

# UC San Diego

## UC San Diego Electronic Theses and Dissertations

### Title

Random sampling estimates of fourier transforms : antithetical stratified Monte Carlo

### Permalink

<https://escholarship.org/uc/item/4rb242mv>

### Author

Vadrevu, Aditya M.

### Publication Date

2008

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

**Random Sampling Estimates of Fourier Transforms:  
Antithetical Stratified Monte Carlo**

A thesis submitted in partial satisfaction of the requirements for the degree  
Master of Science

in

Electrical Engineering

by

Aditya M. Vadrevu

Committee in charge:

Professor Elias Masry, Chair  
Professor Laurence Milstein  
Professor Bhaskar Rao

2008



The dissertation of Aditya M. Vadrevu is approved, and it is  
acceptable in quality and form for publication on microfilm:

---

---

---

Chair

University of California, San Diego

2008

## DEDICATION

To mother, who is my life and my all. To Parul, a gem of a person who I met in San Diego, and who continues restore my faith in humanity. To many wonderful gems from my undergraduate days who have made my journey so special since I left home : Rahil, Ipshita, Vivek, Shanil, Nitin B. To all at UC San Diego who have given me the most fun filled and enriching times thus far : Anup, Ankit, Nikhil Rasiwasia, Nitin Gupta.

## TABLE OF CONTENTS

Signature Page . . . . .	iii
Dedication . . . . .	iv
Table of Contents . . . . .	v
List of Figures . . . . .	vi
Acknowledgements . . . . .	vii
Abstract . . . . .	viii
Chapter 1: Introduction . . . . .	1
Chapter 2: A Class of Antithetical Sampling Schemes and Its Performance . . . . .	3
Chapter 3: Joint Asymptotic Normality . . . . .	12
Chapter 4: Numerical Results . . . . .	15
Appendix A: Proof of Theorem 1 . . . . .	20
Appendix B: Numerical Results . . . . .	22
References . . . . .	33

## LIST OF FIGURES

Figure 1: Mean-Square Errors of the Antithetical estimator, $\lambda = 0$ . . . . .	23
Figure 2: Mean-square errors of the Antithetical estimator, $\lambda = 1$ . . . . .	24
Figure 3: Mean-square errors of the Antithetical estimator, $\lambda = 3$ . . . . .	25
Figure 4: Comparing exact estimation errors of Antithetical and Regular estimators for the same $N$ , $\lambda = 1$ . . . . .	26
Figure 5: Comparing exact estimation errors of Antithetical and Regular estimators for the same number of sampling points, $\lambda = 1$ . . . . .	27
Figure 6: Optimal design densities $h^*(t, \lambda)$ . . . . .	28
Figure 7: Global optimal design density $h_{av}^*(t)$ . . . . .	29
Figure 8: Ratio $R(\lambda)$ of mean-square estimation errors for optimal design to uniform design . . . . .	30
Figure 9: Bandpass signal, comparison of mean-square estimation errors for $N = 2$ . . . . .	31
Figure 10: Bandpass signal, comparison of mean-square estimation errors for $N = 10$ . . . . .	32

## ACKNOWLEDGEMENTS

I gratefully acknowledge the guidance and patience of Professor Elias Masry, without whom this work would not have been possible.

ABSTRACT OF THE THESIS

**Random Sampling Estimates of Fourier Transforms:  
Antithetical Stratified Monte Carlo**

by

Aditya M. Vadrevu

Master of Science in Electrical Engineering

University of California, San Diego, 2008

Professor Elias Masry, Chair

This work estimates the Fourier transform of continuous-time signals on the basis of  $N$  discrete-time nonuniform observations. We introduce a class of antithetical stratified random sampling schemes and we obtain the performance of the corresponding estimates. We show that when the underlying function  $f(t)$  has a continuous second-order derivative, the rate of mean square convergence is  $1/N^5$ , which is considerably faster than the rate of  $1/N^3$  for stratified sampling and the rate of  $1/N$  for standard Monte Carlo integration. In addition, we establish joint asymptotic normality for the real and imaginary parts of the estimate and give an explicit expression for the asymptotic covariance matrix. The theoretical results are illustrated by examples for lowpass and highpass signals.

# Chapter 1:

## Introduction

This thesis considers the estimation of Fourier transforms of square integrable deterministic functions on the basis of discrete-time observations taken at appropriately chosen random sampling points. This work is a continuation of prior works on the subject: Standard Monte Carlo integration was considered in (15) and regular stratified sampling was considered in (12). It was noted in (12) that the mean-square estimation error for standard Monte Carlo estimates has a rate of convergence of  $1/N$  where  $N$  is the sample size and that this rate cannot be improved even if the function  $f(t)$  is smooth. In contrast, it was shown in (12) that regular stratified sampling for functions with one continuous derivative the rate of mean-square convergence is  $1/N^3$ . In this paper we consider a modified sampling scheme, using antithetical sampling points, and show that the rate of mean-square convergence is  $1/N^5$  for functions with two continuous derivatives. The idea of antithetical sampling is due to (7). In addition, we establish joint asymptotic normality for the estimates and determine the explicit expression for the asymptotic covariance matrix. For integral of random processes, (6) treats the case of antithetical sampling where mean-square convergence is considered (no asymptotic normality result was established there). The relationship to alias-free sampling of random processes and the spectral estimation of continuous-time processes from randomly-sampled observations are discussed in (12) (see in particular (11) and (9)). We mention

here that there has been extensive works in the engineering literature on the subject of randomized sampling (3) as a method for digital alias-free signal processing (DASP), which was developed by Bilinskis (3) (4) (5) and investigated by other researchers, in particular (13) (2) (15). The reader is directed to these works for further details.

The organization of the paper is as follows: In Chapter 2, we investigate the mean-square estimation error for a class of antithetical stratified sampling schemes. We show that these estimates will outperform regular stratified sampling for any  $f(t)$ , any  $N$ , and any frequency  $\lambda$ . We further show that if  $f(t)$  has a continuous second-order derivative, then the rate of mean-square convergence is  $1/N^5$ . We provide exact expressions for the bias and variance. We further optimize over the class of sampling schemes in order to obtain the best performance. In Chapter 3 we establish the joint asymptotic normality of the real and imaginary parts of the estimates for large sample size  $N$ . This shows that for large sample size  $N$ , the estimation error is approximately Gaussian. In Chapter 4 we provide numerical results for both low-pass and high-pass signals.

## Chapter 2:

# A Class of Antithetical Sampling Schemes and Its Performance

Let  $f(t)$  be a deterministic real valued function with finite energy. Its Fourier transform is given by

$$F(\lambda) := \int_{-\infty}^{\infty} e^{-it\lambda} f(t) dt.$$

We use the mathematical notation  $:=$  to mean that the left side is *defined* by the right side. If  $f(t)$  is observed over the interval  $[0, T]$ , and  $w(t)$  is an averaging window, one would like to obtain the Fourier transform

$$F_w(\lambda) := \int_0^T e^{-it\lambda} f(t) w(t) dt. \quad (1)$$

The properties of different windows can be found in (1). The integral in (1) can be approximated by the sum

$$\hat{F}_w(\lambda) := \sum_{j=1}^N e^{-it_j\lambda} f(t_j) w(t_j) \Delta_j \quad (2)$$

where  $t_j$  are the sampling points and  $\Delta_j = t_j - t_{j-1}$ . Note that (2) is an estimate of  $F_w(\lambda)$ , not of  $F(\lambda)$  (the selection of  $T$  and an appropriate window  $w(t)$  is a purely deterministic problem: Essentially, given  $f(t)$ , one chooses  $T$  and a window  $w(t)$  such that

$F_w(\lambda)$  is close to  $F(\lambda)$ ). We now introduce our sampling scheme and the corresponding estimates. The estimate is based on  $2N$  random samples of the function, obtained as follows: Let  $0 = \tau_{N,0} < \tau_{N,1} < \dots < \tau_{N,N} = T$  be a partition of the observation interval  $[0, T]$ , defined by a continuous, strictly positive, probability density function  $h(t)$  on  $[0, T]$  such that

$$\int_0^{\tau_{N,j}} h(t) dt = \frac{j}{N}, \quad j = 0, 1, \dots, N. \quad (3)$$

For  $j = 1, \dots, N$ , set

$$A_{N,j} := [\tau_{N,j-1}, \tau_{N,j}), \quad \Delta\tau_{N,j} := \tau_{N,j} - \tau_{N,j-1}. \quad (4)$$

Note that  $h(t) = 1/T$  on  $[0, T]$  yields an equally spaced partition of the interval  $[0, T]$  (in which case  $\tau_{N,j} = j(T/N)$  and  $\Delta\tau_{N,j} = T/N$ ). We shall demonstrate later how the quality of the estimate can be improved by selecting an optimal design density  $h(t)$ . The sampling points  $\{t_{N,j}\}_{j=1}^N$  are selected in the following manner: The random variables  $\{t_{N,j}\}$  are independent such that  $t_{N,j}$  is uniformly distributed on the subinterval  $A_{N,j}$ . Denote the point *antithetical* to  $t_{N,j}$  by  $t'_{N,j}$ :

$$t'_{N,j} := 2c_{N,j} - t_{N,j} \quad (5)$$

where  $c_{N,j}$  is the midpoint of the subinterval  $A_{N,j}$ ,

$$c_{N,j} := \frac{\tau_{N,j} + \tau_{N,j-1}}{2} \quad j = 1, \dots, N. \quad (6)$$

For simplicity of analysis, we define

$$g(t) := e^{-it\lambda} f(t) w(t). \quad (7)$$

The estimate of the Fourier transform is then given by

$$\hat{F}_w(\lambda) := \sum_{j=1}^N \left( \frac{g(t_{N,j}) + g(t'_{N,j})}{2} \right) \Delta\tau_{N,j}. \quad (8)$$

Note that in the case of an equally spaced partition,  $h(t) = 1/T$  on  $[0, T]$ ,

$$\hat{F}_w(\lambda) = \frac{T}{N} \sum_{j=1}^N \left( \frac{g(t_{N,j}) + g(t'_{N,j})}{2} \right).$$

Our first result shows that the estimate (8) is unbiased and we obtain an expression for its variance for every  $N \geq 1$ . The proof is given in the Appendix.

**Theorem 1**     *i.*  $E[\hat{F}_w(\lambda)] = F_w(\lambda)$ .

*ii.*  $\text{Var} \left[ \hat{F}_w(\lambda) \right] =$

$$\sum_{j=1}^N \left( \frac{\Delta\tau_{N,j}}{2} \int_{A_{N,j}} \left[ |f(t)w(t)|^2 + \cos((2t - 2c_{N,j})\lambda) f(t)w(t)f(2c_{N,j} - t)w(2c_{N,j} - t) \right] dt - \left| \int_{A_{N,j}} e^{-it\lambda} f(t)w(t)dt \right|^2 \right). \quad (9)$$

We will show later that the variance (9) is always upper bounded by the variance of the regular stratified random sampling estimate considered in (12) for the same value of  $N$ . The variance (9) clearly depends on the density  $h(t)$  via the partitions  $\{A_{N,j}\}$ . We now determine the exact rate of decay of the variance. Note that in the following derivations  $g(t)$  defined in (7) depends on the frequency  $\lambda$ , which is assumed to be arbitrary but fixed. Set

$$Z_{N,j} := \frac{g(t_{N,j}) + g(t'_{N,j})}{2} \Delta\tau_{N,j} \quad (10)$$

Then,

$$E[Z_{N,j}] = \frac{1}{2} \int_{A_{N,j}} g(t)dt + \frac{1}{2} \int_{A_{N,j}} g(2c_{N,j} - t)dt. \quad (11)$$

We now assume that  $f(t)w(t)$  has two continuous derivatives and so is  $g(t)$ . Then, we can expand  $g(t)$  ( $t \in A_{N,j}$ ) about  $c_{N,j}$  in a Taylor series as follows:

$$g(t) = g(c_{N,j}) + g'(c_{N,j})(t - c_{N,j}) + \frac{1}{2}g''(c_{N,j})(t - c_{N,j})^2 + o(|t - c_{N,j}|^2). \quad (12)$$

With  $t' = 2c_{N,j} - t$ , it follows that

$$\frac{g(t) + g(t')}{2} = g(c_{N,j}) + \frac{1}{4}g''(c_{N,j})[(t - c_{N,j})^2 + (t' - c_{N,j})^2] + o(|t - c_{N,j}|^2). \quad (13)$$

The term involving  $g'(c_{N,j})$  drops out since  $t - c_{N,j} + t' - c_{N,j} = 0$ . Also  $(t' - c_{N,j})^2 = (t - c_{N,j})^2$ . Thus,

$$E[Z_{N,j}] = \int_{A_{N,j}} \left( g(c_{N,j}) + \frac{1}{2}g''(c_{N,j})(t - c_{N,j})^2 + o(|t - c_{N,j}|^2) \right) dt.$$

We now note that

$$\begin{aligned} a_k &:= \int_{A_{N,j}} (t - c_{N,j})^k dt = \int_{-\frac{\Delta\tau_{N,j}}{2}}^{\frac{\Delta\tau_{N,j}}{2}} y^k dy \\ &= \begin{cases} 0, & k \text{ odd} \\ \frac{1}{2^{k(k+1)}}(\Delta\tau_{N,j})^{k+1}, & k \text{ even} \end{cases} \end{aligned} \quad (14)$$

and similarly

$$b_k := \int_{A_{N,j}} |t - c_{N,j}|^k dt = \frac{1}{2^k(k+1)}(\Delta\tau_{N,j})^{k+1} \quad (15)$$

Using (14)-(15),

$$E[Z_{N,j}] = g(c_{N,j})\Delta\tau_{N,j} + \frac{1}{24}g''(c_{N,j})(\Delta\tau_{N,j})^3 + o((\Delta\tau_{N,j})^3). \quad (16)$$

It now follows from (12)-(16) that

$$\begin{aligned} Z_{N,j} - E[Z_{N,j}] &= \frac{\Delta\tau_{N,j}}{2}g''(c_{N,j}) \left[ (t_{N,j} - c_{N,j})^2 - \frac{1}{12}(\Delta\tau_{N,j})^2 \right] \\ &\quad + \left\{ \Delta\tau_{N,j} o(|t_{N,j} - c_{N,j}|^2) + o((\Delta\tau_{N,j})^3) \right\}. \end{aligned} \quad (17)$$

The second moment of (17) gives the variance of  $Z_{N,j}$ ,

$$\begin{aligned} \text{Var}[Z_{N,j}] &= \frac{\Delta\tau_{N,j}}{4} |g''(c_{N,j})|^2 \int_{A_{N,j}} \left[ (t - c_{N,j})^2 - \frac{1}{12}(\Delta\tau_{N,j})^2 \right]^2 dt \\ &\quad + \Delta\tau_{N,j} \int_{A_{N,j}} o(|t - c_{N,j}|^4) dt + o((\Delta\tau_{N,j})^6) + \text{cross term}. \end{aligned} \quad (18)$$

The cross term can be handled using the Cauchy-Schwarz inequality. Evaluating the integral in the first term on the right side of (18) yields (using (14)-(15))

$$J := \int_{A_{N,j}} \left[ (t - c_{N,j})^2 - \frac{1}{12}(\Delta\tau_{N,j})^2 \right]^2 dt = (\Delta\tau_{N,j})^5 \left( \frac{1}{80} - \frac{1}{72} + \frac{1}{144} \right) = \frac{4}{720}(\Delta\tau_{N,j})^5.$$

Similarly, it is seen that the second term on the right side of (18) is  $o((\Delta\tau_{N,j})^6)$ . The cross terms are seen to be of order  $o((\Delta\tau_{N,j})^6)$ . Thus

$$\text{Var} [Z_{N,j}] = \frac{1}{720} |g''(c_{N,j})|^2 (\Delta\tau_{N,j})^6 + o((\Delta\tau_{N,j})^6). \quad (19)$$

We now note from (3) and the mean value theorem that

$$\frac{1}{N} = \int_{A_{N,j}} h(t)dt = h(int_j)\Delta\tau_{N,j} \quad (20)$$

where  $int_j$  is an intermediate point in  $A_{N,j}$ . Since we assume  $h(t) > 0$  on  $[0, T]$ , there exists an  $\varepsilon > 0$  such that  $h(t) > \varepsilon$ . Then (20) implies that  $\Delta\tau_{N,j} < (N\varepsilon)^{-1}$  uniformly in  $j$ . Therefore, the second term in (19) is of the order  $o(1/N^6)$ . Since the random variables  $\{Z_{N,j}\}$  are independent, we have

$$\begin{aligned} \text{Var} [\hat{F}_w(\lambda)] &= \sum_{j=1}^N \text{Var} [Z_{N,j}] \\ &= \frac{1}{720N^5} \sum_{j=1}^N \frac{|g''(c_{N,j})|^2 (\Delta\tau_{N,j})}{h^5(int_j)} + o(1/N^5) \end{aligned} \quad (21)$$

Now, by Riemann integration we have

$$\lim_{N \rightarrow \infty} N^5 \text{Var} [\hat{F}_w(\lambda)] = \frac{1}{720} \int_0^T \frac{|g''(t)|^2}{h^5(t)} dt.$$

We have thus established the following result:

**Theorem 2** *Assume that the function  $f(t)w(t)$  has a continuous second-order derivative. Then, the antithetical stratified random sampling estimator (8) satisfies*

$$\lim_{N \rightarrow \infty} (2N)^5 \text{Var} [\hat{F}_w(\lambda)] = C^2(h, \lambda)$$

where

$$C^2(h, \lambda) := \frac{2}{45} \int_0^T \left\{ \frac{1}{h^5(t)} [(fw)''(t) - \lambda^2(fw)(t)]^2 + 4\lambda^2[(fw)'(t)]^2 \right\} dt. \quad (22)$$

Since the estimate (8) is unbiased as shown in Theorem 1, we have the following corollary.

**Corollary 1** *Under the assumptions of the above theorem, the mean-square error of the antithetical random sampling estimator (8) satisfies*

$$\lim_{N \rightarrow \infty} (2N)^5 E \left| \hat{F}_w(\lambda) - F_w(\lambda) \right|^2 = C^2(h, \lambda)$$

where  $C^2(h, \lambda)$  is given by (22).

We remark that the variance of the estimate given above is a function of the frequency  $\lambda$ . One may be interested in its global behavior: Let  $Q(\lambda)$  be a nonnegative weight function satisfying

$$\int_{-\infty}^{\infty} Q(\lambda) d\lambda = 1, \quad r_4 := \int_{-\infty}^{\infty} \lambda^4 Q(\lambda) d\lambda < \infty. \quad (23)$$

Let

$$r_2 := \int_{-\infty}^{\infty} \lambda^2 Q(\lambda) d\lambda.$$

Consider the weighted integrated mean-square error (IMSE)

$$\text{IMSE}(N, h) := \int_{-\infty}^{\infty} Q(\lambda) E \left[ \left| \hat{F}_w(\lambda) - F_w(\lambda) \right|^2 \right] d\lambda.$$

We obtain an asymptotic expression for  $\text{IMSE}(N, h)$ , as follows: First integrate (21) with respect to  $\lambda$  with weight function  $Q(\lambda)$ . We obtain

$$\int_{-\infty}^{\infty} Q(\lambda) \text{Var} \left[ \hat{F}_w(\lambda) \right] d\lambda = \frac{1}{720N^5} \sum_{j=1}^N \frac{|s^2(c_{N,j})|^2}{h^5(\text{int}_j)} (\Delta \tau_{N,j}) + o(1/N^5) \quad (24)$$

where

$$s^2(t) := [(fw)''(t)]^2 - 2r_2 \{ (fw)(t)(fw)''(t) - 2[(fw)'(t)]^2 \} + r_4 (fw)^2(t). \quad (25)$$

Then by using Riemann integration we obtain

**Theorem 3** *Assume that the function  $f(t)w(t)$  has continuous second order derivatives. Then the antithetical stratified random sampling estimator (8) satisfies*

$$\lim_{N \rightarrow \infty} (2N)^5 \text{IMSE}(N, h) = C_{av}^2(h) \quad (26)$$

where

$$C_{av}^2(h) := \frac{2}{45} \int_0^T \left\{ \frac{s^2(t)}{h^5(t)} \right\} dt. \quad (27)$$

with  $s^2(t)$  given by (25).

We now discuss the implications of the above theorems and corollary.

1. The rate of mean square convergence of the antithetical random sampling estimator (8) is precisely  $1/N^5$  for functions that have two continuous derivatives. The rate is valid for all design densities  $h(t)$ . In particular, it holds for an equally-spaced partition  $h(t) = 1/T$  on  $[0, T]$ . The approximation

$$E \left| \hat{F}_w(\lambda) - F_w(\lambda) \right|^2 \simeq \frac{C^2(h, \lambda)}{(2N)^5}$$

holds for moderate values of  $N$ . This is supported by numerical results in the Chapter 4.

2. The asymptotic constant  $C^2(h, \lambda)$  of (22) depends on the frequency  $\lambda$  (the rate of convergence is  $1/N^5$  for each fixed frequency).
3. We now optimize over the density  $h(t)$  to minimize the constant  $C^2(h, \lambda)$  (we get a different optimal design density for each frequency  $\lambda$ ) or minimize the global asymptotic constant  $C_{av}^2(h)$  for all frequencies. Calculus of variations argument, under the constraint  $\int_{-\infty}^{\infty} h(t) dt = 1$ , yields that the optimal density  $h(t)$  that minimizes  $C^2(h, \lambda)$  for each fixed frequency  $\lambda$ , is given by

$$h^*(t, \lambda) = \frac{|g''(t, \lambda)|^{\frac{1}{3}}}{\int_0^T |g''(x, \lambda)|^{\frac{1}{3}} dx}, \quad t \in [0, T] \quad (28)$$

where

$$|g''(t, \lambda)| = \{[(fw)''(t) - \lambda^2(fw)(t)]^2 + 4\lambda^2[(fw)'(t)]^2\}^{1/2}. \quad (29)$$

Similarly the global optimal design density  $h_{av}^*(t)$  is given by

$$h_{av}^*(t) = \frac{|s(t)|^{\frac{1}{3}}}{\int_0^T |s(x)|^{\frac{1}{3}} dx}, \quad t \in [0, T] \quad (30)$$

where  $s^2(t)$  is given by (25).

4. The smallest asymptotic constants while using the optimal design density  $h^*(t, \lambda)$  or  $h_{av}^*(t)$  are given by

$$\begin{aligned} (C^2)^*(\lambda) &= \frac{2}{45} \left( \int_0^T |g''(t, \lambda)|^{\frac{1}{3}} dt \right)^6 \\ (C_{av}^2)^* &= \frac{2}{45} \left( \int_0^T |s(t)|^{\frac{1}{3}} dt \right)^6. \end{aligned} \quad (31)$$

Note that the optimal design density  $h^*(t, \lambda)$  requires the knowledge of the underlying function  $f(t)$ . If  $f(t)$  is unknown, one can choose equally spaced partitions (uniform  $h(t)$ ). Also note that the rate of mean-square convergence (Theorem 2) of the estimate is the same regardless of whether one uses the optimal design density, or the uniform design density; only the asymptotic constant  $C^2(h, \lambda)$  is different.

5. The use of asymptotically optimal design can significantly reduce the value of the mean-square estimation error. If we compare the performance when  $h^*(t, \lambda)$  is used with a uniform partition  $h(t) = 1/T$ , the improvement is given by the ratio of the asymptotic constants:

$$R(\lambda) := \frac{(C^2)^*(\lambda)}{(C^2)(h = \frac{1}{T}, \lambda)} = \frac{\left( \int_0^T |g''(t, \lambda)|^{\frac{1}{3}} dt \right)^6}{T^5 \int_0^T |g''(t, \lambda)|^2 dt}. \quad (32)$$

This will be illustrated in the Chapter 4. Similar conclusions holds when comparing the global constants  $(C_{av}^2)^*$  and  $(C_{av}^2)(h = 1/T)$  which do not depend on  $\lambda$ .

6. We now compare the performance of regular stratified sampling with antithetical stratified sampling. Both estimates are unbiased and hence we compare their variances. For simplicity, assume a uniform partition of  $[0, T]$ . We show that for the same value of  $N$ , we have

$$\text{Var} \left[ \hat{F}_w(\lambda) \right]_{anti.} \leq \text{Var} \left[ \hat{F}_w(\lambda) \right]_{reg.} \quad (33)$$

for any  $f(t)w(t)$ ,  $N$  and  $\lambda$ . By Theorem 1, we have for an equally-spaced partition

$$\begin{aligned} \text{Var} \left[ \hat{F}_w(\lambda) \right]_{anti} &= \frac{T}{2N} \sum_{j=1}^N \int_{A_{N,j}} [|f(t)w(t)|^2 \\ &\quad + \cos((2t - 2c_{N,j})\lambda) f(t)w(t) f(2c_{N,j} - t)w(2c_{N,j} - t)] dt \\ &\quad - \sum_{j=1}^N \left| \int_{A_{N,j}} e^{-it\lambda} f(t)w(t) dt \right|^2 \end{aligned}$$

It is seen that

$$\begin{aligned} &\left| \int_{A_{N,j}} [\cos((2t - 2c_{N,j})\lambda) f(t)w(t) f(2c_{N,j} - t)w(2c_{N,j} - t)] dt \right| \\ &\leq \int_{A_{N,j}} |f(t)w(t) f(2c_{N,j} - t)w(2c_{N,j} - t)| dt \\ &\leq \left\{ \int_{A_{N,j}} (fw)^2(t) dt \int_{A_{N,j}} (fw)^2(2c_{N,j} - t) dt \right\}^{1/2} \\ &= \int_{A_{N,j}} (fw)^2(t) dt \end{aligned}$$

where we have used the Cauchy-Schwarz inequality for integrals. Thus

$$\begin{aligned} \text{Var} \left[ \hat{F}_w(\lambda) \right]_{anti} &\leq \frac{T}{N} \sum_{j=1}^N \int_{A_{N,j}} |f(t)w(t)|^2 dt \\ &\quad - \sum_{j=1}^N \left| \int_{A_{N,j}} e^{-it\lambda} f(t)w(t) dt \right|^2 \\ &= \text{Var} \left[ \hat{F}_w(\lambda) \right]_{reg} \end{aligned}$$

where the last step follows from (12). It should be noted that given the value of  $N$ , the antithetical estimator uses  $2N$  sampling points whereas the stratified sampling estimator uses  $N$  points. It does not appear possible to obtain an analytical comparison of the two estimates when they are using the same number of sampling points. However, this comparison is carried out computationally in the example of Chapter 4.

## Chapter 3:

# Joint Asymptotic Normality

In this chapter we establish the joint asymptotic normality of the real and imaginary parts of the estimate (8) and provide an explicit expression for the covariance matrix of the asymptotic distribution.

**Theorem 4** *Assume that the function  $f(t)w(t)$  has a continuous second-order derivative. Then, the scaled real and imaginary parts of the antithetical stratified estimator (8)*

$$N^{5/2}\Re\left[\hat{F}_w(\lambda) - F_w(\lambda)\right] \quad , \quad N^{5/2}\Im\left[\hat{F}_w(\lambda) - F_w(\lambda)\right]$$

are jointly asymptotically normal with zero means and covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_1^2(\lambda) & \sigma_{12}(\lambda) \\ \sigma_{12}(\lambda) & \sigma_2^2(\lambda) \end{bmatrix}$$

with

$$\begin{aligned} \sigma_1^2(\lambda) &= \frac{1}{720} \int_0^T \frac{[g_1''(t)]^2}{h^5(t)} dt \\ \sigma_2^2(\lambda) &= \frac{1}{720} \int_0^T \frac{[g_2''(t)]^2}{h^5(t)} dt \\ \sigma_{12}(\lambda) &= \frac{1}{720} \int_0^T \frac{[g_1''(t)g_2''(t)]}{h^5(t)} dt \end{aligned}$$

where

$$g_1(t) := \cos(t\lambda)f(t)w(t); g_2(t) := \sin(t\lambda)f(t)w(t). \quad (34)$$

Thus, for large sample size  $N$ , the real and imaginary parts of the estimation error are Gaussian. This allows us to compute the probability of any event involving the estimate. In particular, confidence intervals for the estimate can be computed.

*Proof:* Set

$$S(\lambda) = \sum_{j=1}^N [Z_{N,j} - E[Z_{N,j}]] \quad (35)$$

where  $Z_{N,j} - E[Z_{N,j}]$  is given in (17) but with  $g(t)$  now given by

$$g(t) = a_1 g_1(t) + a_2 g_2(t) \quad (36)$$

where the  $a_i$ 's are arbitrary real numbers. (this is the standard Crámer device for proving multivariate central limit theorem: it is sufficient to prove the asymptotic normality of  $S(\lambda)$ ). The classical central limit theorem does not apply here since the random variables  $\{Z_{N,j}\}$  are independent but not identically distributed. We instead use the Lyapunov condition (14), which requires us to show that

$$I_N := \frac{\sum_{j=1}^N E|Z_{N,j} - E[Z_{N,j}]|^\nu}{\left(\sum_{j=1}^N \text{Var}[Z_{N,j}]\right)^{\frac{\nu}{2}}} \rightarrow 0 \quad \text{as } N \rightarrow \infty \quad (37)$$

for some  $\nu > 2$ . By (17) and the  $c_r$  inequality (10) ( $|a + b|^\nu \leq 2^{\nu-1}(|a|^\nu + |b|^\nu)$ ),

$$\begin{aligned} E[|Z_{N,j} - E[Z_{N,j}]|^\nu] &\leq 2^{\nu-1} \left\{ \frac{(\Delta\tau_{N,j}|g''(c_{N,j})|)^\nu}{2^\nu} E\left[ (t_{N,j} - c_{N,j})^2 - \frac{1}{12}(\Delta\tau_{N,j})^2 \right]^\nu \right. \\ &\quad \left. + E\left[ \left( \Delta\tau_{N,j} o(|t_{N,j} - c_{N,j}|^2) + o((\Delta\tau_{N,j})^3) \right)^\nu \right] \right\} \end{aligned} \quad (38)$$

Applying the  $c_r$  inequality again, we obtain

$$\begin{aligned} E[|Z_{N,j} - E[Z_{N,j}]|^\nu] &\leq \\ &2^{2(\nu-1)} \left\{ \frac{(\Delta\tau_{N,j}|g''(c_{N,j})|)^\nu}{2^\nu} \left( E|t_{N,j} - c_{N,j}|^{2\nu} \right) + \frac{1}{(12)^\nu} (\Delta\tau_{N,j})^{2\nu} \right. \\ &\quad \left. + (\Delta\tau_{N,j})^\nu \left( E[o(|t_{N,j} - c_{N,j}|^{2\nu})] \right) + o((\Delta\tau_{N,j})^{3\nu}) \right\}. \end{aligned} \quad (39)$$

Taking expectation using (14)-(15)

$$\begin{aligned} E[|Z_{N,j} - E[Z_{N,j}]|^\nu] &\leq \\ &\text{const. } |g''(c_{N,j})|^\nu (\Delta\tau_{N,j})^{3\nu} + O((\Delta\tau_{N,j})^{3\nu}) + o((\Delta\tau_{N,j})^{3\nu}) \end{aligned} \quad (40)$$

where  $const.$  is a generic constant depending on  $\nu$ . Noting that by the argument following (20) we have  $(\Delta\tau_{N,j}) = O(1/N)$  uniformly in  $j$  and that  $g''(t)$  is continuous on  $[0, T]$  and thus bounded, we obtain

$$\sum_{j=1}^N E|Z_{N,j} - E[Z_{N,j}]|^\nu \leq \frac{const.}{N^{3\nu-1}} + O\left(\frac{1}{N^{3\nu-1}}\right) \leq \frac{const}{N^{3\nu-1}}. \quad (41)$$

For the denominator of (37), we need precise behavior. Since

$$\sum_{j=1}^N \text{Var}[Z_{N,j}] = \text{Var}[S(\lambda)],$$

we have as in the proof of Theorem 2 that

$$\lim_{N \rightarrow \infty} N^5 \text{Var}[S(\lambda)] = C_S(h, \lambda)$$

where

$$C_S(h, \lambda) := \frac{1}{720} \int_0^T \frac{|g''(t)|^2}{h^5(t)} dt \quad (42)$$

and  $g(t)$  is now given by (36). Hence we have for the ratio  $I_N$ ,

$$I_N \leq \text{const.} \frac{1}{N^{(\nu/2)-1}} \rightarrow 0$$

as  $N \rightarrow \infty$  since  $\nu > 2$ . This implies that  $S(\lambda)$  is asymptotically Gaussian with zero mean and asymptotic variance  $C_S(h, \lambda)$ . Since the coefficients  $\{a_i\}$  are arbitrary, the theorem follows and the asymptotic covariance matrix  $\Sigma$  is obtained by identifying the coefficients of  $a_1^2$ ,  $a_2^2$ , and  $a_1 a_2$  in the asymptotic variance  $C_S(h, \lambda)$  in (42). ■

# Chapter 4:

## Numerical Results

In this chapter we provide numerical results illustrating the analytical performance established in the previous chapters. We first remark that the estimators  $\hat{F}_w(\lambda)$  of (8) estimate  $F_w(\lambda)$ , not  $F(\lambda)$  (see statement in Chapter 2 on the selection of  $w(t)$  and  $T$ ). Thus one could simply refer to the product  $f(t)w(t)$  as the function whose Fourier transform is being estimated from an observation of length  $T$ . We establish the following:

- a. The finite sample size performance of the antithetical stratified estimate (8) for moderate values of  $N$  is well approximated by the asymptotic results of Theorem 2.
- b. The optimal design density is quite different from being a uniform density. As a consequence, the improvement in performance over equally-spaced partition could be substantial.

For our first example, we select a lowpass signal and we carry all computations analytically. Let

$$f(t) = \frac{\alpha^3}{2} t^2 e^{-\alpha t}; \quad t \geq 0, \quad \alpha > 0. \quad (43)$$

This function has two continuous derivative on  $(0, \infty)$  which is square integrable. Its Fourier transform is given by  $F(\lambda) = \alpha^3/(\alpha + i\lambda)^3$ . The 3 dB down one-sided bandwidth is  $\lambda_0 = \alpha$ . The Fourier transform over  $[0, T]$  is given by

$$F_w(\lambda) = \frac{F(\lambda)}{2} [2 - e^{-\gamma} e^{-i\lambda T} (2 + 2T(\alpha + i\lambda) + T^2(\alpha + i\lambda)^2)].$$

(in effect  $w(t) = 1$  over  $[0, T]$ ). We select this function for illustration for the following reasons: 1) It has two continuous derivatives which is square integrable. 2) we can compute the variance of the estimate analytically providing us with a closed-form expression. We start with a uniform partition  $h(t) = 1/T$  over  $[0, T]$ . Then  $\tau_{N,j} = jT/N$  and the partition interval  $A_{N,j} = ((j-1)T/N, jT/N)$ . The exact variance of the stratified estimator is given by Theorem 1 with  $\Delta\tau_{N,j} = T/N$ . Computing the integrals involved yields the following results:

$$\frac{T}{2N} \int_0^T f^2(t) dt = \frac{\gamma}{32N} \left\{ 3 - e^{-2\gamma} [2\gamma^4 + 4\gamma^3 + 6\gamma^2 + 6\gamma + 3] \right\}$$

where  $\gamma := \alpha T$ . For the second term, the expression for  $\lambda \neq 0$  is equal to

$$\begin{aligned} & \int_{A_{N,j}} f(t) f((2j-1)T/N - t) \cos(2t\lambda - (2j-1)T\lambda/N) dt \\ &= \frac{\alpha^6}{64} e^{-\alpha(2j-1)T/N} \left[ J_1(\lambda) - 2(T(2j-1)/N)^2 J_2(\lambda) + (T(2j-1)/N)^4 J_3(\lambda) \right] \end{aligned}$$

where  $J_1(\lambda) =$

$$\frac{(T/N)^4 \sin(\lambda T/N)}{\lambda} - \frac{4}{\lambda} \left[ \frac{3\lambda^2 (T/N)^2 - 6}{\lambda^4} \sin(\lambda T/N) - \frac{\lambda^2 (T/N)^3 - 6T/N}{\lambda^3} \cos(\lambda T/N) \right]$$

and

$$\begin{aligned} J_2(\lambda) &= \frac{(T/N)^2}{\lambda} \sin(\lambda T/N) + \frac{2T/N}{\lambda^2} \cos(\lambda T/N) - \frac{2}{\lambda^3} \sin(\lambda T/N) \\ J_3(\lambda) &= \frac{1}{\lambda} \sin(\lambda T/N). \end{aligned}$$

For  $\lambda = 0$  we obtain

$$\begin{aligned} & \int_{A_{N,j}} f(t) f((2j-1)T/N - t) \cos(2t\lambda - (2j-1)T\lambda/N) dt \\ &= \frac{\alpha^6}{32} (T/N)^5 e^{-\alpha(2j-1)T/N} \left[ 0.2 - (2/3)(2j-1)^2 + (2j-1)^4 \right]. \end{aligned}$$

For the last term we have

$$\begin{aligned} & \int_{A_{N,j}} e^{-it\lambda} f(t) dt \\ &= \frac{\alpha^3}{2(\alpha + i\lambda)^3} e^{-(\alpha+i\lambda)(j-1)T/N} \left\{ (\alpha + i\lambda)^2 \left[ ((j-1)T/N)^2 - (jT/N)^2 e^{-(\alpha+i\lambda)T/N} \right] \right. \\ & \left. + 2(\alpha + i\lambda) \left[ (j-1)T/N - (jT/N) e^{-(\alpha+i\lambda)T/N} \right] + 2 \left[ 1 - e^{-(\alpha+i\lambda)T/N} \right] \right\}. \end{aligned}$$

We shall compare this exact expression of the variance of the estimator with the asymptotic expression given by

$$\text{Var}[\hat{F}_w(\lambda)] \simeq \frac{C^2(h = 1/T, \lambda)}{(2N)^5}$$

where  $C^2(\lambda, h = 1/T)$ , given by (22), can be shown to be equal to

$$\begin{aligned} C^2(h = 1/T; \lambda) &:= \frac{2T^5}{45} \int_0^T \{ [f''(t) - \lambda^2 f(t)]^2 + 4\lambda^2 (f'(t))^2 \} dt \\ &= \frac{\gamma^5 \alpha}{90} (U_0 + U_1 + U_2 + U_3 + U_4) \end{aligned}$$

where

$$\begin{aligned} U_0 &= \frac{2}{\alpha} [1 - e^{-2\gamma}] \\ U_1 &= -\frac{4}{\alpha} [1 - e^{-2\gamma}(1 + 2\gamma)] \\ U_2 &= \frac{20\alpha^2 + 12\lambda^2}{4\alpha^3} \left\{ 1 - e^{-2\gamma}(1 + 2\gamma + 2\gamma^2) \right\} \\ U_3 &= -\frac{(\alpha^2 + \lambda^2)}{2\alpha^3} \left\{ 6 - e^{-2\gamma}(6 + 12\gamma + 12\gamma^2 + 8\gamma^3) \right\} \\ U_4 &= \frac{(\alpha^2 + \lambda^2)^2}{4\alpha^5} \left\{ 3 - e^{-2\gamma}(3 + 6\gamma + 6\gamma^2 + 4\gamma^3 + 2\gamma^4) \right\}. \end{aligned}$$

In Figures 1-6 we set  $\alpha = 1$ , so that the 3 dB down one-sided bandwidth of  $|F(\lambda)|$  is then  $\lambda_0 = 1$ . This is just a normalization. We also set  $T = 8$  (for which  $F(\lambda)$  is very close to  $F_w(\lambda)$ ). In Figures 1-3 we compare the exact and asymptotic mean-square errors for the antithetical stratified estimator as a function of  $N$  for 3 values of frequencies  $\lambda = 0$ ,  $\lambda = 1$  and  $\lambda = 3$ . It is evident from these figures that for moderate values of  $N \geq 15$  the exact and asymptotic expressions for the mean-square error of

the antithetical estimator (8) are very close. We remark that  $\lambda = 3$  corresponds to three times the value of  $\lambda_0$  (the 3dB down one-sided bandwidth). Figure 4 compares the exact variance of the antithetical estimator and the regular stratified estimator for  $\lambda = 1$  when both estimators use the same  $N$ . It is seen that the antithetical estimator outperforms the regular stratified estimator by two orders of magnitude for large  $N$ . Figure 5 shows the relative performance for  $\lambda = 1$  when both estimators use the same number of sampling points. Again the antithetical estimator outperforms the regular stratified estimator by more than an order of magnitude for large  $N$ .

In the previous numerical results, we assumed that the design density is uniform over  $[0, T]$ . We now obtain the optimal design density of the antithetical estimator for each fixed frequency. From (28) - (29) we find that

$$h^*(t, \lambda) = \frac{a(t, \lambda)}{\int_0^T a(x, \lambda) dx}$$

where

$$a(t, \lambda) = \frac{\alpha}{2^{1/3}} e^{-\alpha t/3} \left[ [2 - 4\alpha t + (\alpha^2 - \lambda^2)t^2]^2 + 4\lambda^2 t^2 (2 - \alpha t)^2 \right]^{1/6}.$$

In Figure 6 we show these optimal design densities for  $\lambda = 0, 1, 3$ . It is evident that these optimal densities are quite different from a uniform density over  $[0, T]$ . Figure 6 implies that for each  $N$ , the partition intervals  $\{A_{N,j}\}_{j=1}^N$  tend to cluster toward the left end of the interval  $[0, T]$ . For the global optimal density we select a Gaussian  $Q(\lambda)$  with zero mean and unit variance. Then  $r_2 = 1$  and  $r_4 = 3$  in which case the global optimal design density  $h_{av}^*(t)$  is plotted in Fig. 7 and is seen again to be far from uniform. Finally we consider the improvement that is expected when using optimal design densities over a uniform design density. We consider the improvement factor on the basis of the asymptotic expressions given in Theorem 1. Thus we compute the ratios of the asymptotic constants

$$R(\lambda) := \frac{(C^2)^*(\lambda)}{C^2(h = 1/T, \lambda)}$$

and

$$R_{av} := \frac{(C_{av}^2)^*(\lambda)}{C_{av}^2(h = 1/T)}.$$

In Figure 8, we set  $\alpha = 1$  and plot  $R(\lambda)$  as a function of  $\gamma = \alpha T$  for three values of  $\lambda = 0, 1, 3$ . It is seen that optimal design can provide significant improvement over uniform design for every value of  $\gamma > 0$  and that this improvement increases with  $\gamma$  (about two orders of magnitude for large  $\gamma$ ). The improvement factor is largest for  $\lambda = 0$  and smaller for  $\lambda > 0$ .

Next we consider high frequency signals and compare the performance of the antithetical estimator with that of regular stratified estimate. Let  $F(\lambda)$  be given by

$$F(\lambda) = \begin{cases} 1, & |\lambda \pm \lambda_0| \leq B \\ 0, & \text{otherwise} \end{cases} \quad (44)$$

where  $\lambda_0$  is the center frequency and  $B$  is the one-sided bandwidth. The corresponding function  $f(t)$  is given by

$$f(t) = \frac{2B \sin Bt}{\pi Bt} \cos \lambda_0 t. \quad (45)$$

This signal is infinitely differentiable and was also used in (12). We select the center frequency  $\lambda_0 = 10^9$  rad/sec and the one-sided bandwidth  $B$  to be half a percent of  $\lambda_0$ ,  $B = 0.5 \times 10^7$  rad/sec. This is therefore a very high frequency signal. Note that the envelope function  $\sin(Bt)/(Bt)$  has its first zero at  $\pi/B$  so we select  $T = 2(\pi/B) = 4\pi \times 10^{-7}$  sec in order to capture most of the energy of the signal. We select the window function  $w(t) = 0.5(1 + \cos(\pi t/T))$  corresponding to the Hann window which yields a smooth  $F_w(\lambda)$  fairly close to  $F(\lambda)$ . For both the antithetical sampling and regular stratified estimators we select equally-spaced partition ( $h(t) = 1/T$ ) and the sampling point  $t_{N,j}$  is uniformly distributed over the subinterval  $A_{N,j}$ . Because of the highly oscillatory nature of the signal, computations of the variance as given in Theorem 1 requires careful numerical integration. The results for both sampling schemes are displayed in Figures 9 (for  $N = 2$ ) and in Figure 10 (for  $N = 10$ ) over a frequency range four times larger than the bandwidth. It is seen that the antithetical sampling estimator outperforms the regular stratified estimator for all frequencies displayed in Figures 9 and 10. These results are consistent with the analytical observations made in Remark 6 in Chapter 2.

# Appendix A:

## Proof of Theorem 1

*Proof of Theorem 1:*

$$E \left[ \hat{F}_w(\lambda) \right] = E \left[ \sum_{j=1}^N \left( \frac{g(t_{N,j}) + g(t'_{N,j})}{2} \right) \Delta\tau_{N,j} \right].$$

Since  $t_{N,j}$  is uniformly distributed on  $A_{N,j}$ , it is clear from (5) that  $t'_{N,j}$  is also uniformly distributed on  $A_{N,j}$ . Then,

$$\begin{aligned} E \left[ \hat{F}_w(\lambda) \right] &= \sum_{j=1}^N E [g(t_{N,j}) \Delta\tau_{N,j}] \\ &= \sum_{j=1}^N \int_{A_{N,j}} g(t) dt \\ &= \sum_{j=1}^N \int_{A_{N,j}} e^{-it\lambda} f(t) w(t) dt \\ &= \int_0^T e^{-it\lambda} f(t) w(t) dt = F_w(\lambda). \end{aligned}$$

Now since the summands are independent,

$$\begin{aligned} \text{Var} \left[ \hat{F}_w(\lambda) \right] &= \text{Var} \left[ \sum_{j=1}^N \frac{g(t_{N,j}) + g(t'_{N,j})}{2} \Delta\tau_{N,j} \right] \\ &= \sum_{j=1}^N \frac{(\Delta\tau_{N,j})^2}{4} \text{Var} [g(t_{N,j}) + g(t'_{N,j})] \end{aligned} \quad (46)$$

where  $t'_{N,j} := 2c_{N,j} - t_{N,j}$ ,  $j = 0, 1, \dots, N$ .

$$\begin{aligned} \text{Var} [g(t_{N,j}) + g(t'_{N,j})] &= E \left[ |g(t_{N,j}) + g(t'_{N,j})|^2 \right] - \left| E [g(t_{N,j}) + g(t'_{N,j})] \right|^2 \\ E \left[ |g(t_{N,j}) + g(t'_{N,j})|^2 \right] &= E \left[ |g(t_{N,j})|^2 \right] + E \left[ |g(t'_{N,j})|^2 \right] + 2E \left[ \Re [g(t_{N,j})g^*(t'_{N,j})] \right] \end{aligned}$$

which we compute as

$$\begin{aligned} &\frac{1}{\Delta\tau_{N,j}} \int_{A_{N,j}} |f(t)w(t)|^2 dt + \frac{1}{\Delta\tau_{N,j}} \int_{A_{N,j}} |f(2c_{N,j} - t)w(2c_{N,j} - t)|^2 dt \\ &+ \frac{1}{\Delta\tau_{N,j}} \int_{A_{N,j}} 2\Re \left\{ e^{-i(2t-2c_{N,j})\lambda} f(t)f(2c_{N,j} - t)w(t)w(2c_{N,j} - t) \right\} dt \quad . \end{aligned}$$

The second term in the above expression is identical to the first term (change of variable:  $u = 2c_{N,j} - t$ ). Thus

$$\begin{aligned} E \left[ |g(t_{N,j}) + g(t'_{N,j})|^2 \right] &= \frac{2}{\Delta\tau_{N,j}} \int_{A_{N,j}} |f(t)w(t)|^2 dt + \\ &\frac{2}{\Delta\tau_{N,j}} \int_{A_{N,j}} \cos((2t - 2c_{N,j})\lambda) f(t)w(t)f(2c_{N,j} - t)w(2c_{N,j} - t) dt. \quad (47) \end{aligned}$$

We proved the following fact while verifying that the estimator is unbiased:

$$E [g(t_{N,j}) + g(t'_{N,j})] = \frac{2}{\Delta\tau_{N,j}} \int_{A_{N,j}} g(t) dt. \quad (48)$$

Using (47) and (48) in (46) yields the desired result:

$$\begin{aligned} \text{Var} \left[ \hat{F}_w(\lambda) \right] &= \sum_{j=1}^N \left( \frac{\Delta\tau_{N,j}}{2} \int_{A_{N,j}} \left[ |f(t)w(t)|^2 + \right. \right. \\ &\quad \left. \left. \cos((2t - 2c_{N,j})\lambda) f(t)w(t)f(2c_{N,j} - t)w(2c_{N,j} - t) \right] dt \right. \\ &\quad \left. - \left| \int_{A_{N,j}} e^{-it\lambda} f(t)w(t) dt \right|^2 \right). \quad (49) \end{aligned}$$

■

## **Appendix B:**

### **Numerical Results**

We now provide in this appendix, plots illustrating the analytical performance of the antithetical estimator that was established in this study.

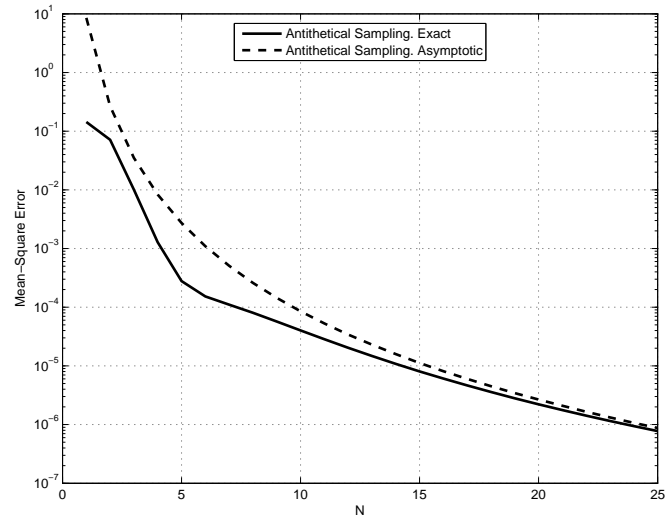


Figure 1: Mean-square errors of the Antithetical estimator,  $\lambda = 0$

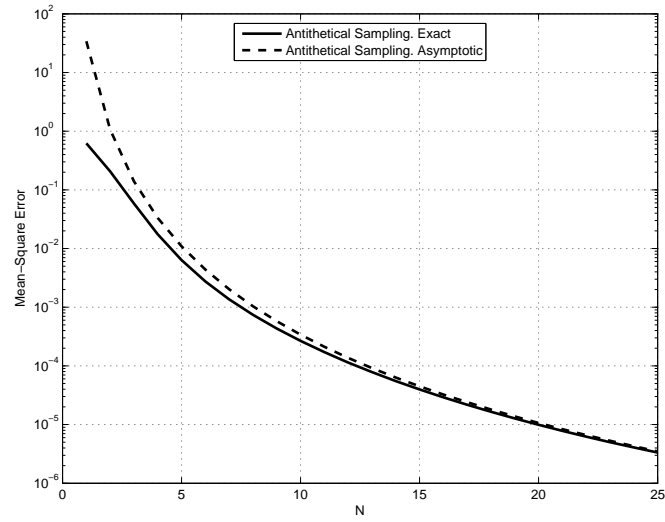


Figure 2: Mean-square errors of the Antithetical estimator,  $\lambda = 1$

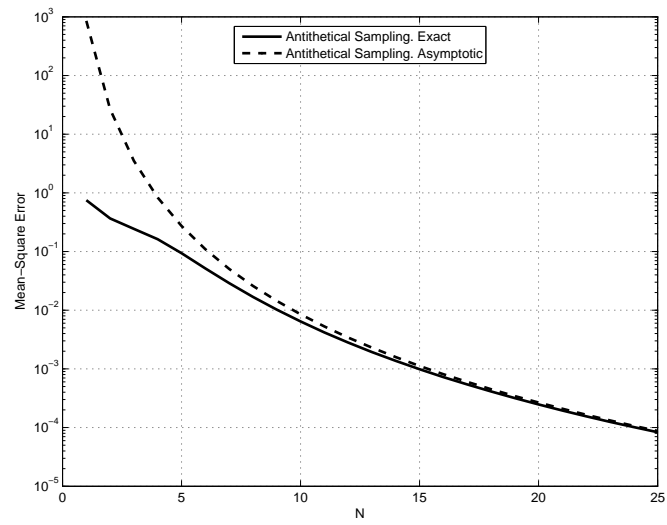


Figure 3: Mean-square errors of the Antithetical estimator,  $\lambda = 3$

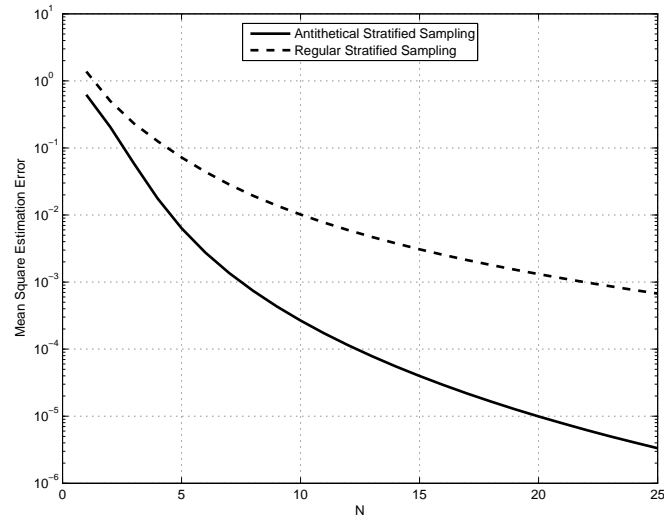


Figure 4: Comparing exact estimation errors of Antithetical and Regular estimators for the same  $N$ ,  $\lambda = 1$

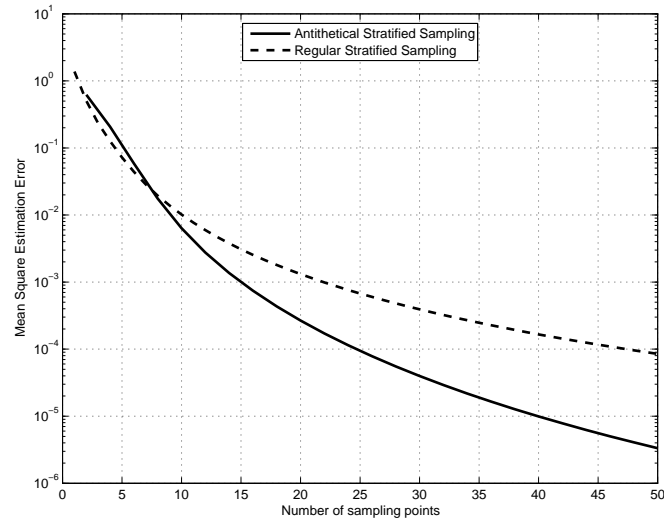


Figure 5: Comparing exact estimation errors of Antithetical and Regular estimators for the same number of sampling points,  $\lambda = 1$

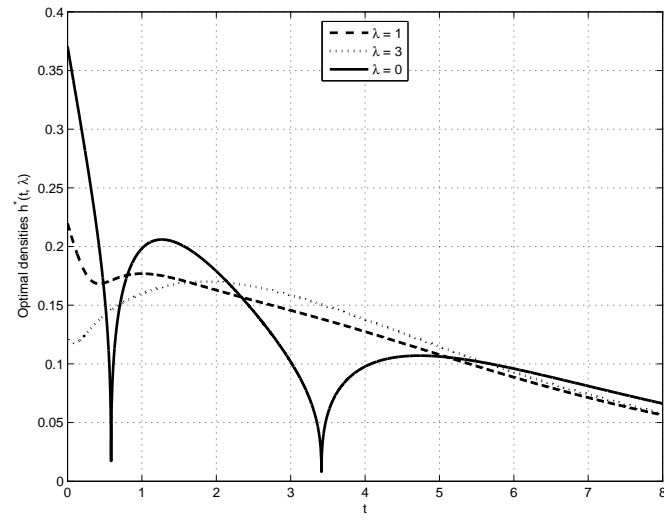


Figure 6: Optimal design densities  $h^*(t, \lambda)$

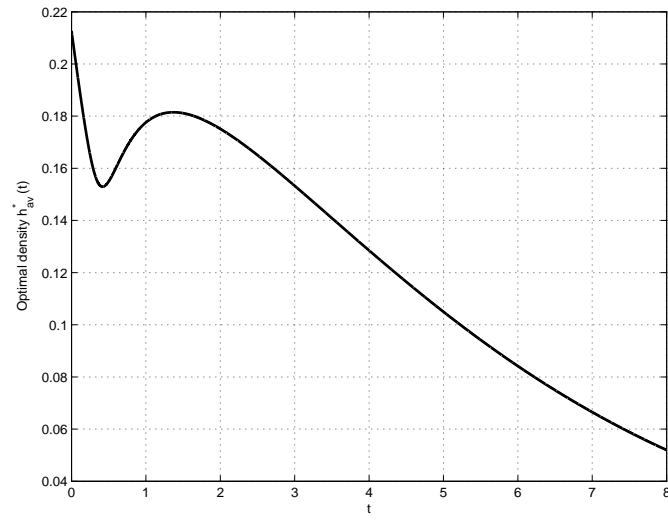


Figure 7: Global optimal design density  $h_{av}^*(t)$

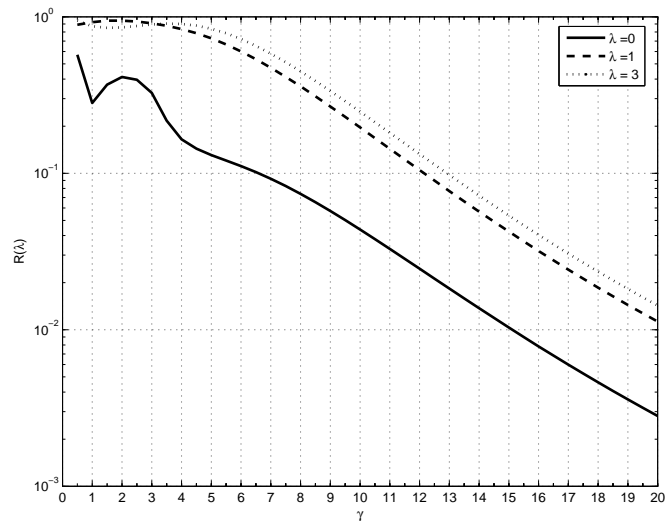


Figure 8: Ratio  $R(\lambda)$  of mean-square estimation errors for optimal design to uniform design

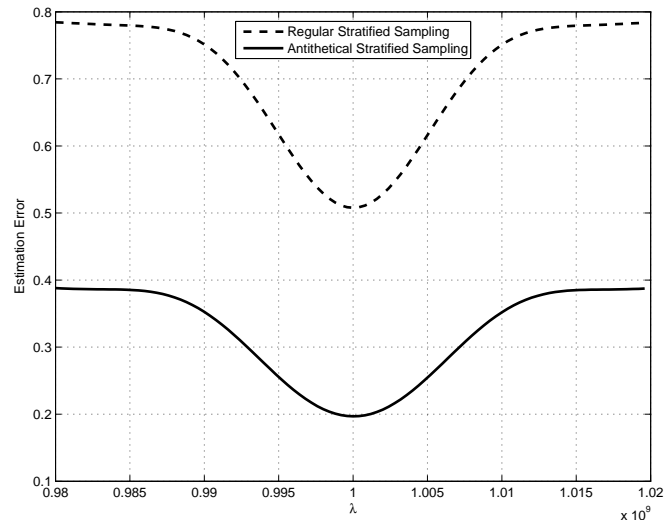


Figure 9: Bandpass signal, comparison of mean-square estimation errors for  $N = 2$

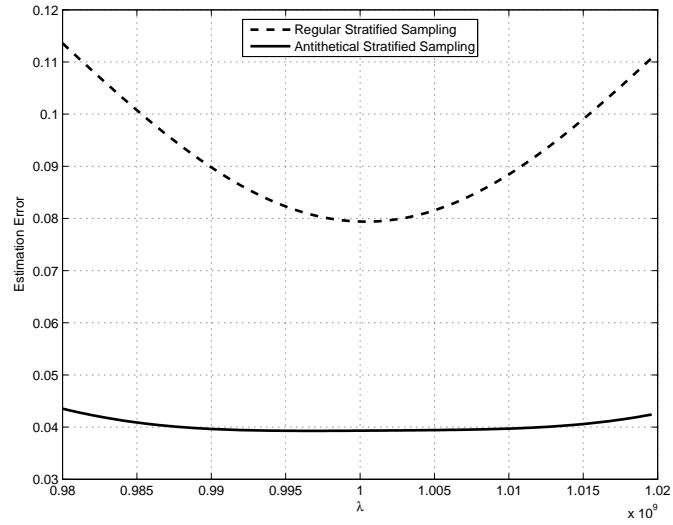


Figure 10: Bandpass signal, comparison of mean-square estimation errors for  $N = 10$

# References

- [1] A. Antoniou. *Digital Filters: Analysis, Design and Applications*. New York: McGraw Hill, 1993.
- [2] Y. Artyukh, A. Ribakov, and V. Vedin, "Evaluation of pseudorandom sampling processes," *Proc. Workshop Sampling Theory Applics.*, 1997, pp. 361-363.
- [3] I. Bilinskis and M. Mikelsons. *Randomized Signal Processing*. London, UK: Prentice Hall, 1992.
- [4] I. Bilinskis and M. Mikelsons "Applications of randomized or irregular sampling as an anti-aliasing technique", in *Signal Processing V: Theories and Applications*. L. Torres, E. Masgrau, and M. A. Lagunas, Eds. New York: Elsevier, 1990, pp. 357-360.
- [5] I. Bilinskis and G.D. Cain, "Digital alias-free signal processing in the GHz frequency range," in *IEE Colloq. Advanced Signal Process. Microwave Applics.*, Nov. 1996, Dig. 1996/236, pp. 6/1-6/6.
- [6] S. Cambanis and E. Masry, "Trapezoidal stratified Monte Carlo integration," *SIAM J. Numerical Analysis*, Vol. 29 (1992), pp. 284-301.
- [7] S. Haber, "A modified Monte-Carlo quadrature. II," *Mathematics of Computation*, vol. 21 (1967), pp. 388-397.
- [8] J.M. Hammersley and D.C. Handscomb. *The Monte Carlo Method*. London, UK: Methuen, 1964.

- [9] K.S. Lii and E. Masry, "Spectral estimation of continuous-time stationary processes from random samples," *Stochastic Processes and their Applics.*, vol. 52 (1993), pp. 39-64.
- [10] M. Loeve, *Probability Theory*. Princeton, N.J.: Van Nostrand Company, 1963.
- [11] E. Masry "Alias-free sampling: An alternative conceptualization and its applications", *IEEE Transaction of Information Theory*, IT-24 (1978), pp. 317-324.
- [12] E. Masry, "Random sampling of deterministic signals: Statistical analysis of Fourier transforms estimates," *IEEE Trans. Signal Processing*, vol. 54 (2006), pp. 1750-1761.
- [13] I. Mednieks "Methods for spectral analysis of non uniformly sampled signals," *Proc. Workshop Sampling Theory Applics.*, 1999, pp. 190-193.
- [14] R.J. Serfling. *Approximation Theorems of Mathematical Statistics*. New York, NY: Wiley, 1980.
- [15] A. Tarczynski and N. Allay, "Spectral analysis of randomly sampled signals: suppression of aliasing and sampler jitter," *IEEE Trans. Signal Processing*, vol. SP-52 (December 2004), pp. 3324-3334.