Stochastic Processes and their Applications 46 (1993) 47-89
North-Holland

Sampling designs for estimation of a random process

Yingcai Su and Stamatis Cambanis

Department of Statistics, University of North Carolina, Chapel Hill, NC, USA

Received 8 November 1991
Revised 3 April 1992

A random process $X(t)$, $t \in [0, 1]$, is sampled at a finite number of appropriately designed points. On the basis of these observations, we estimate the values of the process at the unsampled points and we measure the performance by an integrated mean square error. We consider the case where the process has a known, or partially or entirely unknown mean, i.e., when it can be modeled as $X(t) = m(t) + N(t)$, where $m(t)$ is nonrandom and $N(t)$ is random with zero mean and known covariance function. Specifically, we consider (1) the case where $m(t)$ is known, (2) the semiparametric case where $m(t) = \beta_1 f_1(t) + \cdots + \beta_q f_q(t)$, the β_i's are unknown coefficients and the f_i's are known regression functions, and (3) the nonparametric case where $m(t)$ is unknown. Here $f_i(t)$ and $m(t)$ are of comparable smoothness with the purely random part $N(t)$, and $N(t)$ has no quadratic mean derivative. Asymptotically optimal sampling designs are found for cases (1), (2) and (3) when the best linear unbiased estimator (BLUE) of $X(t)$ is used (a nearly BLUE in case (3)), as well as when the simple nonparametric linear interpolator of $X(t)$ is used. Also it is shown that the mean has no effect asymptotically, and several examples are considered both analytically and numerically.

sampling designs * interpolation of random processes * effect of the mean

1. Introduction, results and examples

This paper deals with the following problem of estimating a random process from a finite number of observations, which arises in statistical communication theory and signal processing as well as in geology (Journel and Huijbregts, 1978) and environmental science (Christakos, 1991).

Suppose a random process $X(t)$, $t \in [0, 1]$, is sampled at a finite number of appropriately designed points. On the basis of these observations, we want to estimate the values of the process at the unsampled points and we measure the performance by an integrated mean square error (IMSE).

The process can be modeled as

$$X(t) = m(t) + N(t), \quad t \in [0, 1]. \tag{1.1}$$

Correspondence to: Dr. Yingcai Su, Department of Statistics, University of Arizona, Tucson, AZ 85721, USA.
Research supported by the Air Force Office of Scientific Research Contract No. 91-0030.

Here $m(t)$ is the nonrandom large-scale mean structure and we consider (1) the case where $m(t)$ is known or, equivalently equals zero; (2) the semiparametric (regression) model where the mean can be modeled as $m(t)=\beta_1 f_1(t)+\cdots+\beta_q f_q(t)$, where the β_i's are unknown coefficients and the f_i's are known (regression) functions; and (3) the nonparametric case where the macroscopic mean structure $m(t)$ is unknown. $N(t)$ is the small-scale random structure which models the temporal dependence and has zero means and known covariance function $R(t,s)=\mathscr{E}N(t)N(s)$. The centered process N is assumed to have no quadratic mean derivative and the functions $m(t)$ and $f_i(t)$ are of comparable smoothness with the microscopic purely random part $N(t)$ (specifically, $m(t)$ and $f_i(t)$ are of the form $\int_0^1 R(t,s)\psi(s)\,\mathrm{d}s$).

There are three findings. The main one is that simple sampling designs are found which are asymptotically optimal as the sample size increases to infinity. This is done for a variety of estimators. First the best linear unbiased estimator (BLUE) of $X(t)$ is used in cases (1) and (2) and a nearly BLUE in the nonparametric case (3) (Theorems 1, 2 and 4). The second finding is that asymptotically the mean has no effect on the overall performance and can therefore be neglected (Theorems 3 and 4). This quantifies the discussions in Journel and Rossi (1989) and Sacks et al. (1989, p. 415). However, an example (Example 2) shows that the mean function may cause some perturbation on the optimal sampling design points. The third finding is that the very simple nonparametric linear interpolation also leads to an asymptotically optimal performance (Theorem 6).

If the centered process $N(t)$ has exactly k $(k=1,2,\ldots)$ quadratic mean derivatives, the convergence rate of the IMSE for the corresponding BLUE estimators is likely to be $n^{-(k+1)}$ (compared with Theorem 1), but we do not investigate further this conjecture.

The basic setup

For the model (1.1), data

$$X'_{T_n}=(X(t_{n,1}),\ldots,X(t_{n,n}))$$

are observed at sampling points $T_n=\{t_{n,i}\}_1^n$, $t_{n,i}\in[0,1]$, $i=1,\ldots,n$, and it is desired to estimate $X(t)$ at every (unsampled) point $t\in[0,1]$ by a linear estimator

$$X_{T_n}(t)=\sum_{i=1}^n C_i(t)X(t_{n,i})\triangleq C'_{T_n}(t)X_{T_n}$$

with coefficient functions $C_i(t)$, $i=1,\ldots,n$, where $C'_{T_n}(t)=(C_1(t),\ldots,C_n(t))$. The mean square estimation error of $X(t)$ by $X_{T_n}(t)$ can be written as

$$\mathrm{MSE}_{T_n}(t)=\mathscr{E}\{X_{T_n}(t)-X(t)\}^2=V_{T_n}(t)+B^2_{T_n}(t),$$

where

$$V_{T_n}(t)=R(t,t)-2C'_{T_n}(t)r_{T_n}(t)+C'_{T_n}(t)R_{T_n}C_{T_n}(t),$$

$$B_{T_n}(t)=C'_{T_n}(t)m_{T_n}-m(t)$$

are the variance and bias of the estimator $X_{T_n}(t)$, respectively, and the following notations are used:

$$r'_{T_n}(t)=(R(t_{n,1},t),\ldots,R(t_{n,n},t)),$$

$$m'_{T_n}=(m(t_{n,1}),\ldots,m(t_{n,n})),\qquad R_{T_n}=(R(t_{n,i},t_{n,j}))_{n\times n}.$$

For every fixed sampling design T_n, the best linear unbiased estimator corresponds to those coefficients $\hat{C}_{T_n}(t)$ which minimize the variance $V_{T_n}(t)$ subject to some unbiasedness condition which takes different forms in cases (1), (2) and (3). Consequently, the BLUE and its MSE have different expressions depending on the form of the mean, which will be specified for these three cases later in this section.

For fixed t, the $\mathrm{MSE}_{T_n}(t)$ will of course vanish for any choice of sampling points T_n containing the point t. However, we are interested in designing sampling points T_n with small estimation error over the entire interval $[0,1]$ of estimation. We thus use as performance criterion an integrated mean square error (IMSE) with weight function $W(t)$,

$$\begin{aligned}\mathrm{IMSE}_{T_n}(W)&=\int_0^1 \mathrm{MSE}_{T_n}(t)\,W(t)\,\mathrm{d}t\\&=\int_0^1 V_{T_n}(t)\,W(t)\,\mathrm{d}t+\int_0^1 B^2_{T_n}(t)\,W(t)\,\mathrm{d}t\\&\triangleq V_{T_n}+B^2_{T_n},\end{aligned}\tag{1.2}$$

where $W(t)$ is a positive continuous probability density function on $(0,1)$, and the sampling points T_n are so chosen that the IMSE is as close to zero as possible. (Sacks, Schiller and Welsh (1989) found numerically two-dimensional sampling design points which minimize $\mathrm{IMSE}_{T_n}(1)$ for the semiparametric regression model (2) with Ornstein–Uhlenbeck error process and certain values of the sample size n and also provided some interesting discussion. However, here we consider only the one-dimensional case.)

For fixed n, it is not generally easy to find n design points T_n which minimize the $\mathrm{IMSE}_{T_n}(W)$. To avoid this problem we adopt the techniques of Sacks and Ylvisaker (1966) to find an asymptotically optimal sequence of sampling designs $\{T_n^*\}_1^\infty$ satisfying

$$\lim_{n\to\infty}\mathrm{IMSE}_{T_n^*}(W)/\inf_{|T_n|=n}\mathrm{IMSE}_{T_n}(W)=1,\tag{1.3}$$

where the infimum is taken over all designs T_n of sample size n.

Recall that the regular sampling designs determined by a density function h on $[0,1]$ are $T_n(h)=\{t_{n,i}\}_1^n$ with $t_{n,1}=0$ and

$$\int_{t_{n,i}}^{t_{n,i+1}} h(t)\,\mathrm{d}t=1/(n-1),\quad i=1,\ldots,n-1,\tag{1.4}$$

namely, the regular sampling design points divide the area enclosed by h (equal to one) into $n-1$ subregions each with area $1/(n-1)$.

Conditions on the covariance function

We consider centered random processes $N(t)$ with no quadratic mean derivative, such as Wiener and Ornstein-Uhlenbeck processes, and need the following technical assumptions on their covariance function.

Assumption (C1). The centered process $X(t)-m(t)=N(t)$ has no quadratic mean derivative (i.e., $R^{1,1}(t,s)$ does not exist at the diagonal) but its covariance function $R(t,s)$ has continuous and bounded mixed derivatives up to order two off the diagonal, at the diagonal the limits $R^{p,q}(u,u\pm)=\lim_{(t,s)\to(u,u\pm)} R^{p,q}(t,s)$ from below $(t>s)$ and from above $(t<s)$ exist for $0\leqslant p+q\leqslant 2$ and are continuous functions of $u\in[0,1]$, and the continuous jump function

$$\alpha(t)=R^{0,1}(t,t-)-R^{0,1}(t,t+)$$

is nonnegative and not identically zero on $[0,1]$. Also the matrix $R_{T_n}=(R(t_{n,i},t_{n,j}))_{n\times n}$ is invertible for every $T_n=\{t_{n,i}\}_1^n$.

Assumption (C2). For each $t\in[0,1]$, the function $R^{0,2}(\cdot,t)$ belongs to RKHS(R), the reproducing kernel Hilbert space of $R(\cdot,\cdot)$, and its RKHS norm $\|R^{0,2}(\cdot,t)\|_R$ is bounded over $[0,1]$.

Assumption (C1) contains the usual regularity conditions needed in the asymptotic analysis of sampling design problems (Cambanis, 1985; Sacks and Ylvisaker, 1966). Assumption (C2) simplifies the proofs of Lemma 2 and Theorem 2 in the next section, but as Sacks and Ylvisaker (1966) point out, it is a rather restrictive assumption in the presence of (C1) and it is not clear whether it is necessary for our results.

The simplest examples of zero mean processes which satisfy Assumptions (C1) and (C2) are Wiener process with $R(t,s)=\sigma^2\min(t,s)$, for which $\alpha(t)=\sigma^2$ and $R^{0,2}(\cdot,t)\equiv 0$; the process with triangular covariance function $R(t,s)=1-\mu|t-s|$ if $|t-s|\leqslant 1/\mu$ and $R(t,s)=0$, otherwise, for which $\alpha(t)=2\mu$ and $R^{0,2}(\cdot,t)\equiv 0$; and Ornstein-Uhlenbeck process with $R(t,s)=\sigma^2 e^{-\mu|t-s|}$, for which $\alpha(t)=2\mu\sigma^2$ and $R^{0,2}(t,s)=\mu^2\sigma^2 e^{-\mu|t-s|}=\mu^2R(t,s)$. Sacks and Ylvisaker (1966) discuss some further interesting classes of examples.

(1) Zero mean ($m\equiv 0$)

Here, the mean square error contains only the variance term $V_{T_n}(t)$ and the minimum variance estimator $X^0_{T_n}(t)$ of $X(t)$ has coefficients $r'_{T_n}(t)R^{-1}_{T_n}$, which minimize $V_{T_n}(t)$ for any fixed sampling design T_n, i.e.,

$$X^0_{T_n}(t)=r'_{T_n}(t)R^{-1}_{T_n}X_{T_n}, \tag{1.5}$$

where the superscript '0' indicates the zero mean model. The corresponding MSE and IMSE are

$$\mathrm{MSE}^0_{T_n}(t) = V_{T_n}(t) = R(t,t) - r'_{T_n}(t) R^{-1}_{T_n} r_{T_n}(t), \tag{1.6}$$

$$\mathrm{IMSE}^0_{T_n}(W) = \int_0^1 \{R(t,t) - r'_{T_n}(t) R^{-1}_{T_n} r_{T_n}(t)\} W(t)\,\mathrm{d}t. \tag{1.7}$$

We have the following results.

Theorem 1. *When $m \equiv 0$ and Assumptions* (C1) *and* (C2) *hold, the following are true.*

(i) *If the function $(\alpha W)^{1/2}$ is Riemann integrable, then*

$$\lim_{n\to\infty} n \inf_{|T_n|=n} \mathrm{IMSE}^0_{T_n}(W) = \frac{1}{6}\left\{\int_0^1 (\alpha(t) W(t))^{1/2}\,\mathrm{d}t\right\}^2,$$

where the infimum is taken over all sampling designs of size n.

(ii) *If the function $\alpha W/h$ is Riemann integrable, then*

$$\lim_{n\to\infty} n\, \mathrm{IMSE}^0_{T_n(h)}(W) = \frac{1}{6}\int_0^1 \frac{\alpha(t)}{h(t)} W(t)\,\mathrm{d}t. \tag{1.8}$$

(iii) *The regular sequence of sampling designs $\{T_n(h_o)\}_1^\infty$ determined by the density function*

$$h_o(t) = \{\alpha(t) W(t)\}^{1/2} \Big/ \int_0^1 \{\alpha(u) W(u)\}^{1/2}\,\mathrm{d}u$$

is asymptotically optimal, provided $(\alpha W)^{1/2}$ is Riemann integrable.

When $W(t) = 1$ and the process is stationary, the asymptotically optimal sampling design is uniform, as one would have expected.

(2) Semiparametric (regression) model

Here, the mean $m(t)$ is specified as follows. For some finite integer q,

$$m(t) = \sum_{i=1}^{q} \beta_i f_i(t), \tag{1.9}$$

where the β_i's are unknown coefficients and the known (regression) functions f_i are of the form

$$f_i(t) = \int_0^1 R(t,s)\phi_i(s)\,\mathrm{d}s, \quad t \in [0,1],\ i = 1, \ldots, q, \tag{1.10}$$

where each $\phi_i(\cdot)$ is a continuous function on [0, 1]. Then the BLUE of $X(t)$ is

$$X^q_{T_n}(t) = \{r'_{T_n}(t) - [F'_{T_n} R^{-1}_{T_n} r_{T_n}(t) - F(t)]'(F'_{T_n} R^{-1}_{T_n} F_{T_n})^{-1} F'_{T_n}\} R^{-1}_{T_n} X_{T_n}, \tag{1.11}$$

where

$$F'(t)=(f_1(t),\dots,f_q(t))$$

and

$$F_{T_n}=\{f_j(t_{n,i})\}_{j=1,\dots,q}^{i=1,\dots,n}=(f_{1,T_n},\dots,f_{q,T_n})=(F(t_{n,1}),\dots,F(t_{n,n}))'$$

(see Stein, 1989). The error can be written as

$$\begin{aligned}X(t)-X^q_{T_n}(t)&=[N(t)-r'_{T_n}(t)R^{-1}_{T_n}N_{T_n}]\\&\quad+[F'_{T_n}R^{-1}_{T_n}r_{T_n}(t)-F(t)]'(F'_{T_n}R^{-1}_{T_n}F_{T_n})^{-1}F'_{T_n}R^{-1}_{T_n}N_{T_n}.\end{aligned}\tag{1.12}$$

Note that the first term in brackets is the same as the estimation error in case (1) and $r'_{T_n}(t)R^{-1}_{T_n}N_{T_n}$ is the projection of $N(t)$ onto the linear space of N_{T_n}. Thus, the two terms in (1.12) are orthogonal. It then follows that the MSE of $X^q_{T_n}(t)$ is

$$\mathrm{MSE}^q_{T_n}(t)=\mathscr{E}[X(t)-X^q_{T_n}(t)]^2=\mathrm{MSE}^0_{T_n}(t)+G_{T_n}(t)$$

where

$$G_{T_n}(t)=[F'_{T_n}R^{-1}_{T_n}r_{T_n}(t)-F(t)]'(F'_{T_n}R^{-1}_{T_n}F_{T_n})^{-1}[F'_{T_n}R^{-1}_{T_n}r_{T_n}(t)-F(t)],$$

and its IMSE is

$$\mathrm{IMSE}^q_{T_n}(W)=\mathrm{IMSE}^0_{T_n}(W)+G_{T_n}(W)\tag{1.13}$$

where $\mathrm{IMSE}^0_{T_n}(W)$ is given in (1.7) and

$$G_{T_n}(W)=\int_0^1 G_{T_n}(t)W(t)\,\mathrm{d}t.\tag{1.14}$$

We will show that for regular sampling designs $T_n(h)$, the term $G_{T_n(h)}(W)$ converges to zero with rate n^{-4}. Thus, asymptotically, $\mathrm{IMSE}^0_{T_n}(W)$ is the dominant term of $\mathrm{IMSE}^q_{T_n}(W)$. More specifically, we have the following results.

Theorem 2. *When the mean* $m(t)$ *is as in* (1.9)-(1.10) *and Assumptions* (C1)-(C2) *hold, the following are true.*

(a) *The results* (i) (ii) *and* (iii) *in Theorem* 1 *remain valid for the estimator* $X^q_{T_n}(t)$.

(b) *If the function* α^2W^2/h^4 *is Riemann integrable, then*

$$\begin{aligned}&\lim_{n\to\infty} n^4\{\mathrm{IMSE}^q_{T_n(h)}(W)-\mathrm{IMSE}^0_{T_n(h)}(W)\}\\&\quad=\lim_{n\to\infty} n^4G_{T_n(h)}(W)\\&\quad=\frac{1}{120}\int_0^1\frac{\alpha^2(t)}{h^4(t)}\phi'(t)S^{-1}\phi(t)W^2(t)\,\mathrm{d}t,\end{aligned}\tag{1.15}$$

and when $h=h_0$, *the asymptotic constant becomes*

$$\frac{1}{120}\left\{\int_0^1(\alpha(t)W(t))^{1/2}\,\mathrm{d}t\right\}^4\int_0^1\phi'(t)S^{-1}\phi(t)\,\mathrm{d}t,\tag{1.16}$$

where $\phi'(t) = (\phi_1(t), \ldots, \phi_q(t))$ *and* $S = (s_{ij})_{q\times q}$ *is a* $q \times q$ *matrix with elements*

$$s_{ij} = \int_0^1 \int_0^1 \phi_i(t) R(t, s) \phi_j(s)\, \mathrm{d}t\, \mathrm{d}s, \quad i, j = 1, \ldots, q.$$

Remark 1. The results in Theorem 2 can be extended to more general (regression) functions than those specified by (1.10), namely of the form

$$f_i(t) = \int_0^1 R(t, s)\phi_i(s)\, \mathrm{d}s + \sum_{j=1}^{J_i} b_{i,j} R(t, a_{i,j}), \quad i = 1, \ldots, q, \tag{1.10$'$}$$

where the $b_{i,j}$'s are known coefficients and the $a_{i,j}$'s are known points in $[0, 1]$. For the model (1.10)$'$, the results in Theorem 2 still hold for the estimator $X^q_{T_n}(t)$ as in (1.11) with the sampling points T_n augmented by the set of points $\{a_{i,j};\, i = 1, \ldots, q;\, j = 1, \ldots, J_i\}$.

From (a) of Theorem 2, it follows that even though the mean structure (1.9) enters prominently in the expression (1.11) of the estimator $X^q_{T_n}(t)$, asymptotically it has no contribution to its performance. This suggests exploring what happens if we use the simpler estimator $X^0_{T_n}(t)$, which is the best linear unbiased estimator of $X(t)$ for the zero mean model, that is, if we proceed as if $m(t) = 0$.

Here, $X^0_{T_n}(t)$ is biased in the presence of the mean as in (1.9). In view of (1.5) and (1.11), we can write

$$X^q_{T_n}(t) = X^0_{T_n}(t) + M^q_{T_n}(t)$$

where the term due to the mean is

$$M^q_{T_n}(t) = -[F'_{T_n} R^{-1}_{T_n} r_{T_n}(t) - F(t)]'(F'_{T_n} R^{-1}_{T_n} F_{T_n})^{-1} F'_{T_n} R^{-1}_{T_n} X_{T_n}.$$

It is straightforward to verify that $M^q_{T_n}(t)$ is orthogonal to $X(t) - X^0_{T_n}(t)$: $\mathscr{E}[(X(t) - X^0_{T_n}(t)) M^q_{T_n}(t)] = 0$. It follows that the MSE of $X^q_{T_n}(t)$ can be written as follows:

$$\mathrm{MSE}^q_{T_n}(t) = \mathscr{E}[X(t) - X^q_{T_n}(t)]^2 = \mathscr{E}[X(t) - X^0_{T_n}(t)]^2 + \mathscr{E}[M^q_{T_n}(t)]^2.$$

Hence, *even though* $X^0_{T_n}(t)$ *is a biased estimator of* $X(t)$ *it nevertheless has smaller MSE than the BLUE* $X^q_{T_n}(t)$. Its MSE, by direct computation, is

$$\mathrm{MSE}^{q,0}_{T_n}(t) = E[X(t) - X^0_{T_n}(t)]^2 = \mathrm{MSE}^0_{T_n}(t) + H_{T_n}(t), \tag{1.17}$$

where the double superscript $(q, 0)$ indicates the MSE of the estimator $X^0_{T_n}(t)$ for the model (1.9)–(1.10) with q unknown (regression) parameters, $\mathrm{MSE}^0_{T_n}(t)$ is as in (1.6) and

$$H_{T_n}(t) = \left\{ \sum_{i=1}^{q} \beta_i [r'_{T_n}(t) R^{-1}_{T_n} f_{i,T_n} - f_i(t)] \right\}^2,$$

and its IMSE is

$$\mathrm{IMSE}^{q,0}_{T_n}(W) = \mathrm{IMSE}^0_{T_n}(W) + H_{T_n}(W) \tag{1.18}$$

where $\mathrm{IMSE}^0_{T_n}(W)$ is as in (1.7) and

$$H_{T_n}(W) = \int_0^1 H_{T_n}(t)\,W(t)\,\mathrm{d}t.$$

For this setup, we have the following results.

Theorem 3. *Under Assumptions* (C1) *and* (C2), *if the estimator* $X^0_{T_n}(t)$ *is used in the model* (1.9)-(1.10), *then the following are true.*

(a) *The results* (i), (ii) *and* (iii) *in Theorem* 1 *remain valid.*

(b) *If the function* $\alpha^2 W^2/h^4$ *is Riemann integrable, then*

$$\lim_{n\to\infty} n^4\{\mathrm{IMSE}^{q,0}_{T_n(h)}(W) - \mathrm{IMSE}^0_{T_n(h)}(W)\}$$

$$= \lim_{n\to\infty} n^4 H_{T_n(h)}(W)$$

$$= \frac{1}{120}\int_0^1 \frac{\alpha^2(t)}{h^4(t)}\left[\sum_{i=1}^q \beta_i\phi_i(t)\right]^2 W^2(t)\,\mathrm{d}t, \tag{1.19}$$

and when $h = h_0$, *the asymptotic constant is*

$$\frac{1}{120}\left\{\int_0^1 (\alpha(t)W(t))^{1/2}\,\mathrm{d}t\right\}^4 \int_0^1\left[\sum_{i=1}^q \beta_i\phi_i(t)\right]^2 \mathrm{d}t. \tag{1.20}$$

Thus, asymptotically, up to first order term, the simpler biased estimator $X^0_{T_n}(t)$ has the same performance as the BLUE estimator $X^q_{T_n}(t)$.

(3) Nonparametric mean

Here no specific knowledge about the mean is assumed except for its general form which is as in (1.9). Specifically, for some unknown continuous function $\psi(t)$ on $[0, 1]$,

$$m(t) = \int_0^1 R(t, s)\psi(s)\,\mathrm{d}s, \quad t\in[0, 1]. \tag{1.21}$$

If $\{\lambda_i\}_1^\infty$ and $\{e_i(t)\}_1^\infty$ are the eigenvalues and eigenfunctions of the covariance function R, i.e., the e_i's are orthonormal in $L_2[0, 1]$ and satisfy

$$\int_0^1 R(t, s)e_i(s)\,\mathrm{d}s = \lambda_i e_i(t), \quad t\in[0, 1],\ i = 1, 2, \ldots,$$

then (1.21) can be written as

$$m(t) = \sum_{i=1}^\infty \psi_i\lambda_i e_i(t)$$

where $\psi_i = \int_0^1 \psi(t)e_i(t)\,dt$, $\sum_{i=1}^{\infty} \psi_i^2 < \infty$, and the series converges in $L_2[0,1]$ as well as for all $t \in [0,1]$. The functions $f_i(t) = \lambda_i^{1/2} e_i(t)$, $i = 1, 2, \ldots$, form a complete orthonormal set in RKHS(R) and putting $\beta_i = \psi_i \lambda_i^{1/2}$ we also have

$$m(t) = \sum_{i=1}^{\infty} \beta_i f_i(t), \tag{1.22}$$

where the series converges in RKHS(R) and for all t in $[0,1]$, and $\sum_{i=1}^{\infty} \beta_i^2 \lambda_i^{-1} < \infty$. Note that $\psi(t) = \sum_{i=1}^{\infty} \beta_i \lambda_i^{-1/2} e_i(t)$ where the series converges in $L_2[0,1]$. As in (1.22) the functions $\{f_i(t)\}_1^{\infty}$ are known based on the covariance function (else some other complete set in RKHS(R) could be used) and the coefficients $\{\beta_i\}_1^{\infty}$ are unknown, this nonparametric case can be viewed as an extension of the semiparametric case to $q = \infty$.

As examples, we list the eigenvalues and eigenfunctions for Wiener process, the process with triangular covariance function and Ornstein-Uhlenbeck process. For the Wiener process with $R(t,s) = \min(t,s)$,

$$e_i(t) = \sqrt{2}\sin(i-\tfrac{1}{2})\pi t, \quad \lambda_i = [(i-\tfrac{1}{2})\pi]^{-2}, \qquad i = 1, 2, \ldots.$$

For the process with triangular covariance function $R(t,s) = 1 - |t-s|$ if $|t-s| < 1$ and $R(t,s) = 0$, otherwise,

$$e_i(t) = \frac{1}{\nu_i}\{2\sin \nu_i t + \nu_i \cos \nu_i t\}, \quad \lambda_i = 2/\nu_i^2,$$

where ν_i solves the following equation:

$$\tan(\tfrac{1}{2}\nu_i) = 2/\nu_i, \quad \nu_i \in [(i-\tfrac{1}{2})\pi, (i+\tfrac{1}{2})\pi], \qquad i = 0, 1, \ldots$$

(see Kailath, 1966). The eigenfunctions and eigenvalues of the Ornstein-Uhlenbeck process with $R(t,s) = e^{-|t-s|}$ are

$$e_i(t) = \left(\frac{2}{u_i^2+3}\right)^{1/2}\{\sin u_i t + u_i \cos u_i t\}, \quad \lambda_i = 2/(1+u_i^2),$$

where u_i solves the following equation:

$$\tan u_i = 2u_i/(u_i^2-1), \quad u_i \in [(i-\tfrac{1}{2})\pi, (i+\tfrac{1}{2})\pi], \qquad i = 0, 1, \ldots$$

(see Hawkins, 1989).

For the model (1.22) with an infinite number of parameters $\{\beta_i\}_1^{\infty}$ it makes no sense to attempt to estimate all of them on the basis of a finite number of observations; instead, for sample size n, we estimate a finite number q of parameters and derive the asymptotic performance of the corresponding estimator.

For fixed sampling designs T_n and coefficients $C_{T_n}(t)$, write the bias term $B_{T_n}(t)$ as

$$B_{T_n}(t) = \sum_{i=1}^{q} \beta_i\{C'_{T_n}(t) f_{i,T_n} - f_i(t)\} + \sum_{i=q+1}^{\infty} \beta_i\{C'_{T_n}(t) f_{i,T_n} - f_i(t)\}.$$

Then, minimizing the variance $V_{T_n}(t)$ subject to the constraints $C'_{T_n}(t)f_{i,T_n}=f_i(t)$, $i=1,\ldots,q$, yields, as in the derivation of the BLUE for the model (1.9), the same estimator $X^q_{T_n}(t)$ as in (1.11), except that here $X^q_{T_n}(t)$ is not an unbiased estimator of $X(t)$. In order to emphasize that this estimator is applied here to the case where (1.9) is replaced by (1.22), i.e., the number of parameters is $q=\infty$ in (1.22), we denote its IMSE as $\mathrm{IMSE}^{\infty,q}_{T_n}$ which can be written as

$$\mathrm{IMSE}^{\infty,q}_{T_n}(W)=\mathrm{IMSE}^0_{T_n}(W)+G_{T_n}(W)+Q_{T_n}(W) \tag{1.23}$$

where $\mathrm{IMSE}^0_{T_n}$ is given by (1.7), $G_{T_n}(W)$ by (1.14) and

$$Q_{T_n}(W)=\int_0^1\left\{\sum_{i=q+1}^{\infty}\beta_i[\hat{C}'_{T_n}(t)f_{i,T_n}-f_i(t)]\right\}^2 W(t)\,\mathrm{d}t,$$

where $\hat{C}_{T_n}(t)$ are the coefficients of $X^q_{T_n}(t)$ (cf. (1.11)). For the above setup, we have the following results.

Theorem 4. *Under Assumptions* (C1) *and* (C2), *if the estimator $X^q_{T_n}(t)$ is used in the model* (1.21)–(1.22), *then the results* (i), (ii) *and* (iii) *in Theorem* 1 *hold. Furthermore, if the function α^2W^2/h^4 is Riemann integrable, then for $q=1,2,\ldots$,*

$$\lim_{n\to\infty} n^4\{\mathrm{IMSE}^{\infty,q}_{T_n(h)}(W)-\mathrm{IMSE}^0_{T_n(h)}(W)\}$$

$$=\frac{1}{120}\int_0^1\frac{\alpha^2(t)}{h^4(t)}\left\{\sum_{k=1}^{q}\lambda_k^{-1}e_k^2(t)+\left[\psi(t)-\sum_{k=1}^{q}\beta_k\lambda_k^{-1/2}e_k(t)\right]^2\right\}W^2(t)\,\mathrm{d}t, \tag{1.24}$$

and when $h=h_0$, the asymptotic constant is

$$\frac{1}{120}\left\{\int_0^1(\alpha(t)W(t))^{1/2}\,\mathrm{d}t\right\}^4\left\{\sum_{k=1}^{q}\lambda_k^{-1}+\sum_{k=q+1}^{\infty}\beta_k^2\lambda_k^{-1}\right\}. \tag{1.25}$$

Letting $q=0$ in (1.25), i.e., using the simpler estimator $X^0_{T_n(h_0)}(t)$ in the model (1.22), yields the following asymptotic constant:

$$\frac{1}{120}\left\{\int_0^1(\alpha(t)W(t))^{1/2}\,\mathrm{d}t\right\}^4\sum_{i=1}^{\infty}\beta_i^2\lambda_i^{-1}.$$

By comparing this with (1.25), one can see that if $\sum_{i=1}^q\beta_i^2\lambda_k^{-1}<\sum_{i=1}^q\lambda_k^{-1}$, then $X^0_{T_n(h_0)}(t)$ has a better asymptotic performance than $X^q_{T_n(h_0)}(t)$, while if the reverse inequality holds, $X^q_{T_n(h_0)}(t)$ is better than $X^0_{T_n(h_0)}(t)$. Therefore, when the mean function m is of the form (1.21) for some (unknown) continuous function ψ, i.e. $m\in\mathrm{Range}(R)$, the simpler estimator $X^0_{T_n}(t)$ is recommended since there may not be any benefit from using the more complicated estimator $X^q_{T_n(h_0)}(t)$, which requires the evaluation of the eigenvalues and eigenfunctions of $R(t,s)$.

For a mean function $m\notin\mathrm{Range}(R)$, it is not clear whether the simpler estimator $X^0_{T_n}(t)$ is still asymptotically optimal. Here, by using the estimator $X^{q_n}_{T_n}(t)$ and letting q_n increase appropriately as the sample size n tends to infinity, we show that for a certain class of mean functions belonging to an infinite dimensional subset of RKHS(R), the estimator $X_{T_n}^{q_n}(t)$ is asymptotically optimal. Note that $m \in$ RKHS(R) iff it is of the form (1.22) with $\sum_{i=1}^{\infty} \beta_i^2 < \infty$.

Theorem 5. *Under Assumptions* (C1) *and* (C2) *and for the class of mean functions given by* (1.22) *with* $\sum_{i=1}^{\infty} |\beta_i| < \infty$, *if the estimator* $X_{T_n}^{q_n}(t)$ *is used and* q_n *tends to infinity in such a way that* q_n/n *remains bounded and*

$$\lim_{n\to\infty} n^{-3} \sum_{i=1}^{q_n} \lambda_i^{-1} = 0, \tag{1.26}$$

then the results (i) *and* (iii) *in Theorem* 1 *hold. Furthermore, if the function* $(\alpha W/h^2)(t)$ *is bounded on* $[0, 1]$, *then* (ii) *in Theorem* 1 *also holds.*

Condition (1.26) provides a constraint on the number of parameters q_n to be estimated on the basis of n observations. For Wiener process,

$$\sum_{i=1}^{q} \lambda_i^{-1} = \sum_{i=1}^{q} \{(i-\tfrac{1}{2})\pi\}^2 \sim \tfrac{1}{3}\pi^2 q^3,$$

and thus, the constraint (1.26) is equivalent to $n^{-3}q_n^3 \to 0$, i.e., $q_n = o(n)$. Similarly, one can justify that for a process with triangular covariance function and an Ornstein–Uhlenbeck process, (1.26) is equivalent to $q_n = o(n)$.

Linear interpolation

The BLUE's, considered so far, involve the evaluation of the inverse of a covariance matrix and more significantly, require the precise knowledge of the covariance function. This leads us to try a simpler nonparametric estimator. Here we consider the sample function of the stochastic process as a real valued function and we estimate its values in between consecutive samples by linear interpolation, i.e. estimate $X(t)$ over each interval $[t_{n,k}, t_{n,k+1}]$, $k = 1, \ldots, n-1$, by

$$X_{T_n}^{\ell}(t) = \{(t_{n,k+1} - t)X(t_{n,k}) + (t - t_{n,k})X(t_{n,k+1})\}/(t_{n,k+1} - t_{n,k}). \tag{1.27}$$

It turns out that this procedure has an asymptotically optimal performance when the regular sampling designs $T_n(h_o)$ are used! This happens, even though for each fixed n the linear interpolator $X_{T_n}^{\ell}(t)$ is generally different from the BLUE.

Here, we do not assume a specific form for the mean structure as in (1.21). Instead, we assume the mean function $m(t)$ satisfies a Hölder condition

$$|m(t) - m(s)| \leq C|t-s|^p, \quad t, s \in [0, 1], \tag{1.28}$$

where $p \in (\tfrac{1}{2}, 1)$ and $0 < C < \infty$. For this setup, we have the following results.

Theorem 6. *If the mean* $m(t)$ *satisfies* (1.28), *the covariance function* $R(t, s)$ *satisfies Assumption* (C1) *and* $\alpha W/h$ *is Riemann integrable, then for the linear interpolator*

$X^{\ell}_{T_n(h)}(t)$ *as in* (1.27) *we have*

$$\lim_{n\to\infty} n\,\mathrm{IMSE}_{T_n(h)}(W) = \frac{1}{6}\int_0^1 \frac{\alpha(t)}{h(t)}\,W(t)\,\mathrm{d}t, \tag{1.29}$$

and clearly $X^{\ell}_{T_n(h_0)}(t)$ *is asymptotically optimal.*

Theorem 6 implies that, when the centered process N is not quadratic-mean differentiable, a linear interpolator as in (1.27) has the same asymptotic performance as the BLUE with the same sampling density. In particular, when the centered process is stationary, the asymptotically optimal estimator $X^{\ell}_{T_n(h_0)}(t)$ is entirely nonparametric, namely, completely independent of the covariance function R.

Remark 2. For the linear interpolator it is possible to identify the higher order terms of the IMSE under additional smoothness assumptions. (This is rather complex for the BLUE's.) For instance the second order term can be identified, and is in fact shown to vanish, when in addition to the assumptions made in Theorem 6, the functions $m(t)$, $\alpha(t)$, $h(t)$ are continuously differentiable and the functions $\alpha h^{(1)}h^{-3}$, $(\alpha/h)^{(1)}h^{-1}$ are Riemann integrable; namely as $n\to\infty$ we have

$$\mathrm{IMSE}_{T_n(h)}(W) = \frac{1}{6(n-1)}\int_0^1 \frac{\alpha(t)}{h(t)}\,W(t)\,\mathrm{d}t + \mathrm{o}(n^{-2}), \tag{1.29$'$}$$

i.e. the coefficient of the term $(n-1)^{-2}$ always vanishes. The higher order terms $(n-1)^{-3}$, etc. (under appropriate additional smoothness assumptions) generally do not vanish, as is seen in Example 3 (equation (1.46)).

Examples

We first consider an example with mean $m=0$, triangular covariance and weight function $W(t)\equiv 1$. Here, for certain values of the parameter μ in the covariance function, we are able to compute numerically the optimal design for every finite sample size n. For other values of the parameter μ, the MSE is worse near the edges of the interval, and we will show how to select a weight function to reduce this discrepancy in the MSE of approximation between the edges and the middle of the interval.

Example 1. *Zero mean process with triangular covariance.* We consider the model (1.1) with triangular covariance function $R(t,s)=1-\mu|t-s|$ if $|t-s|\le 1/\mu$ and $R(t,s)=0$, otherwise, where μ is a positive parameter.

When $\mu\le 1$, for any sampling designs $T_n=\{t_{n,i}\}_1^n$, the BLUE estimator $X^0_{T_n}(t)$ is

$$X^0_{T_n}(t) = \begin{cases} \dfrac{[2-\mu(t_{n,n}-t)]X(t_{n,1})-\mu(t_{n,1}-t)X(t_{n,n})}{2-\mu(t_{n,n}-t_{n,1})}, & \text{if } 0\le t\le t_{n,1},\\[2ex] \dfrac{(t_{n,k+1}-t)X(t_{n,k})+(t-t_{n,k})X(t_{n,k+1})}{t_{n,k+1}-t_{n,k}}, & \text{if } t_{n,k}\le t\le t_{n,k+1},\ k=1,\dots,n-1,\\[2ex] \dfrac{[2-\mu(t-t_{n,1})]X(t_{n,n})+\mu(t_{n,n}-t)X(t_{n,1})}{2-\mu(t_{n,n}-t_{n,1})}, & \text{if } t_{n,n}\le t\le 1, \end{cases} \tag{1.30}$$

and clearly, it is a linear interpolation of observations between successive sampling points except for t in the two end intervals. When the weight function $W(t)\equiv 1$, for every fixed sample size n (≥ 2) the optimal sampling designs are $T_n^{\mathrm{o}}=\{t_{n,i}^{\mathrm{o}}\}_1^n$ with

$$\begin{aligned} t_{n,1}^{\mathrm{o}} &= 1-t_{n,n}^{\mathrm{o}}=\tfrac{1}{2}[1-(n-1)\rho_n], \\ t_{n,i}^{\mathrm{o}} &= t_{n,1}^{\mathrm{o}}+(i-1)\rho_n, \quad i=2,\dots,n-1, \end{aligned} \tag{1.31}$$

where $\rho_n\in(1/n, 1/(n-1))$ satisfies the following equation:

$$\begin{aligned} &4n(n-1)^2\mu^2\rho_n^3-(n-1)\{3(\mu+6)n-(2+3\mu)\}\mu\rho_n^2 \\ &\quad +4\{3(\mu+2)n-(2+3\mu)\}\rho_n-\mu^2+6\mu-24=0, \end{aligned} \tag{1.32}$$

i.e. the points $t_{n,1}^{\mathrm{o}},\dots,t_{n,n}^{\mathrm{o}}$ are periodically spaced with period ρ_n, while the equal edges $t_{n,1}^{\mathrm{o}}$, $1-t_{n,n}^{\mathrm{o}}$ have length smaller than $\frac{1}{2}\rho_n$.

Since both the weight function $W(t)$ and the jump function $\alpha(t)$ here are constant, the regular sampling designs determined by the asymptotically optimal sampling density ($h_{\mathrm{o}}\equiv 1$) are uniform including the end points, i.e., $T_n(h_{\mathrm{o}})=\{(i-1)/(n-1)\}_1^n$. To see the difference between T_n^{o} and $T_n(h_{\mathrm{o}})$, we plot for $n=5$, the corresponding points in Figure 1, with ×'s denoting the regular sampling points and dots the optimal sampling points, which are tabulated in Table 1. The MSE of $X_{T_n^{\mathrm{o}}}^{\mathrm{o}}(t)$ is

$$\mathrm{MSE}_{T_n^{\mathrm{o}}}^{0}(t)=\begin{cases} 2\mu(t_{n,1}^{\mathrm{o}}-t)[2-\mu(t_{n,n}^{\mathrm{o}}-t)]/[2-\mu(n-1)\rho_n], & \text{if } 0\leq t\leq t_{n,1}^{\mathrm{o}}, \\ 2\mu(t-t_{n,k}^{\mathrm{o}})(t_{n,k+1}^{\mathrm{o}}-t)/\rho_n, & \text{if } t_{n,k}^{\mathrm{o}}\leq t\leq t_{n,k+1}^{\mathrm{o}},\ k=1,\dots,n-1, \\ 2\mu(t-t_{n,n}^{\mathrm{o}})[2-\mu(t-t_{n,1}^{\mathrm{o}})]/[2-\mu(n-1)\rho_n], & \text{if } t_{n,n}^{\mathrm{o}}\leq t\leq 1, \end{cases} \tag{1.33}$$

and its IMSE is

$$\begin{aligned} \mathrm{IMSE}_{T_n^{\mathrm{o}}}^{0}(1)&=\frac{\mu}{6[2-\mu(n-1)\rho_n]}\{-2n(n-1)^2\mu\rho_n^3 \\ &\quad +[3\mu(n-1)+6n-2](n-1)\rho_n^2-12(n-1)\rho_n+6-\mu\}. \end{aligned} \tag{1.34}$$

Fig. 1. Optimal (·) and regular (×) sampling design of size 5.

Table 1

Uniform sampling points	Optimal sampling points
0.00	0.075
0.25	0.288
0.50	0.500
0.75	0.712
1.00	0.925

The BLUE of $X(t)$ with the regular sampling design $T_n(h_o)$ and its IMSE take simpler forms:

$$X^0_{T_n(h_o)}(t)=[k-(n-1)t]X\left(\frac{k-1}{n-1}\right)+[1-k+(n-1)t]X\left(\frac{k}{n-1}\right),$$

$$k-1\leq(n-1)t\leq k,\ k=1,\dots,n-1, \tag{1.35}$$

$$\mathrm{IMSE}^0_{T_n(h_o)}(1)=\mu/[3(n-1)]. \tag{1.36}$$

We plot the $\mathrm{IMSE}^0_{T_n^o}(1)$, denoted by IMSE(opt), and the $\mathrm{IMSE}^0_{T_n(h_o)}(1)$, denoted by IMSE(reg), in Figure 2(a) for sample size up to 20 and for $\mu=1$. It should be pointed out that all values of μ in $(0,1]$ produce a similar picture to the one in Figure 2(a). From this figure, one can see that IMSE(reg) approaches IMSE(opt) quickly. To achieve the same IMSE error (10^{-1}, 10^{-2}, 10^{-3}), the sample sizes needed for the optimal sampling design and for the regular design are listed in Table 2.

When $\mu>1$, it is more complicated to obtain the optimal sampling points T_n^o for fixed sample size n. Thus, in this case we use the regular sampling designs $T_n(h_o)$, i.e. uniform sampling with period $1/(n-1)$. For the sampling design $T_n(h_o)$ and fixed parameter μ, if there is an integer k satisfying

$$k/(n-1)=1/\mu, \tag{1.37}$$

then the BLUE $X^0_{T_n(h_o)}(t)=r'_{T_n(h_o)}(t)R^{-1}_{T_n(h_o)}X_{T_n(h_o)}$ takes a much simpler form. Indeed, if (1.37) holds, then the elements of the vector $r_{T_n(h_o)}(t)$ are either 0 or linear

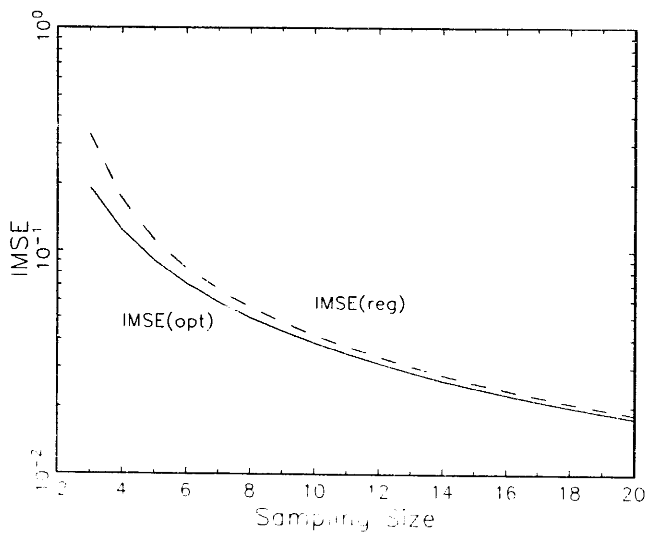

Fig. 2(a). IMSE vs. sample size.

Table 2

IMSE	n_{reg}	n_{opt}
10^{-1}	5	4
10^{-2}	35	34
10^{-3}	335	334

functions of t, and, as a result, the coefficients $r'_{T_n(h_0)}(t)R^{-1}_{T_n(h_0)}$ of $X^0_{T_n(h_0)}(t)$ are either 0 or linear functions of t. By the uniqueness of the linear interpolation satisfying the conditions $X^0_{T_n(h_0)}(t_{n,i}) = X(t_{n,i})$, $i = 1, \ldots, n$, it follows that

$$X^0_{T_n(h_0)}(t) = (n-1)\{[i/(n-1) - t]X[(i-1)/(n-1)] \\ + [t-(i-1)/(n-1)]X[i/(n-1)]\},$$

when $i-1 \leq (n-1)t \leq i$, $i = 1, \ldots, n-1$, namely, $X^0_{T_n(h_0)}(t)$ is a linear interpolation of observations of $X(t)$ between adjacent sampling points. As expected, the pattern of $\mathrm{MSE}^0_{T_n(h_0)}(t)$ is of the same form in each subinterval $[(i-1)/(n-1), i/(n-1)]$, $i = 1, \ldots, n-1$.

To see the pattern of $\mathrm{MSE}^0_{T_n(h_0)}(t)$ when (1.37) is not satisfied, we plotted it for $\mu = 4.56$, which makes the MSE relatively larger, and for a variety of sample sizes n. We found that the plots of $\mathrm{MSE}^0_{T_n(h_0)}(t)$ display some variability. When the condition (1.37) is approximately satisfied, i.e. $(n-1)/4.56$ is very close to an integer, the variability in the pattern of $\mathrm{MSE}^0_{T_n(h_0)}(t)$ is not significant; for instance when $n = 15$, $(n-1)/4.56 = 3.07 \approx 3$ and the pattern of $\mathrm{MSE}^0_{T_{15}(h_0)}(t)$ is close to periodic as shown in Figure 2(b); and likewise when $n = 6$ for which $(n-1)/4.56 = 1.096 \approx 1$ and when $n = 24$ for which $(n-1)/4.56 = 5.043 \approx 5$, etc. We also plotted in Figure 2(c) the MSE for $n = 12$ for which $(n-1)/4.56 = 2.412$. One can see that the peaks of the MSE over the two intervals near the end points are about 12% higher than the peaks over the middle intervals. The corresponding IMSE is 0.1048.

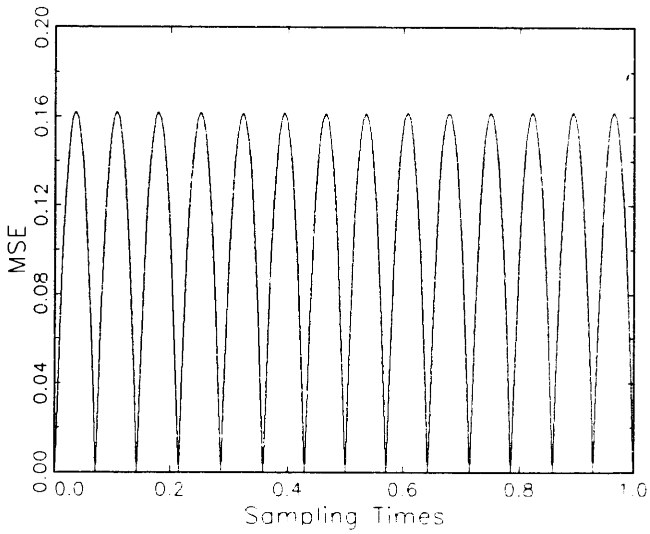

Fig. 2(b). MSE vs. time with regular sampling design ($n = 15$).

If it is desirable that the peaks of $\mathrm{MSE}^0_{T_{12}(h_0)}(t)$ should be more uniform over the entire interval, then one could select a sample size n nearly satisfying (1.37) or else one could consider a non-constant weight function. For instance, by looking at Figure 2(c), it appears that a weight function of the following form is suitable, i.e. a continuous function $W(t)$ taking a constant value (x) over say $[0, \frac{2}{11}]$ and $[\frac{9}{11}, 1]$, a constant (y) over $[\frac{3}{11}, \frac{8}{11}]$, and linear over $[\frac{2}{11}, \frac{3}{11}]$ and $[\frac{8}{11}, \frac{9}{11}]$, where $\int_0^1 W(t)\,\mathrm{d}t = 1$

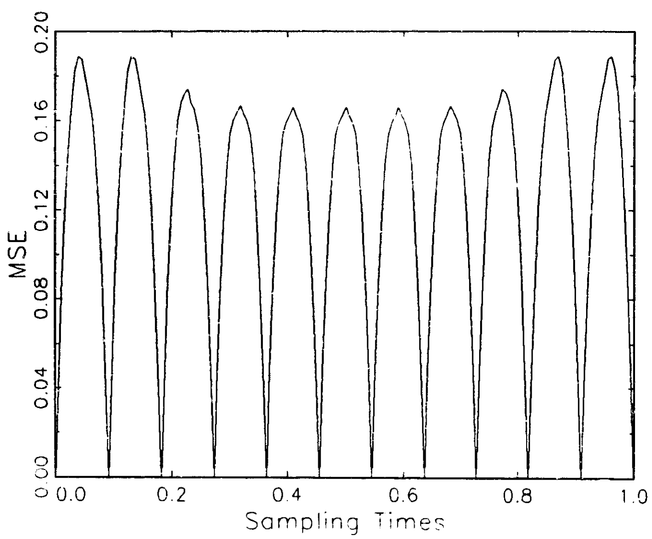

Fig. 2(c). MSE vs. time with regular sampling design ($n = 12$).

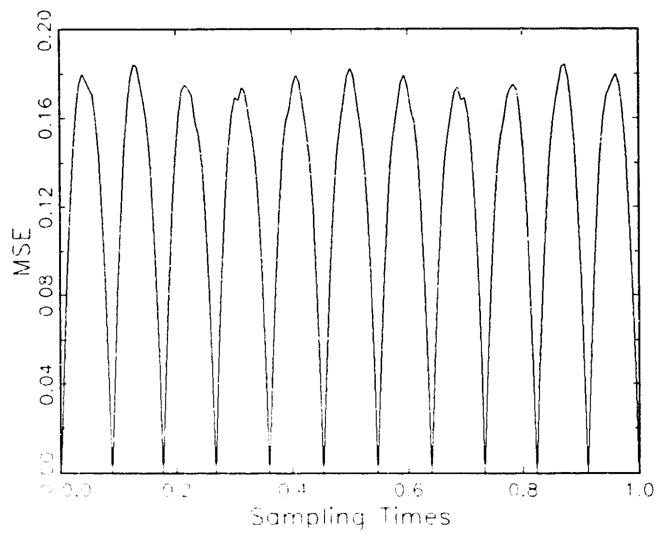

Fig. 2(d). MSE vs. time with regular sampling design ($n = 12$).

gives y in terms of x: $y = \frac{1}{6}(11 - 5x)$. We plotted the corresponding MSE(t) for several values of x and found that when x is close to 1.062, the high and low peaks are closer, indeed, the high peaks of the MSE are only about 5% higher than the low peaks. The plot corresponding to $x = 1.064$ is shown in Figure 2(d). The pattern is now more oscillatory and the low peaks are higher but the corresponding IMSE is 0.1053, a slight increase of about 0.5%.

Next, we consider an example of linear regression in Wiener noise. Here, we are able to compute the optimal sampling design for every sample size n when the weight fuction $W(t) \equiv 1$ is used.

Example 2. *Linear regression in Wiener noise.* Consider the model (1.1), (1.9) with $R(t, s) = \min(t, s)$, $q = 1$, and the regression function $f(t) = t = \min(t, 1)$, which is in the form (1.10)′ with $J = 1$, $\phi(t) \equiv 0$, $b_1 = 1$ and $a_1 = 1$.

For any sampling design T_n, the BLUE of $X(t)$ is

$$X^1_{T_n}(t)=\begin{cases} \dfrac{t}{t_{n,1}}X(t_{n,1}), & \text{if } 0\leq t\leq t_{n,1},\\[2ex] X(t_{n,k})+\dfrac{t-t_{n,k}}{t_{n,k+1}-t_{n,k}}\{X(t_{n,k+1})-X(t_{n,k})\}, & \begin{array}{l}\text{if } t_{n,k}\leq t\leq t_{n,k+1},\\ \quad k=1,\ldots,n-1,\end{array}\\[2ex] \dfrac{t}{t_{n,n}}X(t_{n,n}), & \text{if } t_{n,n}\leq t\leq 1,\end{cases} \tag{1.38}$$

which is a linear interpolation over $(0, t_{n,n})$, and the corresponding MSE and IMSE are, with $t_{n,0}=0$,

$$\mathrm{MSE}^1_{T_n}(t)=\begin{cases}(t-t_{n,k})(t_{n,k+1}-t)/(t_{n,k+1}-t_{n,k}), & \begin{array}{l}\text{if } t_{n,k}\leq t\leq t_{n,k+1},\\ \quad k=0,1,\ldots,n-1,\end{array}\\[2ex] t(t-t_{n,n})/t_{n,n}, & \text{if } t_{n,n}\leq t\leq 1,\end{cases} \tag{1.39}$$

$$\mathrm{IMSE}^1_{T_n}(1)=\tfrac16\sum_{k=0}^{n-1}(t_{n,k+1}-t_{n,k})^2+\tfrac16 t_{n,n}^{-1}(2+t_{n,n})(1-t_{n,n})^2. \tag{1.40}$$

For every sample size n, the optimal sampling design is $T^{\mathrm{o}}_n=\{t^{\mathrm{o}}_{n,k}\}_1^n$, with

$$t^{\mathrm{o}}_{n,k}=k\{n^2(n+1)\}^{-1/3},\quad k=1,\ldots,n. \tag{1.41}$$

Replacing T_n by T^{o}_n in (1.40), yields

$$\mathrm{IMSE}^1_{T^{\mathrm{o}}_n}(1)=\tfrac12\{(1+n^{-1})^{1/3}-1\}=\tfrac16 n^{-1}-\tfrac1{18}n^{-2}+o(n^{-2}).$$

When $W(t)\equiv 1$, then $h_{\mathrm{o}}(t)\equiv 1$, and hence the regular sampling design generated by h_{o} is peridodic $T_n(h_{\mathrm{o}})=\{(k-1)/(n-1)\}_1^n$. By (1.40) we obtain

$$\mathrm{IMSE}^1_{T_n(h_{\mathrm{o}})}(1)=[6(n-1)]^{-1}=\tfrac16 n^{-1}+\tfrac16 n^{-2}+o(n^{-2}),$$

which implies that

$$\mathrm{IMSE}^1_{T_n(h_{\mathrm{o}})}(1)-\mathrm{IMSE}^1_{T^{\mathrm{o}}_n}(1)=\tfrac29 n^{-2}+o(n^{-2}).$$

For $n=5$, the optimal sampling points and the regular sampling points are plotted in Figure 3 as dots and ×'s, respectively. The MSE corresponding to T^{o}_n is a periodic function of t except for the last interval and the MSE of $T_n(h_{\mathrm{o}})$ is periodic over $[0,1]$ as expected. They are plotted in Figure 4(a) and (b), respectively, and their IMSE's are 0.0147 and 0.0167. To see the difference in performance between the optimal and the asymptotically optimal designs, we plot the IMSE with the optimal design T^{o}_n, IMSE(opt), and the IMSE with the regular sampling design $T_n(h_{\mathrm{o}})$, IMSE(reg), in Figure 4(c) from which one can see that IMSE(reg) quickly approaches the IMSE(opt). To achieve the same IMSE error ε^2, the sample size

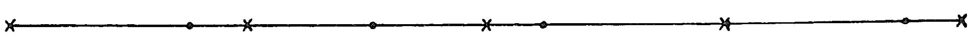

Fig. 3. Optimal (·) and regular (×) sampling design of size 5.

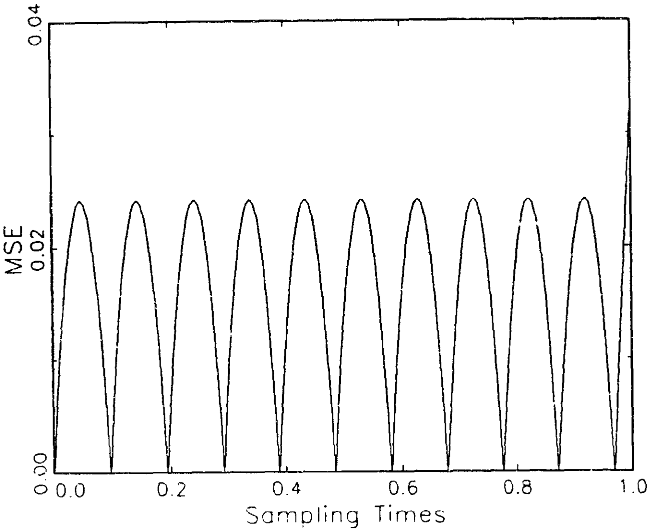

Fig. 4(a). MSE vs. time with optimal sampling design ($n = 10$).

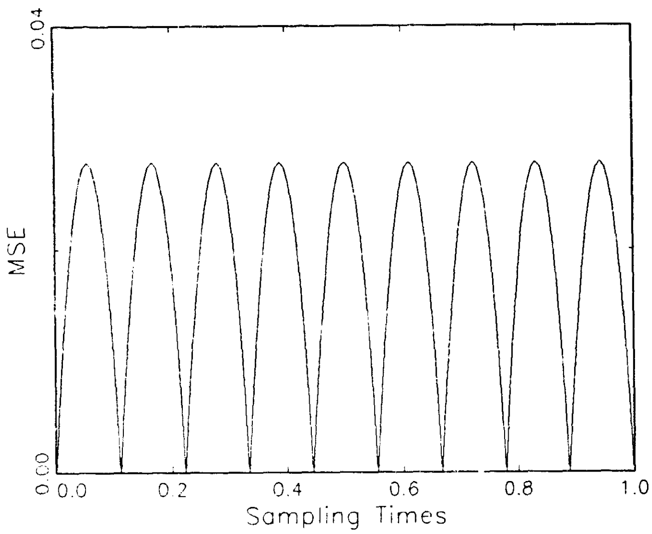

Fig. 4(b). MSE vs. time with regular sampling design ($n = 10$).

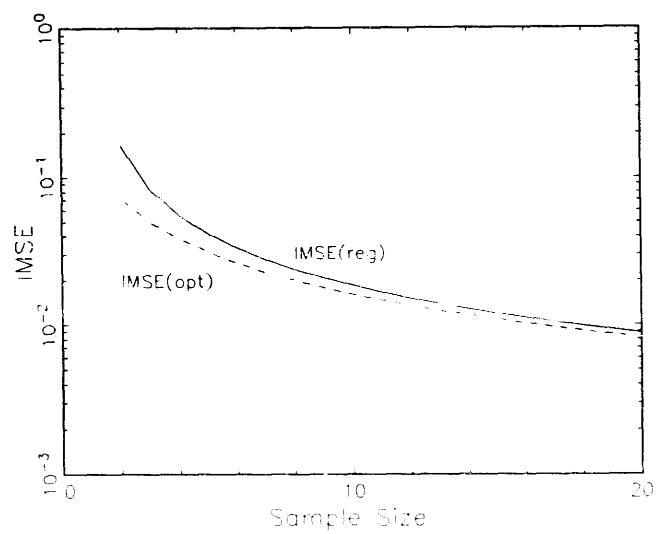

Fig. 4(c). IMSE vs. sample size.

Table 3

IMSE	n_{opt}	n_{reg}
10^{-1}	2	3
10^{-2}	17	18
10^{-3}	167	168

needed for the optimal design is $n_{opt}(\varepsilon^2)=1/[(1+2\varepsilon^2)^3-1]$ and the sample size needed for the regular sampling design is $n_{reg}(\varepsilon^2)=1+1/(6\varepsilon^2)$. For some values of ε^2, these sample sizes are listed in Table 3, which show how efficient the equidistant regular sampling design is in this case.

To see the impact on the sampling designs of the mean function, we take $f(t)=0$ here. Then, the IMSE is

$$\text{IMSE}^0_{T_n}(1)=\tfrac{1}{6}\sum_{k=0}^{n-1}(t_{n,k+1}-t_{n,k})^2+\tfrac{1}{2}(1-t_{n,n})^2,$$

and the corresponding optimal sampling design is $T^\circ_n=\{t^*_{n,k}\}_1^n$ with

$$t^*_{n,k}=3k/(3n+1),\quad k=1,\ldots,n.$$

Note that for $k=1,\ldots,n$,

$$\begin{aligned} t^\circ_{n,k}-t^*_{n,k} &= k\{[(n+1)n^2]^{-1/3}-3/(3n+1)\} \\ &= k\left\{\frac{1}{n}-\frac{1}{3n^2}+\frac{2}{9n^3}-\frac{1}{n}+\frac{1}{3n^2}-\frac{1}{9n^3}+o(n^{-3})\right\} \\ &= k\left\{\frac{1}{9n^3}+o(n^{-3})\right\}, \end{aligned}$$

so that a linear mean function perturbs the kth optimal sampling points by $k9^{-1}n^{-3}+o(n^{-3})$.

Example 3. *Random processes with stable-type covariance.* We consider a random process with covariance function $R(t,s)=e^{-\mu|t-s|^\nu}$, with parameters $\mu>0$ and $0<\nu\leq 2$. When $\nu=1$, it is an Ornstein-Uhlenbeck process. The asymptotically optimal regular sampling design has $h_o\equiv 1$, namely $T_n(h_o)=\{(i-1)/(n-1)\}_1^n$, when the constant weight function $W(t)\equiv 1$ is used.

First, consider the case where the mean is zero. To see the pattern of the $\text{MSE}^0_{T_n(h_o)}(t)$, we plotted it for a variety of values of ν and μ and found that μ has no impact on its shape but ν does. Specifically, when ν is between 0 and about 0.5 the pattern of the $\text{MSE}^0_{T_n(h_o)}(t)$ is as shown in Figure 5(a) for $\nu=0.15$ and $\mu=1.5$, with some, but not significant, variability. When ν is approximately between 0.5 and 1.5 the pattern of the $\text{MSE}^0_{T_n(h_o)}(t)$ is very close to periodic and exactly periodic when $\nu=1$. And when ν is approximately between 1.5 and 2 the pattern of the $\text{MSE}^0_{T_n(h_o)}(t)$ is similar to the plot in Figure 5(b) which corresponds to $\nu=1.82$ and

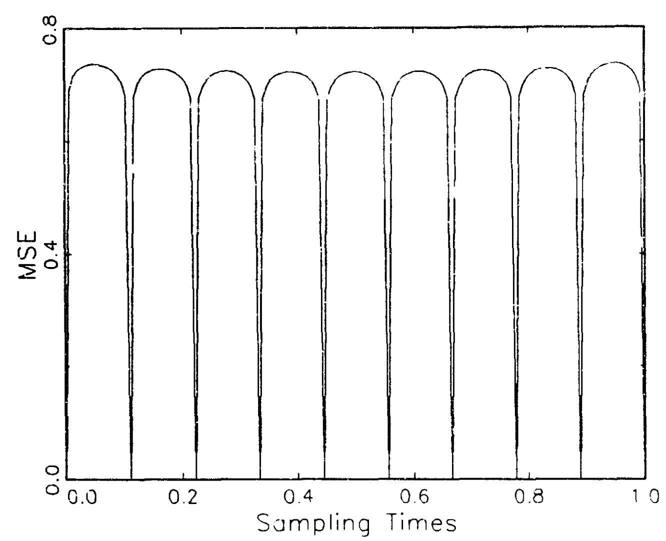

Fig. 5(a). MSE vs. time with regular sampling design ($n = 10$), $\nu = 0.15$, $\mu = 1.5$.

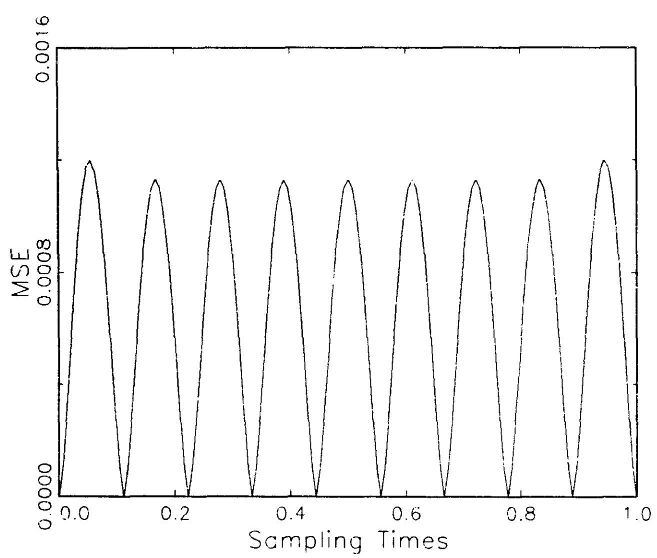

Fig. 5(b). MSE vs. time with regular sampling design ($n = 10$), $\nu = 1.82$, $\mu = 1.5$.

$\mu = 1.5$ and also shows the near periodicity except for the two end intervals where the MSE is more peaked; here the magnitude of $\mathrm{MSE}^{0}_{T_n(h_n)}(t)$ is negligible, i.e. less than 1.2×10^{-3}.

Next, we consider only the important case of Ornstein–Uhlenbeck covariance function, i.e. $\mu = \nu = 1$, with an unknown constant mean $m(t) = \beta$, which is of the form (1.10) with $q = 1$ and

$$f_1(t) \equiv 1 = \frac{1}{2}\left\{ \int_0^1 R(t, s)\, \mathrm{d}s + R(t, 0) + R(t, 1) \right\}.$$

In this case, we have three candidate estimators, the BLUE estimator $X^1_{T_n}(t)$ given by (1.11), the estimator given by (1.5) which is the BLUE when the mean is zero but is biased here, and the simple linear interpolator $X^{\ell}_{T_n}(t)$ as in (1.27). We will compare their performance when the asymptotically optimal (uniform) sampling design $T_n(h_o)=\{(i-1)/(n-1)\}_1^n$ is used.

The estimator $X^0_{T_n(h_o)}(t)$ is given for $t\in[t_{n,k}, t_{n,k+1}]$, $k=1,\dots,n-1$, by

$$X^0_{T_n(h_o)}(t)=\frac{(e^{t_{n,k+1}-t}-e^{-(t_{n,k+1}-t)})X(t_{n,k})+(e^{t-t_{n,k}}-e^{-(t-t_{n,k})})X(t_{n,k+1})}{e^{1/(n-1)}-e^{-1/(n-1)}},\tag{1.42}$$

and its IMSE is

$$\text{IMSE}^0_{T_n(h_o)}(1)=V_{T_n(h_o)}+\beta^2B^2_{T_n(h_o)},\tag{1.43}$$

where

$$V_{T_n(h_o)}=(e^{2/(n-1)}+1)/(e^{2/(n-1)}-1)-(n-1),$$
$$B^2_{T_n(h_o)}=\{(4-3n)\,e^{2/(n-1)}+4\,e^{1/(n-1)}+3n-2\}/(1+e^{1/(n-1)})^2.$$

The BLUE $X^1_{T_n(h_o)}(t)$ is given by (1.11), and can be simplified for $t\in[t_{n,k}, t_{n,k+1}]$, $k=1,\dots,n-1$, as

$$X^1_{T_n(h_o)}(t)=X^0_{T_n(h_o)}(t)-\left\{\frac{e^{t_{n,k+1}-t}-e^{-(t_{n,k+1}-t)}+e^{t-t_{n,k}}-e^{-(t-t_{n,k})}}{e^{1/(n-1)}-e^{-1/(n-1)}}-1\right\}$$
$$\cdot\frac{X(t_{n,1})+(1-e^{-1/(n-1)})\sum_{i=2}^{n-1}X(t_{n,i})+X(t_{n,n})}{(n-2)(1-e^{-1/(n-1)})+2},\tag{1.44}$$

where $X^0_{T_n(h_o)}(t)$ is as in (1.42), and its IMSE is

$$\begin{aligned}\text{IMSE}^1_{T_n(h_o)}(1)&=V_{T_n(h_o)}+\frac{1+e^{1/(n-1)}}{2+n(e^{1/(n-1)}-1)}B^2_{T_n(h_o)}\\&=\frac{1}{3}\left\{\frac{1}{n-1}+\frac{1}{(n-1)^2}+\frac{1}{5(n-1)^3}+O(n^{-4})\right\},\end{aligned}\tag{1.45}$$

where $V_{T_n(h_o)}$ and $B^2_{T_n(h_o)}$ are as in (1.43).

Here the linear interpolator given by (1.27) is also unbiased, and through some straightforward calculations, its IMSE is

$$\begin{aligned}\text{IMSE}^{\ell}_{T_n(h_o)}(1)&=\tfrac{1}{3}\{6(3-2n)+[12(n-1)^2-1](1-e^{-1/(n-1)})\}\\&=\frac{1}{3}\left\{\frac{1}{n-1}-\frac{1}{15(n-1)^3}+O(n^{-4})\right\}.\end{aligned}\tag{1.46}$$

To see the pattern of convergence (to zero) of the IMSE for these three estimators, we plotted their IMSE's versus the sample size n up to 20 for different values of the mean β and found that when β is small, say $\beta<5$, nearly no difference is displayed and the plot corresponding to $\beta=3.5$ is shown in Figure 5(c) with $\text{IMSE}^1_{T_n(h_o)}(1)$, $\text{IMSE}^0_{T_n(h_o)}(1)$ and $\text{IMSE}^{\ell}_{T_n(h_o)}(1)$, denoted by $\text{IMSE}^1_{T_n}$, $\text{IMSE}^0_{T_n}$ and $\text{IMSE}^{\ell}_{T_n}$, respectively; and when β is large the estimator $X^0_{T_n}(t)$ has poor performance for small sample sizes while the other two estimators have, approximately, the same performance, indeed, $\text{IMSE}^{\ell}_{T_n}-\text{ISME}^1_{T_n}=-3^{-1}n^{-2}+O(n^{-3})$, and the plot corresponding to $\beta=25$ is in Figure 5(d).

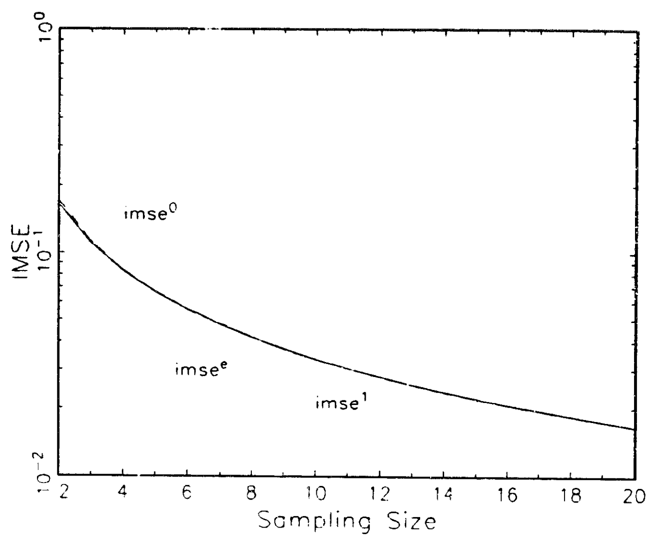

Fig. 5(c). IMSE vs. sample size ((beta = 3.5)

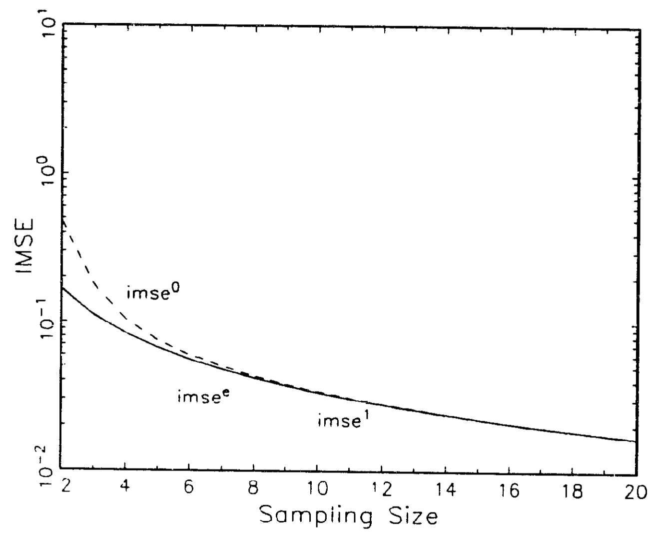

Fig. 5(d). IMSE vs. sample size (beta = 25).

So far, we have considered examples with covariances having constant jump function $\alpha(t)$. In the following example, we consider an error process with a non-constant jump function in order to illustrate the difference in performance between the uniform sampling design and the regular design determined by the asymptotically optimal sampling density.

Example 4. *A process with independent but nonstationary increments and non-constant jump function.* Here we consider the model (1.1) with mean zero, i.e. $m(t) \equiv 0$, and covariance function

$$R(t, s) = \int_0^{\min(t,s)} g(u)\, \mathrm{d}u$$

where $g(\cdot)$ is a nonnegative Riemann integrable function on [0, 1]; i.e. $X(t) = \int_0^t \sqrt{g(u)}\, \mathrm{d}W(u)$, where $W(\cdot)$ is a Wiener process with convariance function

$\min(t, s)$. Clearly, when $g(u) \equiv 1$, $X(t)$ is the Wiener process discussed in Example 2 with $R(t, s) = \min(t, s)$.

It is straightforward to obtain the jump function

$$\alpha(t) = g(t), \quad t \in [0, 1].$$

For simplicity, we choose $g(t) = t^{\beta}$ where $\beta > 0$ is a constant, in which case

$$R(t, s) = \{\min(t, s)\}^{\beta+1}/(\beta+1).$$

For any sampling design T_n, the corresponding BLUE of $X(t)$ is

$$X^0_{T_n}(t) = \frac{(t^{\beta+1}_{n,k+1} - t^{\beta+1})X(t_{n,k}) + (t^{\beta+1} - t^{\beta+1}_{n,k})X(t_{n,k+1})}{t^{\beta+1}_{n,k+1} - t^{\beta+1}_{n,k}}, \tag{1.47}$$

if $t_{n,k} \leqslant t \leqslant t_{n,k+1}$, $k = 0, 1, \ldots, n-1$, where $t_{n,0} = 0$, and $X^0_{T_n}(t) = X(t_{n,n})$, if $t_{n,n} \leqslant t \leqslant 1$. This is $(\beta+1)$-power interpolation: linear interpolation when $\beta = 0$, quadratic interpolation when $\beta = 1$, etc. Here, the asymptotically optimal sampling density function is $h_o(t) = (1 + \beta/2)t^{\beta/2}$ and the regular sampling designs determined by h_o are $T_n(h_o) = \{t_{n,k}\}_1^n$ with

$$t_{n,k} = \{(k-1)/(n-1)\}^{2/(2+\beta)}, \quad k = 1, \ldots, n.$$

To compare the performance of the uniform design $T_n(1)$ and the regular design $T_n(h_o)$, we plotted the IMSE of $X^0_{T_n(1)}(t)$, denoted by IMSE(unif), and the IMSE of $X^0_{T_n(h_o)}(t)$, denoted by IMSE(reg), versus the sample size n for values up to 20 and for a variety of values of the parameter β. We found that the larger β, the more significant the improvement of $T_n(h_o)$ over $T_n(1)$ is, which is intuitively clear because when β is small the regular design $T_n(h_o)$ is close to uniform. For $\beta = 2$, the plots are shown in Figure 6, from which one can see the considerable improvement of $T_n(h_o)$ over $T_n(1)$.

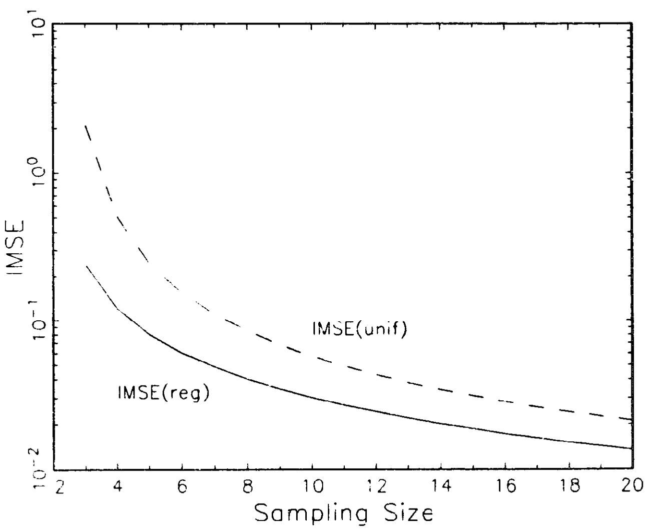

Fig. 6. IMSE vs. sample size.

2. Proofs

For any positive weight function $W(t)$ and random process $X(t)$ in (1.1), the random process $Y(t)=\sqrt{W(t)}\,X(t)$ has a mean function $\sqrt{W(t)}\,m(t)$ and centered component $\sqrt{W(t)}\,N(t)$. It is easily verified that the BLUE of $Y(t)$ is $\sqrt{W(t)}$ times the BLUE of $X(t)$ and furthermore, the weighted IMSE of the BLUE of $X(t)$ with the weight function $W(t)$ equals the unweighted IMSE of the BLUE of $Y(t)$ (i.e. with weight$\equiv 1$). Thus, without loss of generality, we can take $W(t)\equiv 1$ in the following proofs. For notational simplicity we will write IMSE for IMSE(1) throughout the proofs.

Before turning to the proofs of the theorems, we establish the following lemmas.

Lemma 1. *Under Conditions* (C1)-(C2), *if* $T_n=\{t_{n,i}\}_1^n$ *is a sequence of designs, then* $\mathrm{IMSE}_{T_n}^0\to 0$ *implies*

$$\max_{1\leqslant i\leqslant n-1}(t_{n,i+1}-t_{n,i})\to 0. \tag{2.1}$$

Proof. Write

$$f(t)=\int_0^1 R(t,s)\,\mathrm{d}s=\mathscr{E}N(t)\int_0^1 N(s)\,\mathrm{d}s.$$

Then it follows from the Cauchy-Schwarz inequality that

$$\begin{aligned}
&\mathrm{IMSE}_{T_n}^0\int_0^1\int_0^1 R(t,s)\,\mathrm{d}t\,\mathrm{d}s\\
&=\int_0^1 \mathscr{E}\{N(t)-r'_{T_n}(t)R_{T_n}^{-1}N_{T_n}\}^2\,\mathrm{d}t\;\mathscr{E}\left(\int_0^1 N(s)\,\mathrm{d}s\right)^2\\
&\geqslant\int_0^1\left\{\mathscr{E}(N(t)-r'_{T_n}(t)R_{T_n}^{-1}N_{T_n})\int_0^1 N(s)\,\mathrm{d}s\right\}^2\mathrm{d}t\\
&=\int_0^1\{f(t)-r'_{T_n}(t)R_{T_n}^{-1}f_{T_n}\}^2\,\mathrm{d}t\\
&\geqslant\left\{\int_0^1(f(t)-r'_{T_n}(t)R_{T_n}^{-1}f_{T_n})\,\mathrm{d}t\right\}^2=\|f-P_{T_n}f\|_R^4
\end{aligned}$$

where P_{T_n} is the projection of f onto span$\{R(\cdot,t_{n,i}),\ i=1,\ldots,n\}$. This together with Lemma 3.3 in Sacks and Ylvisaker (1966) establishes (2.1). □

Lemma 1 implies that any asymptotically optimal sequence of designs satisfies (2.1). Therefore, it suffices to consider sequences of designs satisfying (2.1).

Let $\hat{N}_{T_n}(t)$ be the projection of $N(t)$ onto span$\{N(t_{n,i}),\ i=1,\ldots,n\}$ and $N'_{T_n}=(N(t_{n,1}),\ldots,N(t_{n,n}))$. Then

$$\hat{N}_{T_n}(t)=r'_{T_n}(t)R_{T_n}^{-1}N_{T_n}\triangleq\hat{C}'_{T_n}(t)N_{T_n},$$

where $\hat{C}'_{T_n}(t)=(\hat{C}_1(t),\ldots,\hat{C}_n(t))=r'_{T_n}(t)R_{T_n}^{-1}$.

Define

$$p(t)=R(t,t),$$

$$p_{T_n}(t)=\mathscr{E}(\hat{N}_{T_n}(t))^2=r'_{T_n}(t)R_{T_n}^{-1}r_{T_n}(t)=\hat{C}'_{T_n}(t)r_{T_n}(t).$$

Then, (1.7) can be written as

$$\mathrm{IMSE}^0_{T_n}=\sum_{i=1}^{n}\int_{t_{n,i}}^{t_{n,i+1}}[p(t)-p_{T_n}(t)]\,\mathrm{d}t. \tag{2.2}$$

Since $\hat{N}_{T_n}(t_{n,i})=N(t_{n,i})$, we have

$$p_{T_n}(t_{n,i})=p(t_{n,i}),\quad t_{n,i}\in T_n,\ i=1,\ldots,n, \tag{2.3}$$

and since each $\hat{C}_j(t)$ is a linear combination of $R(t_{n,i},t)$, $i=1,\ldots,n$, it follows that $\hat{C}_j(t)$ is piecewise continuously differentiable up to order two with knots $t_{n,i}$, $i=1,\ldots,n$.

Moreover, it follows from $\hat{N}_{T_n}(t_{n,i})=N(t_{n,i})$ and the invertibility of R_{T_n} that

$$\hat{C}_j(t_{n,i})=\begin{cases}1, & \text{when } j=i,\\ 0, & \text{otherwise,}\end{cases}\quad i,j=1,\ldots,n. \tag{2.4}$$

Note that for each $t\in[0,1)$,

$$\begin{aligned}p^{(1)}(t+)&=\lim_{\Delta t\searrow 0}\{R(t+\Delta t,t+\Delta t)-R(t,t)\}/\Delta t\\&=\lim_{\Delta t\searrow 0}\left\{\frac{R(t+\Delta t,t+\Delta t)-R(t+\Delta t,t)}{\Delta t}+\frac{R(t+\Delta t,t)-R(t,t)}{\Delta t}\right\}\\&=R^{0,1}(t,t-)+R^{1,0}(t+,t)=R^{0,1}(t,t-)+R^{0,1}(t,t+)\end{aligned}$$

where the third equality follows from Assumption (C1) and the last one from the symmetry of $R(t,s)$. Likewise, for each $t\in(0,1]$, we find

$$p^{(1)}(t-)=\lim_{\Delta t\nearrow 0}\{R(t+\Delta t,t+\Delta t)-R(t,t)\}/\Delta t=R^{0,1}(t,t-)+R^{0,1}(t,t+).$$

Thus, $p(t)$ is differentiable at each $t\in(0,1)$ and

$$p^{(1)}(t)=R^{0,1}(t,t-)+R^{0,1}(t,t+). \tag{2.5}$$

By direct calculations, we have, for $t\notin T_n$,

$$p^{(1)}_{T_n}(t)=2\hat{C}'_{T_n}(t)r^{(1)}_{T_n}(t)$$

and by (2.4),

$$p_{T_n}^{(1)}(t_{n,i}+)=2\sum_{j=1}^{n}\hat{C}_j(t_{n,i})R^{0,1}(t_{n,j},t_{n,i}+)=2R^{0,1}(t_{n,i},t_{n,i}+).$$

Thus,

$$p^{(1)}(t_{n,i})-p_{T_n}^{(1)}(t_{n,i}+)=R^{0,1}(t_{n,i},t_{n,i}-)-R^{0,1}(t_{n,i},t_{n,i}+)=\alpha(t_{n,i}). \tag{2.6}$$

The second derivative of $p_{T_n}(t)$ is expressed in the following lemma.

Lemma 2. *Under Assumptions* (C1) *and* (C2), *we have, for any* $t\in(t_{n,i},t_{n,i+1})$,

$$p_{T_n}^{(2)}(t)=2\alpha(t_{n,i})/(t_{n,i+1}-t_{n,i})+\mathrm{O}(1),\quad i=1,\ldots,n-1, \tag{2.7}$$

where $\mathrm{O}(1)$ *is uniformly bounded in* t *and* n.

Proof. For convenience, we denote $t_{n,i}$ by t_i. Note that for $t\in(t_i,t_{i+1})$,

$$\begin{aligned}p_{T_n}^{(2)}(t)&=2\mathrm{d}[\hat{C}'_{T_n}(t)r_{T_n}^{(1)}(t)]/\mathrm{d}t\\&=2\sum_{j=1}^{n}\hat{C}_j^{(1)}(t)R^{0,1}(t_j,t)+2\sum_{j=1}^{n}\hat{C}_j(t)R^{0,2}(t_j,t).\end{aligned} \tag{2.8}$$

For the first term, using the Taylor expansions

$$R(t_j,t_i)=R(t_j,t)+R^{0,1}(t_j,t)(t_i-t)+\int_t^{t_i}(t_i-u)R^{0,2}(t_j,u)\,\mathrm{d}u,$$

$$R(t_j,t_{i+1})=R(t_j,t)+R^{0,1}(t_j,t)(t_{i+1}-t)+\int_t^{t_{i+1}}(t_{i+1}-u)R^{0,2}(t_j,u)\,\mathrm{d}u,$$

we have

$$\begin{aligned}R^{0,1}(t_j,t)d_i=\{R(t_j,t_{i+1})-R(t_j,t_i)\}-\bigg\{\int_t^{t_{i+1}}&(t_{i+1}-u)R^{0,2}(t_j,u)\,\mathrm{d}u\\&-\int_t^{t_i}(t_i-u)R^{0,2}(t_j,u)\,\mathrm{d}u\bigg\},\end{aligned} \tag{2.9}$$

where $d_i=t_{i+1}-t_i$, and similarly,

$$\begin{aligned}\hat{C}_j^{(1)}(t)d_i&=\{\hat{C}_j(t_{i+1})-\hat{C}_j(t_i)\}\\&\quad-\bigg\{\int_t^{t_{i+1}}(t_{i+1}-u)\hat{C}_j^{(2)}(u)\,\mathrm{d}u-\int_t^{t_i}(t_i-u)\hat{C}_j^{(2)}(u)\,\mathrm{d}u\bigg\}.\end{aligned} \tag{2.10}$$

For the second term in (2.8), as well as for the three of four terms resulting from multplying out (2.9) and (2.10), we proceed as follows.

By Cauchy-Schwarz inequality, for $t\in(t_i, t_{i+1})$, $s\in[0,1]$,

$$\left|\sum_{j=1}^{n}\int_{t}^{t_{i+1}}(t_{i+1}-u)R^{0,2}(t_j,u)\,\mathrm{d}u\cdot\hat{C}_j(s)\right|$$

$$=\left|\int_{t}^{t_{i+1}}(t_{i+1}-u)r_{T_n}^{(2)\prime}(u)R_{T_n}^{-1}r_{T_n}(s)\,\mathrm{d}u\right|$$

$$\leq\int_{t}^{t_{i+1}}(t_{i+1}-u)|r_{T_n}^{(2)\prime}(u)R_{T_n}^{-1}r_{T_n}(s)|\,\mathrm{d}u$$

$$\leq\int_{t}^{t_{i+1}}(t_{i+1}-u)\{r_{T_n}^{(2)\prime}(u)R_{T_n}^{-1}r_{T_n}^{(2)}(u)\cdot r_{T_n}'(s)R_{T_n}^{-1}r_{T_n}(s)\}^{1/2}\,\mathrm{d}u$$

$$\leq\int_{t}^{t_{i+1}}(t_{i+1}-u)\|R^{0,2}(\cdot,u)\|_R\,\mathrm{d}u\cdot R^{1/2}(s,s)$$

$$\leq\tfrac{1}{2}\sup_{0\leq t\leq 1}\|R^{0,2}(\cdot,t)\|_R\cdot\sup_{0\leq s\leq 1}R^{1/2}(s,s)\cdot(t_{i+1}-t)^2\leq\text{const.}\ d_i^2, \tag{2.11}$$

where the second and the third inequalities follow from the fact that $r_{T_n}^{(\nu)\prime}(t)R_{T_n}^{-1}r_{T_n}^{(\nu)}(t)$ is the norm of the projection of $R^{0,\nu}(\cdot,t)$ onto span$\{R(t_i,\cdot),\ i=1,\dots,n\}$, $\nu=0,2$, and the last one from the continuity of $R(\cdot,\cdot)$ and Assumption (C2). Likewise, for all $t\in(t_i,t_{i+1})$, $s\in[0,1]$,

$$\begin{aligned}
&\left|\sum_{j=1}^{n}\int_{t_i}^{t}(t_i-u)R^{0,2}(t_j,u)\,\mathrm{d}u\cdot\hat{C}_j(s)\right|\leq\text{const.}\ d_i^2,\\
&\left|\sum_{j=1}^{n}\int_{t}^{a}(a-u)\hat{C}_j^{(2)}(u)\,\mathrm{d}u\cdot R(t_j,s)\right|\leq\text{const.}\ d_i^2,\\
&\left|\sum_{j=1}^{n}\int_{t}^{a}(a-u)R^{0,2}(t_j,u)\,\mathrm{d}u\cdot\int_{t}^{b}(b-v)\hat{C}_j^{(2)}(v)\,\mathrm{d}v\right|\leq\text{const.}\ d_i^4,
\end{aligned}\tag{2.12}$$

where $a,b=t_i$ or t_{i+1}. By (2.8), (2.9)-(2.12) and (2.4),

$$\begin{aligned}
p_{T_n}^{(2)}(u)d_i^2&=2\sum_{j=1}^{n}d_i\hat{C}_j^{(1)}(u)\cdot d_iR^{0,1}(t_j,u)+2d_i^2\sum_{j=1}^{n}\hat{C}_j(u)R^{0,2}(t_j,u)\\
&=2\sum_{j=1}^{n}\{\hat{C}_j(t_{i+1})-\hat{C}_j(t_i)\}\{R(t_j,t_{i+1})-R(t_j,t_i)\}+d_i^2\mathrm{O}(1)\\
&=2\{R(t_{i+1},t_{i+1})-R(t_{i+1},t_i)-R(t_i,t_{i+1})+R(t_i,t_i)\}+d_i^2\mathrm{O}(1).
\end{aligned}$$

Using Taylor expansions and Assumption (C2) repeatedly, we obtain

$$R(t_{i+1},t_{i+1})-R(t_{i+1},t_i)=R^{0,1}(t_{i+1},t_i)d_i+d_i^2\mathrm{O}(1)=R^{0,1}(t_i,t_i-)d_i+d_i^2\mathrm{O}(1)$$

and likewise,

$$R(t_i,t_{i+1})-R(t_i,t_i)=R^{0,1}(t_i,t_i+)d_i+d_i^2\mathrm{O}(1).$$

Therefore,

$$p_{T_n}^{(2)}(t)d_i^2=2\{R^{0,1}(t_i,t_i-)-R^{0,1}(t_i,t_i+)\}d_i+d_i^2\mathrm{O}(1)=2\alpha(t_i)d_i+d_i^2\mathrm{O}(1),$$

from which (2.7) follows. □

Proof of Theorem 1. The IMSE is given by (2.2). By Taylor expansion, (2.3) and (2.6), we have for $t\in(t_i, t_{i+1})$,

$$\begin{aligned}p(t)-p_{T_n}(t)&=\{p^{(1)}(t_i)-p_{T_n}^{(1)}(t_i+)\}(t-t_i)+\int_{t_i}^{t}(t-u)\{p^{(2)}(u)-p_{T_n}^{(2)}(u)\}\,\mathrm{d}u\\&=\alpha(t_i)(t-t_i)+\int_{t_i}^{t}(t-u)p^{(2)}(u)\,\mathrm{d}u-\int_{t_i}^{t}(t-u)p_{T_n}^{(2)}(u)\,\mathrm{d}u.\end{aligned}$$

Proceeding as in the derivation of (2.5), we can show that $p(\cdot)$ has a second derivative at each $t\in(0,1)$ given by

$$p^{(2)}(t)=R^{0,2}(t,t-)+2R^{1,1}(t-,t+)+R^{0,2}(t,t+). \tag{2.13}$$

By (2.13), the boundness assumption in (C2) and Lemma 2, we obtain

$$p(t)-p_{T_n}(t)=\alpha(t_i)(t-t_i)+(t-t_i)^2\mathrm{O}(1)-(t-t_i)^2\alpha(t_i)/d_i,$$

which yields

$$\int_{t_i}^{t_{i+1}}[p(t)-p_{T_n}(t)]\,\mathrm{d}t=\tfrac{1}{2}\alpha(t_i)d_i^2-\tfrac{1}{3}\alpha(t_i)d_i^2+d_i^3\mathrm{O}(1)=\tfrac{1}{6}\alpha(t_i)d_i^2+d_i^3\mathrm{O}(1)$$

and then by (2.2),

$$\mathrm{IMSE}_{T_n}^{0}=\tfrac{1}{6}\sum_{i=1}^{n}\alpha(t_i)d_i^2+\sum_{i=1}^{n}d_i^3\mathrm{O}(1), \tag{2.14}$$

where O(1) is bounded uniformly in n and t. By Hölder's inequality, for large n,

$$n\sum_{i=1}^{n}\alpha(t_i)d_i^2\geq\left\{\sum_{i=1}^{n}\sqrt{\alpha(t_i)}\,d_i\right\}^2=\left\{\int_0^1\sqrt{\alpha(t)}\,\mathrm{d}t\right\}^2+\mathrm{o}(1)$$

and

$$\sum_{i=1}^{n}d_i^3\geq\left\{\sum_{i=1}^{n}d_i\right\}^3\left\{\sum_{i=1}^{n}1^{3/2}\right\}^{-2}=n^{-2}.$$

Thus,

$$\liminf_n n\,\mathrm{IMSE}_{T_n}^{0}\geq\frac{1}{6}\left\{\int_0^1\sqrt{\alpha(t)}\,\mathrm{d}t\right\}^2, \tag{2.15}$$

which provides the asymptotic lower bound of $n\ \mathrm{IMSE}^0_{T_n}$. On the other hand, it follows from (1.4) and the mean value theorem of integrals that $d_i = 1/[h(w_i)(n-1)]$, where $w_i \in [t_i, t_{i+1}]$, $i=1,\ldots,n-1$. By (2.14), we can write $\mathrm{IMSE}^0_{T_n}$ as

$$\mathrm{IMSE}^0_{T_n(h)} = \frac{1}{6(n-1)} \sum_{i=1}^{n-1} \frac{\alpha(t_i)}{h(w_i)} d_i + \mathrm{O}(n^{-2}).$$

Then, by the Riemann integrability of the function $(\alpha/h)(\cdot)$, we have

$$\lim_{n\to\infty} n\ \mathrm{IMSE}^0_{T_n(h)} = \frac{1}{6}\int_0^1 \frac{\alpha(t)}{h(t)}\,\mathrm{d}t.$$

Hence, (ii) holds.

Replacing h by h_o, then clearly,

$$\lim_{n\to\infty} n\ \mathrm{IMSE}^0_{T_n(h_o)} = \frac{1}{6}\int_0^1 \frac{\alpha(t)}{h_o(t)}\,\mathrm{d}t = \frac{1}{6}\left\{\int_0^1 \sqrt{\alpha(t)}\,\mathrm{d}t\right\}^2. \tag{2.16}$$

The results (i) and (iii) follow from (2.15) and (2.16). □

Proof of Theorem 2. By (1.13), since $G_{T_n} \geq 0$, we have the following inequality

$$\inf_{|T_n|=n} \mathrm{IMSE}^0_{T_n} \leq \inf_{|T_n|=n} \mathrm{IMSE}^q_{T_n} \leq \mathrm{IMSE}^0_{T_n(h)} + G_{T_n(h)}$$

and thus using Theorem 1 and taking $h = h_o$, we obtain

$$\begin{aligned}\lim_{n\to\infty} n\ \mathrm{IMSE}^0_{T_n(h_o)} &= \lim_{n\to\infty} n \inf_{|T_n|=n} \mathrm{IMSE}^0_{T_n} \leq \lim_{n\to\infty} n \inf_{|T_n|=n} \mathrm{IMSE}^q_{T_n} \\ &\leq \lim_{n\to\infty} n\ \mathrm{IMSE}^0_{T_n(h_o)} + \lim_{n\to\infty} n\, G_{T_n(h_o)}.\end{aligned}$$

Therefore, to show Theorem 2, it suffices to show (1.15).

Note that the (i,j) element of the matrix $F'_{T_n} R^{-1}_{T_n} F_{T_n}$ is $f'_{i,T_n} R^{-1}_{T_n} f_{j,T_n}$ and as $n\to\infty$, it tends to $s_{ij} = \int_0^1\int_0^1 \phi_i R \phi_j$ as follows from

$$\begin{aligned}&\left| f'_{i,T_n} R^{-1}_{T_n} f_{j,T_n} - \int_0^1\int_0^1 \phi_i R\phi_j \right| \\ &= \left| \mathscr{E}\left(f'_{i,T_n} R^{-1}_{T_n} N_{T_n} - \int_0^1 N\phi_i \right)\cdot \int_0^1 N\phi_j \right| \\ &\leq \|f_i - P_{T_n} f_i\|_R \left(\int_0^1\int_0^1 \phi_j R \phi_j \right)^{1/2} \leq \text{const.}\ \sigma_i s_{jj}^{1/2} \max_{1\leq k\leq n-1} d_k,\end{aligned} \tag{2.17}$$

where the last inequality follows from Lemma 3.1 in Sacks and Ylvisaker (1966) and $\sigma_i = \sup_{0\leq t\leq 1} |\phi_i(t)|$. Furthermore, if the regular sampling design $T_n(h)$ is used, (2.17) implies

$$\begin{aligned} f'_{i,T_n(h)} R^{-1}_{T_n(h)} f_{j,T_n(h)} - \int_0^1\int_0^1 \phi_i R\phi_j &= \mathrm{O}(1)\sigma_i s_{jj}^{1/2}(n-1)^{-1} \max_{0\leq t\leq 1} \frac{1}{h(t)} \\ &= \mathrm{O}(1)\sigma_i s_{jj}^{1/2}(n-1)^{-1},\end{aligned} \tag{2.18}$$

where $\mathrm{O}(1)$ is uniform in i, j and n, which means the convergence rate is n^{-1}.

By using the inequality $\|A^{-1}-B^{-1}\| \leq \|A^{-1}\|^2\|A-B\|\{1+\mathrm{O}(\|A-B\|)\}$, where A and B are two invertible matrices and $\|\cdot\|$ denotes the usual matrix norm, for instance $\|A\| = \max_i \sum_j |a_{ij}|$ with $A=(a_{ij})$, and (2.18) we can conclude that the matrix $(F'_{T_n(h)}R^{-1}_{T_n(h)}F_{T_n(h)})^{-1}$ tends to the matrix $(s_{ij})^{-1}=S^{-1}$ elementwise, as $n\to\infty$.

By Assumptions (C1)-(C2), similar to the derivation of (3.43) in Sacks and Ylvisaker (1966), we have, for every $k=1,\dots,n-1$, $i=1,\dots,q$ and $t\in(t_k,t_{k+1})$,

$$\begin{aligned}
V_{i,T_n}(t) &\triangleq r'_{T_n}(t)R^{-1}_{T_n}f_{i,T_n}-f_i(t)\\
&=\frac{1}{2}\Big\{-(\alpha\phi_i)(a_{i,k})+f'_{i,T_n}R^{-1}_{T_n}r^{(2)}_{T_n}(a_{i,k})\\
&\qquad -\int_0^1 R^{2,0}(a_{i,k},s)\phi_i(s)\,\mathrm{d}s\Big\}d_k(t-t_k)\\
&\quad+\frac{1}{2}\Big\{(\alpha\phi_i)(\tau_{i,k})+f'_{i,T_n}R^{-1}_{T_n}r^{(2)}_{T_n}(\tau_{i,k})\\
&\qquad -\int_0^1 R^{2,0}(\tau_{i,k},s)\phi_i(s)\,\mathrm{d}s\Big\}(t-t_k)^2,
\end{aligned}\tag{2.19}$$

where $a_{i,k}$, $\tau_{i,k}\in(t_k,t)$ and depend continuously on t. Since for every $t\in[0,1]$, $R^{0,2}(\cdot,t)\in\mathrm{RKHS}(R)$, there is a random variable $\xi_t\in\mathrm{span}\{N(s),0\leq s\leq 1\}$ such that $R^{0,2}(\cdot,t)=\mathscr{E}N(\cdot)\xi_t$ and $\|R^{0,2}(\cdot,t)\|^2_R=\mathscr{E}\xi_t^2$. Then by using Cauchy-Schwarz's inequality, Assumptions (C1)-(C2) and Lemma 3.1 in Sacks and Ylvisaker (1966), we have, for each $t\in(t_k,t_{k+1})$,

$$\begin{aligned}
&\left|f'_{i,T_n}R^{-1}_{T_n}r^{(2)}_{T_n}(t)-\int_0^1 R^{2,0}(t,s)\phi_i(s)\,\mathrm{d}s\right|\\
&=\left|\mathscr{E}\xi_t\left(f'_{i,T_n}R^{-1}_{T_n}N_{T_n}-\int_0^1 N(s)\phi_i(s)\,\mathrm{d}s\right)\right|\\
&\leq\left\{\mathscr{E}\xi_t^2\cdot\left(\int_0^1\int_0^1\phi_iR\phi_i-f'_{i,T_n}R^{-1}_{T_n}f_{i,T_n}\right)\right\}^{1/2}\\
&=\|R^{0,2}(\cdot,t)\|_R\cdot\|f_i-P_{T_n}f_i\|_R\leq\text{const.}\ \sigma_i\sup_{1\leq j\leq n-1}d_j,
\end{aligned}\tag{2.20}$$

where the const. is independent of i, k and t and σ_i is as in (2.17). From (2.19) and (2.20) we have, for each $k=1,\dots,n-1$, $i=1,\dots,q$ and $t\in(t_k,t_{k+1})$,

$$\begin{aligned}
V_{i,T_n}(t)&=-\tfrac{1}{2}(\alpha\phi_i)(a_{i,k})d_k(t-t_k)\\
&\quad+\tfrac{1}{2}(\alpha\phi_i)(\tau_{i,k})(t-t_k)^2+\mathrm{O}(1)\sigma_id_k^2\sup_{1\leq j\leq n-1}d_j,
\end{aligned}\tag{2.21}$$

where $a_{i,k}$, $\tau_{i,k}\in(t_k,t_{k+1})$ and $\mathrm{O}(1)$ is independent of i, k and t. Thus, for $i,j=1,\dots,q$,

$$\int_0^1 V_{i,T_n(h)}(t)V_{j,T_n(h)}(t)\,dt$$

$$= \sum_{k=1}^{n-1}\int_{t_k}^{t_{k+1}}\left\{-\tfrac{1}{2}(\alpha\phi_i)(a_{i,k})d_k(t-t_k)+\tfrac{1}{2}(\alpha\phi_i)(\tau_{i,k})(t-t_k)^2+\mathrm{O}(1)\sigma_i d_k^2 \sup_{1\le l\le n-1} d_l\right\}$$

$$\cdot\left\{-\tfrac{1}{2}(\alpha\phi_j)(a_{j,k})d_k(t-t_k)+\tfrac{1}{2}(\alpha\phi_j)(\tau_{j,k})(t-t_k)^2+\mathrm{O}(1)\sigma_j d_k^2 \sup_{1\le l\le n-1} d_l\right\}dt$$

$$= \sum_{k=1}^{n-1}\left\{\tfrac{1}{12}(\alpha\phi_i)(a_{i,k})(\alpha\phi_j)(a_{j,k})d_k^5-\tfrac{1}{16}(\alpha\phi_i)(a_{i,k})(\alpha\phi_j)(\tau_{j,k})d_k^5\right.$$

$$\left.-\tfrac{1}{16}(\alpha\phi_i)(\tau_{i,k})(\alpha\phi_j)(a_{j,k})d_k^5+\tfrac{1}{20}(\alpha\phi_i)(\tau_{i,k})(\alpha\phi_j)(\tau_{j,k})d_k^5+\mathrm{O}(1)d_k^5 \sup_{1\le l\le n-1} d_l\right\}$$

$$= \sum_{k=1}^{n-1}\left\{\tfrac{1}{12}(\alpha\phi_i)(a_{i,k})(\alpha\phi_j)(a_{j,k})-\tfrac{1}{16}(\alpha\phi_i)(a_{i,k})(\alpha\phi_j)(\tau_{j,k})-\tfrac{1}{16}(\alpha\phi_i)(\tau_{i,k})(\alpha\phi_j)(a_{j,k})\right.$$

$$\left.+\tfrac{1}{20}(\alpha\phi_i)(\tau_{i,k})(\alpha\phi_j)(\tau_{j,k})+\mathrm{O}(1)\sup_{1\le l\le n-1} d_l\right\}\frac{1}{n^4h^4(w_k)}d_k$$

and by the Riemann integrability of the function $\alpha^2\phi_i\phi_j h^{-4}$,

$$n^4\int_0^1 V_{i,T_n(h)}(t)V_{j,T_n(h)}(t)\,dt$$

$$\to \int_0^1 \alpha^2(t)\phi_i(t)\phi_j(t)\{\tfrac{1}{12}-\tfrac{1}{16}-\tfrac{1}{16}+\tfrac{1}{20}\}\frac{1}{h^4(t)}\,dt$$

$$= \frac{1}{120}\int_0^1 \frac{\alpha^2(t)}{h^4(t)}\phi_i(t)\phi_j(t)\,dt. \tag{2.22}$$

Finally, writing $S^{-1}=(\nu_{ij})_{q\times q}$, yields

$$n^4G_{T_n(h)}\to \sum_{i=1}^{q}\sum_{j=1}^{q}\nu_{ij}\frac{1}{120}\int_0^1\frac{\alpha^2(t)}{h^4(t)}\phi_i(t)\phi_j(t)\,dt$$

$$= \frac{1}{120}\int_0^1\frac{\alpha^2(t)}{h^4(t)}\phi'(t)S^{-1}\phi(t)\,dt, \tag{2.23}$$

and (1.15) follows. □

Proof of Theorem 3. Note that by putting $A=(\beta_i\beta_j)_{q\times q}$, we can write

$$H_{T_n}(t)=(F'_{T_n}R_{T_n}^{-1}r_{T_n}(t)-F(t))'A(F'_{T_n}R_{T_n}^{-1}r_{T_n}(t)-F(t)),$$

which is of the same form as $G_{T_n}(t)$ except that here the matrix A is constant (i.e. does not depend on n). Thus, (1.19) follows immediately from (2.23). □

Proof of Theorem 4. If we can show that

$$\lim_{n\to\infty} n^4 G_{T_n(h)} = \frac{1}{120}\int_0^1 \frac{\alpha^2(t)}{h^4(t)}\left\{\sum_{k=1}^{q} \lambda_k^{-1} e_k^2(t)\right\} \mathrm{d}t, \tag{2.24}$$

and

$$\lim_{n\to\infty} n^4 Q_{T_n(h)} = \frac{1}{120}\int_0^1 \frac{\alpha^2(t)}{h^4(t)}\left\{\psi(t) - \sum_{k=1}^{q} \beta_k \lambda_k^{-1/2} e_k(t)\right\}^2 \mathrm{d}t, \tag{2.25}$$

then (1.24) follows. Indeed for the regular sampling design $T_n(h)$, by Theorem 1, (1.23) and (2.24)-(2.25), we have

$$\begin{aligned}\lim_{n\to\infty} n\,\mathrm{IMSE}^0_{T_n(h_0)} &= \lim_{n\to\infty} n \inf_{|T_n|=n} \mathrm{IMSE}^0_{T_n} \leq \lim_{n\to\infty} n \inf_{|T_n|=n} \mathrm{IMSE}^{\infty,q}_{T_n} \\ &\leq \lim_{n\to\infty} n\,\mathrm{IMSE}^0_{T_n(h)} + \lim_{n\to\infty} nG_{T_n(h)} + \lim_{n\to\infty} nQ_{T_n(h)} \\ &= \lim_{n\to\infty} n\,\mathrm{IMSE}^0_{T_n(h)},\end{aligned}$$

which implies that (i) in Theorem 1 holds for the estimator $X^q_{T_n}(t)$. The results (ii) and (iii) follow from (1.23) and (2.24)-(2.25). Therefore, it remains to show (2.24) and (2.25).

Proof of (2.24). By the orthonormality of the f_i's in the RKHS(R), we have, for $i, j = 1, \ldots, q$, $\langle f_i, f_j\rangle_R = \delta_{ij}$, and moreover, by putting $\phi_i(t) = \lambda_i^{-1/2} e_i(t)$ in (2.18) and using

$$\begin{aligned}\sup_{0\le t\le 1} |\phi_i(t)| &= \lambda_i^{-3/2} \sup_{0\le t\le 1} |\lambda_i e_i(t)| \\ &= \lambda_i^{-3/2} \sup_{0\le t\le 1} \left|\mathscr{E} N(t) \int_0^1 N(s) e_i(s)\,\mathrm{d}s\right| \\ &\leq \lambda_i^{-3/2} \sup_{0\le t\le 1} R^{1/2}(t,t)\left\{\int_0^1\int_0^1 e_i R e_i\right\}^{1/2} = \text{const.}\ \lambda_i^{-1},\end{aligned}$$

we have

$$f'_{i,T_n(h)} R^{-1}_{T_n(h)} f_{j,T_n(h)} - \delta_{ij} = \mathrm{O}(1)\lambda_i^{-1}(n-1)^{-1}, \tag{2.26}$$

which implies, as in the proof of Theorem 2, that

$$(F'_{T_n(h)} R^{-1}_{T_n(h)} F_{T_n(h)})^{-1} \xrightarrow[n\to\infty]{} I \quad \text{elementwise.} \tag{2.27}$$

Now using (2.22) with $\phi_i(t) = \lambda_i^{-1/2} e_i(t)$, (2.24) follows.

Proof of (2.25). By the unbiasedness constraints on the parameters β_i, $i=1,\ldots,q$, and the form of $\hat{C}_{T_n(h)}(t)$ as in (1.11), we can write

$$
\begin{aligned}
Q_{T_n(h)} &= \int_0^1 \{\hat{C}'_{T_n(h)}(t)m_{T_n(h)} - m(t)\}^2\,\mathrm{d}t \\
&= \int_0^1 \{r'_{T_n(h)}(t)R^{-1}_{T_n(h)}m_{T_n(h)} - m(t)\}^2\,\mathrm{d}t \\
&\quad -2\int_0^1 \{r'_{T_n(h)}(t)R^{-1}_{T_n(h)}m_{T_n(h)} - m(t)\} \\
&\qquad \cdot\{F'_{T_n(h)}R^{-1}_{T_n(h)}r_{T_n(h)}(t) - F(t)\}' \\
&\qquad \cdot(F'_{T_n(h)}R^{-1}_{T_n(h)}F_{T_n(h)})^{-1}F'_{T_n(h)}R^{-1}_{T_n(h)}m_{T_n(h)}\,\mathrm{d}t \\
&\quad +\int_0^1 \{(F'_{T_n(h)}R^{-1}_{T_n(h)}r_{T_n(h)}(t) - F(t))' \\
&\qquad \cdot(F'_{T_n(h)}R^{-1}_{T_n(h)}F_{T_n(h)})^{-1}F'_{T_n(h)}R^{-1}_{T_n(h)}m_{T_n(h)}\}^2\,\mathrm{d}t \\
&\triangleq D_{1,T_n(h)} + D_{2,T_n(h)} + D_{3,T_n(h)}.
\end{aligned}
$$

For the first term $D_{1,T_n(h)}$, by using (2.22) with $\phi_i=\psi$, we have

$$n^4 D_{1,T_n(h)} \to \frac{1}{120}\int_0^1 \frac{\alpha^2(t)}{h^4(t)}\psi^2(t)\,\mathrm{d}t.$$

For the second term $D_{2,T_n(h)}$, note that by the orthonormality of f_i's, as $n\to\infty$,

$$F'_{T_n(h)}R^{-1}_{T_n(h)}m_{T_n(h)} \to (\beta_1,\ldots,\beta_q)' \quad \text{elementwise,}$$

and then from (2.22) and (2.27) with $\phi_i=\psi$ and $\phi_j=\lambda_j^{-1/2}e_j$,

$$n^4 D_{2,T_n(h)} \to -\frac{2}{120}\int_0^1 \frac{\alpha^2(t)}{h^4(t)}\psi(t)\left\{\sum_{k=1}^{q}\beta_k\lambda_k^{-1/2}e_k(t)\right\}\mathrm{d}t.$$

Likewise,

$$n^4 D_{3,T_n(h)} \to \frac{1}{120}\int_0^1 \frac{\alpha^2(t)}{h^4(t)}\left\{\sum_{k=1}^{q}\beta_k\lambda_k^{-1/2}e_k(t)\right\}^2\mathrm{d}t.$$

Therefore,

$$n^4 Q_{T_n(h)} \to \frac{1}{120}\int_0^1 \frac{\alpha^2(t)}{h^4(t)}\left\{\psi(t)-\sum_{k=1}^{q}\beta_k\lambda_k^{-1/2}e_k(t)\right\}^2\mathrm{d}t. \qquad \square$$

Proof of Theorem 5. By a similar analysis as at the beginning of the proof of Theorem 4, it can be verified that to show Theorem 5, it suffices to show the following:

$$\lim_{n\to\infty} nG_{T_n(h)} = 0, \tag{2.28}$$

$$\lim_{n\to\infty} nQ_{T_n(h)} = 0. \tag{2.29}$$

By the orthonormality of f_i's and the following inequality,

$$|f'_{i,T_n(h)}R^{-1}_{T_n(h)}f_{j,T_n(h)}| \leq \{f'_{i,T_n(h)}R^{-1}_{T_n(h)}f_{i,T_n(h)} \cdot f'_{j,T_n(h)}R^{-1}_{T_n(h)}f_{j,T_n(h)}\}^{1/2} \leq \|f_i\|_R\,\|f_j\|_R = 1,$$

we have for $i, j = 1, \ldots, q$, as $n \to \infty$,

$$f'_{i,T_n(h)}R^{-1}_{T_n(h)}f_{j,T_n(h)} = \delta_{ij} + \mathrm{o}(1), \tag{2.30}$$

where o(1) is uniform in i and j. By (2.30) and straightforward computations, it is verified that

$$(F'_{T_n(h)}R^{-1}_{T_n(h)}F_{T_n(h)})^{-1} = (I + \mathrm{o}(1)UU')^{-1} = I - \{\mathrm{o}(1)/[1+\mathrm{o}(1)q_n]\}UU',$$

where $U = (1, \ldots, 1)'$. Thus, $G_{T_n(h)}$ can be written as

$$G_{T_n(h)} = \sum_{i=1}^{q_n} \int_0^1 V^2_{i,T_n(h)}(t)\,\mathrm{d}t - \{\mathrm{o}(1)/[1+\mathrm{o}(1)q_n]\} \int_0^1 \left(\sum_{i=1}^{q_n} V_{i,T_n(h)}(t)\right)^2 \mathrm{d}t,$$

which together with (2.24) and the boundedness of α^2/h^4, imply that for large n,

$$\begin{aligned} nG_{T_n(h)} &= \frac{n^{-3}}{120} \sum_{i=1}^{q_n} \lambda_i^{-1} \int_0^1 \frac{\alpha^2(t)}{h^4(t)} e_i^2(t)\,\mathrm{d}t \\ &\quad + \mathrm{o}(1)n^{-3} \int_0^1 \frac{\alpha^2(t)}{h^4(t)} \left(\sum_{i=1}^{q_n} \lambda_i^{-1} e_i(t)\right)^2 \mathrm{d}t + \mathrm{o}(n^{-3}) \\ &= \mathrm{const.}(1+\mathrm{o}(1))n^{-3} \sum_{i=1}^{q_n} \lambda_i^{-1}, \end{aligned}$$

thus (2.28) follows from (1.26).

By use of (2.30) and the matrix inversion formula above, we have for $i = q+1, \ldots,$

$$\begin{aligned} &\hat{C}'_{T_n(h)}(t) f_{i,T_n(h)} - f_i(t) \\ &= V_{i,T_n(h)}(t) + \sum_{k=1}^{q_n} V_{k,T_n(h)}(t) f'_{k,T_n(h)} R^{-1}_{T_n(h)} f_{i,T_n(h)} \\ &\quad + [\mathrm{o}(1)/(1+\mathrm{o}(1)q_n)] \sum_{k=1}^{q_n} V_{k,T_n(h)}(t) \left(\sum_{l=1}^{q_n} f'_{l,T_n(h)} R^{-1}_{T_n(h)} f_{i,T_n(h)} \right) \\ &= V_{i,T_n(h)}(t) + \mathrm{o}(1) \sum_{k=1}^{q_n} V_{k,T_n(h)}(t) \\ &\quad + [\mathrm{o}^2(1)q_n/(1+\mathrm{o}(1)q_n)] \sum_{k=1}^{q_n} V_{k,T_n(h)}(t) \\ &= V_{i,T_n(h)}(t) + \mathrm{o}(1) \sum_{k=1}^{q_n} V_{k,T_n(h)}(t), \end{aligned}$$

where o(1) is uniform in i. Thus $Q_{T_n(h)}$ can be written as

$$Q_{T_n(h)}=\int_0^1\left\{\sum_{i=q_n+1}^{\infty}\beta_i\left[V_{i,T_n(h)}(t)+\mathrm{o}(1)\sum_{k=1}^{q_n}V_{k,T_n(h)}(t)\right]\right\}^2\mathrm{d}t$$

$$=\int_0^1\left\{\sum_{i=q_n+1}^{\infty}\beta_iV_{i,T_n(h)}(t)\right\}^2\mathrm{d}t$$

$$+2\mathrm{o}(1)\sum_{i=q_n+1}^{\infty}\beta_i\int_0^1\left\{\sum_{j=q_n+1}^{\infty}\beta_jV_{j,T_n(h)}(t)\right\}\sum_{k=1}^{q_n}V_{k,T_n(h)}(t)\,\mathrm{d}t$$

$$+\mathrm{o}^2(1)\left(\sum_{i=q_n+1}^{\infty}\beta_i\right)^2\int_0^1\left\{\sum_{k=1}^{q_n}V_{k,T_n(h)}(t)\right\}^2\mathrm{d}t$$

$$\triangleq E_{1,T_n(h)}+E_{2,T_n(h)}+E_{3,T_n(h)}.$$

For the first term $E_{1,T_n(h)}$, we use the inequality

$$|V_{i,T_n(h)}(t)|=\left|\mathscr{E}[r'_{T_n(h)}(t)R^{-1}_{T_n(h)}N_{T_n(h)}-N(t)]\lambda_i^{-1/2}\int_0^1 Ne_i\right|$$

$$\leqslant\{R(t,t)-r'_{T_n(h)}(t)R^{-1}_{T_n(h)}r_{T_n(h)}(t)\}^{1/2}\cdot\left\{\lambda_i^{-1}\int_0^1\int_0^1 e_iRe_i\right\}^{1/2}$$

$$=\{\mathrm{MSE}^0_{T_n(h)}(t)\}^{1/2},$$

to produce

$$nE_{1,T_n(h)}\leqslant\left\{\sum_{i=q_n+1}^{\infty}|\beta_i|\right\}^2\cdot n\ \mathrm{IMSE}^0_{T_n(h)}\to 0$$

as$n\to\infty$, since $\sum_{i=1}^{\infty}|\beta_i|<\infty$. Using the Cauchy-Schwarz inequality and $|V_{i,T_n(h)}(t)|\leqslant\{\mathrm{MSE}^0_{T_n(h)}(t)\}^{1/2}$ we obtain

$$n|E_{2,T_n(h)}|\leqslant\mathrm{o}(1)n\sum_{i=q_n+1}^{\infty}|\beta_i|\left\{\int_0^1\left(\sum_{j=q_n+1}^{\infty}\beta_nV_{j,T_n(h)}(t)\right)^2\mathrm{d}t\right.$$

$$\left.\cdot\int_0^1\left(\sum_{k=1}^{q_n}V_{k,T_n(h)}(t)\right)^2\mathrm{d}t\right\}^{1/2}$$

$$\leqslant\mathrm{o}(1)\left(\sum_{i=q_n+1}^{\infty}|\beta_i|\right)^2\{n\ \mathrm{IMSE}^0_{T_n(h)}\}^{1/2}$$

$$\cdot\left\{n\sum_{k=1}^{q_n}\sum_{j=1}^{q_n}\int_0^1V_{k,T_n(h)}(t)V_{j,T_n(h)}(t)\,\mathrm{d}t\right\}^{1/2}.$$

Then

$$nE_{2,T_n(h)}\to 0$$

will follow if we show that $n\sum_{k=1}^{q_n}\sum_{j=1}^{q_n}\int_0^1 V_kV_j$ is bounded. From (2.21)-(2.22) it follows that

$$\int_0^1 V_kV_j=\frac{n^{-4}}{120}\int_0^1\frac{\alpha^2}{h^4}\phi_k\phi_j+o(n^{-5})\sigma_k\sigma_j,$$

and since $\sigma_k\leqslant \text{Const.}\,\lambda_k^{-1/2}$ and α/h^2 is bounded we have

$$n\sum_{k=1}^{q_n}\sum_{j=1}^{q_n}\int_0^1 V_kV_j\leqslant n^{-3}\,\text{Const.}\sum_{k=1}^{q_n}\lambda_k^{-1}+o(n^{-4})\left\{\sum_{k=1}^{q_n}\lambda_k^{-1/2}\right\}^2.$$

The first term tends to 0 by (1.26), and the second is upper bounded by $o(n^{-4})q_n\sum_{k=1}^{q_n}\lambda_k^{-1}$, which also tends to zero by (1.26). Similar to the derivation of (2.28), it is shown that

$$nE_{3,T_n(h)}\to 0.$$

Thus (2.29) holds. □

Proof of Theorem 6 and Remark 2. For $k=1,\ldots,n-1$, introduce the notation

$$a_k(t)=(t-t_k)/d_k,\quad t\in[t_k,t_{k+1}].$$

Then the linear interpolator (1.27) can be written as

$$X_{T_n}^{\ell}(t)=X(t_k)+a_k(t)\{X(t_{k+1})-X(t_k)\},\quad t\in[t_k,t_{k+1}],$$

its MSE is

$$\begin{aligned}\text{MSE}_{T_n}(t)&=\{m(t_k)-m(t)+a_k(t)[m(t_{k+1})-m(t_k)]\}^2\\&\quad+\{R(t_k,t_k)-2R(t,t_k)+R(t,t)\}\\&\quad+2\{R(t_k,t_{k+1})-R(t_k,t_k)-R(t,t_{k+1})+R(t,t_k)\}a_k(t)\\&\quad+\{R(t_{k+1},t_{k+1})-2R(t_k,t_{k+1})+R(t_k,t_k)\}a_k^2(t)\\&\triangleq A_{T_n}(t)+\sum_{j=1}^{3}B_{j,T_n}(t),\quad t\in[t_k,t_{k+1}],\end{aligned}$$

and its IMSE is

$$\text{IMSE}_{T_n}=\int_0^1 A_{T_n}(t)\,\mathrm{d}t+\sum_{j=1}^{3}\int_0^1 B_{j,T_n}(t)\,\mathrm{d}t\triangleq A_{T_n}+\sum_{j=1}^{3}B_{j,T_n}.\tag{2.31}$$

For the bias term, using (1.28), we have

$$\begin{aligned}A_{T_n}&\leqslant\sum_{k=1}^{n-1}\int_{t_k}^{t_{k+1}}\{C^2(t-t_k)^{2p}+2C^2a_k(t)(t-t_k)^p(t_{k+1}-t_k)^p\\&\qquad+C^2a_k^2(t)(t_{k+1}-t_k)^{2p}\}\,\mathrm{d}t\\&\leqslant\sum_{k=1}^{n-1}C^2\left\{d_k^{2p+1}+2d_k^{2p}\int_{t_k}^{t_{k+1}}a_k(t)\,\mathrm{d}t+d_k^{2p}\int_{t_k}^{t_{k+1}}a_k^2(t)\,\mathrm{d}t\right\}\\&=C^2\sum_{k=1}^{n-1}d_k^{2p}\{d_k+d_k+\tfrac{1}{3}d_k\}=\tfrac{7}{3}C^2\sum_{k=1}^{n-1}d_k^{2p+1},\end{aligned}$$

which yields

$$n^{2p}A_{T_n(h)} \leq \tfrac{7}{3}C^2 \frac{n^{2p}}{(n-1)^{2p}} \sum_{k=1}^{n-1} \frac{d_k}{h^{2p}(\text{int.})}$$

$$\to \tfrac{7}{3}C^2 \int_0^1 \frac{\mathrm{d}t}{h^{2p}(t)} \quad \text{as } n \to \infty.$$

Since $2p > 1$,

$$nA_{T_n(h)} \to 0 \quad \text{as } n \to \infty. \tag{2.32}$$

Moreover, if $m(t)$ is continuously differentiable, simple calculations show that

$$n^2 A_{T_n(h)} \to 0 \quad \text{as } n \to \infty. \tag{2.33}$$

We will establish Remark 2 (instead of the simpler Theorem 6).

For the terms involving the covariance function R, we Taylor-expand around the point t_k. Then, for $t \in [t_k, t_{k+1}]$, we have

$$R(t_k, t) = R(t_k, t_k) + R^{0,1}(t_k, t_k+)(t - t_k) + \tfrac{1}{2}R^{0,2}(t_k, u_k(t))(t - t_k)^2 \tag{2.34}$$

and using Taylor expansion twice and (2.34), we obtain

$$\begin{aligned} R(t, t_{k+1}) = {} & R(t_k, t_k) + R^{0,1}(t_k, t_k+)d_k + R^{0,1}(t_k, t_k-)(t - t_k) \\ & + \tfrac{1}{2}R^{0,2}(t_k, w_k)d_k^2 + R^{1,1}(t_k-, t_k+)d_k(t - t_k) \\ & + \tfrac{1}{2}R^{2,0}(v_k(t), t_{k+1})(t - t_k)^2, \end{aligned} \tag{2.35}$$

where $t_k < u_k(t)$, $v_k(t) < t$ and $t_k < w_k < t_{k+1}$ and both $u_k(t)$ and $v_k(t)$ depend continuously on t. Also, for $t \in [t_k, t_{k+1}]$,

$$R(t, t) = p(t) = p(t_k) + p^{(1)}(t_k)(t - t_k) + \tfrac{1}{2}p^{(2)}(y_k(t))(t - t_k)^2, \tag{2.36}$$

where $t_k < y_k(t) < t$ depends continuously on t and $p^{(1)}$, $p^{(2)}$ are given in (2.5) and (2.13) respectively.

In light of (2.34)-(2.36), we have

$$B_{1,T_n}(t) = \alpha(t_k)(t - t_k) + \tfrac{1}{2}\{p^{(2)}(y_k(t)) - 2R^{0,2}(t_k, u_k(t))\}(t - t_k)^2, \tag{2.37}$$

$$\begin{aligned} B_{2,T_n}(t)/[2a_k(t)] = {} & -\alpha(t_k)(t - t_k) - R^{1,1}(t_k-, t_k+)d_k(t - t_k) \\ & + \tfrac{1}{2}\{R^{0,2}(t_k, u_k(t)) - R^{2,0}(v_k(t), t_{k+1})\}(t - t_k)^2, \end{aligned} \tag{2.38}$$

$$B_{3,T_n}(t)/a_k^2(t) = \alpha(t_k)d_k + \tfrac{1}{2}\{p^{(2)}(y_k(t_{k+1})) - 2R^{0,2}(t_k, u_k(t_{k+1}))\}d_k^2. \tag{2.39}$$

If a function $g(t)$ is differentiable on [0, 1] and $g^{(1)}(t)/h(t)$ is Riemann integrable, then for regular sampling designs $T_n(h) = \{t_k\}_1^n$, we have

$$\sum_{k=1}^{n-1} g(t_k)d_k = \int_0^1 g(t)\,\mathrm{d}t - \frac{n^{-1}}{2}\int_0^1 \frac{g^{(1)}(t)}{h(t)}\,\mathrm{d}t + \mathrm{o}(n^{-1}). \tag{2.40}$$

If h is differentiable, then Taylor expansion at the point t_k gives $h(t) = h(t_k) + h^{(1)}(\text{int.})(t - t_k)$, which together with the definition of regular designs yield

$$h^{(1)}(\text{int.})d_k^2 + 2h(t_k)d_k - \frac{2}{n-1} = 0.$$

Solving this equation, we have

$$d_k = \frac{1}{h(t_k)}\frac{1}{n-1} - \frac{h^{(1)}(\text{int.})}{2h^3(t_k)}\frac{1}{(n-1)^2} + \mathrm{O}[(n-1)^{-3}]. \tag{2.41}$$

Now we derive the asymptotics for B_{j,T_n}, $j = 1, 2, 3$. Employing (2.37) and the mean value theorem of integrals, yields, for any sampling design T_n,

$$B_{1,T_n} = \sum_{k=1}^{n-1}\left\{\alpha(t_k)\int_{t_k}^{t_{k+1}}(t - t_k)\,\mathrm{d}t + \frac{1}{2}\int_{t_k}^{t_{k+1}}[p^{(2)}(y_k(t)) - 2R^{0,2}(t_k, u_k(t))](t - t_k)^2\,\mathrm{d}t\right\}$$

$$= \sum_{k=1}^{n-1}\{\tfrac{1}{2}\alpha(t_k)d_k^2 + \tfrac{1}{6}[p^{(2)}(y_k) - 2R^{0,2}(t_k, u_k)]d_k^3\}$$

and then using (2.41), yields,

$$B_{1,T_n(h)} = \tfrac{1}{2}(n-1)^{-1}\sum_{k=1}^{n-1}\frac{\alpha(t_k)}{h(t_k)}d_k - \tfrac{1}{4}(n-1)^{-2}\sum_{k=1}^{n-1}\alpha(t_k)\frac{h^{(1)}(\text{int.})}{h^3(t_k)}$$
$$+ \tfrac{1}{6}(n-1)^{-2}\sum_{k=1}^{n-1}[p^{(2)}(y_k') - 2R^{0,2}(t_k, u_k'+)]\frac{d_k}{h^2(t_k)} + \mathrm{o}(n^{-2}).$$

Finally, by Riemann integrability of the functions α/h, $\alpha h^{(1)}/h^3$, $p^{(2)}/h^2$ and $R^{0,2}(\cdot, \cdot+)/h^2(\cdot)$ and putting $g = \alpha/h$ in (2.40), we have, for large n,

$$B_{1,T_n(h)} = \tfrac{1}{2}(n-1)^{-1}\int_0^1\frac{\alpha}{h}(t)\,\mathrm{d}t - \tfrac{1}{4}(n-1)^{-2}\int_0^1\frac{1}{h(t)}\left(\frac{\alpha}{h}\right)^{(1)}(t)\,\mathrm{d}t$$
$$- \tfrac{1}{4}(n-1)^{-2}\int_0^1\alpha(t)\frac{h^{(1)}(t)}{h^3(t)}\,\mathrm{d}t$$
$$+ \tfrac{1}{6}(n-1)^{-2}\int_0^1\{p^{(2)}(t) - 2R^{0,2}(t, t+)\}\frac{\mathrm{d}t}{h^2(t)} + \mathrm{o}(n^{-2}). \tag{2.42}$$

Likewise, by (2.38) and (2.40)-(2.41), we obtain

$$B_{2,T_n(h)} = -\tfrac{2}{3}(n-1)^{-1}\int_0^1\frac{\alpha}{h}(t)\,\mathrm{d}t + \tfrac{1}{3}(n-1)^{-2}\int_0^1\frac{1}{h(t)}\left(\frac{\alpha}{h}\right)^{(1)}(t)\,\mathrm{d}t$$
$$+ \tfrac{1}{3}(n-1)^{-2}\int_0^1\alpha(t)\frac{h^{(1)}(t)}{h^3(t)}\,\mathrm{d}t - \tfrac{2}{3}(n-1)^{-2}\int_0^1 R^{1,1}(t-, t+)\frac{\mathrm{d}t}{h^2(t)}$$
$$+ \tfrac{1}{4}(n-1)^{-2}\int_0^1\{R^{0,2}(t, t+) - R^{0,2}(t, t-)\}\frac{\mathrm{d}t}{h^2(t)} + \mathrm{o}(n^{-2}) \tag{2.43}$$

and by (2.39)-(2.41),

$$B_{3,T_n(h)} = \tfrac{1}{3}(n-1)^{-1}\int_0^1 \frac{\alpha}{h}(t)\,\mathrm{d}t - \tfrac{1}{6}(n-1)^{-2}\int_0^1 \frac{1}{h(t)}\left(\frac{\alpha}{h}\right)^{(1)}(t)\,\mathrm{d}t$$

$$-\tfrac{1}{6}(n-1)^{-2}\int_0^1 \alpha(t)\frac{h^{(1)}(t)}{h^3(t)}\,\mathrm{d}t$$

$$+\tfrac{1}{6}(n-1)^{-2}\int_0^1 \{p^{(2)}(t) - 2R^{0,2}(t,t+)\}\frac{\mathrm{d}t}{h^2(t)} + \mathrm{o}(n^{-2}). \tag{2.44}$$

Therefore, by (2.42)-(2.44) and (2.13), we find

$$\sum_{j=1}^{3} B_{j,T_n(h)} = \tfrac{1}{6}(n-1)^{-1}\int_0^1 \frac{\alpha}{h}(t)\,\mathrm{d}t - \tfrac{1}{12}(n-1)^{-2}\int_0^1 \frac{1}{h(t)}\left(\frac{\alpha}{h}\right)^{(1)}(t)\,\mathrm{d}t$$

$$-\tfrac{1}{12}(n-1)^{-2}\int_0^1 \alpha(t)\frac{h^{(1)}(t)}{h^3(t)}\,\mathrm{d}t$$

$$+\tfrac{1}{12}(n-1)^{-2}\int_0^1 \{R^{0,2}(t,t-) - R^{0,2}(t,t+)\}\frac{\mathrm{d}t}{h^2(t)} + \mathrm{o}(n^{-2}),$$

$$= \tfrac{1}{6}(n-1)^{-1}\int_0^1 \frac{\alpha}{h}(t)\,\mathrm{d}t + \mathrm{o}(n^{-2}), \tag{2.45}$$

which together with (2.31)-(2.32) yield (1.29). □

Proof of (1.30)–(1.34). It is straightforward to verify that the inverse of the covariance matrix R_{T_n} is $R_{T_n}^{-1} = (a_{ij})_{n\times n}$ with

$$a_{11} = \frac{1}{2\mu}\{1/d_2 + 1/b_n\}, \quad a_{1n} = a_{n1} = 1/(2\mu b_n), \quad a_{nn} = \frac{1}{2\mu}\{1/d_n + 1/b_n\},$$

$$a_{ii} = \frac{1}{2\mu}\{1/d_i + 1/d_{i+1}\}, \quad i = 2,\dots,n-1,$$

$$a_{i,i-1} = a_{i-1,i} = -1/(\mu d_i), \quad i = 2,\dots,n,$$

$$a_{ij} = 0, \quad \text{all other } i, j,$$

where $d_i = t_i - t_{i-1}$, $i = 2,\dots,n$ and $b_n = 2/\mu - (t_n - t_1)$. This yields that for any functions f and g,

$$f'_{T_n}R_{T_n}^{-1}g_{T_n} = \frac{1}{2\mu}\sum_{i=1}^{n-1}\{f(t_{i+1}) - f(t_i)\}\{g(t_{i+1}) - g(t_i)\}/(t_{i+1} - t_i)$$

$$+\tfrac{1}{2}\{f(t_1) + f(t_n)\}\{g(t_1) + g(t_n)\}/[2 - \mu(t_n - t_1)]. \tag{2.46}$$

Letting for fixed t, $f(s) = R(t,s)$ and $g(s) = X(s)$ in (2.46), yields (1.30).

For fixed t, taking $f(s)=g(s)=R(t,s)$ in (2.46), we obtain the MSE for the BLUE $X^0_{T_n}(t)$,

$$\mathrm{MSE}^0_{T_n}(t)=\begin{cases}2\mu(t_1-t)[2-\mu(t_n-t)]/[2-\mu(t_n-t_1)], & \text{if } 0\leq t\leq t_1,\\ 2\mu(t-t_k)(t_{k+1}-t)/(t_{k+1}-t_k), & \text{if } t_k\leq t\leq t_{k+1},\ k=1,\ldots,n-1,\\ 2\mu(t-t_n)[2-\mu(t-t_1)]/[2-\mu(t_n-t_1)], & \text{if } t_n\leq t\leq 1,\end{cases} \tag{2.47}$$

which gives the IMSE with the uniform weight function $W(t)\equiv 1$,

$$\begin{aligned}\mathrm{IMSE}^0_{T_n}(1)&=1-\tfrac{1}{2}\mu(t_n-t_1)+\tfrac{1}{3}\mu\sum_{i=1}^{n-1}(t_{i+1}-t_i)^2\\&\quad-\frac{1}{3\mu}\{2-\mu(t_n-t_1)\}\{1+\mu(t_n-t_1)\}\\&\quad+\frac{2-\mu}{6\mu}\{4[\mu+(1-\mu)^2]-6\mu^2(t_1-t_n)\\&\qquad+3\mu^2(t_1+t_n)^2\}/[2-\mu(t_n-t_1)].\end{aligned} \tag{2.48}$$

To find the minimizer of $\mathrm{IMSE}_{T_n}(1)$, set $\partial\,\mathrm{IMSE}^0_{T_n}(1)/\partial t_i=0$, $i=2,\ldots,n-1$, which yield the following equations:

$$d_2=\cdots=d_n\triangleq\rho_n,\quad\text{or}\quad t_i=\tfrac{1}{2}(t_{i-1}+t_{i+1}),\quad i=2,\ldots,n-1, \tag{2.49}$$

and one-step more calculation gives $\partial^2\,\mathrm{IMSE}^0_{T_n}(1)/\partial t_i^2=4>0$. Thus, $\mathrm{IMSE}^0_{T_n}(1)$ achieves its minimum in the region $0\leq t_1<t_2<\cdots<t_n\leq 1$. Using (2.48)-(2.49), we have

$$\begin{aligned}&\mathrm{IMSE}^0_{T_n}(1)\\&=1-\tfrac{1}{2}\mu(n-1)\rho_n(1-\tfrac{2}{3}\rho_n)-\frac{1}{3\mu}\{2-\mu(n-1)\rho_n\}\{1+\mu(n-1)\rho_n\}\\&\quad\frac{2-\mu}{6\mu}\,\frac{4[\mu+(1-\mu)^2]-6\mu^2[2t_1+(n-1)\rho_n]+3\mu^2[2t_1+(n-1)\rho_n]^2}{2-\mu(n-1)\rho_n}.\end{aligned}$$

To determine t_1 and ρ_n, let $\partial\,\mathrm{IMSE}^0_{T_n}(1)/\partial t_1=0$ and $\partial\,\mathrm{IMSE}^0_{T_n}(1)/\partial\rho_n=0$, which together with (2.48) and $t_n\leq 1$ lead to $t_1=1-t_n=\tfrac{1}{2}[1-(n-1)\rho_n]$ and the following equation

$$\begin{aligned}&4n(n-1)^2\mu^2\rho_n^3-(n-1)\{3(\mu+6)n-(2+3\mu)\}\mu\rho_n^2\\&\quad+4\{3(\mu+2)n-(2+3\mu\}\rho_n-\mu^2+6\mu-24=0,\\&\rho_n\in(1/n,1/(n-1)).\end{aligned} \tag{2.50}$$

It is straightforward to verify that the second derivative of $\mathrm{IMSE}^0_{T_n}(1)$ with respect to t is positive everywhere and with respect to ρ_n is positive at the point specified by $t_1 = \frac{1}{2}[1-(n-1)\rho_n]$. The uniqueness of the solution to the eq. (2.50) is justified by the following arguments. Denote the cubic polynomial in ρ_n in (2.50) by $G(\rho_n)$. Then $G(1/n) = -(2-\mu)[2n(1-\mu)+\mu]/n^2 < 0$ and $G(1/(n-1)) = 4(\mu-2)^2 > 0$, which implies that there is a solution to the equation in $(1/n, 1/(n-1))$. Moreover, the equation $G^{(1)}(\rho_n) = 0$ has two distinct roots and the smaller one is $[3n(\mu+2)-(2+3\mu)]/[6\mu n(n-1)]$, which is greater than $1/(n-1)$. Thus, $G^{(1)}(\rho_n)$ is positive in $(1/n, 1/(n-1))$ and $G(\rho_n)$ is strictly increasing in $(1/n, 1/(n-1))$, which guarantees (2.50) has a unique solution. Thus, for any $0<\mu\leq 1$ and every sample size n (≥ 2), the optimal sampling designs are specified by (1.31)-(1.32).

The expression (1.33) follows from (2.47) and (1.31), and (1.34) follows from (2.48) and (1.31). □

Proof of (1.38)–(1.41). It is known that for any functions u and v and any sampling designs T_n,

$$u'_{T_n} R_{T_n}^{-1} v_{T_n} = \frac{u(t_1)v(t_1)}{t_1} + \sum_{k=1}^{n-1} \frac{[u(t_{k+1})-u(t_k)][v(t_{k+1})-v(t_k)]}{t_{k+1}-t_k}. \tag{2.51}$$

Taking $u(t) = v(t) = f(t) = t$ in (2.51), we obtain $f'_{T_n} R_{T_n}^{-1} f_{T_n} = t_n$, and taking for fixed t, $u(s) = R(s, t)$ and $v(s) = s$, we have $r'_{T_n}(t) R_{T_n}^{-1} f_{T_n} = \min(t_n, t)$. By using these results, (1.11) and (2.51) with $u(s) = R(s, t)$ for fixed t, $v(s) = X(s)$, we obtain (1.38). Similarly, (1.39) can be obtained, and by direct integration (1.40) follows.

To find the optimal design, set $\partial\, \mathrm{IMSE}^1_{T_n}/\partial t_k = 0$, $k = 1, \ldots, n$. This yields

$$2t_k - (t_{k-1} + t_{k+1}) = 0, \quad k = 1, \ldots, n-1, \quad \text{and} \quad 2t_n - t_n^{-2} - t_{n-1} = 0,$$

where $t_0 = 0$. One more calculation yields $\partial^2\, \mathrm{IMSE}^1_{T_n}/\partial t_k^2 = 2 > 0$, $k = 1, \ldots, n-1$ and $\partial^2\, \mathrm{IMSE}^1_{T_n}/\partial t_n^2 = 2 + t_n^{-3} > 0$, which imply that $\mathrm{IMSE}^1_{T_n}$ achieves its minimum in the region $0 \leq t_0 < t_1 < \cdots < t_n \leq 1$ and the minimizer of $\mathrm{IMSE}^1_{T_n}$ is given by the solution to the above equations. By solving these equations, we obtain (1.41). □

Proof of (1.42)–(1.46). Here, it is known (see Anderson, 1965) that for any functions u and v and any sampling design $T_n = \{t_i\}_1^n$,

$$\begin{aligned} u'_{T_n} R_{T_n}^{-1} v_{T_n} = {} & \frac{u(t_1)v(t_1)}{1-\mathrm{e}^{-2(t_2-t_1)}} + \frac{u(t_n)v(t_n)}{1-\mathrm{e}^{-2(t_n-t_{n-1})}} \\ & + \sum_{i=2}^{n-1} \frac{u(t_i)v(t_i)[1-\mathrm{e}^{-2(t_{i+1}-t_{i-1})}]}{[1-\mathrm{e}^{-2(t_i-t_{i-1})}][1-\mathrm{e}^{-2(t_{i+1}-t_i)}]} \\ & - \sum_{i=1}^{n-1} \frac{[u(t_i)v(t_{i+1}) + u(t_{i+1})v(t_i)]\, \mathrm{e}^{(t_{i+1}-t_i)}}{1-\mathrm{e}^{-2(t_{i+1}-t_i)}}. \end{aligned} \tag{2.52}$$

For fixed t, putting $u(s) = R(s, t)$ and $v(s) = X(s)$ in (2.52) and through some straightforward calculations, $X^0_{T_n(h_0)}(t)$ can be simplified as in (1.42).

By direct calculations, the MSE of $X^0_{T_n(h_0)}(t)$ is

$$\mathrm{MSE}^0_{T_n(h_0)}(t) = \frac{e^{2/(n-1)} + 1 - e^{2/(n-1)}\left(e^{-2(t_{k+1}-t)} + e^{-2(t-t_k)}\right)}{e^{2/(n-1)} - 1} + \beta^2 \left\{ \frac{e^{t_{k+1}-t} - e^{-(t_{k+1}-t)} + e^{t-t_k} - e^{-(t-t_k)}}{e^{1/(n-1)} - e^{-1/(n-1)}} - 1 \right\}^2, \tag{2.53}$$

if $t_k \leqslant t \leqslant t_{k+1}$, which yields (1.43).

By (2.52), we have for $f(t) \equiv 1$,

$$r'_{T_n}(t) R^{-1}_{T_n} f_{T_n} - 1 = \frac{e^{t_{k+1}-t} - e^{-(t_{k+1}-t)} + e^{t-t_k} - e^{-(t-t_k)}}{e^{1/(n-1)} - e^{-1/(n-1)}} - 1, \tag{2.54}$$

and putting $u(t) = v(t) = f(t) \equiv 1$, in (2.52), yields

$$f'_{T_n} R^{-1}_{T_n} f_{T_n} = \{2 + (n-2)(1 - e^{-1/(n-1)})\}/(1 + e^{-1/(n-1)}). \tag{2.55}$$

Putting $u(t) = f(t) \equiv 1$ and $v(t) = X(t)$ in (2.52), gives

$$f'_{T_n} R^{-1}_{T_n} X_{T_n} = \frac{X(t_1) + (1 - e^{-1/(n-1)}) \sum_{i=2}^{n-1} X(t_i) + X(t_n)}{1 + e^{-1/(n-1)}}. \tag{2.56}$$

Then, (1.44) follows from (1.11) and (2.54)-(2.56). By using (1.44) and through some straightforward calculations, (1.45) follows.

Through direct calculation, we obtain for $k = 1, \ldots, n-1$ and $t \in [t_k, t_{k+1}]$,

$$\tfrac{1}{2}\mathrm{MSE}^{\ell}_{T_n}(t) = \left(\frac{t_{k+1}-t}{t_{k+1}-t_k}\right)^2 (1 - e^{-(t-t_k)}) + \left(\frac{t-t_k}{t_{k+1}-t_k}\right)^2 (1 - e^{-(t_{k+1}-t)}) + \left(\frac{t_{k+1}-t}{t_{k+1}-t_k}\right)\left(\frac{t-t_k}{t_{k+1}-t_k}\right)(1 - e^{-(t-t_k)} - e^{-(t_{k+1}-t)} + e^{-(t_{k+1}-t_k)}),$$

which yields (1.46). □

Proof of (1.47). It can be verified that for any functions u and v and any sampling designs $T_n = \{t_k\}_1^n$ with $t_1 > 0$,

$$u'_{T_n} R^{-1}_{T_n} v_{T_n} = \frac{u(t_1)v(t_1)}{t_1^{\beta+1}} + \sum_{k=1}^{n-1} \frac{\{u(t_{k+1}) - u(t_k)\}\{v(t_{k+1}) - v(t_k)\}}{t_{k+1}^{\beta+1} - t_k^{\beta+1}}.$$

Then, by taking for fixed t, $u(s) = R(s, t)$ and $v(s) = X(s)$, we obtain (1.47). □

References

T.W. Anderson, Some stochastic process models for intelligence test scores, in: K.J. Arrow et al., eds., Mathematical Methods in the Social Sciences (Stanford Univ. Press, Stanford, CA, 1960) pp. 205-220.

S. Cambanis, Sampling designs for time series, in: E.J. Hannan et al., eds., Handbook of Statistics, Vol. 5: Time Series in Time Domain (North-Holland, Amsterdam, 1985) pp. 337-362.

G. Christakos, Random Fields Models in Earth Sciences (Academic Press, San Diego, CA, 1992).

D.L. Hawkins, Some practical problems in implementing a certain sieve estimation of the Gaussian mean function, Comm. Statist.—Simulation Comput. 18 (1989) 481-500.

A.G. Journel and C.J. Huijbregts, Mining Geostatistics (Academic Press, New York, 1978).

A.G. Journel and M.E. Rossi, When do we need a trend model in kriging?, Math. Geology 21 (1989) 715-739.

T. Kailath, Some integral equations with "nonrational" kernels, IEEE Trans. Inform. Theory IT-12 (1966) 442-447.

J. Sacks, W.J. Welch, T.J. Mitchell and H.P. Wynn, Design and analysis of computer experiments, Statist. Sci. 4 (1989) 409-435.

J. Sacks and D. Ylvisaker, Designs for regression problems with correlated errors, Ann. Math. Statist. 37 (1966) 66-89.

M.L. Stein, Asymptotic distributions of minimum norm quadratic estimators of the covariance function of a Gaussian random field, Ann. Statist. 17 (1989) 980-1000.