ANTITHETICAL RANDOM SAMPLING: STATISTICAL ANALYSIS OF FOURIER TRANSFORMS ESTIMATORS

Elias Masry and Aditya Vadrevu

Department of Electrical and Computer Engg.
University of California
San Diego, USA
Email: masry@ece.ucsd.edu, avadrevu@ucsd.edu

Abstract. We consider the estimation of the Fourier transform of continuous-time signals from a finite set N of discrete-time nonuniform observations. We introduce a class of antithetical stratified random sampling schemes and we obtain the performance of the corresponding estimates. For functions $f(t)$ with two continuous derivatives, we show that the rate of mean square convergence is $1/N^5$, which is considerably faster that 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. The theoretical results are illustrated by examples for lowpass and highpass signals.

Keywords Fourier transforms estimates, non-uniform sampling, rates of mean-square convergence, asymptotic normality.

1. INTRODUCTION

In this paper we establish the statistical properties of discrete-time estimates of the Fourier transform of deterministic functions using antithetical random sampling schemes. Standard Monte Carlo integration was considered in [12] and regular stratified sampling was considered in [9]. It was shown in [9] that regular stratified sampling for functions with one continuous derivative, the rate of mean-square convergence is $1/N^3$. In this paper we employ antithetical sampling schemes based on $2N$ sampling points and show that the rate of mean-square convergence is $1/N^5$ for functions with two continuous derivatives. In addition, we establish joint asymptotic normality for the estimates and determine the explicit expression for the asymptotic covariance matrix. This result could be used to obtain confidence intervals for the estimates. The notion of antithetical sampling is due to Haber [7]. Related work for integrals of random processes can be found in [6]. The relationship of random sampling schemes to the notion of alias-free sampling is thoroughly discussed in [9]. 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 [11] [2] [12]. The reader is directed to these works for further details.

The organization of the paper is as follows: In Section II, we consider the mean-square estimation error for a class of antithetical stratified sampling schemes. We show that these estimates outperform regular stratified sampling for any $f(t)$, any N, and any frequency λ. We further show that if $f(t)$ has a continuous second-oder 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 Section III we establish the joint asymptotic normality of the real and imaginary parts of the estimates for large sample size N. In Section IV we provide numerical results for both low-pass and high-pass signals. Proofs of all the theorems stated in this paper can be found in the full paper [10].

2. A CLASS OF ANTITHETICAL SAMPLING SCHEMES

Consider real-valued functions $f(t)$ with finite energy. The Fourier transform of $f(t)$ is given by

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

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

$$F_w(\lambda) := \int_0^T e^{-it\lambda} f(t)w(t)dt. \tag{2.1}$$

The properties of different windows can be found in [1]. The integral in (2.1) can be approximated by the sum

$$\hat{F}_w(\lambda) := \sum_{j=1}^{N} e^{-it_j\lambda} f(t_j)w(t_j)\Delta_j \tag{2.2}$$

where t_j are the sampling points and $\Delta_j = t_j - t_{j-1}$. Note that (2.2) is an estimate of $F_w(\lambda)$, not of $F(\lambda)$. 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 $0 = \tau_{N,0} < \tau_{N,1} < \ldots < \tau_{N,N} = T$ 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

$$\int_0^{\tau_{N,j}} h(t)dt = \frac{j}{N}, \quad j = 0,1,\ldots,N. \tag{2.3}$$

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

$$A_{N,j} := [\tau_{N,j-1}, \tau_{N,j}), \quad \Delta\tau_{N,j} := \tau_{N,j} - \tau_{N,j-1}. \tag{2.4}$$

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 that the quality of the estimate can be significantly 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 $t'_{N,j}$:

$$t'_{N,j} := 2c_{N,j} - t_{N,j} \tag{2.5}$$

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

$$c_{N,j} \quad := \quad \frac{\tau_{N,j} + \tau_{N,j-1}}{2} \quad j = 1,\ldots,N. \tag{2.6}$$

For simplicity of analysis, we define

$$g(t) := e^{-it\lambda} f(t)w(t). \tag{2.7}$$

The estimate of the Fourier transform is then given by

$$\hat{F}_w(\lambda) := \sum_{j=1}^N \left(\frac{g(t_{N,j}) + g(t'_{N,j})}{2} \right) \Delta\tau_{N,j}. \tag{2.8}$$

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

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

Our first result shows that the estimate (2.8) is unbiased and we obtain an expression for its variance for every $N \geq 1$.

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

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

$$\begin{aligned} &= \sum_{j=1}^N \Bigg(\frac{\Delta\tau_{N,j}}{2} \int_{A_{N,j}} \Big[|f(t)w(t)|^2 \\ &+ \cos((2t - 2c_{N,j})\lambda) f(t)w(t) \\ &\times f(2c_{N,j} - t) w(2c_{N,j} - t) \Big] dt \\ &- \left| \int_{A_{N,j}} e^{-it\lambda} f(t)w(t)dt \right|^2 \Bigg). \end{aligned} \tag{2.9}$$

We will show later that the variance (2.9) is always upper bounded by the variance of the regular stratified random sampling estimate considered in [9] for the same value of N. We now determine the exact rate of decay of the variance.

Theorem 2 *Assume that the function $f(t)w(t)$ has a continuous second-order derivative. Then, the antithetical stratified random sampling estimator (2.8) satisfies*

$$lim_{N\to\infty} (2N)^5 \; Var \left[\hat{F}_w(\lambda)\right] = C^2(h,\lambda)$$

where

$$\begin{aligned} C^2(h,\lambda) := \frac{2}{45} \int_0^T & \left\{ \frac{1}{h^5(t)} \left[(fw)''(t) - \lambda^2 (fw)(t) \right]^2 \right. \\ & \left. + 4\lambda^2 [(fw)'(t)]^2 \right\} dt. \end{aligned} \tag{2.10}$$

Since the estimate (2.8) 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 (2.8) satisfies*

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

where $C^2(h,\lambda)$ is given by (2.10).

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

$$\int_{-\infty}^{\infty} Q(\lambda)d\lambda = 1, \quad r_4 := \int_{-\infty}^{\infty} \lambda^4 Q(\lambda)d\lambda < \infty. \tag{2.11}$$

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 have,

Theorem 3 *Assume that the function $f(t)w(t)$ has continuous second order derivatives. Then the antithetical stratified random sampling estimator (2.8) satisfies*

$$lim_{N\to\infty} (2N)^5 IMSE(N,h) = C^2_{av}(h) \tag{2.12}$$

where

$$C^2_{av}(h) := \frac{2}{45} \int_0^T \left\{ \frac{s^2(t)}{h^5(t)} \right\} dt. \tag{2.13}$$

with $s^2(t)$ given by

$$\begin{aligned} s^2(t) := [(fw)''(t)]^2 - 2r_2\{(fw)(t)(fw)''(t) - 2[(fw)'(t)]^2\} \\ + r_4 (fw)^2(t). \end{aligned} \tag{2.14}$$

The above results have the following implications:

1. The rate of mean square convergence of the antithetical random sampling estimator (2.8) 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
$$E\left|\hat{F}_w(\lambda) - F_w(\lambda)\right|^2 \simeq \frac{C^2(h,\lambda)}{(2N)^5} \qquad (2.15)$$
holds for moderate values of N. This is supported by numerical results in the Section IV.

2. The asymptotic constant $C^2(h, \lambda)$ of (2.10) depends on the frequency λ (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 λ) or minimize the global asymptotic constant $C^2_{av}(h)$ for all frequencies. The optimal density $h(t)$ that minimizes $C^2(h, \lambda)$ for each fixed frequency λ, is given by
$$h^*(t,\lambda) = \frac{|g''(t,\lambda)|^{\frac{1}{3}}}{\int_0^T |g''(x,\lambda)|^{\frac{1}{3}} dx}, \quad t \in [0,T] \quad (2.16)$$
where
$$|g''(t,\lambda)| = \left\{[(fw)''(t) - \lambda^2 (fw)(t)]^2 + 4\lambda^2 [(fw)'(t)]^2\right\}^{1/2}. \qquad (2.17)$$
Similarly the global optimal design density $h^*_{av}(t)$ is given by
$$h^*_{av}(t) = \frac{|s(t)|^{\frac{1}{3}}}{\int_0^T |s(x)|^{\frac{1}{3}} dx}, \quad t \in [0,T] \qquad (2.18)$$
where $s^2(t)$ is given by (2.14).

4. The smallest asymptotic constants while using the optimal design density $h^*(t, \lambda)$ or $h^*_{av}(t)$ are given by
$$\begin{aligned} (C^2)^*(\lambda) &= \frac{2}{45}\left(\int_0^T |g''(t,\lambda)|^{\frac{1}{3}} dt\right)^6 \\ (C^2_{av})^* &= \frac{2}{45}\left(\int_0^T |s(t)|^{\frac{1}{3}} dt\right)^6. \end{aligned} \qquad (2.19)$$
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. Compared to a uniform sampling $h(t) = 1/T$, the improvement is given by the ratio of the asymptotic constants:
$$R(\lambda) := \frac{(C^2)^*(\lambda)}{(C^2)(h=\frac{1}{T},\lambda)} = \frac{\left(\int_0^T |g''(t,\lambda)|^{\frac{1}{3}} dt\right)^6}{T^5 \int_0^T |g''(t,\lambda)|^2 dt}. \qquad (2.20)$$
This will be illustrated in the Section IV. Similar conclusions hold when comparing the global constants $(C^2_{av})^*$ and $(C^2_{av})(h = 1/T)$ which do not depend on λ.

6. Comparing the performance of regular stratified sampling with antithetical stratified sampling. Both estimates are unbiased and hence we compare their variances. Assume a uniform partition of $[0, T]$. Then for the same value of N, we have
$$\text{Var}\left[\hat{F}_w(\lambda)\right]_{anti.} \leq \text{Var}\left[\hat{F}_w(\lambda)\right]_{reg.} \qquad (2.21)$$
for any $f(t)w(t)$, N and λ.

3. JOINT ASYMPTOTIC NORMALITY

We have the following result establishing the joint asymptotic normality of the real and imaginary parts of the estimate (2.8) 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 (2.8)*
$$N^{5/2}\Re\left[\hat{F}_w(\lambda) - F_w(\lambda)\right] \quad , \quad N^{5/2}\Im\left[\hat{F}_w(\lambda) - F_w(\lambda)\right]$$
are jointly asymptotically normal with zero means and covariance matrix specified in [10].

4. NUMERICAL RESULTS

In this section we provide numerical results illustrating the analytical performance established in the previous sections. Because of space limitation, we summarize the conclusions drawn from the ten plots rather than displaying the plots themselves.

For our first example, we select a lowpass signal and we carry all computations analytically. Let
$$f(t) = \frac{\alpha^3}{2} t^2 e^{-\alpha t};\ t \geq 0,\ \alpha > 0. \qquad (4.1)$$
This function has two continuous derivatives on $(0, \infty)$ which is square integrable. We set $\alpha = 1$ and select $T = 8$ for

which $F_w(\lambda)$ is fairly close to $F(\lambda)$ with $w(t) = 1$. We start with a uniform partition $h(t) = 1/T$ over $[0, T]$. The exact variance of the stratified estimator is given by Theorem 1 with $\Delta\tau_{N,j} = T/N$. We compare this exact expression of the variance of the estimator with the asymptotic expression given by (2.15). The plots show that for moderate values of $N \geq 15$ the exact and asymptotic expressions for the mean-square error of the antithetical estimator (2.8) are very close. We also compare the exact variance of the antithetical estimator and the regular stratified estimator for $\lambda = 1$ when both estimators use the same N. The antithetical estimator outperforms the regular stratified estimator by two orders of magnitude for $N \geq 15$.

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. The plot shows that these optimal densities are quite different from a uniform density over $[0, T]$. Further, that for each N, the partition intervals $\{A_{N,j}\}_{j=1}^{N}$ tend to cluster toward the left end of the interval $[0, T]$. Finally we consider the improvement that is expected when using optimal design densities over a uniform design density on the basis of the asymptotic expressions given in Theorem 2. The optimal design provides significant improvement over uniform design for every value of $\gamma := \alpha T > 0$ and that this improvement increases with γ (about two orders of magnitude for large γ). 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

$$F(\lambda) = \begin{cases} 1, & |\lambda \pm \lambda_0| \leq B \\ 0, & \text{otherwise} \end{cases} \tag{4.2}$$

where λ_0 is the center frequency and B is the one-sided bandwidth. The corresponding function $f(t)$ is given by

$$f(t) = \frac{2B}{\pi} \frac{\sin Bt}{Bt} \cos \lambda_0 t. \tag{4.3}$$

This signal is infinitely differentiable. We select the center frequency $\lambda_0 = 10^9$ rad/sec and the one-sided bandwidth B to be half a percent of λ_0, $B = 0.5 \times 10^7$ rad/sec. Note that the envelope function $\sin(Bt)/(Bt)$ has its first zero at π/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 $A_{N,j}$. The plots show that the antithetical sampling estimator outperforms the regular stratified estimator for all frequencies displayed in the plots.

The numerical results of this section are consistent with the analytical observations made in Section II.

5. 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 Colleq. 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] E. Masry, "Random sampling of deterministic signals: Statistical analysis of Fourier transforms estimates," *IEEE Trans. Signal Processing*, vol. 54 (2006), pp. 1750-1761.

[10] E. Masry and A. Vadrevu, "Random sampling estimates of Fourier transforms: Antithetical stratified Monte Carlo," submitted to *IEEE Trans. Signal Processing*, October 2007.

[11] I. Mednieks "Methods for spectral analysis of non uniformly sampled signals," *Proc. Workshop Sampling Theory Applics.*, 1999, pp. 190-193.

[12] 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.