Uncl

SECURI

AD-A205 840

DTIC ELECTE MAR 08 1989 S H D

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: U	1b. RESTRICTIVE MARKINGS: None
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 89-0256
6a. NAME OF PERFORMING ORGANIZATION: Dr. Jay H. Beder, Uni. of Wisconsin-Milwaukee; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: Mathematical and Information Sciences, Air Force Office of Scientific Research
6c. ADDRESS (City, State and ZIP Code): Department of Mathematical Sciences, P. O. Box 413, Milwaukee, WI 53201	7b. ADDRESS (City, State and ZIP Code): Bld 410, Bolling Air Force Base, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR; 8b. OFFICE SYMBOL (If applicable): NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR 84-0329
8c. ADDRESS (City, State and ZIP Code): Bldg. 410, Bolling AFB, DC 20332-6448	10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
61102F	2304	A5	

11. TITLE (Include Security Classification) Sieves, Signal Extraction, and Design (Unclassified)

12. PERSONAL AUTHOR(S) Beder, Jay H.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Final Technical Report	FROM 9-30-84 TO 8-31-88	1-4-89	5

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Confounding, consistency, factorial experiment, Gaussian process, Hadamard matrix, reproducing kernel Hilbert space, sieve, simulation, stochastic signal, zero-one law. (LD)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This report reviews the research completed under this grant. The areas covered include: (a) sieve estimation for the mean and the covariance of a Gaussian process; (b) stochastic signal extraction and a zero-one law of M. Driscoll; and (c) the problem of confounding in factorial experiments. Parts (a) and (b) have been conducted without any assumptions whatever on the "time" parameter underlying the process. Work on part (c) has similarly made no a priori assumption about the algebraic or geometric structure of the set of levels of each factor; in certain cases, the problem of confounding is shown to be related to the Hadamard matrix problem.

DISTRIBUTION STATEMENT A
Approved for public release;

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION: Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL: Maj. Brian Woodruff	22b. TELEPHONE NUMBER (Include Area Code): 202-767-5027; 22c. OFFICE SYMBOL: NM

DD FORM 1473, 83 APR — EDITION OF 1 JAN 73 IS OBSOLETE.

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

AFOSR-TR- [illegible] 56

Sieves, Signal Extraction, and Design (Unclassified)

1. Summary: See DD 1473, item 19.

2,3. Research objectives, status of work: In the description that follows, citations indicated by [1], [2], etc., refer to the list given in item (4) below.

a. Sieve estimation for Gaussian processes.

Let $X = \{X_t, t \in T\}$ be a Gaussian process with mean function m and covariance function R. The PI has developed sieve estimators $\hat{m}$ and $\hat{R}$ to estimate m when R is known, and R when m is known. The construction and properties of these estimators allow T to be an arbitrary set. Thus this method has very wide applicability.

In the case of estimating m, the only special assumption needed is that m belong to a known countably-dimensioned subspace of the reproducing kernel Hilbert space (RKHS) $\mathcal{H}(R)$. For the estimation of R when m is known, an additional assumption is needed concerning the bivariate expansion of R. In either case the data consists of n replicates (trajectories) of the process. A discussion of this as a design constraint is given in [7], an extended version of [4].

In both cases, the estimators are strongly consistent in a natural norm as long as a sieve parameter d (roughly, the number of terms in the estimator) grows as o(n), and are consistent in probability and in mean square locally if d grows as O(n). They are also asymptotically unbiased as d goes to infinity. Details are given in papers [2] and [4]. Paper [1] shows how the two estimators can be combined, and applies the method in the Banach space setting.

There is a good deal of latitude in the construction of these estimators, depending as they do on an arbitrary choice of "basis" functions in an appropriate Hilbert space. Using this and elementary tools (essentially Gram-Schmidt orthogonalization), it is easy to construct an estimator of the mean which is computationally simple and which requires only d time "sites" $t_1, \ldots, t_d$ for observing the process. Computer simulations (still underway) show that this method allows one to recover significant features of the underlying mean function (the "trend" or "signal") easily. (Theory requires that d be limited by n, the number of replicates observed; greater accuracy can be attained by using more replicates. The simulation study so far indicates that d should be about $n^{.9}$.)

b. On a zero-one law of Driscoll, and optimal signal extraction.

We are given an observable signal $X_t = S_t + N_t$, $t \in T$, where S and N are independent zero-mean Gaussian processes (signal and noise) with known covariances. Under certain conditions, Driscoll [1975] has shown that an optimal method to extract the signal S from X is to use the process $\hat{S}$ defined by $\hat{S}_t = E(S_t | \mathcal{A}_X)$, where $\mathcal{A}_X$ is the sigma-algebra generated by $\{X_t, t \in T\}$. These conditions concern the topology of T and the continuity of the covariances.

The PI has generalized this framework by letting T be an arbitrary set and by not assuming any particular relationship between X and S, so that the stochastic signal can be corrupted by noise in a more arbitrary fashion. Optimality can still be defined as in Driscoll by requiring that the sample paths $S_\cdot$ and $\hat{S}_\cdot$ belong to a RKHS $\mathcal{H}(R)$ for and appropriate positive definite kernel R. Closeness is now defined by $\|S_\cdot - \hat{S}_\cdot\|$, using the norm of $\mathcal{H}(R)$. To do this, we take R to dominate the covariance K of S in the sense that $\mathcal{H}(K) \subset \mathcal{H}(R)$; we denote this by $K << R$. If R n-dominates K (Fortet [1974]), we write $K \overset{n}{<<} R$. The plan is to show:

i. If $K << R$ then $P(S_\cdot \in \mathcal{H}(R)) = 0$ or 1, and is 1 iff $K \overset{n}{<<} R$;

ii. if $K \overset{n}{<<} R$ then $P(\hat{S}_\cdot \in \mathcal{H}(R)) = 1$; and

iii. $E\|S_\cdot - \hat{S}_\cdot\|^2 \le E\|S_\cdot - h\|^2$ a.s. for all $h \in \mathcal{H}(R)$; and to give the value for the left-hand side of (*).

Again, in all cases, no assumptions are made on T.

The "zero" part of item (i) has been proved in full generality, while the "one" part and item (ii) have been proved under a separability assumption (see the AFOSR technical report Beder [1986]) that may not be removable. The value of the left-hand side of (*) has been derived as the trace of an operator (Beder [1986]), but the inequality remains to be proved.

This work also has application to certain stochastic design problems; see Ylvisaker [1987].

c. <u>Experimental Design</u>.

The PI has given a definitive analysis [5] of confounding patterns in two-factor experiments in which one factor has a prime number of levels. The existence of a generalized Hadamard matrix implies the existence of an orthogonal decomposition of interaction of a certain type. Moreover, the converse holds at least in the case that the prime (number of levels) is 2 or 3. When the prime is 2, then, the problem of confounding is equivalent to the classical Hadamard matrix conjecture. The method used makes no assumption a

A-1

priori about the algebraic or geometric structure of the set of levels attached to each factor in the experiment. Background to the method is given in [8].

It is of particular interest that the problem of confounding in the simplest two-factor asymmetric design is equivalent to a difficult combinatorial conjecture, and thus far deeper than the problem in classical symmetric factorials. This depth is not revealed by standard methods of analyzing confounding.

The classical Hadamard matrix problem may be viewed as that of constructing a maximal set of nx1 vectors consisting of 1's and -1's. Such a maximal set can of course have no more than n elements. It is known that if a maximal set of size n exists, then n = 2 or n is a multiple of 4; the conjecture is the converse, namely that if n is a multiple of 4 then there exists a maximal set of size n. The PI has shown that if n = 2 (mod 4) then all maximal sets are of size 2. Also, a computer search routine written by the PI indicates that if n is a multiple of 4 then the size of any maximal set is also a multiple of 4. The empirical evidence from this search is for n = 4, 8, ..., 24, values for which the conjecture is known. These results involve maximal orthogonal sets of possibly non-maximum cardinality, a new approach. Whether this will shed further light on the conjecture remains to be seen.

d. Miscellaneous.

Paper [3] and the technical report [6] were also completed under this grant.

e. References.

Beder, J. H. (1986). Sieves and signal extraction. AFOSR Annual Technical Report.

Driscoll, M. F. (1975). The signal-noise problem - a solution for the case that signal and noise are Gaussian and independent. J. Appl. Prob. 12: 183-187.

Fortet, R. (1974). Espaces à noyau reproduisant et lois de probabilité de fonctions aléatoires. C. R. Acad. Sci. Paris, A, t.278: 1439-1440.

Ylvisaker, D. (1987). Prediction and design. Ann. Statist. 15: 1-19.4.

4. Written Publications:

[1] Antoniadis, A., and Beder, J. H. (1989). Joint estimation of the mean and the covariance of a Banach valued Gaussian vector. Statistics, 20, no. 1 (to appear).

[2] Beder, J. H. (1987). A sieve estimator for the mean of a Gaussian process. Ann. Statist. 15: 59-78.

[3] Beder, J. H. (1988). Estimating a covariance function having an unknown scale parameter. Commun. Statist. Theory and Methods, A17: 323-340.

[4] Beder, J. H. (1988). A sieve estimator for the covariance of a Gaussian process. Ann. Statist. 16: 648-660.

[5] Beder, J. H. (1989). The problem of confounding in two-factor experiments. Commun. Statist. Theory and Methods, A18, no.2 (to appear).

Technical Reports:

[6] Beder, J. H. (1987). Simultaneous diagonalization of two covariance kernels.

[7] Beder, J. H. (1987). A sieve estimator for the covariance of a Gaussian process: Theory and background.

[8] Beder, J. H., and O'Laughlin, M. (1985). The cell-means interpretation of confounding.

5. Professional Personnel:

Ms. Julie Letellier, a graduate student at UW-Milwaukee, assisted with simulation studies during the summer of 1987.

6. Interactions:

a. Papers presented:

(i) "A new sieve estimator for a Gaussian mean," Colloquium, Dept. of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, April 3, 1985.

(ii) "Sieve estimation for Gaussian processes," Joint Statistical Meetings, Las Vegas, NV, August 7, 1985. (IMS Abs. 193-50).

(iii) "Sieve estimation for Gaussian processes," International Symposium on Foundations of Statistical Inference, Tel Aviv, December 15-19, 1985.

(iv) "Sieve estimation for Gaussian processes," Colloquium, Dept. of Mathematics, University of California, Irvine, June 10, 1986.

(v) "On the zero-one laws of Kallianpur and Driscoll," Colloquium, Dept. of Mathematics, U.C. Irvine, March 4, 1987.

(vi) "Sieve estimation for Gaussian processes," Colloquium, Div. of Statistics, U.C. Davis, March 5, 1987.

(vii) "On the zero-one laws of Kallianpur and Driscoll," IMS Regional Meeting, Dallas, TX, March 24, 1987. (IMS Abs. 199-17).

(viii) "Inference for Gaussian processes," Dept. of Statistics, University of Chicago, May 16, 1988.

(ix) "A sieve estimator for the mean of a Gaussian process," International Symposium on Forecasting, Amsterdam, June 13, 1988.

(x) "Confounding and the Hadamard matrix problem," Centennial Meeting of Amer. Math. Soc., Providence, RI, Aug. 9, 1988 (AMS Abs. 844-62-113).

b. Consultative and advisory functions: Joint research with A. Antoniadis on sieve estimation. Dept. of Mathematics, U.C.-Irvine, June 9-11, 1986.

7. New discoveries, inventions, patents: none.