by the maximum value of the expression

$$Err_{Simp1/3} = -\frac{h^5}{90} f^{(4)}(\xi), \tag{2.11}$$

where $f^{(4)}$ is the fourth derivative of $f(x)$ and ξ is a number in the open interval (a, b).

To better approximate the function $f(x)$ and decrease the estimation error in (2.11), a larger sequence of samples has to be used. Consequently, the whole interval $[a, b]$ is partitioned into n equally spaced subintervals of $2h$-wide each, where h is the segment's width. Then, the regular Simpson's 1/3 rule is successively applied (n times) to calculate the AUC of $f(x)$ again, this time with more accurate estimate. Assume that the new sequence of samples comprises N values, then the segment width $h = \frac{b-a}{N-1} = \frac{b-a}{2n}$, where $a = x_0$ and $b = x_{2n}$, as shown in Fig. 5.

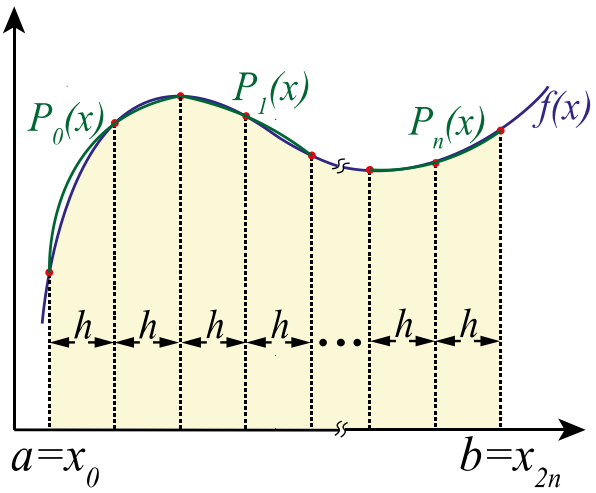


Fig. 5. Composite Simpson's 1/3 rule, where multiple Lagrange polynomials have been used to approximate $f(x)$ over the whole interval from a to b. Note that the notation for a and b here is different than those of Fig. 4.

The new numerical estimate of the AUC of $f(x)$, within the interval $[a, b]$ using composite Simpson's 1/3 rule, is given by

$$\int_a^b f(x)dx \approx \frac{h}{3}\left(f(x_0) + 4\sum_{i=1}^{n} f(x_{2i-1}) + 2\sum_{j=1}^{n-1} f(x_{2j}) + f(x_{N-1}) \right). \tag{2.12}$$

The error of estimation using this composite rule is apparently the sum of all individual errors from each subinterval $[x_{2i-2}, x_{2i}]$, $i = 1, \dots, n$. Thus

$$Err_{CSimp1/3} = -\frac{h^5}{90}\sum_{i=1}^{n} f^{(4)}(\xi_i), \tag{2.13}$$

where $\xi_i \in (x_{2i-2}, x_{2i})$, provided that the fourth derivative exist and bounded.

If the worst-case error (i.e. the absolute maximum error) of a specific subinterval is to be generalised over all other subintervals, the total maximum absolute estimated error is simplified in (2.14). Alternatively, we may be interested in finding the average estimated error across all subintervals, $\overline{f^{(4)}(\xi)}$, then (2.13) will be simplified to (2.15). Indeed, the total estimation error shown in (2.15) is more accurate than that of (2.14), since for large number of subintervals, individual estimated errors may cancel each other for their signs could be positive or negative. Lastly, the error in (2.15) is rewritten in a new form, (2.16), to indicate its relationship with the total number of samples N,

$$Max.Abs.Err_{CSimp1/3} = \frac{h^4(b-a)}{180}\max\left|f^{(4)}(\xi)\right|, \tag{2.14}$$

$$Av.Err_{CSimp1/3} = -\frac{h^4(b-a)}{180}\overline{f^{(4)}(\xi)}, \tag{2.15}$$

$$Av.Err_{CSimp1/3} = -\frac{(b-a)^5}{180(N-1)^4}\overline{f^{(4)}(\xi)}, \tag{2.16}$$

where $\xi \in (a, b)$.

2.4. Nonuniform Sampling

Signals are sampled nonuniformly either for accidental (not deliberate) or intentional reasons. On the one hand, there exist some signals that naturally have no uniform presence in the time domain (or any other 1D domain). For example, in astronomy, star luminosity can't be always tracked because of weather, geophysical or equipment failure conditions. In IT and computer networks, packets are received and queued in

bursts [38]. In stock market, prices can be event triggered, and this often occur nonuniformly. In WSN, Ad hoc networks and other multiple access systems that share common channel, data transmission occur in random times. Furthermore, NUS is used in some applications in medicine, such as computed tomography (CT) scan generated by the magnetic resonance imaging (MRI), where signals under consideration are not available at an equidistant samples [39].

NUS is also used in case of missing data problem in wireless communications systems [40]–[42]. This problem occurs when some samples are lost from originally uniformly sampled signals due to poor signal to noise ratio (SNR), multipath fading or receiver hardware malfunction.

Sometimes, the hardware sampling circuit suffers from high temperature or manufacturing imperfection that leads to what so-called sampling jitter [43], where the intended uniform samples deviate from the equidistant time instant pattern with a small random time [44], [45]. The overall sampling instants in this case look like random ones, hence, NUS is employed here to overcome this issue and process the signal further. However, Tarczynski et al showed in [44] that the sampler jitter worsens the accuracy of FT estimation, especially at higher frequencies. They also suggested some techniques to decrease the negative effects of the sampling jitter.

On the other hand, we may intentionally utilise nonuniform sampling in certain situations and applications to gain technical advantages and/or reduce the cost [46]. Such applications are found in areas of wireless communications, signal processing, filter design, Fourier transform, wideband spectrum sensing, radar and automotive industry [47], [48]. The main reasons for performing nonuniform sampling in such cases are to relax some restrictions imposed by the classical DSP hardware or software implementations, or to post-process already collected and stored data. For example, to save memory and storage resources, NUS can be used to compress the data, hence, reducing the overall cost.

Yen [23], Shapiro and Silverman [24], Masry [25], Bilinskis and Mikelsons [12], Marvasti [47], and Tarczynski and Allay [44] proved that carefully designed NUS schemes and algorithms can supress aliasing even if the sampling rates do not conform to the Nyquist rate. In such literature, proofs were presented for when a given sampling

method could result in aliasing or not, depending on how the random sampling process is performed, its probability density function (PDF) and characteristic function, the number of samples, and the observation window interval.

2.4.1. Alias-Free Random Sampling Techniques

As stated above, aliasing happens when two different continuous-time signals have the same sequence of samples. So, a sampling scheme plays a vital role in mitigating the aliasing problem, with random sampling techniques are more immune to aliasing than others [24], [25], [49]–[51]. In general, any stationary random sampling point process that satisfies some basic conditions, such as sampling all parts of the signal of interest with equal probability across the whole observation window [12, pp. 75–77], should overcome aliasing or reduce its harmful effects significantly even if the processed signal's spectrum extends beyond half the average sampling rate. Other definitions for alias-free sampling can be found in [25] where a stationary point process, characterising a specific random sampling scheme, is considered alias-free if no two continuous-time signals, belonging to the class of processed signals, have the same covariance measure.

In this thesis, three main nonuniform sampling techniques are heavily studied. I shed light on the research literature to give a brief description about total random (ToRa), stratified (StSa) and antithetical stratified (AnSt) random sampling techniques.

2.4.1.1. ToRa

In ToRa sampling technique, the sampling instants are randomly distributed across an observation time interval $I = [0, T]$, as shown in Fig. 6. The random time process has a uniform distribution PDF $p_{\tau\_ToRa}(\tau) = 1/T$ if $\tau \in I$ and zero elsewhere. In Chapter 3, we will see how selecting such a PDF guarantees unbiasedness of the ToRa estimator.

This method is sometimes called simple random sampling (SRS) or simple Monte Carlo sampling [52]. Anyhow, I adopt here the ToRa name for consistency with my published articles and papers. Remark that this sampling scheme can be viewed from two different perspectives which lead to the "same" overall result.

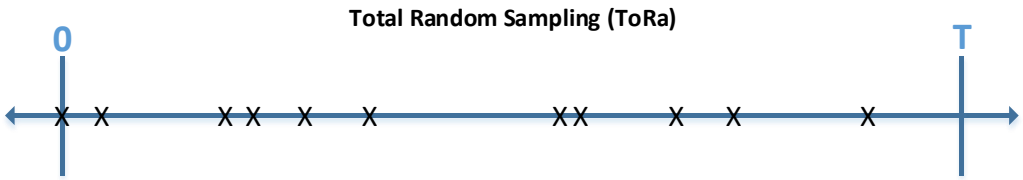


Fig. 6. An 11-point ToRa example, where the sample points are taken randomly from the time period of consideration using a uniform distribution random time process.

On the one hand, all the N time instants of the sample points within the observation window, I, are *randomly* selected at once, i.e. simultaneously, by using a uniform distribution random process with $p_{\tau\_ToRa}(\tau)$ PDF. Then, a suitable processing algorithm uses these randomly sampled points and estimate a particular sought-after value.

On the other hand, only one sampling point is randomly collected across the I window with the same PDF, as before, and then this procedure is repeated N-times to compute the average value of the estimate, hence the Monte Carlo name. Both methods are analytically identical.

In a discrete-only case (on-grid oversampling), ToRa can be implemented by *pseudorandomly* selecting N samples from otherwise uniform dense grid (with a total of N_g equally spaced possible samples) with a uniform probability mass function (PMF) of $p_{m\_ToRa}(m) = 1/N_g$. This specific case is also considered in Chapter 3.

2.4.1.2. StSa and AnSt

These two NUS techniques depend on the notion of what so-called *stratification*, in which the observation time interval, I, is divided into a number of strata (subintervals). Strata could have equal or different lengths, it depends on the relevant sampling scheme and what kind of signal it is. A function used to calculate the strata lengths is suggested in [52]. Though, equidistant strata approach could be practically used, especially in the case of unknown signals, and still exhibit excellent estimation results. Indeed, this will be the adopted approach for all discussions about stratification-based techniques. Fig. 7 illustrates how the sample points in these alias-free random schemes

are selected, taking into account that the strata borders (limits) themselves are chosen equally spaced.

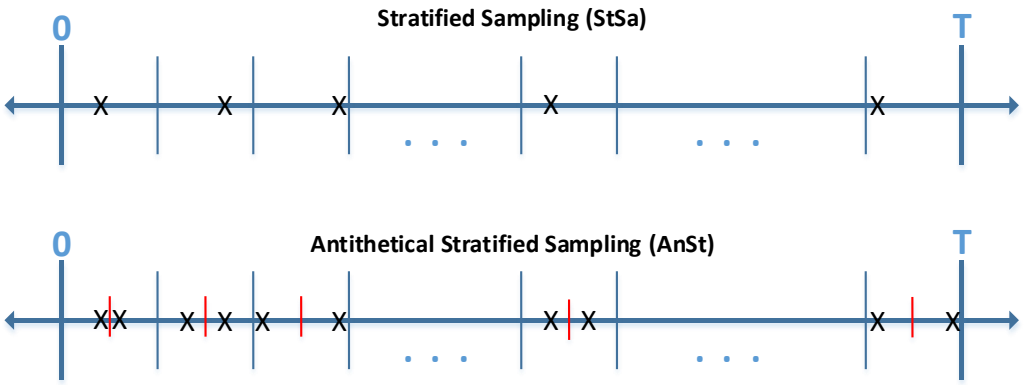


Fig. 7. StSa and AnSt random sampling stratification-based techniques.

The main differences between these two NUS methods are in the number and location of sample points within each stratum. While StSa uses only one randomly selected point inside each stratum, AnSt uses two points per single stratum [53]. One of these points is selected the same way as in StSa, while the second point is selected such that it is exactly located at the mirror reflection point of the first one when taking the centre of the stratum as the mirror line (i.e. line of symmetry), as indicated by the red lines in Fig. 7. It should be noted that AnSt uses double the number of sample points as StSa does if they have the same number of strata within a given observation window $[0, T]$. However, to have a fair comparison between the two sampling techniques, using the same observation window, I should select the same number of sample points for both techniques. Consequently, the stratum length in StSa should be half of AnSt's, in this case.

I will examine these two stratified techniques in all the following chapters, where these techniques are used to acquire the needed discrete-time data for the filtering estimation analyses in both main cases for the input signal/filter impulse response: continuous-time and piecewise-continuous functions.

2.4.1.3. Other Nonuniform Sampling Schemes

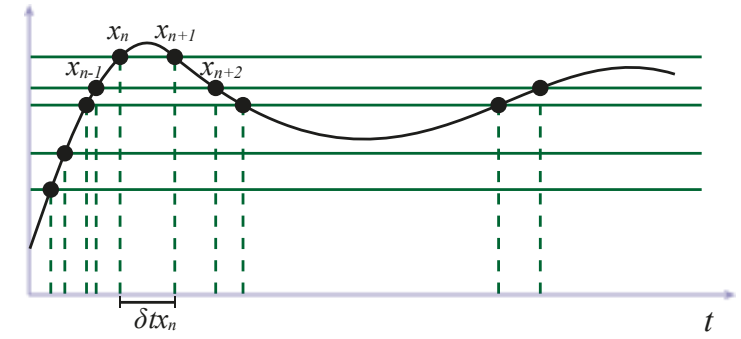
A variety of other nonuniform sampling techniques were also introduced in the research literature. However, not all of them are random/pseudorandom. Meaning, they could be nonuniform but deterministic, or having a mixture of uniform and random/pseudorandom samples as indicated in Fig. 1, above. Hence, their effectiveness in mitigating the problem of aliasing has to be investigated separately, but, certainly, they have other advantages in terms of power dissipation, computational complexity, or implementation simplicity.

2.4.2. Irregular Quantisation

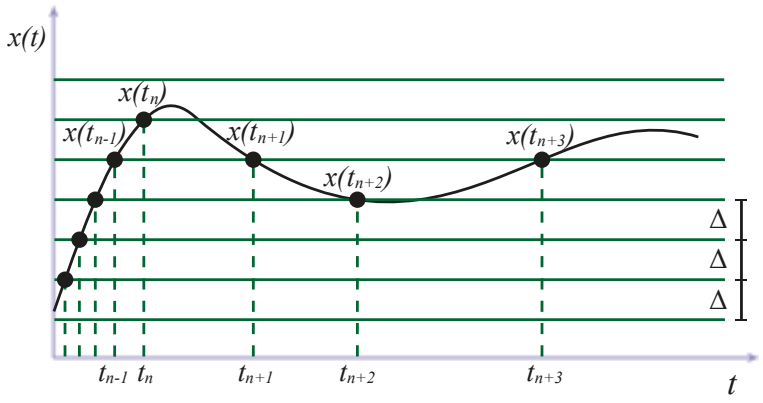
All nonuniform sampling schemes discussed, so far, depends on the randomness of the time instants at which the sample points of an input analog signal are acquired. Nonetheless, there are also several nonuniform quantisation techniques that have been investigated in the literature. I explore a few of them below for the sake of inclusiveness and comparison with time-based random processes.

2.4.2.1. Level Crossing

In level crossing sampling technique, the signal is sampled on the events when its amplitude crosses specific levels or thresholds, and not on uniformly timed fashion [54], [55]. This sampling method is best fit for applications where high-resolution timers are available, and so, it relaxes the need for precision amplitude quantizers. Fig. 8(a) shows an illustration of level crossing sampling with five predefined amplitude levels. The principle of amplitude-based event sampling is not new. Actually, it is another form of send-on-delta sampling scheme, widely used in wireless sensor networks to save energy consumption, as depicted in Fig. 8(b) [56]. The genuine difference between the two amplitude-based sampling schemes is the distribution of amplitude levels. They are nonuniformly distributed in level crossing scheme, whereas the send-on-delta scheme’s levels are equally spaced. In addition, both of them inherited the concept of amplitude-based event triggering from the well-known zero crossing sampling technique [57], [58], which has the same principle of sampling but with only one amplitude threshold; the zero level.



(a)



(b)

Fig. 8. The principle of amplitude-based sampling: (a) level crossing and (b) send-on-delta sensor reporting.

Note that the randomness of samples' time instants in level crossing scheme depends on the sampled signal itself. For the example depicted in Fig. 8(a), it is obvious that time instants of the samples are random. However, imagine that the signal being sampled is a fixed periodic sawtooth signal or even a sinusoid! Definitely, we will obtain a nonuniform sampling sequence, but not random as well. Therefore, its ability to mitigate aliasing problem is then questionable, although it certainly helps reduce the overall power consumption of the sampling circuit, which can be achieved through triggering the sampling events only when there is a change in amplitude and enters a "sleep mode" if the signal's amplitude is constant or changes slightly.

2.4.2.2. Peak Detection Sampling

This sampling scheme depends only on the local minima or maxima of a given signal, but not on time or amplitude levels [55]. It is good for sampling of signals where certain important peaks have to be detected, such as ECG heart beat signals. At the same time, it is highly sensible to noise, where it generates too much samples even if there is no real change in the original sampled signal. Moreover, it may miss-detect a lot of considerable signal changes when it is monotonically increasing or decreasing.

A combination of level crossing and peak detection sampling schemes is also considered in [55], where it is shown that this arrangement would benefit from the advantages of both schemes together. For example, it is more immune to small noise due to the way it is carried out, where level crossing is applied first, and then peak detection. This helps get rid of the noise ripple before detecting the peaks. However, it still suffers from some drawbacks presented in any of the two sampling schemes, especially when there are small changes of signal amplitude and no minima or maxima are available within a considerable time period of the signal.

2.4.2.3. Slope Sampling or Linear Decimation.

Both terms refer to the same principle of delta-surface-area sampling, where it is assumed that pre-samples are already exist (but not necessarily kept or stored) [55]. Fig. 9 shows an illustration of slope sampling, where initially three samples are taken to form a triangle. If the surface area of the triangle is more than a pre-determined threshold, then the first and second samples are kept, and the second sample will serve as the initial sample for the next surface area calculation. But if the triangle area is under the threshold, a fourth sample has to be considered, and the polygon area calculated and compared to the threshold again, and so on. While it is more effective than level crossing when the signal amplitude variations are small thanks to the concept of small errors add up, but still has a drawback of the necessity to collect too many samples at the first glance (even that some of them will be neglected), and also, it needs more computational complexity.

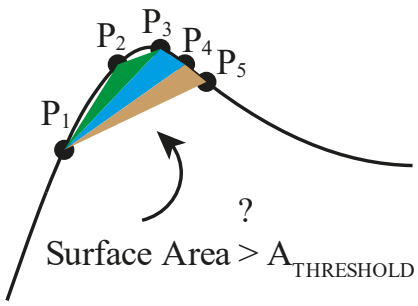


Fig. 9. Slope sampling (or linear decimation) principle. First, three sample points are collected and the triangle $P_1P_2P_3$ surface area is calculated and compared to a preassigned threshold area, $A_{THRESHOLD}$. If it is not greater than the threshold area, another point P_4 is collected and the new polygon $P_1P_2P_3P_4$ area is examined to see if it is greater than $A_{THRESHOLD}$, otherwise, P_5 is acquired, and so on.

2.4.3. Digital Alias-free Signal Processing (DASP)

DASP is a promising technology to sample and process signals digitally without the harmful effect of aliasing. Bilinskis et al works [11] introduced new techniques and algorithms in an attempt to overcome/decrease the harmful effects of aliasing. Noticeable advancements have been achieved in this area by utilising specific randomised/pseudorandomised sampling and quantisation techniques [59]–[61]. Therefore, wide range of frequency spectrum have been made detectable with reduced sampling frequency (lower than the Nyquist rate). But on the other hand, randomization of sampling and/or quantisation has led to other types of errors (statistical) in the digitising process. Therefore, the main focus of DASP is to maximize the advantages of randomisation and minimise the drawbacks of such statistical errors.

The statistical errors in DASP are mainly produced as a result of sampling irregularities [12, pp. 58–59]. That is why nonuniform deterministic sampling has been introduced, where some deterministic approaches are embedded within the sampling process to decrease the statistical and probabilistic errors.

Furthermore, Bilinskis and his colleagues have not only proposed theoretical framework for DASP, but they developed and analysed structures, models and algorithms that have been built on general-purpose embedded systems [11, p. 396]. For example, a method called sequential component extraction (SECOEX), which can

be fit inside a firmware of an embedded system, was presented and evaluated. Significant improvements in estimating the power spectrum using this method were achieved depending on the nonuniform sampling approach. Estimation of the most powerful spectrum coefficients takes place in a repetitive manner. Each time the strongest components are stiffed out (subtracted) from the remaining signal, starting for the first time with the original signal and repeating until the remaining power is less than a specific level, where the estimation process comes to the end.

2.4.4. Reconstruction of Nonuniformly Sampled Signals

Reconstruction of nonuniformly sampled signals has attracted many researchers in the last few decades [7], [62]–[65]. Inspired by their reduced computational costs, NUS-based circuit elements and software algorithms were developed and employed in many DASP applications. The conventional signal reconstruction methods based on uniform sampling techniques, shown above, cannot be used directly to the NUS case. Either some amendments should be added, or new designs must be provided.

A NUS generalised reconstruction filter based on a kernel $K(t, t_n)$, or $g(t, t_n)$ [66], was introduced in [67], where the performance of nine NUS reconstruction algorithms were investigated and compared. The proposed reconstruction formula is

$$x(t) = 2B \sum_{n=-\infty}^{\infty} c_n \, g(t - t_n) = 2B \sum_{n=-\infty}^{\infty} c_n \, sinc(2B(t - t_n)) \tag{2.17}$$

where B is the single sided bandwidth of the lowpass (baseband) signal, and c_n are coefficients to be determined by a specific reconstruction algorithm. Or equivalently, in terms of the kernel $K(t, t_n)$ which is a unique reciprocal basis of $\{ g(t, t_n)\}$,

$$x(t) = \sum_{n=-\infty}^{\infty} x(t_n) K(t - t_n) \tag{2.18}$$

Recovery of nonuniformly sampled signals in the presence of noise was examined in [65] based on quasi-random sampling (or quasi-Monte Carlo). Pawlak et al introduced a consistent algorithm to reconstruct a noisy signal from its irregularly sampled points. Nonetheless, this will be further investigated in a future work, as listed in Chapter 7.

Other reconstruction algorithms for multiband and non-bandlimited signals, either with or without a priori knowledge of their spectral support functions, have also been investigated in [8], [9] . Although digging deeply into these algorithms are out of the scope of this thesis, but in general, there are a variety of kernels involved in the reconstruction process that can serve quite similar to the popular *sinc* kernel, such as the Gaussian and exponential spline kernels. Conditions for guaranteed reconstruction have been introduced in each case. For instance, in spectrum-blind reconstruction of multiband signals [8], it is proved that a multiband signal with known f_L and f_U, having a non-zero fraction upper-bounded spectral occupancy rate, and sampled with a universal pattern, similar to the PNS, achieving the minimum Landau rate can be reconstructed even though the full spectral support is not known in advance. Whereas in [9], the blind reconstruction scheme of the multiband signal requires that the number of bands and their widths are known beforehand, and the algorithm is based on compressive sensing framework.

2.4.5. Compressive Sensing

Blind and non-blind sparse-spectrum sensing techniques, based on nonuniform sampling, have also been considered in the research literature. Compressive sensing (or compressed sensing) is used to directly compress the sampled signal during the sampling process itself [68]–[71]. In many image, audio and video applications, a lot of data samples that have been collected using regular sampling techniques contain no much details, and so, can be neglected or squeezed. Indeed, this is what usually done when conducting image compression or zipping, for example. The idea of compressive sensing is to decrease the number of samples and compress the data just before the samples being collected.

Compressive sensing is also used to reconstruct a signal from its sparse-sample representation, even if the number of samples doesn't fulfil Nyquist rate. There are two conditions to make the reconstruction possible: *high signal sparsity* and *incoherence*. However, several theoretical prerequisites for this approach are needed before it can be implemented practically. The high computational cost of the reconstruction algorithm and the high SNR are just two of them.

2.4.6. Spectrum Estimation

Tarczynski et al in [44] have inspected nonuniform sampling techniques, algorithms and applications with regard to spectral estimation of Fourier transform. They presented the weighted sample (WS) and the weighted probability (WP) density functions, as well as other methods to tackle aliasing and reduce the effect of sampling jitter. Mathematical expressions for the discrete nonuniformly sampled signal's spectral estimation have been derived, along with formulae for the mean-squared error (MSE) of the estimated spectra in each case. A guaranteed rate of uniform convergence of $1/N$ for ToRa technique was proved. This was a measure of how fast the number of nonuniform sample points, $2N$, can affect the accuracy of the estimation. In addition, supressing of aliasing and improved discrete Fourier transform (DFT) estimation has been shown in the simulation results.

Periodic nonuniform sampling (PNS) and weighted periodic nonuniform sampling (WPNS) are investigated in [5], [6], [72], [73], where using of repetitive patterns and weighted repetitive patterns of nonuniform sampling showed improved spectrum estimation and alias suppression results. Formulae for finding the optimal sampling sequences have been introduced, despite the fact that finding such optimal solutions in real-time applications is considered as time-consuming and may increase the complexity of sampling process. PNS is also known as multicoset sampling in literature, and it is more useful in multiband signal sampling with sparse spectrum.

Masry presented the utilisation of random StSa technique to estimate the FT of a continuous-time deterministic signal [52]. He proved that StSa-based estimator of the FT converges at a rate of $1/N^3$, provided that the processed analog signal, $x(t)$, and its first-order derivative are continuous-time smooth functions. Later, Masry et al published another article [53], this time they employed the AnSt random sampling technique to estimate the FT. A faster rate of convergence, $1/N^5$, was achieved for the variance of the estimator if $x(t)$ and its FOD and SOD are all continuous.

A new NUS method to estimate the FT was then proposed by Ahmad et al [74]. They called it hybrid-stratified (HySt) sampling technique. The new technique was proved to be unbiased and fast converging. Indeed, it was shown that the HySt FT estimator converges uniformly at a rate of $1/N^5$ if the first three derivatives of the analysed

signal, $x(t)$, exist and continuous. Furthermore, they showed that other estimators' (ToRa, StSa and AnSt) rates of uniform convergence are all the same and exactly equal to $1/N$, but the rate of *pointwise* convergence is the same as proposed in the respected literature papers. Indeed, Ahmad et al showed that the uniform convergence of StSa and AnSt estimators is only established after a certain number of sample points that is mainly dependent on the highest frequency of the sampled signal.

The idea behind HySt sampling scheme is also based on stratification, like both StSa and AnSt. However, in HySt, N random sample points are taken exactly as StSa, in addition to other N+1 deterministic sample points selected as the strata borders themselves, forming a mixture of nonuniform and uniform sample points, hence, the name hybrid. In the provided analytical and numerical comparisons of the four FT estimators, just mentioned above, it was clearly shown that HySt is the most effective one in the sense of computation performance and precise FT estimation results for the same number of sample points. Moreover, a closed form mathematical expression for the optimal stratifying function is derived. This function is used to calculate strata time-limits (borders) for all stratification-based sampling schemes. However, since more than half of the sample points in this approach are equally spaced, aliasing may occur if the total number of sample points is relatively small.

2.4.7. Filter Output Estimation

Practical designs of uniform-based digital filters started with the launch of DSP applications, in the middle of last century [75], [76]. Research and applications in this vital area never stopped throughout the decades till now. Filter designs, techniques, algorithms, and implementations are developed continuously [77], [78]. However, the case is different for NUS, where less dense research, designs and implementations are available in the literature that directly address the topic of filtering nonuniformly sampled signals.

In [79], new algorithms of NUS-based FIR filter was introduced. Thanks to modifying filters' coefficients on each sampling time instant, the proposed filter was efficient in the sense that it suppressed aliasing and minimized the energy of the error signal of the filter's output.

On the grounds of an interesting class of irregular sampling (asynchronous) ADCs [80], an analogue for asynchronous FIR filtering was derived in [81]. The implemented NUS-based filter combined the concepts of asynchronous technology and irregular sampling. A significant reduction in the power consumption of the proposed asynchronous filter design was achieved, in addition, the computational complexity of the new design was proven to be much lower than that of the regular FIR filter. The notion behind this new filter architecture was to properly compute the convolution between NUS input signal and the impulse response of the filter based on a resampling scheme of their both time instants. This was required to synchronise the time instants of the two sequences, leading to feasible and sensible convolution for the filtering process.

Other researches also introduced algorithms and techniques for building NUS filters [82]–[84]. The proposed designs or frameworks mainly based on randomised quantisation techniques, such as level crossing, peak detection, and slope sampling, or lookup tables of the discrete impulse response. Each approach of interpolation used in these papers defines its own concept of dealing with irregularly spaced samples and depends on some presumed conditions.

One of the challenges in filtering NUS signals is how to align the time instants of signal samples with those of the impulse response of a given filter. Different methods to deal with this issue have been proposed in the papers just cited above. Although the overall performance of the introduced filter designs and algorithms was good and, to some extent, alias-free, I believe there is still a lot to do. Hence, this work is to further advance the research in this area and bridge unresolved gaps.

2.4.8. Interpolation of Unequally Spaced Samples

Lagrange interpolation polynomials can be generalised to include unequally spaced nodes (samples) [47, p. 124]. The continuous and differentiable function $f(x)$ with $n+1$ nonuniformly distributed nodes $\big(x_0, f(x_0)\big), \big(x_1, f(x_1)\big), \dots, \big(x_n, f(x_{\mathrm{n}})\big)$ can be approximated with a general n-degree polynomial

$$P_n(x) = \sum_{k=0}^{n} P_k(x), \tag{2.19}$$

where each $P_k(x)$ is calculated by

$$P_k(x) = f(x_{\mathrm{k}}) \prod_{\substack{i=0 \\ i \neq k}}^{n} \frac{x - x_i}{x_k - x_i}. \tag{2.20}$$

This form of nonuniform interpolation was used in the research literature to reconstruct bandlimited continuous-time signals from their irregularly spaced discrete-time sequences [49].

However, the optimal least-squares interpolation of a sequence of nonuniformly spaced samples of a bandlimited signal $x(t)$, i.e. $\{x(t_n)\}_{n=1}^{N}$, with bandwidth B is proved by Yen [23] as to be equal to

$$\hat{x}(t) = \sum_{n=1}^{\mathrm{N}} \left(x(t_n) \sum_{m=1}^{N} \gamma_{mn}\, K(t, t_m) \right), \tag{2.21}$$

where $K(t, t_m) = sinc\big(2B(t - t_m)\big)$ and the coefficient γ_{mn} is the (m, n)th element of the inverse of the matrix $\boldsymbol{K}$ whose elements are $K(t_n, t_m),\ n, m = 1,2, \dots, N$.

2.5. Pros and Cons of Randomised Sampling Techniques

The following table, Table I, shows a quick comparison between common random sampling techniques addressed in literature. Note that this table excludes any randomisation techniques rely on quantisation and level crossing for two reasons: first, it is out of the scope of this thesis; and second, comparison between two different concepts is neither fair nor accurate. Hence, I focus in Table I on pros and cons of sampling techniques that acquire sample points at nonuniform time-instants generated by either deterministic or random time processes.

TABLE I: PROS & CONS OF TIME-BASED NONUNIFORM SAMPLING TECHNIQUES FOR SMOOTH SAMPLED SIGNALS AND THEIR DERIVATIVES

Sampling Technique	Pros	Cons
ToRa	• Easy to implement. • Convergence behaviour begins immediately with even low N. • Nonsmoothness of ZOD has no effect on the convergence rate. • Can mitigate aliasing.	• Slow uniform convergence rate, N^{-1}. • Requires sorting time instants before implementation.
StSa	• Implementation is easier than StSa. • Faster convergence rate than ToRa, N^{-3}. • Can mitigate aliasing.	• Requires more implementation cost than ToRa, (stratification). • Fast convergence behaviour begins after certain number of points depending on the sampled signal characteristics.
AnSt	• Implementation is easier than HySt and CS. • Faster convergence rate than StSa, N^{-5}. • Can mitigate aliasing.	• Requires more implementation cost than StSa, (antithetical sampling). • Fast convergence behaviour begins after certain number of points, usually after StSa does for the same sampled signal characteristics.
HySt	• Implementation is easier than CS. • More immune to signal high frequencies than others. • Faster convergence rate than StSa, N^{-5}. • Can mitigate aliasing, however, it is not guaranteed at lower average random sampling rates.	• Requires more implementation cost than AnSt, (offline coefficients). • Fast convergence behaviour begins after certain number of points, usually after StSa does for the same sampled signal characteristics.
CS	• Blind sampling and reconstruction are viable. • Faster convergence rate than HySt at low signal frequency and high spectrum sparsity. • Can mitigate aliasing.	• Requires extensive implementation cost, (iterative algorithm). • Convergence behaviour degrades as signal frequency increases or spectrum sparsity decreases.

2.6. DSP and Signal Smoothness

The research done so far on processing unsmooth signals is really scarce. Discontinuities in the processed signals may occur in several real-life applications. In many electronics applications, power signals are converted from AC to DC. The conversion procedure normally involves rectification of an analog continuous-time waveform into unsmooth waveform (i.e. in terms of its derivatives). Meaning, discontinuities appear in one or more orders of the derivatives. Similar phenomenon applies when clipping signals by some electronic components like diodes and transistors. Sawtooth, square-wave and other types of sharp-transition-based signals contain many discontinuities. Broader applications comprise signals with discontinuities can also be found in communication signals (BPSK, QAM), digital data, event-triggered signals, sampling and quantisation [85], and stock market response to global events such as COVID-19 pandemic [86].

Filtering unsmooth signals were barely considered in the research literature. Some publications have been found in digital audio processing [13] and image up-sampling [87]. Chapter 4 and Chapters 5 extend the research on this field by proposing StSa- and AnSt-based filtering estimation examples of piecewise-smooth signals. The asymptotic behaviours of such filter estimators are considered, where the statistical properties such as the mean, bias, variance, convergence rate and consistency are investigated. Several special cases for the nonsmoothness characteristics of piecewise-smooth input signal, windowing function, impulse response of the filter, or their derivatives are examined.

Part I

Continuous-Time Integrand/Summand Functions of the Filter Convolution Operation

CHAPTER 3

Randomised Digital Filtering of Continuous-Time Input Signals

3.1. Overview

In this chapter, I consider filtering of randomly sampled continuous-time analog signals using ToRa, StSa and AnSt random sampling techniques. As stated above, ToRa is simply a Monte Carlo (MC)-based averaging technique across the whole observation window and a specific number of iterations. While both StSa and AnSt rely on the notion of stratification of the observation window before conducting the random sampling procedure.

First, I introduce the filter model and associated notation. Next, three filter estimators, based on the above random sampling techniques, are established, assuming equally spaced strata for both StSa- and AnSt-based estimators. A generalised form that encompasses the three estimators is also provided for conciseness reasons, where applicable. The estimators are then analysed, and statistically assessed. Particularly, I devise mathematical expressions for the variance and the convergence rate in each case of the three estimators. The estimators are proven to be unbiased and consistent but converging to the true value of the filter output at different speeds.

At the end of the chapter, I demonstrate numerical examples that covers several types of continuous-time input signals and filters. It is clear that the simulation results reflect the analytical findings, especially, the uniform convergence rates of the estimators after certain amount of sample points.

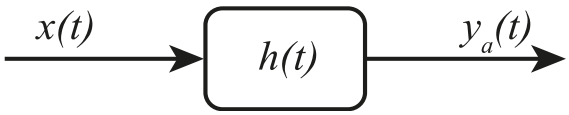
By comparing the performance of the estimators in terms of asymptotic convergence rates, we will see that the AnSt-based estimator is the fastest, with a rate of N^{-5}, where N is the number of strata of the AnSt technique (i.e. twice the number of samples). Whereas the ToRa-based estimator is the slowest, with a convergence rate of only

N^{-1}. Finally, the middle one is the StSa-based estimator, which converges at a speed of N^{-3}.

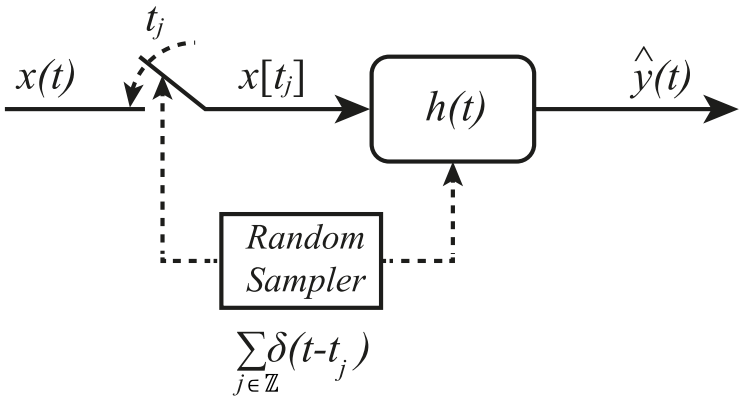
3.2. Analog Filtering Model

The input analog signal, $x(t)$, is assumed to be real-valued, bandlimited, and integrable. As a bandlimited signal, $x(t)$ is smooth and so are its derivatives (specifically, ZOD, FOD, and SOD are all continuous-time functions). Assume also the filter is a continuous-time, linear, time-invariant (CT LTI) system with bounded and symmetric impulse response, $h(t)$, as depicted in Fig. 10a. Then, the output analog signal, $y_a(t)$, is given by (3.1), where $*$ denotes the convolution operation.

$$y_a(t) = (x * h)(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau = \int_{-\infty}^{\infty} x(t-\tau)h(\tau)d\tau, \tag{3.1}$$



(a)



(b)

Fig. 10. Filtering models: (a) Analog filtering, where $x(t)$, $h(t)$, and $y_a(t)$ are the continuous-time input signal, the filter impulse response, and the output signal respectively; and (b) Discrete random filter estimator. Note that in both models the output signal, in time domain, is the convolution of the input signal and the impulse response.

Let us assume that we are interested in a truncated version of (3.1) where the input signal is continuously observed within a T-length sliding window, i.e. $[t-T, t]$. Thus, we have

$$y(t) = \int_{t-T}^{t} x(\tau)h(t-\tau)d\tau. \tag{3.2}$$

For the sake of simplicity, the integrand function of (3.2) is denoted by $f(t,\tau)$, i.e.

$$f(t,\tau) := x(\tau)h(t-\tau). \tag{3.3}$$

Hence, (3.2) simplifies to

$$y(t) = \int_{t-T}^{t} f(t,\tau)d\tau. \tag{3.4}$$

3.3. FIR Filter Estimators

The output signal, $y(t)$, in (3.4) can be approximated using Tora-, StSa-, and AnSt-Based filter estimators, as shown in Fig. 10b. Such estimators are assumed to use causal, linear phase, symmetric FIR filter designs with a finite duration of T sec. The input to the filters is a sequence of randomly sampled points of the integrand function, $f(t,\tau)$. Since the three considered filter estimators have different sampling schemes, then I assume that the total number of sample points is fixed for all estimators and is equal to $2N$. Thus, comparison between the estimators' performance would be fairer and more realistic.

Sampling the integrand function, at any given t time delay, is carried out by a random point process. This is basically equivalent to multiplying $f(t,\tau)$ with a finite train of irregularly spaced Dirac delta functions, as given in (3.5),

$$\{f(t,\tau_j)\}_{j=1}^{2N} = \sum_{j=1}^{2N} f(t,\tau)\delta(\tau-\tau_j). \tag{3.5}$$

But how to select the random time instants, $\{\tau_j\}_{j=1}^{2N}$, themselves? To answer this question, I need to know how each of the three considered random sampling techniques works. Although this is briefly introduced in the literature review chapter (Chapter 2), I demonstrate here the mechanism of each scheme with much detail and illustrations.

The ToRa-based estimator uses a total of $2N$ samples chosen randomly from the observation window, the StSa-based estimator also utilises $2N$ randomly sampled points, but each point should be selected from each of the $2N$ different strata, whereas in AnSt-based estimator case, only half of the points ($= N$) are randomly selected from the N strata (one sample per each stratum), while the other half are essentially their correlated antithetical counterparts, as shown in Fig. 11. Note that I can interpret ToRa as if it is a stratification-based technique using just one stratum, but averaged across the whole $2N$ samples, i.e. Monte Carlo averaging. If this is so then when conducting the numerical examples, I carry out *double* Monte Carlo simulations for ToRa-based estimator; the first one is for the $2N$ samples and the other is for averaging over a specific number of iterations, say V. However, only one Monte Carlo averaging is carried out for both StSa and AnSt-based estimators, that is averaging over the V iterations.

Now, it is time to introduce the three filter estimators, as per the assumptions above, the notation given at the beginning of the thesis, and the notation listed in Table II. Note that $\tau_{((k-1)L+j)} \sim \mathrm{U}(S_{j-1}, S_j)$ is a random variable that has a uniform distribution with PDF equals to

$$p_j(\tau) = \begin{cases} 1/\Delta, & \tau \in A_j \\ 0\ , & \text{elsewhere} \end{cases}. \tag{3.6}$$

Indeed, $p_j(\tau_j)$ has the same format for the three considered estimators but with different interpretations for Δ (= the stratum length), as illustrated in Table II. Remark that C_j in $\tau^a_{((k-1)L+j)}$, the antithetical sampling instant, is $C_j = \left((j-1)\frac{T}{N} + j\frac{T}{N}\right)/2 = \left(j - \frac{1}{2}\right)\frac{T}{N} = \left(j - \frac{1}{2}\right)\Delta$. Note also that the subintervals are given by, $A_j = [S_{j-1}, S_j) \equiv [(j-1)\Delta, j\Delta) \equiv \left[C_j - \frac{\Delta}{2}, C_j + \frac{\Delta}{2}\right)$, which means that in ToRa scheme we have only

one subinterval and is equal to the whole observation interval, i.e. $A_j = A_1 = [S_0, S_1) = [t - T, t)$.

TABLE II: NOTATION OF THE THREE FILTER ESTIMATORS FOR EQUAL STRATA LENGTHS AND T-LENGTH OBSERVATION WINDOW

	Symbol	**ToRa**	**StSa**	**AnSt**
Monte Carlo (MC) iterations	M	$2N$	1	1
Number of sample points per MC iteration	R	1	$2N$	$2N$
Number of strata	L	1	$2N$	N
Stratum length	$\Delta = \frac{T}{L}$	T	$\frac{T}{2N}$	$\frac{T}{N}$
Number of sample points per stratum	$P = \frac{R}{L}$	1	1	2
MC averaging	$\frac{1}{M}$	$\frac{1}{2N}$	1	1
Random sampling instant in the j-th stratum and k-th MC iteration, where $k = 1, \dots, M$ and $j = 1, \dots, L$	$\tau_{(k-1)L+j}$	$\tau_{(k-1)L+j}$	$\tau_{(k-1)L+j}$	$\tau_{(k-1)L+j}$
Antithetical sampling instant in the j-th stratum (for AnSt only)	$\tau^a_{(k-1)L+j}$	---	---	$2C_j - \tau_{(k-1)L+j}$
Subinterval(s)	A_j	$[S_0, S_1)$	$[S_{j-1}, S_j)$	$[S_{j-1}, S_j)$

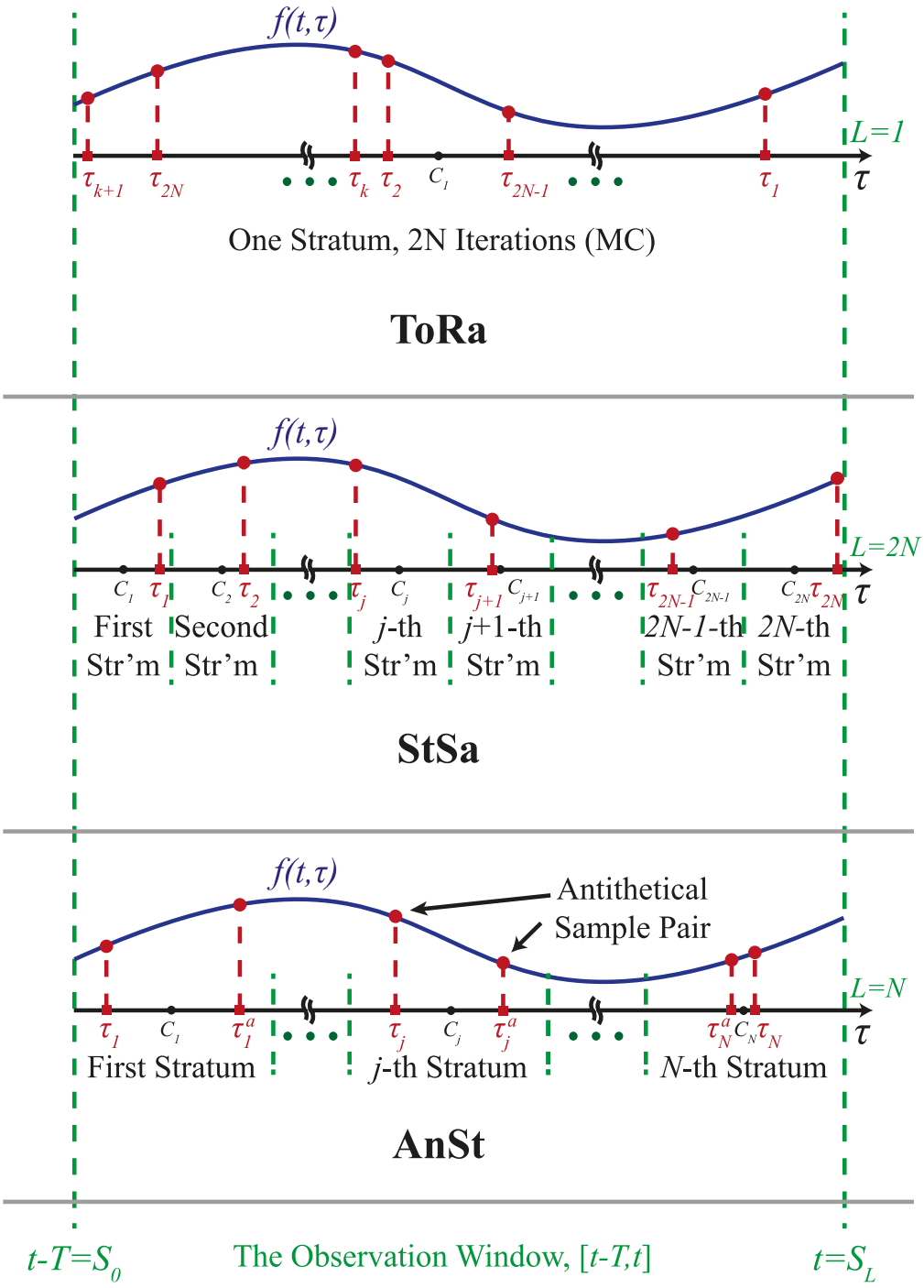


Fig. 11. The three considered random sampling schemes with the integrand function (solid dark blue). ToRa is considered as a unity stratum scheme but with $2N$ Monte Carlo averaging iterations. The StSa scheme uses $2N$ strata with one random sample per stratum. While AnSt scheme uses N strata and randomly selects the first sample in each stratum, say τ_j, and the second one, τ_j^a, is its antithetical counterpart. C_j's are the centres of the strata, with $j = 1$, $j = 1, \dots, 2N$, and $j = 1, \dots, N$ for ToRa, StSa, and AnSt respectively. The solid dark blue line is the smooth function $f(t, \tau)$ at a given t-shift time.

The ToRa-based estimator:

$$\hat{y}(t) = \frac{T}{2N}\sum_{k=1}^{2N}\sum_{j=1}^{1} f\left(t, \tau_{(k-1)L+j}\right), \tag{3.7a}$$

$$\hat{y}(t) = \frac{T}{2N}\sum_{k=1}^{2N} f(t, \tau_k). \tag{3.7b}$$

The StSa-based estimator:

$$\hat{y}(t) = \frac{T}{2N}\sum_{k=1}^{1}\sum_{j=1}^{2N} f\left(t, \tau_{(k-1)L+j}\right), \tag{3.8a}$$

$$\hat{y}(t) = \frac{T}{2N}\sum_{j=1}^{2N} f(t, \tau_j). \tag{3.8b}$$

The AnSt-based estimator:

$$\hat{y}(t) = \frac{T}{N}\sum_{k=1}^{1}\sum_{j=1}^{N}\left(\frac{f\left(t, \tau_{(k-1)L+j}\right) + f\left(t, \tau^{a}_{(k-1)L+j}\right)}{2}\right), \tag{3.9a}$$

$$\hat{y}(t) = \frac{T}{N}\sum_{j=1}^{N}\left(\frac{f\left(t, \tau_j\right) + f\left(t, \tau^{a}_{j}\right)}{2}\right). \tag{3.9b}$$

Based on the three individual estimators above, I propose a generalised form of the filter estimator that encompasses all these estimators. This general form is useful when deriving the mathematical expressions for the statistical properties of the estimators under consideration.

Generalised form of the filter estimator:

$$\hat{y}(t) = \frac{\Delta}{MP}\sum_{n=1}^{ML}\big(f(t,\tau_n) + (P-1)f(t,\tau_n^a)\big), \tag{3.10a}$$

$$\hat{y}(t) = \frac{\Delta}{MP}\sum_{k=1}^{M}\sum_{j=1}^{L}\Big(f\big(t,\tau_{(k-1)L+j}\big) + (P-1)f\big(t,\tau^a_{(k-1)L+j}\big)\Big). \tag{3.10b}$$

Remark that the subscript $n = (k-1)L + j$ never exceeds the total number of random samples in any selected approach, i.e. $n = 1, 2, 3, \dots, 2N$.

3.4. Bias Check

The following theorem proves that the estimator $\hat{y}(t)$ is unbiased, and the expected value of the estimator is equal to the true filter output signal. That is, $\mathbb{E}[\hat{y}(t)] = y(t)$.

Theorem 3.1. *The general filter estimator, given in (3.10b), is unbiassed for all t.*

Proof:

From (3.10b), we have

$$\mathbb{E}[\hat{y}(t)] = \mathbb{E}\left[\frac{\Delta}{MP}\sum_{k=1}^{M}\sum_{j=1}^{L}\Big(f\big(t,\tau_{(k-1)L+j}\big) + (P-1)f\big(t,\tau^a_{(k-1)L+j}\big)\Big)\right], \tag{3.11a}$$

$$= \frac{\Delta}{MP}\sum_{k=1}^{M}\sum_{j=1}^{L}\int_{-\infty}^{\infty} p_j(\tau)\Big(\big(f(t,\tau) + (P-1)f(t,\tau^a)\big)\Big)\,d\tau, \tag{3.11b}$$

$$= \frac{1}{MP}\sum_{k=1}^{M}\sum_{j=1}^{L}\int_{S_{j-1}}^{S_j} f(t,\tau)d\tau + \frac{(P-1)}{MP}\sum_{k=1}^{M}\sum_{j=1}^{L}\int_{S_{j-1}}^{S_j} f(t,\tau^a)d\tau \tag{3.11c}$$

Note that $\int_{S_{j-1}}^{S_j} f(t,\tau^a)d\tau = \int_{S_{j-1}}^{S_j} f(t,\tau^a)d\tau^a$. Hence, (3.11c) becomes

$$\mathbb{E}[\hat{y}(t)] = \frac{1}{MP}\sum_{k=1}^{M}\sum_{j=1}^{L}\int_{S_{j-1}}^{S_j} f(t,\tau)d\tau + \frac{(P-1)}{MP}\sum_{k=1}^{M}\sum_{j=1}^{L}\int_{S_{j-1}}^{S_j} f(t,\tau^a)d\tau^a \quad (3.11d)$$

$$= \frac{1}{MP}\sum_{k=1}^{M}\int_{t-T}^{t} f(t,\tau)d\tau + \frac{(P-1)}{MP}\sum_{k=1}^{M}\int_{t-T}^{t} f(t,\tau^a)d\tau^a, \quad (3.11e)$$

$$= \frac{My(t)}{MP} + \frac{M(P-1)y(t)}{MP}, \quad (3.11f)$$

$$\mathbb{E}[\hat{y}(t)] = y(t). \quad (3.11g)$$

■

The following corollary builds on top of Theorem 3.1 and concludes the unbiasedness of all proposed filter estimators.

Corollary 3.1. *The proposed ToRa, StSa, and AnSt filter estimators are all unbiassed.*

Proof:

As per the results of Theorem 3.1, the generalised form filter estimator is unbiased. Consequently, every single estimator comprising the generalised filter estimator is unbiased, as well.

■

3.5. Variance of the Generalised Form Filter Estimator

We have seen in the previous section that the considered estimators are unbiased. Therefore, the variance of the estimators is identical to the mean-squared error (MSE). Below, I devise mathematical expressions for the variance of each single estimator and determine its uniform convergence rate.

Remark that the generalised form filter estimator, (3.10a), can also be rewritten as

$$\hat{y}(t) = \frac{1}{M}\sum_{n=1}^{ML} \phi_n, \tag{3.12a}$$

$$\hat{y}(t) = \frac{1}{M}\sum_{k=1}^{M}\sum_{j=1}^{L} \phi_{(k-1)L+j}, \tag{3.12b}$$

where $\phi_n = \phi_{(k-1)L+j}$ is the k-th MC iteration and j-th stratum's contribution to the overall value of the estimator. In this section, I refer to ϕ_n as the n-th sub-estimator and is given by

$$\phi_n := \frac{\Delta}{P}\big(f(t,\tau_n) + (P-1)f(t,\tau_n^a)\big). \tag{3.13a}$$

$$\phi_{(k-1)L+j} := \frac{\Delta}{P}\Big(f\big(t,\tau_{(k-1)L+j}\big) + (P-1)f\big(t,\tau^a_{(k-1)L+j}\big)\Big). \tag{3.13b}$$

For conciseness and simplicity of analysis, I would like to supress the t (time shift) argument of all functions using it unless it is explicitly required. So, for example, the succinct form of the n-th sub-estimator is

$$\phi_n = \frac{\Delta}{P}\big(f(\tau_n) + (P-1)f(\tau_n^a)\big). \tag{3.13c}$$

$$\phi_n = \phi_{(k-1)L+j} = \frac{\Delta}{P}\Big(f\big(\tau_{(k-1)L+j}\big) + (P-1)f\big(\tau^a_{(k-1)L+j}\big)\Big). \tag{3.13d}$$

Since all sub-estimators are independent from each other, I find the variance of the general estimator, (3.13d), by adding up the individual variances of all sub-estimators. Therefore, I need to calculate the variance of $\phi_{(k-1)L+j}$, in the first place.

The expected value of ϕ_n is

$$\mathbb{E}\big[\phi_{(k-1)L+j}\big] = \mathbb{E}\left[\frac{\Delta}{P}\Big(f\big(\tau_{(k-1)L+j}\big) + (P-1)f\big(\tau^a_{(k-1)L+j}\big)\Big)\right], \tag{3.14a}$$

$$= \frac{1}{P}\int_{-\infty}^{\infty} p_j(\tau)\big(f(\tau) + (P-1)f(\tau^a)\big)\Delta\, d\tau, \tag{3.14b}$$

$$= \frac{1}{P}\int_{S_{j-1}}^{S_j} f(\tau)d\tau + \frac{P-1}{P}\int_{S_{j-1}}^{S_j} f(\tau^a)d\tau, \tag{3.14c}$$

$$= \frac{1}{P}\int_{S_{j-1}}^{S_j} f(\tau)d\tau + \frac{P-1}{P}\int_{S_{j-1}}^{S_j} f\big(2C_j - \tau\big)d\tau, \tag{3.14d}$$

The integrand function, $f(\cdot)$, is assumed to be continuous and square-integrable, and its first two derivatives exist and are continuous. So, it is plausible to use Taylor series expansion to approximate $f(\tau)$ and $f(\tau^a) = f\big(2C_j - \tau\big)$ about C_j, the central time instant of the j-th stratum. Consequently, equation (3.14d) can be rewritten as

$$\begin{aligned}\mathbb{E}\big[\phi_{(k-1)L+j}\big] = \frac{1}{P}\int_{S_{j-1}}^{S_j} \Big(& f\big(C_j\big) + \big(\tau - C_j\big)f'\big(C_j\big) + \frac{1}{2}\big(\tau - C_j\big)^2 f''\big(C_j\big) \\ & + o\left(\big|\tau - C_j\big|^2\right)\Big)d\tau \\ & + \frac{P-1}{P}\int_{S_{j-1}}^{S_j} \Big(f\big(C_j\big) + \big(C_j - \tau\big)f'\big(C_j\big) + \frac{1}{2}\big(C_j - \tau\big)^2 f''\big(C_j\big) \\ & + o\left(\big|C_j - \tau\big|^2\right)\Big)d\tau\end{aligned} \tag{3.15}$$

where $o(\cdot)$ is the little-o notation and $f'\big(C_j\big)$ and $f''\big(C_j\big)$ are the FOD and SOD, respectively, of $f(t,\tau)$ with respect to τ at $\tau = C_j$. Remark that $S_{j-1} = (j-1)\Delta = C_j - \frac{\Delta}{2}$ and $S_j = j\Delta = C_j + \frac{\Delta}{2}$ for equidistant stratification. Working out the integral in (3.15), we get

$$\mathbb{E}\big[\phi_{(k-1)L+j}\big] = \Delta f\big(C_j\big) + \frac{\Delta^3}{24}f''\big(C_j\big) + o(\Delta^3). \tag{3.16}$$

Denote by $e_n := e_{(k-1)L+j}$ the part of estimation error related to the n-th sub-estimator, so we have

$$e_{(k-1)L+j} = \phi_{(k-1)L+j} - \mathbb{E}\left[\phi_{(k-1)L+j}\right], \tag{3.17a}$$

$$\begin{aligned} & e_{(k-1)L+j} \\ & = \frac{\Delta}{P}\left(f\left(\tau_{(k-1)L+j}\right) + (P-1)f\left(\tau^{a}_{(k-1)L+j}\right)\right) \\ & - \left(\Delta f\left(C_j\right) + \frac{\Delta^3}{24}f''\left(C_j\right) + o(\Delta^3)\right). \end{aligned} \tag{3.17b}$$

The MSE of the sub-estimator (i.e. the variance $\mathbb{V}\left[\phi_{(k-1)L+j}\right]$) can be computed by finding the second moment of $e_{(k-1)L+j}$, taking into consideration that all elements (functions, derivatives, constants and arguments) of the n-th error term in (3.17b) are real-valued. Hence,

$$\mathbb{V}\left[\phi_{(k-1)L+j}\right] = \mathbb{E}\left[\left(e_{(k-1)L+j}\right)^2\right] = \int_{-\infty}^{\infty} p_j(\tau)\left(e_{(k-1)L+j}\right)^2 d\tau. \tag{3.18a}$$

$$\begin{aligned} & \mathbb{V}\left[\phi_{(k-1)L+j}\right] \\ & = \int_{(j-1)\Delta}^{j\Delta} \frac{1}{\Delta}\left(\frac{\Delta}{P}\left(f(\tau) + (P-1)f(\tau^{a})\right) - \Delta f\left(C_j\right) - \frac{\Delta^3}{24}f''\left(C_j\right)\right. \\ & \left. - o(\Delta^3)\right)^2 d\tau. \end{aligned} \tag{3.18b}$$

I expand both $f(\tau)$ and $f(\tau^{a})$ of (3.18b) about C_j using Taylor series approximation. After doing the expansion, rearranging, simplifying, and integrating we get the final expression for the variance of the generalised filter sub-estimator,

$$\mathbb{V}\left[\phi_{(k-1)L+j}\right] = \frac{(P-2)^2}{12P^2}\left(f'\left(C_j\right)\right)^2\Delta^4 + \frac{1}{720}\left(f''\left(C_j\right)\right)^2\Delta^6 + o(\Delta^6). \tag{3.19}$$

For ToRa and StSa, $P = 1$, so we deduce from (3.19) that the variance of the n-th sub-estimator is

$$\mathbb{V}[\phi_n] = \mathbb{V}\left[\phi_{(k-1)L+j}\right] = \frac{1}{12}\left(f'\left(C_j\right)\right)^2\Delta^4 + \frac{1}{720}\left(f''\left(C_j\right)\right)^2\Delta^6 + o(\Delta^6). \tag{3.20a}$$

$$\mathbb{V}\left[\phi_{(k-1)L+j}\right] = \frac{1}{12}\left(f'\left(C_j\right)\right)^2 \Delta^4 + o(\Delta^4). \tag{3.20b}$$

Remark that (3.20b) is obtained from (3.20a) as a consequence of the fact that

$$\frac{1}{720}\left(f''\left(C_j\right)\right)^2 \Delta^6 + o(\Delta^6) = o(\Delta^5) = o(\Delta^4). \tag{3.21}$$

In the case of AnSt estimator, $P = 2$, therefore (3.19) simplifies to

$$\mathbb{V}\left[\phi_{(k-1)L+j}\right] = \frac{1}{720}\left(f''\left(C_j\right)\right)^2 \Delta^6 + o(\Delta^6). \tag{3.22}$$

The devised general expression, (3.19), and its approach-specific forms, (3.20b) and (3.22), are associated with the j-th stratum and k-th MC iteration. Hence, to calculate the value of the overall variance across the whole observation window and the total Monte Carlo averaging iterations as given in Table II above, I need to accumulate the values of all sub-estimator variances, which requires us to consider each case separately.

3.5.1. Variance of the ToRa Estimator

The following theorem shows an original and exact mathematical expression for the variance of the ToRa filter estimator.

Theorem 3.2. *The variance of the ToRa filter estimator is* $\mathbb{V}[\hat{y}] = \frac{T^4}{24N}\left(f'\left(\frac{T}{2}\right)\right)^2 + o(N^{-1})$.

Proof:

Since all sub-estimators are independent, the variance of the sum of sub-estimators is equal to the sum of their variances. Consequently, we have from (3.12b) and (3.20b)

$$\mathbb{V}[\hat{y}] = \mathbb{V}\left[\frac{1}{M}\sum_{k=1}^{M}\sum_{j=1}^{L}\phi_{(k-1)L+j}\right] = \frac{1}{M^2}\sum_{k=1}^{M}\sum_{j=1}^{L}\mathbb{V}\left[\phi_{(k-1)L+j}\right], \tag{3.23a}$$

$$\mathbb{V}[\hat{y}] = \frac{1}{M^2}\sum_{k=1}^{M}\sum_{j=1}^{L}\left(\frac{1}{12}\left(f'(C_j)\right)^2\Delta^4 + o(\Delta^4)\right). \quad (3.23b)$$

The configuration parameters for ToRa listed in Table II assert that $M = 2N$, $L = 1$, and $\Delta = T$. This clearly indicates that there is only one stratum for all ToRa MC iterations, as j in (3.23b) is always equal to 1. Hence, the centre of the stratum for all MC iterations is $C_j = C_1 = \frac{T}{2}$. Substituting these parameters into (3.23b), we get

$$\mathbb{V}[\hat{y}] = \frac{1}{(2N)^2}\sum_{k=1}^{2N}\left(\frac{1}{12}\left(f'\left(\frac{T}{2}\right)\right)^2 T^4 + o(T^4)\right), \quad (3.24a)$$

$$\mathbb{V}[\hat{y}] = \frac{T^4}{24N}\left(f'\left(\frac{T}{2}\right)\right)^2 + o(N^{-1}), \quad (3.24b)$$

where (3.24b) is obtained from (3.24a) by observing that all arguments of the summation are independent of k and $\frac{2N \times o(T^4)}{(2N)^2} = o(N^{-1})$, as T is constant.

■

3.5.2. StSa Estimator's Variance

Theorem 3.3 below provides an exact expression for the variance of the StSa filter estimator.

Theorem 3.3. *The variance of the StSa filter estimator is* $\mathbb{V}[\hat{y}] = \frac{T^3}{96N^3}\sum_{j=1}^{2N}\left(\left(f'(C_j)\right)^2\Delta\right) + o(N^{-3})$.

Proof:

From (3.20b), which also applies to the StSa case, I obtained (3.23b). Back to StSa column of Table I, we see that $M = 1$, $\Delta = T/2N$, and $L = 2N$. Therefore, from (3.23b) we get

$$\mathbb{V}[\hat{y}] = \frac{T^3}{96N^3}\sum_{j=1}^{2N}\left(\left(f'(C_j)\right)^2\Delta + o(N^{-4})\right), \tag{3.25a}$$

$$\mathbb{V}[\hat{y}] = \frac{T^3}{96N^3}\sum_{j=1}^{2N}\left(\left(f'(C_j)\right)^2\Delta\right) + o(N^{-3}). \tag{3.25b}$$

■

3.5.3. AnSt Estimator's Variance

The following theorem is established for the variance of the AnSt filter estimator.

Theorem 3.4. *The variance of the AnSt filter estimator is* $\mathbb{V}[\hat{y}] = \frac{T^5}{720N^5}\sum_{j=1}^{N}\left(\left(f''(C_j)\right)^2\Delta\right) + o(N^{-5})$.

Proof:

Referring to (3.22) and (3.23a), and considering AnSt settings in Table II, especially, $M = 1$, $\Delta = T/N$, and $L = N$, I compute the overall variance of the AnSt-based filter estimator as

$$\mathbb{V}[\hat{y}] = \sum_{j=1}^{L}\left(\frac{1}{720}\left(f''(C_j)\right)^2\Delta^6 + o(\Delta^6)\right), \tag{3.26a}$$

$$\mathbb{V}[\hat{y}] = \frac{T^5}{720N^5}\sum_{j=1}^{N}\left(\left(f''(C_j)\right)^2\Delta\right) + o(N^{-5}). \tag{3.26b}$$

■

3.6. Consistency

Having devised the mathematical expressions for variances of the three considered filter estimators, I would now like to check the consistency of the estimators. Hence, the following theorem is established.

Theorem 3.5. *All considered filter estimators (ToRa, StSa, and AnSt) are consistent.*

Proof:

For an estimator to be consistent, its variance should converge to zero as the number of sample points, $2N$, approaches infinity. Thus, it is required to verify that the limits of expressions in (3.24b), (3.25b), and (3.26b) are all decaying to zero as $2N \to \infty$, which is equivalent to $N \to \infty$,

$$\textbf{ToRa: } \lim_{N\to\infty} \mathbb{V}[\hat{y}] = \lim_{N\to\infty} \left(\frac{T^4}{24N} \left(f'\left(\frac{T}{2}\right) \right)^2 + o(N^{-1}) \right) = 0, \tag{3.27a}$$

$$\textbf{StSa: } \lim_{N\to\infty} \mathbb{V}[\hat{y}] = \lim_{N\to\infty} \left(\frac{T^3}{96N^3} \sum_{j=1}^{2N} \left(\left(f'(C_j) \right)^2 \Delta \right) + o(N^{-3}) \right) = 0, \tag{3.27b}$$

$$\textbf{AnSt: } \lim_{N\to\infty} \mathbb{V}[\hat{y}] = \lim_{N\to\infty} \left(\frac{T^5}{720N^5} \sum_{j=1}^{N} \left(\left(f''(C_j) \right)^2 \Delta \right) + o(N^{-5}) \right) = 0. \tag{3.27c}$$

Since ToRa, StSa, and AnSt filter estimators are decaying to zero as the sample size is increasing to infinity, then all estimators are consistent.

■

So far, the estimators are verified to be unbiassed and consistent, which means that all of them are accurately approximating the filter output as the number of utilised sample points increases. Having said this, however, not all of them are converging at the same rate. Next, I find the exact decaying rate for each single estimator.

3.7. Convergence Rates

The convergence rates of the three filter estimators (ToRa, StSa and AnSt), as the number of sample points, $2N$, approaches infinity, are established in the following theorem.

Theorem 3.6. *As the number of sample points $2N$ approaches ∞, ToRa, StSa, and AnSt filter estimators converge at rates of N^{-1}, N^{-3}, and N^{-5} respectively.*

Proof:

Again, if $2N \to \infty$, then $N \to \infty$, as well. By using Riemann integral, it follows from (3.27a-c) that

ToRa:

$$\lim_{N\to\infty}(2N \times \mathbb{V}[\hat{y}]) = \lim_{N\to\infty} 2N \times \left(\frac{T^4}{24N}\left(f'\left(\frac{T}{2}\right)\right)^2 + o(N^{-1})\right), \tag{3.28a}$$

$$= \frac{T^4}{12}\left(f'\left(\frac{T}{2}\right)\right)^2. \tag{3.28b}$$

StSa:

$$\lim_{N\to\infty}((2N)^3 \times \mathbb{V}[\hat{y}])$$

$$= \lim_{N\to\infty}(2N)^3 \times \left(\frac{T^3}{96N^3}\sum_{j=1}^{2N}\left(\left(f'(C_j)\right)^2\Delta\right) + o(N^{-3})\right), \tag{3.29a}$$

$$= \frac{T^3}{12}\int_{t-T}^{t}\left(f'(\tau)\right)^2 d\tau. \tag{3.29b}$$

AnSt:

$$\lim_{N\to\infty}((2N)^5 \times \mathbb{V}[\hat{y}])$$

$$= \lim_{N\to\infty}(2N)^5 \times \left(\frac{T^5}{720N^5}\sum_{j=1}^{N}\left(\left(f''(C_j)\right)^2\Delta\right) + o(N^{-5})\right), \tag{3.30a}$$

$$= \frac{2T^5}{45}\int_{t-T}^{t}\left(f''(\tau)\right)^2 d\tau \equiv \frac{2T^5}{45}\int_{t-T}^{t}\left(f''(t,\tau)\right)^2 d\tau. \tag{3.30b}$$

As verified above, we notice that the AnSt-based filter estimator is converging at the fastest rate of N^{-5}, outperforming both StSa- and ToRa-based estimators. Moreover, the ToRa-based estimator is the slowest converging one amongst the three estimators

with a rate of only N^{-1}. Finally, StSa is the middle estimator in terms of speed of convergence, where N^{-3} decaying rate is achieved.

■

It should be noted that variance values of StSa and AnSt estimators depend on the FOD-squared and ZOD-squared. This means that if the underlying integrand function comprises some high frequency components (or ripple), then the limit values in (3.29b) and (3.30b) will be greater. This would shift the asymptotic line of convergence upward to indicate increased absolute MSE values. However, the convergence rates wouldn't be affected.

3.8. Almost Sure (Strong) Convergence

It is possible to prove that StSa and AnSt filter estimators, which converge in the mean faster than N^{-1}, converge strongly (almost surely, a.s.) to the filter output, whereas the ToRa filter estimator does not a.s. (strongly) converge to the target output.

The Borel-Cantelli lemma states that if

$$\sum_{N=1}^{\infty} \Pr(|\hat{y}_N - y| > \varepsilon) < \infty, \tag{3.31}$$

then

$$\Pr\left(\lim_{N\to\infty} \sup(|\hat{y}_N - y| > \varepsilon)\right) = 0, \tag{3.32}$$

where $\Pr(\cdot)$ denotes the probability, $\hat{y}_N$ is the filter estimator using N sample points, and $\varepsilon > 0$ is an arbitrary small number.

I use this lemma to analyse the convergence of the filter estimators $\hat{y}_N$ to the true output value y. According to this lemma, if I can show that (3.31) holds for any selected ε, then the probability of the sequence of events, $\{|\hat{y}_N - y| > \varepsilon\}$, being infinitely long is zero. Hence, I will observe it, at most, a finite number of times.

To prove that the last inequality, (3.31), is satisfied, I use the Chebyshev inequality, which states that the probability that a random variable $\hat{y}_N$ with variance σ_N^2 will differ from its mean value by more than ε is upper-bounded by

$$\Pr(|\hat{y}_N - y| \geq \varepsilon) \leq \frac{\sigma_N^2}{\varepsilon^2}. \tag{3.33}$$

In the case of my estimators' variance $\sigma_N^2 = \frac{A}{N^k} + o(N^{-k})$, where $k = 1,3,5$ and A is a constant related to the properties of the input signal and the impulse response of the implemented filter (i.e. related to the integrand function and its derivatives). Let us calculate the infinite sum of the inequality (3.33),

$$\sum_{N=1}^{\infty} \Pr(|\hat{y}_N - y| \geq \varepsilon) \leq \sum_{N=1}^{\infty} \frac{\sigma_N^2}{\varepsilon^2} = \frac{A}{\varepsilon^2} \sum_{N=1}^{\infty} \frac{1}{N^k} + B, \tag{3.33}$$

where $B = \sum_{N=1}^{\infty} o(N^{-k})$.

We note that when $k = 1$ (i.e. ToRa estimator), $\sum_{N=1}^{\infty} \frac{1}{N^k} = \sum_{N=1}^{\infty} \frac{1}{N} = \infty$. Hence the assumption of the Borel-Cantelli lemma is not satisfied and a.s. convergence cannot be proven. However, if $k = 3$ or 5, which are the cases for StSa and AnSt, respectively, then $\sum_{N=1}^{\infty} \frac{1}{N^k} < \infty$, and B is finite, as well. Therefore, the assumption of the Borel-Cantelli lemma is satisfied, which implies $\Pr\left(\lim_{N\to\infty} \sup(|\hat{y}_N - y| > \varepsilon)\right) = 0$, and the a.s. convergence occurs.

3.9. When to Use Which Estimator

We have seen that some estimators converge faster than the others. The question now is how to pick a specific estimator for a given application? The easy answer is AnSt-based estimator, since it has the fastest decaying rate of the three considered estimators. While this is true in general, there are a few conditions that have to be met for this estimator to have such a converging rate:

1. The zero-order derivative (ZOD) of the integrand function, i.e. $f(\tau)$, must be continuous, bounded and integrable.

2. The first-order derivative (FOD) of the integrand function, i.e. $f'(\tau)$, must exist and be continuous and bounded.

3. The second-order derivative (SOD) of the integrand function, i.e. $f''(\tau)$, should also exist and be continuous and bounded.

4. The integrand function can be deliberately sampled as per the AnSt scheme dictates, that is, sample points should be acquired in antithetical pairs for all strata. However, for continuous-time integrand functions (input signals and impulse responses) this condition can usually be satisfied.

5. The stochastic point process used to generate the random time instants of the sample points should have a PDF that guarantees the estimator to be unbiased and consistent [12, p. 38]. For instance, the uniform distribution with constant PDF fulfils this condition, as I have proven above.

6. The size of the sample sequence should be above a certain number for the asymptotic behaviour of the estimator to establish.

Some of these conditions are also required for other estimators. For example, both ToRa and StSa estimators require conditions 1,2, 5, and 6 to be fulfilled, as well, in top of their own other prerequisites. Indeed, conditions 1 and 2 to expand using Taylor series, condition 5 to guarantee estimators unbiasedness and consistency, and condition 6 to guarantee the declared convergence rates. Moreover, for StSa case, the sampling circumstances allow for stratification to be carried out, with deliberate random selection of only one sample point per stratum is affordable. In ToRa case, only one condition must also apply. That is, the total random sampling procedure should be deliberately viable.

Following this discussion, we can see that ToRa is the most relaxing sampling scheme, next is the StSa scheme, with AnSt technique having the toughest restrictions, i.e. absolutely the opposite ranking with regards to the speed of convergence.

A remaining issue still need to be addressed is when all or part of the aforementioned conditions are not satisfied. Apparently, in this case, there is no guarantee on the consistency, unbiasedness, variance values, and rates of convergence of the considered estimators. One needs to seek other estimators most relevant to the specific

situation encountered. For example, if there are no control over the time instants at which sample points are collected, i.e. unintentional nonuniform sampling, in such case, the preceding estimators are either not valid or some of their statistical characteristics will be different.

An exception to the first three conditions could be tackled with while applying the three considered estimators is still valid, of course, with an impact on some of their statistical properties. This is exactly what I am going to address in the second part of this thesis, where I propose the use of the same estimators but with discontinuities exist on either the SOD, FOD, or ZOD of the integrand function.

3.10. Computational Cost

One of the main goals of this thesis is to prove that adopting randomisation in digital sampling through DASP would provide, under certain conditions, cost-effective filtering techniques. Applications that are suitable for such DASP random techniques should be error-tolerating, though. Otherwise, traditional DSP techniques seem to be inevitable even if they would cost more. DASP, as emphasized elsewhere in this thesis, should be understood as a supplement technology to DSP and not a complete alternative. However, if the spectrum of the processed signal is wide and unknown, and there are limitations on using extra analog components/blocks (e.g. antialiasing analog filters) for any reason, then the sparser and closer to the dc-component the spectrum is the more cost-saving would be achieved with random sampling and DASP techniques. By cost-saving I mean reduced computational complexity and power consumption, which in turn lead to less processing time and money.

Since I am dealing with digital filtering, I would like to compare my filter estimators' computational complexity with classical DSP filters'. To this end, I introduce below, (3.34a-b), the conventional FIR filter model commonly used in DSP applications. For the comparison to be fair, I assume that the sampling rate of the DSP filter is just fulfilling the required Nyquist rate (i.e. not oversampled) and the filtering is conducted for the same parameters and circumstances: same input signal, $x(t)$; a uniformly sampled version of the same analog filter impulse response, $h(t)$; observation

window, $[t-T, t]$; and using an equivalent notation for the number of samples, $2N_c$, where the subscript c denotes classical DSP. Thus,

$$y_c(t) = T_s \sum_{n=0}^{2N_c-1} x(t-nT_s)h(nT_s)\,, \tag{3.34a}$$

$$y_c(t) = T_s \sum_{k=N_0-2N_c}^{N_0-1} x(kT_s)h(t-kT_s)\,, \tag{3.34b}$$

$$y_c(t) = T_s \sum_{k=N_0-2N_c}^{N_0-1} f(t, kT_s), \tag{3.34c}$$

where $T_s = \frac{T}{2N_c} = \frac{1}{F_s}$ is the uniform sampling period and $N_0 = \left\lfloor \frac{t}{T_s} \right\rfloor$. Note that the sampling frequency here, F_s, should at least be equal to the Nyquist rate,

A quick glance at (3.34c) suggests, with no surprise, that it is analogous to the formulas of my randomised filter estimators, (3.7a), (3.8a), and (3.9a). Hence, the computational complexity, in terms of multiplications and additions per input sample, for digital filtering in both cases (uniform and random) depends only on the utilised number of sample points. Moreover, the power dissipation of the ADCs in such cases could be significantly decreased by acquiring a smaller number of samples, as well. Therefore, for randomised filtering to be more cost-effective than the uniform approach, it is only required to show that it is capable of yielding a *good* estimate of the filter output while using a randomly sampled sequence of size $2N$ where N satisfies the condition $N < N_c$.

Indeed, I have already proved the two requirements in the analysis above. According to my proposed estimators and their statistical features results, it is explicitly found that all estimators are consistent and unbiassed in (3.11g) and (3.27a-c). The other requirement, $N < N_c$, can be implicitly deduced from the analytical derivation of the randomised filter estimators' formulas where unbiasedness and consistency do not cease to be valid even with lesser values of N. Together with the fact that randomisation in sampling and/or quantisation can mitigate the adverse effects of

aliasing, as demonstrated in the previous two chapters, it is now evident that the proposed random based estimators can be more cost-effective than the classical DSP filtering if the conditions for proper randomisation and/or quantisation are satisfied. In the next section, I further validate my results by demonstrating numerical examples and simulations.

First, I would like to summarise the required sampling rates for conventional uniform sampling of bandlimited baseband, bandpass and multiband signals when the detailed spectral support function (SSF) of the signal of interest is not fully known, as listed in Table III.

TABLE III: MINIMUM REQUIRED RATES FOR UNIFORMLY SAMPLING SIGNALS WITH UNKNOWN SPECTRAL SUPPORT FUNCTIONS

	Min Freq.	**Max Freq.**	**No. of Sub-bands**	**Bandwidth Or Sub-bandwidths**	**Nyquist Rate, F_s (Uniform)**
Baseband	0	f_U	1	f_U	$2f_U$
Bandpass	f_L	f_U	1	$f_U - f_L$	$2f_U/n$ See note[1]
Multiband	f_L	f_U	M	$B_1, \dots, B_M$	$F_s = 4\sum_{i=1}^{M} B_i$ See note[2]

I am aware that, in most cases, there should be a minimum knowledge about the signal to be sampled. In case of real-valued bandlimited baseband signals, the lower frequency, f_L, is known to be zero (i.e. the dc-component), but the upper frequency, f_U, has to be known in advance, or at least anticipated somewhere in the spectrum according to the application being considered, otherwise, I end up sampling the whole range of the spectrum, which is unrealistic. For real-valued bandlimited passband and

[1] n is the maximum integer number of the quotient $f_U/(f_U - f_L)$, and $\frac{2f_U}{n} \le F_s \le \frac{2f_L}{n-1}$.

[2] This value $4\sum_{i=1}^{M} B_i$ is double the Landau minimum rate. For known spectrum support, the Nyquist rate can approach the Landau minimum rate, i.e. $2\sum_{i=1}^{M} B_i$ if proper uniform sampling schemes are put in place. If all bandwidths of the multiband signal are equal to B_s, for example, then ($F_s = 2MB_s$).

multiband signals, both f_L and f_U, or centre frequency and bandwidth, are required before conducting the sampling process. Note that if the lower frequency, f_L, is unknown, both bandpass and multiband signals should be treated as if they are baseband signals in the sense that f_L may be arbitrarily changing and, at some point, it may approach the zero dc-component. Moreover, precise number of sub-bandwidths (if they are equal) or their individual lengths (if they are not the same) of a multiband signal have to be known, otherwise, it is treated as a bandpass signal with known f_L and f_U only.

Remark that for the case of random sampling, there will always be statistical errors due to sampling irregularities apart from the utilised average sampling rate, even if the SSF is fully known and the average sampling rate[3] is exceeding the equivalent Nyquist rate for uniform sampling. Said that, this doesn't mean that random sampling is useless, or it is not worth it. On the contrary, it may help us a lot when conventional DSP solutions fall short for technical or financial grounds, especially when the SSF is sparse and unknown. However, to provide a comparative indicator with uniform sampling, as illustrated in Table III, the rule of thumb for random sampling is that for signals with known SSF, the average sampling frequency is the minimum Landau rate. Whereas if the SSF is unknown, then the sampling rate doubles. This is just a guidance for *good* random sampling practice. Designers can decide what random sampling rate is suitable for a specific application and how much error could be tolerated.

3.11. Implementation Algorithm

Fig. 12 below shows an implementation example of ToRa estimator. In this example, the observation window interval, T, is equal to the length of the filter impulse response. The input signal is the top curve, $x(\tau)$, and the sinc function represents $h(\tau)$. The random sample points of the input signals are acquired at time instants $\tau_0, \tau_1, \tau_2, \ldots$. The filter is assumed to be symmetric FIR filter with linear phase and has a group delay of $T/2$. The output signal sample spacing, according to this particular example, is not equidistant, rather, they are random with a constant offset of $T/2$. This

[3] Average random sampling rate can be computed by dividing the total number of samples by the observation window interval.

timing arrangement of calculating the out signal is not an obligatory and it can be any time sequence including equidistant.

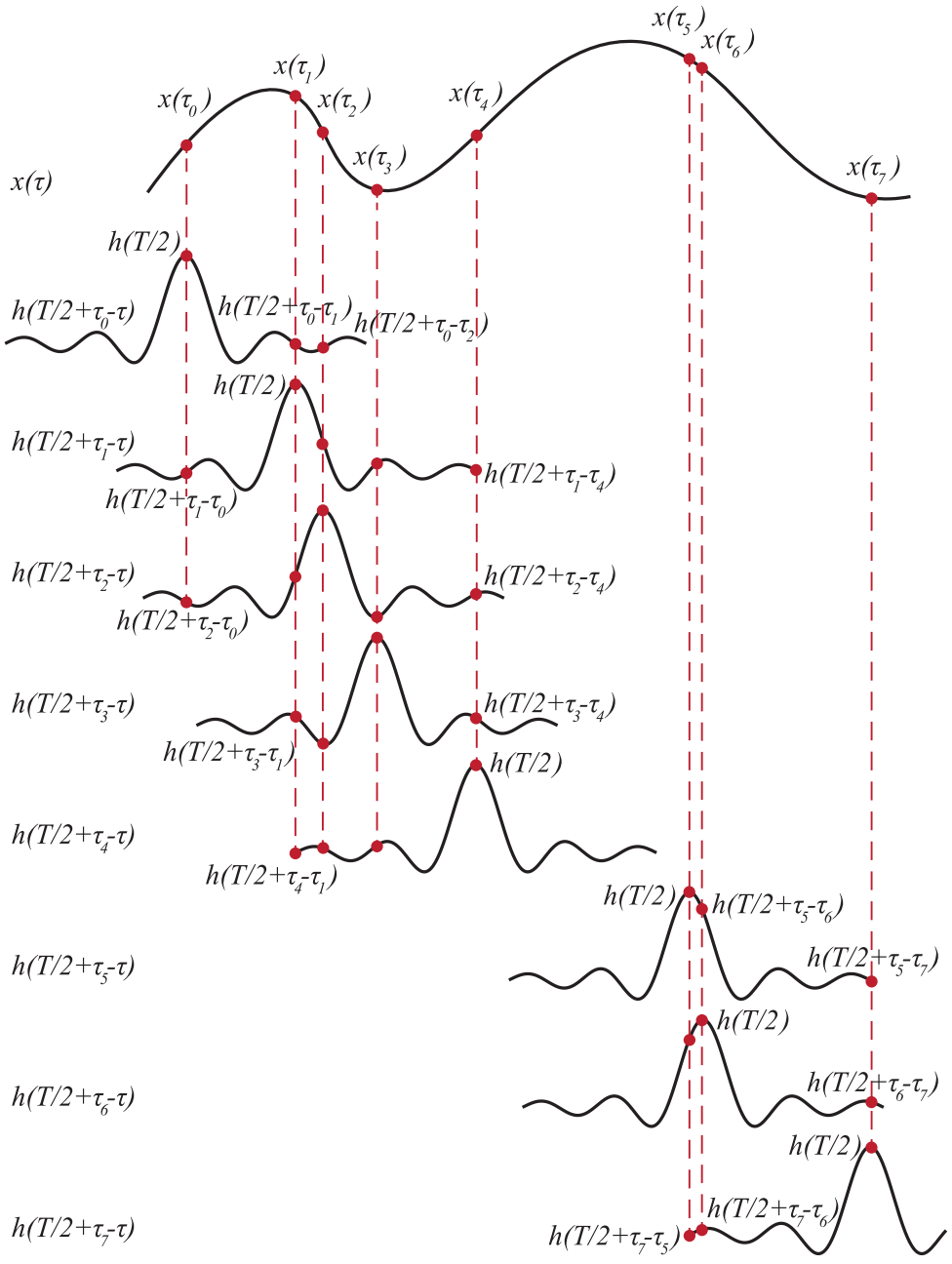


Fig. 12. Implementation algorithm of a ToRa filter estimator, where random samples of shifted replicas of the filter impulse response are multiplied with the corresponding random sample points of the input signal. Simultaneous sampling of $x(\tau)$ and $h(\tau)$ is the key point in this algorithm, and then convolution is carried out nonuniformly.

Each row of the sinc function represents one output sample. For instance,

$$y\left(\frac{T}{2}+\tau_0\right)=x(\tau_0)h\left(\frac{T}{2}\right)+x(\tau_1)h\left(\frac{T}{2}+\tau_0-\tau_1\right)+x(\tau_2)h\left(\frac{T}{2}+\tau_0-\tau_2\right), \quad (3.35a)$$

$$y\left(\frac{T}{2}+\tau_1\right)=x(\tau_0)h\left(\frac{T}{2}+\tau_1-\tau_0\right)+x(\tau_1)h\left(\frac{T}{2}\right)+x(\tau_2)h\left(\frac{T}{2}+\tau_1-\tau_2\right), \quad (3.35b)$$

and so on. For this to work properly, each output sample requires new synchronised impulse response sampling with the input signal's samples within the sliding observation window.

3.12. Methodology for StSa and AnSt Estimators

For StSa, the implementation methodology can be summarised as follows:

1. A sliding observation window of length T seconds is determined according to a given application. The Window is divided into $2N$ equidistant strata, where $2N/T$ will be the intended average random sampling rate.
2. A software-defined analog FIR filter or an oversampled, grid-based, FIR filter is designed, and pre-saved as a continuous-time function (CTF) or a lookup table (LUT) in the memory buffer of the application.
3. A $2N$-size sequence of sample points is acquired nonuniformly (randomly/pseudorandomly) from the input signal within the observation window, based on a random point process with specific probability density/mass function (PDF/PMF). In this thesis, both belongs to uniform distribution.
4. Spontaneous sequence of impulse response samples (filter coefficients) are "extracted" from the CTF/LUT to exactly match the time instants of the random input samples within the current sliding window.
5. Convolution of the two discrete sequences is carried out to estimate one sample of the filtered output signal.
6. The observation window is shifted one stratum to the right discarding the left-most sample point and causing all other $2N-1$ points to be left-shifted one stratum. A

new random sample point is then acquired from the input signal in the now-empty right-most $2N$-th stratum.

7. Steps 4-6 are repeated as many times as needed, probably till the end of all input signal sample points or for a fixed size of the estimated output signal, depending on the application.

For AnSt, the same methodology is implemented, but instead of acquiring one sample point per stratum, two points are taken. In return, the number of strata would be half of StSa's, i.e. only N strata will be used. Moreover, in step 6. above, the left-most two sample points are discarded, and a new antithetical sample pair is acquired to the right of the sample sequence.

3.13. Numerical Examples

3.13.1. Function Integration and Asymptotic Behaviour of Estimators

To verify my findings numerically, I start my collection of simulation examples with a very simple setup. I want to validate the asymptotic behaviour of the ToRa, StSa, and AnSt estimators for continuous-time signals. Hence, a two-sinusoid smooth function, $y(t) = 10\sin(2\pi \times 17t) + 7\cos(2\pi \times 29t)$, is sampled randomly with the three considered techniques. All sample sequences have the same size, N, at any given iteration. The observation window is set to $[0,1.2]$ seconds. Since all estimators are random and cannot be judged by a single realization, I carry out Monte Carlo simulations of 100 independent iterations for each estimator.

Fig. 13 shows the estimators asymptotic behaviours, where it is seen that AnSt-based estimator is uniformly converging at a rate of N^{-5} after about the 118$^{\text{th}}$ sample point. Whereas, StSa-based estimator is converging uniformly at a slower rate of N^{-3}, and it needs around 44 sample points to establish its asymptotic characteristics. The slowest decaying estimator is the ToRa-based, where it can only converge at a uniform rate of N^{-1}, albeit this behaviour starts as early as 2 sample points. Indeed, ToRa-based estimator reaches its asymptotic behaviour at any value of N since the result in (3.28b) is fixed, for the same observation window and smooth function $f(t,\tau)$, and

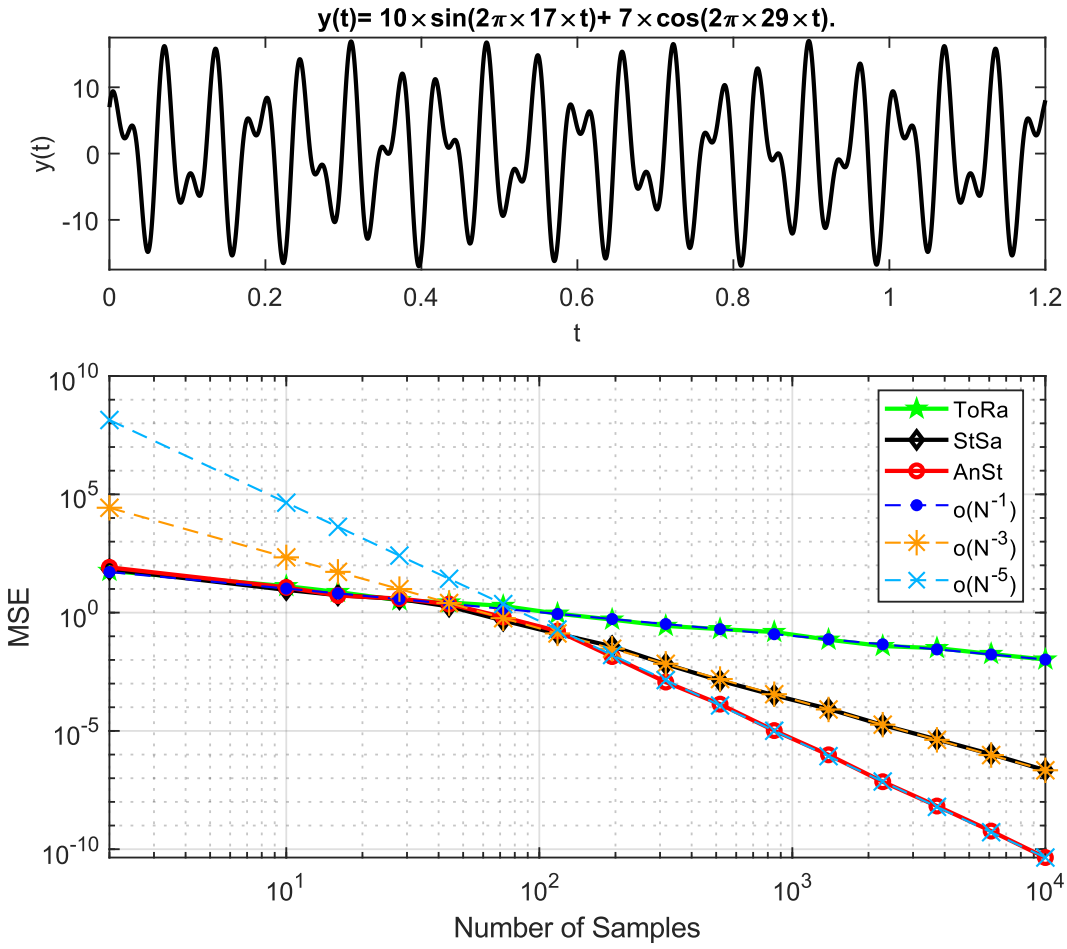


Fig. 13. A smooth function $y(t)$ with continuous, bounded, and square-integrable FOD and SOD (Top). Uniform convergence rates of the ToRa, StSa, and AnSt based estimators (Bottom), where it is evident that, after a specific number of sample points for each estimator, they begin to decay linearly. Namely, ToRa, StSa, and AnSt estimators are uniformly converging at N^{-1}, N^{-3}, and N^{-5} rates, respectively.

does not depend on the number of sample points. Therefore, it is always linear in the logarithmic scale.

Whereas both StSa and AnSt numerical integration techniques are function-dependent regarding the start point of their asymptotic linear convergence in the logarithmic scale, as intuitively remarked from (3.29a-b) and (3.30a-b).

In general, AnSt depends on the concept of adding and subtracting the same amount of error when calculating AUC of the function under consideration. This concept works perfectly when the number of samples is large enough to guarantee the part of the function within each stratum is smooth and linear (or monotonically increasing or

decreasing). This way the error decreases dramatically, since all mistakenly added areas are nearly equal to the incorrectly subtracted ones, leading to overall error to be minimal.

For StSa, there are no such antithetical pairing concept for sample points, therefore, the estimator starts to converge uniformly, according to (3.29b), as long as $f(t,\tau)$ is a smooth function with Riemann integral value of its FOD-squared is almost equal to the summation form of (3.29a).

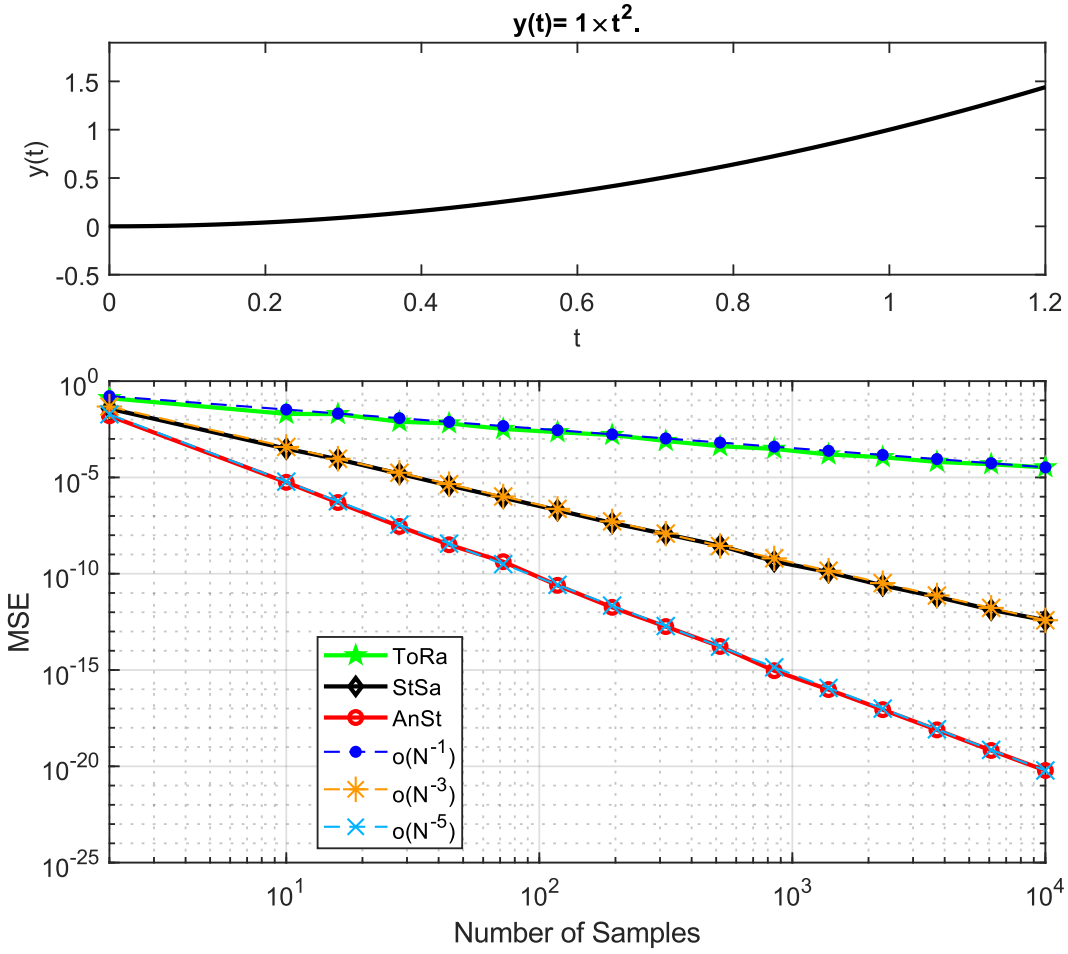


Fig. 14. A smooth function $y(t)$ with continuous, bounded, and square-integrable FOD and SOD (Top). Uniform convergence rates of the ToRa, StSa, and AnSt based estimators (Bottom), where it is evident that, after a specific number of sample points for each estimator, they begin to decay linearly. Namely, ToRa, StSa, and AnSt estimators are uniformly converging at N^{-1}, N^{-3}, and N^{-5} rates, respectively.

The number of sample points required for both StSa and AnSt estimators to establish their logarithmic decaying linearity can be decreased considerably if the whole part of

the function of interest within the observation window is smooth and monotonically increasing or decreasing. Consider, for example, another function, $y(t) = t^2$, observed within the same interval as the previous example, i.e. [0,1.2] seconds. As can be seen in Fig. 14, all estimators are converging at the same rates as expected, but also all of them start to converge as early as 2 sample points, with AnSt estimator having the smallest MSE values. This is because the SOD of $y(t)$ is constant in this example, and so is (3.30b). Similarly, for the StSa estimator, 2 sample points are quite enough to represent a linear FOD of $y(t)$, hence, the term in (3.29b) does not change with increasing number of samples.

3.13.2. Filtering Estimation Examples

In another example, I demonstrate how to approximate a filter output using the three considered estimators and a continuous-time analog input signal. To achieve this purpose, I assume that there is a wireless sensor network (WSN) that comprises several multiplexed sensors and transducers sharing a common communication channel.

The WSN modulates and transmits the shared analog signal, $x(t)$, to a remote location using a specific RF link. The original analog signal is assumed to be a bandlimited signal having a maximum frequency of $f_U = 65.536kHz$ and comprising audio (roughly from $0Hz$ to $20kHz$), ultrasonic (from $22kHz$ up to $44kHz$) and other types of signals with higher frequencies up to $f_U = 65.536kHz$. Let us also assume that the exact SSF of $x(t)$ is sparse and not fully known. The upper and lower frequencies of each frequency band are known, where each band can only have a few sinusoids and/or narrowband *sinc* functions (just for simplicity of simulation) at any given time.

The receiver circuit demodulates the transmitted RF signal by means of suitable analog frontend blocks. The received multiplexed analog signal, $x(t)$, is observed within a time interval $[0, T]$ where $T = 0.0156$ sec, and is given by

$$\begin{aligned} x(t) &= sin(2\pi F_1 t) + 3 \times sin(2\pi F_2 t) + 2 \times sin(2\pi F_3 t) \\ &\quad + 1.5 \times T \times F_4 \times sinc(2F_4 t) \times cos(2\pi F_5 t) - 1.5 \times sin(2\pi F_6 t) \\ &\quad + 2.5 \times sin(2\pi F_7 t), \end{aligned} \tag{3.36}$$

Where $sinc(t) = sin(\pi t)/\pi t$, $F_1 = 2.048kHz$, $F_2 = 8.192kHz$, $F_3 = 25.4kHz$, $F_4 = 1.5kHz$, $F_5 = 51.5kHz$, $F_6 = 42\ kHz$ and $F_7 = 61.9kHz$.

A specific sub-circuit of the receiver is only interested in the frequencies within the audio frequency range (i.e. $0 - 20kHz$), but, for some reason, it does not use an antialiasing analog pre-filtering (potentially because it is dynamically changing as in an SDR system, for instance). Therefore, the input analog signal to this sub-circuit is the same as $x(t)$, which currently has an upper frequency of $F_7 = 61.9kHz$, but would extend to $65.536kHz$ in other observation windows. Hence, according to the classical DSP approach for the existing signal and sub-circuit setup, we need a sampling rate that is at least equal to twice the maximum frequency that might present in the processed analog signal, i.e. the Nyquist rate, which in this case is $F_s = 2 \times 65.536kHz = 131.072kHz$.

The example setup above fits for the random sampling and filtering approaches, discussed earlier in this chapter, since $x(t)$, as shown in Fig. 15 (solid black line), has a sparse and unknown spectrum support, albeit the lower and upper frequencies of each wave range are assumed to be known. So, to mitigate aliasing from other bands while reducing the cost of implementation, I utilise the three considered random filter estimators ToRa, StSa, and AnSt to approximate the filter output signal, $y(t)$.

For comparison purposes, I also consider uniform sampling and filtering scheme, which is the case of classical DSP. Remark that aliasing occurs if the utilised uniform sampling frequency is less than the required Nyquist rate, as depicted in Fig. 19 where only half the Nyquist rate is used, i.e. $Fs = 65.536kHz$.

To receive the acoustic components of $x(t)$ in the audio frequency range, an analog function representing the required FIR filter is designed. Namely, the filter is a boxcar LPF with a bandwidth of $B_{LBF} = 20kHz$, as illustrated in Fig. 15 (dashed blue line). In effect, the impulse response of the filter, $h(t)$, is a shifted version of the *sinc* function multiplied with a Hanning smoothing window, i.e.

$$h(t) = sinc\left(B\left(t - \frac{T}{2}\right)\right) \times 0.5\left(1 - cos\left(\frac{2\pi}{T}t\right)\right), \tag{3.37}$$

where B is the double-sided bandwidth of the LPF filter ($= 2 \times 20kHz$).

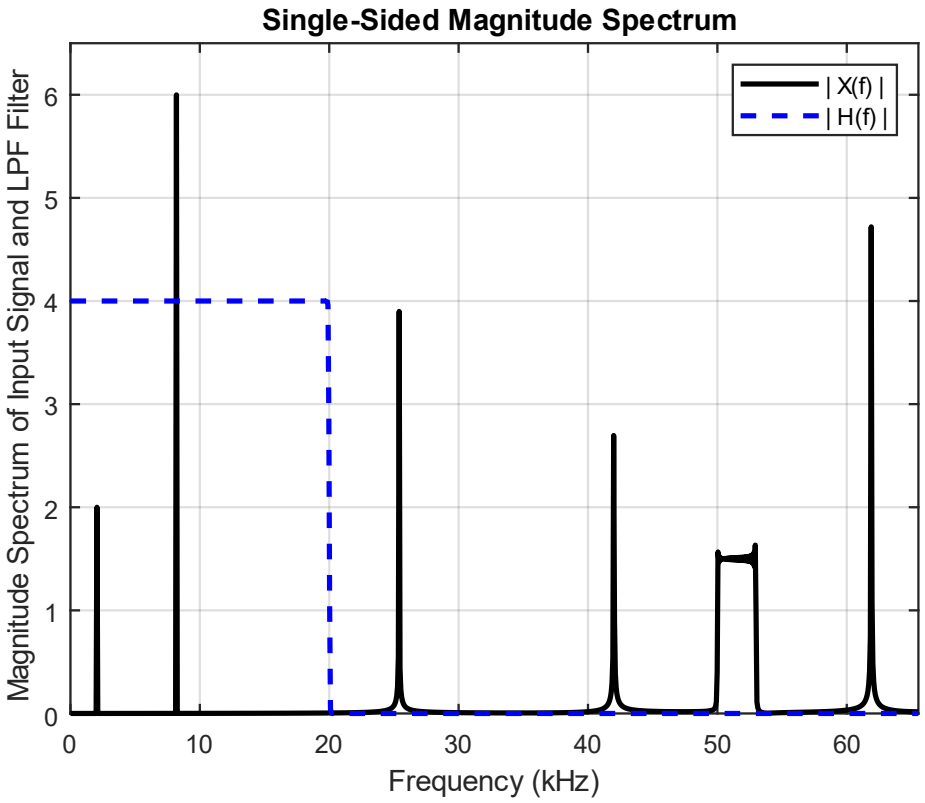


Fig. 15. Single-sided magnitude spectrum of the input analog continuous-time signal (solid black) uniformly sampled at Nyquist rate ($= 131.072kHz$), and the frequency response of the boxcar LPF filter (dashed blue).

$h(t)$ is an analog function of time that can be thought of as a software-defined analog filter. However, for a given application that requires specific filter design which cannot mathematically be represented by an analog function, oversampled on-grid coefficients of the requested filter impulse response can be saved in a lookup table (i.e. inside the receiver buffer), as will be discussed in the next chapter.

The analog input signal, $x(t)$, and the filter impulse response, $h(t)$, are simultaneously sampled using both uniform and random sampling techniques. Then, DSP filtering and DASP filtering are carried out, simulated and compared.

For the traditional DSP case, uniform sampling frequency matching the Nyquist rate is used, and the filtered audio signal is depicted in top left subplots of Figs. 16-18, where it is considered as a reference ("Ref") spectrum to be compared with spectra of random estimators or other different uniform sampling rates. While in the DASP case, I consider equivalent average random sampling frequencies to the utilised uniform

ones to recover the audio signal using the three estimators (ToRa, StSa, and AnSt), as shown in top right, bottom left, and bottom right subplots, respectively, of Figs.16-18. Note that for the sake of simplicity when comparing input/output signals' spectra, all output spectra are scaled down by a factor of $1/4$, since the absolute magnitude of frequency response of the employed LPF in the passband region is 4.

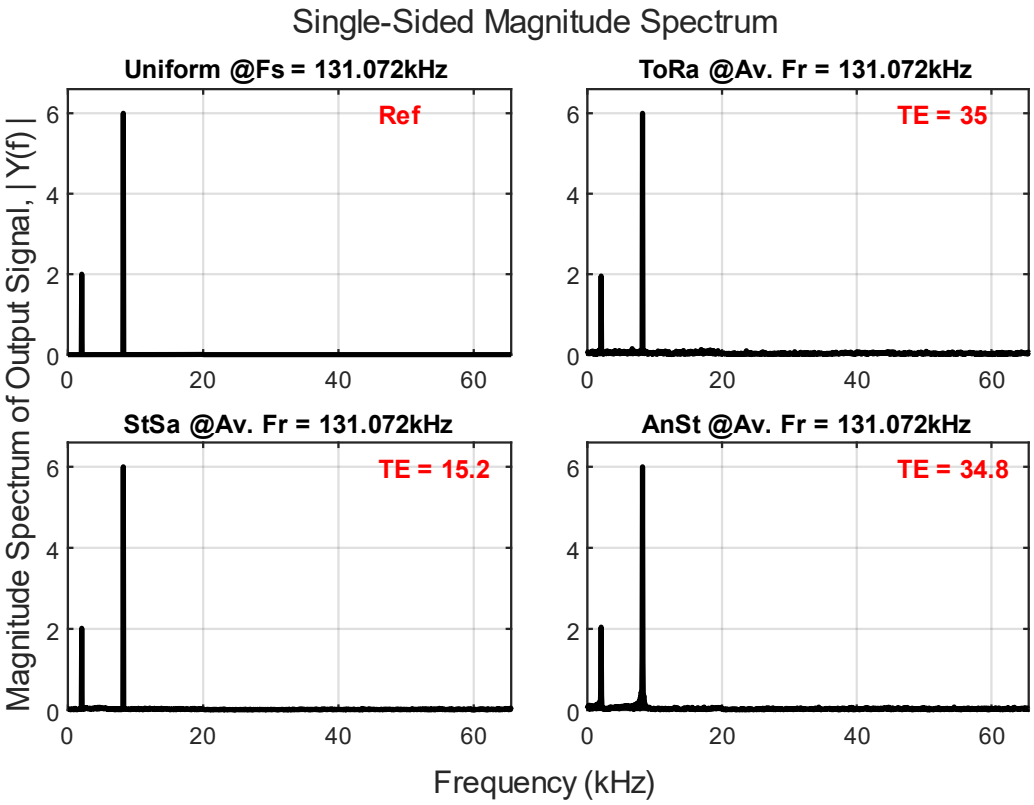


Fig. 16. Single-sided magnitude spectrum of the LPF filter output signal. Sampling frequency is $131.072kHz$ and the number of independent Monte Carlo iterations for the random estimators is 100.

Spectra comparisons are carried out based on the total error (TE) available within the Nyquist frequency range, i.e. $(0, F_{Nyq}/2)$, and is calculated by finding the sum of all absolute differences between the estimated spectra and the "Ref" spectrum.

As can be seen from Fig. 16, spectra of the estimated output signals using random sampling techniques are barely distinguishable from that of the uniformly sampled one using 100 MC simulations. Also, there are less errors in StSa estimator case than the errors of both ToRa and AnSt estimators, with AnSt is slightly better than ToRa.

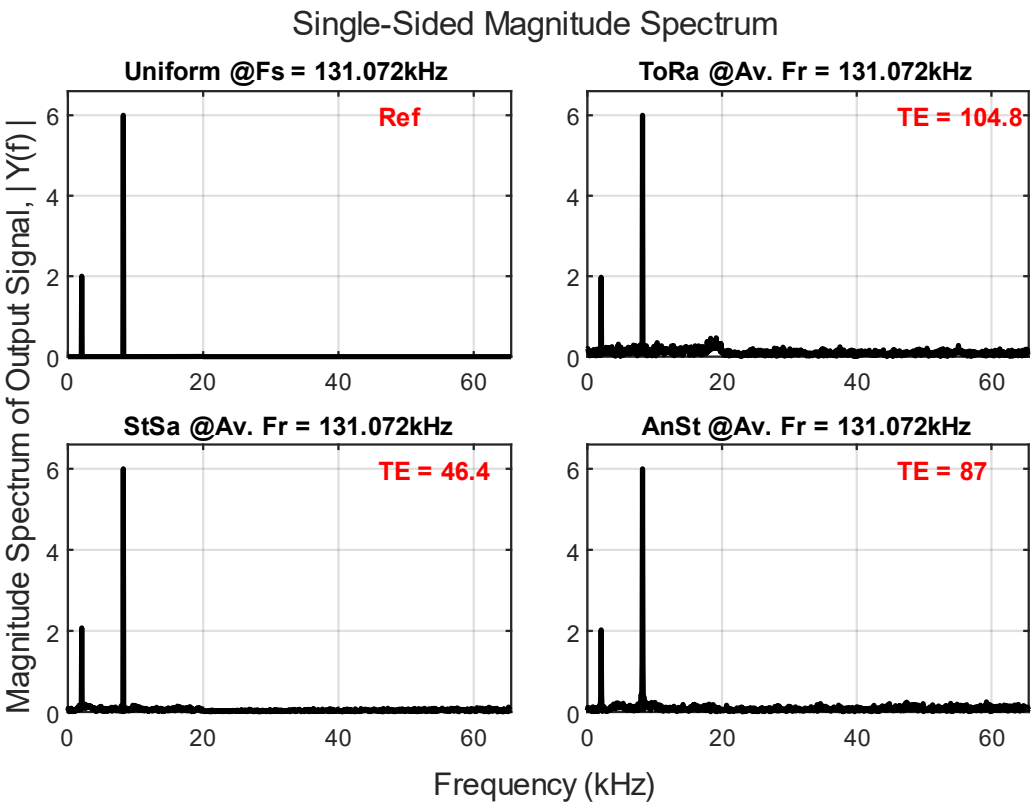


Fig. 17. Single-sided magnitude spectrum of the LPF filter output signal. Sampling frequency is $131.072kHz$ and the number of independent Monte Carlo iterations for the random estimators is 10.

To demonstrate the effect of MC averaging of random estimators, I include here two extra plots for 10 independent MC runs and only one MC run, as illustrated in Fig. 17 and Fig. 18. It is clear that the less MC runs are, the more statistical errors appear. This is absolutely feasible for the estimators when looked at as functions of random variables. That is why I have said beforehand that these estimators cannot be judged by only a single realisation.

Remark that despite the exploited average random sampling frequency in Figs. 16-18 is matching the Nyquist rate, the statistical errors still appear, contrary to uniform sampling case. Actually, this is one of the *drawbacks* of randomised sampling in general, since these errors are due to sampling irregularities and cannot be got rid of completely, but can, at an additional cost, be reduced.

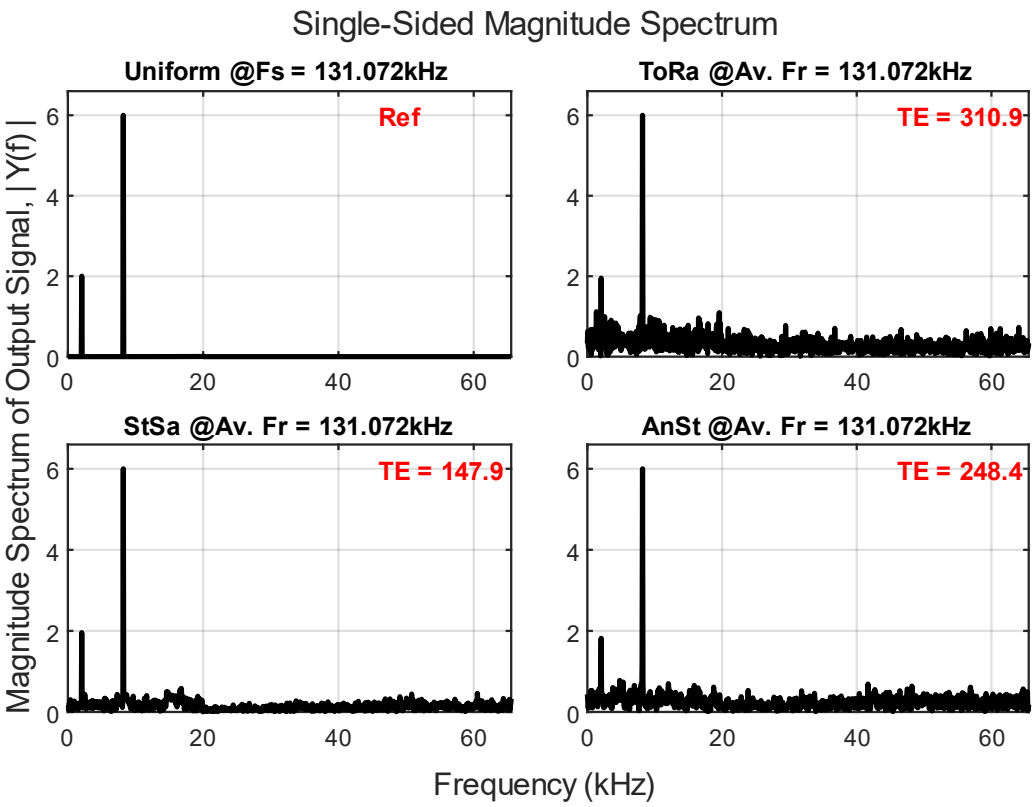


Fig. 18. Single-sided magnitude spectrum of the LPF filter output signal. Sampling frequency is $131.072kHz$ and only one Monte Carlo iteration for the random estimators.

Note that in these simulation results, as well as all other simulations to follow, I have used three layers of randomisation for each one of ToRa-, StSa-, and AnSt-based filter estimators. First, I applied randomisation in the sampling of input signal, $x(t)$, and filter impulse response, $h(t)$. Then, the second layer of randomisation is conducted when estimating the filter output signal $y(t)$. Finally, randomisation is also exploited when estimating the one-sided magnitude spectrum of the output signal, $|Y(f)|$.

3.13.3. Sub-Nyquist Sampling

The following example uses the same input analog signal and LPF filter configurations except for a reduced sampling rate. With only half the required Nyquist rate for uniform sampling, i.e. $65.536kHz$, and equivalent average random sampling frequency, I obtain the results depicted in Figs. 19-21.

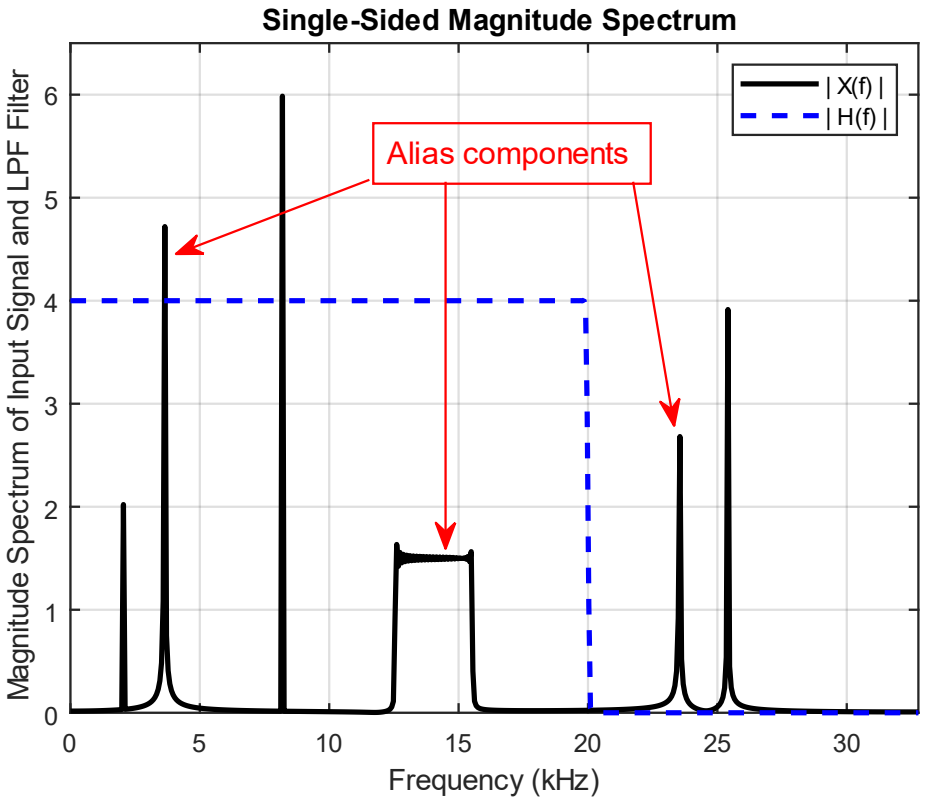


Fig. 19. Single-sided magnitude spectrum of the input analog continuous-time signal (solid black) uniformly sampled at half the Nyquist rate ($F_s = 65.536kHz$), and the frequency response of the boxcar LPF filter (dashed blue). Remark how frequency components of the spectrum of input signal, $|X(f)|$, that are above $32.768kHz$ (= half the new sampling rate, F_s) have been aliased into the range of detectable spectrum, i.e. $\left[-\frac{F_s}{2}, \frac{F_s}{2}\right]$, when using uniform sampling schemes.

In Fig. 19, we can obviously see the aliases of some frequency components of the input signal $x(t)$. Originally, they were above the $32.768kHz$ limit, which is half the new sampling rate. Classical DSP sampling and filtering yield such aliases when there are frequency components higher than the utilised uniform sampling rate. Unfortunately, some of these aliases do already exist within the LPF passband, and they will wrongly appear in the output signal for conventional DSP filters. Whereas the case is different when considering randomised sampling-based techniques and estimators, as can be seen in Fig. 20 and Fig. 21 for MC=100 and MC=1 respectively.

In Fig. 20 and Fig. 21, the spectra of filter output for uniform sampling are clearly showing the aliases of higher frequency components. While the random estimators

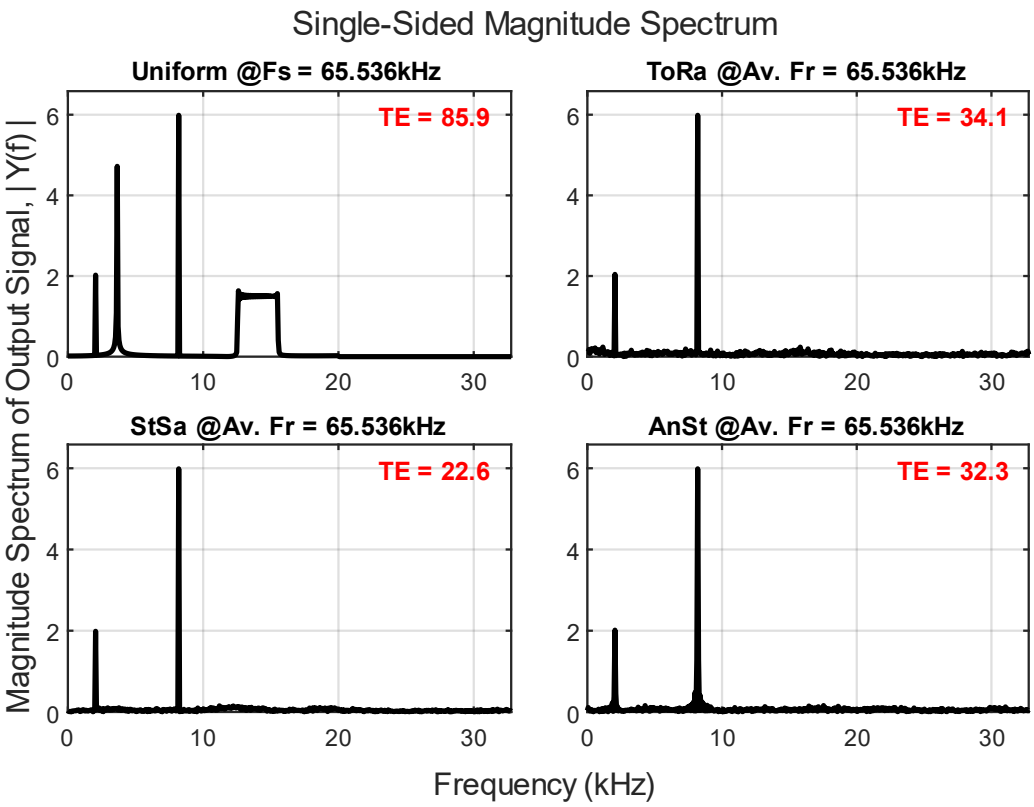


Fig. 20. Spectrum of the output signal. $F_s = 65.536kHz$ and MC=100.

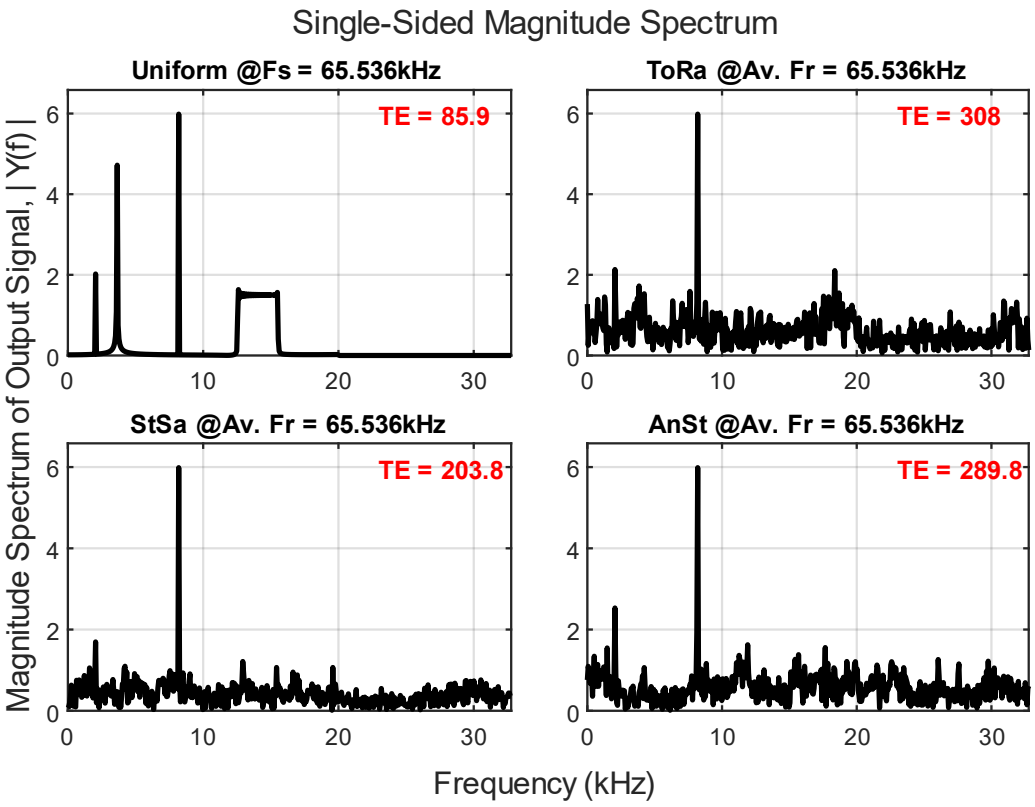


Fig. 21. Spectrum of the output signal. $F_s = 65.536kHz$ and MC=1.

can still reveal the true spectra despite that the amount of statistical errors increase more than the above counterpart figures, i.e. Fig. 16 and Fig. 18.

Furthermore, I examine the same example above, but for lower sampling rates to verify the behaviour of uniform and random sampling and filtering schemes and compare them with the above simulation results. Hence, for the following figures, I set the sampling rates as follows: Fig. 22, $F_s = 32.768kHz$ and Fig. 23, $F_s = 16.384kHz$. Note that the latter is the minimum Landau rate if the SSF of $|Y(f)|$ in this particular example is known. Going under this rate will certainly lead to an unpredictable output results with considerable amount of statistical errors that dominate the spectrum especially for small number of MC runs. For example, if MC=1 and $F_s = 8.192kHz$, then Fig. 24 shows how huge the errors are, and the original audio signal is no longer distinguishable.

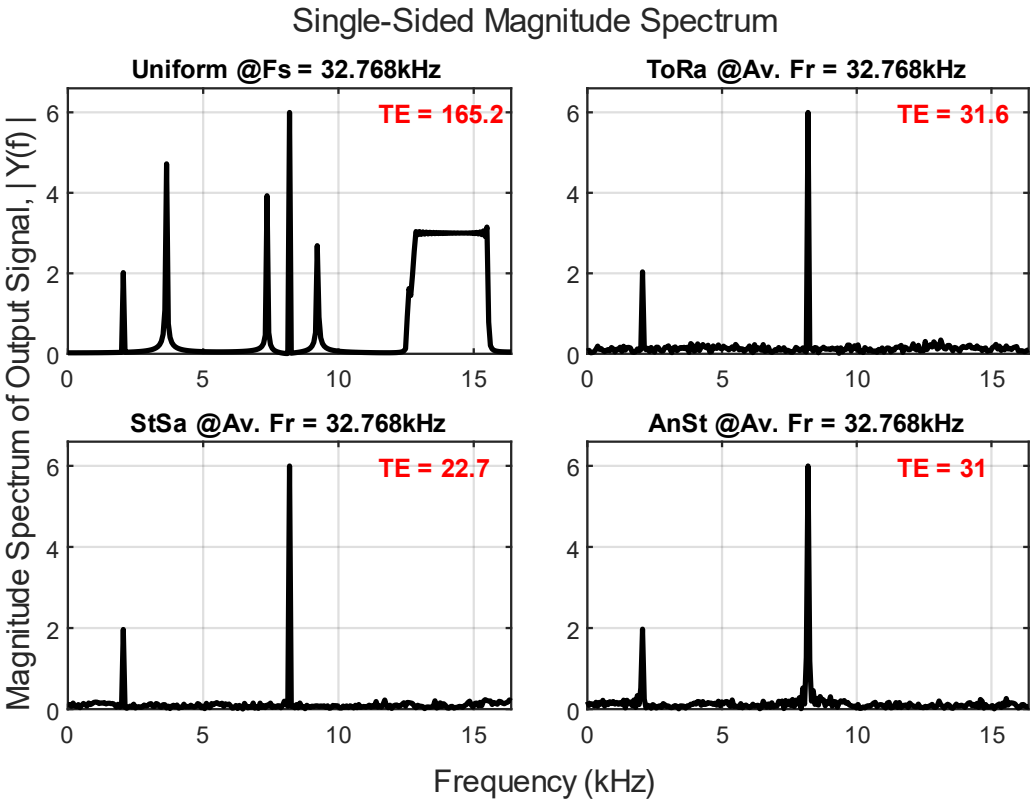


Fig. 22. Spectrum of the output signal. $F_s = 32.768kHz$ and MC=100.

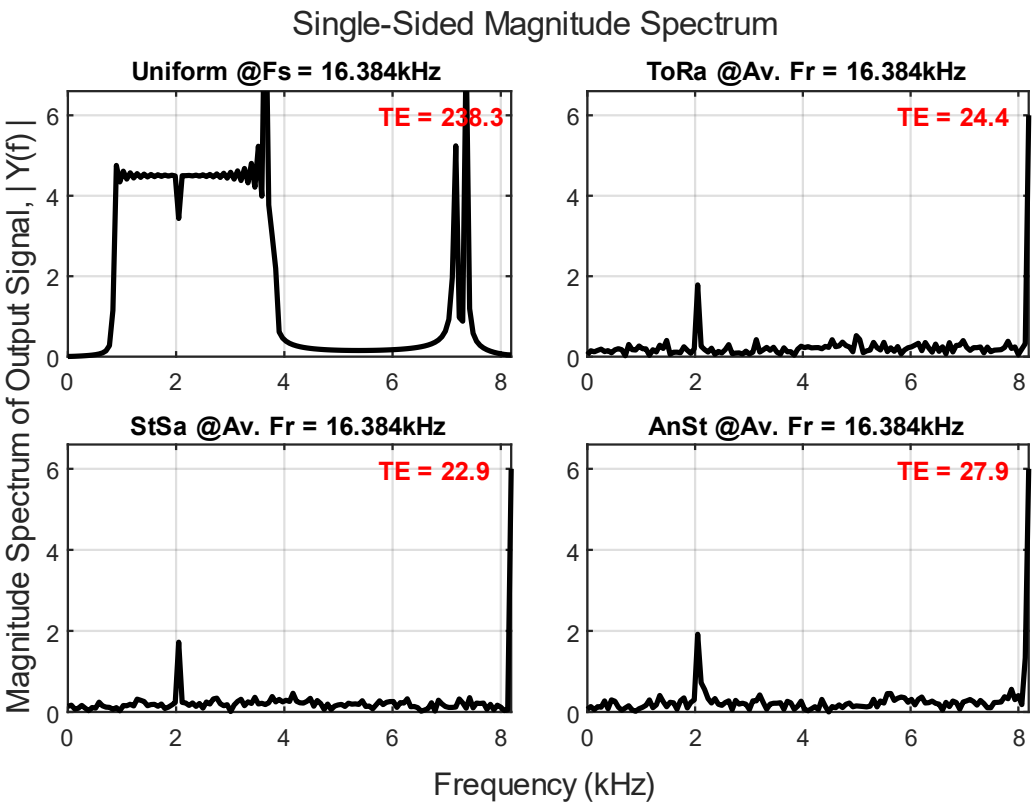


Fig. 23. Spectrum of the output signal. $F_s = 16.384kHz$ and MC=100.

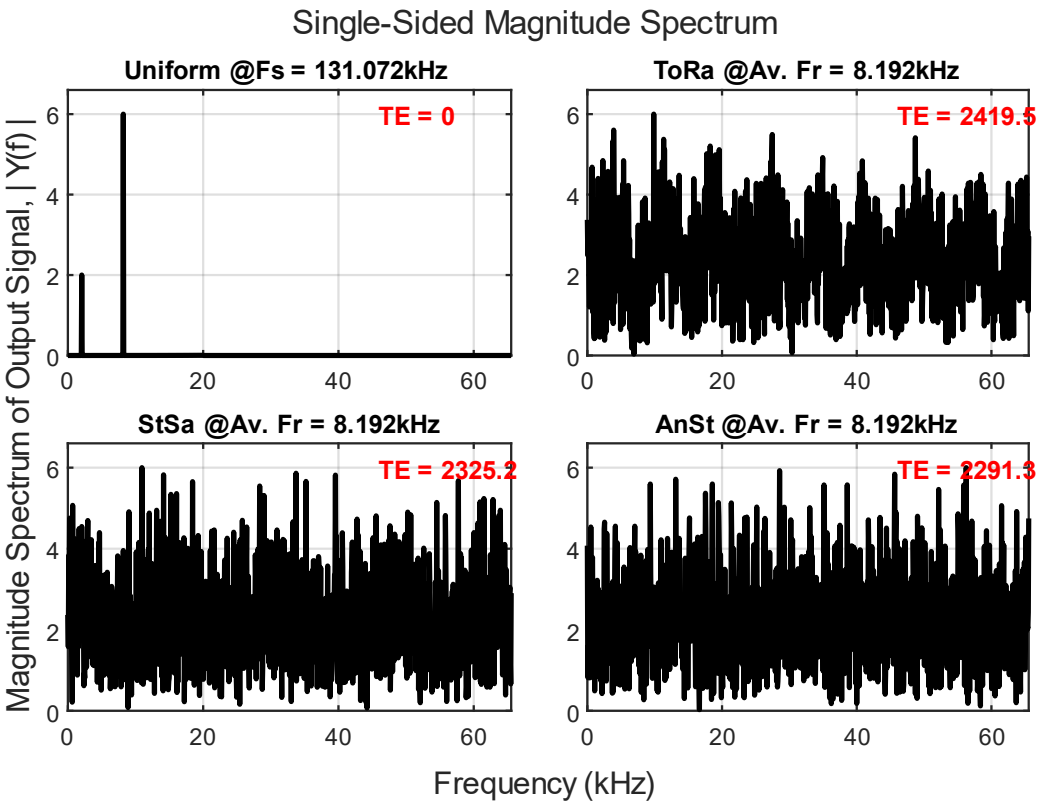


Fig. 24. Spectrum of the output signal. Uniform $F_s = 131.072kHz$. While random average sampling rate $Av.F_r = 8.192kHz$, and MC=1.

Indeed, the Landau rate is an indicator for the minimum density of sample points that can lead to lossless signal reconstruction. This does not contradict the fact that, for some sorts of signals, one can use lower sampling rates than Landau dictates when no full reconstruction of the sampled signal is needed, rather, computing some signal averaging statistics is the goal. The mean value of an ergodic signal is just an example.

3.13.4. BPF Example

To demonstrate the performance of the three randomised estimators when considering bandpass filtering, I redesigned the above brick-wall LPF filter to become a BPF with $22kHz$ bandwidth and $33kHz$ centre frequency, as shown in Fig. 25 below together with the spectrum of the same input analog signal, $x(t)$.

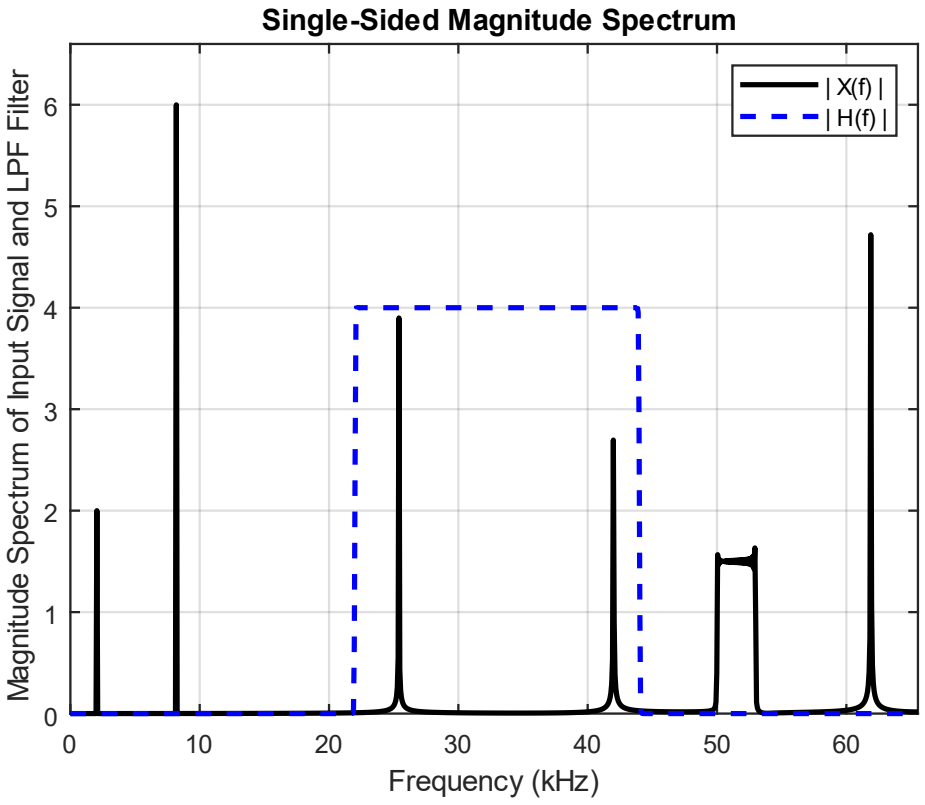


Fig. 25. Spectrum of the input analog signal (solid black) and the frequency response of the BPF filter cantered at $33kHz$ and has a bandwidth of $22kHz$. Uniform sampling rate is equal to the Nyquist rate, $F_s = 131.072kHz$.

In this example, the three random filter estimators, as well as an equivalent uniform conventional DSP filter, are examined against extracting the ultrasonic signals, spanning the frequency range from $22kHz$ to $44kHz$, in a number of cases regarding different average random sampling rates. All spectra shown in Fig. 26 (uniform and random) are fulfilling the Nyquist rate. As can be seen, it is possible to filter out the ultrasonic signals using my random estimators with relatively small amount of statistical errors.

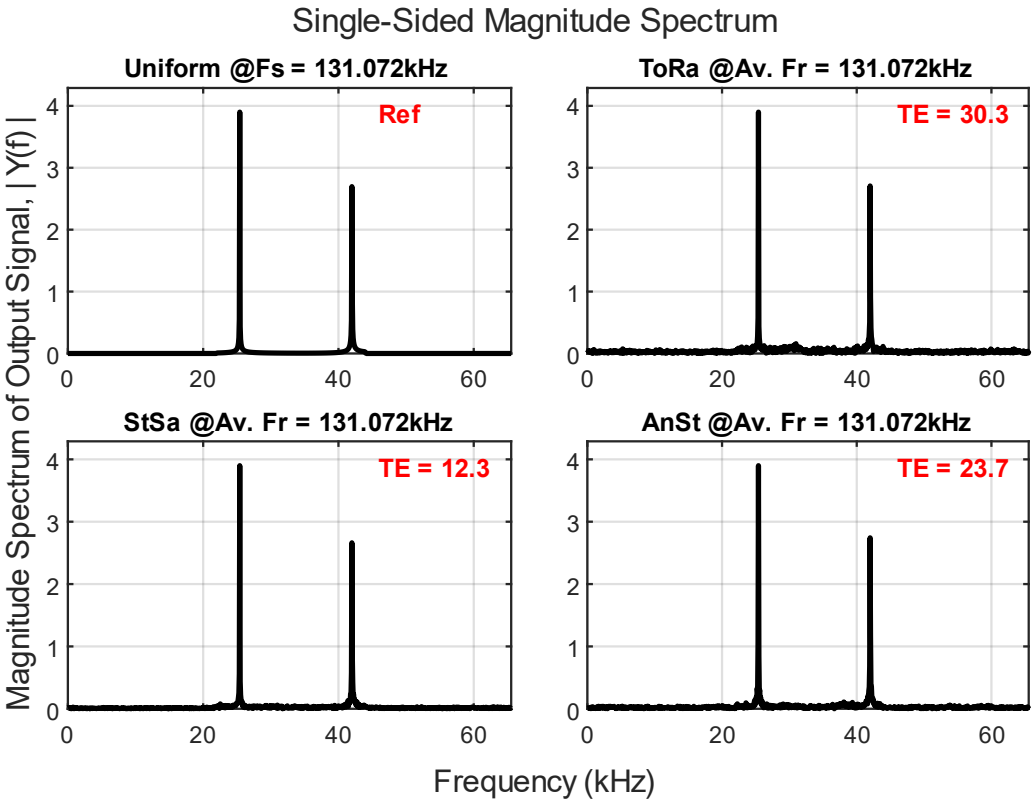


Fig. 26. Spectrum of the filter output signal. Uniform $F_s = 131.072kHz$. Also, random average sampling rate $Av.F_r = 131.072kHz$, and MC=100.

Whereas, in Figs. 27-28, the utilised sampling rates are less than the Nyquist rate. Namely, with only 89.6kHz, I am still capable of estimating the filtered ultrasonic signals quite fine with the random estimators. However, using the uniform approach, high frequency components fold back into the band of interest on a form of aliases,

see top-left sub-plot of Fig. 27. Theses aliases cannot be ruled out using uniform approach and under the assumptions of this example and others in this sub-section.

Another advantage of random estimation approaches is that it is possible to detect wider frequency range than half the average random sampling rate, as depicted in Figs. 28-29. Moreover, Interesting results of the estimators can be spotted in Fig. 29. With only one MC run and an average random sampling frequency, $Av.F_r$ of $42.24kHz$, it is clear that the two sinusoids of the ultrasonic signal range are detectable. Obviously, this rate is less than half the required Nyquist rate using uniform sampling, in this particular example, when no antialiasing pre-analog filtering is employed.

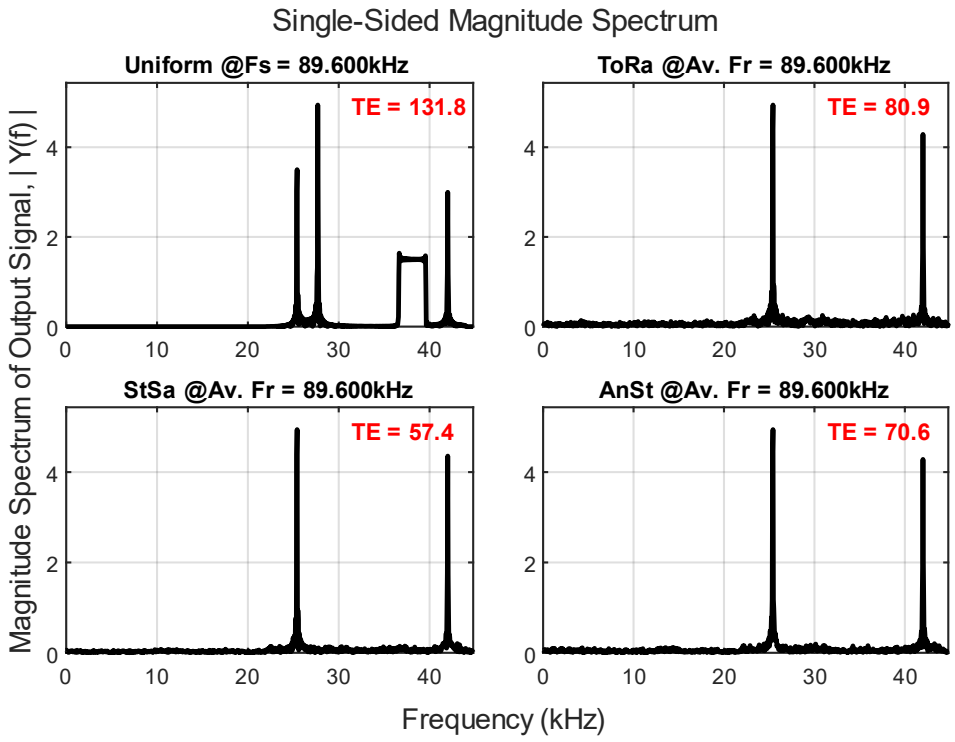


Fig. 27. Spectrum of the filter output signal. Uniform and average random sampling rates are equal to $89.6kHz$ and MC=100. Note the aliasing effect on uniform approach.

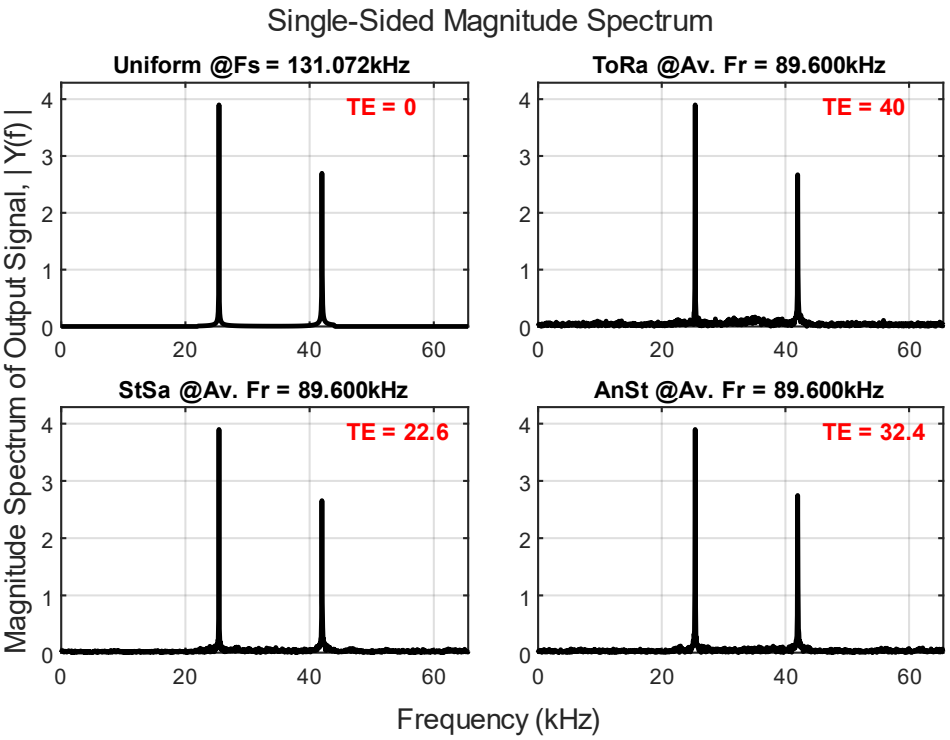


Fig. 28. With random sampling, it is possible to extend the detected spectrum beyond half the average random sampling rate. MC=100.

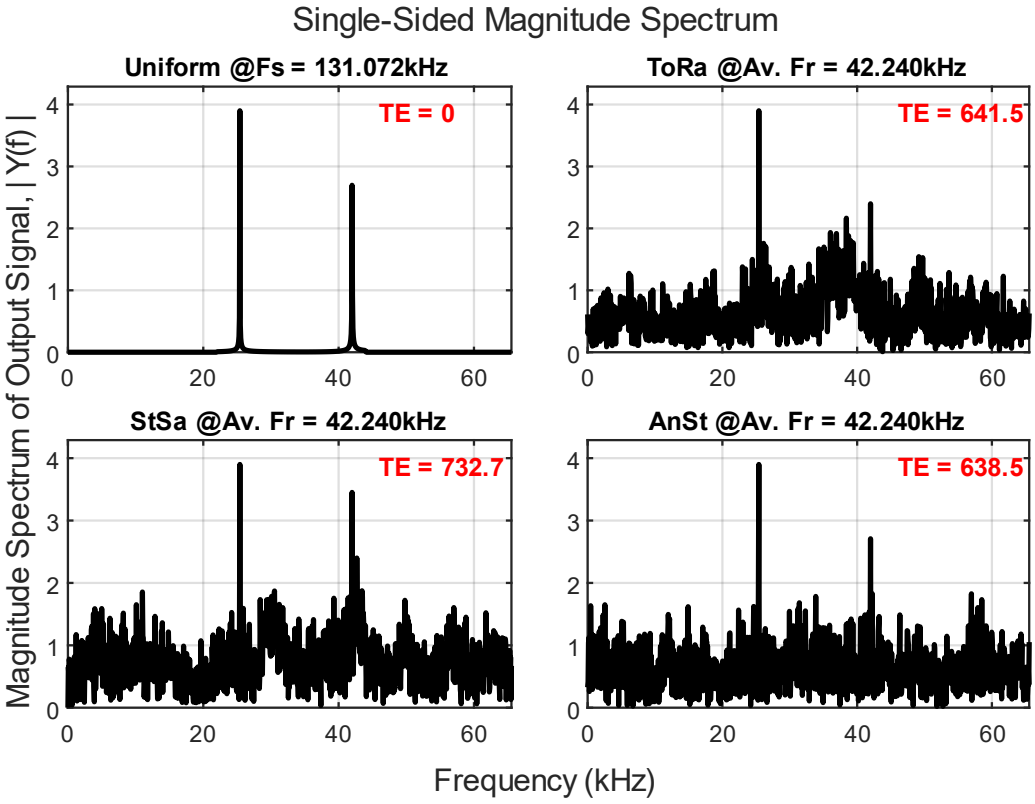


Fig. 29. $|Y(f)|$ with uniform sampling rate $F_s = 131.072kHz$ and average random sampling rate $Av.F_r = 42.24kHz$, MC=1.

3.13.5. Implementation Cost

In all discussed examples in this section, the uniform traditional DSP filtering approach that guarantees no aliasing present in the output signal is using a sampling frequency of $131.072kHz$ within a $[0,0.0156]$ sec observation window. The uniform FIR filter characterises are listed in Table IV, below.

Whereas the implementation cost of each random filter depends on the number of acquired points per the observation window in the corresponding example. This is exactly equal to the utilised average sampling frequency multiplied by $T = 0.0156$ sec. For example, the random estimators of Fig. 21 use $Av.Fr = 65.536kHz$ only, which means that to estimate one point of the filter output signal, a total of $N = 65.536kHz \times 0.0156\ sec \approx 1023$ input random sample points are required. i.e. the computational cost is 1023 multiplications and 1022 additions per each sample point. While for Fig. 28, I need 659 multiplications and 658 additions to estimate one point of output signal since $N = 42.24kHz \times 0.0156\ sec \approx 659$, here. These two examples reveal that a cost saving of about 50% and 68%, respectively, has been achieved when considering the randomised DASP estimators rather than the conventional uniform DSP filtering approach.

TABLE IV: UNIFORM FIR FILTER IMPLEMENTATION COST[4]

Discrete-Time FIR Filter (Real)	
Filter Length[5]	2049
Number of Multipliers	2047
Number of Adders	2046
Multiplications per Input Sample	2047
Additions per Input Sample	2046

[4] I am aware that other possible FIR filter designs may have smaller order than the presented one here. Of course, this has an impact on the impulse response of the uniform FIR filter and may lead to the comparison with random approach to be unfair. Moreover, with the assumptions that spectrum of the input signal is sparse and unknown, together with no analog prefiltering is available, then the most important factor in deciding the cost of FIR filtering between uniform and random is the number of sampling points within the fixed observation window, i.e. the sampling rate. It is the ADC preceding the FIR filter which consumes the most time and power in the whole operation.

[5] Including two zero coefficients at start and end of the filter impulse response.

CHAPTER 4

Pseudorandomised On-Grid Interpolation and Filtering Estimation

4.1. Overview

This chapter introduces the concept of pseudorandomised sampling and filtering [88]. A high-frequency equally spaced grid is designed so that its frequency resolution is an integer-multiple of the required Nyquist rate. Meaning, the smallest value of the integer multiplier is one, and in this case the Nyquist rate is guaranteed. Moreover, the larger the integer multiplier is, the more accurate filtering estimation would be achieved, and then aliasing destructive effects are either eliminated or reduced significantly.

The three random filter estimators introduced in Chapter 3 depend on the notion of instantaneous sampling of both input signal and a locally saved analog function representing the impulse response of a specific filter. This enabled us to have the convolution operation of the filter been carried out smoothly, without the need to search for a synchronised and corresponding impulse response counterpart for a specific randomly/pseudorandomly selected sample point from the input signal.

However, in many real-life applications with limited hardware/software capacity, it is not affordable to have a functionally-converted impulse response, and it can only be designed as a sequence of digital data representing the filter coefficients based on a specific uniform sampling frequency not less than the Nyquist rate required by the application for which this filter is designed to work in.

In this chapter, I explore pseudorandom sampling as a means to mitigating the aliasing problem. So, as the input signal under consideration is being pseudorandomly

sampled, I need a mechanism in which I can synchronise the input sample points with their corresponding filter coefficients. To achieve this goal, I propose the use of a dense and uniform grid of time instants at which pseudorandomised sampling of both the input signal and the impulse response should stick to by using, for instance, a pseudorandom number generator (PRNG) circuit/subroutine. The reciprocal of the grid's time step (spacing) is the grid frequency, and it should be a multi-fold higher than the required Nyquist rate to ensure good randomisation and alias-mitigating results.

For a specific filtering application, an on-grid filter is designed similar to any other conventional DSP filter, with the only exception is to presume the uniform sampling rate at which the filter is designed is way larger than the usual Nyquist rate. Then, the impulse response of the filter is saved on a lookup table inside the sampling circuit's memory resource. The pseudorandom sampling procedure is then carried out simultaneously between the input signal and the filter impulse response, and the convolution sum is calculated to estimate the filter output signal.

The layout of the remaining of this chapter is as follows: next section introduces what do I mean by on-grid sampling and filtering. Section 4.3 illustrates the pseudorandom filtering model. Then, in Section 4.4, ToRa, StSa and AnSt filter estimators based on the simple *Rectangular rule* are proposed. Next, in Section, 4.5 I introduce the nonuniform pseudorandom C3NS interpolation rule. Finally, I validate my findings by demonstrating some simulation examples.

4.2. On-Grid Sampling and Filtering

In a practical digital world, one of the challenges facing theoretical random sampling and filtering is that it is not possible to acquire or process sample points at infinitesimal time instants. So, for a given observation window, the number of sample points should be finite, and there must be a feasible minimum time spacing between the two closest signal samples. By "feasible" I mean the highest hardware sampling capabilities of state-of-the-art discretization circuitry or analog to digital converter (ADC).

Another challenge of filtering randomly sampled signals is the difficulty of working out mathematical operations between two (or more) time-based random variables,

representing some realisations of those signals, or even between one random variable and other equally spaced one since they might not be accurately synchronised, which definitely leads to significant estimation errors. Therefore, extreme attention should be paid to synchronise corresponding variables before being processed.

Pseudorandom sampling and filtering would be suitable to overcome such practical challenges. In essence, pseudorandomisation is a deterministic sampling approach that is based on a nonuniform selection of sample points form otherwise fictionally uniformly oversampled signal/impulse response.

To demonstrate how pseudorandom sampling and filtering would be achieved, a dense and on-grid filter impulse response is actually oversampled (black dots) as illustrated in Fig. 30. The input signal also can be imagined as if it is being uniformly oversampled (grey dots). As can be seen, the time instants of the true input signal samples (red dots) and corresponding impulse response coefficients (red dots, as well) are always aligned to the equally spaced grid, which is designed to have a uniform sampling frequency $F_s \gg F_{Nyq}$, and a grid time spacing $T_s = 1/F_s$. The grid frequency, F_s, depends on the hardware capabilities of a specific application. Nevertheless, the higher F_s is, the more accurate filtering estimation will be achieved.

4.3. Mathematical Model

Suppose that an input continuous-time signal, $x(t)$, is oversampled on a uniform grid with T_s time spacing and observed within $[t-T, t)$ time interval. The resulting uniform discrete-time sequence, $\{x_u(kT_s)\}_{k=0}^{N_u-1}$, which comprises a total of $N_u = T/T_s$ samples, is filtered using an FIR filter with its impulse response, $\{h(kT_s)\}_{k=0}^{N_u-1}$, been sampled on the same grid, i.e. similar to the example illustrated in Fig. 30. The output signal, denoted by $y_u(iT_s)$, is considered as a reference signal for other nonuniformly sampled and filtered signals to be introduced shortly. Thus, $y_u(iT_s)$ is given by

$$y_u(iT_s) = T_s \sum_{k=0}^{N_u-1} x_u(kT_s)\, h(iT_s - kT_s), \tag{4.1}$$

where k and i are both non-negative integers.

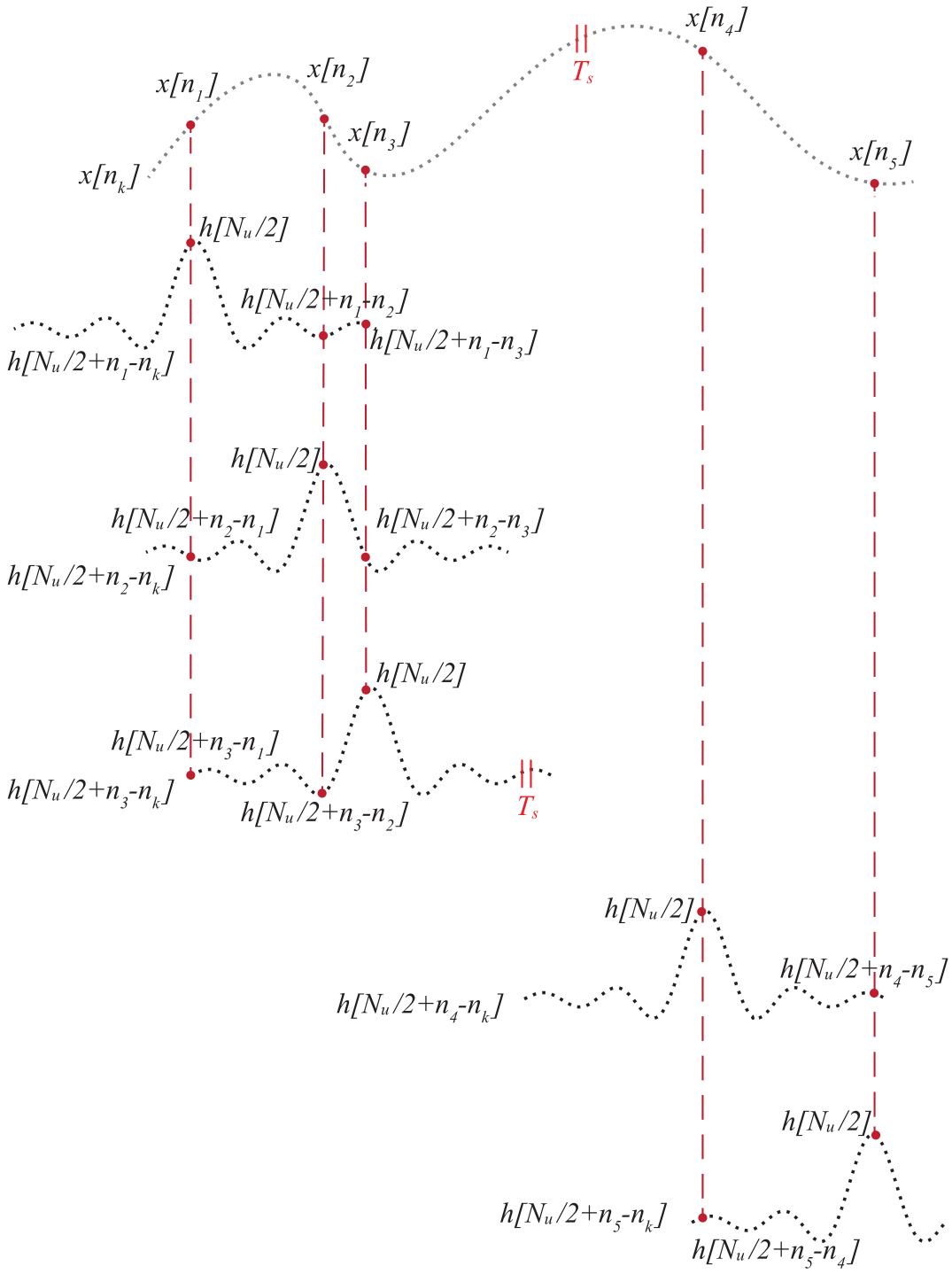


Fig. 30. Uniformly oversampled filter impulse response, $h(kT_s) \equiv$ black dots, and potential on-grid samples of input signal, $x(kT_s) \equiv$ grey dots. The true pseudorandom samples of the input signal, $x(n_kT_s) = x[n_k] \equiv$ red dots, are time-aligned with corresponding subset of impulse response coefficients, $h\left(\frac{T}{2} + n_jT_s - n_kT_s\right) = h[\frac{N_u}{2} + n_j - n_k] \equiv$ red dots also, selected from the whole sequence of coefficients already stored in memory buffer of a sampling and filtering circuit.

Remark that $y_u(iT_s)$ is just a uniform and dense discrete-time form of $y(t)$ given in the previous chapter for the continuous-time filtering case, i.e. $y(t) = \int_{t-T}^{t} x(\tau)h(t-\tau)d\tau$.

4.4. Convolution Estimation Based on Simple Rectangular rule

I aim at estimating the output of an FIR filter when the input signal is pseudorandomly sampled according to the grid setup shown above, then to compare the spectra of output signal and the reference signal.

Grid-based versions of ToRa, StSa, and AnSt are shown in Fig. 31. As illustrated, the time instants of the sample points are precisely aligned to the uniform grid. I assume that the time instants of the underlying grid are denoted by $t_i = iT_s$, where $i = 0, 1, 2, \dots, N_u - 1$, and the pseudorandomised time instants of input signal's sample points are $t_k = n_k T_s$, where $k = 0, 1, 2, \dots, N_r - 1$. Here N_r is the total number of pseudorandom sample points and it must be less than or equal to N_u. Moreover, the nonuniform integers $n_k \in \{0, N_u - 1\}$, and they directly depend on which random/ pseudorandom sampling technique is being used.

Regarding ToRa random sampling scheme, for example, and $[\mathrm{t} - T, t)$ observation window, and N_r i.i.d. pseudorandom sample points of an input signal $x(t)$, denoted by $x(t_k) = x(t)|_{t=t_k}$, then the probability mass function (PMF) of pseudorandom variables t_k is denoted by $p_{t_k}(t_k)$ and given by

$$p_{t_k}(t_k) = \begin{cases} \dfrac{1}{N_u}, & t_k \in G \\ 0, & elsewhere \end{cases}, \tag{4.2}$$

where $G = \{0, T_s, 2T_s, 3T_s, \dots, (N_u - 1)T_s\}$.

The discrete-time estimated output signal, $\hat{y}(t_j)$, filtered out using the aforementioned FIR grid-based filter is

$$\hat{y}(t_j) = \sum_{k=0}^{N_r-1} x(t_k)\, h(t_j - t_k)\, \Delta_k, \tag{4.3}$$

where $t_j = n_j T_s$, i.e. it is an integer multiple of T_s, and Δ_k is the average time spacing between the nonuniform sample points, and so, $\Delta_k = T/N_r$.

It is my choice to decide the time spacing of the output signal (uniform or nonuniform) according to how t_j are being selected. In this chapter, I consider them unequally spaced, i.e. n_j are pseudorandom integers, as well. This would enable us to add a second layer of randomized signal processing by estimating the FT of the estimated and irregularly spaced filter output signal. Thus, (4.3) can be rewritten as

$$\hat{y}(n_j T_s) = \frac{T}{N_r} \sum_{k=0}^{N_r-1} x(n_k T_s)\, h(n_j T_s - n_k T_s)\,. \tag{4.4}$$

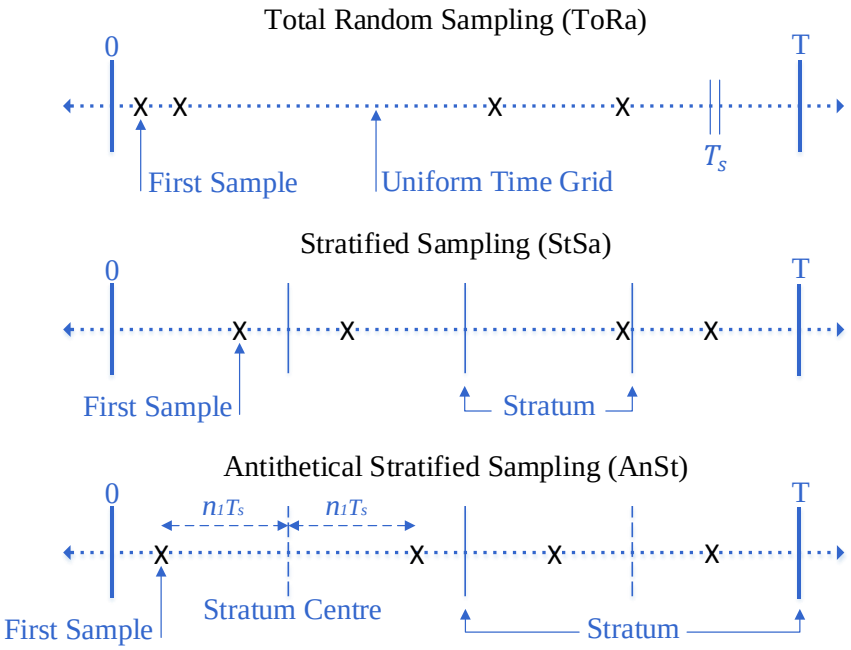


Fig. 31. Four sample points example of grid-based ToRa, StSa and AnSt random sampling techniques, where each sample point (black x) is precisely aligned to one of the uniform grid time instants (blue dots), [88].

Remark that n_k has the same PMF as t_k, i.e. $p_{n_k}(n_k) = 1/N_u$, $n_k \in \{0, 1, 2, \dots, N_u - 1\}$ and zero elsewhere. Hence, the N_r summands in (4.4) are all products of discrete random variables that have the same PMF.

4.5. Expected Value of the Estimator

The following theorem reveals the bias status of the discrete oversampled on-grid ToRa filter estimator.

Theorem 4.1. *The discrete on-grid ToRa filter estimator is unbiassed.*

Proof:

The expected value of the discrete ToRa filter estimator in (4.4) can be calculated by adding up the individual expected values of all components of the summation. For each component, we have

$$\mathbb{E}\big[\hat{y}^c(n_jT_s)\big] = \frac{T}{N_r}\mathbb{E}\big[x(n_kT_s)\, h(n_jT_s - n_kT_s)\big], \tag{4.5a}$$

$$= \frac{T}{N_r}\sum_{k=-\infty}^{\infty} x(n_kT_s)\, h(n_jT_s - n_kT_s)p_{n_k}(n_k). \tag{4.5b}$$

As $p_{n_k}(n_k)$ equals 0 outside $\{0, N_u - 1\}$ and $T = N_uT_s$, (3.5b) becomes

$$\mathbb{E}\big[\hat{y}^c(n_jT_s)\big] = \frac{N_uT_s}{N_r}\sum_{k=0}^{N_u-1} x(kT_s)\, h(n_jT_s - \mathrm{k}T_s)\frac{1}{N_u}, \tag{4.5c}$$

$$= \frac{T_s}{N_r}\sum_{k=0}^{N_u-1} x(kT_s)\, h(n_jT_s - \mathrm{k}T_s) = \frac{1}{N_r}y_u(n_jT_s). \tag{4.5d}$$

For N_r components of $\hat{y}^c(n_jT_s)$, the expectation is equal to

$$\mathbb{E}\big[\hat{y}(n_jT_s)\big] = N_r\left(\frac{1}{N_r}y_u(n_jT_s)\right) = y_u(n_jT_s). \tag{4.5e}$$

Therefore, the ToRa estimator in (4.4) is unbiased.

■

Indeed, similar analyses for other random sampling techniques (StSa and AnSt) can be done to verify the unbiasedness of their corresponding estimators. To cut long story short, I confirm that both estimators are unbiassed, as well, which coincides with the continuous-time case estimators validated in the previous chapter.

4.6. Quality of Estimation

To assess the quality of the pseudorandom estimator (4.4) and check its consistency, I need to verify its variance. As an unbiassed estimator, the MSE is identical to the variance. Furthermore, both input signal and filter impulse response are real-valued. The following theorem unveils the exact convergence rate of the discrete on-grid ToRa filter estimator.

Theorem 4.2. *The discrete on-grid ToRa filter estimator converges at a rate of* $N_r{}^{-1}$*, where* N_r *represents the number of pseudorandom sample points.*

Proof:

The variance is

$$\mathbb{V}[\hat{y}(n_j T_s)] = MSE[\hat{y}(n_j T_s)] = \mathbb{E}\left[\left(\hat{y}(n_j T_s)\right)^2\right] - \left(\mathbb{E}[\hat{y}(n_j T_s)]\right)^2. \tag{4.6}$$

First, I calculate $\mathbb{E}\left[\left(\hat{y}(n_j T_s)\right)^2\right]$,

$$\mathbb{E}\left[\left(\hat{y}(n_j T_s)\right)^2\right] = E\left\{\left(\frac{T}{N_r}\sum_{k=0}^{N_r-1} x(n_k T_s)\, h(n_j T_s - n_k T_s)\right)^2\right\}. \tag{4.7}$$

For simplicity, I omit the T_s notation from the arguments of functions in (4.7) and will be added when required. Now, the square value $\left(\hat{y}(n_j)\right)^2$ can be written as

$$\left(\hat{y}(n_j)\right)^2 = \left(\frac{T}{N_r}\right)^2 \sum_{k=0}^{N_r-1} \left(x(n_k)\, h(n_j - n_k)\right)^2 + \left(\frac{T}{N_r}\right)^2 \sum_{k=0}^{N_r-1} \sum_{\substack{m=0 \\ m\neq k}}^{N_r-1} \left(x(n_k)\, h(n_j - n_k)\right) \left((x(n_m)\, h(n_j - n_m)\right). \tag{4.8}$$

So, the expected value of $\left(\hat{y}(n_j)\right)^2$ is

$$\mathbb{E}\left[\left(\hat{y}(n_j)\right)^2\right] = \left(\frac{T}{N_r}\right)^2 \mathbb{E}\left[\sum_{k=0}^{N_r-1} \left(x(n_k)\, h(n_j - n_k)\right)^2\right] + \left(\frac{T}{N_r}\right)^2 \mathbb{E}\left[\sum_{k=0}^{N_r-1} \sum_{\substack{m=0 \\ m\neq k}}^{N_r-1} \left(x(n_k)\, h(n_j - n_k)\right) \left((x(n_m)\, h(n_j - n_m)\right)\right]. \tag{4.9}$$

I benefit from the law of the unconscious statistician to use the same PMF of n_k to calculate the expectation terms in the RHS of (4.9). Moreover, since the double summations in the second expectation term are statistically independent of each other, i.e. $m \neq k$ means $n_m \neq n_k$, then the expectation of this term can be calculated by multiplying the two individual expectations. That is

$$\mathbb{E}\left[\left(\hat{y}(n_j)\right)^2\right] = \left(\frac{T}{N_r}\right)^2 N_r \sum_{k=-\infty}^{\infty} x^2(n_k)\, h^2(n_j - n_k)\, p_{n_k}(n_k) + \left(\frac{T}{N_r}\right)^2 \mathbb{E}\left[\sum_{k=0}^{N_r-1} \left(x(n_k)\, h(n_j - n_k)\right)\right] \times \mathbb{E}\left[\sum_{\substack{m=0 \\ m\neq k}}^{N_r-1} \left((x(n_m)\, h(n_j - n_m)\right)\right], \tag{4.10a}$$

$$\mathbb{E}\left[\left(\hat{y}(n_j)\right)^2\right] = \left(\frac{N_u T_s}{N_r}\right)^2 \frac{N_r}{N_u} \sum_{k=0}^{N_u-1} x^2(n_k) h^2(n_j - n_k)$$

$$+ \left(\frac{N_u T_s}{N_r}\right)^2 \frac{N_r}{N_u} \sum_{k=0}^{N_u-1} x^2(n_k) h^2(n_j - n_k) \tag{4.10b}$$

$$\times \frac{(N_r - 1)}{N_u} \sum_{k=0}^{N_u-1} x^2(n_k) h^2(n_j - n_k),$$

$$\mathbb{E}\left[\left(\hat{y}(n_j)\right)^2\right] = \frac{N_u}{N_r} {T_s}^2 \sum_{k=0}^{N_u-1} x^2(n_k)\, h^2(n_j - n_k)$$

$$+ \frac{(N_r - 1)}{N_r} \left(T_s \sum_{k=0}^{N_u-1} x(n_k)\, h(n_j - n_k) \right) \tag{4.10c}$$

$$\times \left(T_s \sum_{\substack{m \neq k \\ m=0}}^{N_u-1} x(n_m)\, h(n_j - n_m) \right),$$

$$\mathbb{E}\left[\left(\hat{y}(n_j)\right)^2\right] = \frac{T}{N_r} E_{y_u} + \frac{(Nr - 1)}{N_r} \times {y_u}^2(n_j T_s), \tag{4.10d}$$

where E_{y_u} is the total energy of the output signal y_u in the interval $[t - T, t)$.

Substituting (4.10d) and (4.5e) into (4.6) and explicitly redisplaying T_s yields

$$\mathbb{V}\left[\hat{y}(n_j T_s)\right] = \frac{T}{N_r} E_{y_u} + \frac{(Nr - 1)}{N_r} \times {y_u}^2(n_j T_s) - {y_u}^2(n_j T_s), \tag{4.11a}$$

$$\mathbb{V}\left[\hat{y}(n_j T_s)\right] = \frac{T E_{y_u} - {y_u}^2(n_j T_s)}{N_r}. \tag{4.11b}$$

Multiplying the variance in (4.11b) by N_r and then calculating the limit, we get

$$\lim_{N_r \to \infty} \left(N_r \times \mathbb{V}\left[\hat{y}(n_j T_s)\right]\right) = T E_{y_u} - {y_u}^2(n_j T_s). \tag{4.12}$$

This means that the discrete-time ToRa filter estimator is converging at a rate of ${N_r}^{-1}$.

■

It is obvious from the results of Theorem 4.2 that increasing the number of sample points will result in smaller value for the variance. So, there is a trade-off between the quality of estimation and the number of pseudorandom sample points.

Same analyses are carried out for the other two estimators, StSa and AnSt, in the discrete-time pseudorandom form, and I found that they converge at rates of $N_r{}^{-3}$ and $N_r{}^{-5}$ respectively.

4.7. Composite 3-Nonuniform-Sample (C3NS) Interpolation Rule

The filter convolution sum (4.3) for a specific output sample is nothing more than calculating an area under the curve (AUC) of the summand function. It usually uses the simple rectangle (or midpoint) rule to do so. In this section, nonetheless, I propose another method that can estimate AUC faster than the Rectangular rule, depending on Lagrange unequal interpolation techniques. The so-called *composite 3-nonuniform-sample* (C3NS) rule is introduced here, and it utilises the high-resolution uniform grid structure of the impulse response and the PRNG-based sampling of the input signal. Basically, it somehow mimics the traditional composite Simpson's 1/3 rule, introduced in Chapter 2, which interpolates three uniform sample points to generate a second-order polynomial. My proposed C3NS rule deals with interpolating three pseudorandom on-grid unequally spaced sample points.

4.7.1. Area Under the Curve Estimation

Given a continuous-time function, $f(t)$ for instance. It is requested to estimate AUC of $f(t)$ in the time interval $[0, T)$ by using C3NS nonuniform interpolation rule. To this end, the time interval $[0, T)$ is split into subintervals according to the number of pseudorandom sample points. Every three consecutive points represent one subinterval. For example, the first subinterval is depicted in Fig. 32. The borders of any generic subinterval (i.e. the two sample points at the left and right ends of the subinterval) are shared with neighbor subintervals to compose the C3NS rule. The total area from 0 to T is then estimated by adding up the definite integrals of interpolated polynomials in all subintervals

To further illustrate how C3NS works, suppose that the first three pseudorandom sample points, which comprise only one subinterval, are acquired at time instants $t_0 = 0$, t_1 and t_2. Since I am pseudorandomly picking up samples from a uniform dense grid having a time step of T_s, then all acquired sample points are spaced from each other by an integer multiple of T_s. Now, I introduce two integer numbers, n_1 and n_2, as follows: $n_1 T_s = t_1 - t_0$ and $n_2 T_s = t_2 - t_1$. Area under the curve of $f(t)$ from t_0 to $t_2 = (n_1 + n_2)T_s$ is estimated by interpolating the three samples $f(t_0)$, $f(t_1)$ and $f(t_2)$ using a quadratic Lagrange parabola $P(t) = a_0 + a_1 t + a_2 t^2$, where a_0, a_1 and a_2 are just the coefficients of the polynomial. Then, I integrate the second-order polynomial $P(t)$ from t_0 to t_2. I find that the result of the definite integral in this subinterval is equal to

$$\int_{t_0}^{t_2} f(t)dt \approx \frac{T_s(n_1 + n_2)}{6n_1 n_2}[(2n_1 n_2 - {n_2}^2)f(t_0) + (n_1 + n_2)^2 f(t_1) + (2n_1 n_2 - {n_1}^2)f(t_2)]. \tag{4.13}$$

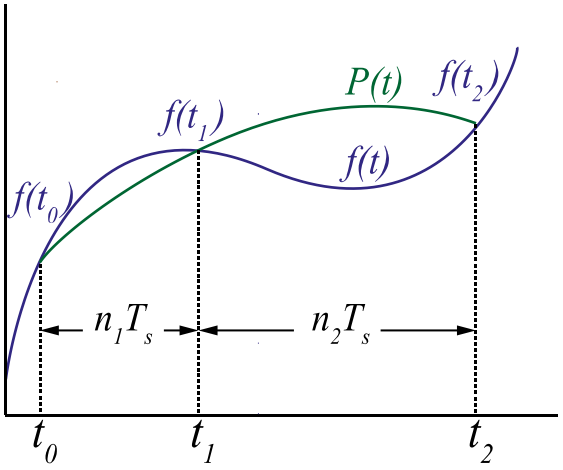


Fig. 32. One subinterval, $[t_0, t_2]$, of the proposed C3NS rule. Original function is $f(t)$, and the interpolated parabola is $P(t)$. n_1 and n_2 are pseudorandom integers, and T_s is the uniform grid time resolution, [88].

The total area of $f(t)$ within $[0, T)$ can now be estimated by accumulating results of

integration of all n subintervals. If the total number of pseudorandom sample points is N_r, then $N_r = 2n + 1$. Let n_{i1} and n_{i2} denote the corresponding integers n_1 and n_2, respectively, of the i-th subinterval. Thus, we get

$$\int_0^T f(t)dt \approx T_s \sum_{i=1}^{n} \left(\frac{(n_{i1} + n_{i2})}{6n_{i1}n_{i2}} \left((2n_{i1}n_{i2} - {n_{i2}}^2) f(t_{2i-2}) + (n_{i1} + n_{i2})^2 f(t_{2i-1}) + (2n_{i1}n_{i2} - {n_{i1}}^2) f(t_{2i}) \right) \right). \tag{4.14}$$

4.7.2. Error Analysis

The estimation error associated with the first subinterval, e_1, is

$$e_1 = \int_0^{(n_1+n_2)T_s} f(t)dt - \frac{T_s(n_1 + n_2)}{6n_1n_2} [(2n_1n_2 - {n_2}^2) f(t_0) + (n_1 + n_2)^2 f(t_1) + (2n_1n_2 - {n_1}^2) f(t_2)]. \tag{4.15}$$

Since $f(t)$ is assumed to be continuous and differentiable, I can expand it using Taylor series expansion at $t = t_1 = n_1T_s$. Thus,

$$f(t) = f(t_1) + (t - n_1T_s) f^{(1)}(t_1) + \frac{1}{2}(t - n_1T_s)^2 f^{(2)}(t_1) + \frac{1}{6}(t - n_1T_s)^3 f^{(3)}(t_1) + \frac{1}{24}(t - n_1T_s)^4 f^{(4)}(t_1) + O((t - n_1T_s)^5), \tag{4.16}$$

where $f^{(i)}(t_1)$ is the i-th order derivative of $f(t)$ about t_1, and $O(\cdot)$ is the big-O notation.

Considering (4.16), I now find the values of $f(t_0 = 0)$, $f(t_1)$, and $f(t_2)$,

$$f(t_0) = f(t_1) + (n_1T_s) f^{(1)}(t_1) + \frac{1}{2}(n_1T_s)^2 f^{(2)}(t_1) + \frac{1}{6}(n_1T_s)^3 f^{(3)}(t_1) + \frac{1}{24}(n_1T_s)^4 f^{(4)}(t_1) + O((n_1T_s)^5), \tag{4.17}$$

$$f(t_1) = f(t_1) \tag{4.18}$$

$$\begin{aligned} f(t_2) = f(t_1) + n_2 T_s f^{(1)}(t_1) + \frac{1}{2}(n_2 T_s)^2 f^{(2)}(t_1) + \frac{1}{6}(n_2 T_s)^3 f^{(3)}(t_1) \\ + \frac{1}{24}(n_2 T_s)^4 f^{(4)}(t_1) + O((n_2 T_s)^5). \end{aligned} \tag{4.19}$$

By substituting (4.16)-(4.19) into the expression of e_1 given in (4.15), we obtain

$$\begin{aligned} e_1 = \int_0^{(n_1+n_2)T_s} \Bigg(f(t_1) + (t - n_1 T_s) f^{(1)}(t_1) + \frac{1}{2}(t - n_1 T_s)^2 f^{(2)}(t_1) \\ + \frac{1}{6}(t - n_1 T_s)^3 f^{(3)}(t_1) + \frac{1}{24}(t - n_1 T_s)^4 f^{(4)}(t_1) + O((t - n_1 T_s)^5) \Bigg) dt \\ - \frac{T_s(n_1 + n_2)}{6 n_1 n_2} [(2 n_1 n_2 - {n_2}^2) f(t_0) + (n_1 + n_2)^2 f(t_1) \\ + (2 n_1 n_2 - {n_1}^2) f(t_2)], \end{aligned} \tag{4.20}$$

$$\begin{aligned} e_1 = (n_1 + n_2) T_s f(t_1) + \frac{1}{2}({n_2}^2 - {n_1}^2) {T_s}^2 f^{(1)}(t_1) \\ + \frac{1}{6}({n_2}^3 + {n_1}^3) {T_s}^3 f^{(2)}(t_1) + \frac{1}{24}({n_2}^4 - {n_1}^4) {T_s}^4 f^{(3)}(t_1) \\ + \frac{1}{120}({n_2}^5 + {n_1}^5) {T_s}^5 f^{(4)}(t_1) + O(({n_2}^6 - {n_1}^6) {T_s}^6) \\ - \frac{T_s(n_1 + n_2)}{6 n_1 n_2} \Bigg[(2 n_1 n_2 - {n_2}^2) \Bigg(f(t_1) - n_1 T_s f^{(1)}(t_1) \\ + \frac{1}{2}(n_1 T_s)^2 f^{(2)}(x_1) - \frac{1}{6}(n_1 T_s)^3 f^{(3)}(t_1) + \frac{1}{24}(n_1 T_s)^4 f^{(4)}(t_1) \\ + O((n_1 T_s)^5) \Bigg) + (n_1 + n_2)^2 f(t_1) \\ + (2 n_1 n_2 - {n_1}^2) \Bigg(f(t_1) + n_2 T_s f^{(1)}(t_1) + \frac{1}{2}(n_2 T_s)^2 f^{(2)}(t_1) \\ + \frac{1}{6}(n_2 T_s)^3 f^{(3)}(t_1) + \frac{1}{24}(n_2 T_s)^4 f^{(4)}(t_1) + O((n_2 T_s)^5) \Bigg) \Bigg]. \end{aligned} \tag{4.21}$$

By simplifying the expression in (4.21), we obtain the following result for the error in the first subinterval,

$$e_1 = \frac{{T_s}^4(n_1 + n_2)^3(n_1 - n_2)}{72} f^{(3)}(t_1) - \frac{{T_s}^5(n_1 + n_2)^3(4n_1^2 - 7n_1n_2 + 4n_2^2)}{720} f^{(4)}(t_1) + O((n_m T_s)^5), \tag{4.22}$$

where n_m is the maximum of n_1 and n_2, i.e. the dominant. However, max $(n_m T_s) \leq \frac{2T}{N_r}$ for StSa-based C3NS pseudorandom sampling, for example. So, $O((n_a T_s)^5) = O({N_r}^{-5})$. Remark that $O({N_r}^{-5})$ is also in $O({N_r}^{-4})$. Therefore, if $n_1 \neq n_2$, then (4.22) can be further simplified to

$$e_1 = \frac{{T_s}^4(n_1 + n_2)^3(n_1 - n_2)}{72} f^{(3)}(t_1) + O({N_r}^{-4}). \tag{4.23}$$

If $n_1 = n_2 = n_{\mathrm{m}}$ in (4.22), then the error term for this subinterval is significantly reduced to

$$e_{1,\, n_1 = n_2} = -\frac{{n_{\mathrm{m}}}^5 {T_s}^5}{90} f^{(4)}(t_1) + O({N_r}^{-5}). \tag{4.24a}$$

$$e_{1,\, n_1 = n_2} = -\frac{16T^5}{45{N_r}^5} f^{(4)}(t_1) + O({N_r}^{-5}). \tag{4.24b}$$

The expression in (4.24a) is the same as the error for uniform Simpson's 1/3 rule as found in the literature, but without n_{m}, as $n_{\mathrm{m}} T_s$ here is identical to T_s over there, and both represent the time spacing between the equidistant sample points.

As can be noticed from (4.22), there is a trade-off in selecting identical or different n_1 and n_2. If they are similar, then this means the sampling is uniform, hence, aliasing will appear if the utilised sampling frequency does not fulfil the Nyquist rate. Whereas different n_1 and n_2 values lead to nonuniform sampling, which helps mitigate aliasing problems, but, at the same time, produce larger error value.

The error term in (4.22) is for one subinterval only. To calculate the total composite error for n subintervals, e_{C3NS}, we need to add up the errors of all subintervals,

$$e_{\text{C3NS}} = \sum_{i=1}^{n} \left(\frac{{T_s}^4 (n_{i1} + n_{i2})^3 (n_{i1} - n_{i2})}{72} f^{(3)}(t_{2i-1}) \right.$$
$$- \frac{{T_s}^5 (n_{i1} + n_{i2})^3 (4n_{i1}^2 - 7n_{i1}n_{i2} + 4n_{i2}^2)}{720} f^{(4)}(t_{2i-1}) \qquad (4.25)$$
$$\left. + O((n_m T_s)^5) \right)$$

On the other hand, the MSE in C3NS nonuniform interpolation, based on StSa sampling approach, can be calculated by computing the expected value of the squared error term in (4.23). In a quick glance, I can deduce that the MSE will be converging at N_r^{-7} rate since squaring (4.23) leads to terms in $O\left({N_r}^{-8}\right)$. Then, multiplying by the PDF ($= 1/N_r$) is cancelled by the N_r-proportional definite integral of the expectation. Finally, I sum up the errors of $N_r/2$ subintervals leading eventually to $O(N_r^{-7})$ convergence rate. Simulation results below validate this rate, as well as all the rates for the above pseudorandom based estimators.

4.8. Simulation Results

Consider a test function $f(t) = \pi\, sinc(22t) + sin(2\pi \times 17t) + cos(2\pi \times 21.4t) + t^3$ that is observed in the interval [0,1] sec. This function is continuous and differentiable. I have numerically integrated this function using pseudorandom on-grid sampling based on two estimation criteria: simple Rectangular rule based on ToRa, StSa, AnSt estimators; and nonuniform interpolation rule C3NS based on StSa. Numerical examples use different grid frequencies and different number of Monte Carlo iterations. The simulation results approve my abovementioned findings and clearly show that the grid-based pseudorandom estimators are uniformly converging to the true value of the estimated function's AUC at speeds of N^{-1}, N^{-3}, N^{-5}, and N^{-7} for ToRa, StSa, AnSt, and C3NS respectively, as shown in Figs. 33-39.

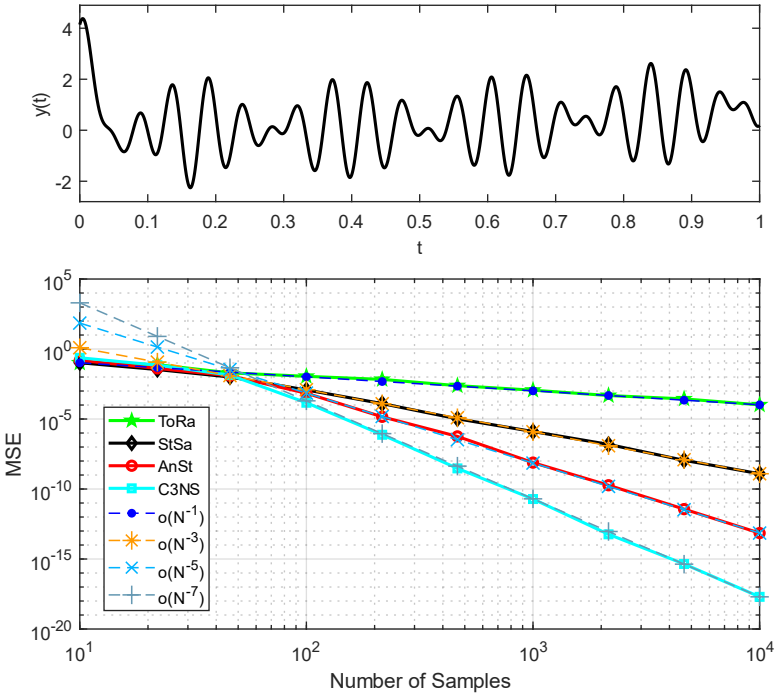


Fig. 33. Pseudorandom on-grid Rectangular rule (ToRa, StSa, and AnSt) estimators and nonuniform interpolation rule (C3NS) estimator. Grid frequency$= 2MHz$ and MC $= 100$ iterations. The plots show the variance of estimating AUC of the function $f(t)$ within [0,1] sec interval.

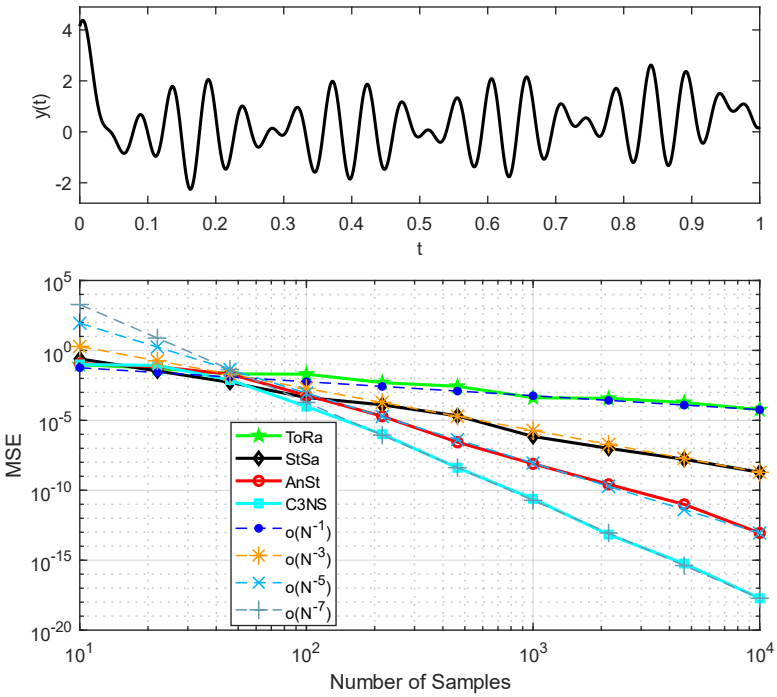


Fig. 34. Grid frequency $= 2MHz$ and MC $= 10$ iteration.

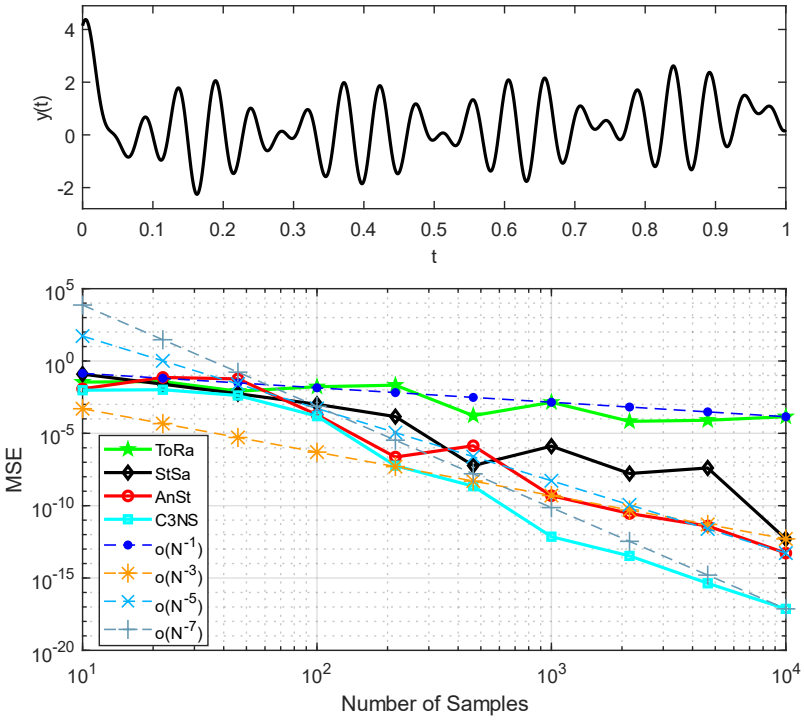


Fig. 35. Grid frequency $= 2MHz$ and MC $= 1$ iteration. It is obvious that the non-smoothness in the curves is because they are a result of only one realisation of the pseudorandom time instants of the sample points utilised by the estimators.

As already known, a single or a few realisations of the pseudorandom sample points does not reflect the smooth asymptotic behaviour of any random estimator. This is exactly what can be deduced by comparing Figs. 33-35, as they all reflect the same example and configurations except for the number of MC iterations. The higher number of iterations is, the smoother the asymptotic behaviour of the estimator can be obtained.

In the following set of figures, I demonstrate how the resolution of grid frequency can affect estimator's statistical features. As long as the grid frequency is large enough to ensure that the integers n_1 and n_2, in every subinterval, to have a big pool of pseudorandom integers to choose from, then the estimation is quite good, as the case of Fig. 33, and Figs. 36-37. Further decrease of grid frequency makes estimation errors to increase, as shown in Figs. 38-39.

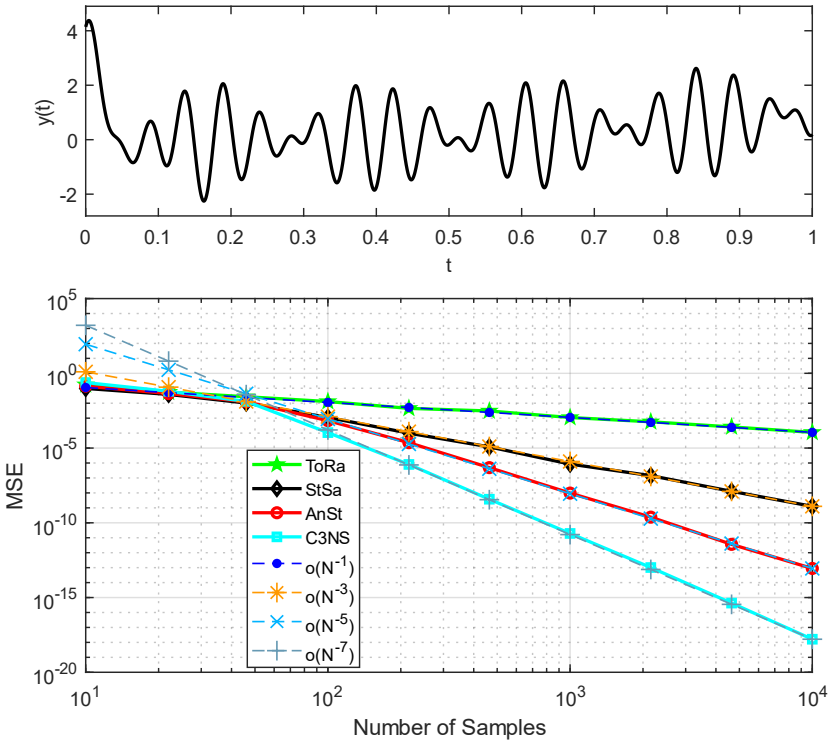


Fig. 36. Grid frequency resolution $= 1MHz$ and MC $= 100$ iterations.

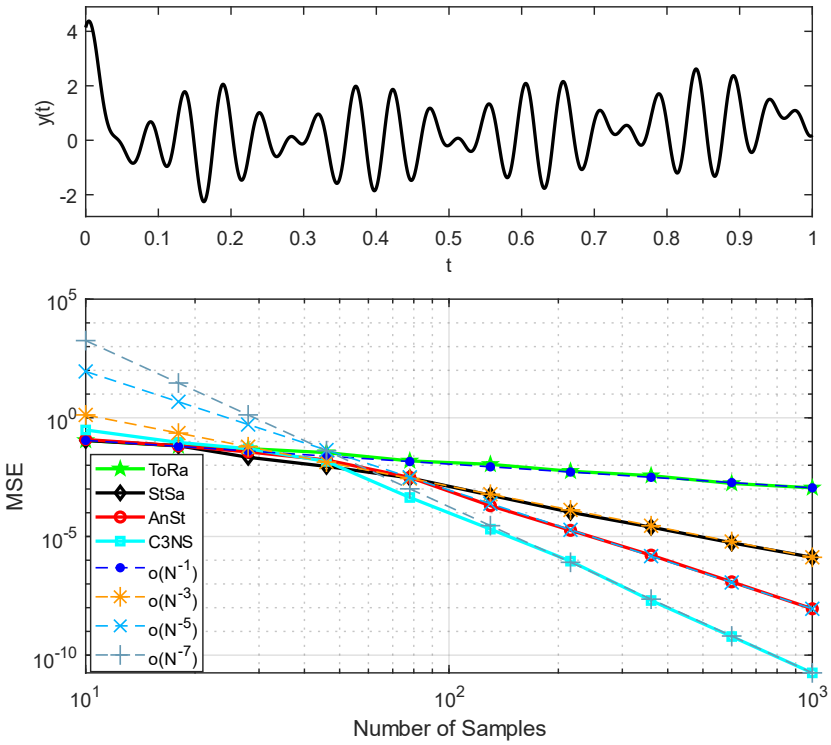


Fig. 37. Grid frequency resolution $= 0.1MHz$ and MC $= 100$ iteration.

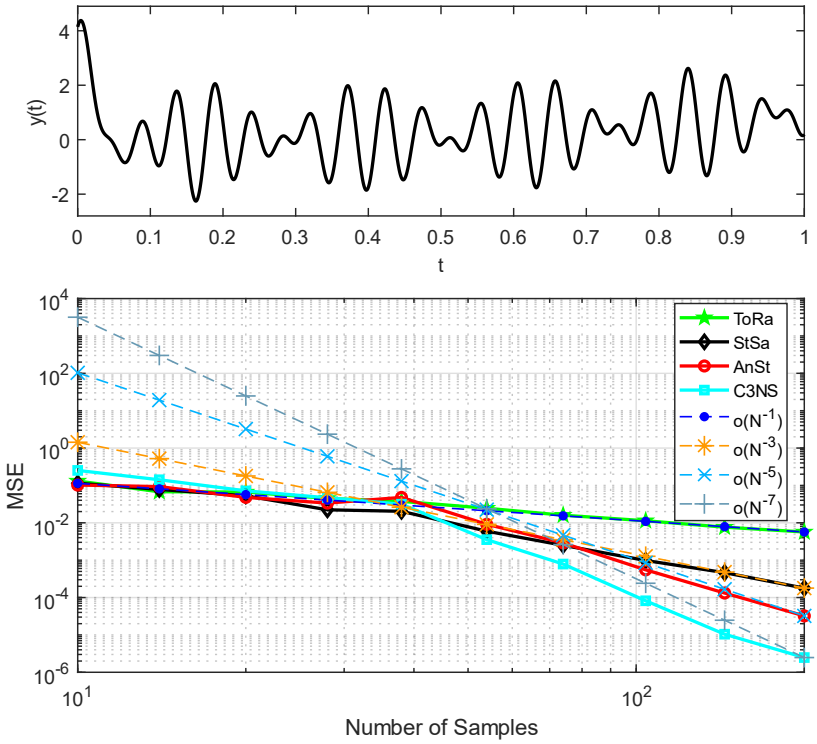


Fig. 38. Grid frequency resolution $= 1kHz$ and MC $= 100$ iteration.

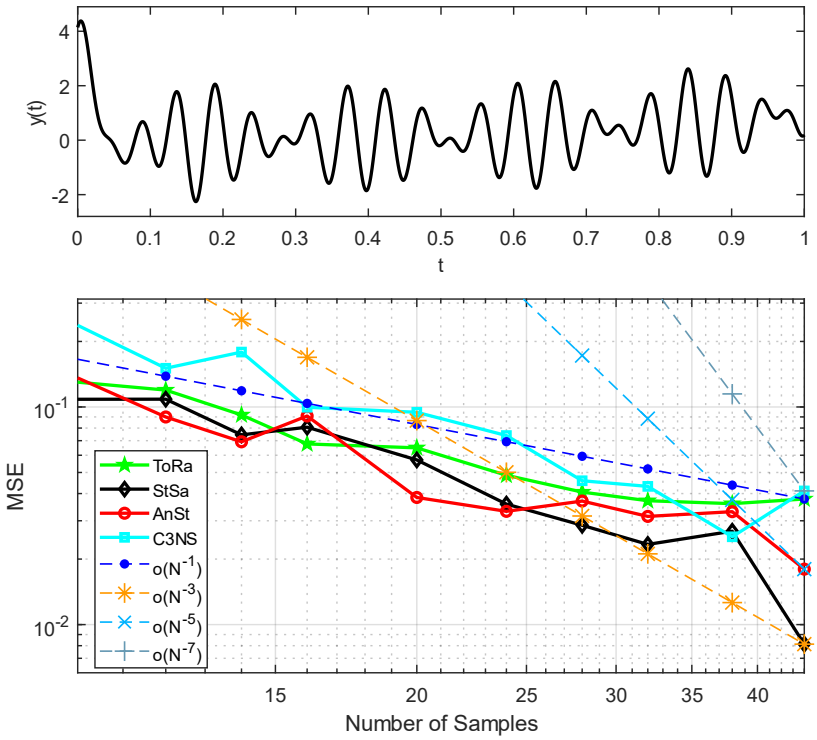


Fig. 39. Grid frequency resolution $= 500Hz$ and MC $= 100$ iteration.

Part II

Non-Smooth Integrand Functions

CHAPTER 5

Stratified-Sampling Estimator: Piecewise-Continuous Case

5.1. Overview

In Part I of this thesis, I proposed the StSa filter estimator and investigated its various statistical features for the case when the integrand function, $f(t,\tau)$, of the convolution operation is continuous, bounded and square-integrable, as well as its first two derivatives. The conditions on continuity of the zero-order derivative (ZOD), first-order derivative (FOD), and second-order derivative (SOD) were necessary to apply the Taylor series expansion of the integrand function where needed.

However, in real-life applications, the integrand's ZOD, FOD and SOD are not always continuous. On the contrary, there are many examples where discontinuities in the input signal $x(t)$, the impulse response of the filter $h(t)$, or both do exist, as well as their derivatives. In power electronics, for example, discontinuities appear on the rectified or clipped signals even if the original AC signal is smooth and bounded [13]. In communication, phase-shift keying (PSK), binary phase-shift keying (BPSK), digital data, and signals alike, all involve discontinuities in one or more orders of the derivative. In control systems, transient signals are normally unsmooth. In stock market, financial data frequently shows discontinuities, especially at serious global or local events, such as wars, political unrests, natural catastrophes, and pandemics (e.g. COVID-19 [86]).

Regarding the impulse response of the utilised filter, discontinuities may present in the function(s) representing it or its (their) derivatives. Furthermore, several filter designs include window functions to smooth out the frequency response of the filter by decreasing the overshoot and ripples in the passband. Some common window functions are intrinsically discontinuous or piecewise-continuous, such as exponential,

hamming, triangular, and rectangular windows. Theoretically, any suitable mathematical function can be used as a window function for some designed filter. In SDR, for example, providing flexible software-based filter designs widen the frequency band of operation and make more choices available for specific set of applications. Therefore, piecewise-continuous (or even discontinuous) filter impulse responses are likely to be designed.

Such potential non-smoothness in the integrand function (i.e. in its constituents $x(t)$, $h(t)$, or both) motivates us to study its impact on the statistical properties of the StSa filter estimator, where the integrand function, as a whole, cannot be expanded using Taylor series any more for it is no longer assumed to be continuous, as was the case in Chapter 3.

Moreover, in Chapter 3, the StSa filter estimator was introduced in (3.9a-b) for the continuous integrand function case. It was then proved that this estimator is unbiased, consistent and converging at a rate of N^{-3}, where N is half the total number of sample points acquired using StSa random sampling scheme.

In this chapter, on the other hand, which revises my published paper [89], I assess the decaying rate of the StSa estimator as a function of the number of sample points, $2N$, taking into account two main cases of non-smoothness of the integrand function. The main considered cases of $f(t,\tau)$, with respect to τ, are:

1. piecewise-continuous FOD;
2. piecewise-continuous ZOD.

The respective derivative in each single case is assumed to be bounded and has limited number of discontinuities. In the sequel, I prove that the StSa estimator uniformly converges to $y(t)$, i.e. the true value of the filter output signal, at a rate of N^{-3} in the first case (i.e. case number 1), and N^{-2} in the second case.

The rest of this chapter is composed of four sections: in the next section (Section 5.2), I briefly revisit the StSa random sampling technique to make the subsequent analysis look intact and integrated. Section 5.3 introduces the StSa filter estimator for non-smooth integrand functions, where a comprehensive study of the impact of potential discontinuities in the integrand function or its first two derivatives is analytically

presented. On the other hand, Section 5.4 summarises the rate of convergence of the ToRa estimator in the presence of discontinuities in FOD and ZOD, because it is directly related to proofs of theorems presented in this chapter. Finally, numerical examples and simulations are provided in Section 5.5 to validate the analytical findings and derivations.

5.2. StSa Technique

Fig. 40 shows how StSa sampling technique works. An analog input signal is randomly sampled using $2N$ sample points within an observation window $[t-T, t]$, where $2N$ is also equal to the number of strata in the whole window. This means that every stratum contributes one sample point, only. The time instant of the j-th sample point, τ_j, is selected randomly from the stratum's time span, $A_j = [S_{j-1}, S_j)$.

Strata lengths could be equal, or not. It depends on the relevant application and how much a priori knowledge is available. For the sake of this thesis, considering equidistant strata is more than enough to unveil the StSa estimator's main statistical properties, and also it simplifies the analysis. However, this form of equal partitioning of strata needs not to be the optimum one for any given application. For example, if a sufficient information about the signal to be sampled is available, we may concentrate more strata near sharp rises and falls of the signal. For advanced partitioning of strata, reader is referred to [52]. In this thesis, I assume that there is no sufficient information about the sampled-to-be signal and, therefore, all strata have the same length, which is equal to $T/2N$.

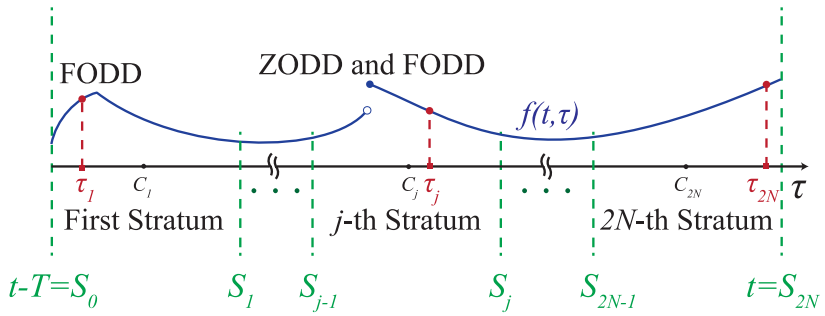


Fig. 40. StSa sampling technique, where the first sampling instant in the j-th stratum, τ_j, is chosen randomly from the stratum's time interval $[S_{j-1}, S_j)$. The last D letter in the acronyms ZODD and FODD denotes discontinuity.

5.3. StSa Estimator Using Non-Smooth Functions

Back to my filtering model (3.2) of Chapter 3, where an LTI system with $h(t)$ as its CT impulse response has been used to filter out an input signal $x(t)$, within a time interval $[t-T, t)$, to yield the output signal $y(t)$, I would like to borrow a few equations from that chapter for the sake of this analysis, as follows:

$$y(t) = \int_{t-T}^{t} f(t,\tau)d\tau, \tag{5.1}$$

where $f(t,\tau)$ was defined as

$$f(t,\tau) := x(\tau)h(t-\tau). \tag{5.2}$$

The equidistant-strata StSa filter estimator introduced in Chapter 3 was

$$\hat{y}(t) = \Delta \sum_{j=1}^{2N} f(t,\tau_j), \tag{5.3}$$

where $2N$ is the number of strata (= the total number of utilised sample points), Δ is the stratum length (= $T/2N$), and τ_j is the time-instant of the j-th stratum's sample point. All τ_j time instants are i.i.d random variables with a uniform distribution PDF of

$$p_j(\tau) = \begin{cases} 1/\Delta, & \tau \in A_j \\ 0\ , & \text{elsewhere} \end{cases}. \tag{5.4}$$

As per the two discontinuity cases suggested above in section 5.1, where there are some finite and bounded discontinuities in the integrand function or its FOD, an illustration example of a generic j-th stratum for which $f(t,\tau)$ is not smooth is depicted in Fig. 41. In the sequel, it is assumed that in any given stratum, there is no more than one discontinuity in $f(t,\tau)$ or its FOD, $f'(t,\tau)$. However, multiple different-order derivative discontinuities are allowed per some specific strata if they occur exactly at the same time instant, as shown in Fig. 41, where both ZODD and FODD exist at τ_{D_j}.

The two subfunctions, $f_{j,L}(t,\tau)$ and $f_{j,R}(t,\tau)$ that appear in Fig. 41 represent the left- and right-hand portions of $f(t,\tau)$ in the j-th stratum, respectively. Both subfunctions are smooth and differentiable. Remark that the δ-wide fictional extensions, depicted in dashed brown, are just extrapolations of both subfunctions for incremental periods of time to guarantee that the subfunctions are Taylor series expandable about the very time instant of the discontinuity, i.e. τ_{D_j}.

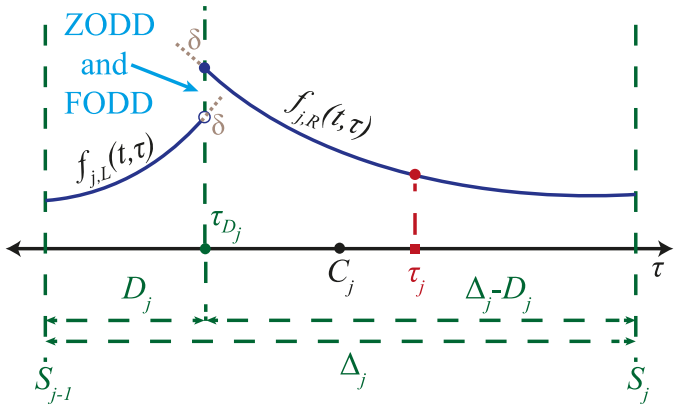


Fig. 41. j-th stratum of a non-smooth $f(t,\tau)$, where a discontinuity in both FOD and ZOD occurs at a time instant τ_{D_j}. Two smooth subfunctions, $f_{j,L}(t,\tau)$ and $f_{j,R}(t,\tau)$, are also shown.

Mathematically, the two smooth subfunctions, $f_{j,L}(t,\tau)$ and $f_{j,R}(t,\tau)$, can be defined as

$$f_{j,L}(t,\tau) := \begin{cases} f(t,\tau), & \tau \in A_{j,L} \\ f_{j,xL2R}(t,\tau), & \tau_{D_j} \leq \tau \leq \tau_{D_j} + \delta \end{cases}, \tag{5.5}$$

$$f_{j,R}(t,\tau) := \begin{cases} f(t,\tau), & \tau \in A_{j,R} \\ f_{j,xR2L}(t,\tau), & \tau_{D_j} - \delta \leq \tau < \tau_{D_j} \end{cases}, \tag{5.6}$$

where $f_{j,xL2R}(t,\tau)$ and $f_{j,xR2L}(t,\tau)$ are the fictional extrapolated extensions. Consequently, the integrand function, $f(t,\tau)$, in the whole observation interval can be rewritten in terms of both subfunctions as

$$f(t,\tau)=\sum_{j=1}^{2N}\begin{cases} f_{j,L}(t,\tau), & \tau\in A_{j,L} \\ f_{j,R}(t,\tau), & \tau\in A_{j,R}\end{cases}. \tag{5.7}$$

The main impact of non-smooth integrand function (i.e. discontinuities present in its ZOD) on the statistical properties of the filter estimator is to slow down its convergence rate. Whereas, no change in the decaying rate of the estimator if the discontinuities are only in the FOD, provided they are finite and bounded.

Let us define the amplitudes of the FOD and ZOD discontinuities (jumps), if any, in the j-th stratum as F_{1j} and F_{0j} respectively,

$$F_{1j}:=f'_{j,L}\left(t,\tau_{D_j}\right)-f'_{j,R}\left(t,\tau_{D_j}\right). \tag{5.8a}$$

$$F_{0j}:=f_{j,L}\left(t,\tau_{D_j}\right)-f_{j,R}\left(t,\tau_{D_j}\right). \tag{5.8b}$$

5.3.1. Piecewise-Continuous FOD

Assume that the FOD (= $f'(t,\tau)$) of the integrand function has an M bounded jumps, M is a positive integer, while both $f'(t,\tau)$ and $f(t,\tau)$ are smooth and continuous. Suppose the FOD discontinuities occur at time instants $\{\tau_{D_j}\}_{j\in I_M}$, where $I_M=\{i_1,i_2,i_3,\dots,i_M\}$ is a set of indices for the M strata with discontinuities. Accordingly, the following equations hold,

$$f_{j,L}\left(t,\tau_{D_j}\right)=f_{j,R}\left(t,\tau_{D_j}\right)=f\left(t,\tau_{D_j}\right), \tag{5.9a}$$

$$f'_{j,L}\left(t,\tau_{D_j}\right)\neq f'_{j,R}\left(t,\tau_{D_j}\right). \tag{5.9b}$$

Meaning, $F_{0j}=0$ and $F_{1j}\neq 0$.

Now, I am going to verify the consistency of the StSa estimator, (5.3), in the discontinuous case. First, I need to find the variance of the estimator to check if it is converging to the true value of the filter output signal, $y(t)$, as the sample size increases to infinity. To this end, I formulate my j-th sub-estimator, ϕ_j, as

$$\phi_j = f(t,\tau_j)\Delta \tag{5.10a}$$

$$\phi_j = \Delta \begin{cases} f_{j,L}(t,\tau_j), & \tau_j \in A_{j,L} \\ f_{j,R}(t,\tau_j), & \tau_j \in A_{j,R} \end{cases} \tag{5.10b}$$

According to the law of the unconscious statistician (LOTUS), all subfunctions comprising ϕ_j have the same PDF ($\equiv p_j(\tau)$). Hence, the expected value of ϕ_j can be calculated as

$$E[\phi_j] = \int_{-\infty}^{\infty} p_j(\tau) f(t,\tau)\Delta \, d\tau, \tag{5.11a}$$

$$= \int_{S_{j-1}}^{S_j} f(t,\tau) d\tau, \tag{5.11b}$$

$$= \int_{A_{j,L}} f_{j,L}(t,\tau) d\tau + \int_{A_{j,R}} f_{j,R}(t,\tau) d\tau. \tag{5.11c}$$

I can now expand $f_{j,L}(t,\tau)$ and $f_{j,R}(t,\tau)$ about $\tau = \tau_{D_j}$ using Taylor series since all of them are continuous and differentiable. Thus, the expanded form of (5.11c) is

$$\begin{aligned} E[\phi_j] = &\int_{A_{j,L}} \left(f_{j,L}\left(t,\tau_{D_j}\right) + \left(\tau - \tau_{D_j}\right) f'_{j,L}\left(t,\tau_{D_j}\right) + o\left(\left|\tau - \tau_{D_j}\right|\right)\right) d\tau \\ &+ \int_{A_{j,R}} \left(f_{j,R}\left(t,\tau_{D_j}\right) + \left(\tau - \tau_{D_j}\right) f'_{j,R}\left(t,\tau_{D_j}\right) + o\left(\left|\tau - \tau_{D_j}\right|\right)\right) d\tau, \end{aligned} \tag{5.12a}$$

$$\begin{aligned} E[\phi_j] = &f_{j,R}\left(t,\tau_{D_j}\right)\Delta - \frac{1}{2}{K_j}^2 \left(f'_{j,L}\left(t,\tau_{D_j}\right) - f'_{j,R}\left(t,\tau_{D_j}\right)\right)\Delta^2 \\ &+ \frac{1}{2}(1 - 2K_j) f'_{j,R}\left(t,\tau_{D_j}\right)\Delta^2 + o(\Delta^2). \end{aligned} \tag{5.12b}$$

Equation (5.12b) is obtained by evaluating the integrals in (5.12a) and considering equations (5.9a-b). Also, under the equality (5.9a) and a specific time shift t, the point $f_{j,R}\left(t,\tau_{D_j}\right)$ can be interchangeably replaced by $f_{j,L}\left(t,\tau_{D_j}\right)$ anytime within this section, if necessary, to make analysis easier.

The j-th error term, e_j, can be calculated by subtracting (5.12b) from (5.10b), that is

$$e_j = \phi_j - E[\phi_j], \tag{5.13a}$$

$$e_j = \begin{cases} f_{j,L}(t,\tau_j)\Delta - E[\phi_j], & \tau_j \in A_{j,L} \\ f_{j,R}(t,\tau_j)\Delta - E[\phi_j], & \tau_j \in A_{j,R} \end{cases} = \begin{cases} e_{j,L}, & \tau_j \in A_{j,L} \\ e_{j,R}, & \tau_j \in A_{j,R} \end{cases}, \tag{5.13b}$$

where

$$\begin{aligned} e_{j,L} = &\left(f_{j,L}(t,\tau_j) - f_{j,L}\left(t,\tau_{D_j}\right)\right)\Delta + \frac{1}{2}{K_j}^2 F_{1j}\,\Delta^2 \\ &-\frac{1}{2}(1-2K_j)f'_{j,R}\left(t,\tau_{D_j}\right)\Delta^2 - o(\Delta^2), \end{aligned} \tag{5.13c}$$

$$\begin{aligned} e_{j,R} = &\left(f_{j,R}(t,\tau_j) - f_{j,R}\left(t,\tau_{D_j}\right)\right)\Delta + \frac{1}{2}{K_j}^2 F_{1j}\,\Delta^2 \\ &-\frac{1}{2}(1-2K_j)f'_{j,R}\left(t,\tau_{D_j}\right)\Delta^2 - o(\Delta^2), \end{aligned} \tag{5.13d}$$

The j-th sub-estimator's variance, $\mathbb{V}[\phi_j]$, is equal to the expected value of the error-squared in the j-th stratum. That is

$$\mathbb{V}[\phi_j] = \int_{-\infty}^{\infty} p_j(\tau)\left(e_j\right)^2 d\tau, \tag{5.14a}$$

$$\mathbb{V}[\phi_j] = \frac{1}{\Delta}\int_{A_{j,L}} \left(e_{j,L}\right)^2 d\tau + \frac{1}{\Delta}\int_{A_{j,R}} \left(e_{j,R}\right)^2 d\tau. \tag{5.14b}$$

I present the following theorem for the variance of the StSa filter estimator. For the proof, see Appendix A.

Theorem 5.1. *Assume that a total of $2N$ sample points are randomly acquired from a continuous, real-valued, and bounded integrand function $f(t,\tau)$ over an observation window, $[t-T,T]$ using StSa technique. If the integrand function's FOD, $f'(t,\tau)$, is piecewise-continuous with a finite number of bounded discontinuities, M. Then,*

1. *the StSa filter estimator's variance, $\mathbb{V}[\hat{y}(t)]$, is*

$$\mathbb{V}[\hat{y}(t)] = \frac{T^4}{12(2N)^4}\sum_{j\in I_M}\left(3{\beta_j}^4{F_{1j}}^2 + \left(f'_{j,L}\left(t,\tau_{D_j}\right)\right)^2\right) + \frac{T^3}{12(2N)^3}\sum_{\substack{j=1\\ j\notin I_M}}^{2N}\left(\left(f'\left(t,C_j\right)\right)^2\Delta\right) + o(N^{-3}), \tag{5.15a}$$

where $\beta_j = K_j - 1$, *and*

2. *the StSa filter estimator converges uniformly at a rate of* N^{-3} *and satisfies*

$$\lim_{N\to\infty}((2N)^3 \times \mathbb{V}[\hat{y}(t)]) = \frac{T^3}{12}\sum_{k=1}^{M+1}\int_{T_{k-1}}^{T_k}\left(f'(t,\tau)\right)^2 d\tau, \tag{5.15b}$$

where $\{T_k\}_{k=1}^{M}$ *is a set of time instants at which* $f'(t,\tau)$ *has jump discontinuities, whereas* $T_0 := t - T$ *and* $T_{M+1} := t$.

It is evident from (5.15b) of Theorem 5.1 that if $M = 0$, i.e. the FOD has no discontinuities at all, then the estimator still uniformly converges at a speed of N^{-3}, and (5.15b) will simplify to (5.16), which is identical to (3.30b) for the continuous case of the StSa filter estimator, as discussed in Chapter 3 of this thesis,

$$\lim_{N\to\infty}((2N)^3 \times \mathbb{V}[\hat{y}(t)]) = \frac{T^3}{12}\int_{t-T}^{t}\left(f'(t,\tau)\right)^2 d\tau. \tag{5.16}$$

The results of Theorem 5.1 emphasise that even if the integrand function is not smooth over the whole observation window, then the convergence rate is identical to the continuous case under the presumed conditions. This particular finding is stronger than that in [52] for the StSa Fourier transform (FT) estimator, which stipulated that smoothness of the integrand function is a necessary condition for the StSa FT estimator to work properly.

5.3.2. Piecewise-Continuous ZOD

I now consider the second case regarding the smoothness of the integrand function. If the integrand function itself is not smooth and has a limited and bounded discontinuities at M time instants, $\{\tau_{D_j}\}_{j\in I_M}$, then

$$f_{j,L}\left(t,\tau_{D_j}\right) \neq f_{j,R}\left(t,\tau_{D_j}\right), \tag{5.17a}$$

$$f'_{j,L}\left(t,\tau_{D_j}\right) \neq f'_{j,R}\left(t,\tau_{D_j}\right). \tag{5.17b}$$

Similar analysis to the previous sub-section is conducted here, taking into account the jump discontinuities given in (5.17a-b). To cut long story short, I conclude the final results of the analysis in the following theorem, Theorem 5.2. Appendix B presents the proof of this theorem.

Theorem 5.2. *Assume there are M bounded ZOD discontinuities in $f(t,\tau)$ within an observation interval $[t-T,T]$, where M is a finite integer. Suppose that the set of integers $I_M = \{i_1, i_2, i_3, \dots, i_M\}$ represents the strata indices where such discontinuities happen. Then,*

1. the variance of StSa filter estimator utilising $2N$ sample points is

$$\begin{aligned}\mathbb{V}[\hat{y}(t)] = \sum_{j\in I_M}\Bigg(&{\beta_j}^2\,{F_{0j}}^2\Delta^2 - {\beta_j}^3 F_{0j}\,F_{1j}\,\Delta^3 + \frac{1}{4}{\beta_j}^4{F_{1j}}^2\Delta^4 \\ &+ \frac{1}{12}\left(f'_{j,L}\left(t,\tau_{D_j}\right)\right)^2\Delta^4\Bigg) + \sum_{\substack{j=1\\ j\notin I_M}}^{2N}\left(\frac{1}{12}\left(f'\left(t,C_j\right)\right)^2\Delta^4\right) + o(\Delta^3),\end{aligned} \tag{5.18a}$$

where $\beta_j = (K_j - 1)$, and

2. the StSa filter estimator converges uniformly at an exact rate of N^{-2}, and satisfies

$$\lim_{N\to\infty}\left((2N)^2 \times \mathbb{V}[\hat{y}(t)]\right) = T^2\sum_{j\in I_M}{\beta_j}^2\,{F_{0j}}^2. \tag{5.18b}$$

The results of Theorem 5.2 emphasise my findings in the previous sub-section. Namely, if $\{F_{0j} = 0\}_{j\in I_M}$, that is, the integrand function's ZOD discontinuities do not exist at all, then (5.18a) will exactly reduce to (5.15a). Consequently, the StSa estimator will then be converging at N^{-3} rate.

5.4. ToRa Estimator and the Presence of Discontinuities

Based on (3.23a) and (A.3), I can confirm that if there are FOD discontinuities then ToRa filter estimator still converges at a speed of N^{-1}. More interestingly, if the integrand function itself is also piecewise-continuous, i.e. there is a finite number of bounded ZOD discontinuities, then it follows from (3.23a) and (B.6) that the ToRa estimator will still be converging at the same rate, N^{-1}.

5.5. Numerical Results

To validate my analytical derivations and findings, two sets of simulation examples have been conducted. First, simple abstract functions are considered, where either there are no discontinuities at all (i.e. infinitely differentiable functions) or there are a finite number of bounded discontinuities in the FOD. I estimate the integrals of the abstract functions using my StSa estimator, then compare the results with the true integral values. In the second set of numerical examples, however, the StSa estimator is tested on true filter designs and smooth and non-smooth input signals.

5.5.1. Abstract Functions and the StSa Estimator

I have randomly sampled the following functions using StSa technique within an observation window $[0,0.5)$ sec,

$$f_1(t) = 2.5\,sin(18\pi t) - 3.8\,cos(42\pi t), \tag{5.19a}$$

$$f_2(t) = f_1(t) - 18 \times SAWTOOTH(24\pi(t - 0.1825),0.5), \tag{5.19b}$$

$$f_3(t) = f_1(t) - 18 \times SAWTOOTH(60\pi(t - 0.262),1), \tag{5.19c}$$

where $SAWTOOTH(g(t), m)$ is the built-in *sawtooth* function in the MATLAB. if m is equal to either zero or one, then the $SAWTOOTH$ is identical to $g(t) - \lfloor g(t) \rfloor$, where the $\lfloor . \rfloor$ sign denotes the standard *floor* function. Whereas, setting m between zero and one renders the $SAWTOOTH$ function as the *triangle* wave function, where $triangle(g(t)) = \frac{2}{\pi} arcsin(sin(g(t)))$.

Note that $f_1(t)$ is smooth with no jumps at any order of the derivative. The FOD of $f_2(t)$ is piecewise-continuous with twelve jump discontinuities. And finally, $f_3(t)$ is piecewise-continuous and it has fifteen ZOD jumps. Figs. 42-45 depict the functions' plots together with their respective StSa estimators' variances.

To guarantee that the depicted performance is not associated with a specific realisation of the StSa random sampling settings, I have conducted 100 Monte Carlo simulations (except for Fig. 43 which comprises one MC iteration only, and it is included here intentionally for comparison purposes) for each single function. Thus, the demonstrated figures are reliable and reflects the true behaviour of the StSa estimator.

Remark that the asymptotic convergence rates of the StSa estimator, as shown in Figs. 42-45, are N^{-3}, N^{-3}, N^{-3}, and N^{-2} for the respective abstract functions $f_1(t)$ and MC=100, $f_1(t)$ and MC=1, $f_2(t)$, and $f_3(t)$. This precisely confirms the analytical results established in the previous two sub-sections.

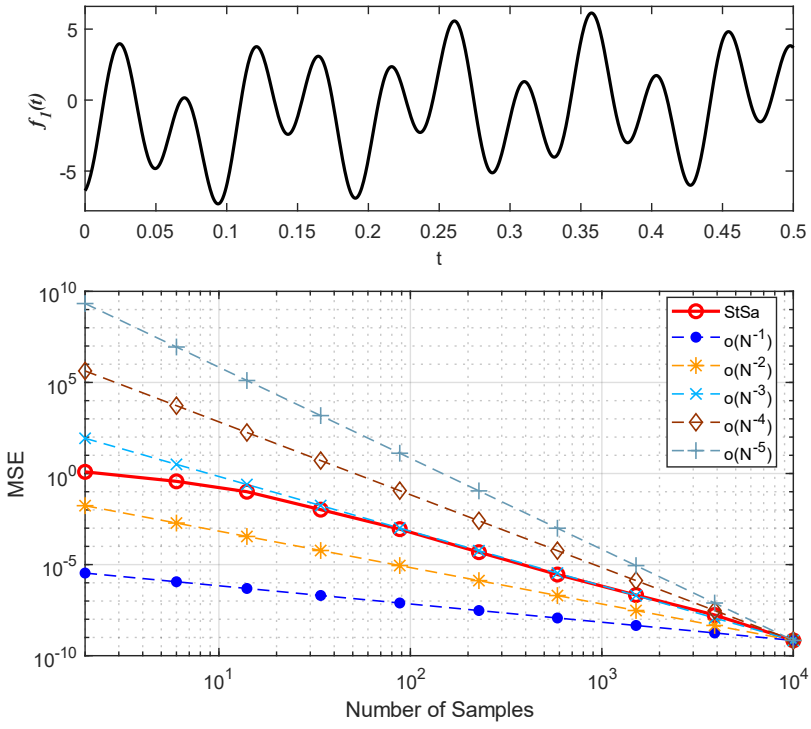


Fig. 42. The abstract function $f_1(t)$ is smooth, and there are no FOD discontinuities at all. Remark the uniform convergence rate of StSa estimator, which is equal to N^{-3}. MC=100 iterations have been carried out independently.

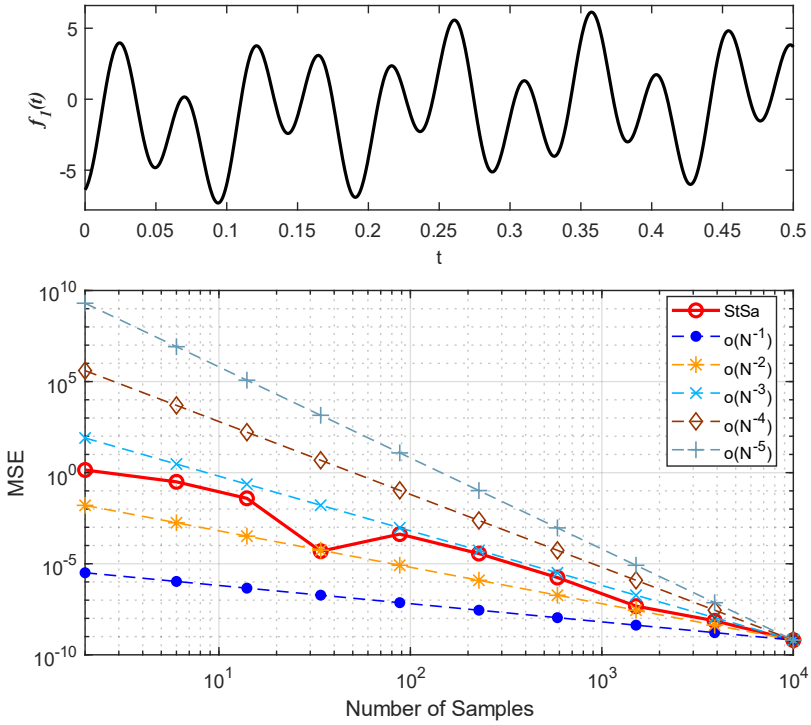


Fig. 43. The abstract function $f_1(t)$ is smooth, and there are no FOD discontinuities at all. The uniform convergence rate of StSa estimator is N^{-3}. Only one MC iteration is used.

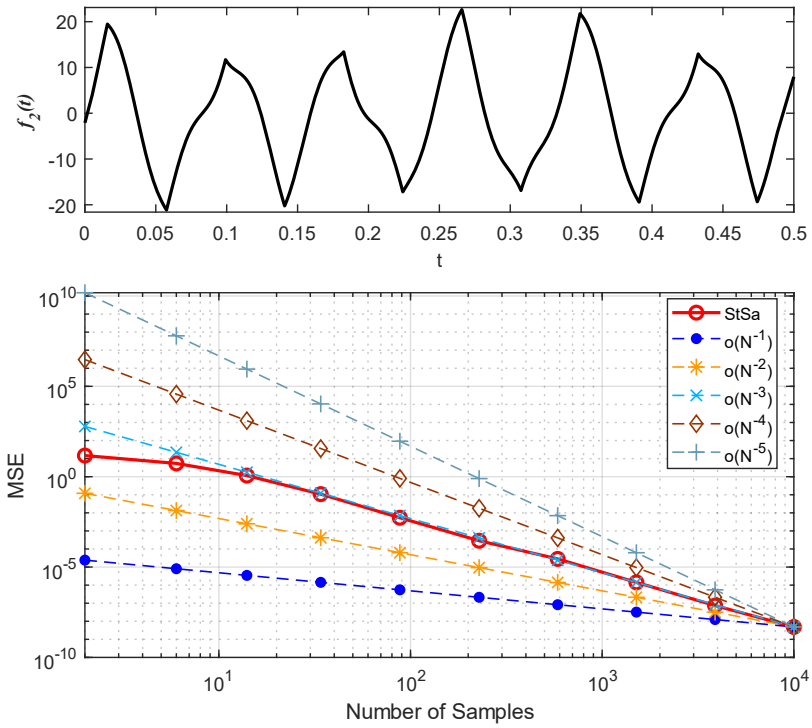


Fig. 44. The FOD of the abstract function $f_2(t)$ is piecewise-continuous. Indeed, there are 12 jumps in the FOD, but the function itself is continuous. Hence, the StSa estimator converges at is N^{-3} rate.

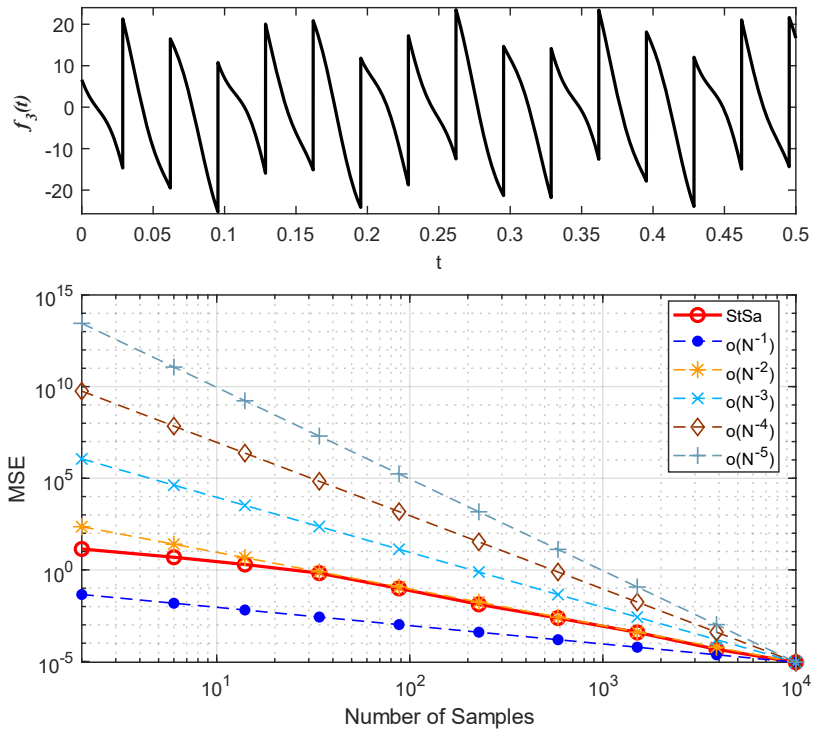


Fig. 45. Here, the function $f_3(t)$ is non-smooth, and so is its FOD. 15 ZOD jumps occur in the function itself. Consequently, the StSa estimator is converging at a slower speed of N^{-2}.

5.5.2. StSa Estimation of FIR Filter Output

Three analog input signals, $x_1(t)$, $x_2(t)$, and $x_3(t)$, need to be filtered out using a $10kHz$ bandwidth LBF filter. The first signal is smooth and infinitely differentiable. Whereas, the other two or their derivatives are non-smooth. More precisely, $x_2(t)$ has a piecewise-continuous FOD, and $x_3(t)$ has a piecewise-continuous ZOD.

The utilised LPF is similar to the analog filter designed in Chapter 3 (boxcar filter smoothed out by a Hanning window function). However, this time the filter has a cutoff frequency of $10kHz$, as shown in Fig. 46 below with the dashed blue line.

The output signals, $y_1(t)$, $y_2(t)$, and $y_3(t)$, are estimated using a StSa filter estimator. The estimates, $\hat{y}_1(t)$, $\hat{y}_2(t)$, and $\hat{y}_3(t)$ are then used again to estimate the FT of the output signals, $\hat{Y}_1(f)$, $\hat{Y}_2(f)$, and $\hat{Y}_3(f)$ respectively. This is a two-fold DASP processing. The behaviour of the estimator is examined for each input signal using different average random sampling frequencies. To observe the difference between traditional DSP approach and StSa DASP approach, Fig. 46 also shows the FT of

uniformly sampled versions of the input signals with a sampling frequency ($F_s = 131.072kHz$) fulfilling the Nyquist rate, and Fig. 47 depicts the filtered output signals using uniform sampling approach and the same sampling rate as F_s.

$$x_1(t) = sin(2\pi F_1 t) + 3 \times sin(2\pi F_2 t) + 2 \times sin(2\pi F_3 t) + 1.5 \times T \times F_4 \times sinc(2F_4 t) \times cos(2\pi F_5 t) - 1.5 \times sin(2\pi F_6 t) + 2.5 \times sin(2\pi F_7 t), \quad (5.20a)$$

$$x_2(t) = sin(2\pi F_1 t) + 3 \times sin(2\pi F_2 t) + 2 \times sin(2\pi F_3 t) + 1.5 \times T \times F_4 \times sinc(2F_4 t) \times cos(2\pi F_5 t) - 1.5 \times sin(2\pi F_6 t) + 2.5 \times sin(2\pi F_7 t) + 2 \times SAWTOOTH(1280\pi(t - 0.007),0.5), \quad (5.20b)$$

$$x_3(t) = sin(2\pi F_1 t) + 3 \times sin(2\pi F_2 t) + 2 \times sin(2\pi F_3 t) + 1.5 \times T \times F_4 \times sinc(2F_4 t) \times cos(2\pi F_5 t) - 1.5 \times sin(2\pi F_6 t) + 2.5 \times sin(2\pi F_7 t) + 2 \times SAWTOOTH(1280\pi(t - 0.007), 1), \quad (5.20c)$$

where $F_1 = 2.048kHz$, $F_2 = 8.192kHz$, $F_3 = 25.4kHz$, $F_4 = 2.5kHz$, $F_5 = 33kHz$, $F_6 = 42kHz$, $F_7 = 61.9kHz$.

Assuming no analog antialiasing pre-filtering is used, the traditional DSP filter is working fine using uniform sampling rates at least matching the required Nyquist rate. Whereas the case is different if lower sampling rate is considered. Fig. 48 shows how alias components appear in the spectra of uniformly sampled input signals at a rate of only $19.2kHz$. However, no aliasing exists when the DASP StSa filter estimator is utilised instead of DSP's with this sub-Nyquist average sampling rate, as illustrated in Fig. 49.

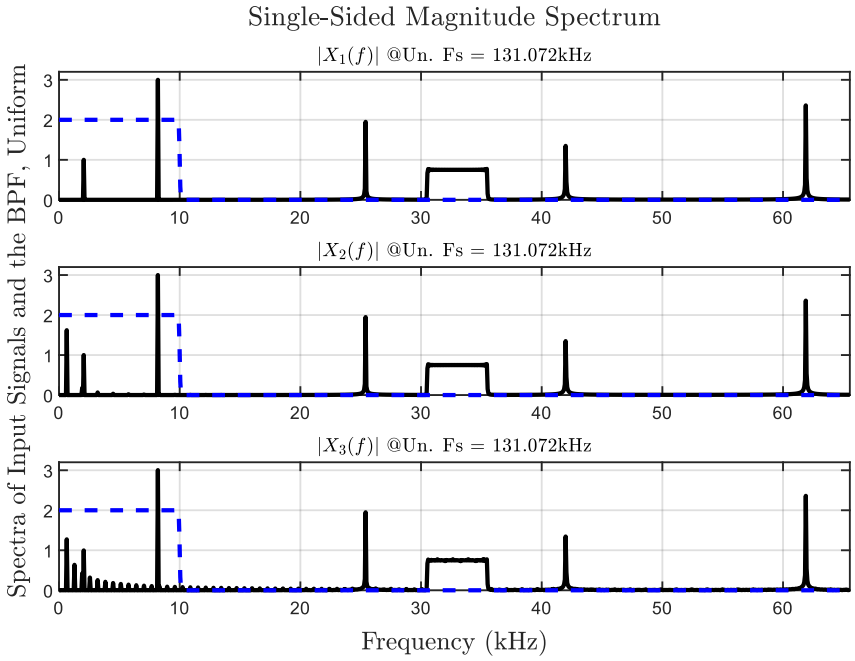


Fig. 46. Spectra of input signals (solid black) and the LPF (dashed blue) sampled uniformly at $F_s = 131.072kHz$. $x_1(t)$ is continuous and smooth, $x_2(t)$ is piecewise-continuous in FOD, and $x_3(t)$ is piecewise-continuous in ZOD. The bandwidth of the LPF is $10kHz$.

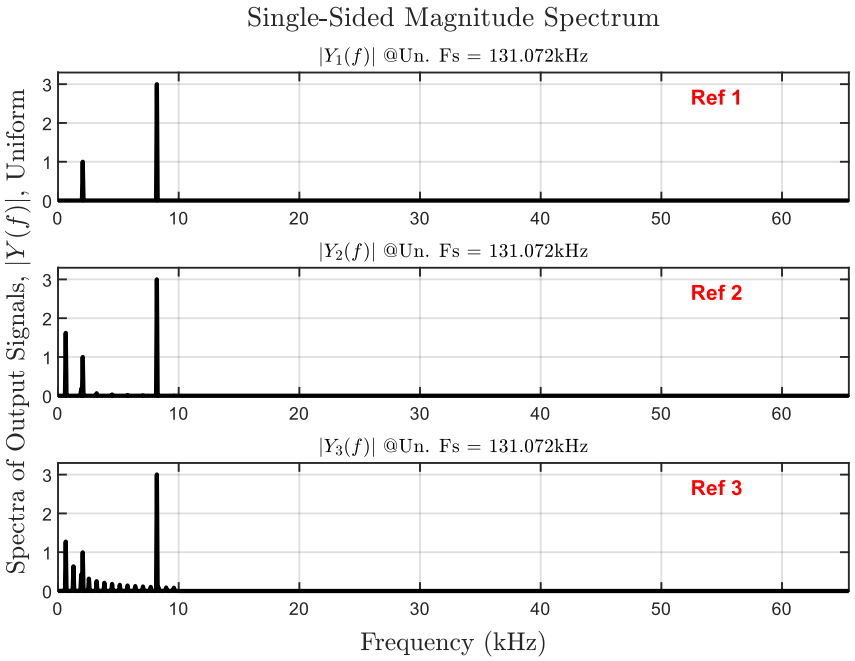


Fig. 47. Spectra of filtered output signals for the uniformly sampled input signals at $F_s = 131.072kHz$. These uniform output spectra serve as references for corresponding subsequent figures. Ref $1 = |Y_1(f)|$ is the reference spectrum for smooth signals. Ref $2 = |Y_2(f)|$ is the reference spectrum for signals with piecewise-continuous FOD, and Ref $3 = |Y_3(f)|$ is the reference spectrum for signals with piecewise-continuous ZOD.

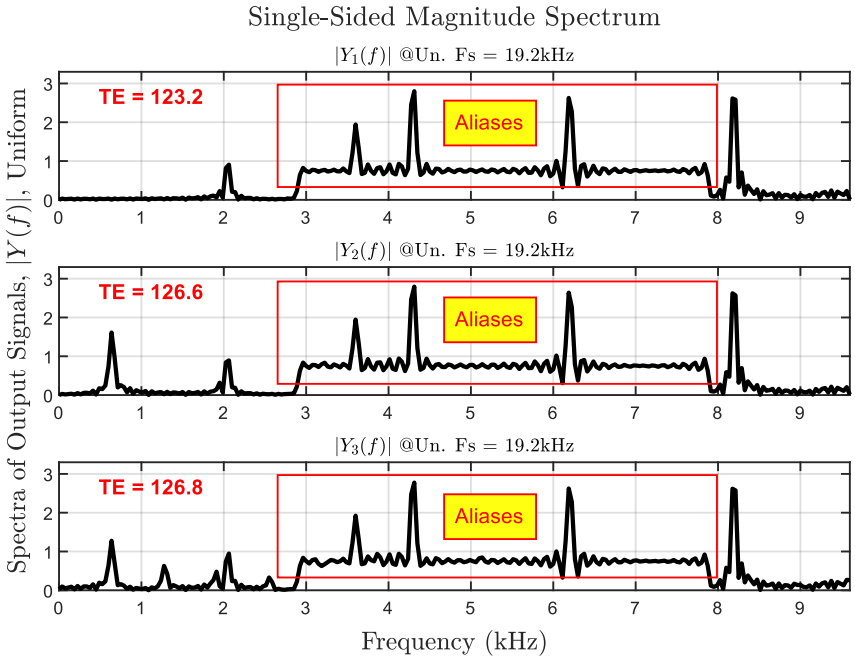


Fig. 48. Spectra of filtered output signals for the uniformly sampled input signals at $F_s = 19.2kHz$. No antialiasing prefiltering is used.

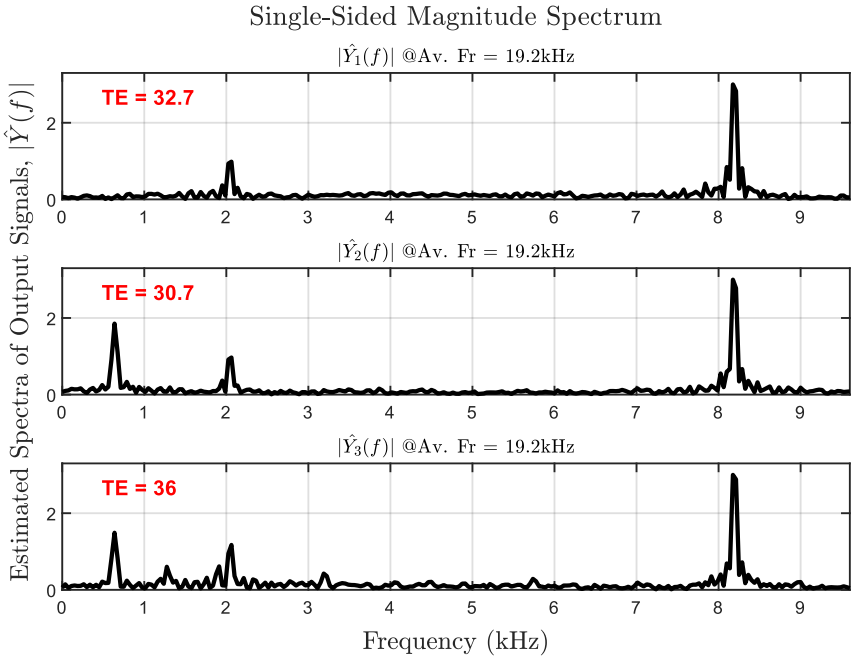


Fig. 49. Estimated spectra of output signals using DASP StSa filter estimator and an average random sampling frequency of $Av.F_r = 19.2kHz$. 100 independent MC iterations are carried out to average out the results. Remark how aliasing has been mitigated.

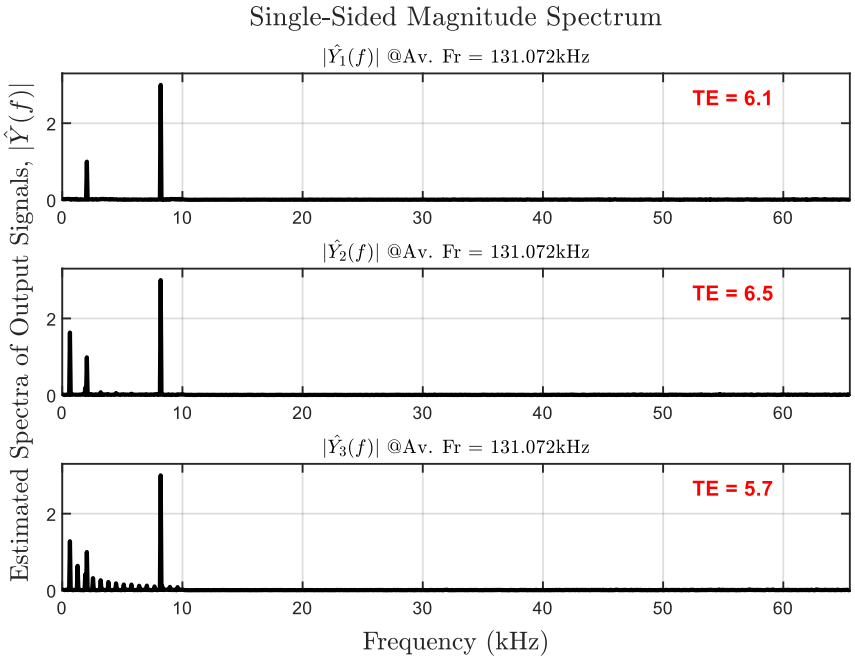


Fig. 50. Estimated spectra of output signals using DASP StSa filter estimator and an average random sampling frequency of $Av.F_r = 131.072kHz$. MC =100 independent iterations. The quantification metric TE denotes the total error between these spectra and the corresponding reference spectra shown in Fig. 47.

If the utilised average random frequency increases to match the Nyquist rate, then the results further enhance, as shown in Fig. 50. Nothing is perfect except Allah, the MC averaging for the StSa estimator plays a vital role in providing a level of guarantee that the depicted plots are not due to a single realisation of the random variables comprising the estimator. Nonetheless, the statistical errors accompanying randomisation can sometimes be large, and this is actually the main weakness of DASP in general, as indicated elsewhere in this thesis, despite there are a few methods to reduce these errors, where MC averaging is just one of them. Another method is SECOEX discussed in [13, pp. 401-403], which depends on sequential component extraction of frequency pins with high magnitude until reaching a preassigned threshold. This iterative method normally consumes more time to extract all frequency components especially if the spectrum is not sparse enough.

To spot the drawback of sampling randomisation, Fig. 51 and Fig. 52 show the spectra of StSa estimator when only one MC iteration is considered using average sampling frequencies of $Av.F_r = 19.2kHz$ and $Av.F_r = 131.072kHz$ respectively.

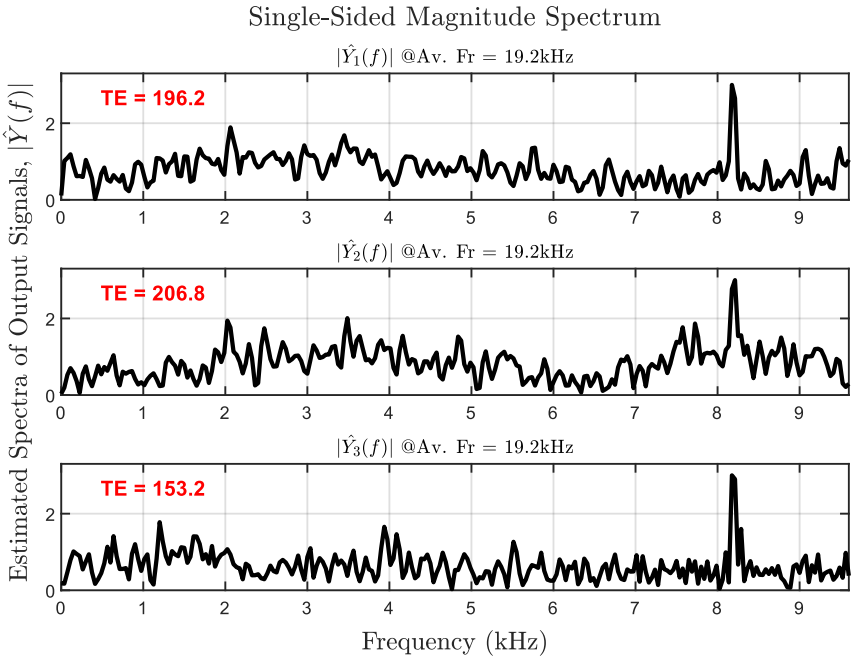


Fig. 51. Spectra of estimated filter output signals using a single realisation (MC=1) and an average random sampling frequency $Av.F_s = 19.2kHz$.

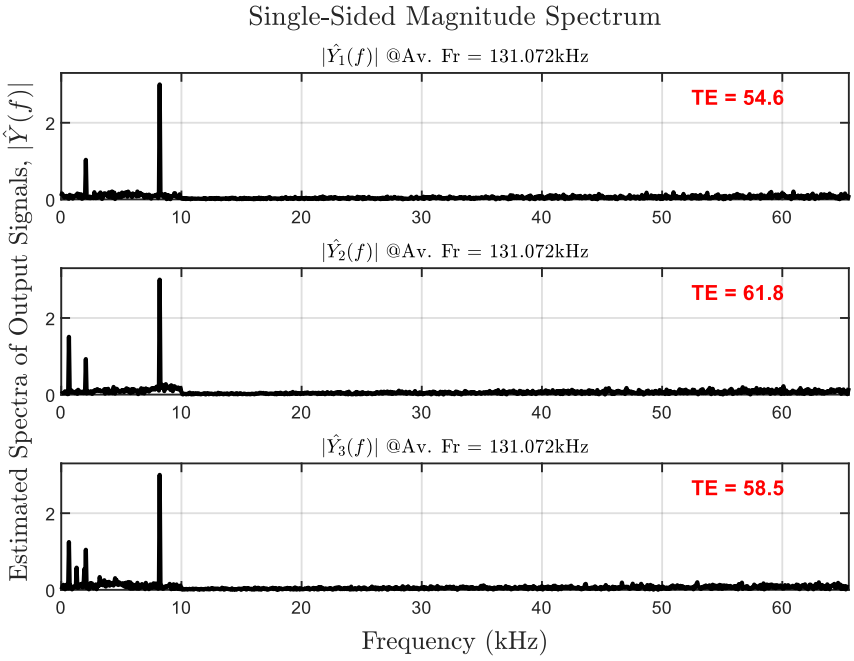


Fig. 52. Estimated spectra of filter output signals using a single realisation (MC=1) and an average random sampling frequency $Av.F_s = 131.072kHz$.

CHAPTER 6

Antithetical Stratified Estimator: Piecewise-Continuous Case

6.1. Overview

In Chapter 3, it was proved that the AnSt filter estimator is unbiased, consistent, and its rate of uniform convergence is equal to N^{-5}. All these findings were based on the assumptions that the integrand function, $f(t,\tau)$, and its first two derivatives are continuous and bounded. The assumptions were necessary for the analytical derivation of various statistical properties of the estimator since applying Taylor series expansion of the integrand function at any stratum requires that the function under consideration is continuous (i.e. smooth ZOD), and so are its FOD and SOD.

On the other hand, in many situations, the integrand function and/or its first two derivatives might not be smooth. This chapter investigates the impact of such non-smoothness of $f(t,\tau)$ on the statistical features of the AnSt filter estimator. Remark that since $f(t,\tau)$, $f'(t,\tau)$, or $f''(t,\tau)$ are now assumed to be piecewise-continuous, they cannot be directly expanded using Taylor series, and so, the analysis that has been carried out in Chapter 3 needs to be re-done again for the non-smoothness case.

This chapter revises my paper, [90], on this regard. Basically, I examine the AnSt filter estimator's variance and asymptotic behaviour in three main cases characterising the integrand function:

1. piecewise-continuous SOD;
2. piecewise-continuous FOD;
3. piecewise-continuous ZOD.

In any of the three cases above, the number of jump discontinuities are assumed to be finite and bounded.

In the next section, I quickly demonstrate how the AnSt sampling scheme works, as this is vital for addressing subsequent sections. In Section 6.3, the AnSt filter estimator in the case of discontinuities is introduced and its statistical characteristics are verified. Lastly, in Section 6.4, I validate my analytical findings and results by carrying out some simulation examples.

6.2. Revisiting AnSt Random Sampling Technique

In AnSt random sampling scheme with N strata, $2N$ signal sample points are collected within an observation window $[t-T, t]$. Two sample points per stratum are collected, I call them the antithetical sample pair (ASP).

Considering the j-th stratum, the time instant of the first point, τ_j, is randomly selected from the stratum's interval, $A_j = [S_{j-1}, S_j)$, while the time instant of the other point in the ASP is precisely the antithetical counterpart of the first one. Consequently, if C_j is the centre of the j-th stratum, then the time instants of the j-th ASP would be $\{\tau_j, \tau_j^a\} = \{\tau_j, 2C_j - \tau_j\}$, as shown in Fig. 53. The strata lengths are assumed to be equal in this chapter, as well. Meaning that the length of any stratum is T/N.

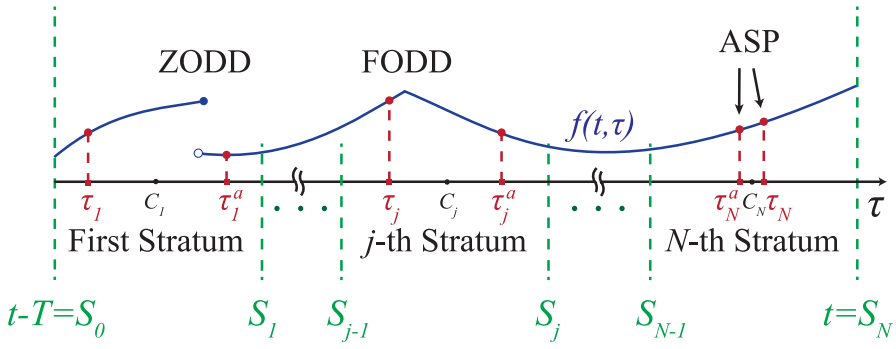


Fig. 53. AnSt sampling scheme example, where the first sampling instant in the j-th stratum, τ_j, is randomly selected from the stratum's time interval. Whereas the second sampling instant, τ_j^a, is its antithetical counterpart. Note that τ_j needs not to be less than τ_j^a.

6.3. Non-Smoothness and the AnSt Estimator

The AnSt filter estimator presented in Chapter 3 was

$$\hat{y}(t) = \frac{\Delta}{2}\sum_{j=1}^{N}\Big(f\big(t,\tau_j\big) + f\big(t,\tau_j^a\big)\Big), \tag{6.1}$$

where Δ is the stratum length ($= T/N$), and τ_j and τ_j^a are the ASP in the j-th stratum. All random variables $\{\tau_j\}_{j=1}^{N}$ are i.i.d and having a PDF

$$p_j(\tau) = \begin{cases} 1/\Delta, & \tau \in A_j \\ 0\ , & \text{elsewhere} \end{cases}. \tag{6.2}$$

with $A_j = \big[S_{j-1}, S_j\big)$ is the subinterval associated with the j-th stratum. Note that $p_j(\tau)$ is also the PDF of τ_j^a, as it is directly correlated with τ_j.

Fig. 54 depicts an exemplar j-th stratum that involves a discontinuity at time instant τ_{D_j} in both integrand function's FOD and SOD. In the analysis to follow, it is assumed that only one discontinuity would be present within any given stratum, except for the case when multiple different-order-derivative discontinuities occur at the same time instant, i.e. there might be FOD and SOD, or ZOD, FOD, and SOD at a given τ_{D_j}.

Similar to the approach in the previous chapter, I introduce two subfunctions, $f_{j,L}(t,\tau)$ and $f_{j,R}(t,\tau)$, as they appear in Fig. 54. Both subfunctions are continuous and differentiable, and therefore are Taylor series expandable about any time instant within their time span, including the discontinuity point, τ_{D_j}.

The two subfunctions, $f_{j,L}(t,\tau)$ and $f_{j,R}(t,\tau)$, are defined in (5.5) and (5.6) respectively. This make it plausible to define the integrand function, $f(t,\tau)$, across the whole observation window as

$$f(t,\tau) = \sum_{j=1}^{N}\begin{cases} f_{j,L}(t,\tau), & \tau \in A_{j,L} \\ f_{j,R}(t,\tau), & \tau \in A_{j,R} \end{cases}. \tag{6.3}$$

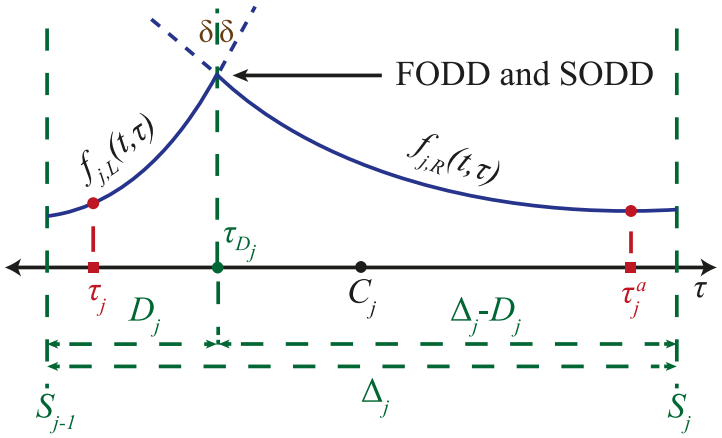


Fig. 54. An exemplar of a j-th stratum with a discontinuity in both FOD and SOD of $f(t,\tau)$. $f_{j,L}(t,\tau)$ and $f_{j,R}(t,\tau)$ subfunctions and their respective first two derivatives are smooth and differentiable.

The impact of non-smooth integrand function, with discontinuities in its ZOD, FOD, or SOD, is to be investigated in this section. So, let us begin with the following definitions for the magnitude of the jumps in different order of the derivative of $f(t,\tau)$ which are going to be used frequently in my analytical derivation below,

$$F_{2j} := f''_{j,L}\left(t,\tau_{D_j}\right) - f''_{j,R}\left(t,\tau_{D_j}\right), \tag{6.4a}$$

$$F_{1j} := f'_{j,L}\left(t,\tau_{D_j}\right) - f'_{j,R}\left(t,\tau_{D_j}\right), \tag{6.4b}$$

$$F_{0j} := f_{j,L}\left(t,\tau_{D_j}\right) - f_{j,R}\left(t,\tau_{D_j}\right). \tag{6.4c}$$

6.3.1. Non-smooth SOD

If $f''(t,\tau)$ involves M bounded discontinuities whereas both $f'(t,\tau)$ and $f(t,\tau)$ do not have any discontinuity across the observation window, and if the discontinuities appear at time instants $\{\tau_{D_j}\}_{j=1}^{M}$, then

$$f_{j,L}\left(t,\tau_{D_j}\right) = f_{j,R}\left(t,\tau_{D_j}\right) = f\left(t,\tau_{D_j}\right), \tag{6.5a}$$

$$f'_{j,L}\left(t,\tau_{D_j}\right)=f'_{j,R}\left(t,\tau_{D_j}\right)=f'\left(t,\tau_{D_j}\right), \tag{6.5b}$$

$$f''_{j,L}\left(t,\tau_{D_j}\right)\neq f''_{j,R}\left(t,\tau_{D_j}\right). \tag{6.5c}$$

Equivalently, I have $F_{0j}=F_{1j}=0$, but $F_{2j}\neq 0$.

To validate the consistency of the AnSt estimator, (6.1), for non-smooth SOD of the integrand function, it is essential to examine the statistical behaviour of the estimator as the sample size increases towards infinity. To check this behaviour, I start by defining the j-th sub-estimator, ϕ_j, as

$$\phi_j := \frac{1}{2}\Big(f\big(t,\tau_j\big)+f\big(t,\tau_j^a\big)\Big)\,\Delta \tag{6.6a}$$

$$\phi_j=\frac{\Delta}{2}\begin{Bmatrix} f_{j,L}\big(t,\tau_j\big), & \tau_j\in A_{j,L}\\ f_{j,R}\big(t,\tau_j\big), & \tau_j\in A_{j,R}\end{Bmatrix}+\frac{\Delta}{2}\begin{Bmatrix} f_{j,L}\big(t,\tau_j^a\big), & \tau_j^a\in A_{j,L}\\ f_{j,R}\big(t,\tau_j^a\big), & \tau_j^a\in A_{j,R}\end{Bmatrix} \tag{6.6b}$$

$$\phi_j=\frac{\Delta}{2}\begin{Bmatrix} f_{j,L}\big(t,\tau_j\big)+f_{j,L}\big(t,\tau_j^a\big), & \tau_j,\tau_j^a\in A_{j,L}\\ f_{j,L}\big(t,\tau_j\big)+f_{j,R}\big(t,\tau_j^a\big), & \tau_j\in A_{j,L}\,,\tau_j^a\in A_{j,R}\\ f_{j,R}\big(t,\tau_j\big)+f_{j,L}\big(t,\tau_j^a\big), & \tau_j\in A_{j,R}\,,\tau_j^a\in A_{j,L}\\ f_{j,R}\big(t,\tau_j\big)+f_{j,R}\big(t,\tau_j^a\big), & \tau_j\,,\tau_j^a\in A_{j,R}\end{Bmatrix} \tag{6.6c}$$

Now, I calculate the expected value of ϕ_j. Thus

$$E\big[\phi_j\big]=\frac{1}{2}\int_{-\infty}^{\infty}p_j(\tau)\big(f(t,\tau)+f(t,\tau^a)\big)\Delta\,d\tau, \tag{6.7a}$$

$$=\frac{1}{2}\int_{S_{j-1}}^{S_j}f(t,\tau)d\tau+\frac{1}{2}\int_{S_{j-1}}^{S_j}f(t,\tau^a)d\tau, \tag{6.7b}$$

$$= \frac{1}{2}\int_{A_{j,L}} f_{j,L}(t,\tau)d\tau + \frac{1}{2}\int_{A_{j,R}} f_{j,R}(t,\tau)d\tau + \frac{1}{2}\int_{A_{j,L}} f_{j,L}(t,\tau^a)d\tau^a + \frac{1}{2}\int_{A_{j,R}} f_{j,R}(t,\tau^a)d\tau^a. \tag{6.7c}$$

Since the subfunctions $f_{j,L}(t,\tau)$, $f_{j,R}(t,\tau)$, $f_{j,L}(t,\tau^a)$ and $f_{j,R}(t,\tau^a)$ are all assumed to be smooth and differentiable, I can expand them about $\tau = \tau_{D_j}$ using Taylor. Thus, (6.7c) can be rewritten as

$$\begin{aligned} E[\phi_j] = {} & \frac{1}{2}\int_{A_{j,L}} \left(f_{j,L}\left(t,\tau_{D_j}\right) + \left(\tau-\tau_{D_j}\right) f'_{j,L}\left(t,\tau_{D_j}\right) \right. \\ & \left. + \frac{1}{2}\left(\tau-\tau_{D_j}\right)^2 f''_{j,L}\left(t,\tau_{D_j}\right) + o\left(\left|\tau-\tau_{D_j}\right|^2\right)\right) d\tau \\ & + \frac{1}{2}\int_{A_{j,R}} \left(f_{j,R}\left(t,\tau_{D_j}\right) + \left(\tau-\tau_{D_j}\right) f'_{j,R}\left(t,\tau_{D_j}\right) \right. \\ & \left. + \frac{1}{2}\left(\tau-\tau_{D_j}\right)^2 f''_{j,R}\left(t,\tau_{D_j}\right) + o\left(\left|\tau-\tau_{D_j}\right|^2\right)\right) d\tau \\ & + \frac{1}{2}\int_{A_{j,L}} \left(f_{j,L}\left(t,\tau_{D_j}\right) + \left(\tau^a-\tau_{D_j}\right) f'_{j,L}\left(t,\tau_{D_j}\right) \right. \\ & \left. + \frac{1}{2}\left(\tau^a-\tau_{D_j}\right)^2 f''_{j,L}\left(t,\tau_{D_j}\right) + o\left(\left|\tau^a-\tau_{D_j}\right|^2\right)\right) d\tau^a \\ & + \frac{1}{2}\int_{A_{j,R}} \left(f_{j,R}\left(t,\tau_{D_j}\right) + \left(\tau^a-\tau_{D_j}\right) f'_{j,R}\left(t,\tau_{D_j}\right) \right. \\ & \left. + \frac{1}{2}\left(\tau^a-\tau_{D_j}\right)^2 f''_{j,R}\left(t,\tau_{D_j}\right) + o\left(\left|\tau^a-\tau_{D_j}\right|^2\right)\right) d\tau^a, \end{aligned} \tag{6.8}$$

By calculating the integrals in (6.8) while considering (6.5a-c), we get

$$\begin{aligned} E[\phi_j] = {} & \frac{1}{6}\left(K_j^3 f''_{j,L}\left(t,\tau_{D_j}\right) - \left(K_j^3 - 3K_j^2 + 3K_j - 1\right) f''_{j,R}\left(t,\tau_{D_j}\right)\right)\Delta^3 \\ & + \frac{1}{2}\left(1-2K_j\right) f'_{j,R}\left(t,\tau_{D_j}\right)\Delta^2 + f_{j,R}\left(t,\tau_{D_j}\right)\Delta + o(\Delta^3). \end{aligned} \tag{6.9}$$

All constituents of (6.9) are real-valued, either by their nature or by assumption. Hence, the error of estimation related to the j-th stratum, e_j, can be found by subtracting (6.9) from (6.6b), so

$$e_j = \phi_j - E[\phi_j], \tag{6.10a}$$

$$\begin{aligned} e_j = & \frac{\Delta}{2}\begin{Bmatrix} f_{j,L}(t,\tau_j), & \tau_j \in A_{j,L} \\ f_{j,R}(t,\tau_j), & \tau_j \in A_{j,R} \end{Bmatrix} + \frac{\Delta}{2}\begin{Bmatrix} f_{j,L}(t,\tau_j^a), & \tau_j^a \in A_{j,L} \\ f_{j,R}(t,\tau_j^a), & \tau_j^a \in A_{j,R} \end{Bmatrix} \\ & - \left(\frac{1}{6}\left(K_j^3 f''_{j,L}\left(t,\tau_{D_j}\right) - c_{1j} f''_{j,R}\left(t,\tau_{D_j}\right)\right)\Delta^3 + c_{2j} f'_{j,R}\left(t,\tau_{D_j}\right)\Delta^2 \right. \\ & \left. + f_{j,R}\left(t,\tau_{D_j}\right)\Delta\right) + o(\Delta^3), \end{aligned} \tag{6.10b}$$

where c_{1j} and c_{2j} are $c_{1j} = K_j^3 - 3K_j^2 + 3K_j - 1$ and $c_{2j} = (1 - 2K_j)/2$.

The error term in (6.10b) may have different values according to the time-instant of the discontinuity instant ($= \tau_{D_j}$) with respect to the time-instants of the j-th stratum's ASP, i.e. τ_j and τ_j^a. Apart from the centre and end points of the stratum's interval, e_j has four possible results, as indicated by (6.6c). Nevertheless, they are almost alike. Their differences don't affect the convergence rate of the estimator as all of them characterise identical degree polynomial of Δ (the critical variable to determine the speed of estimator convergence.) The core differences between the error polynomials are in their constant coefficients. To avoid repetition, only one case will be considered here, namely, when the discontinuity happens anywhere between the two time-instants of the ASP, that is when $\tau_j < \tau_{D_j} < \tau_j^a$, i.e. $\tau_j \in A_{j,L}$, $\tau_j^a \in A_{j,R}$, so

$$\begin{aligned} e_j = & \frac{1}{4}\left(f''_{j,L}\left(t,\tau_{D_j}\right) + f''_{j,R}\left(t,\tau_{D_j}\right)\right)\left(\tau_j - \tau_{D_j}\right)^2 \Delta \\ & - c_{2j}\, f''_{j,R}\left(t,\tau_{D_j}\right)\left(\tau_j - \tau_{D_j}\right)\Delta^2 \\ & + \frac{1}{12}\left(c_{3j} f''_{j,R}\left(t,\tau_{D_j}\right) - 2K_j^3 f''_{j,L}\left(t,\tau_{D_j}\right)\right)\Delta^3 + o(\Delta^3), \end{aligned} \tag{6.11}$$

where $c_{3j} = 1 - 6K_j + 6K_j^2 + 2K_j^3$.

By calculating the second moment of the error term in the j-th stratum, e_j, I obtain the value of the variance associated with the j-th sub-estimator, $\mathbb{V}[\phi_j]$,

$$\mathbb{V}[\phi_j] = \int_{-\infty}^{\infty} p_j(\tau)(e_j)^2 d\tau. \tag{6.12}$$

The following theorem can now be established for the whole variance of the AnSt filter estimator. Note that the proof of this theorem is provided in Appendix C.

Theorem 6.1. *Assume that the integrand function, $f(t,\tau)$, and its FOD, $f'(t,\tau)$, are continuous, real-valued, and bounded over an observation interval, $[t-T,T]$, whereas the SOD, $f''(t,\tau)$, is piecewise-continuous and involves a limited number of bounded discontinuities, M, within the whole observation window. Then,*

3. *the variance of the AnSt filter estimator, $\mathbb{V}[\hat{y}(t)]$, using $2N$ sample points is*

$$\mathbb{V}[\hat{y}(t)] = \frac{T^6}{720(\mathrm{N})^6}\sum_{j\in I_M}\left(c_{4j}{F_{2j}}^2 + f''_{j,L}\left(t,\tau_{D_j}\right)f''_{j,R}\left(t,\tau_{D_j}\right)\right) + \frac{T^5}{720(\mathrm{N})^5}\sum_{\substack{j=1\\ j\notin I_M}}^{N}\left(\left(f''\left(t,C_j\right)\right)^2\Delta\right) + o(N^{-5}), \tag{6.13a}$$

where $I_M = \{i_1, i_2, i_3, \dots, i_M\}$ is a set of indices for the M strata with discontinuities and $c_{4j} = 9 - 45K_j + 90K_j^2 - 110K_j^3 + 105K_j^4 - 60K_j^5 + 20K_j^6$, and

4. *the uniform convergence rate of the AnSt filter estimator is exactly N^{-5} and satisfies*

$$\lim_{N\to\infty}((2N)^5 \times \mathbb{V}[\hat{y}(t)]) = \frac{2T^5}{45}\sum_{k=1}^{M+1}\int_{T_{k-1}}^{T_k}\left(f''(t,\tau)\right)^2 d\tau, \tag{6.13b}$$

where $\{T_k\}_{k=1}^{M}$ is a set of time instants at which $f''(t,\tau)$ has jump discontinuities, whereas $T_0 := t - T$ and $T_{M+1} := t$.

If $M = 0$, which means $f''(t,\tau)$ is also smooth and has no discontinuities at all, then it is obvious from (6.13b) that the estimator is still converging at a speed of N^{-5}, and (6.13b) will simplify to (6.14), which is equal to (3.30b), the convergence value of the AnSt filter estimator in the continuous case, of discussed in Chapter 3,

$$\lim_{N\to\infty}((2N)^5 \times \mathbb{V}[\hat{y}(t)]) = \frac{2T^5}{45}\int_{t-T}^{t}\left(f''(t,\tau)\right)^2 d\tau. \tag{6.14}$$

According to the results of Theorem 6.1, the AnSt estimator converges at the same rate regardless of the smoothness or non-smoothness of the integrand function's SOD. This very conclusion is stronger than the AnSt FT estimator discussed in [53], which basically emphasised that the integrand function should be smooth for the analytical derivation of the AnSt FT estimator's statistical features to be valid.

6.3.2. Non-smooth FOD

If the FOD of the integrand function is piecewise-continuous (and implicitly the SOD, as well) but the ZOD (= the integrand function itself) is continuous, and if the discontinuities in the FOD and SOD occur exactly at some τ_{D_j} time instants, then we have

$$f_{j,L}\left(t,\tau_{D_j}\right) = f_{j,R}\left(t,\tau_{D_j}\right) = f\left(t,\tau_{D_j}\right), \tag{6.15a}$$

$$f'_{j,L}\left(t,\tau_{D_j}\right) \neq f'_{j,R}\left(t,\tau_{D_j}\right), \tag{6.15b}$$

$$f''_{j,L}\left(t,\tau_{D_j}\right) \neq f''_{j,R}\left(t,\tau_{D_j}\right). \tag{6.15c}$$

The variance of the AnSt filter estimator in this case is adversely affected by such non-smoothness of the integrand function. This is clearly seen in the new variance and convergence rate of the estimator. Theorem 6.2 concludes the changes and provides new mathematical expressions for these two statistical features of the AnSt estimator. For the proof, reader is referred to Appendix D.

Theorem 6.2. *Assume that the integrand function, $f(t,\tau)$, is smooth and real-valued over an observation interval $[t-T,t)$, whereas its first two derivatives are piecewise-continuous and bounded. Assume also that the number of discontinuities in both the FOD and SOD is finite and is equal to M.Then,*

1. *the variance of the AnSt filter estimator, $\mathbb{V}[\hat{y}(t)]$, using $2N$ sample points is*

$$\mathbb{V}[\hat{y}(t)] = \frac{\Delta^4}{12} \sum_{j \in I_M} \left(c_{5j} {F_{1j}}^2 \right) + \frac{\Delta^5}{720} \sum_{\substack{j=1 \\ j \notin I_M}}^{N} \left(\left(f''(t, C_j) \right)^2 \Delta \right) + o(\Delta^5) \,, \tag{6.16a}$$

where $c_{5j} = 1 - 3\,K_j + 6\,K_j^2 - 6\,K_j^3 + 3\,K_j^4$, *and*

2. the uniform convergence rate is exactly N^{-4} *and satisfies*

$$\lim_{N \to \infty} ((2N)^4 \times \mathbb{V}[\hat{y}(t)]) = \frac{4T^4}{3} \sum_{j \in I_M} c_{5j}\, {F_{1j}}^2, \tag{6.16b}$$

Remark that the Δ high-power terms of the variance in (6.16a) can be ignored compared to Δ^{-4} as $\Delta \to 0$. Therefore, additional simplification of (6.16a) can be obtained as

$$\mathbb{V}[\hat{y}(t)] = \frac{T^4}{12N^4} \sum_{j \in I_M} \left(c_{5j}\, {F_{1j}}^2 \right) + o(N^{-4}). \tag{6.17}$$

Moreover, it is rational to get a faster convergence rate of N^{-5} for the variance in (6.16a) when all F_{1j} values are equal to zero, i.e. smooth FOD, which matches the speed of convergence of AnSt estimator in the previous sub-section.

6.3.3. Piecewise-Continuous ZOD

If the integrand function itself is non-smooth and has a limited and bounded discontinuities at M time instants, $\{\tau_{D_j}\}_{j=1}^{M}$, then

$$f_{j,L}\left(t, \tau_{D_j}\right) \neq f_{j,R}\left(t, \tau_{D_j}\right), \tag{6.18a}$$

$$f'_{j,L}\left(t, \tau_{D_j}\right) \neq f'_{j,R}\left(t, \tau_{D_j}\right), \tag{6.18b}$$

$$f''_{j,L}\left(t, \tau_{D_j}\right) \neq f''_{j,R}\left(t, \tau_{D_j}\right). \tag{6.18c}$$

Considering the new ZOD jump discontinuities equations as given in (6.18a-c), analogous investigation to the previous sub-section is carried out here. To avoid

repetition, I conclude the final findings of the analysis in Theorem 6.3. Appendix E presents the proof of this theorem.

Theorem 6.3. *Assume there are M bounded ZOD discontinuities in $f(t,\tau)$ within an observation interval, where M is a finite integer. Suppose that the set of integers $I_M = \{i_1, i_2, i_3, \dots, i_M\}$ represents the strata indices where such discontinuities happen. Then,*

1. the variance of AnSt filter estimator utilising $2N$ sample points is

$$\begin{aligned}\mathbb{V}[\hat{y}(t)] \;=\; & \sum_{j\in I_M}\Big(c_{2j}{}^2\,F_{0j}{}^2\,\Delta^2 + \frac{1}{2}c_{2j}\,c_{7j}\,F_{1j}F_{0j}\Delta^3 \\ & + \frac{1}{12}\big(c_{5j}\,F_{1j}{}^2 + 4c_{2j}{}^2c_{8j}\,F_{2j}F_{0j}\big)\Delta^4 + \frac{1}{48}c_{6j}\,F_{2j}F_{1j}\Delta^5 \\ & + \frac{1}{720}c_{4j}\,F_{2j}{}^2\,\Delta^6 + \frac{1}{720}f''_{j,L}\left(t,\tau_{D_j}\right)f''_{j,R}\left(t,\tau_{D_j}\right)\Delta^6\Big) \\ & + \sum_{\substack{j=1\\ j\notin I_M}}^{N}\left(\frac{1}{720}\Big(f''\big(t,C_j\big)\Big)^2\Delta^6\right) + o(\Delta^5).\end{aligned} \tag{6.19a}$$

where $c_{7j} = 1 - 2K_j + 2K_j^2$ and $c_{8j} = 1 - K_j + K_j^2$, and

2. the AnSt filter estimator converges uniformly at an exact rate of N^{-2} and satisfies

$$\lim_{N\to\infty}\left((2N)^2 \times \mathbb{V}[\hat{y}(t)]\right) = 16T^2\sum_{j\in I_M}\left(c_{2j}{}^2\,F_{0j}{}^2\right), \tag{6.19b}$$

The results of Theorem 6.3 emphasise my findings in the previous sub-sections. Namely, if $\{F_{0j} = 0\}_{j\in I_M}$, that is, the integrand function's ZOD discontinuities don't exist at all, then (6.19a) will exactly reduce to (6.16a). Therefore, the AnSt estimator will be converging precisely at N^{-4} rate. Whereas, if $\{F_{0j} = F_{1j} = 0\}_{j\in I_M}$, then (6.19a) shows that the convergence rate of AnSt estimator will be N^{-5}.

6.4. Numerical and Simulation Examples

AnSt estimation of non-smooth functions is examined in the next sub-section, where theoretical abstract functions are numerically integrated to calculate AUC within a definite interval. The samples of the integrated functions are randomly selected according to AnSt sampling technique. Then, a comparison is held between the estimated AUC and the actual AUC based on the definite integral of the function of interest. Variance versus number of samples are then plotted to demonstrate the uniform convergence rate of the AnSt estimator and when this asymptotic behaviour starts to happen.

Almost similar examples are carried out again, but this time the piecewise-continuous functions are set as input signals to a BPF. The output of the filter is estimated using AnSt approach and then compared to a uniform version of the output signal when the Nyquist rate is respected.

In both sets of examples, the simulation results, as will be seen shortly, emphasise on my analytical findings.

6.4.1. AnSt Estimation of Abstract Functions

The following set of functions have been randomly sampled based on the AnSt sampling technique within an observation window $[0,0.4)$ sec,

$$f_1(t) = -3.5\,sin(42\pi t) + 2.9\,cos(35\pi t), \tag{6.20a}$$

$$\begin{aligned} f_2(t) = {} & f_1(t) + 60 \times (t - 0.1) \times |t - 0.1| - 90 \times (t - 0.27) \times |t - 0.27| \\ & + 180 \times (t - 0.35) \times |t - 0.35|, \end{aligned} \tag{6.20b}$$

$$f_3(t) = -9\,cos(6.3\pi t) + 20 \times SAWTOOTH(30\pi(t - 0.1),0.4), \tag{6.20c}$$

$$f_4(t) = -9\,cos(63\pi t) + 20 \times SAWTOOTH(45\pi(t - 0.21), 1), \tag{6.20d}$$

where $SAWTOOTH(g(t), m)$ is the MATLAB built-in *sawtooth* function, with more details are given in the sub-section 5.4.1 of previous chapter. Note that the smoothness status of the above four functions are given in Table V below.

TABLE V: SMOOTHNESS STATUS OF INTEGRATED FUNCTIONS

Function	Smoothness Status	Discontinuities
$f_1(t)$	ZOD, FOD, and SOD are all smooth	None
$f_2(t)$	ZOD and FOD are smooth, but SOD is piecewise-continuous.	3 SODDs
$f_3(t)$	ZOD is smooth, but FOD and SOD are piecewise-continuous.	12 FODDs
$f_4(t)$	ZOD, FOD, and SOD are all non-smooth	9 ZODDs

After estimating the numerical definite integrals of the given functions using AnSt random sampling scheme and calculating the mean-squared error (MSE = variance since the estimator is unbiased), I get the results illustrated in Figs. 55-58, where 100 independent Monte Carlo simulations is conducted per each figure to guarantee it is not a result of only one specific realisation of the randomly sampled function.

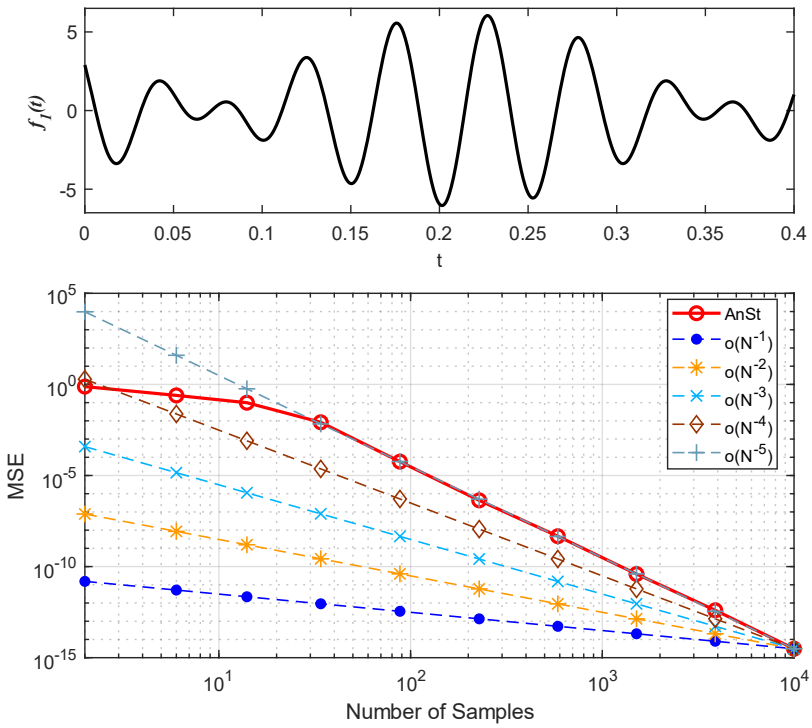


Fig. 55. The abstract function $f_1(t)$ is smooth, and there are no ZOD, FOD, and SOD discontinuities at all. The uniform convergence rate of AnSt estimator is N^{-5}.

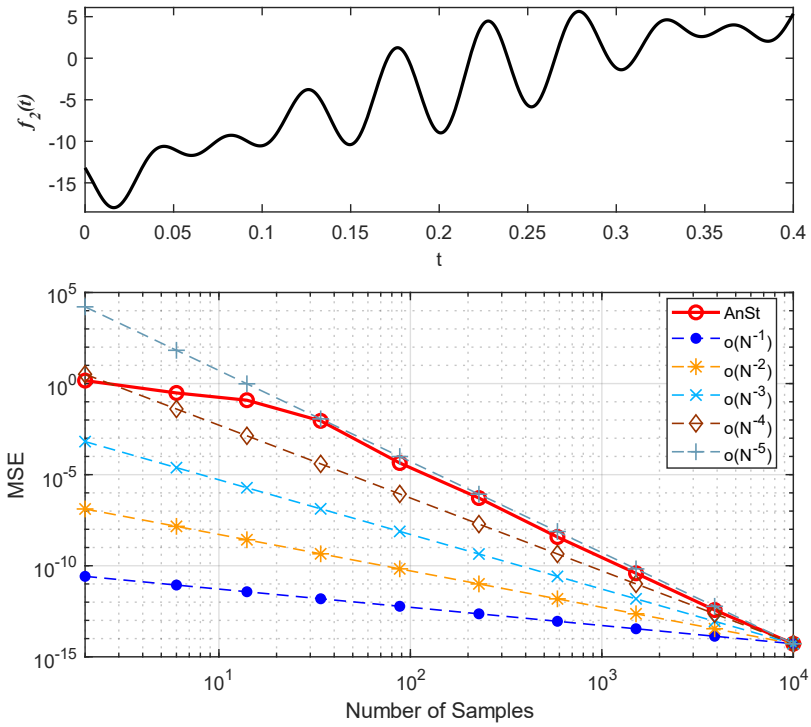


Fig. 56. The abstract function $f_2(t)$ and its FOD are continuous, whereas the SOD is piecewise-continuous with three jumps at $t = 0.1, 0.27,$ and 0.35 sec. The AnSt estimator is still converging at N^{-5} rate.

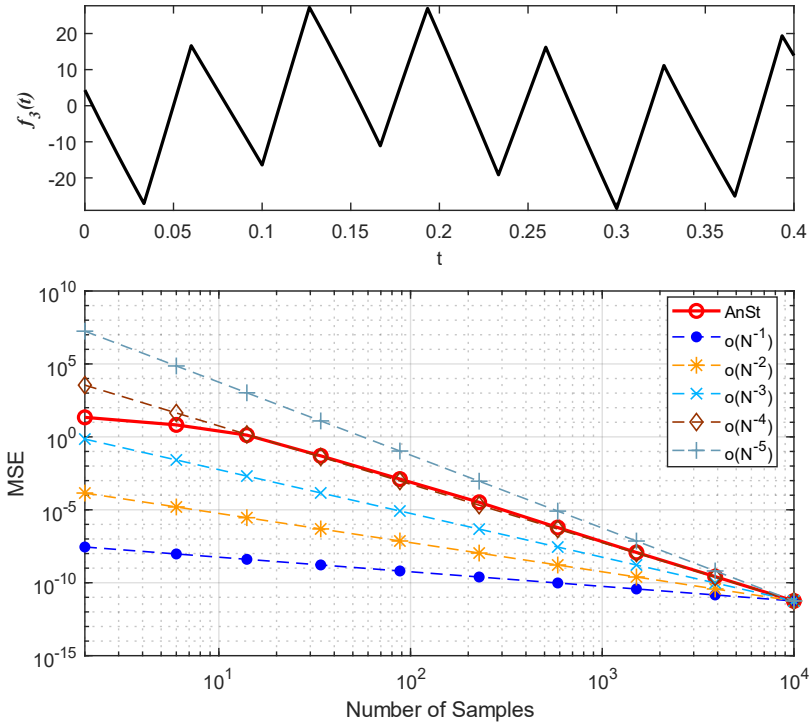


Fig. 57. The FOD of the abstract function $f_3(t)$ is piecewise-continuous. Indeed, there are twelve jumps in the FOD, but the function itself is continuous. Hence, the AnSt estimator convergence rate is N^{-4}.

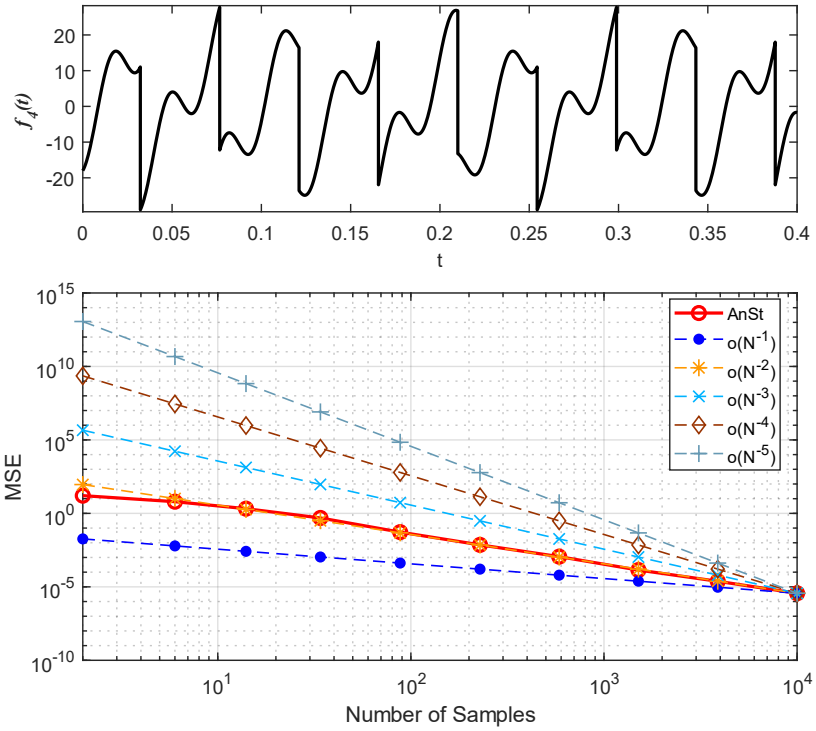


Fig. 58. Here, the function $f_4(t)$ is non-smooth, and so are its FOD and SOD. Nine ZOD jumps occur in the function itself. Consequently, the AnSt estimator is converging at its slowest speed, i.e. N^{-2}.

As can be clearly seen, AnSt estimator uniform convergence rates are N^{-5}, N^{-5}, N^{-4}, and N^{-2} for the respective abstract functions $f_1(t)$, $f_2(t)$, $f_3(t)$, and $f_4(t)$. These rates are exactly what I expect as per my analytical expressions devised in the previous three sub-sections.

6.4.2. FIR AnSt Filter Estimation

Similar to the BPF estimation example in the sub-section 3.10.2, but this time the input analog signal, $x(t)$, has been changed to include non-smooth functions. For each case of non-smoothness discussed above (i.e. piecewise-continuous SOD, piecewise-continuous FOD, piecewise-continuous ZOD), I examine the behaviour of the AnSt filter estimator using different average random sampling frequencies. Therefore, I consider the following three input signals ($x_2(t)$, $x_3(t)$, and $x_4(t)$) for the three non-smoothness cases respectively, together with a smooth signal, $x_1(t)$, just for comparison purposes,

$$x_1(t) = sin(2\pi F_1 t) + 3 \times sin(2\pi F_2 t) + 2 \times sin(2\pi F_3 t) + 1.5 \times T \times F_4 \times sinc(2F_4 t) \times cos(2\pi F_5 t) - 1.5 \times sin(2\pi F_6 t) + 2.5 \times sin(2\pi F_7 t), \quad (6.21a)$$

$$x_1(t) = sin(2\pi F_1 t) + 3 \times sin(2\pi F_2 t) + 2 \times sin(2\pi F_3 t) + 1.5 \times T \times F_4 \times sinc(2F_4 t) \times cos(2\pi F_5 t) - 1.5 \times sin(2\pi F_6 t) + 2.5 \times sin(2\pi F_7 t) + (t - 0.002) \times |t - 0.002| - 1.5 \times (t - 0.005) \times |t - 0.005| + 3 \times (t - 0.013) \times |t - 0.013|, \quad (6.21b)$$

$$x_2(t) = sin(2\pi F_1 t) + 3 \times sin(2\pi F_2 t) + 2 \times sin(2\pi F_3 t) + 1.5 \times T \times F_4 \times sinc(2F_4 t) \times cos(2\pi F_5 t) - 1.5 \times sin(2\pi F_6 t) + 2.5 \times sin(2\pi F_7 t) + 2 \times SAWTOOTH(1280\pi(t - 0.007),0.5), \quad (6.21c)$$

$$x_3(t) = sin(2\pi F_1 t) + 3 \times sin(2\pi F_2 t) + 2 \times sin(2\pi F_3 t) + 1.5 \times T \times F_4 \times sinc(2F_4 t) \times cos(2\pi F_5 t) - 1.5 \times sin(2\pi F_6 t) + 2.5 \times sin(2\pi F_7 t) + 2 \times SAWTOOTH(1280\pi(t - 0.007), 1), \quad (6.21d)$$

where $F_1 = 2.048kHz$, $F_2 = 8.192kHz$, $F_3 = 25.4kHz$, $F_4 = 2.5kHz$, $F_5 = 33kHz$, $F_6 = 42kHz$, $F_7 = 61.9kHz$.

The bandwidth of the BPF is $22kHz$ and its centre frequency is $33kHz$. Fig. 59 depicts the spectra of input signals $X_1(f)$, …, $X_4(f)$ uniformly sampled at the requested Nyquist rate (i.e. $F_s = 131.072kHz$). The frequency response of the BPF is also included in dashed blue line. The spectra of the corresponding filter output signals using uniform sampling are shown in Fig. 60. This is included here to be compared with the randomly estimated one to follow.

Under the assumption of no antialiasing prefiltering is available, that conventional DSP filtering works well only if the utilised uniform sampling rate is not less than the Nyquist rate, which is equal to $131.072kHz$ in this example. Whereas no such restriction on the sampling rate is imposed if random sampling approaches are used, instead. Fig. 61 shows how alias components appear in the spectra of uniformly sampled input signals at a rate of only $89.6kHz$. However, no aliasing exists when the

DASP AnSt filter estimator is utilised instead of DSP's, even with this sub-Nyquist average sampling rate, as illustrated in Fig. 62. While Fig. 63 shows an enhanced estimate of filter output since the utilised average random sampling rate is now identical to the Nyquist rate.

On the other hand, the statistical errors yielded as a result of randomisation can sometimes be significant. To spot this drawback, I have conducted two further simulation examples using AnSt filter estimator and considering only one MC iteration. In the first example, I used an average sampling frequency of $Av.F_r = 89.6kHz$, and in the second one, the average sampling rate was $Av.F_r = 131.072kHz$. Fig. 64 and Fig. 65 depicts the spectra of the estimated output signals for the two examples, respectively.

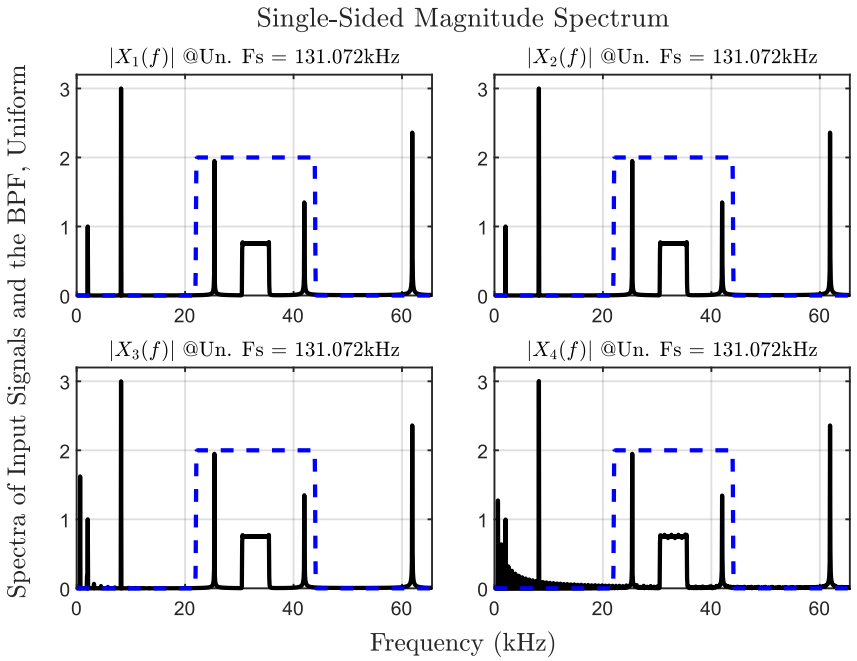


Fig. 59. Spectra of input signals (solid black) and the BPF (dashed blue) sampled uniformly at $F_s = 131.072kHz$. $x_1(t)$ is continuous and smooth, $x_2(t)$ is piecewise-continuous in SOD, $x_3(t)$ is piecewise-continuous in FOD, and $x_4(t)$ is piecewise-continuous in ZOD. The bandwidth of the BPF is $22kHz$ cantered at $33kHz$, i.e. spanning the frequency range from $22kHz$ to $44kHz$.

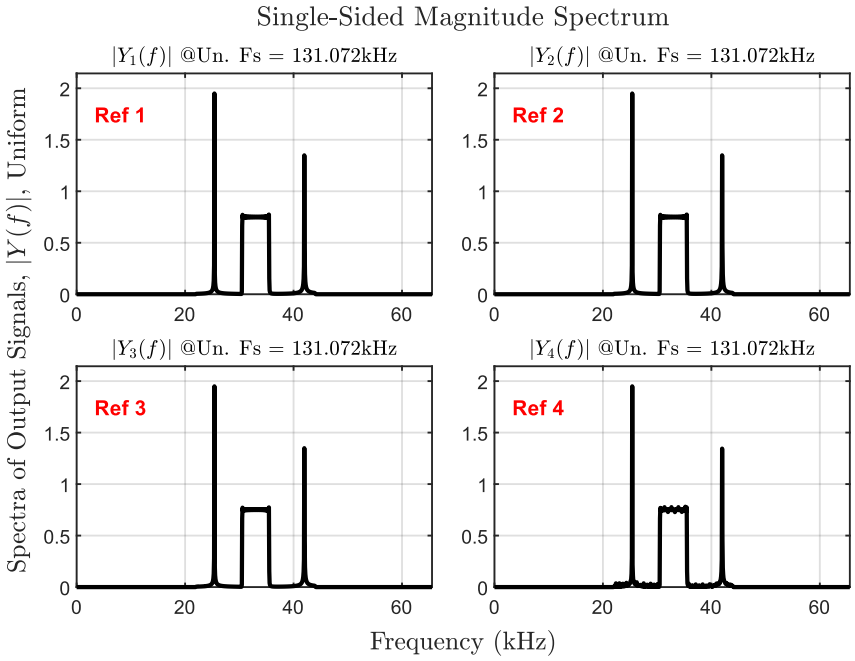


Fig. 60. Spectra of output signals for the uniformly sampled input signals with a sampling rate matching the Nyquist rate, i.e. $F_s = 131.072kHz$. Ref 1 to Ref 4 are the reference spectra for the subsequent corresponding estimated spectra.

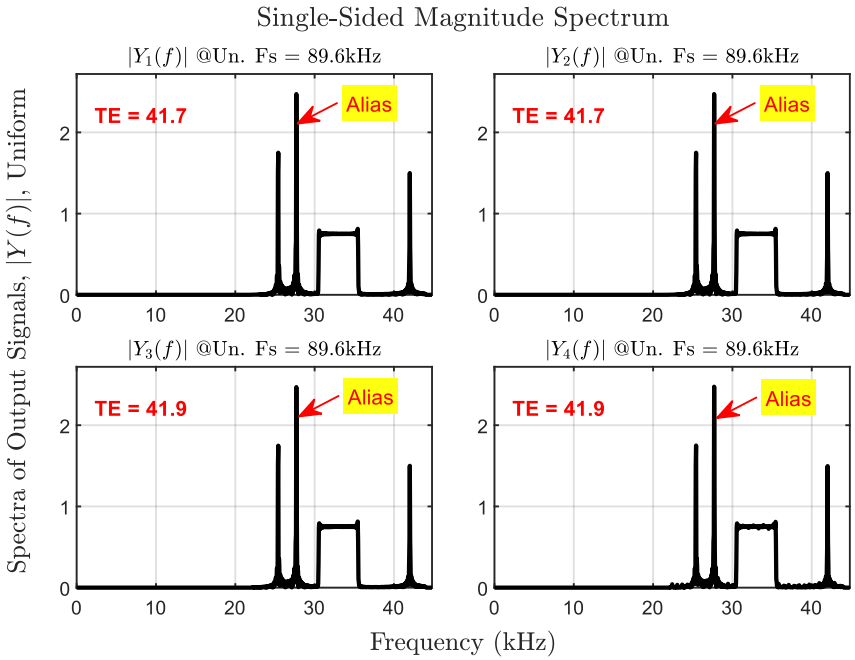


Fig. 61. Spectra of filtered output signals for the uniformly sampled input signals. In this figure, the utilised sampling rate is $F_s = 89.6kHz$. No antialiasing analog prefiltering is used, therefore aliases appear when the sampling rate is less than the Nyquist rate.

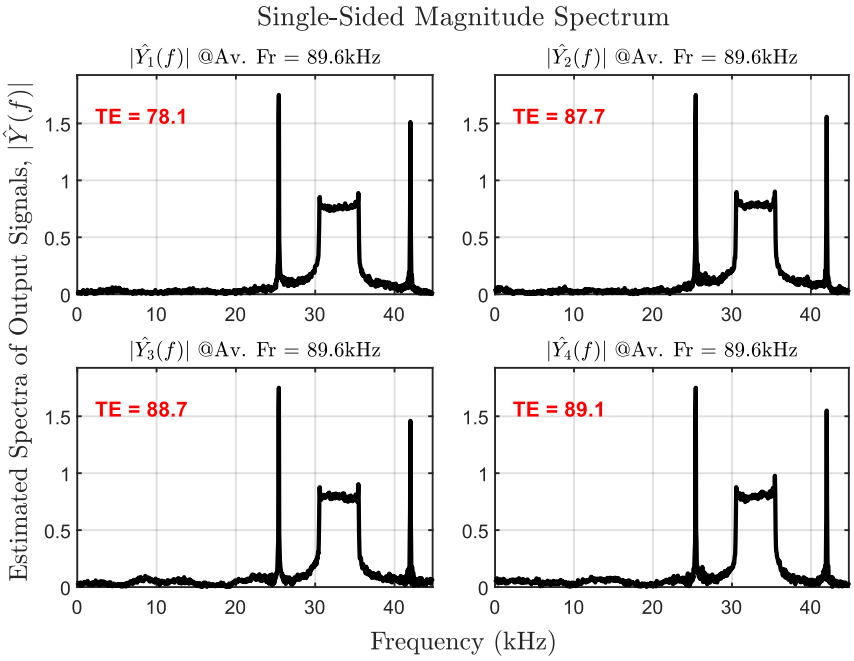


Fig. 62. Estimated spectra of output signals using DASP AnSt filter estimator and an average random sampling frequency of $Av.F_r = 89.6kHz$. 100 independent MC iterations are carried out. Remark that aliasing components are wiped out.

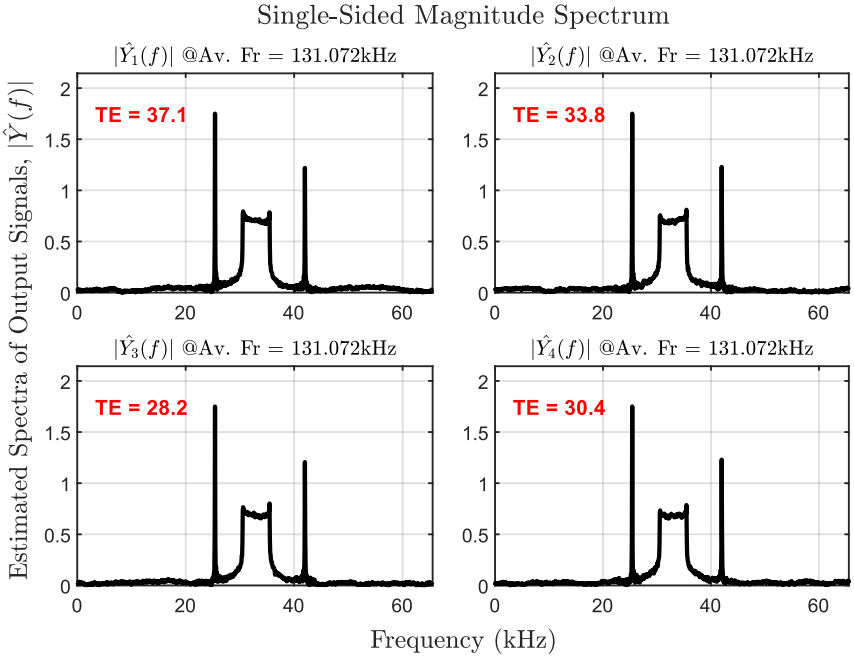


Fig. 63. Estimated spectra of output signals using DASP AnSt filter estimator and an average random sampling frequency of $Av.F_r = 131.072kHz$. 100 independent MC iterations are conducted to average out the results.

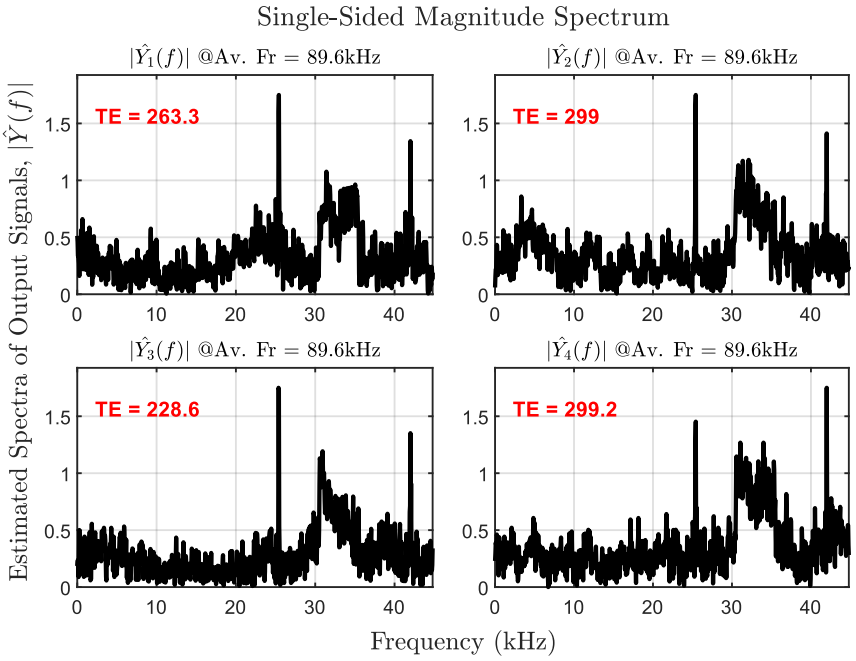


Fig. 64. Spectra of estimated filter output signals using a single realisation (MC=1) and an average random sampling frequency $Av.F_s = 89.6kHz$.

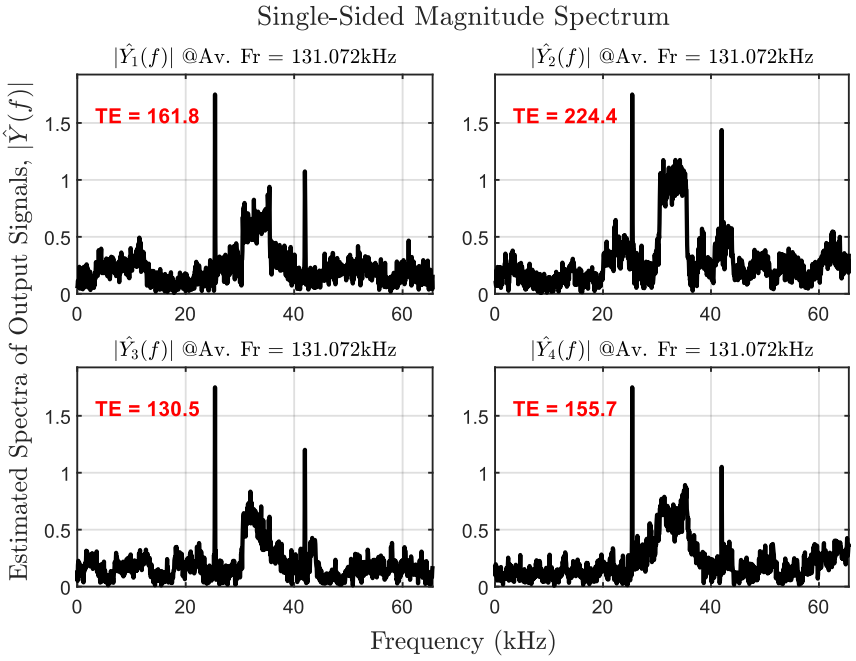


Fig. 65. Estimated spectra of filter output signals using a single realisation (MC=1) and an average random sampling frequency $Av.F_s = 131.072kHz$.

CHAPTER 7

Conclusions and Future Work

In this chapter, I outline the main conclusions of the thesis as per each chapter from Chapter 3 to Chapter 6. Some conclusions in a specific chapter may look like other chapter(s). But the ground truth is that each conclusion listed below is genuinely related to the respective chapter, and it was either introduced, devised, found, verified, or concluded within that very chapter. Moreover, I include some research areas and extensions to be addressed in any future work.

7.1. Conclusions

Up to the author knowledge, the conclusions summarised in the following list reflect the original contributions to knowledge that were made by this thesis:

Chapter 3

Three random sampling techniques (ToRa, StSa, and AnSt) were introduced and investigated as filter estimators in digital alias-free signal processing environment under the assumption that the input analog signal and/or filter impulse response are smooth continuous-time functions, and so are their derivatives. Furthermore, a generalised form filter estimator that encompasses the three estimators together was also established to make analysis concise and help developers/designers to implement such estimators in one general inclusive code.

The filter estimators were examined in terms of their main statistical characteristics, where they were all found to be unbiased and consistent, and their variances converge uniformly after certain numbers of sample points, depending mainly on the type of estimator. I also found that the uniform convergence rates vary according to which estimator is being considered. The fastest filter estimator is AnSt which converges at

a rate of N^{-5}, where $2N$ is the total number of sample points. The second one is StSa estimator with a converging speed of N^{-3}, however, in general, it starts to converge at a smaller number of sample points than the AnSt estimator does. The slowest estimator of all is ToRa with only N^{-1} decaying rate, but this rate is constant across any number of sample points. Meaning, it is the earliest estimator of the three to establish such a converging speed.

ToRa filter estimator was also proposed in this chapter in such a way that have never been addressed before in literature. I dealt with ToRa as if it is a special form of stratification-based sampling technique with only one stratum but $2N$ Monte Carlo iterations. This enabled us to devise a new variance expression for ToRa estimator, which is novel and original to the literature knowledge, even for other applications of ToRa estimator, such as Fourier transform.

Additionally, as analytically and numerically demonstrated in Chapter 3, we have seen how using suitable random sampling and filtering techniques helped in estimating filter output signal, even by using average random sampling rates less than the required Nyquist rate. This important advantage of the proposed random filter estimators cannot be achieved with the conventional DSP approaches. When the spectral support of the analog input signal is sparse and not fully known, and when utilising antialiasing analog pre-filtering is not viable, for any reason, then random filtering approaches are more cost-effective solutions than equivalent uniform ones.

The final conclusion in this chapter is that I proposed a mechanism for obtaining synchronised random sample points for both the input analog signal and the filter impulse response by using software defined analog filter representation, where the impulse response is formulated as an analog continuous-time function that could be instantly sampled together with the input signal and, therefore, no need for large bulk of memory to save high order filter coefficients.

Chapter 4

An oversampled on-grid filter impulse response is proposed in this chapter. The high-resolution impulse response is to be saved on a lookup table in the sampling circuit.

This arrangement is suitable for implementation on hardware testbeds, such as FPGAs, ASICs, or even DSP microcontrollers. The key purpose of oversampling on a uniform grid is to make use of pseudorandom sampling of both input signal and impulse response simultaneously to implement filter output estimation using my three random estimators, ToRa, StSa, and AnSt.

Furthermore, a new non-equidistant interpolation rule, composite 3-nonuniform sampling (C3NS), has been proposed. This rule relies on Lagrange interpolation polynomials to estimate an integrand/summand function AUC within a given time interval. It was used along with StSa approach to provide the fastest estimator of all considered ones that mainly rely on the simple Rectangular rule.

All proposed on-grid filter estimators are again proved to be unbiased and consistent. Though, their rates of uniform convergence are not identical. The C3NS estimator is the fastest to converge, after the asymptotic behaviour is established, with a rate of N^{-7}. The other Rectangular rule-based on-grid estimators, ToRa, StSa, and AnSt, have the same convergence speeds as in the previous chapter, i.e. N^{-1}, N^{-3}, and N^{-5} respectively. This is an interesting finding, since it widens the applications of filtering estimation, not only by using mixed hardware and software defined analog filtering (based on computers) solutions, but also standalone hardware ones (microcontrollers, FPGAs, ASICs, etc).

Chapter 5

I investigated the effect of non-smooth input signals and/or impulse response of a given FIR filter (which both comprise the integrand function of the filtering convolution operation) on the statistical features of estimated filter output signals using StSa random sampling technique. The StSa filter estimator is an unbiassed estimator, as proven in Chapter 3, and this is also applicable for the non-smoothness case. In Chapter 5, I verified its consistency, and devised exact mathematical expressions for its variance in two main cases regarding smoothness of the integrand function: piecewise-continuous first-order derivative; and piecewise-continuous zero-order derivative (i.e. the integrand function itself is non-smooth).

I found that despite the discontinuities that might present in the input signal/impulse response and/or their derivatives, the StSa filter estimator is still converging uniformly, after a certain number of sample points, to the true filter output at a rate depending on the order of the derivative where the discontinuities happen. If they appear on the first-order derivative, then the decaying rate is proportional to the negative third-power of the number of utilised input sample points as long as the number of discontinuities is finite, and they are bounded in magnitude. This converging rate slows down to the negative second-power of the sample sequence size if the discontinuities occur in the zero-order derivative. This means that the slowest converging speed of StSa filter estimator happens when either the analog input signal or filter impulse response, or both, is (are) piecewise-continuous, i.e. discontinuous at some limited number of time-instants.

I also conclude that it is possible to speed up the convergence rate of StSa filter estimator if, somehow, I was able to detect and eliminate jump discontinuities mainly from the input signal, since the impulse response of the filter can be managed and controlled locally. This very point is also added to the future works suggested to extend this research.

Chapter 6

In this chapter, I investigated the AnSt filter estimator performance and asymptotic behaviour in the non-smooth integrand function and/or its first two derivatives case. As an unbiassed estimator, AnSt was also proven to be consistent but with different converging rates according to the nonsmoothness status of the input signal/impulse response.

My mathematical derivations showed that estimation errors increase largely if the discontinuities appear in the input signal/impulse response functions themselves. If this is the case, then the convergence rate of the AnSt filter estimator is proportional to N^{-2} where $2N$ is the random sample sequence size.

The convergence rate of the estimator is greatly enhanced if the input signal/impulse response functions are continuous, but their first-order derivative(s) is (are) piecewise-

continuous. A faster converging rate of N^{-4} is achieved in this case, outperforming the previous case.

The last case addressed in this chapter was when the second-order derivative of the integrand function (input signal/impulse response) is piecewise-continuous involving limited number of jump discontinuities. The asymptotic behaviour of the AnSt filter estimator had not been affected with such non-smoothness, where the convergence rate was found to be equal to the case of continuous and smooth integrand function and its derivatives. That is, the rate of convergence is still be equal to N^{-5} as analytically validated.

Moreover, the demonstrated examples for AnSt filter estimator showed that despite the non-smoothness of the input signal, it was possible to carry out the filtering estimation with reduced average random sampling frequencies compared to the required Nyquist rate for the case of uniform filtering approaches. However, this advantage is not for free. Statistical errors appeared on the estimated outputs, especially when a single realisation of the randomly sampled input signal was considered. This could greatly be tackled with by means of Monte Carlo averaging, as also illustrated in this chapter.

7.2. Future Work

The work in this thesis could be extended to address the following areas/topics:

- **Additive White Gaussian Noise (AWGN).** It is possible to advance this research and verify the performance of the filter estimators under the presence of AWGN. As expected, this would increase estimation errors in general (i.e. the absolute MSE of the estimator), however, the exact new convergence rates of the estimators and their unbiasedness and consistency need to be investigated again.

  In this particular area of potential extension of this research, I would build on top of Pawlak et al's works [30], [65] for sampling and reconstruction of noisy signals based on equidistributed sequences or quasi-Monte Carlo algorithms and techniques.

- **Non-equidistant Stratification.** Although StSa and AnSt filter estimators discussed in this thesis is based on equal partitioning of strata, it is by no means restricted to this type of stratification setup. Indeed, both estimators can work with even enhanced performance if sufficient information about the input signal, or part of it, is available in advance. Thereafter, non-equidistant partitioning could be better to consider. If this is the case, then all devised expressions and decaying rates mentioned above will still be valid. However, the main difference is that the asymptotic convergence of the StSa estimator will be established earlier than equidistant stratification. Meaning, the absolute value of the variance will be smaller for the same number of sample points.

- **Infinite Impulse Response (IIR) Filters.** One may also investigate the case of IIR filters instead of FIR's. I addressed FIR filters mainly for their stability, causality and phase-linearity, which make filter estimation less prone to errors on top of the statistical errors due to sampling irregularities. However, lower-order IIR filters can be designed to achieve almost similar frequency response to higher-order FIR filters. If the drawbacks of IIR filters for a specific application can be tolerated, then considering them in filtering estimation would be more cost-effective than FIRs, especially in the case of on-grid pseudorandomisation.

- **Random Quantisation.** I addressed some irregular quantisation techniques in Chapter 2 which was dedicated to literature review. Nonetheless, I did not investigate any filter estimator, either in continuous or discontinuous cases, based on nonuniform/random quantisation. Advancing this research in this track would be plausible.

- **Detection and Elimination of Discontinuities.** In the second part of this thesis, I discussed filtering estimation in the presence of finite and bounded discontinuities in the integrand function or its first two derivatives. Though, I did not address techniques or algorithms to potentially detect and/or eliminate such discontinuities from the input signal to speed up the convergence rate of estimation. It is worth it to dig deeply into this area to find possible opportunities.

- **Hardware Implementation.** Although I have implemented many filter examples throughout this thesis by using software (i.e. MATLAB simulations), which is also

easy to be converted to any software-based solution (SDR, CR, etc.), I think one could try to implement filter estimators considered above by using hardware circuits. As we have pointed out earlier, FPGA, ASIC or any other microcontroller-based circuit is a capable device to host such filtering estimation applications.

Appendices

Appendix A: Proof of Theorem 5.1

From (5.13c-d) and (5.14b), we get

$$\begin{aligned}\mathbb{V}[\phi_j] = \frac{1}{\Delta}\int_{S_{j-1}}^{\tau_{D_j}} \Bigg(\Big(f_{j,L}(t,\tau) - f_{j,L}\left(t,\tau_{D_j}\right)\Big)\Delta + \frac{1}{2}{K_j}^2 F_{1j}\,\Delta^2 \\ - \frac{1}{2}\big(1-2K_j\big) f'_{j,R}\left(t,\tau_{D_j}\right)\Delta^2 - o(\Delta^2)\Bigg)^2 d\tau \\ + \frac{1}{\Delta}\int_{\tau_{D_j}}^{S_j} \Bigg(\Big(f_{j,R}(t,\tau) - f_{j,R}\left(t,\tau_{D_j}\right)\Big)\Delta + \frac{1}{2}{K_j}^2 F_{1j}\,\Delta^2 \\ - \frac{1}{2}\big(1-2K_j\big) f'_{j,R}\left(t,\tau_{D_j}\right)\Delta^2 - o(\Delta^2)\Bigg)^2 d\tau.\end{aligned} \tag{A.1}$$

By expanding $f_{j,L}(t,\tau)$ and $f_{j,R}(t,\tau)$ about τ_{D_j} using Taylor series and noticing that the absolute time instant of potential discontinuity is $\tau_{D_j} = S_{j-1} + K_j\Delta = S_j - \big(1-K_j\big)\Delta$, then (A.1) becomes

$$\begin{aligned}\mathbb{V}[\phi_j] = \frac{1}{\Delta}\int_{S_{j-1}}^{S_{j-1}+K_j\Delta} \Bigg(\Big(\big(\tau-\tau_{D_j}\big) f'_{j,L}\left(t,\tau_{D_j}\right) + o\left(\left|\tau-\tau_{D_j}\right|\right)\Big)\Delta \\ + \frac{1}{2}{K_j}^2 F_{1j}\,\Delta^2 - \frac{1}{2}\big(1-2K_j\big) f'_{j,R}\left(t,\tau_{D_j}\right)\Delta^2 - o(\Delta^2)\Bigg)^2 d\tau \\ + \frac{1}{\Delta}\int_{S_j-(1-K_j)\Delta}^{S_j} \Bigg(\Big(\big(\tau-\tau_{D_j}\big) f'_{j,R}\left(t,\tau_{D_j}\right) + o\left(\left|\tau-\tau_{D_j}\right|\right)\Big)\Delta \\ + \frac{1}{2}{K_j}^2 F_{1j}\,\Delta^2 - \frac{1}{2}\big(1-2K_j\big) f'_{j,R}\left(t,\tau_{D_j}\right)\Delta^2 - o(\Delta^2)\Bigg)^2 d\tau.\end{aligned} \tag{A.2}$$

By calculating the definite integral in (A.2), we obtain

$$\mathbb{V}[\phi_j] = \frac{1}{4}{\beta_j}^4 {F_{1j}}^2 \Delta^4 + \frac{1}{12}\left(f'_{j,L}\left(t, \tau_{D_j}\right)\right)^2 \Delta^4 + o(\Delta^4), \text{ where } \beta_j = K_j - 1. \quad \text{(A.3)}$$

(A.3) represents the part of variance associated with the sub-estimator of the j-th stratum in which there is a discontinuity in the integrand function's FOD at time instant τ_{D_j}. For other strata with no discontinuities at all, we have from Chapter 3 for the continuous integrand function and its derivatives case that the sub-variance associated with any j-th stratum is equal to $\frac{1}{12}\left(f'(C_j)\right)^2 \Delta^4 + o(\Delta^4)$, cf. (3.20b), where C_j is the centre of the stratum. To find the total value of the variance, we need to sum up the sub-variance values of M sub-estimators with discontinuities and $2N - M$ sub-estimators with no discontinuities. Thus,

$$\begin{aligned}\mathbb{V}[\hat{y}(t)] &= \sum_{j=1}^{2N} \mathbb{V}[\phi_j] \\ &= \sum_{j \in I_M} \left(\frac{1}{4}{\beta_j}^4 {F_{1j}}^2 \Delta^4 + \frac{1}{12}\left(f'_{j,L}\left(t, \tau_{D_j}\right)\right)^2 \Delta^4 + o(\Delta^4)\right) \\ &+ \sum_{\substack{j=1 \\ j \notin I_M}}^{2N} \left(\frac{1}{12}\left(f'(t, C_j)\right)^2 \Delta^4 + o(\Delta^4)\right),\end{aligned} \quad \text{(A.4)}$$

$$\begin{aligned}\mathbb{V}[\hat{y}(t)] &= \frac{T^4}{12(2N)^4} \sum_{j \in I_M} \left(3{\beta_j}^4 {F_{1j}}^2 + \left(f'_{j,L}\left(t, \tau_{D_j}\right)\right)^2\right) \\ &+ \frac{T^3}{12(2N)^3} \sum_{\substack{j=1 \\ j \notin I_M}}^{2N} \left(\left(f'(t, C_j)\right)^2 \Delta\right) + o(N^{-3}),\end{aligned} \quad \text{(A.5)}$$

where $I_M = \{i_1, i_2, i_3, \dots, i_M\}$ is a finite set of indices of size M for the strata with FOD discontinuities.

Remark that when $M = 0$, i.e. I_M is an empty set, then there are no discontinuities at all and (A.5) simplifies to $\frac{T^3}{12(2N)^3} \sum_{j=1}^{2N} \left(\left(f'(t, C_j)\right)^2 \Delta\right) + o(N^{-3})$ which is the same as the variance of StSa estimator in the case of smooth integrand function.

By using Riemann integration of the variance in (A.5) to find the convergence rate of the StSa filter estimator, we get

$$\begin{aligned}\lim_{N\to\infty}((2N)^3 \times \mathbb{V}[\hat{y}(t)]) &= \lim_{N\to\infty}\Bigg((2N)^3 \\ &\times\Bigg(\frac{T^4}{12(2N)^4}\sum_{j\in I_M}\left(3{\beta_j}^4{F_{1j}}^2+\left(f'_{j,L}\left(t,\tau_{D_j}\right)\right)^2\right) \\ &+\frac{T^3}{12(2N)^3}\sum_{\substack{j=1\\ j\notin I_M}}^{2N}\left(\left(f'\left(t,C_j\right)\right)^2\Delta\right)+o(N^{-3})\Bigg)\Bigg).\end{aligned} \tag{A.6}$$

The first summation in (A.6) drops out since it is finite and approaches to zero at a faster rate than the other summation. Moreover, since there are only M discontinuities in the FOD of the integrand function within the whole observation interval, this means that $f'(t,\tau)$ is a piecewise-continuous function having exactly $M+1$ integrable pieces. Thus,

$$\lim_{N\to\infty}((2N)^5 \times \mathbb{V}[\hat{y}(t)]) = \frac{T^3}{12}\sum_{k=1}^{M+1}\int_{T_{k-1}}^{T_k}\left(f'(t,\tau)\right)^2 d\tau. \tag{A.7}$$

This completes the proof of Theorem 5.1. ■

Appendix B: Proof of Theorem 5.2

Evaluating the integrals in (5.12a) under the assumptions, (5.17a-b), regarding non-smoothness of the ZOD of the integrand function, I obtain this new expression for the expected value of the j-th sub-estimator,

$$E[\phi_j] = f_{j,R}\left(t, \tau_{D_j}\right)\Delta + K_j F_{0j}\Delta + c_{2j} f'_{j,R}\left(t, \tau_{D_j}\right)\Delta^2 - \frac{1}{2}K_j^2\, F_{1j}\,\Delta^2 + o(\Delta^2). \tag{B.1}$$

We compute $e_j = \phi_j - E[\phi_j] = \begin{cases} e_{j,L}, & \tau_j \in A_{j,L} \\ e_{j,R}, & \tau_j \in A_{j,R} \end{cases}$ by subtracting (B.1) from (5.10b). Then we get

$$e_{j,L} = \left(f_{j,L}(t, \tau_j) - f_{j,R}\left(t, \tau_{D_j}\right)\right)\Delta - K_j F_{0j}\Delta - c_{2j} f'_{j,R}\left(t, \tau_{D_j}\right)\Delta^2 + \frac{1}{2}K_j^2\, F_{1j}\,\Delta^2 - o(\Delta^2) \tag{B.2}$$

$$e_{j,R} = \left(f_{j,R}(t, \tau_j) - f_{j,R}\left(t, \tau_{D_j}\right)\right)\Delta - K_j F_{0j}\Delta - c_{2j} f'_{j,R}\left(t, \tau_{D_j}\right)\Delta^2 + \frac{1}{2}K_j^2\, F_{1j}\,\Delta^2 - o(\Delta^2) \tag{B.3}$$

The variance value associated with the j-th sub-estimator is equal to the second moment of the error term e_j. That is, $\mathbb{V}[\phi_j] = \int_{-\infty}^{\infty} p_j(\tau)(e_j)^2 d\tau = \frac{1}{\Delta}\int_{S_{j-1}}^{S_j}(e_j)^2 d\tau$. Hence,

$$\mathbb{V}[\phi_j] = \int_{-\infty}^{\infty} p_j(\tau)(e_j)^2 d\tau = \frac{1}{\Delta}\int_{S_{j-1}}^{S_j}(e_j)^2 d\tau \tag{B.4}$$

$$\mathbb{V}[\phi_j] = \frac{1}{\Delta}\int_{A_{j,L}}(e_{j,L})^2 d\tau + \frac{1}{\Delta}\int_{A_{j,R}}(e_{j,R})^2 d\tau \tag{B.5}$$

$$\mathbb{V}[\phi_j] = {\beta_j}^2\, {F_{0j}}^2\Delta^2 - {\beta_j}^3 F_{0j}\, F_{1j}\,\Delta^3 + \frac{1}{4}{\beta_j}^4 {F_{1j}}^2\Delta^4 + \frac{1}{12}\left(f'_{j,L}\left(t, \tau_{D_j}\right)\right)^2\Delta^4 + o(\Delta^4) \tag{B.6}$$

The variance of the j-th sub-estimator found in (B.6) is for those strata which involve ZOD discontinuity exactly at time instant τ_{D_j}. Whereas for other strata which do not include any discontinuity, I again apply the j-th sub-estimator in (3.20b) as $f(t, \tau)$ is continuous within those strata. To find the whole variance of the StSa filter estimator

in the case of non-smooth integrand function, I add up all smooth and non-smooth variance values of the $2N$ sub-estimators. Therefore,

$$\begin{aligned}\mathbb{V}[\hat{y}(t)] = \sum_{j\in I_M}\Bigg(&\beta_j{}^2\, F_{0j}{}^2\Delta^2 - \beta_j{}^3 F_{0j}\, F_{1j}\,\Delta^3 + \frac{1}{4}\beta_j{}^4 F_{1j}{}^2\Delta^4 \\ &+ \frac{1}{12}\left(f'_{j,L}\left(t,\tau_{D_j}\right)\right)^2\Delta^4\Bigg) + M\times o(\Delta^4) \\ &+ \sum_{\substack{j=1\\ j\notin I_M}}^{2N}\left(\frac{1}{12}\left(f'\left(t,C_j\right)\right)^2\Delta^4\right) + o(\Delta^3),\end{aligned} \tag{B.7}$$

Note that $M\times o(\Delta^4)$ is still equal to $o(\Delta^4)$ since M is finite and does not depend on N, hence it drops out in the presence of $o(\Delta^3)$. Thus,

$$\begin{aligned}\mathbb{V}[\hat{y}(t)] = \sum_{j\in I_M}\Bigg(&\beta_j{}^2\, F_{0j}{}^2\Delta^2 - \beta_j{}^3 F_{0j}\, F_{1j}\,\Delta^3 + \frac{1}{4}\beta_j{}^4 F_{1j}{}^2\Delta^4 \\ &+ \frac{1}{12}\left(f'_{j,L}\left(t,\tau_{D_j}\right)\right)^2\Delta^4\Bigg) + \sum_{\substack{j=1\\ j\notin I_M}}^{2N}\left(\frac{1}{12}\left(f'\left(t,C_j\right)\right)^2\Delta^4\right) + o(\Delta^3).\end{aligned} \tag{B.8}$$

However, the dominant term in (B.8) as $N\to\infty$ is the first term, $\beta_j{}^2\, F_{0j}{}^2\Delta^2$, as it involves the least-power term of $\Delta = \frac{T}{2N}$. So, the variance of the estimator (B.8) simplifies to

$$\mathbb{V}[\hat{y}(t)] = \sum_{j\in I_M}\left(\beta_j{}^2\, F_{0j}{}^2\Delta^2\right) + o(\Delta^2) = \frac{T^2}{(2N)^2}\sum_{j\in I_M}\left(\beta_j{}^2\, F_{0j}{}^2\right) + o(\Delta^2). \tag{B.9}$$

It is clear from (B.9) that the uniform convergence rate of the StSa filter estimator is precisely N^{-2}, since the limit of (B.9) as $N\to\infty$ is

$$\lim_{N\to\infty}\left((2N)^2\,\mathbb{V}[\hat{y}(t)]\right) = \lim_{N\to\infty}\left((2N)^2\left(\frac{T^2}{(2N)^2}\sum_{j\in I_M}\left(\beta_j{}^2\, F_{0j}{}^2\right) + o(\Delta^2)\right)\right), \tag{B.10}$$

$$= T^2 \sum_{j \in I_M} \left({\beta_j}^2 \, {F_0}_j{}^2 \right)$$

This completes the proof of Theorem 5.2. ■

Appendix C: Proof of Theorem 6.1

From (6.16) and (6.15), we get

$$\begin{aligned} \mathbb{V}[\phi_j] = \int_{S_{j-1}}^{S_j} \frac{1}{\Delta} \Bigg(\frac{1}{4} \Big(f''_{j,L}\left(t, \tau_{D_j}\right) + f''_{j,R}\left(t, \tau_{D_j}\right) \Big) \left(\tau_j - \tau_{D_j}\right)^2 \Delta \\ - c_{2j} \, f''_{j,R}\left(t, \tau_{D_j}\right) \left(\tau_j - \tau_{D_j}\right) \Delta^2 \\ + \frac{1}{12} \Big(c_{3j} f''_{j,R}\left(t, \tau_{D_j}\right) - 2K_j^3 f''_{j,L}\left(t, \tau_{D_j}\right) \Big) \Delta^3 + o(\Delta^3) \Bigg)^2 d\tau. \end{aligned} \tag{C.1}$$

By expanding the integrand in (C.1) and calculating the definite integral, we obtain

$$\mathbb{V}[\phi_j] = \frac{1}{720} \Big(c_{4j} {F_{2j}}^2 + f''_{j,L}\left(t, \tau_{D_j}\right) f''_{j,R}\left(t, \tau_{D_j}\right) \Big) \Delta^6 + \, o(\Delta^6), \tag{C.2}$$

where $c_{4j} = 9 - 45K_j + 90K_j^2 - 110K_j^3 + 105K_j^4 - 60K_j^5 + 20K_j^6$, and F_{2j} is given in (6.8a), and denotes the difference between values of the left- and right-hand SOD sub-functions at the discontinuity time-instant of the j-th stratum ($= \tau_{D_j}$).

(C.2) is related to those M strata which have SOD discontinuities of the integrand function, whereas other strata involving smooth parts of the integrand function's SOD, the j-th sub-estimator's variance is given in (3.22), which is associated with the continuous part discussed in Chapter 3. Now, the whole variance can be calculated by adding M terms of (C.2) and $N - M$ terms of (3.22). So, we get

$$\mathbb{V}[\hat{y}(t)] = \sum_{j\in I_M}\left(\frac{1}{720}\left(c_{4j}{F_{2j}}^2 + f''_{j,L}\left(t,\tau_{D_j}\right)f''_{j,R}\left(t,\tau_{D_j}\right)\right)\Delta^6 + o(\Delta^6)\right) + \sum_{\substack{j=1\\ j\notin I_M}}^{N}\left(\frac{1}{720}\left(f''\left(t,C_j\right)\right)^2\Delta^6 + o(\Delta^6)\right), \tag{C.3}$$

$$\mathbb{V}[\hat{y}(t)] = \frac{T^6}{720(\mathrm{N})^6}\sum_{j\in I_M}\left(c_{4j}{F_{2j}}^2 + f''_{j,L}\left(t,\tau_{D_j}\right)f''_{j,R}\left(t,\tau_{D_j}\right)\right) + \frac{T^5}{720(\mathrm{N})^5}\sum_{\substack{j=1\\ j\notin I_M}}^{N}\left(\left(f''\left(t,C_j\right)\right)^2\Delta\right) + o(N^{-5}), \tag{C.4}$$

$$\mathbb{V}[\hat{y}(t)] = \frac{T^5}{720(\mathrm{N})^5}\sum_{\substack{j=1\\ j\notin I_M}}^{N}\left(\left(f''\left(t,C_j\right)\right)^2\Delta\right) + o(N^{-5}). \tag{C.5}$$

where $I_M = \{i_1, i_2, i_3, \dots, i_M\}$ is a set of indices for the M strata with SOD discontinuities. Remark that the first summation in (C.4) comprises finite items and depends not on N. Hence, as $N \to \infty$ it tends to zero at a rate of N^{-6}, i.e. faster than the other summation. Thus, it can be embedded in $o(N^{-5})$ term appearing in (C.5).

By using Riemann integration, the variance of the AnSt filter estimator, (C.5), can be verified for its convergence rate. Thus,

$$\lim_{N\to\infty}((2N)^5 \times \mathbb{V}[\hat{y}(t)]) = \lim_{N\to\infty}\left((2N)^5 \times \left(\frac{T^5}{720(\mathrm{N})^5}\sum_{\substack{j=1\\ j\notin I_M}}^{N}\left(\left(f''\left(t,C_j\right)\right)^2\Delta\right) + o(N^{-5})\right)\right). \tag{C.6}$$

Note that $\Delta = T/N$, and $\lim_{N\to\infty}\left((2N)^5 \times o(N^{-5})\right) = 0$. Moreover, since there are only M discontinuities in the SOD within the whole observation interval, this means that $f''(t,\tau)$ is a piecewise-continuous function having exactly $M+1$ pieces. Thus,

$$= \frac{2T^5}{45}\sum_{k=1}^{M+1}\int_{T_{k-1}}^{T_k}\left(f''(t,\tau)\right)^2 d\tau. \tag{C.8}$$

$\lim_{N\to\infty}((2N)^5$

$\times \mathbb{V}[\hat{y}(t)])$

This completes the proof of Theorem 6.1. ■

Appendix D: Proof of Theorem 6.2

Recalculating the integrals in (6.12), this time taking (6.19a-c) into consideration instead of (6.9a-c), we get a new expression for the expected value of the j-th sub-estimator,

$$\begin{aligned} E[\phi_j] \;&= \frac{1}{6}\left(K_j^3 f''_{j,L}\left(t,\tau_{D_j}\right) - \left(K_j^3 - 3K_j^2 + 3K_j - 1\right) f''_{j,R}\left(t,\tau_{D_j}\right)\right)\Delta^3 \\ &+ \frac{1}{2}\left(1-2K_j\right) f'_{j,R}\left(t,\tau_{D_j}\right)\Delta^2 \\ &- \frac{1}{2}K_j^2\left(f'_{j,L}\left(t,\tau_{D_j}\right) - f'_{j,R}\left(t,\tau_{D_j}\right)\right)\Delta^2 + f_{j,R}\left(t,\tau_{D_j}\right)\Delta + o(\Delta^3). \end{aligned} \tag{D.1}$$

$$\begin{aligned} &= \frac{1}{6}\left(K_j^3 f''_{j,L}\left(t,\tau_{D_j}\right) - c_{1j}\, f''_{j,R}\left(t,\tau_{D_j}\right)\right)\Delta^3 + c_{2j}\, f'_{j,R}\left(t,\tau_{D_j}\right)\Delta^2 \\ &- \frac{1}{2}K_j^2\, F_{1j}\,\Delta^2 + f_{j,R}\left(t,\tau_{D_j}\right)\Delta + o(\Delta^3). \end{aligned} \tag{D.2}$$

where $c_{1j} = K_j^3 - 3K_j^2 + 3K_j - 1$, $c_{2j} = \frac{1}{2}\left(1-2K_j\right)$, and $F_{1j} := f'_{j,L}\left(t,\tau_{D_j}\right) - f'_{j,R}\left(t,\tau_{D_j}\right)$.

Computing the new error term $e_j = \phi_j - E[\phi_j]$ from (6.10b) and (D.1), for the case when $\tau_j < \tau_{D_j} < \tau_j^a$, we get

$$
\begin{aligned}
e_j \;=& \frac{1}{2}\Delta\Bigg(2f_{j,R}\left(t,\tau_{D_j}\right) - f'_{j,L}\left(t,\tau_{D_j}\right)\left(\tau_{D_j} - \tau_j\right) \\
&- f'_{j,R}\left(t,\tau_{D_j}\right)\left(\tau_j - \tau_{D_j} + 2\Delta\left(K_j - \frac{1}{2}\right)\right) + \frac{1}{2}f''_{j,L}\left(t,\tau_{D_j}\right)\left(\tau_{D_j} - \tau_j\right)^2 \\
&+ \frac{1}{2}f''_{j,R}\left(t,\tau_{D_j}\right)\left(\tau_j - \tau_{D_j} + 2\Delta\left(K_j - \frac{1}{2}\right)\right)^2\Bigg) \\
&+ \frac{1}{6}\Delta\left(-3\Delta f'_{j,R}\left(t,\tau_{D_j}\right)\left(K_j - 1\right)^2 + 6f_{j,R}\left(t,\tau_{D_j}\right)\left(K_j - 1\right)\right. \\
&\left.+ \Delta^2 f''_{j,R}\left(t,\tau_{D_j}\right)\left(K_j - 1\right)^3\right) \\
&- \frac{1}{6}\Delta K_j\left(6f_{j,R}\left(t,\tau_{D_j}\right) - 3\Delta K_j f'_{j,L}\left(t,\tau_{D_j}\right) + \Delta^2 K_j^2 f''_{j,L}\left(t,\tau_{D_j}\right)\right) + o(\Delta^3).
\end{aligned}
\tag{D.3}
$$

Since the error term, e_j, is real-valued and the estimator is unbiassed, the expected value associated with the j-th sub-estimator is equal to the variance of the sub-estimator, i.e. $\mathbb{V}[\phi_j] = \int_{-\infty}^{\infty} p_j(\tau)\left(e_j\right)^2 d\tau = \frac{1}{\Delta}\int_{S_{j-1}}^{S_j}\left(e_j\right)^2 d\tau$. Evaluating such integral here would require substantial space, as it is very long. However, it is easy to calculate for it is a collection of various degree polynomials. Hence, I present here the final result,

$$
\begin{aligned}
\mathbb{V}[\phi_j] =& \frac{c_{5j}}{12}{F_{1j}}^2\Delta^4 + \frac{c_{6j}}{48}F_{2j}F_{1j}\Delta^5 + \frac{c_{4j}}{720}{F_{2j}}^2\Delta^6 \\
&+ \frac{1}{720}f''_{j,L}\left(t,\tau_{D_n}\right)f''_{j,R}\left(t,\tau_{D_n}\right)\Delta^6 + o(\Delta^6),
\end{aligned}
\tag{D.4}
$$

where $c_{4j} = 9 - 45K_j + 90K_j^2 - 110K_j^3 + 105K_j^4 - 60K_j^5 + 20K_j^6$, $c_{5j} = 1 - 3\,K_j + 6\,K_j^2 - 6\,K_j^3 + 3\,K_j^4$, and $c_{6j} = 2c_{2j}\left(3 - 6\,K_j + 10\,K_j^2 - 8\,K_j^3 + 4\,K_j^4\right)$.

The whole variance of the AnSt filter estimator, $\mathbb{V}[\hat{y}(t)]$, utilising N-strata (i.e. $2N$ sample points) in the case of non-smooth FOD with M bounded discontinuities is simply the summation of the individual sub-variances in each stratum, as they are all calculated using statistically independent random variables. However, only M strata have such FOD discontinuities, whereas N-M strata do not. Therefore, I need to add a mix of (D.4) and (3.22), as follows

$$\mathbb{V}[\hat{y}(t)] = \sum_{j \in I_M} \left(\frac{c_{5j}}{12} {F_{1j}}^2 \Delta^4 + \frac{c_{6j}}{48} F_{2j} F_{1j} \Delta^5 + \frac{c_{4j}}{720} {F_{2j}}^2 \Delta^6 \right.$$
$$+ \frac{1}{720} f''_{j,L}\left(t, \tau_{D_j}\right) f''_{j,R}\left(t, \tau_{D_j}\right) \Delta^6 + o(\Delta^6) \Big) \tag{D.5}$$
$$+ \sum_{\substack{j=1 \\ j \notin I_M}}^{N} \left(\frac{1}{720} \left(f''\left(t, C_j\right) \right)^2 \Delta^6 + o(\Delta^6) \right),$$

$$\mathbb{V}[\hat{y}(t)] = \frac{\Delta^4}{12} \sum_{j \in I_M} \left(c_{5j} {F_{1j}}^2 \right) + \frac{\Delta^5}{720} \sum_{\substack{j=1 \\ j \notin I_M}}^{N} \left(\left(f''\left(t, C_j\right) \right)^2 \Delta \right) + o(\Delta^5)\,, \tag{D.6}$$

To compute the convergence rate of the variance in (D.6), I take the limit as $N \to \infty$. Note that the second and third terms of the RHS of (D.6) will then be equal to zero. Thus,

$$\lim_{N \to \infty} \left((2N)^4 \times \mathbb{V}[\hat{y}(t)] \right) = \lim_{N \to \infty} \left((2N)^4 \times \left(\frac{T^4}{12N^4} \sum_{j \in I_M} \left(c_{5j} {F_{1j}}^2 \right) \right) \right), \tag{D.7}$$

$$= \frac{4T^4}{3} \sum_{j \in I_M} c_{5j}\, {F_{1j}}^2. \tag{D.8}$$

Which means that the AnSt estimator converges at a rate of N^{-4} if the integrand function's first-order derivative is non-smooth (i.e. piecewise-continuous).

This completes the proof of Theorem 6.2. ■

Appendix E: Proof of Theorem 6.3

Evaluating the integrals in (6.12) under the new assumptions, (6.22a-c), regarding non-smoothness of the ZOD of the integrand function, I obtain this new expression for the expected value of the j-th sub-estimator,

$$E[\phi_j] = \frac{1}{6}\left(K_j^3 f''_{j,L}\left(t,\tau_{D_j}\right) - c_{1j} f''_{j,R}\left(t,\tau_{D_j}\right)\right)\Delta^3 + c_{2j} f'_{j,R}\left(t,\tau_{D_j}\right)\Delta^2 - \frac{1}{2}K_j^2\, F_{1j}\,\Delta^2 + f_{j,R}\left(t,\tau_{D_j}\right)\Delta + K_j F_{0j}\Delta + o(\Delta^3). \quad \text{(E.1)}$$

Similar to the previous two appendices, I compute $e_j = \phi_j - E[\phi_j]$ when $\tau_j < \tau_{D_j} < \tau_j^a$, since other possibilities for the location of τ_{D_j} with respect to τ_j and τ_j^a has no effect on the convergence rate of the estimator, excluding the stratum centre and borders. Consequently, by subtracting (E.1) from (6.10b) and doing some algebra, the error term e_j simplifies to

$$\begin{aligned} e_j = {} & \frac{1}{2}\Delta\Bigg(f_{j,L}\left(t,\tau_{D_j}\right) + f_{j,R}\left(t,\tau_{D_j}\right) - f'_{j,L}\left(t,\tau_{D_j}\right)\left(\tau_{D_j} - \tau_j\right) \\ & - f'_{j,R}\left(t,\tau_{D_j}\right)\left(\tau_j - \tau_{D_j} + 2\Delta\left(K_j - \frac{1}{2}\right)\right) + \frac{1}{2}f''_{j,L}\left(t,\tau_{D_j}\right)\left(\tau_{D_j} - \tau_j\right)^2 \\ & + \frac{1}{2}f''_{j,R}\left(t,\tau_{D_j}\right)\left(\tau_j - \tau_{D_j} + 2\Delta\left(K_j - \frac{1}{2}\right)\right)^2\Bigg) \\ & + \frac{1}{6}\Delta(K_j - 1)\left(-3\Delta f'_{j,R}\left(t,\tau_{D_j}\right)(K_j - 1) + 6 f_{j,R}\left(t,\tau_{D_j}\right)\right. \\ & \left. + \Delta^2 f''_{j,R}\left(t,\tau_{D_j}\right)(K_j - 1)^2\right) \\ & - \frac{1}{6}\Delta K_j\left(6 f_{j,L}\left(t,\tau_{D_j}\right) - 3\Delta K_j f'_{j,L}\left(t,\tau_{D_j}\right) + \Delta^2 K_j^2 f''_{j,L}\left(t,\tau_{D_j}\right)\right) + o(\Delta^3). \end{aligned} \quad \text{(E.2)}$$

The variance value associated with the j-th sub-estimator is equal to the second moment of the error e_j. That is, $\mathbb{V}[\phi_j] = \int_{-\infty}^{\infty} p_j(\tau)(e_j)^2 d\tau = \frac{1}{\Delta}\int_{S_{j-1}}^{S_j}(e_j)^2 d\tau$. Hence,

$$\begin{aligned} \mathbb{V}[\phi_j] = {} & {c_{2j}}^2\, {F_{0j}}^2\,\Delta^2 + \frac{1}{2}c_{2j}\, c_{7j}\, F_{1j}F_{0j}\Delta^3 \\ & + \frac{1}{12}\left(c_{5j}\,{F_{1j}}^2 + 4{c_{2j}}^2 c_{8j}\, F_{2j}F_{0j}\right)\Delta^4 + \frac{1}{48}c_{6j}\, F_{2j}F_{1j}\Delta^5 \\ & + \frac{1}{720}c_{4j}\,{F_{2j}}^2\,\Delta^6 + \frac{1}{720}f''_{j,L}\left(t,\tau_{D_j}\right)f''_{j,R}\left(t,\tau_{D_j}\right)\Delta^6 + o(\Delta^6), \end{aligned} \quad \text{(E.3)}$$

where $c_{6j} = 2c_{2j}\left(3 - 6\,K_j + 10\,K_j^2 - 8\,K_j^3 + 4\,K_j^4\right)$, $c_{7j} = 1 - 2K_j + 2K_j^2$ and $c_{8j} = 1 - K_j + K_j^2$.

To find the whole variance of the AnSt filter estimator in the case of non-smooth integrand function, I add up all variance values of the N sub-estimators. So, I have M strata with ZOD discontinuities and $N - M$ strata with smooth ZOD. Hence,

$$\begin{aligned}\mathbb{V}[\hat{y}(t)] = \sum_{j \in I_M} \Big(& {c_{2j}}^2 \, {F_{0j}}^2 \, \Delta^2 + \frac{1}{2} c_{2j} \, c_{7j} \, F_{1j} F_{0j} \Delta^3 \\ & + \frac{1}{12} \left(c_{5j} \, {F_{1j}}^2 + 4 {c_{2j}}^2 c_{8j} \, F_{2j} F_{0j} \right) \Delta^4 + \frac{1}{48} c_{6j} \, F_{2j} F_{1j} \Delta^5 \\ & + \frac{1}{720} c_{4j} \, {F_{2j}}^2 \, \Delta^6 + \frac{1}{720} f''_{j,L}\left(t, \tau_{D_j}\right) f''_{j,R}\left(t, \tau_{D_j}\right) \Delta^6 \Big) \\ & + \sum_{\substack{j=1 \\ j \notin I_M}}^{N} \left(\frac{1}{720} \left(f''(t, C_j) \right)^2 \Delta^6 \right) + o(\Delta^5).\end{aligned} \tag{E.4}$$

Since the term ${c_{2j}}^2 \, {F_{0j}}^2 \, \Delta^2$ in the first summation of (E.4) involves second-power of $\Delta = \frac{T}{N}$, while all other terms contain Δ raised to higher powers, then the first term will be the dominant as $N \to \infty$. Therefore, the variance of the estimator (E.4) simplifies to

$$\mathbb{V}[\hat{y}(t)] = \sum_{j \in I_M} \left({c_{2j}}^2 \, {F_{0j}}^2 \, \Delta^2 \right) + o(\Delta^2) = \frac{T^2}{N^2} \sum_{j \in I_M} \left({c_{2j}}^2 \, {F_{0j}}^2 \right) + o(N^{-2}). \tag{E.5}$$

Remark that $F_{0j} = f_{j,L}\left(t, \tau_{D_j}\right) - f_{j,R}\left(t, \tau_{D_j}\right)$ is a nonzero value only for those M strata having ZOD discontinuities in the integrand function, that is why the summation in (E.5) includes only M terms with $j \in I_M$.

It is clear from (E.5) that the uniform convergence rate of the AnSt filter estimator is precisely N^{-2}, since the limit of (E.5) when the number of strata approaches infinity is

$$\lim_{N \to \infty} \left((2N)^4 \, \mathbb{V}[\hat{y}(t)] \right) = \lim_{N \to \infty} \left((2N)^4 \left(\frac{T^2}{N^2} \sum_{j \in I_M} \left({c_{2j}}^2 \, {F_{0j}}^2 \right) + o(N^{-2}) \right) \right), \tag{E.6}$$

$$= 16T^2 \sum_{j \in I_M} \left({c_{2j}}^2 \, {F_{0j}}^2 \right)$$

This completes the proof of Theorem 6.3. ■

References

[1] C. E. Shannon, “Communication in the Presence of Noise,” *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, Jan. 1949.

[2] H. Nyquist, “Certain Topics in Telegraph Transmission Theory,” *Transactions of the American Institute of Electrical Engineers*, vol. 47, no. 2, pp. 617–644, Apr. 1928.

[3] H. J. Landau, “Necessary density conditions for sampling and interpolation of certain entire functions,” *Acta Math.*, vol. 117, no. 1, pp. 37–52, Jul. 1967.

[4] R. G. Vaughan, N. L. Scott, and D. R. White, “The theory of bandpass sampling,” *IEEE Transactions on Signal Processing*, vol. 39, no. 9, pp. 1973–1984, Sep. 1991.

[5] Y. P. Lin and P. P. Vaidyanathan, “Periodically nonuniform sampling of bandpass signals,” *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II-ANALOG AND DIGITAL SIGNAL PROCESSING*, vol. 45, no. 3, pp. 340–351, Mar. 1998.

[6] S. C. Scoular and W. J. Fitzgerald, “Periodic nonuniform sampling of multiband signals,” *Signal Processing*, vol. 28, no. 2, pp. 195–200, Aug. 1992.

[7] R. Venkataramani and Y. Bresler, “Perfect reconstruction formulas and bounds on aliasing error in sub-Nyquist nonuniform sampling of multiband signals,” *IEEE Transactions on Information Theory*, vol. 46, no. 6, pp. 2173–2183, 2000.

[8] P. Feng and Y. Bresler, “Spectrum-blind minimum-rate sampling and reconstruction of multiband signals,” in *1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings*, May 1996, vol. 3, pp. 1688–1691 vol. 3.

[9] M. Mishali and Y. C. Eldar, “Blind multiband signal reconstruction: Compressed sensing for analog signals,” *IEEE Transactions on signal processing*, vol. 57, no. 3, pp. 993–1009, 2009.

[10] M. Mishali and Y. C. Eldar, “From Theory to Practice: Sub-Nyquist Sampling of Sparse Wideband Analog Signals,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 375–391, Apr. 2010.

[11] I. Bilinskis, *Digital alias-free signal processing*. Chichester ; Hoboken, N.J: John Wiley, 2007.

[12] I. Bilinskis and A. K. Mikelson, *Randomized Signal Processing*. New York: Prentice Hall, 1992.

[13] F. Esqueda, S. Bilbao, and V. Valimaki, “Aliasing Reduction in Clipped Signals,” *IEEE Trans. Signal Process.*, vol. 64, no. 20, pp. 5255–5267, Oct. 2016.

[14] K. Gulati and H. Lee, “Reconfigurable analog-to-digital converter,” Feb. 03, 2004.

[15] C. R. Schlottmann, S. Shapero, S. Nease, and P. Hasler, “A digitally enhanced dynamically reconfigurable analog platform for low-power signal processing,” *IEEE Journal of Solid-State Circuits*, vol. 47, no. 9, pp. 2174–2184, 2012.

[16] K. Kiela, M. Jurgo, V. Macaitis, and R. Navickas, “Wideband Reconfigurable Integrated Low-Pass Filter for 5G Compatible Software Defined Radio Solutions,” *Electronics*, vol. 10, no. 6, Art. no. 6, Jan. 2021.

[17] Alejandro. G. Gener, J. Valverde, J. A. Otero, and P. J. Harris, “A Fast Prototyping Workflow for Reconfigurable SDR Applications,” in *2019 14th International Symposium on Reconfigurable Communication-centric Systems-on-Chip (ReCoSoC)*, Jul. 2019, pp. 66–73.

[18] P. Ramakrishna and K. H. Kishore, “A low power reconfigurable ADC for bioimpedance monitroing system,” *Int J Speech Technol*, Mar. 2021.

[19] H. Y. Darawsheh and A. Jamoos, “Performance Analysis of Energy Detector Over α-μ Fading Channels with Selection Combining,” *Wireless personal communications*, vol. 77, no. 2, pp. 1507–1517, 2014.

[20] H. Y. Darawsheh and A. Jamoos, “Selection diversity combining analysis of energy detector over α-μ generalized fading channels,” in *2013 The International Conference on Technological Advances in Electrical, Electronics and Computer Engineering (TAEECE)*, 2013, pp. 563–567.

[21] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “Wireless sensor networks: a survey,” *Computer Networks*, vol. 38, no. 4, pp. 393–422, Mar. 2002.

[22] R. R. Ernst, G. Bodenhausen, and A. Wokaun, *Principles of nuclear magnetic resonance in one and two dimensions*, vol. 14. Clarendon press Oxford, 1987.

[23] J. Yen, “On Nonuniform Sampling of Bandwidth-Limited Signals,” *IRE Transactions on Circuit Theory*, vol. 3, no. 4, pp. 251–257, Dec. 1956.

[24] H. S. Shapiro and R. A. Silverman, “Alias-Free Sampling of Random Noise,” *Journal of the Society for Industrial and Applied Mathematics*, vol. 8, no. 2, pp. 225–248, 1960.

[25] E. Masry, “Alias-free sampling: An alternative conceptualization and its applications,” *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 317–324, May 1978.

[26] V. A. Kotelnikov, “On the transmission capacity of the ‘ether’and of cables in electrical communications,” 1933.

[27] P. L. Butzer, “A survey of the Whittaker-Shannon sampling theorem and some of its extensions,” in *Journal of Mathematical Research and Exposition*, 1983, vol. 3, no. 1.

[28] C. B. Feldman and W. R. Bennett, “Band width and transmission performance,” *The Bell System Technical Journal*, vol. 28, no. 3, pp. 490–595, Jul. 1949.

[29] T. Schanze, “Sinc interpolation of discrete periodic signals,” *IEEE Transactions on Signal Processing*, vol. 43, no. 6, pp. 1502–1503, Jun. 1995.

[30] M. Pawlak, E. Rafajlowicz, and A. Krzyzak, “Postfiltering versus prefiltering for signal recovery from noisy samples,” *IEEE Transactions on Information Theory*, vol. 49, no. 12, pp. 3195–3212, Dec. 2003.

[31] J.-B.-J. Fourier, *Théorie analytique de la chaleur*. Paris: F. Didot, 1822.

[32] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, vol. 55. US Government printing office, 1964.

[33] H. Jeffreys, *Methods of Mathematical Physics, Third Edition*, 3rd edition. Cambridge, U.K. ; New York: Cambridge University Press, 1999.

[34] P. O. J. Scherer, “Numerical Integration,” in *Computational Physics: Simulation of Classical and Quantum Systems*, P. O. J. Scherer, Ed. Berlin, Heidelberg: Springer, 2010, pp. 37–46.

[35] P. J. Davis, *Interpolation and approximation*. Courier Corporation, 1975.

[36] A. Quarteroni, R. Sacco, and F. Saleri, *Numerical mathematics*, vol. 37. Springer Science & Business Media, 2010.

[37] H. Y. Darawsheh and A. Tarczynski, “Comparison Between Uniform and Nonuniform Interpolation Techniques for Digital Alias-free FIR Filtering,” in *International Conference on Digital Image & Signal Processing (DISP’19)*, Oxford, United Kingdom, May 2019, pp. 1–5.

[38] F. Eng, “Non-uniform sampling in statistical signal processing,” PhD Thesis, Institutionen för systemteknik, 2007.

[39] J. J. Benedetto and P. J. S. G. Ferreira, *Modern Sampling Theory: Mathematics and Applications*. Springer Science & Business Media, 2012.

[40] A. A. Afifi and R. M. Elashoff, “Missing Observations in Multivariate Statistics I. Review of the Literature,” *Journal of the American Statistical Association*, vol. 61, no. 315, pp. 595–604, Sep. 1966.

[41] M. A. Woodbury, “A Missing Information Principle: Theory and Applications,” Duke University Medical Center Durham United States, Duke University Medical Center Durham United States, Jun. 1970. Accessed: Oct. 19, 2017. [Online]. Available: http://www.dtic.mil/docs/citations/AD1022173.

[42] P. J. S. G. Ferreira, “Incomplete sampling series and the recovery of missing samples from oversampled band-limited signals,” *IEEE Transactions on Signal Processing*, vol. 40, no. 1, pp. 225–227, Jan. 1992.

[43] J. Ma, R. Tao, Y. Li, and X. Kang, “Fractional Spectrum Analysis for Nonuniform Sampling in the Presence of Clock Jitter and Timing Offset,” *IEEE Transactions on Signal Processing*, vol. 68, pp. 4148–4162, 2020.

[44] A. Tarczynski and N. Allay, “Spectral analysis of randomly sampled signals: suppression of aliasing and sampler jitter,” *IEEE Transactions on Signal Processing*, vol. 52, no. 12, pp. 3324–3334, Dec. 2004.

[45] E. Masry, “Analysis of Signal Reconstruction With Jittered Sampling,” *IEEE Transactions on Signal Processing*, vol. 59, no. 1, pp. 27–34, Jan. 2011.

[46] P. Martínez-Nuevo, “Nonuniform Sampling Rate Conversion:An Efficient Approach,” *IEEE Transactions on Signal Processing*, vol. 69, pp. 2913–2922, 2021.

[47] F. Marvasti, *Nonuniform Sampling: Theory and Practice*. Springer Science & Business Media, 2012.

[48] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, “Spectrum sensing for cognitive radio: State-of-the-art and recent advances,” *IEEE signal processing magazine*, vol. 29, no. 3, pp. 101–116, 2012.

[49] F. J. Beutler, “Error-Free Recovery of Signals from Irregularly Spaced Samples,” *SIAM Rev.*, vol. 8, no. 3, pp. 328–335, Jul. 1966.

[50] F. Beutler, “Alias-free randomly timed sampling of stochastic processes,” *IEEE Transactions on Information Theory*, vol. 16, no. 2, pp. 147–152, Mar. 1970.

[51] E. Masry, “Random sampling and reconstruction of spectra,” *Information and Control*, vol. 19, no. 4, pp. 275–288, Nov. 1971.

[52] E. Masry, “Random sampling of deterministic signals: statistical analysis of Fourier transform estimates,” *IEEE Transactions on Signal Processing*, vol. 54, no. 5, pp. 1750–1761, May 2006.

[53] E. Masry and A. Vadrevu, “Random Sampling Estimates of Fourier Transforms: Antithetical Stratified Monte Carlo,” *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 194–204, Jan. 2009.

[54] N. Sayiner, H. V. Sorensen, and T. R. Viswanathan, “A level-crossing sampling scheme for A/D conversion,” *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 43, no. 4, pp. 335–339, 1996.

[55] B. Bidegaray-Fesquet and L. Fesquet, “Levels, peaks, slopes… which sampling for which purpose?,” Jun. 2016, pp. 1–6.

[56] M. Miskowicz, “Send-on-delta concept: An event-based data reporting strategy,” *sensors*, vol. 6, no. 1, pp. 49–63, 2006.

[57] R. M. Haralick, “Digital step edges from zero crossing of second directional derivatives,” in *Readings in Computer Vision*, Elsevier, 1987, pp. 216–226.

[58] S. Kay and R. Sudhaker, “A zero crossing-based spectrum analyzer,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, no. 1, pp. 96–104, 1986.

[59] I. Bilinskis, E. Boole, and K. Sudars, “Combination of periodic and alias-free non-uniform signal sampling for wideband signal digitizing and compressed transmitting based on picosecond-resolution event timing,” in *2017 Signal Processing Symposium (SPSympo)*, Sep. 2017, pp. 1–6.

[60] A. Tarczynski, “Alias-free Discrete-time FIR System Realisation Using Hybrid Stratified Sampling,” in *2020 6th International Conference on Event-Based Control, Communication, and Signal Processing (EBCCSP)*, Sep. 2020, pp. 1–8.

[61] M. Al-Ani, A. T. Tarczynski, and B. I. Ahmad, “High-Order Hybrid Stratified Sampling: Fast Uniform-Convergence Fourier Transform Estimation,” in *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, Oct. 2018, pp. 1019–1023.

[62] B. Sankur and L. A. Gerhardt, “Reconstruction of signals from nonuniform samples,” 1973.

[63] Y. C. Eldar and A. V. Oppenheim, “Filterbank reconstruction of bandlimited signals from nonuniform and generalized samples,” *IEEE Transactions on Signal Processing*, vol. 48, no. 10, pp. 2864–2875, 2000.

[64] S. Maymon and A. V. Oppenheim, “Sinc interpolation of nonuniform samples,” *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4745–4758, 2011.

[65] M. Pawlak and E. Rafajłowicz, “Quasi-random sampling for signal recovery,” *Nonlinear Analysis: Theory, Methods & Applications*, vol. 71, no. 10, pp. 4357–4363, Nov. 2009.

[66] J. Higgins, “A sampling theorem for irregularly spaced sample points (corresp.),” *IEEE Transactions on Information Theory*, vol. 22, no. 5, pp. 621–622, 1976.

[67] Y.-R. Sun, “Nonuniform bandpass sampling in radio receivers,” PhD Thesis, 2004.

[68] D. L. Donoho, “Compressed sensing,” *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[69] R. G. Baraniuk, “Compressive sensing [lecture notes],” *IEEE signal processing magazine*, vol. 24, no. 4, pp. 118–121, 2007.

[70] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, “Model-based compressive sensing,” *IEEE Transactions on information theory*, vol. 56, no. 4, pp. 1982–2001, 2010.

[71] M. Mishali and Y. C. Eldar, “Sub-nyquist sampling,” *IEEE Signal Processing Magazine*, vol. 28, no. 6, pp. 98–124, 2011.

[72] A. Tarczynski and D. Qu, “Optimal periodic sampling sequences for nearly-alias-free digital signal processing,” in *2005 IEEE International Symposium on Circuits and Systems*, 2005, pp. 1425–1428.

[73] D. Qu and A. Tarczynski, “A novel spectral estimation method by using periodic nonuniform sampling,” in *2007 Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers*, 2007, pp. 1134–1138.

[74] B. I. Ahmad and A. Tarczynski, “A novel sub-Nyquist Fourier transform estimator based on alias-free hybrid stratified sampling,” in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2016, pp. 4473–4477.

[75] R. M. Stewart, “Statistical design and evaluation of filters for the restoration of sampled data,” *Proceedings of the IRE*, vol. 44, no. 2, pp. 253–257, 1956.

[76] L. Young, “Microwave filter design using an electronic digital computer,” *IRE Transactions on Microwave Theory and Techniques*, vol. 7, no. 1, pp. 99–101, 1959.

[77] S. K. Mitra and J. F. Kaiser, *Handbook for digital signal processing*. John Wiley & Sons, Inc., 1993.

[78] L. B. Jackson, *Digital Filters and Signal Processing: With MATLAB® Exercises*. Springer Science & Business Media, 2013.

[79] A. Tarczynski, V. Valimaki, and G. D. Cain, “FIR filtering of nonuniformly sampled signals,” in *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Apr. 1997, vol. 3, pp. 2237–2240 vol.3.

[80] E. Allier, G. Sicard, L. Fesquet, and M. Renaudin, “A new class of asynchronous A/D converters based on time quantization,” in *Ninth International Symposium on Asynchronous Circuits and Systems, 2003. Proceedings.*, 2003, pp. 196–205.

[81] F. Aeschlimann, E. Allier, L. Fesquet, and M. Renaudin, “Asynchronous FIR filters: towards a new digital processing chain,” in *10th International Symposium on Asynchronous Circuits and Systems, 2004. Proceedings.*, 2004, pp. 198–206.

[82] S. Mian Qaisar, L. Fesquet, and M. Renaudin, “Adaptive rate sampling and filtering based on level crossing sampling,” *EURASIP Journal on Advances in Signal Processing*, vol. 2009, pp. 1–12, 2009.

[83] B. Bidégaray-Fesquet and L. Fesquet, “A fully nonuniform approach to FIR filtering,” in *SAMPTA’09-International Conference on Sampling Theory and Applications*, 2009, pp. 129–1.

[84] W. B. Ye and Y. J. Yu, “An efficient FIR filtering technique for processing non-uniformly sampled signal,” in *2015 IEEE International Conference on Digital Signal Processing (DSP)*, Jul. 2015, pp. 182–186.

[85] M. Vetterli, P. Marziliano, and T. Blu, “Sampling signals with finite rate of innovation,” *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, Jun. 2002.

[86] S. R. Baker, N. Bloom, S. J. Davis, K. J. Kost, M. C. Sammon, and T. Viratyosin, “The Unprecedented Stock Market Impact of COVID-19,” National Bureau of Economic Research, w26945, Apr. 2020.

[87] X. Wei and P. L. Dragotti, “Sampling piecewise smooth signals and its application to image up-sampling,” in *2015 IEEE International Conference on Image Processing (ICIP)*, 2015, pp. 4293–4297.

[88] H. Y. Darawsheh and A. Tarczynski, “Filtering Nonuniformly Sampled Grid-Based Signals,” in *2018 4th International Conference on Frontiers of Signal Processing (ICFSP)*, Sep. 2018, pp. 56–60.

[89] H. Y. Darawsheh and A. Tarczynski, “FIR Filtering of Discontinuous Signals: A Random-Stratified Sampling Approach,” in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2020, pp. 5800–5804.

[90] H. Y. Darawsheh and A. Tarczynski, “Antithetical stratified sampling estimator for filtering signals with discontinuities,” *Signal Processing*, vol. 181, p. 107910, Apr. 2021.