

# Estimating Random Integrals from Noisy Observations: Sampling Designs and Their Performance

JAMES A. BUCKLEW, MEMBER, IEEE, AND STAMATIS CAMBANIS, SENIOR MEMBER, IEEE

**Abstract**—The problem of estimating a weighted average of a random process from noisy observations at a finite number of sampling points is considered. The performance of sampling designs with optimal or suboptimal, but easily computable, estimator coefficients is studied. Several examples and special cases are studied including additive independent noise, nonlinear distortion with noise, and quantization noise.

## I. INTRODUCTION

THE PROBLEM of estimating a weighted integral of a random process from observations of the process at a finite number of sampling points has been studied by several authors (see the survey [2]). It is an important problem of interest in several areas of communications, information theory, statistics, and signal processing. The usual questions of interest are to find the optimal sampling design of size  $n$ , or sampling designs which are asymptotically optimal as the sample size tends to infinity. Coupled with these is the problem of estimator design and the study of how the mean square estimation error tends to zero as the sample size tends to infinity.

In this paper we consider these problems for the case where the observations are corrupted by noise. We allow the noise to be possibly dependent upon the random process whose integral we are trying to estimate, henceforth called the signal process. In this case, as the number of sampling points increases to infinity, the mean square approximation error no longer tends to zero but instead to some positive least possible value. We consider estimators which use optimal coefficients as well as suboptimal (but simple) coefficients.

As far as the authors are aware, the only case of noisy observations considered in the literature is in [5], [6], where the observation noise is assumed white and the signal Gauss–Markov. The optimal sampling designs are determined in [5], and the rates of convergence of the mean square estimation error are found in [6] to be  $1/n$  with noise and  $1/n^2$  with no noise.

Manuscript received February 5, 1985; revised February 10, 1987. This paper was presented at the International Symposium on Information Theory, Brighton, England, June 24–28, 1985. This research was supported by AFOSR Grant No. F49620 82 C 0009.

J. A. Bucklew is with the Electrical and Computer Engineering Department, University of Wisconsin, Madison, WI 53706.

S. Cambanis is with the Department of Statistics, University of North Carolina, Chapel Hill, NC 27514.

IEEE Log Number 8718722.

One of the main contributions of this paper is to show that these mean square estimation errors and their rate of convergence to least possible values depend crucially on the solution of a certain Wiener–Hopf integral equation. If the solution to the integral equation is smooth and contains at most Dirac delta functions, but not derivatives of delta functions, then asymptotically optimal sampling designs can be chosen for both kinds of estimators. Otherwise, the rate of convergence of the optimal coefficient estimator is not known. Fortunately the rate of convergence of simple coefficient estimators can still be found even in this case, but their asymptotic mean square estimation error is *not* the least possible, even though it can be made arbitrarily close.

In Section II we develop the general set-up and solution to the problem. In Section III we consider in more detail the cases of additive observation noise, of nonlinear signal distortion plus additive noise, and of quantization noise. In particular, we see that while additive noise of comparable smoothness with the signal does not affect the rates of convergence (but does affect the asymptotic constants), quantization noise also reduces the convergence rate.

Throughout the paper we consider in detail the case of random processes with no quadratic mean derivative, both for simplicity of exposition and because several questions still remain unresolved, especially when quadratic derivatives of order two and higher exist.

We consider only two kinds of (nonrandom) sampling design. They generalize periodic sampling, which includes the endpoints  $a$  and  $b$ , and periodic sampling which does not include the endpoints but is symmetrically spaced in the observation interval. Choose a continuous positive probability density  $h$  on the interval  $[a, b]$ . *Regular sampling* chooses for each  $n$  as the  $n$  sampling points  $T_n = \{t_{n1}, \dots, t_{nn}\}$  all  $(n-1)^{-1}$  percentiles of  $h$ :

$$\int_a^{t_{nk}} h(t) dt = \frac{k-1}{n-1}, \quad k=1, 2, \dots, n,$$

and we refer to  $\{T_n\}$  as a sequence of regular sampling designs generated by the density  $h$ . *Median sampling* chooses for each  $n$  as the  $n$  sampling points  $T_n = \{t_{n1}, \dots, t_{nn}\}$  the medians of a regular sequence of designs:

$$\int_a^{t_{nk}} h(t) dt = \frac{2k-1}{2n}, \quad k=1, 2, \dots, n,$$

and we refer to  $\{T_n\}$  as a sequence of median sampling designs generated by the density  $h$ . When  $h$  is the uniform density, both regular and median sampling become periodic; the former includes the interval endpoints while the latter does not.

We will use the following notation to simplify the text. With  $T = (t_1, \dots, t_n)$  an  $n$ -point subset of  $[a, b]$  and with functions  $f(t)$  defined on  $[a, b]$  and  $R(s, t)$  defined on  $[a, b] \times [a, b]$ , we will write  $f'_T$  for the  $n$ -vector  $(f(t_1), \dots, f(t_n))$  and  $R_T$  for the  $n \times n$  matrix  $\{R(t_i, t_j)\}_{i,j=1}^n$ . We also frequently delete the range of integration as well as the argument, writing, e.g.,  $\int \int R f f$  for  $\int_a^b \int_a^b R(s, t) f(s) f(t) ds dt$ .

## II. GENERAL CASE

We consider the problem of estimating the weighted integral of a random process  $X = \{X(t), a \leq t \leq b\}$ :

$$I = \int_a^b X(t) f(t) dt$$

from "noisy" observations of the random process  $Y = \{Y(t), a \leq t \leq b\}$  at  $n$  sample points  $T = \{t_k\}_{k=1}^n$ . The processes  $X$  and  $Y$  are assumed to have continuous correlation functions  $R_X(s, t)$  and  $R_Y(s, t)$  and cross-correlation function  $R_{YX}(s, t)$ , and the weighting function  $f$  is assumed to be continuous. We restrict attention to linear estimates with weights  $c'_T = (c_{T,1}, \dots, c_{T,n})$ :

$$I_T = \sum_{k=1}^n c_{T,k} Y(t_k) = c'_T Y_T,$$

whose mean square approximation error is

$$e_T^2 = E(I - I_T)^2 = \sigma^2 - 2c'_T g_T + c'_T R_{Y,T} c_T \quad (1)$$

where

$$\begin{aligned} \sigma^2 &= EI^2 = \int \int R_X f f, \\ s(t) &= \int_a^b R_X(t, u) f(u) du, \quad (s = R_X f), \\ g(t) &= \int_a^b R_{YX}(t, u) f(u) du, \quad (g = R_{YX} f). \end{aligned} \quad (2)$$

If the observation process  $Y$  could be observed over the entire interval  $[a, b]$ , then the minimum mean square approximation error  $\epsilon_\infty^2$  would be achieved by the projection  $\hat{I}$  of  $I$  onto the linear span of  $Y$ , which is determined by

$$E[ IY(t) ] = E[ \hat{I}Y(t) ]$$

for all  $a \leq t \leq b$ , or equivalently by

$$g(t) = (R_{YX} f)(t) = E[ \hat{I}Y(t) ].$$

It follows that  $g$  belongs to the reproducing kernel Hilbert space (RKHS)  $\mathcal{R}(R_Y)$  of  $Y$  and

$$\epsilon_\infty^2 = EI^2 - E\hat{I}^2 = \sigma^2 - \|g\|_{\mathcal{R}(R_Y)}^2. \quad (3)$$

Of course when noiseless observations are available, i.e., when  $Y(t) = X(t)$ , then  $g = s$  and  $\epsilon_\infty^2 = 0$ . In the general case we always have  $\epsilon_\infty^2 \geq 0$ .

Our goal is to choose the sampling points  $T$  and the estimator weights  $c_T$  in such a way that the resulting mean square estimation error  $e_T^2$  should be as close to  $\epsilon_\infty^2$  as possible.

### Optimal Coefficients

For a fixed sample  $T$ , the optimal coefficients  $\hat{c}_T$  are those which minimize the mean square approximation error  $e_T^2$  of (1), or equivalently those which make  $c'_T Y_T$  the projection of  $I$  onto the linear span of  $Y_T$ . They are given by  $\hat{c}'_T = g'_T R_{Y,T}^{-1}$ , and thus the optimal estimator and its mean square approximation error are

$$\begin{aligned} \hat{I}_T &= g'_T R_{Y,T}^{-1} Y_T \\ \epsilon_T^2 &= \sigma^2 - g'_T R_{Y,T}^{-1} g_T = \sigma^2 - \|P_T g\|_{\mathcal{R}(R_Y)}^2 \end{aligned} \quad (4)$$

where  $P_T g$  is the projection of  $g$  to the subspace of  $\mathcal{R}(R_Y)$  generated by  $\{R_Y(\cdot, t), t \in T\}$ .

The optimal sampling design of size  $n$  (if it exists) thus maximizes  $\|P_T g\|_{\mathcal{R}(R_Y)}^2$  over all sampling designs  $\Delta_n$  of size  $n$ :  $T = \{a \leq t_1 < t_2 < \dots < t_n \leq b\}$ . The performance of the optimal designs tends to  $\epsilon_\infty^2$  as the sample size tends to infinity, since

$$\inf_{T \in \Delta_n} \epsilon_T^2 = \sigma^2 - \sup_{T \in \Delta_n} \|P_T g\|_{\mathcal{R}(R_Y)}^2 \rightarrow \sigma^2 - \|g\|_{\mathcal{R}(R_Y)}^2 = \epsilon_\infty^2.$$

Optimal sampling designs may not exist, and even when they exist, it may be difficult to determine them.

We now consider *regular* sequences  $\{T_n\}$  of sampling designs  $T_n$  generated by a density  $h$ , and write  $\hat{I}_{r,n}$  and  $\epsilon_{r,n}$  for  $\hat{I}_{T_n}$  and  $\epsilon_{T_n}$ . As the sampling size  $n$  increases, they satisfy

$$\epsilon_{r,n}^2 = \sigma^2 - \|P_{T_n} g\|_{\mathcal{R}(R_Y)}^2 = \epsilon_\infty^2 + \|g - P_{T_n} g\|_{\mathcal{R}(R_Y)}^2 \rightarrow \epsilon_\infty^2.$$

Precise rates of convergence follow from the work of Sacks and Ylvisaker [9], [10], [11] where the observation process  $Y$  has exactly  $k$  quadratic mean derivatives, under the additional assumption that the function  $g$  of (2), which is in the RKHS of  $R_Y$ , actually belongs to the smaller space  $r(R_Y)$ , the range of the integral type operator with kernel  $R_Y$ , i.e.,

$$g(t) = \int_a^b R_Y(t, u) \phi(u) du, \quad (g = R_Y \phi), \quad (5)$$

with  $\phi$  a continuous function. Specifically, under certain regularity conditions,

$$n^{2k+2} (\epsilon_{r,n}^2 - \epsilon_\infty^2) \rightarrow C_k \int_a^b \frac{\alpha_{Y,k}(t) \phi^2(t)}{h^{2k+2}(t)} dt \quad (6)$$

where  $\alpha_{Y,k}$  is the jump along the diagonal of the derivative of the  $(k, k)$  partial derivative of  $R_Y$ ,  $\alpha_{Y,k}(t) = R_Y^{k,k+1}(t, t-0) - R_Y^{k,k+1}(t, t+0) > 0$  (superscripts denoting partial derivatives). The regularity conditions are specified in [11] and need not be repeated here. It should be noted, however, that for  $k \geq 2$  this result has not been established for as broad a class of covariances  $R_Y$  as when  $k = 0, 1$  (cf. [2, Section 6.1]). By choosing the density  $h^*(t)$  which minimizes the right-hand side of (6), i.e., proportional to

$[\alpha_{Y,k}(t)\phi^2(t)]^{1/(2k+3)}$ , we obtain a regular sequence  $\{T_n^*\}$  of sampling designs which is *asymptotically optimal*. That is, it satisfies

$$\frac{\epsilon_{r^*,n}^2}{\inf_{T \in \Delta_n} \epsilon_T^2} \rightarrow 1,$$

and

$$n^{2k+2}(\epsilon_{r^*,n}^2 - \epsilon_\infty^2) \rightarrow C_k \left\{ \int_a^b [\alpha_{Y,k}(t)\phi^2(t)]^{1/(2k+3)} dt \right\}^{2k+3}$$

The regularity conditions are satisfied by stationary processes with rational spectral densities, the stationary process with triangular covariance, the Wiener process, etc., and in all these cases the jumps  $\alpha_k(t)$  of the covariance derivative are constant. The value of the constant is  $C_k = |B_{2k+2}|/(2k+2)!$  where  $B_m$  is the  $m$ th Bernoulli coefficient, and  $C_0 = 1/12$ ,  $C_1 = 1/720$  (see the discussion in [2, p. 351]).

For simplicity we will consider from now on examples with no quadratic mean derivative, i.e.,  $k = 0$ , such as the stationary Gauss–Markov process, the stationary process with triangular covariance, or the Wiener process, in which case we have

$$n^2(\epsilon_{r,n}^2 - \epsilon_\infty^2) \rightarrow \frac{1}{12} \int_a^b \frac{\alpha_{Y,0}(t)\phi^2(t)}{h^2(t)} dt. \quad (7)$$

An asymptotically optimal regular sequence of designs  $\{T_n^*\}$  is generated by the density  $h^*(t)$  proportional to  $[\alpha_{Y,0}(t)\phi^2(t)]^{1/3}$ , with

$$n^2(\epsilon_{r^*,n}^2 - \epsilon_\infty^2) \rightarrow \frac{1}{12} \left\{ \int_a^b [\alpha_{Y,0}(t)\phi^2(t)]^{1/3} dt \right\}^3. \quad (8)$$

#### The Wiener–Hopf Equation (5)

From (2) and (5) it follows that a sufficient condition for these precise rates is the existence of a continuous function  $\phi$  such that

$$\int_a^b R_Y(t, u)\phi(u) du = \int_a^b R_{YX}(t, u)f(u) du (=g(t)), \quad (9)$$

$a \leq t \leq b$ , or  $R_Y\phi = R_{YX}f$  ( $=g$ ). It may be desirable to first check the existence of a square integrable solution, and then check its continuity. In this connection it is of interest to note that a square integrable solution  $\phi$  of the Wiener–Hopf integral equation (9) will exist for every square integrable  $f$ , if (and only if) any of the following equivalent conditions are satisfied:

- a)  $r(R_{YX}) \subset r(R_Y)$ ;
- b)  $\|R_{YX}e\|^2 \leq C\|R_Ye\|^2$  for all square integrable functions  $e$  and some finite constant  $C$  (where the norm is  $L_2$ );
- c) for some finite constant  $C$ ,

$$C \int_a^b R_Y(t, u)R_Y(s, u) du - \int_a^b R_{YX}(t, u)R_{YX}(s, u) du$$

is a nonnegative definite function of  $t, s$ ;

- d) the minimum mean square error linear estimate  $\hat{I}$  of  $I$  based on  $\{Y(t), a \leq t \leq b\}$  is of the form (with  $\phi \in L_2$ )

$$\hat{I} = \int_a^b Y(t)\phi(t) dt.$$

In this case the expression of the asymptotic mean square error (3) can be written as follows:

$$\epsilon_\infty^2 = \int \int R_X ff - \int \int R_Y \phi \phi.$$

#### Simple Coefficients

Here we consider the very simple choice of (nonoptimal) coefficients of the following form: for each sample of size  $n$ ,  $T_n = \{t_{nk}\}_{k=1}^n$ , we take

$$c_{T_n, k} = \frac{1}{n} c(t_{nk})$$

for some continuous function  $c(t)$ . Thus the coefficient  $c_{T_n, k}$  of  $Y(t_{nk})$  depends only on the sample point  $t_{nk}$  via an appropriate function  $c$ , unlike the optimal coefficients which depend on the entire sample  $T_n$ .

For a sequence of median sampling designs  $\{T_n\}$  generated by the density  $h$  we then have

$$\begin{aligned} I_{m,n} &= \frac{1}{n} \sum_{k=1}^n c(t_{nk})Y(t_{nk}) = \frac{1}{n} c'_{T_n} Y_{T_n}, \\ e_{m,n}^2 &= E(I - I_{m,n})^2 \\ &= \int \int R_X ff - \frac{2}{n} c'_{T_n} g_{T_n} + \frac{1}{n^2} c'_{T_n} R_{Y, T_n} c_{T_n}, \end{aligned}$$

and by [4, Lemma 2],

$$\begin{aligned} e_{m,\infty}^2 &= \lim_{n \rightarrow \infty} e_{m,n}^2 \\ &= \int \int R_X ff - 2 \int chg + \int \int R_Y(ch)(ch) \\ &= \int \int R_X ff - 2 \int \int R_{YX}(ch)(f) + \int \int R_Y(ch)(ch) \\ &= E \left( \int Xf - \int Ych \right)^2. \end{aligned} \quad (10)$$

A reasonable way of choosing the weighting function  $c$  is by minimizing the limiting value of the mean square error  $e_{m,\infty}^2$ . It follows from (10) that  $e_{m,\infty}^2 = \epsilon_\infty^2$  if and only if  $\int Ych = \hat{I}$  or equivalently (by d) in the preceding subsection) if  $R_Y\phi = R_{YX}f$  has an  $L_2$  solution  $\phi$ , in which case  $c$  is determined by

$$c(t)h(t) = \phi(t), \quad a \leq t \leq b. \quad (11)$$

Also, any smoothness requirements imposed on  $c$ , such as continuity (which has already been assumed) or twice continuous differentiability (which will be required shortly) would have to be satisfied as well by the  $L_2$  solution  $\phi$  to the Wiener–Hopf integral equation. When  $R_Y\phi = R_{YX}f$  does not have a continuous (resp. twice continuously differentiable) or even an  $L_2$  solution  $\phi$ , then one can find a

continuous (resp. twice continuously differentiable)  $c$  with corresponding mean square error  $e_{m,\infty}^2$  exceeding  $\epsilon_\infty^2$  by an arbitrarily small amount (and in the latter case no minimizing  $L_2$  function  $c$  exists). This is because random variables of the form  $\int Y\psi$  with  $\psi$  continuous (resp. twice continuously differentiable) form a dense set in the linear space of  $Y$ , so that given any  $\delta > 0$  we can find a continuous (resp. twice continuously differentiable)  $\psi$  such that  $E(\hat{I} - \int Y\psi)^2 \leq \delta^2$  and choosing  $ch = \psi$  we have

$$\begin{aligned} e_{m,\infty} &= E^{1/2} \left( \int Xf - \int Y\psi \right)^2 \\ &\leq E^{1/2} \left( \int Xf - \hat{I} \right)^2 + E^{1/2} \left( \hat{I} - \int Y\psi \right)^2 \\ &\leq \epsilon_\infty + \delta. \end{aligned}$$

We are thus led to consider the following two cases.

Case 1:  $R_Y\phi = R_{YX}f$  has a twice continuously differentiable  $L_2$  solution  $\phi$ . Then  $c$  is chosen by (11) and

$$I_{m,n} = \frac{1}{n} \sum_{k=1}^n \frac{\phi(t_{nk})}{h(t_{nk})} Y(t_{nk}), \quad (12)$$

$$\begin{aligned} e_{m,n}^2 - \epsilon_\infty^2 &= \iint R_Y\phi\phi - \frac{2}{n} \left( \frac{\phi}{h} \right)'_{T_n} (R_Y\phi)_{T_n} \\ &\quad + \frac{1}{n^2} \left( \frac{\phi}{h} \right)'_{T_n} R_{Y,T_n} \left( \frac{\phi}{h} \right)_{T_n}. \end{aligned}$$

Assuming that the observation process  $Y$  has no quadratic mean derivatives ( $k=0$ ), that  $R_Y$  satisfies the same regularity conditions required for (7), and that  $\phi/h$  is twice continuously differentiable, we obtain from [4, p. 94],

$$n^2(e_{m,n}^2 - \epsilon_\infty^2) \rightarrow \frac{1}{12} \int_a^b \frac{\alpha_{Y,0}(t)\phi^2(t)}{h^2(t)} dt. \quad (13)$$

Thus choosing  $h^*(t)$  proportional to  $[\alpha_{Y,0}(t)\phi^2(t)]^{1/3}$ , we obtain a sequence of median sampling designs  $\{T_n^*\}$  whose corresponding estimators  $\{I_{m,n}^*\}$  are asymptotically optimal. Comparing the asymptotically optimal sequence of estimators using median sampling and nonoptimal coefficients with the asymptotically optimal sequence of estimators using regular sampling and optimal coefficients, we see that in both cases the design is determined by the same density  $h^*(t)$ , while the estimator coefficients require solving an integral equation in the former case and inverting an  $n \times n$  matrix in the latter.

Case 2:  $R_Y\phi = R_{YX}f$  has no twice continuously differentiable solution  $\phi$ . In this case,

$$e_{m,n}^2 \rightarrow e_{m,\infty}^2 > \epsilon_\infty^2$$

with strict inequality for all choices of twice continuously differentiable  $c$ . Thus the asymptotic performance is always inferior to that achieved by using optimal coefficients. It turns out, however, that the rate of convergence of  $e_{m,n}^2$  can be found, while under the present conditions no rates are generally known for optimal coefficient estimators.

To see this, we write

$$\begin{aligned} e_{m,n}^2 - e_{m,\infty}^2 &= \left\{ \frac{1}{n^2} c'_{T_n} R_{Y,T_n} c_{T_n} - \iint R_Y(ch)(ch) \right\} \\ &\quad - 2 \left\{ \frac{1}{n} c'_{T_n} (R_{YX}f)_{T_n} - \iint R_{YX}(ch)(f) \right\}. \quad (14) \end{aligned}$$

Assuming as in Case 1 that  $Y$  has no quadratic mean derivative ( $k=0$ ), that  $c$  and  $f/h$  are twice continuously differentiable, that  $R_Y$  satisfies the same regularity conditions required for (7), and that  $R_{YX}$  satisfies similar regularity conditions, we show in Appendix I that

$$\begin{aligned} n^2(e_{m,n}^2 - e_{m,\infty}^2) &\rightarrow \frac{1}{12} \int \left[ \alpha_{Y,0}c^2 - \beta_{YX}\frac{cf}{h} + A_{YX}f - A_Ych \right] \quad (15) \end{aligned}$$

where

$$\begin{aligned} \beta_{YX}(t) &= [R_{YX}^{1,0}(t-, t) - R_{YX}^{1,0}(t+, t)] \\ &\quad - [R_{YX}^{0,1}(t, t-) - R_{YX}^{0,1}(t, t+)], \\ A_Y(t) &= h^{-1}(b) [c'(b)R_Y(b, t) + c(b)R_Y^{1,0}(b, t)] \\ &\quad - h^{-1}(a) [c'(a)R_Y(a, t) + c(a)R_Y^{1,0}(a, t)], \end{aligned}$$

and  $A_{YX}$  likewise with  $R_Y$  replaced by  $R_{YX}$ . It should be noticed that  $\beta_{YX}(t) \equiv 0$  when  $X$  and  $Y$  are jointly stationary, as well as when  $R_{YX}(t, s)$  is a symmetric function of  $t$  and  $s$ ; the latter is the case when  $Y = X$ , or  $Y = X +$  independent noise (see Section III-A), or even when  $Y$  is a zero-memory nonlinear transformation of  $X$  (see Section III-C) possibly plus independent noise (see Section III-B).

As has been pointed out in the paragraph following (10), in this case the estimator weights  $c(t)$  should be chosen in such a way that the resulting asymptotic mean square error  $e_{m,\infty}^2$  of (10) would be close to its minimum value  $\epsilon_\infty^2$ . This is the primary consideration in choosing  $c$ , and any further considerations such as those resulting from the form of the asymptotic constant in (15) are only of secondary importance. Since the asymptotic constant in (15) is not necessarily positive, some choices of  $c$  may produce a negative value indicating that for sufficiently large  $n$  the performance would be at least as good as  $e_{m,\infty}^2$ , whose value would perhaps be appreciably larger than its minimum. Also, judicious choices of  $c$  may exist which render the asymptotic constant in (15) equal to zero, indicating a faster rate of convergence which is not currently known. However, the dependence of  $A_Y, A_{YX}$  on the boundary values of  $c$  and  $c'$  complicates the question of existence of such  $c$ 's, and more significantly, even though one can derive the general form of such choices of  $c$ , it does not seem feasible to determine how close the resulting asymptotic mean square error  $e_{m,\infty}^2$  can be to its minimum  $\epsilon_\infty^2$ . We therefore do not pursue this matter any further. Incidentally, the estimator weight which minimizes the asymptotic constant in (15) is given by

$$c(t)h(t) = \frac{1}{2\alpha_{Y,0}(t)} \{ \beta_{YX}(t)f(t) + A_Y(t)h^2(t) \}$$

for  $a < t < b$ , with appropriately determined values of

$c(a)$ ,  $c(b)$ ,  $c'(a)$ ,  $c'(b)$  via a system of four linear equations resulting from the dependence of  $A_Y$ ,  $A_{YX}$  on these, and with corresponding minimum value of the asymptotic constant in (15),

$$-\frac{1}{48} \int \frac{1}{\alpha_{Y,0}} \left( \beta_{YX} \frac{f}{h} + A_Y h \right)^2 + \frac{1}{12} \int A_{YX} h.$$

Recall that in the noiseless case  $Y = X$ , we have  $\phi = f$  and the appropriate choice of  $c$  is, by (11),  $ch = f$ . It is therefore of interest to determine the best constant multiplier of  $f$  as a possible value of  $ch$ . Thus putting

$$c(t)h(t) = \lambda f(t), \quad a \leq t \leq b, \quad (16)$$

in the expression (10) of  $e_{m,\infty}^2$  we find that the value of  $\lambda$  which minimizes the asymptotic mean square error  $e_{m,\infty}^2$  is

$$\lambda = \frac{\iint R_{YX} ff}{\iint R_Y ff}, \quad (17)$$

with resulting estimator in this case of no quadratic mean derivative

$$I_{m,n} = \frac{\lambda}{n} \sum_{k=1}^n \frac{f(t_{nk})}{h(t_{nk})} Y(t_{nk}), \quad (18)$$

minimum asymptotic mean square error value

$$e_{m,\infty}^2 = \iint R_X ff - \frac{\left( \iint R_{YX} ff \right)^2}{\iint R_Y ff}, \quad (19)$$

and asymptotics

$$n^2(e_{m,n}^2 - e_{m,\infty}^2) \rightarrow \frac{1}{12} \left\{ \lambda \int (\lambda \alpha_{Y,0} - \beta_{YX}) \frac{f^2}{h^2} + \int (A_{YX} - \lambda A_Y) f \right\}. \quad (20)$$

The estimator (18) has the advantages of being generally applicable and fairly nonparametric, in that it depends on the correlation functions only via the integrals in (17). In sharp contrast, the estimator (12) requires the solution of the integral equation and thus fairly detailed knowledge of the correlation functions, and so does of course the optimal coefficient estimator. The estimator (18) is therefore more robust with respect to inaccuracies in our knowledge of the required correlation functions than are the other two estimators. It can be used for its simplicity and robustness, instead of the estimator (12), even when the integral equation has a twice continuously differentiable solution, i.e., in Case 1, and in this case one would want to know how much performance is lost asymptotically because the limiting mean square error  $e_{m,\infty}^2$  of (19) exceeds  $\epsilon_\infty^2$ . One would want to compute

$$\frac{e_{m,\infty}^2 - \epsilon_\infty^2}{\epsilon_\infty^2} = \frac{\iint R_Y ff \iint R_{YX} f \phi - \left( \iint R_{YX} ff \right)^2}{\iint R_Y ff \left( \iint R_X ff - \iint R_{YX} f \phi \right)}.$$

### Observation Process with Rational Spectral Density

When the observation process  $Y$  is stationary and has rational spectral density, the integral equation (5) or (9) has a solution which contains delta functions and their derivatives at the endpoints of the observation interval [8, Ch. III, Sect. 7]. In particular, when  $Y$  has no quadratic mean derivatives ( $k=0$ ), then the solution contains only delta functions so that

$$g(t) = \int_a^b R_Y(t-u) \phi_0(u) du + AR_Y(t-a) + BR_Y(t-b)$$

where  $\phi_0$  is a continuous function [8]. In this case Sacks and Ylvisaker [9] show that their asymptotic result is still valid with  $\phi_0$  playing the role of  $\phi$ :

$$n^2(\epsilon_{r,n}^2 - \epsilon_\infty^2) \rightarrow \frac{\alpha_{Y,0}}{12} \int_a^b \frac{\phi_0^2}{h^2}. \quad (21)$$

It is also straightforward to check that if we adjust the simple coefficient estimator using median sampling designs  $\{T_n\}$  by adjoining the endpoints with appropriate weights, and if we choose  $c(t)$  from  $c(t)h(t) = \phi_0(t)$ :

$$\tilde{I}_{m,n} = AY(a) + BY(b) + \frac{1}{n-2} \sum_{k=1}^{n-2} \frac{\phi_0(t_{n-2,k})}{h(t_{n-2,k})} Y(t_{n-2,k}) \quad (22)$$

( $n \geq 2$ ), its mean square error  $e_{m,n}^2 = E(I - \tilde{I}_{m,n})^2$  satisfies

$$n^2(e_{m,n}^2 - \epsilon_\infty^2) \rightarrow \frac{\alpha_{Y,0}}{12} \int_a^b \frac{\phi_0^2}{h^2}. \quad (23)$$

It converges to  $\epsilon_\infty^2$  (rather than to the larger  $e_{m,\infty}^2$ ) and with the same rate as  $\epsilon_{r,n}^2$ , provided  $\phi_0/h$  is twice continuously differentiable. Thus, in this case the modified simple coefficient estimator using median sampling is asymptotically optimal.

We should note that, while the assumption of rational spectral density is frequently reasonable, such as when the observation is signal plus noise, there are important cases where it is unlikely to be satisfied. Some are considered in Parts B and C of Section III.

### III. SPECIAL CASES AND EXAMPLES

In this section we consider several special cases of noisy observations of interest.

#### A. Independent Signal Plus Noise

Suppose that  $Y(t) = X(t) + N(t)$ ,  $a \leq t \leq b$ , where the noise process  $N$  is independent of the signal  $X$  and has zero mean. In this case  $R_{YX} = R_X$  and  $R_Y = R_X + R_N$ . We consider in more detail the following two special cases.

**Gauss-Markov Signal and Noise:** Suppose we desire to estimate the average of a Gauss-Markov signal over the unit interval, when observed in additive independent Gauss-Markov noise, i.e.,

$$\begin{aligned} R_X(\tau) &= \sigma_X^2 \exp[-a_X |\tau|], \\ R_N(\tau) &= \sigma_N^2 \exp[-a_N |\tau|], \\ f(t) &= 1, \quad 0 \leq t \leq 1. \end{aligned}$$

In this case,

$$g(t) = \int_0^1 R_X(t-u) du = \frac{\sigma_X^2}{a_X} (2 - e^{-a_X t} - e^{-a_X(1-t)}),$$

$0 \leq t \leq 1$ . When  $a_X \neq a_N$  the integral equation (5) has a solution  $\phi$  containing delta functions

$$\phi(t) = \phi_0(t) + M\{\delta(t) + \delta(t-1)\} \quad (24)$$

where

$$\phi_0(t) = M_1 + M_2\{\exp[-\alpha t] + \exp[-\alpha(1-t)]\} \quad (25)$$

and the values of the constants  $M$ ,  $M_1$ ,  $M_2$ , and  $\alpha > 0$  are given in Appendix II.

The optimal mean square estimation error is computed from (3),  $\epsilon_\infty^2 = \sigma^2 - \|g\|_{\mathcal{H}(R_Y)}^2$ , where

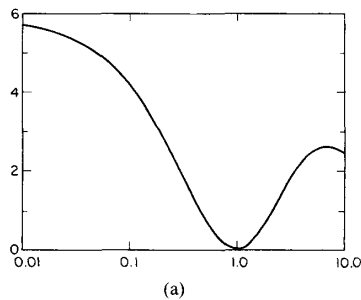
$$\begin{aligned} \sigma^2 &= \int_0^1 \int_0^1 R_X(t-u) dt du \\ &= (2\sigma_X^2/a_X^2)(e^{-a_X} - 1 + a_X) \triangleq r_X^2 \end{aligned}$$

and  $\|g\|_{\mathcal{H}(R_Y)}^2 = \int_0^1 g(t)\phi(t) dt$  is easily calculated.

This optimal error  $\epsilon_\infty^2$  is the limiting (large sample size) value of the mean square estimation error when using optimal coefficients and regular sampling, cf. (21), or the adjusted simple coefficient estimator of (22) and median sampling, cf. (23). When using median sampling with the simple coefficient estimator (18) with the optimal constant  $\lambda$  of (17), the limiting value of the mean square estimation error of (19) is given in this case by

$$e_{m,\infty}^2 = \frac{r_X^2 r_N^2}{r_X^2 + r_N^2}$$

and the optimal scaling constant by  $\lambda = r_X^2/(r_X^2 + r_N^2)$ . Asymptotically, the loss of performance by using the estimator (18) versus the adjusted simple coefficient estimator (22) can be measured by the ratio  $100(e_{m,\infty}^2 - \epsilon_\infty^2)/\epsilon_\infty^2$ , which is plotted in Fig. 1 for a range of parameter values. It is surprising that over a very large range of parameter values this ratio does not exceed nine percent. This means there is only a moderate loss of performance asymptotically when the simple coefficient median sampling scheme is not adjusted by including the interval endpoints with appropriate weights. Only as  $a_N/a_X$  approaches zero does the loss of performance become large, indicating a substantially improved performance of the adjusted scheme.



In Fig. 2 we show how the asymptotic mean square error decreases with increasing sample size  $n$  in the following cases by plotting the corresponding asymptotic expression of the mean square error.

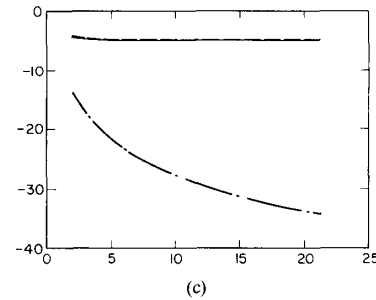
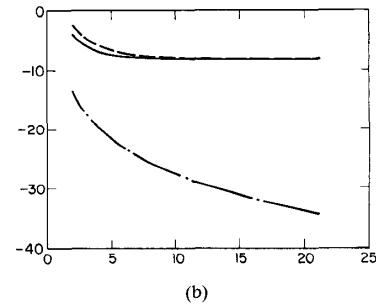
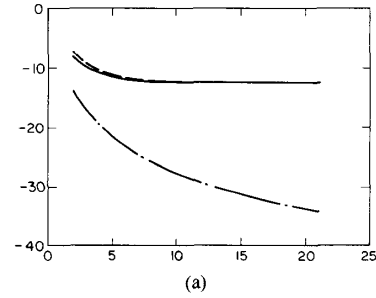


Fig. 2. Plot of asymptotic mean square error (in dB) versus number of samples  $n$  ( $n: 2-21$ ). — · — · no noise case, — noise present simple coefficients, — noise present optimal coefficients, - - - noise present nonadjusted simple coefficients. (a)  $\sigma_X^2 = a_X = 1$ ,  $\sigma_N^2 = 0.1$ ,  $a_N = 2$ . (b)  $\sigma_X^2 = \sigma_N^2 = a_X = 1$ ,  $a_N = 10$ . (c)  $\sigma_X^2 = \sigma_N^2 = a_X = 1$ ,  $a_N = 2$ .

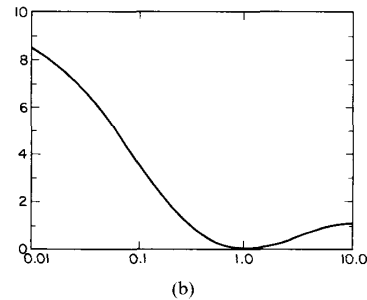


Fig. 1. Plot of  $100(e_{m,\infty}^2 - \epsilon_\infty^2)/\epsilon_\infty^2$  versus  $a_N$ . (a)  $\sigma_X^2 = \sigma_N^2 = a_X = 1$ ,  $a_N: 0.01-10$ . (b)  $\sigma_X^2 = a_X = 1$ ,  $\sigma_N^2 = 10$ ,  $a_N: 0.01-10$ .

- a) No noise, uniform sampling, optimal or simple coefficient estimator:

$$10 \log_{10} \{ (a_X/6) n^{-2} \} \text{ (with dash-dot line).} \quad (26)$$

- b) Noise present, uniform sampling, optimal or adjusted simple coefficient estimator (cf. (21) and (23)):

$$10 \log_{10} \left\{ \epsilon_\infty^2 + \left[ \frac{1}{6} (a_X + a_N) \int \phi_0^2 \right] n^{-2} \right\} \quad (27)$$

(with solid line).

- c) Noise present, optimal estimator with optimal regular sampling, or simpler coefficient estimator with optimal median sampling (cf. (21) and (23) with optimal  $h^*$  proportional to  $|\phi_0|^{2/3}$ ):

$$10 \log_{10} \left\{ \epsilon_\infty^2 + \left[ \frac{1}{6} (a_X + a_N) \left( \int |\phi_0|^{2/3} \right)^3 \right] n^{-2} \right\} \quad (28)$$

(with solid line).

- d) Noise present, uniform sampling, nonadjusted simple coefficient estimator (cf. (20)):

$$10 \log_{10} \{ e_{m,\infty}^2 + C n^{-2} \} \text{ (with dotted line)}$$

where

$$C = \frac{\lambda}{6} \left\{ \lambda (a_X + a_N) - (1 - \lambda) \sigma_X^2 (1 - e^{-a_X}) + \lambda \sigma_N^2 (1 - e^{-a_N}) \right\}.$$

As it turns out, the constant term  $M_1$  of  $\phi_0$  is much larger than the  $M_2$  term for the values of the parameters we plotted. Hence  $\phi_0$  is nearly always flat and the optimal sampling density  $h^*$  in (c) is nearly uniform, i.e.,  $(\int |\phi_0|^{2/3})^3$  nearly equals  $\int \phi_0^2$ , and in our plots cases b) and c) were apparently identical. Note also that the expressions we plotted in Fig. 2 for  $n = 2$  to 21 are asymptotic values valid only for large  $n$ .

In Fig. 3 we also plot one of the asymptotic curves—the (b) curve given by (27)—to compare the actual with the also indicate the level of the asymptotic mean square error  $\epsilon_\infty^2 = .3173$  or  $-4.985$  dB, which is the mean square error of the optimal continuous-time estimator. The portion of the mean square error over the level corresponding to  $\epsilon_\infty^2$  represents the mean square error due to sampling and thus the degradation of performance (due to sampling) over the performance of the optimal continuous-time estimator. If this performance degradation should not exceed 5 percent, then  $n = 5$  samples suffice for all sampling schemes, and  $n = 6$  samples suffice for 2.5 percent, and  $n = 12$  samples suffice for 1 percent performance degradation.

In Fig. 3 we also plot one of the asymptotic curves—the (b) curve given by (27)—to compare the actual with the

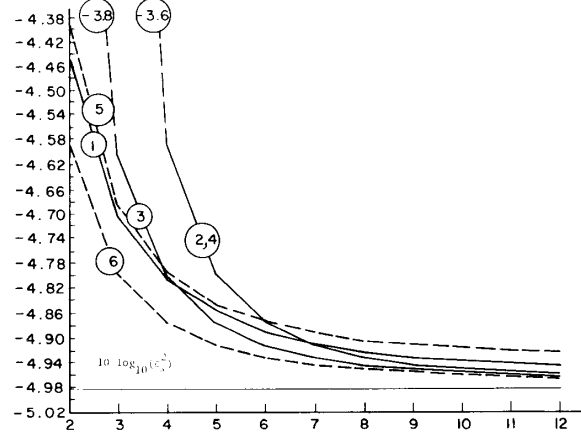


Fig. 3. Plots of mean square error (in dB) for the following schemes. ① Uniform sampling, optimal coefficients. ② Uniform sampling, adjusted simple coefficients. ③ Optimal regular sampling, optimal coefficients. ④ Optimal median sampling, adjusted simple coefficients. ⑤ Uniform sampling, simple coefficients. ⑥ Asymptotic expression for uniform sampling optimal coefficients or uniform sampling adjusted simple coefficients.

asymptotic values. It is seen that the mean square error due to sampling reaches its asymptotic performance very slowly. For instance, for uniform sampling with optimal coefficients, the mean square error due to sampling at  $n = 5$  is about 2 times its asymptotic value and at  $n = 8$  is about 2.44 times its asymptotic value! Even though the mean square error due to sampling reaches its asymptotic performance very slowly, because it is only a small fraction of the overall mean square error, the asymptotic values of the overall mean square error plotted in Fig. 2 are very close to the actual values for  $n \geq 5$ .

It is of interest to know how many samples are required asymptotically to attain a given performance, say for error  $(1 + \beta)\epsilon_\infty^2$ . When no noise is present and uniform sampling is used with optimal or simpler coefficient estimator, expression (26) gives

$$n_{NN,U}^2 = \frac{a_X}{6(1 + \beta)\epsilon_\infty^2}.$$

With noise present and uniform sampling with optimal or adjusted simpler coefficient estimator, expression (27) gives

$$n_{N,U}^2 = \frac{a_X + a_N}{6\beta\epsilon_\infty^2} \int \phi_0^2.$$

The ratio

$$\frac{n_{N,U}}{n_{NN,U}} = \left\{ (1 + a_X/a_N)(1 + 1/\beta) \int \phi_0^2 \right\}^{1/2}$$

summarizes the effect of noise upon the necessary sampling rate. We have plotted this ratio in Fig. 4 for some representative values of the parameters. Again, as  $a_N/a_X$  approaches zero, the effect of noise becomes much more marked.

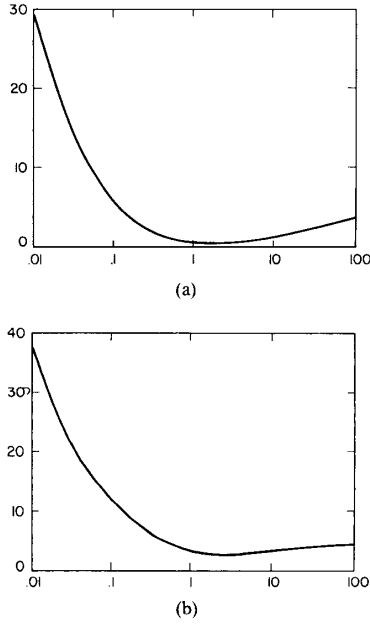


Fig. 4. Plot of sampling ratio  $n_{N,U}/n_{NN,U}$  versus  $a_N$ . (a)  $\sigma_X^2 = a_X = 1$ ,  $\sigma_N^2 = 10$ ,  $\beta = .05$ ,  $a_N$ : 0.01–100. (b)  $\sigma_X^2 = \sigma_N^2 = a_X = 1$ ,  $\beta = 0.05$ ,  $a_N$ : 0.01–100

**Proportional Signal and Noise Correlations:** We now consider the case where  $R_N = \gamma R_X$  for some positive constant  $\gamma$ . In this case  $R_Y = (1 + \gamma)R_X$  and  $g = s = R_X f = (1 + \gamma)^{-1} R_Y f$  so that the integral equation always has a continuous solution  $\phi = (1 + \gamma)^{-1} f$ . We also have  $a_{Y,0} = (1 + \gamma)a_{X,0}$  and

$$\begin{aligned} \epsilon_\infty^2 &= \iint R_X f f - \iint R_Y \phi \phi \\ &= \left(1 - \frac{1}{1 + \gamma}\right) \iint R_X f f = \frac{\gamma}{1 + \gamma} \sigma^2, \end{aligned}$$

so that

$$n^2 \left( \epsilon_{r,n}^2 - \frac{\gamma}{1 + \gamma} \sigma^2 \right) \rightarrow \frac{1}{12(1 + \gamma)} \int \frac{a_{X,0} f^2}{h^2} \triangleq \frac{A^2}{1 + \gamma}$$

and similarly for  $\epsilon_{m,n}^2$ , provided  $f/h$  is twice continuously differentiable. The asymptotically optimal sampling density  $h^*$  is proportional to  $(a_{X,0} f^2)^{1/3}$  and is therefore independent of the noise (i.e., of  $\gamma$ ). For large sample size  $n$ , both  $\epsilon_{r,n}^2$  and  $\epsilon_{m,n}^2$  are approximately equal to

$$\frac{1}{1 + \gamma} \left( \gamma \sigma^2 + \frac{A^2}{n^2} \right). \quad (28)$$

As in the previous case it is of interest to know how many samples are needed asymptotically to achieve a certain mean square error  $\epsilon^2$ :  $\gamma(1 + \gamma)^{-1} \sigma^2 < \epsilon^2 < \sigma^2$ . It follows from the asymptotic expression (28) of the mean square error that

$$n^2(\gamma) = \frac{A^2}{\epsilon^2 - \gamma(\sigma^2 - \epsilon^2)},$$

and with no noise present,

$$n^2(0) = \frac{A^2}{\epsilon^2}.$$

The ratio

$$\frac{n(\gamma)}{n(0)} = \frac{\epsilon}{\{\epsilon^2 - \gamma(\sigma^2 - \epsilon^2)\}^{1/2}}$$

increases with  $\gamma$  from 1 to infinity as  $\gamma$  approaches  $\epsilon^2/(\sigma^2 - \epsilon^2)$ . The first-order (linear) approximations in  $\gamma$  are for the weak noise case, i.e.,  $\gamma \sim 0$ ,

$$\frac{n(\gamma)}{n(0)} \sim 1 + \frac{\gamma}{2} \left( \frac{\sigma^2}{\epsilon^2} - 1 \right),$$

and for the comparable noise and signal case, i.e.,  $\gamma \sim 1$ ,

$$\frac{n(\gamma)}{n(0)} \sim \frac{\epsilon}{\sqrt{2\epsilon^2 - \sigma^2}} \left\{ 1 + \frac{\gamma - 1}{2} \cdot \frac{\sigma^2 - \epsilon^2}{2\epsilon^2 - \sigma^2} \right\}.$$

### B. Nonlinear Distortion Plus Noise

Suppose that the signal  $X(t)$  has suffered some nonlinear distortion, in addition to being corrupted by noise, i.e., that the observation process  $Y(t)$  is of the form

$$Y(t) = A(X(t)) + N(t)$$

where  $X(t)$  is stationary and Gaussian with mean zero,  $A(\cdot)$  is a memoryless nonlinearity such that  $EA^2(X(0)) < \infty$ , and  $N(t)$  is an independent, zero-mean, wide sense stationary noise. Then  $Y(t)$  is wide sense stationary and, assuming for simplicity that  $R_X(0) = 1$ , we can write

$$R_Y(\tau) = \sum_{k=0}^{\infty} \frac{a_k^2}{k!} R_X^k(\tau) + R_N(\tau)$$

where  $a_k = E[A(X(t))H_k(X(t))]$  and  $\{H_k(x)\}_k$  are the Hermite polynomials. Also, by the cross-covariance property, we have

$$R_{YX}(\tau) = dR_X(\tau)$$

where  $d = E[A(X(t))X(t)]$  (see [1]). Therefore  $g = R_{YX}f = dR_X f = ds$  and  $\epsilon_\infty^2 = \sigma^2 - d^2 \|s\|_{\mathcal{H}(R_Y)}^2$ .

It should be noted that if  $A$  is an even function, then  $d = 0$  and  $R_{YX} \equiv 0$ ; i.e., the observation process  $Y$  is orthogonal to the signal process  $X$ , and thus no linear estimate based on  $Y$  is better than zero as an estimate of the integral of  $X$ . We therefore assume throughout that  $A$  is such that  $d \neq 0$  (e.g., an odd function).

When optimal estimator coefficients and regular sampling are used, the asymptotics are determined by (7) with rate of convergence  $1/n^2$ , provided the integral equation  $R_Y \phi = dR_X f$  has a continuous solution  $\phi$ , and  $R_Y(\tau)$  has the needed differentiability properties, i.e., a finite positive jump  $-\alpha_{Y,0}$  in its derivative at zero. When the simple estimator coefficients and median sampling are used, then the rate of convergence is  $1/n^2$ , provided  $R_Y(\tau)$  has the needed differentiability properties, and the precise asymptotics are determined by (13) when the integral equation



$R_Y \phi = dR_X f$  has a twice continuously differentiable solution  $\phi$  and the estimator (12) is used, and by (20) when the estimator (18) is used.

*Smooth Limiters:* As an example, let us consider the so-called smooth limiters

$$A_v(x) = \frac{2K}{\sqrt{2\pi}v} \int_0^x e^{-z^2/2v^2} dz$$

for which

$$d = \frac{2K}{\sqrt{2\pi(1+v^2)}} \\ R_Y(\tau) = \frac{2K^2}{\pi} \sin^{-1} \left( \frac{R_X(\tau)}{1+v^2} \right) + R_N(\tau).$$

It is easily seen that if  $R_X$  and  $R_N$  have the needed differentiability properties to define  $\alpha_{X,0}$  and  $\alpha_{N,0}$  then so does  $R_Y$ , provided  $v > 0$ , and in fact

$$\alpha_{Y,0} = \frac{2K^2 \alpha_{X,0}}{\pi \sqrt{v(v+2)}} + \alpha_{N,0}.$$

Also  $\alpha_{YX,0} = d\alpha_{X,0}$ . Thus, provided  $v > 0$ , the rates of convergence remain  $1/n^2$ , as in the absence of nonlinear distortion. *Hard Limiter:* The case of a hard limiter

$$A(x) = K \operatorname{sgn}(x) = A_0(x)$$

requires special attention since, assuming  $R_X$  and  $R_N$  have the needed differentiability properties to define  $\alpha_{X,0}$  and  $\alpha_{N,0}$ , it follows that  $\alpha_{Y,0} = +\infty$  indicating a rate of convergence slower than  $1/n^2$ . This is because

$$R_Y(\tau) = \frac{2K^2}{\pi} \sin^{-1}(R_X(\tau)) + R_N(\tau)$$

implies  $R'_Y(0-) = +\infty$ ,  $R'_Y(0+) = -\infty$ . While the first derivatives of  $R_Y$  at  $0\pm$  are not finite, we notice that the "one-half derivatives" are. Indeed, we have

$$\begin{aligned} R_Y^{(1/2)}(0-) &\triangleq \lim_{\tau \uparrow 0} \frac{R_Y(0) - R_Y(\tau)}{(-\tau)^{1/2}} \\ &= \frac{2K^2}{\pi} \lim_{\tau \uparrow 0} \frac{-\frac{R'_X(\tau)}{[1-R_X^2(\tau)]^{1/2}}}{-\frac{1}{2}(-\tau)^{-1/2}} \\ &= \frac{2K^2}{\pi} \lim_{\tau \uparrow 0} \frac{2R'_X(\tau)}{\left\{ \frac{1-R_X(\tau)}{-\tau} \cdot [1+R_X(\tau)] \right\}^{1/2}} \\ &= \frac{2K^2}{\pi} \frac{2R'_X(0-)}{\{2R'_X(0-)\}^{1/2}} \\ &= \frac{2K^2}{\pi} \{2R'_X(0-)\}^{1/2} \\ &= \frac{2K^2}{\pi} \alpha_{X,0}^{1/2}. \end{aligned}$$

Similarly,

$$R_Y^{(1/2)}(0+) \triangleq \lim_{\tau \downarrow 0} \frac{R_Y(\tau) - R_Y(0)}{\tau^{1/2}} = -\frac{2K^2}{\pi} \alpha_{X,0}^{1/2}$$

and thus,

$$\alpha_{Y,1/2} \triangleq R_Y^{(1/2)}(0-) - R_Y^{(1/2)}(0+) = \frac{4K^2}{\pi} \alpha_{X,0}^{1/2}.$$

The expression (14) for the mean square error when the simpler coefficients are used becomes, in this case,

$$\begin{aligned} e_{m,n}^2 - e_{m,\infty}^2 &= \left\{ \frac{1}{n^2} c'_{T_n} R_{Y,T_n} c_{T_n} - \iint R_Y(ch)(ch) \right\} \\ &\quad - 2d \left\{ \frac{1}{n} c'_{T_n} (R_X f)_{T_n} - \iint R_X(ch)(f) \right\} \\ &\triangleq (1\text{st term}) - 2d(2\text{nd term}). \end{aligned} \quad (29)$$

For the second term we have as in (15) (cf. Eq. (54)),  $n^2(2\text{nd term}) \rightarrow \text{finite constant}$ , provided  $c$  and  $f/h$  are twice continuously differentiable. The rate of convergence of the first term can be found in a similar way. Instead of using, say for  $\tau > 0$ ,  $R_Y(\tau) = R_Y(0) + \tau R'_Y(0+) + o(\tau)$ , we now use  $R_Y(\tau) = R_Y(0) + \sqrt{\tau} R_Y^{(1/2)}(0+) + o(\sqrt{\tau})$  (and similarly for  $\tau < 0$ ). It is shown in Appendix III that, provided  $c$  and  $h$  are twice continuously differentiable, and  $R_X, R_N$  have finite one-sided second derivatives at 0, we have

$$n^{3/2}(1\text{st term}) \rightarrow \gamma \alpha_{Y,1/2} \int \frac{(ch)^2}{h^{3/2}} \quad (30a)$$

where

$$\gamma = \frac{1}{3} \left( \frac{4}{5} - \frac{1}{16} \zeta \left( \frac{3}{2} \right) \right) \approx .21225 \quad (30b)$$

and  $\zeta(3/2) = \sum_{k=1}^{\infty} k^{-3/2} \approx 2.612$  is Riemann's zeta function evaluated at  $3/2$ . The first term in (29) is thus the dominant one, and in this case of signal and noise with no quadratic mean derivative we have

$$n^{3/2}(e_{m,n}^2 - e_{m,\infty}^2) \rightarrow \gamma \alpha_{Y,1/2} \int \frac{(ch)^2}{h^{3/2}}.$$

Hard limiting thus reduces the rate of convergence by a one-half power to  $1/n^{3/2}$  from the rate of  $1/n^2$  of soft limiting or no nonlinear distortion. If  $ch$  is chosen proportional to  $f$ , then the sampling density  $h^*$  which minimizes the asymptotic constant is proportional to  $|f|^{4/5}$ .

### C. Quantization

In applications one frequently has access only to quantized data. Here we assume that our observation process is

$$Y(t) = Q(X(t)) \triangleq QX(t)$$

where  $Q$  is an  $L$ -level quantizer with  $Q(x) = y_l$  when  $x_l < x \leq x_{l+1}$  where  $-\infty = x_1 < y_1 < x_2 < y_2 < \dots < x_L < y_L < x_{L+1} = +\infty$  and  $x_l = (y_l + y_{l-1})/2$  for  $l = 2, \dots, L$ . The process  $X(t)$  is stationary and Gaussian with mean zero and variance one. (More generally,  $X(t)$  can be taken to be any wide sense stationary process whose

bivariate densities have a diagonal expansion, so that the cross-covariance property holds—see [1].)

*Optimal Coefficients:* We first consider the case where the optimal coefficients estimate based on the quantized samples is used:

$$\hat{I}_{Q,T} = \hat{c}'_{Q,T} QX_T = g'_{Q,T} R_{QX,T}^{-1} QX_T$$

where  $g_Q = R_{QX} Xf = d_Q R_X f = d_Q g$ ,  $g = R_X f$ ,  $d_Q = E[QX(t)X(t)]$ . Denoting as usual by  $\hat{I}_T$  the optimal coefficients estimate based on the samples  $X_T$ :  $\hat{I}_T = \hat{c}'_T X_T = g'_T R_{XX,T}^{-1} X_T$ , we can write the mean square error as

$$\begin{aligned} \epsilon_{Q,T}^2 &= E(I - \hat{I}_{Q,T})^2 \\ &= E(I - \hat{I}_T)^2 + E(\hat{I}_T - \hat{I}_{Q,T})^2 \\ &\quad + 2E[(I - \hat{I}_T)(\hat{I}_T - \hat{I}_{Q,T})]. \end{aligned}$$

The third term vanishes as  $E[(I - \hat{I}_T)\hat{I}_T] = 0$ , since  $\hat{I}_T$  is the projection of  $I$  onto the linear span of  $X_T$ , and, by the cross-covariance property,

$$\begin{aligned} E[(I - \hat{I}_T)\hat{I}_{Q,T}] &= E\left[\left(\int Xf - \hat{c}'_T X_T\right) \hat{c}'_{Q,T} QX_T\right] \\ &= \sum_i c_{Q,T,i} \left\{ \int E[X(t)QX(t_i)] f(t) dt \right. \\ &\quad \left. - \sum_j c_{T,j} E[X(t_j)QX(t_i)] \right\} \\ &= d_Q \sum_i c_{Q,T,i} \left\{ \int E[X(t)X(t_i)] f(t) dt \right. \\ &\quad \left. - \sum_j c_{T,j} E[X(t_j)X(t_i)] \right\} \\ &= d_Q E[(I - \hat{I}_T) \hat{c}'_{Q,T} X_T] = 0. \end{aligned}$$

Thus the mean square error decomposes into two components:

$$\epsilon_{Q,T}^2 = E(I - \hat{I}_T)^2 + E(\hat{I}_T - \hat{I}_{Q,T})^2 \triangleq \epsilon_T^2 + e_{Q,T}^2.$$

The first term is due to sampling and has been discussed in Section II; its expression is  $\epsilon_T^2 = \sigma^2 - \|P_T g\|_{\mathcal{H}(R_X)}^2$ . The second term is due to quantization of the samples used to estimate the integral, and is given by

$$\begin{aligned} e_{Q,T}^2 &= E(g'_T R_{X,T}^{-1} X_T - d_Q g'_T R_{QX,T}^{-1} QX_T)^2 \\ &= g'_T R_{X,T}^{-1} g_T - d_Q^2 g'_T R_{QX,T}^{-1} g_T \\ &= \|P_T g\|_{\mathcal{H}(R_X)}^2 - d_Q^2 \|P_T g\|_{\mathcal{H}(R_{QX})}^2 \end{aligned}$$

where the cross-covariance property has been used. Thus when using a regular sequence  $\{T_n\}$  of sampling designs we have

$$\epsilon_{Q,r,n}^2 = \sigma^2 - d_Q^2 \|P_{T_n} g\|_{\mathcal{H}(R_{QX})}^2 \rightarrow \sigma^2 - d_Q^2 \|g\|_{\mathcal{H}(R_{QX})}^2 \triangleq \epsilon_{Q,\infty}^2.$$

(That  $g \in \mathcal{H}(R_{QX})$  follows from  $d_Q g = g_Q \in \mathcal{H}(R_{QX})$ , cf.

statement preceding Eq. (3).) It follows that

$$\epsilon_{Q,r,n}^2 - \epsilon_{Q,\infty}^2 = d_Q^2 \|g - P_T g\|_{\mathcal{H}(R_{QX})}^2,$$

and the conjecture here is that, as in the case of hard limiter, its rate of convergence to zero is  $n^{-3/2}$ , when  $X$  has no quadratic mean derivative ( $k=0$ ) and the derivative of  $R_X$  has positive jump  $\alpha_{X,0}$  at the origin, and where  $R_{QX}\phi = R_X f$  has a continuous solution  $\phi$  (but no proof is currently available).

*Simple Coefficients:* Next we consider the case where the simple coefficients are used along with a sequence of median sampling designs generated by the sampling density  $h_s$ . The resulting estimate is

$$I_{m,Q,n} = \frac{1}{n} \sum_k c(t_{n,k}) QX(t_{n,k})$$

with

$$\begin{aligned} e_{m,Q,n}^2 &= E(I - I_{m,Q,n})^2 \\ &= \iint R_X f f - \frac{2}{n} c'_T(R_{QX} Xf) r_n + \frac{1}{n^2} c'_T R_{0X} r_n c_T \\ &\rightarrow \iint R_X f f - 2 \iint R_{QX} X(f)(ch_s) + \\ &\quad \iint R_{QX}(ch_s)(ch_s) \\ &= E\left[\int Xf - \int (QX)(ch_s)\right]^2 \triangleq e_{Q,\infty}^2, \end{aligned} \quad (31)$$

by [4, Lemma 2]. The asymptotic error  $e_{Q,\infty}^2$  is minimized for some  $c$ ,  $h_s$  if and only if the projection of  $I = \int Xf$  onto the span of  $QX$  is of the form  $\int QX \cdot \phi$  for some  $\phi \in L_2$  (in which case then  $ch_s = \phi$ ), or equivalently if and only if there is an  $L_2$  solution  $\phi$  to the integral equation  $d_Q \cdot R_X f = R_{QX} \phi$ . From (31) it is clear that as  $L \rightarrow \infty$ ,  $e_{Q,L,\infty}^2 \rightarrow E[\int X(f - ch_s)]^2$ . Thus, asymptotically for large number  $L$  of levels of quantization, the best choice for  $c$  and  $h_s$  is

$$c(t)h_s(t) = f(t), \quad a \leq t \leq b. \quad (32)$$

Throughout the rest of this subsection we assume this to be the case, so that

$$e_{Q,\infty}^2 = E\left[\int (X - QX)f\right]^2 = \iint R_{X-QX} f f.$$

Under this assumption the mean square error is again, at least asymptotically as the number of samples  $n$  tends to infinity, the sum of a mean square error due only to sampling and of one due to the quantization of the samples. Indeed, with

$$I_{m,n} = \frac{1}{n} \sum_k c(t_{n,k}) X(t_{n,k})$$

we obtain

$$\begin{aligned} e_{m,Q,n}^2 &= E(I - I_{m,n})^2 + E(I_{m,n} - I_{m,Q,n})^2 \\ &\quad + 2E[(I - I_{m,n})(I_{m,n} - I_{m,Q,n})]. \end{aligned} \quad (33)$$

By the cross-covariance property,

$$E[(I - I_{m,n})I_{m,Q,n}] = d_Q E[(I - I_{m,n})I_{m,n}]$$

and

$$\begin{aligned} E[(I - I_{m,n})I_{m,n}] &= \frac{1}{n} c'_{T_n}(R_X f)_{T_n} - \frac{1}{n^2} c'_{T_n} R_{X,T_n} c_{T_n} \\ &\rightarrow \iint R_X(f)(ch_s) - \iint R_X(ch_s)(ch_s) \\ &= 0, \end{aligned}$$

by (32). Thus the cross term in (33) tends to 0 with  $n$ .

We now study the rate of convergence of  $e_{m,Q,n}^2$  to  $e_{Q,\infty}^2$  as  $n \rightarrow \infty$ , assuming that  $R_X$  has finite one-sided derivatives at 0 required for the definition of  $\alpha_{X,0}$ . As in Section III-B, we have

$$\begin{aligned} e_{m,Q,n}^2 - e_{Q,\infty}^2 &= \iint R_X ff - 2d_Q \frac{1}{n} c'_{T_n}(R_X f)_{T_n} + \frac{1}{n^2} c'_{T_n} R_{QX,T_n} c_{T_n} \\ &\quad - \iint (R_X - 2d_Q R_X + R_{QX}) ff \\ &= \left\{ \frac{1}{n^2} c'_{T_n} R_{QX,T_n} c_{T_n} - \iint R_{QX} ff \right\} \\ &\quad - 2d_Q \left\{ \frac{1}{n} c'_{T_n}(R_X f)_{T_n} - \iint R_X ff \right\} \\ &\triangleq (\text{1st term}) - 2d_Q (\text{2nd term}). \end{aligned} \quad (34)$$

Assuming that  $f, h, f/h$  are twice continuously differentiable, we have for the second term, as in (15) (cf. Eq. (54)),

$$n^2 (\text{2nd term}) \rightarrow \frac{1}{24} \int A_X f. \quad (35)$$

It is shown in Appendix IV that

$$\alpha_{QX,1/2} = (2\alpha_{X,0}/\pi)^{1/2} B_Q \quad (36a)$$

where

$$B_Q = (2\pi)^{-1/2} \sum_{l=2}^L (y_l - y_{l-1})^2 \exp\left\{-\frac{(y_l - y_{l-1})^2}{8}\right\} \quad (36b)$$

and that as  $\tau \rightarrow 0$ ,

$$|\tau|^{3/2} R''_{QX}(\tau) \rightarrow \frac{1}{8} \alpha_{QX,1/2}. \quad (37)$$

This case generalizes the hard limiter case considered in Section III-B. In fact, an inspection of the proof of (30) in Appendix III shows that the relationship between  $X$  and  $Y$  affects the asymptotics described by (30) only via  $\alpha_{Y,1/2}$  and (57b), which is identical with (37) above. It then follows from Appendix III, or (30), that provided  $R_X$  has finite one-sided second derivatives at 0, we have

$$n^{3/2} (\text{1st term}) \rightarrow \gamma \alpha_{QX,1/2} \int \frac{f^2}{h_s^{3/2}} \quad (38)$$

with  $\gamma$  given in (30b). We thus have

$$n^{3/2} (e_{m,Q,n}^2 - e_{Q,\infty}^2) \rightarrow \gamma (2\alpha_{X,0}/\pi)^{1/2} B_Q \int \frac{f^2}{h_s^{3/2}}. \quad (39)$$

Quantization therefore reduces the rate of convergence by a one-half power to  $n^{-3/2}$ , independently of the number of quantization levels; and the optimal sampling density  $h_s^*$ , which minimizes the right-hand side, is proportional to  $|f|^{4/5}$ .

To study the interplay between the number  $L$  of quantization levels of the quantizer  $Q_L$  and the number  $n$  of samples used, asymptotically as  $L$  and  $n$  tend to infinity we keep both terms in (34). From (34), (35), and (38) we have for large  $n$ ,

$$e_{m,Q,n}^2 \approx e_{Q,\infty}^2 + \frac{1}{n^{3/2}} B_Q \gamma (2\alpha_{X,0}/\pi)^{1/2} \int \frac{f^2}{h_s^{3/2}}. \quad (40)$$

Suppose that a “regular” sequence of quantizers  $\{Q_L\}$  based on a continuous density  $h_Q$  is used, i.e., the levels  $y_{L,1} < y_{L,2} < \dots < y_{L,L}$  of quantizer  $Q_L$  are, respectively, the  $1/(2L), 3/(2L), \dots, (2L-1)/(2L)$  quantiles of  $h_Q$ . With  $p(x)$  denoting the standard normal density of  $X(t)$ , choosing  $h_Q$  proportional to  $[p(x)]^{1/3}$ , i.e.,

$$h_Q^*(x) = (6\pi)^{-1/2} \exp(-x^2/6), \quad -\infty < x < \infty,$$

we obtain an asymptotically optimal sequence of quantizers. As in [3], with

$$\frac{1}{L} = \int_{y_{L,l-1}}^{y_{L,l}} h_Q(x) dx = h_Q(z_{L,l})(y_{L,l} - y_{L,l-1})$$

for some  $y_{L,l-1} < z_{L,l} < y_{L,l}$ , we obtain (assuming  $ph_Q^{-1}$  is Riemann integrable over  $(-\infty, \infty)$ ),

$$\begin{aligned} LB_{Q_L} &= \frac{1}{\sqrt{2\pi}} \sum_{l=2}^L \frac{\exp\left\{-\frac{(y_{L,l} - y_{L,l-1})^2}{8}\right\}}{h_Q(z_{L,l})} \\ &\quad \cdot (y_{L,l} - y_{L,l-1}) \\ &\rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\exp\left\{-\frac{(2x)^2}{8}\right\}}{h_Q(x)} dx \\ &= \int_{-\infty}^{\infty} \frac{p(x)}{h_Q(x)} dx. \end{aligned}$$

When  $h_Q^*$  is used the asymptotic constant becomes

$$\int_{-\infty}^{\infty} \frac{p}{h_Q^*} = 3\sqrt{\pi}.$$

Thus for large  $L$ ,

$$B_{Q_L} \approx \frac{1}{L} \int \frac{p}{h_Q}. \quad (41)$$

The precise rate of convergence to zero of  $e_{Q_L,\infty}^2$  seems harder to determine. However, we have the following

bound, assuming  $\int p h_Q^{-2} < \infty$ :

$$\begin{aligned} e_{Q_L, \infty}^2 &= E \left\{ \int_a^b [X(t) - Q_L X(t)] f(t) dt \right\}^2 \\ &= \int_a^b E [X(t) - Q_L X(t)] \\ &\quad \cdot [X(s) - Q_L X(s)] f(t) f(s) dt ds \\ &\leq E [X(t) - Q_L X(t)]^2 \left( \int_a^b |f| \right)^2, \end{aligned}$$

and by [3],

$$\begin{aligned} L^2 e_{Q_L, \infty}^2 &\leq L^2 E [X(t) - Q_L X(t)]^2 \left( \int_a^b |f| \right)^2 \\ &\rightarrow \frac{1}{L} \int_{-\infty}^{\infty} \frac{p}{h_Q^2} \left( \int_a^b |f| \right)^2. \end{aligned}$$

Again, when  $h_Q^*$  is used we have

$$\int_{-\infty}^{\infty} \frac{p(x)}{[h_Q^*(x)]^2} dx = 6\sqrt{3} \pi.$$

Thus for large  $L$ ,

$$e_{Q_L, \infty}^2 \leq \frac{1}{L^2} \frac{1}{12} \int \frac{p}{h_Q^2} \left( \int |f|^2 \right). \quad (42)$$

Also  $d_{Q_L} = E[X(t)Q_L X(t)] \rightarrow E[X^2(t)] = 1$ . From (40), (41), and (42) we see that for large  $L$  and  $n$ ,

$$\begin{aligned} e_{m, Q_L, n}^2 &\leq \frac{1}{L^2} \left\{ \frac{1}{12} \int \frac{p}{h_Q^2} \left( \int |f| \right)^2 \right\} \\ &\quad + \frac{1}{Ln^{3/2}} \left\{ \gamma(2\alpha_{X,0}/\pi)^{1/2} \int \frac{p}{h_Q} \int \frac{f^2}{h_s^{3/2}} \right\} \\ &\quad - \frac{1}{n^2} \frac{1}{12} \int A_X f \\ &\triangleq \frac{C}{L^2} + \frac{D}{Ln^{3/2}} + \frac{E}{n^2}. \end{aligned} \quad (43)$$

The bound (43) can be used to determine the allocation to quantization levels  $L$  and sample size  $n$  to achieve mean square error not exceeding a desirable value. Also, the more interesting problem of minimizing the total number of samples  $L+n$  (or some other function of  $L$  and  $n$  reflecting quantization and sampling costs), subject to mean square error not exceeding a desirable value, can be solved (numerically—analytic expressions being hard to obtain).

#### ACKNOWLEDGMENT

We are very thankful to a referee whose sharp reading of the manuscript helped us correct several inaccuracies and substantially improve the presentation.

#### APPENDIX I PROOF OF (15)

Here we find the rate of convergence and the corresponding asymptotic constants of the two terms on the right-hand side of (14). These results correct (and in fact the second also extends) the statements in [4] displayed between Eqs. (3.36) and (3.37); the final result in [4, Eq. (3.37)] remains correct. For simplicity of exposition throughout the following we do not display the terms of higher order and we use  $\approx$  to indicate equality up to higher order terms.

#### Sampling Points and Subdivision Points

The sampling points  $t_{n,k}$  are determined by  $\int_a^{t_{n,k}} h = (k-1/2)/n$ ,  $k=1, \dots, n$ . Introduce the interval subdivision points  $s_{n+1,k}$  by  $\int_a^{s_{n+1,k}} h = (k-1)/n$ ,  $k=1, \dots, n+1$ , so that each  $t_{n,k}$  is the median of  $(s_{n+1,k}, s_{n+1,k+1})$  with respect to  $h$ . For notational simplicity we drop the subscript  $n$  from  $t_{n,k}$  and  $n+1$  from  $s_{n+1,k}$ .

By the mean value theorem,

$$\frac{1}{n} = \int_{s_k}^{s_{k+1}} h = h(w_k)(s_{k+1} - s_k) \quad (44a)$$

$$\frac{1}{2n} = \int_{s_k}^{t_k} h = h(a_k)(t_k - s_k) \quad (44b)$$

$$\frac{1}{2n} = \int_{t_k}^{s_{k+1}} h = h(b_k)(s_{k+1} - t_k) \quad (44c)$$

where  $s_k < w_k < s_{k+1}$ ,  $s_k < a_k < t_k < b_k < s_{k+1}$ . It follows that as  $n \rightarrow \infty$ ,

$$\frac{t_k - s_k}{s_{k+1} - s_k} = \frac{1}{2} \frac{h(w_k)}{h(a_k)} \rightarrow \frac{1}{2} \quad (45)$$

so that asymptotically as  $n \rightarrow \infty$ ,  $t_k$  is the midpoint of  $(s_{k+1}, s_k)$ . We will need the order of magnitude of the difference of the two pieces,

$$D_k = (s_{k+1} - t_k) - (t_k - s_k).$$

Substituting  $h(u) = h(t_k) + (u - t_k)h'(u, \text{pt.})$  under the integral in (44b) and (44c), using the mean value theorem, and subtracting the resulting expressions, we obtain

$$D_k = -\frac{1}{8h(t_k)} \left\{ \frac{h'(b'_k)}{h^2(b_k)} + \frac{h'(a'_k)}{h^2(a_k)} \right\} \cdot \frac{1}{n^2}. \quad (46)$$

#### The First Term

We first consider the first term on the right-hand side of (14), which can be written as follows with  $K(u, v) = c(u)R_Y(u, v) \cdot c(v)$ , a symmetric function:

$$\begin{aligned} \text{1st term} &\triangleq \frac{1}{n^2} \sum_{k,j} c(t_k) R_Y(t_k, t_j) c(t_j) \\ &\quad - \iint_a^b R_Y(u, v) ch(u) ch(v) dudv \\ &= \sum_{k,j} \int_{s_k}^{s_{k+1}} \int_{s_j}^{s_{j+1}} [K(t_k, t_j) - K(u, v)] h(u) h(v) dudv \\ &\triangleq \sum_{h,j} I_{h,j}. \end{aligned}$$

The integral over each diagonal square is split into the six regions shown in Fig. 5, and by symmetry it suffices to consider

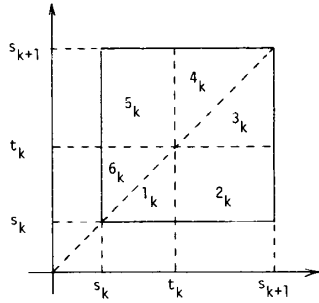


Fig. 5

only regions 1, 2, and 3. We use the Taylor expansion of the integrand:

$$\begin{aligned} K(t_k, t_k) - K(u, v) &= K(t_k, t_k) - K(t_k, v) \\ &\quad - (u - t_k) K^{1,0}(x_k, v) \\ &= -(v - t_k) K^{0,1}(t_k, y_k) \\ &\quad - (u - t_k) K^{1,0}(x_k, v) \end{aligned}$$

where  $x_k$  is between  $u$  and  $t_k$ , and  $y_k$  is between  $v$  and  $t_k$ . Over each region both  $u - t_k$  and  $v - t_k$  have constant sign so the mean value theorem is applicable. For region 1 we have

$$\begin{aligned} \iint_{1_k} &= K^{0,1}(a_k, b_k) h(a_k) h(b_k) \iint_{1_k} -(v - t_k) du dv \\ &\quad + K^{1,0}(a'_k, b'_k) h(a'_k) h(b'_k) \iint_{1_k} -(u - t_k) du dv \\ &= \left\{ \frac{1}{3} K^{0,1}(a_k, b_k) h(a_k) h(b_k) \right. \\ &\quad \left. + \frac{1}{6} K^{1,0}(a'_k, b'_k) h(a'_k) h(b'_k) \right\} (t_k - s_k)^3 \end{aligned}$$

where the points  $(a_k, b_k)$  and  $(a'_k, b'_k)$  belong to  $1_k$ . Using (44a) and (45) we obtain

$$\begin{aligned} n^2 \sum_k \iint_{1_k} &= \sum_k \left\{ \frac{1}{3} K^{0,1}(a_k, b_k) \frac{h(a_k) h(b_k)}{h^2(w_k)} \right. \\ &\quad \left. + \frac{1}{6} K^{1,0}(a'_k, b'_k) \frac{h(a'_k) h(b'_k)}{h^2(w_k)} \right\} \\ &\quad \cdot \left( \frac{t_k - s_k}{s_{k+1} - s_k} \right)^3 (s_{k+1} - s_k) \\ &\rightarrow \int_a^b \left\{ \frac{1}{24} K^{0,1}(t, t-) + \frac{1}{48} K^{1,0}(t+, t) \right\} dt. \end{aligned}$$

A similar argument gives

$$\begin{aligned} n^2 \sum_k \iint_{2_k} &\rightarrow \int_a^b \left\{ \frac{1}{16} K^{0,1}(t, t-) - \frac{1}{16} K^{1,0}(t+, t) \right\} dt \\ n^2 \sum_k \iint_{3_k} &\rightarrow \int_a^b \left\{ -\frac{1}{48} K^{0,1}(t, t-) - \frac{1}{24} K^{1,0}(t+, t) \right\} dt. \end{aligned}$$

Putting these together and using the symmetry of  $K$ , we have for the diagonal terms,

$$\begin{aligned} n^2 \sum_k I_{k,k} &\rightarrow \frac{1}{6} \int_a^b \{ K^{0,1}(t, t-) - K^{0,1}(t, t+) \} dt \\ &= \frac{1}{6} \int_a^b \alpha_{Y,0}(t) c^2(t) dt. \end{aligned} \quad (47)$$

Over each nondiagonal rectangle  $k \neq j$ ,  $K$  has continuous partial derivatives. Thus Taylor expanding  $K$  and retaining only the lowest degree terms, we obtain

$$I_{k,j} \approx \sum_{0 < p+q \leq 2} \frac{1}{p!q!} K^{p,q}(t_k, t_j) \cdot \int_{s_k}^{s_{k+1}} \int_{s_j}^{s_{j+1}} (u - t_k)^p (v - t_j)^q h(u) h(v) du dv$$

where in fact the term with  $p=1=q$  is of higher order. Using  $h(u) = h(t_k) + (u - t_k)h'(\text{int. pt.})$  and (46) we find

$$\begin{aligned} A_k &\triangleq \int_{s_k}^{s_{k+1}} (u - t_k) h(u) du \\ &\approx \left\{ -\frac{1}{16} \left[ \frac{h'(b'_k)}{h^2(b_k)} + \frac{h'(a'_k)}{h^2(a_k)} \right] + \frac{1}{12} \frac{h'(t_k)}{h^2(a_k)} \right\} \\ &\quad \cdot (s_{k+1} - s_k) \frac{1}{n^2} \\ &\approx -\frac{1}{24} \frac{h'(t_k)}{h^2(t_k)} (s_{k+1} - s_k) \frac{1}{n^2} \end{aligned} \quad (48)$$

where  $s_k < \tau_k < s_{k+1}$ , and likewise,

$$\begin{aligned} B_k &\triangleq \int_{s_k}^{s_{k+1}} (u - t_k)^2 h(u) du \approx \frac{1}{12} \frac{h(t_k)}{h^2(a_k)} (s_{k+1} - s_k) \frac{1}{n^2} \\ &\approx \frac{1}{12 h(t_k)} (s_{k+1} - s_k) \frac{1}{n^2}. \end{aligned} \quad (49)$$

It then follows (using (44a)) that

$$\begin{aligned} I_{k,j} &\approx - \left\{ K^{1,0}(t_k, t_j) \left( -\frac{1}{24} \right) \frac{h'(t_k)}{h^2(t_k)} h(w_j) \right. \\ &\quad + K^{0,1}(t_k, t_j) h(w_k) \left( -\frac{1}{24} \right) \frac{h'(t_j)}{h^2(t_j)} \\ &\quad + \frac{1}{2} K^{2,0}(t_k, t_j) \frac{1}{12 h(t_k)} h(w_j) \\ &\quad \left. + \frac{1}{2} K^{0,2}(t_k, t_j) h(w_k) \frac{1}{12 h(t_j)} \right\} \\ &\quad \cdot (s_{k+1} - s_k)(s_{j+1} - s_j) \frac{1}{n^2} \end{aligned}$$

and thus, by the symmetry of  $K$ ,

$$\begin{aligned} n^2 \sum_{k \neq j} I_{k,j} &\rightarrow \frac{1}{12} \iint_{t \neq s} \left\{ K^{0,1}(t, s) \frac{h(t) h'(s)}{h^2(s)} \right. \\ &\quad \left. - K^{0,2}(t, s) \frac{h(t)}{h(s)} \right\} dt ds \\ &= \frac{1}{12} \int_a^b dt h(t) \left( \int_a^t + \int_t^b \right) ds \frac{d}{ds} \left[ -\frac{K^{0,1}(t, s)}{h(s)} \right] \\ &= -\frac{1}{12} \int_a^b [K^{0,1}(t, t-) - K^{0,1}(t, t+)] dt \\ &\quad - \frac{1}{12} \int_a^b \left\{ \frac{K^{0,1}(t, b)}{h(b)} - \frac{K^{0,1}(t, a)}{h(a)} \right\} h(t) dt \\ &= -\frac{1}{12} \int \alpha_{Y,0} c^2 - \frac{1}{12} \int A_Y ch \end{aligned} \quad (50)$$

where  $A_Y$  is defined following (15).

Finally, adding (47) and (50) we obtain

$$n^2 [1st \text{ term}] \rightarrow \frac{1}{12} \int \alpha_{Y,0} c^2 - \frac{1}{12} \int A_Y c h. \quad (51)$$

*The Second Term*

We now consider the second term in (14):

$$\begin{aligned} \text{2nd term} &\triangleq \frac{1}{n} \sum_k c(t_k) \int_a^b R_{YX}(t_k, v) f(v) dv \\ &\quad - \iint_a^b R_{YX}(u, v) c h(u) f(u) dv \\ &= \sum_{k,j} \int_{s_k}^{s_{k+1}} \int_{s_j}^{s_{j+1}} [M(t_k, v) - M(u, v)] h(u) h(v) du dv \\ &\triangleq \sum_{k,j} J_{k,j}, \end{aligned}$$

where  $M(u, v) = c(u) R_{YX}(u, v) f(v) / h(v)$  is a generally non-symmetric function.

The integral  $J_{k,k}$  over each diagonal square is split into the six regions  $1_k$  to  $6_k$ . Over region  $1_k$ , using the Taylor expansion  $M(t_k, v) - M(u, v) = -(u - t_k) M^{1,0}(x_k, v)$ ,  $s_k < v < x_k < t_k$ , and the mean value theorem, we obtain

$$\begin{aligned} \iint_{1_k} &= M^{1,0}(a_k, b_k) h(a_k) h(b_k) \iint_{1_k} -(u - t_k) du dv \\ &= \frac{1}{6} M^{1,0}(a_k, b_k) h(a_k) h(b_k) (t_k - s_k)^3 \end{aligned}$$

where  $s_k < b_k < a_k < t_k$ , and thus by (44a) and (45),

$$\begin{aligned} n^2 \sum_k \iint_{1_k} &= \frac{1}{24} \sum_k M^{1,0}(a_k, b_k) \frac{h(b_k)}{h(a_k)} \frac{t_k - s_k}{s_{k+1} - s_k} (s_{k+1} - s_k) \\ &\rightarrow \frac{1}{48} \int_a^b M^{1,0}(t, t) dt. \end{aligned}$$

A similar argument gives

$$\begin{aligned} n^2 \sum_k \iint_{2_k} &\rightarrow -\frac{1}{16} \int_a^b M^{1,0}(t, t) dt, \\ n^2 \sum_k \iint_{4_k} &\rightarrow -\frac{1}{48} \int_a^b M^{1,0}(t, t) dt, \\ n^2 \sum_k \iint_{5_k} &\rightarrow \frac{1}{16} \int_a^b M^{1,0}(t, t) dt. \end{aligned}$$

The integrals over regions  $3_k$  and  $6_k$  are slightly more complex. For region  $3_k$  we use the Taylor expansion

$$\begin{aligned} M(t_k, v) - M(u, v) &= M(t_k, t_k) + (v - t_k) \\ &\quad \cdot [M^{0,1}(t_k, x_k) - M^{0,1}(u, x_k)] \\ &= -(u - t_k) M^{1,0}(y_k, t_k) + (v - t_k) \\ &\quad \cdot [M^{0,1}(t_k, x_k) - M^{0,1}(u, x_k)] \end{aligned}$$

where  $t_k < x_k$ ,  $y_k < u < x_{k+1}$ , and applying similarly the mean value theorem we obtain

$$\begin{aligned} n^2 \sum_k \iint_{3_k} &\rightarrow \int_a^b \left\{ -\frac{1}{24} M^{1,0}(t, t) \right. \\ &\quad \left. + \frac{1}{48} [M^{0,1}(t, t) - M^{0,1}(t, t)] \right\} dt. \end{aligned}$$

Likewise,

$$\begin{aligned} n^2 \sum_k \iint_{6_k} &\rightarrow \int_a^b \left\{ \frac{1}{24} M^{1,0}(t, t) \right. \\ &\quad \left. - \frac{1}{48} [M^{0,1}(t, t) - M^{0,1}(t, t)] \right\} dt. \end{aligned}$$

Finally, putting all six pieces together, we have for the diagonal terms

$$\begin{aligned} n^2 \sum_k J_{k,k} &\rightarrow \int_a^b \left\{ \frac{1}{12} [M^{1,0}(t, t) - M^{1,0}(t, t)] \right. \\ &\quad \left. - \frac{1}{24} [M^{0,1}(t, t) - M^{0,1}(t, t)] \right\} dt. \quad (52) \end{aligned}$$

Over each nondiagonal rectangle  $k \neq j$ ,  $M$  has continuous partial derivatives. Thus Taylor expanding  $M(t_k, v) - M(u, v)$  into

$$\begin{aligned} &-(u - t_k) M^{1,0}(t_k, t_j) - \frac{1}{2} (u - t_k)^2 M^{2,0}(t_k, t_j) \\ &\quad - (v - t_k)(v - t_j) M^{1,1}(t_k, t_j) \end{aligned}$$

plus third-order terms, and retaining only the lowest degree terms ( $M^{1,0}$  and  $M^{2,0}$ ) we obtain, using (48) and (49),

$$\begin{aligned} J_{k,j} &\approx -M^{1,0}(t_k, t_j) \int_{s_k}^{s_{k+1}} \int_{s_j}^{s_{j+1}} (u - t_k) h(u) h(v) du dv \\ &\quad - \frac{1}{2} M^{2,0}(t_k, t_j) \int_{s_k}^{s_{k+1}} \int_{s_j}^{s_{j+1}} (u - t_k)^2 h(u) h(v) du dv \\ &= \frac{1}{24} \left\{ M^{1,0}(t_k, t_j) \frac{h'(t_k)}{h^2(t_k)} h(w_j) - M^{2,0}(t_k, t_j) \frac{h(w_j)}{h(t_k)} \right\} \\ &\quad \cdot (s_{k+1} - s_k)(s_{j+1} - s_j) \frac{1}{n^2} \end{aligned}$$

and thus

$$\begin{aligned} n^2 \sum_{k \neq j} J_{k,j} &\rightarrow \frac{1}{24} \iint_{t \neq s} \left\{ M^{1,0}(t, s) \frac{h'(t) h(s)}{h^2(t)} \right. \\ &\quad \left. - M^{2,0}(t, s) \frac{h(s)}{h(t)} \right\} dt ds \\ &= -\frac{1}{24} \int_a^b [M^{1,0}(s, s) - M^{1,0}(s, s)] ds \\ &\quad - \frac{1}{24} \int_a^b \left\{ \frac{M^{1,0}(b, s)}{h(b)} - \frac{M^{1,0}(a, s)}{h(a)} \right\} h(s) ds \quad (53) \end{aligned}$$

where the last equality is obtained integrating by parts as in (50).

Finally, adding (52) and (53), we find

$$\begin{aligned} n^2 [2nd \text{ term}] &\rightarrow \frac{1}{24} \int_a^b \left\{ [M^{1,0}(t, t) - M^{1,0}(t, t)] \right. \\ &\quad \left. - [M^{0,1}(t, t) - M^{0,1}(t, t)] \right\} dt \\ &\quad - \frac{1}{24} \int_a^b \left\{ \frac{M^{1,0}(b, t)}{h(b)} - \frac{M^{1,0}(a, t)}{h(a)} \right\} h(t) dt \\ &= \frac{1}{24} \int \beta_{YX} \frac{cf}{h} - \frac{1}{24} \int A_{YX} f \quad (54) \end{aligned}$$

where  $\beta_{YX}$  and  $A_{YX}$  are defined following (15).

Now (15) follows from (14), (51), and (54).

APPENDIX II  
EXPRESSIONS FOR THE CONSTANTS IN (24) AND (25)

Using the method of solving the integral equation (5) described in [8, Ch. III, Sect. 7] we find

$$\begin{aligned}\alpha^2 &= a_X a_N \frac{\sigma_X^2 a_N + \sigma_N^2 a_X}{\sigma_X^2 a_X + \sigma_N^2 a_N}, \\ M_1 &= \frac{\sigma_X^2 a_N}{\sigma_X^2 a_N + \sigma_N^2 a_X}, \\ M_2 &= \frac{\sigma_X^2 a_X \sigma_N^2 a_N (\sigma_X^2 + \sigma_N^2) (a_X^2 - a_N^2)}{(\sigma_X^2 a_X + \sigma_N^2 a_N) (\sigma_X^2 a_N + \sigma_N^2 a_X) S}, \\ M &= \frac{\sigma_X^2 \sigma_N^2 (a_N - a_X) \alpha (1 - e^{-\alpha})}{(\sigma_X^2 a_N + \sigma_N^2 a_X) S},\end{aligned}$$

where  $S = \alpha(1 - e^{-\alpha})(\sigma_X^2 a_X + \sigma_N^2 a_N) + (1 + e^{-\alpha})a_X a_N (\sigma_X^2 + \sigma_N^2)$ .

APPENDIX III  
PROOF OF (30)

With the notation introduced in Appendix I we have (writing  $R$  for  $R_Y$ )

$$\begin{aligned}\text{1st term} &= \frac{1}{n^2} \sum_{k,j} c(t_k) R(t_k - t_j) c(t_j) \\ &\quad - \iint_a^b R(u-v) ch(u) ch(v) dudv \\ &= \sum_{k,j} \int_{s_k}^{s_{k+1}} \int_{s_j}^{s_{j+1}} [R(t_k - t_j) c(t_k) c(t_j) \\ &\quad - R(u-v) c(u) c(v)] h(u) h(v) dudv \\ &\triangleq \sum_{k,j} I_{k,j}.\end{aligned}\quad (55)$$

We first concentrate on the diagonal terms  $I_{k,k}$  and split each square  $(s_k, s_{k+1}) \times (s_k, s_{k+1})$  into its upper and lower triangles. Then

$$\begin{aligned}&\iint_{\text{lower triangle}} [R(0) c^2(t_k) - R(u-v) c(u) c(v)] h(u) h(v) dudv \\ &= \iint_{\text{lower triangle}} [R(0) \{c^2(t_k) - c(u) c(v)\} \\ &\quad - R^{(1/2)}(0+) \sqrt{u-v} c(u) c(v) + o(\sqrt{u-v}) c(u) c(v)] \\ &\quad \cdot h(u) h(v) dudv \\ &\simeq -R^{(1/2)}(0+) \iint_{\text{lower triangle}} \sqrt{u-v} ch(u) ch(v) dudv \\ &= -R^{(1/2)}(0+) ch(u_k) ch(v_k) \iint_{\text{lower triangle}} \sqrt{u-v} dudv \\ &= -\frac{4}{15} R^{(1/2)}(0+) ch(u_k) ch(v_k) (s_{k+1} - s_k)^{5/2}\end{aligned}$$

for some  $(u_k, v_k)$  in the lower triangle, where  $\simeq$  indicates equality up to higher order terms. Similarly, the integral over the upper triangle has main term

$$\frac{4}{15} R^{(1/2)}(0-) ch(a_k) ch(b_k) (s_{k+1} - s_k)^{5/2}$$

for some  $(a_k, b_k)$  in the upper triangle. Using (44a) we then obtain

$$n^{3/2} \sum_k I_{k,k} \rightarrow \frac{4}{15} [R^{(1/2)}(0-) - R^{(1/2)}(0+)] \int \frac{(ch)^2}{h^{3/2}}. \quad (56)$$

For the nondiagonal terms  $k \neq j$ , Taylor-expanding  $R$  and  $c$  we obtain, writing  $c_k$  for  $c(t_k)$ ,  $c'_k$  for  $c'(t_k)$ , etc.,

$$\begin{aligned}I_{k,j} &= -R(t_k - t_j) \left\{ c_k c'_j \frac{A_j}{n} + c'_k c_j \frac{A_k}{n} + \frac{1}{2} c_k c''_j \frac{B_j}{n} + \frac{1}{2} c'_k c'_j \frac{B_k}{n} \right. \\ &\quad \left. + c_k c_j A_k A_j \right\} \\ &\quad - R'(t_k - t_j) \left\{ c_k c_j \left( \frac{A_k}{n} - \frac{A_j}{n} \right) \right. \\ &\quad \left. + c'_k c_j \left( \frac{B_k}{n} - A_k A_j \right) + c_k c'_j \left( A_k A_j - \frac{B_j}{n} \right) \right\} \\ &\quad - R''(t_k - t_j) \frac{1}{2} c_k c_j \left( \frac{B_k}{n} + \frac{B_j}{n} - 2A_k A_j \right) \\ &\quad + \text{higher order terms}\end{aligned}$$

where  $A_k, B_k$  are given in (48), (49). All terms except for those retained below have rate  $n^{-2}$ , so that the dominant term is

$$\sum_{k \neq j} I_{k,j} \simeq -\frac{1}{2} \sum_{k \neq j} R''(t_k - t_j) c(t_k) c(t_j) \left( \frac{B_k}{n} + \frac{B_j}{n} \right)$$

and by (49) and (44a),

$$\sum_{k \neq j} I_{k,j} \simeq -\frac{1}{24n^4} \sum_{k \neq j} R''(t_k - t_j) c(t_k) c(t_j) \cdot \left\{ \frac{1}{h^2(t_k)} + \frac{1}{h^2(t_j)} \right\}.$$

The double sum above over all terms with  $|t_k - t_j| \geq \epsilon$  has rate  $n^{-2}$ , in fact,

$$n^2 \sum_{|t_k - t_j| \geq \epsilon} I_{k,j} \rightarrow -\frac{1}{24} \iint_{|t-s| \geq \epsilon} R''(t-s) c(t) c(s) \cdot \left\{ \frac{h(s)}{h(t)} + \frac{h(t)}{h(s)} \right\} dt ds$$

and the limiting constant is finite. On the other hand, the same argument shows that  $n^2 \sum_{|t_k - t_j| < \epsilon} I_{k,j}$  tends to the same integral as above but over  $|t-s| < \epsilon$ , which is not finite, as integration by parts and  $R'(0\pm) = \infty$  show. Thus its rate is slower and it becomes the dominant term:

$$\sum_{k \neq j} I_{k,j} \simeq -\frac{1}{24n^4} \sum_{|t_k - t_j| < \epsilon} R''(t_k - t_j) c(t_k) c(t_j) \cdot \left\{ \frac{1}{h^2(t_k)} + \frac{1}{h^2(t_j)} \right\}$$

for each  $\epsilon > 0$ .

Differentiating the expression of  $R$  in terms of  $R_X$  and  $R_N$  we obtain

$$R''(\tau) = \frac{2K^2}{\pi} \cdot \frac{R_X(\tau) [R_X'(\tau)]^2 + R_X''(\tau)}{[1 - R_X^2(\tau)]^{1/2}} + R_N''(\tau).$$

Assuming that  $R_X''(0\pm)$  and  $R_N''(0\pm)$  exist and are finite, it follows from

$$\frac{|\tau|}{1 - R_X^2(\tau)} \xrightarrow{\tau \rightarrow 0} \frac{1}{\alpha_{X,0}} \quad (57a)$$

that

$$|\tau|^{3/2} R''(\tau) \xrightarrow{\tau \rightarrow 0} \frac{2K^2}{\pi} \cdot \frac{\alpha_{X,0}^{1/2}}{4} = \frac{1}{8} \alpha_{Y,1/2}. \quad (57b)$$

Thus for very small  $\epsilon$  ( $\epsilon \approx 0$ ),

$$R''(t_k - t_j) \approx \frac{1}{8} \alpha_{Y,1/2} \frac{1}{|t_k - t_j|^{3/2}}.$$

Also for  $\epsilon \approx 0$  and, say,  $k > j$ , we have

$$\frac{1}{n} = \int_{t_i}^{t_{i+1}} h = h(t'_i) \Delta t_i \approx h(t_j) \Delta t_i$$

for  $i = j, \dots, k$ , and thus summing up,

$$\frac{k-j}{n} \approx h(t_j)(t_k - t_j),$$

and

$$R''(t_k - t_j) \approx \frac{1}{8} \alpha_{Y,1/2} \frac{h^{3/2}(t_j) n^{3/2}}{(k-j)^{3/2}}.$$

Likewise, when  $j > k$ ,

$$R''(t_k - t_j) \approx \frac{1}{8} \alpha_{Y,1/2} \frac{h^{3/2}(t_k) n^{3/2}}{(j-k)^{3/2}}.$$

Using also  $c(t_k) \approx c(t_j)$ ,  $h(t_k) \approx h(t_j)$ , we obtain that for  $\epsilon \approx 0$ ,

$$\sum_{k \neq j} I_{k,j} = -\frac{1}{96n^{5/2}} \alpha_{Y,1/2} \left\{ \sum_{0 < t_k - t_j < \epsilon} \frac{1}{(k-j)^{3/2}} \frac{c^2(t_j)}{h^{1/2}(t_j)} + \sum_{0 < t_j - t_k < \epsilon} \frac{1}{(j-k)^{3/2}} \frac{c^2(t_k)}{h^{1/2}(t_k)} \right\}.$$

As the sum extends over  $0 < k-j < nh(t_j)\epsilon$  and  $0 < j-k < nh(t_k)\epsilon$ , we obtain in the limit, in view of (44a),

$$\begin{aligned} n^{3/2} \sum_{k \neq j} I_{k,j} &\rightarrow -\frac{1}{96} \alpha_{Y,1/2} \left\{ \sum_{m=1}^{\infty} \frac{1}{m^{3/2}} \int \frac{c^2}{h^{1/2}} h \right. \\ &\quad \left. + \sum_{m=1}^{\infty} \frac{1}{m^{3/2}} \int \frac{c^2}{h^{1/2}} h \right\} \\ &= -\frac{1}{48} \alpha_{Y,1/2} \zeta(3/2) \int c^2 h^{1/2}. \end{aligned} \quad (58)$$

These arguments for  $\epsilon \approx 0$  can be made precise by writing upper and lower bounds in terms of  $\epsilon$  and then letting  $\epsilon \downarrow 0$  as these hold for all  $\epsilon > 0$ .

Combining (56) with (58) gives (30).

#### APPENDIX IV PROOF OF (36) AND (37)

We have

$$\begin{aligned} R_{QX}(\tau) &= E[Q(X(0))Q(X(\tau))] \\ &= \iint_{-\infty}^{\infty} Q(x)Q(y)p(x,y;R_X(\tau))dxdy \end{aligned}$$

where  $p(x,y;\rho)$  is the bivariate normal density with zero means, variances one, and correlation coefficient  $\rho$ . Thus, by Price's theorem [7], we obtain

$$\begin{aligned} R'_{QX}(\tau) &= \frac{\partial R_{QX}}{\partial R_X} \cdot R'_X(\tau) \\ &= R'_X(\tau) \iint p(x,y;R_X(\tau)) dQ(x) dQ(y) \\ &= R'_X(\tau) \sum_{i,j=2}^L p(x_i, x_j; R_X(\tau)) (y_i - y_{i-1})(y_j - y_{j-1}) \\ &= \frac{R'_X(\tau)}{2\pi\sqrt{1-R_X^2(\tau)}} \left\{ \sum_{i=2}^L (y_i - y_{i-1})^2 \exp\left[-\frac{x_i^2}{1+R_X(\tau)}\right] \right. \\ &\quad \left. + \sum_{i \neq j=2}^L (y_i - y_{i-1})(y_j - y_{j-1}) \right. \\ &\quad \left. \cdot \exp\left[-\frac{x_i^2 + x_j^2 - 2x_i x_j R_X(\tau)}{2(1-R_X^2(\tau))}\right] \right\}. \end{aligned}$$

It follows that

$$\begin{aligned} R_{QX}^{(1/2)}(0+) &= \lim_{\tau \downarrow 0} \frac{R_{QX}(\tau) - R_{QX}(0)}{\sqrt{\tau}} = \lim_{\tau \downarrow 0} 2\sqrt{\tau} R'_{QX}(\tau) \\ &= \frac{R'_X(0+)}{\pi \{-2R'_X(0+)\}^{1/2}} \sum_{i=2}^L (y_i - y_{i-1})^2 e^{-x_i^2/2} \\ &= -\frac{1}{\pi} \{- (1/2) R'_X(0+)\}^{1/2} \\ &\quad \sum_{i=2}^L (y_i - y_{i-1})^2 e^{-(y_i + y_{i-1})^2/8} \\ &\triangleq -\{-R'_X(0+)/\pi\}^{1/2} B_Q = -\{\alpha_{X,0}/2\pi\}^{1/2} B_Q. \end{aligned}$$

Similarly,

$$R_{QX}^{(1/2)}(0-) = \{\alpha_{X,0}/2\pi\}^{1/2} B_Q$$

and thus

$$\alpha_{QX,1/2} = R_{QX}^{(1/2)}(0-) - R_{QX}^{(1/2)}(0+) = \{2\alpha_{X,0}/\pi\}^{1/2} B_Q.$$

This establishes (36). For (37), differentiating once the expression above for  $R'_{QX}(\tau)$ , we obtain

$$\begin{aligned} R''_{QX}(\tau) &= \frac{1}{2\pi} \left\{ \frac{R'_X(\tau)}{[1-R_X^2(\tau)]^{1/2}} + \frac{R_X(\tau)[R'_X(\tau)]^2}{[1-R_X^2(\tau)]^{3/2}} \right\} \left\{ \sum_{i=2}^L (y_i - y_{i-1})^2 \exp\left[-\frac{x_i^2}{1+R_X(\tau)}\right] \right. \\ &\quad \left. + \sum_{i \neq j=2}^L (y_i - y_{i-1})(y_j - y_{j-1}) \exp\left[-\frac{x_i^2 + x_j^2 - 2x_i x_j R_X(\tau)}{2(1-R_X^2(\tau))}\right] \right\} + \frac{R'_X(\tau)}{2\pi[1-R_X^2(\tau)]^{1/2}} \\ &\quad \cdot \left\{ \sum_{i=2}^L (y_i - y_{i-1})^2 \exp\left[-\frac{x_i^2}{1+R_X(\tau)}\right] \frac{x_i^2 R'_X(\tau)}{[1+R_X(\tau)]^2} + \sum_{i \neq j=2}^L (y_i - y_{i-1})(y_j - y_{j-1}) \right. \\ &\quad \left. \cdot \exp\left[-\frac{x_i^2 + x_j^2 - 2x_i x_j R_X(\tau)}{2(1-R_X^2(\tau))}\right] \left[ -\frac{R'_X(\tau)}{1-R_X^2(\tau)} \right] \left[ (x_i^2 + x_j^2) R_X(\tau) - x_i x_j (1+R_X^2(\tau)) \right] \right\} \end{aligned}$$



and using (57a) we have as  $\tau \rightarrow 0$ ,

$$\begin{aligned} |\tau|^{3/2} R''_{QX}(\tau) &\rightarrow \frac{1}{2\pi} \left( \frac{\alpha_{X,0}}{2} \right)^2 \left( \frac{1}{\alpha_{X,0}} \right)^{3/2} \sum_{l=2}^L (y_l - y_{l-1})^2 e^{-x_l^2/2} \\ &= \frac{1}{4} \left\{ \alpha_{X,0}/(2\pi) \right\}^{1/2} B_Q = \frac{1}{8} \alpha_{QX,1/2}, \end{aligned}$$

establishing (37).

## REFERENCES

- [1] F. J. Barrett and D. G. Lampard, "An expansion for some second-order probability distributions and its application to noise problems," *IRE Trans. Inform. Theory*, vol. IT-1, pp. 10–15, Mar. 1955.
- [2] S. Cambanis, "Sampling designs for time series," in *Handbook of Statistics*, Volume 5: Time Series in Time Domain, E. J. Hannan, P. R. Krishnaiah, and M. M. Rao, Eds. Amsterdam: North-Holland, 1985, pp. 337–362.
- [3] S. Cambanis and N. L. Gerr, "A simple class of asymptotically optimal quantizers," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 666–676, Sept. 1983.
- [4] S. Cambanis and E. Masry, "Sampling designs for the detection of signals in noise," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 83–104, Jan. 1983.
- [5] A. E. Jones, "Systematic sampling of continuous parameter populations," *Biometrika*, vol. 35, pp. 283–290, 1948.
- [6] M. G. Kendall, "Continuation of Dr. Jones' paper," *Biometrika*, vol. 35, pp. 291–296, 1948.
- [7] R. Price, "A useful theorem for nonlinear devices having Gaussian inputs," *IRE Trans. Inform. Theory*, vol. IT-4, pp. 69–72, Jan. 1958.
- [8] Yu. A. Rozanov, *Stationary Random Processes*. San Francisco: Holden-Day, 1967.
- [9] J. Sacks and D. Ylvisaker, "Designs for regression problems with correlated errors," *Ann. Math. Statist.*, vol. 37, pp. 68–89, 1966.
- [10] —, "Designs for regression problems with correlated errors III," *Ann. Math. Statist.*, vol. 41, pp. 2057–2074, 1970.
- [11] —, "Statistical designs and integral approximation," in *Proc. Twelfth Biennial Seminar of the Canadian Math. Congress*, pp. 115–136, Canadian Math. Congress, Montreal, 1970.
- [12] C. Schoenfelder, "Random designs for estimating integrals of stochastic processes," *Institute of Statistics Mimeo Series 1201*, University of North Carolina, Chapel Hill, NC, 1978.