DASP Implementation of Continuous-Time, Finite-Impulse-Response Systems

Andrzej Tarczynski
School of Computer Science and Engineering
University of Westminster
London, UK
https://orcid.org/0000-0002-7931-6133

Hikmat Y. Darawsheh
School of Computer Science and Engineering
University of Westminster
London, UK
w1527684@my.westminster.ac.uk

***Abstract*— Digital Alias-free Signal Processing (DASP) uses random sampling to mitigate aliasing. This paper investigates the use of DASP for realization of continuous-time, linear, time-invariant systems with finite-duration impulse response. We propose a random sampling scheme and suitable processing algorithm to produce an estimator of the target output. The estimator is unbiased, and its variance is guaranteed to converge to zero at least at O(*T*) rate, where *T* is the average distance between consecutive sampling instants. If the input signal and system impulse response are piecewise continuous and satisfy some benign conditions, the convergence rate is at least O(*T*^2). But if they are continuous everywhere, the rate increases to O(*T*^3).**

Keywords—alias-free sampling, alias-free digital signal processing, system realization, random sampling, fast convergence

I. INTRODUCTION

DASP is a branch of DSP aiming at alleviating aliasing – a fundamental constraint of classical DSP. Manifestation of aliasing is that, unless the class of processed signals is suitably restricted, for any deterministic sampling pattern, uniform or not, there exist different signals taking identical values at the sampling instants. In order to avoid aliasing, DASP uses random sampling and tailor-made processing algorithms. In DASP, every sampled signal is random. Sampling schemes suitable for DASP guarantee that these processes differ from each other if the continuous-time originals are different. To give a simple example of avoiding aliasing, consider collecting a single signal sample $x(\tau_0)$, where τ_0 is a random variable with the probability density function (PDF) $f_\tau(\tau_0) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{\tau_0^2}{2}\right)$. The expected value of the discrete time process $x_d(t;\tau_0) = \frac{x(\tau_0)}{f_\tau(\tau_0)}\delta(t-\tau_0)$ is $\mathrm{E}\{x_d(t;\tau_0)\} = \int_{-\infty}^{\infty} x_d(t;\tau)f_\tau(\tau)\mathrm{d}\tau = x(t)$, which proves that in this case different continuous-time signals have different discrete-time counterparts – a necessary and sufficient condition for eliminating aliasing. Moreover, if $g(\cdot)$ is the impulse response of a linear-time invariant system, then $\hat{y}(t;\tau_0) = \frac{x(\tau_0)}{f_\tau(\tau_0)}g(t-\tau_0)$ is an unbiased estimator of its output $y(t)$ when the input is $x(\cdot)$. To prove it we note that $\mathrm{E}\{\hat{y}(t;\tau_0)\} = \int_{-\infty}^{\infty}\hat{y}(t;\tau)f_\tau(\tau)\mathrm{d}\tau = \int_{-\infty}^{\infty}x(\tau)g(t-\tau)\mathrm{d}\tau = y(t)$. This short discussion shows that, in principle, aliasing can be avoided.

Although avoidance of aliasing is a desirable feature of DSP, it does not make DASP a perfect replacement of continuous-time signal processing. Random discrete-time signals are normally observed through their single, time-limited realizations. The results of signal processing are derived from this limited data rather than full knowledge of the random process. Therefore, DASP focuses on constructing estimators of the required outcomes, such as the signal Fourier transform or output of a filter, and assessing their accuracy.

The origins of DASP can be traced back to 1960s when Shapiro and Silverman [1] elaborated on estimating power spectrum density of continuous-time random signals from samples collected at random time instants. The work was expanded and adopted for practical application in several papers including [2]-[3]. While the literature on the use of nonuniform sampling is widespread, here we entirely focus on deployment of random sampling for aliasing avoidance. Particularly, we draw attention to two books devoted to this topic [4]-[5] and an ample paper exploring hardware aspects needed for implementation of DASP systems [6].

The work in this paper is chiefly motivated by the approaches developed for DASP estimation of the Fourier Transforms [7], where the authors proposed an unbiased and consistent DASP estimator of the Fourier Transform of deterministic signals, working in an arbitrary range of frequencies. It was demonstrated there that the clock jitter in the sampler reduces the infinite bandwidth of DASP. The larger the jitter, the narrower is the range of frequencies, within which alias-free signal processing can be performed. Subsequent papers [8]-[11] replaced the total random sampling scheme used in [7] with stratified sampling [8], which led to improving the pointwise convergence rate of the estimator from $O(T)$ to $O(T^3)$, and then with antithetical stratified sampling [9] reaching the rate $O(T^5)$. All three estimators [7]-[9], still had the uniform convergence rate of $O(T)$. This was addressed in [10], where the use of the first-order hybrid sampling produced pointwise and uniform convergence rates $O(T^5)$, and [11] where L-th order hybrid sampling increased both convergence rates to $O(T^{2L+3})$. Ample applications of DASP include communication systems [12], nuclear magnetic resonance (NMR) [13], and image processing [14].

There is very limited literature on system realization in DASP. Some preliminary results can be found in [4] and [5].

This paper focuses on DASP realization of continuous, linear, time-invariant systems defined by their finite-duration impulse response $g(\cdot)$, i.e. $g(t) = 0$ if $t < T_D$ or $t \geq T_H$, and $H = T_H - T_D > 0$. The target output $y(\cdot)$, is $y(t) = \int_{t-T_H}^{t-T_D} x(\tau)g(t-\tau)\mathrm{d}\tau = \int_{t-T_H}^{t-T_D} z(\tau;t)\mathrm{d}\tau$, where $z(\tau;t) = x(\tau)g(t-\tau)$. More generally:

$$y(t) = \int_{\alpha(t)}^{\beta(t)} z(\tau;t)\mathrm{d}\tau = \int_{\Gamma} z(\tau;t)\mathrm{d}\tau \qquad (1)$$

where $\alpha(t) \le t - T_H$, $\beta(t) \ge t - T_D$, and $\Gamma = [\alpha(t), \beta(t))$. We sample $x(\cdot)$ using a simplified version of stratified sampling [8]. In our approach, all strata $\Gamma_k = [kT, (k+1)T)$ have the same length T. Each stratum Γ_k comprises one sampling instant $\tau_k = kT + \hat{\tau}_k$, where $\{\hat{\tau}_k\}_{k=-\infty}^{\infty}$ is a sequence of IID random variables with PDF $f_{\hat{\tau}}(\tau) = \begin{cases} \frac{1}{T} & \text{if } \tau \in [0,T) \\ 0 & \text{if } \tau \notin [0,T) \end{cases}$. We select $\alpha(t)$ and $\beta(t)$ as close to each other as possible, and aligned with the strata borders. Let $\alpha(t) = k_b(t)T$ and $\beta(t) = \big(k_f(t)+1\big)T$, where the integers $k_b(t)$ and $k_f(t)$ are indexes of the first and last stratum over which the integral (1) is calculated. To ensure $k_b(t)T \le t - T_H$ and $\big(k_f(t)+1\big)T \ge t - T_D$ we select $k_b(t) = \left\lfloor \frac{t-T_H}{T} \right\rfloor$ and $k_f(t) = \left\lceil \frac{t-T_D}{T} \right\rceil - 1$, where $\lfloor\cdot\rfloor$ and $\lceil\cdot\rceil$ denote the floor and ceiling functions respectively. The integral (1) becomes

$$y(t) = \sum_{k=k_b(t)}^{k_f(t)} \int_{\Gamma_k} z(\tau;t)\mathrm{d}\tau = \sum_{k=k_b(t)}^{k_f(t)} I_k \qquad (2)$$

The number of components in (2) is $K = k_f(t) - k_b(t) + 1$. Since $0 \le \frac{t-T_H}{T} - k_b(t) < 1$ and $0 \le k_f(t) + 1 - \frac{t-T_D}{T} < 1$ we get $0 \le k_f(t) - k_b(t) + 1 - \frac{H}{T} < 2$ and $K = \frac{H}{T} + \Lambda$, where $0 \le \Lambda < 2$. We use $\hat{I}_k = Tz(\tau_k;t)$ to approximate the integrals I_k in (2) and define DASP system realization:

$$y_{DASP}(t) = \sum_{k=k_b(t)}^{k_f(t)} \hat{I}_k = T \sum_{k=k_b(t)}^{k_f(t)} x(\tau_k) g(t - \tau_k) \qquad (3)$$

In the remainder of this paper, we explore the properties of system realization (3). We prove that $y_{DASP}(t)$ is an unbiased and consistent estimator of $y(t)$. We provide explicit measures of its accuracy and investigate how it changes when the stratum length goes to zero. Then we present a few numerical examples demonstrating properties of the DASP realizations. Finally, we summarize the work and outline our plans for future research.

II. Properties of DASP System Realization

A. Causality

The system (3) is causal iff for any t the sampling instants τ_k used for calculating $y_{DASP}(t)$ satisfy $\tau_k \le t$. Theorem 1 provides a necessary and sufficient condition for the causality.

Theorem 1: The DASP system described by (3) is causal iff $T_D \ge T$.

Proof: It follows from (3) that $\tau_{k_f(t)}$ is the largest τ_k used for calculating $y_{DASP}(t)$. Since $\tau_{k_f(t)}$ is a random variable with the upper bound $\left\lceil \frac{t-T_D}{T} \right\rceil T$, the system (3) is causal iff for any t:

$$\left\lceil \frac{t - T_D}{T} \right\rceil \le \frac{t}{T} \qquad (4)$$

Let the integers N_t and N_D satisfy $t = N_t T + \delta_t$ and $T_D = N_D T + \delta_D$, where $\delta_t, \delta_D \in [0,T)$. By substituting these for t and T_D in (4), we get $\left\lceil \frac{\delta_t - \delta_D}{T} \right\rceil - N_D \le \frac{\delta_t}{T}$. If $\delta_t \le \delta_D$ we have $-N_D \le \frac{\delta_t}{T}$, implying $N_D \ge 0$. But if $\delta_t > \delta_D$ then $1 - N_D \le \frac{\delta_t}{T}$, hence $N_D \ge 1$. To ensure that (4) is satisfied for any t we conclude $N_D \ge 1$, which is equivalent to $T_D \ge T$.

□

B. Bias

Theorem 2 below states that $y_{DASP}(\cdot)$ is an unbiased estimator of $y(\cdot)$.

Theorem 2: For any t

$$\mathrm{E}\{y_{DASP}(t)\} = y(t)$$

Proof: It follows from (2) and (3) that Theorem 2 holds if

$$\mathrm{E}\{\hat{I}_k\} = I_k \qquad (5)$$

In fact:

$\mathrm{E}\{\hat{I}_k\} = \mathrm{E}\{Tx(kT + \hat{\tau}_k) g(t - kT - \hat{\tau}_k)\} =$

$\int_0^T f_{\hat{\tau}}(\tau) Tx(kT + \tau) g(t - kT - \tau)\mathrm{d}\tau =$

$\int_0^T x(kT + \tau) g(t - kT - \tau)\mathrm{d}\tau = \int_{\Gamma_k} x(\tau) g(t - \tau)\mathrm{d}\tau = I_k.$

□

C. Consistency

Since $y_{DASP}(\cdot)$ is an unbiased estimator of $y(\cdot)$, the necessary and sufficient condition for its consistency is that its variance goes to zero when $T \to 0$. Theorem 3 states that this actually is true.

Theorem 3: If $Y = \int_{\Gamma} z^2(\tau;t)\mathrm{d}\tau$ is finite then

$$\lim_{T \to 0} \sigma^2\{y_{DASP}(t)\} = 0$$

Proof: Since the sampling instants are independent from each other:

$$\sigma^2\{y_{DASP}(t)\} = \sum_{k=N-k_1}^{N-k_2} \sigma^2\{\hat{I}_k\} \qquad (6)$$

By (5): $\sigma^2\{\hat{I}_k\} = \mathrm{E}\{\hat{I}_k^2\} - I_k^2 \le \mathrm{E}\{\hat{I}_k^2\} = T \int_{\Gamma_k} z^2(\tau;t)\mathrm{d}\tau$. Therefore, $\sigma^2\{y_{DASP}(t)\} \le T \sum_{k=N-k_1}^{N-k_2} \int_{\Gamma_k} z^2(\tau;t)\mathrm{d}\tau = T \int_{\Gamma} z^2(\tau;t)\mathrm{d}\tau = TY$. Since $\sigma^2\{y_{DASP}(t)\}$ cannot be negative, we conclude

$$0 \le \sigma^2\{y_{DASP}(t)\} \le TY \qquad (7)$$

It is clear from (7) that when $T \to 0$ the variance of $y_{DASP}(t)$ converges to zero.

□

D. Rate of convergence

Formula (7) implies that the convergence rate of $y_{DASP}(t)$ to $y(t)$, in the sense of the size of its variance, is not slower than $O(T)$. But if $z(\cdot\,;t)$ satisfies some conditions and T is sufficiently short, this rate is faster. Let $\dot{z}(\tau;t) = \frac{d}{d\tau} z(\tau;t)$. We say that $z(\cdot\,;t) \in \mathcal{C}$ if $z(\cdot\,;t)$ and $\dot{z}(\cdot\,;t)$ are bounded, and $\dot{z}(\cdot\,;t)$ is continuous everywhere in Γ except a finite number of points. Fig. 1 shows a sample $z(\cdot\,;t) \in \mathcal{C}$ and its derivative. At $\tau = A$, $z(\cdot\,;t)$ is continuous but $\dot{z}(\cdot\,;t)$ is not. At $\tau = B$ both functions are discontinuous.

Theorems 4 and 5 state that if $z(\cdot\,;t) \in \mathcal{C}$ then the convergence rate of $y_{DASP}(t)$ is at least $O(T^2)$. If, in addition, $z(\cdot\,;t)$ is continuous in Γ then the convergence rate is $O(T^3)$. Let $d_m, m = 1, \dots, M$ be the sequence of all points where $z(\cdot\,;t)$ is discontinuous in Γ. We define the jump $\Delta(d_m)$ as $\Delta(d_m) = \lim_{\tau \to d_m +} z(\tau;t) - \lim_{\tau \to d_m -} z(\tau;t)$.

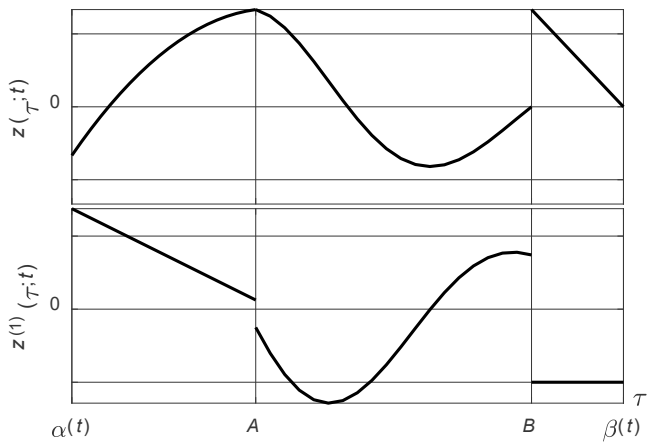

Fig. 1. Sample $z(\cdot\,;t) \in \mathcal{C}$ and its derivative $\dot{z}(\cdot\,;t)$

Theorem 4: If $z(\cdot\,;t) \in \mathcal{C}$ then

$$\limsup_{T\to 0} \frac{\sigma^2_{DASP}\{y_{DASP}(t)\}}{T^2} \le \frac{1}{4}\sum_{m=1}^{M} \Delta^2(d_m)$$

Theorem 5: If $z(\cdot\,;t) \in \mathcal{C}$ is continuous in Γ then

$$\lim_{T\to 0} \frac{\sigma^2_{DASP}\{y_{DASP}(t)\}}{T^3} = \frac{1}{12}\int_{t-T_H}^{t-T_D} \dot{z}^2(\tau;t)\mathrm{d}\tau$$

The proof of both Theorems is presented in the Appendix.

III. Numerical Examples

A. Example 1

We test DASP realization of a system whose impulse response is

$$g(t) = 10^6 \mathrm{sinc}\,(10^6 t - 6) \times 1_{[10^{-6},\ 11\times 10^{-6}]}(t)$$

where $1_X(t)$ is an indicator function of the set X. The system is implemented as described in this paper using strata of length $T = 0.008\mu$s, and then tested on an input signal

$$x(t) = \sum_{n=0}^{19} c_n e^{-\pi\left(t-0.5\times 10^{-6} n\right)^2} + e(t)$$

where the coefficients c_n were selected randomly using Gaussian, $N(0,1)$, random number generator, and then fixed for all repeated experiments. The disturbance $e(t)$ is a sinusoid whose frequency comes from a wide, theoretically infinite, range. In this scenario, any use of a DSP filter with regular sampling could result in aliasing. In our tests, we used $e(t) = 0.2\cos(2.3537 \times 10^9 t)$. Owing to the use of random sampling, the outputs $y_{DASP}(t)$ differ from one experiment to another. To observe the level of variations we present the results of ten independently run simulations. We also show the target output of the filter, $y(t)$, obtained by convolving the signals $g(t)$ and $x(t)$ as well as the results of filtering $x(t)$ using a DSP discrete-time filter defined by

$$y_{DSP}(t) = T\sum_{k=k_b(t)}^{k_f(t)} x(kT)g(t-kT)$$

Fig. 2 shows the results of the tests. Since some plots overlap each other, the inset shows a zoomed-in version of their small fragment. Additionally, Fig 3. Shows the errors $\epsilon_{DASP}(t) = y_{DASP}(t) - y(t)$ and $\epsilon_{DSP}(t) = y_{DSP}(t) - y(t)$.

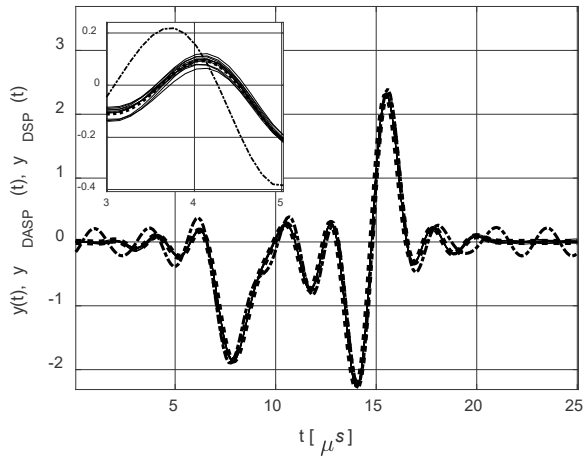

Fig. 2. The target output of the system $y(t)$ (thick dotted line), ten realizations of $y_{DASP}(t)$ (thin continuous lines) and $y_{DSP}(t)$ (medium dashdot line)

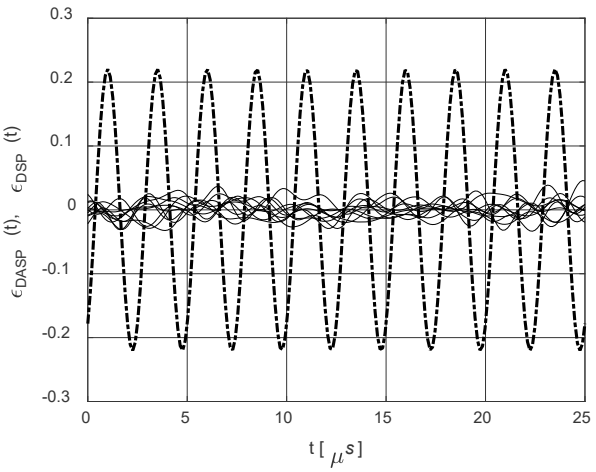

Fig. 3. The output estimation errors of ten realizations of $y_{DASP}(t)$ (thin continuous lines) and of $y_{DSP}(t)$ (medium dashdot line)

It is clear from these plots, that DASP estimator provides much more accurate estimate of the output signal of the system than the DSP filter. Although the results for the DASP system differ from case to case, they are consistently similar. The implementations of DASP and DSP systems were prepared without prior knowledge of the disturbing sinusoid, which put the DSP solution in a disadvantage making it susceptible to aliasing, as observed in Fig. 3.

B. Example 2

In the second example, we test the convergence rates when $T \to 0$. Consider a system with the impulse response

$$g(t) = e^{-(t-0.1)} 1_{[0.1,2.1]}(t)$$

and input

$$x(t) = \sin(2\pi t)$$

We use DASP to estimate the output at $t = 2.35$. Its target value is $y(2.35) = \int_{0.25}^{2.25} \sin(2\pi\tau) e^{\tau-2.25}\,\mathrm{d}\tau = \frac{1-e^{-2}}{(2\pi)^2+1} \approx 0.0214$. The plot of $z(\tau\,;2.35)$ on Fig. 4 reveals discontinuities at 0.25 and 2.25. To illustrate how they affect the convergence rate, we present two versions of this example. In each case we test thirteen different strata lengths taken from

$[10^{-3}, 10^{-1}]$. First, the lengths are $T = \frac{0.25}{n}$, where n takes whole values. This makes both discontinuities fall on the strata borders, masking their existence. Next, we select $T = \frac{0.25}{n+0.5}$, which places both discontinuities exactly in the middle of their strata, maximizing their adverse effect on the size of the estimator variance. See the Appendix part "Case $k \in S_3$" for the justification of these claims. For each value of T, we derive an approximated variance of the estimator using the outcomes of 1000 simulations. We show the results in Fig. 5.

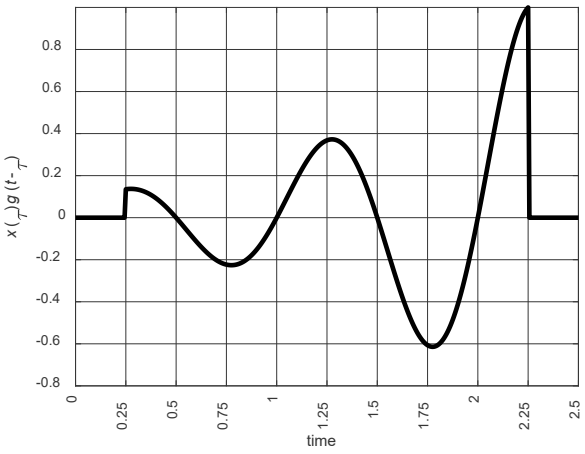

Fig. 4 The plot of $z(\cdot\,;t)$ considered in Example 2

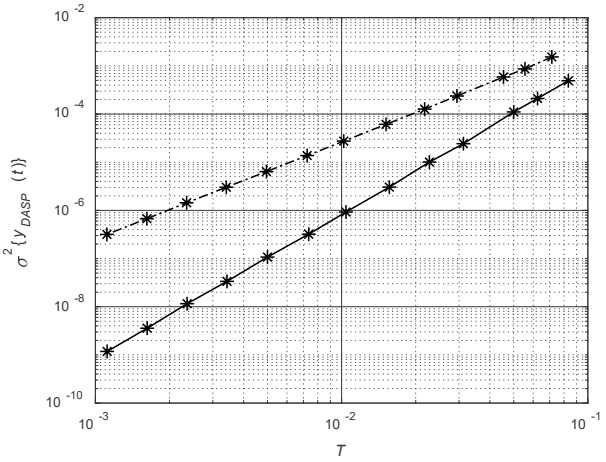

Fig 5. The relation between the stratum length T and $\sigma^2_{DASP}\{y_{DASP}(t)\}$, when the discontinuities of $z(\cdot\,;2.35)$ are obscured (continuous line), and when they are present in the strata centers (dashdot line)

The outcomes of this example confirm the theses of Theorems 4 and 5. When $z(\cdot\,;2.35)$ is continuous (in this case they fall on the strata borders), the convergence rate is $O(T^3)$, but when $z(\cdot\,;2.35)$ is not continuous the rate drops to $O(T^2)$.

IV. Conclusions

We presented preliminary results on DASP realization of continuous time systems defined by their impulse response. The proposed realizations are alias-free. The output of the DASP system is an unbiased estimator of the target output, of analogue system. The variance of the estimator can be pre-assessed using the theoretical results from this paper. Its size can be controlled by selecting a suitable sampling scheme. Although in the presented analyses no constraints on the bandwidths of the impulse response or the input signal were imposed, we expect that, similarly to [7], technical imperfections of system implementation would introduce such limits. The sampling method and processing algorithm used here are similar to those in [8]. However, we have proven a stronger result by showing that in order to achieve the $O(T^3)$ convergence rate, only the integrated function has to be continuous - but not its derivative. In addition, we showed a new result that if the integrated function is piecewise continuous, the convergence rate is not worse than $O(T^2)$ rather than reaching the default $O(T)$. Future work will focus on testing other sampling schemes offering potentially faster convergence rates, and investigation into realization of systems with infinite impulse response.

V. Appendix – Proof of Theorems 4 and 5

First, we show a useful property of the Riemann integral $I = \int_a^b f(t)\mathrm{d}t$. The standard approach to calculating I, is to divide $[a, b)$ into K_n subintervals $T_{k,n}$ of lengths $\langle T_{k,n}\rangle$ satisfying: $\lim_{n\to\infty} \max_k\{\langle T_{k,n}\rangle\} = 0$, $\cup_{k=1}^{K_n} T_{k,n} = [a, b)$, and $k_1 \neq k_2 \Rightarrow T_{k_1,n} \cap T_{k_2,n} = \emptyset$. Next, we arbitrarily select $t_{k,n} \in T_{k,n}$ and calculate $I = \lim_{n\to\infty} \sum_{k=1}^{K_n} \langle T_{k,n}\rangle f(t_{k,n})$. We observe that if $f(\cdot)$ is bounded, i.e. $|f(t)| < D$ then at any stage n, we can remove from the Riemann summation a fixed number of components (say Π) and the sum of the remaining components still converges to I. In fact, let $\mathcal{N}_n$ be a set of Π indexes denoting the components removed from the n^{th} summation, $S_{1,n} = \sum_{k\notin\mathcal{N}_n} \langle T_{k,n}\rangle f(x_{k,n})$ and $S_{2,n} = \sum_{k\in\mathcal{N}_n} \langle T_{k,n}\rangle f(x_{k,n})$. Then $I = \lim_{n\to\infty} S_{1,n} + \lim_{n\to\infty} S_{2,n}$. Since $|S_{2,n}| = |\sum_{k\in\mathcal{N}_n} \langle T_{k,n}\rangle f(x_{k,n})| \le \max_k\{\langle T_{k,n}\rangle\}\, \Pi D$, therefore $\lim_{n\to\infty} S_{2,n} = 0$, and $\int_a^b f(t)\mathrm{d}t = \lim_{n\to\infty} \sum_{k\notin\mathcal{N}_n} \langle T_{k,n}\rangle f(x_{k,n})$.

Now, we are ready to prove Theorems 4 and 5. Without loss of generality, we assume that the strata length T is so short that each stratum contains no more than one point where $\dot{z}(\cdot\,;t)$ is discontinuous. Let $\mathcal{S}_1$ comprises indexes of those strata where $\dot{z}(\cdot\,;t)$ is continuous, $\mathcal{S}_2$ those where $\dot{z}(\cdot\,;t)$ is not continuous but $z(\cdot\,;t)$ is, and $\mathcal{S}_3$ where $z(\cdot\,;t)$ is not continuous, hence containing one of the d_m points. The numbers of strata in sets $\mathcal{S}_2$ and $\mathcal{S}_3$ do not change with T and are denoted by L and M respectively. Set $\mathcal{S}_1$ comprises $P = K - L - M = \frac{H}{T} + \Lambda - L - M$ strata. Hence $P = \frac{1}{O(T)}$. For each stratum, we select $c_k \in \Gamma_k$. If $k \in \mathcal{S}_1$ then $c_k = (k + 0.5)T$. Otherwise, c_k is the discontinuity point of $\dot{z}(\cdot\,;t)$. In the strata belonging to $\mathcal{S}_2$ or $\mathcal{S}_3$ we form

$$z(\tau\,;t) = \begin{cases} z_1(\tau\,;t) & \text{if } \tau < c_k \\ z_2(\tau\,;t) & \text{if } \tau \ge c_k \end{cases} \tag{8}$$

where $z_1(\cdot\,;t)$ and $z_2(\cdot\,;t)$ have continuous first derivatives in Γ_k. We calculate the variances of $I_{DASP,k}$ for different strata sets and use (6) to obtain $\sigma^2_{DASP}\{y_{DASP}(t)\}$. Within each stratum Γ_k, we form Taylor expansions of $z(\cdot\,;t)$ (or $z_1(\cdot\,;t)$ and $z_2(\cdot\,;t)$) about c_k. Let $z_k = z(c_k;t)$, $\dot{z}_k = \dot{z}(c_k;t)$, and, if needed, $z_{l,k} = z_l(c_k;t)$, $\dot{z}_{l,k} = \dot{z}_l(c_k;t)$, $l \in \{1,2\}$. The respective expansions for $\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{S}_3$ are:

$$z(\tau;t) = z_k + (\tau - c_k)\dot{z}_k + o(\tau - c_k) \tag{9}$$

$$z(\tau\,;t) = z_k + \begin{cases} (\tau - c_k)\dot{z}_{1,k} & \text{if } \tau < c_k \\ (\tau - c_k)\dot{z}_{2,k} & \text{if } \tau \ge c_k \end{cases} + o(\tau - c_k) \tag{10}$$

$$z(\tau\,;t)=\begin{cases}z_{1,k} & \text{if } \tau < c_k\\ z_{2,k} & \text{if } \tau \ge c_k\end{cases}+o((\tau-c_k)^0) \qquad (11)$$

Since each expansion is used only within the stratum Γ_k we note $|\tau-c_k|\le T$ and replace $o(\tau-c_k)$ and $o((\tau-c_k)^0)$ in (9) - (11) with $o(T)$ and $o(T^0)$ respectively. For each type of strata we use (9)-(11) to calculate: I_k, $\hat{I}_k$, error $\varepsilon_k=\hat{I}_k-I_k$, and variance $\sigma^2\{I_{DASP,k}\}=\mathrm{E}\{\varepsilon_k^2\}$. We denote $T_1=c_k-kT$ and $T_2=(k+1)T-c_k$.

A. Case $k\in\mathcal{S}_1$

We note $I_k=Tz_k+o(T^2)$, $\hat{I}_k=Tz_k+T(\tau_k-c_k)\dot{z}_k+o(T^2)$, and $\varepsilon_k=T(\tau_k-c_k)\dot{z}_k+o(T^2)$. The dominant part of ε_k is $O(T^2)$. Hence $\varepsilon_k^2=T^2(\tau_k-c_k)^2\dot{z}_k^2+o(T^4)$ and $\sigma^2\{I_{DASP,k}\}=T\int_{\Gamma_k}(\tau-c_k)^2\,\mathrm{d}\tau\,\dot{z}_k^2+o(T^4)=\frac{T^4}{12}\dot{z}_k^2+o(T^4)$. The contribution of all $\mathcal{S}_1$ strata to $\sigma^2_{DASP}\{y_{DASP}(t)\}$ is $\sigma_1^2=\sum_{k\in\mathcal{S}_1}\left(\frac{T^4}{12}\dot{z}_k^2+o(T^4)\right)$. The size of $\mathcal{S}_1$ is $P=\frac{1}{O(T)}$. Hence $\sigma_1^2=\frac{T^3}{12}\sum_{k\in\mathcal{S}_1}T\dot{z}^2(c_k;t)+o(T^3)$. By recalling our earlier observation about Riemann integration, we conclude that

$$\lim_{T\to0}\frac{\sigma_1^2}{T^3}=\frac{1}{12}\int_{t-T_H}^{t-T_D}\dot{z}^2(\tau;t)\mathrm{d}\tau \qquad (12)$$

Since $\sigma_1^2=O(T^3)$ we also note that

$$\limsup_{T\to0}\frac{\sigma_1^2}{T^2}=0 \qquad (13)$$

B. Case $k\in\mathcal{S}_2$

We note $I_k=Tz_k-\frac{T_1^2}{2}\dot{z}_{1,k}+\frac{T_2^2}{2}\dot{z}_{2,k}+o(T^2)$, $\hat{I}_k=Tz_k+\begin{cases}T(\tau_k-c_k)\dot{z}_{1,k} & \text{if } \tau_k<c_k\\ T(\tau_k-c_k)\dot{z}_{2,k} & \text{if } \tau_k\ge c_k\end{cases}+o(T^2)$, and $\varepsilon_k=\begin{cases}\left[T(\tau_k-c_k)+\frac{T_1^2}{2}\right]\dot{z}_{1,k}-\frac{T_2^2}{2}\dot{z}_{2,k} & \text{if } \tau_k<c_k\\ \frac{T_1^2}{2}\dot{z}_{1,k}+\left[T(\tau_k-c_k)-\frac{T_2^2}{2}\right]\dot{z}_{2,k} & \text{if } \tau_k\ge c_k\end{cases}+o(T^2)$. Since $\varepsilon_k\in O(T^2)$, then $\varepsilon_k^2\in O(T^4)$ and $\sigma^2\{I_{DASP,k}\}\in O(T^4)$. The contribution of all $\mathcal{S}_2$ strata to $\sigma^2_{DASP}\{y_{DASP}(t)\}$ is $\sigma_2^2=\sum_{k\in\mathcal{S}_2}\sigma^2\{I_{DASP,k}\}\in O(T^4)$. Consequently

$$\lim_{T\to0}\frac{\sigma_2^2}{T^3}=0 \qquad (14)$$

$$\limsup_{T\to0}\frac{\sigma_2^2}{T^2}=0 \qquad (15)$$

C. Case $k\in\mathcal{S}_3$

For strata in this set we get: $I_k=T_1z_{1,k}+T_2z_{2,k}+o(T)$, $\hat{I}_k=\begin{cases}Tz_{1,k} & \text{if } \tau_k<c_k\\ Tz_{2,k} & \text{if } \tau_k\ge c_k\end{cases}+o(T)$, and $\varepsilon_k=\begin{cases}T_2(z_{1,k}-z_{2,k}) & \text{if } \tau_k<c_k\\ T_1(z_{2,k}-z_{1,k}) & \text{if } \tau_k\ge c_k\end{cases}+o(T)$. Since $\varepsilon_k\in O(T)$ then $\varepsilon_k^2=(z_{2,k}-z_{1,k})^2\times\begin{cases}T_2^2 & \text{if } \tau_k<c_k\\ T_1^2 & \text{if } \tau_k\ge c_k\end{cases}+o(T^2)$. We note that $z_{2,k}-z_{1,k}$ is the jump of $z(\cdot\,;t)$ at c_k, which in turn is one of the d_m points. Therefore, $\varepsilon_k^2=\Delta^2(c_k)\times\begin{cases}T_2^2 & \text{if } \tau<c_k\\ T_1^2 & \text{if } \tau\ge c_k\end{cases}+o(T^2)$. By taking the expected value of ε_k^2 we get $\sigma^2\{I_{DASP,k}\}=T_1T_2\times\Delta^2(c_k)+o(T^2)$. If the discontinuity point is aligned with the stratum border then either $T_1=0$ or $T_2=0$ resulting in $\sigma^2\{I_{DASP,k}\}=o(T^2)$. However, in the worst-case scenario when the discontinuity point is at the center of the stratum: $T_1=T_2=\frac{T}{2}$ and $\sigma^2\{I_{DASP,k}\}=\frac{T^2}{4}\Delta^2(c_k)+o(T^2)$. The contribution of all $\mathcal{S}_3$ strata to $\sigma^2_{DASP}\{y_{DASP}(t)\}$ is $\sigma_3^2\le\frac{T^2}{4}\sum_{k\in\mathcal{S}_3}\Delta^2(c_k)+o(T^2)$, implying

$$\limsup_{T\to0}\frac{\sigma_3^2}{T^2}\le\frac{1}{4}\sum_{k\in\mathcal{S}_3}\Delta^2(c_k)=\frac{1}{4}\sum_{m=1}^{M}\Delta^2(d_m) \qquad (16)$$

By using (13), (15) and (16) we get $\limsup_{T\to0}\frac{\sigma^2_{DASP}\{y_{DASP}(t)\}}{T^2}=\limsup_{T\to0}\frac{\sigma_1^2+\sigma_2^2+\sigma_3^2}{T^2}=\frac{1}{4}\sum_{m=1}^{M}\Delta^2(d_m)$, which completes the proof of Theorem 4. To finalize the proof of Theorem 5 we assume that $z(\cdot\,;t)$ is continuous, which means that the set $\mathcal{S}_3$ is empty. By (12) and (14): $\lim_{T\to0}\frac{\sigma^2_{DASP}\{y_{DASP}(t)\}}{T^3}=\lim_{T\to0}\frac{\sigma_1^2+\sigma_2^2}{T^2}=\frac{1}{12}\int_{t-T_H}^{t-T_D}\dot{z}^2(\tau;t)\mathrm{d}\tau$.

□

REFERENCES

[1] H. S. Shapiro and R. A. Silverman, "Alias-free sampling of random noise," SIAM J. Appl. Math., vol. 8, pp. 225–236, June 1960.

[2] E. Masry, "Alias-free sampling: an alternative conceptualization and its applications," IEEE Trans. Inform. Theory, vol. IT-24, pp. 317–324, May 1978.

[3] F. J. Beutler, "Error-free recovery of signals from irregularly spaced samples," SIAM Review, vol. 8, no. 3, pp. 328-335, July 1966.

[4] I. Bilinskis and A. K. Mikelson, Randomized Signal Processing. Prentice-Hall, Englewood Cliffs, NJ 1992

[5] I. Bilinskis, Digital Alias-free Signal Processing. John Wiley & Sons, Hoboken, NJ 2006.

[6] F. Papenfuss, Y. Artyukh, E. Boole and D. Timmermann, "Nonuniform sampling driver design for optimal ADC utilization", IEEE International Symposium Circuits and Systems, 2003

[7] A. Tarczynski and N. Allay, "Spectral analysis of randomly sampled signals: suppression of aliasing and sampler jitter", IEEE Trans. Signal Processing, vol. 52, no. 12, pp. 3324-3334, December 2004.

[8] E. Masry, "Random sampling of deterministic signals: statistical analysis of Fourier transforms estimates", IEEE Trans. Signal Processing, vol. 54, pp. 1750-1761, 2006.

[9] E. Masry and A. Vadrevu, "Random sampling estimates of Fourier transforms: antithetical stratified Monte Carlo", IEEE Trans. Signal Processing, vol. 57, pp. 149-204, 2009.

[10] A. Tarczynski and B. I. Ahmad, "Estimation of Fourier Transform Using Alias-Free Hybrid-Stratified Sampling", IEEE Trans. Signal Processing, vol. 64, pp. 3065 - 3076, 2016

[11] M. Al-Ani, A. Tarczynski and B. I. Ahmad "High-order hybrid stratified sampling: fast uniform-convergence Fourier transform estimation", 52nd Asilomar Conf. Signals, Systems, and Computers, 2018

[12] B. I. Ahmad and A. Tarczynski, "Spectral analysis of stratified sampling: a means to perform efficient multiband spectrum sensing", IEEE Trans. Wireless Communications, vol. 11, pp. 178 – 187, 2012

[13] K. Kazimierczuk, A. Zawadzka, W. Koźmiński and I. Zhukov, "Random sampling of evolution time space and Fourier transform processing", Journal of Biomolecular NMR, vol. 36, pp. 147-168, 2006

[14] M. Niknejad, J. Bioucas-Dias, and M. A. T. Figueiredo, "External patch-based image restoration using importance sampling", IEEE Trans. on Image Processing, vol. 28, pp. 4460 – 4470, 2019