ORGIA INSTITUTE OF TECHNOLOGY   OFFICE OF CONTRACT ADMINISTRATION

PROJECT ADMINISTRATION DATA SHEET

[X] ORIGINAL

[ ] REVISION NO. ____

roject No. G-37-630   DATE: 6-26-81

roject Director: M. C. Spruill   School/Lab Mathematics

ponsor: National Science Foundation
Washington, D.C. 20550

ype Agreement: Grant No. MCS 8103444

ward Period: From 7-1-81 To 12-31-83 (Performance) 3-31-84 (Reports)

ponsor Amount: $ 21,852   Contracted through:

ost Sharing: $ 6,624   GTRI/GIT

itle: Optimal Experimental Designs

DMINISTRATIVE DATA   OCA CONTACT Don Hasty

) Sponsor Technical Contact: David S. Moore
(202) 357-9764

) Sponsor Admin./Contractual Contact: Philip M. King
(202) 357-9671

eports: See Deliverable Schedule   Security Classification: N/A

efense Priority Rating: N/A

ESTRICTIONS

ee Attached NSF Supplemental Information Sheet for Additional Requirements

ravel: Foreign travel must have prior approval - Contact OCA in each case. Domestic travel requires sponsor approval where total will exceed greater of $500 or 125% of approved proposal budget category.

quipment: Title vests with N/A - None Proposed

OMMENTS:

OPIES TO:

dministrative Coordinator
esearch Property Management
ccounting Office
rocurement/EES Supply Services
Research Security Services
Reports Coordinator (OCA)
Legal Services (OCA)
Library, Technical Reports
EES Research Public Relations
Project File (OCA)
Other: ____

GEORGIA INSTITUTE OF TECHNOLOGY　　OFFICE OF CONTRACT ADMINISTRATION

SPONSORED PROJECT TERMINATION/CLOSEOUT SHEET

Date April 14, 1986

Project No. G-37-630　　School/~~XXX~~ Math

Includes Subproject No.(s) ____

Project Director(s) M. C. Spruill　　GTRC / ~~XXX~~

Sponsor National Science Foundation

Title Optimal Experimental Designs

Effective Completion Date: 12/31/83 (Performance) 3/31/83 (Reports)

Grant/Contract Closeout Actions Remaining:

- [ ] None
- [ ] Final Invoice or Final Fiscal Report
- [ ] Closing Documents
- [X] Final Report of Inventions　Patent Questionnaire
- [ ] Govt. Property Inventory & Related Certificate
- [ ] Classified Material Certificate
- [ ] Other ____

Continues Project No. ____　　Continued by Project No. ____

COPIES TO:

Project Director
Research Administrative Network
Research Property Management
Accounting
Procurement/EES Supply Services
Research Security Services
Reports Coordinator (OCA)
Legal Services

Library
GTRC
Research Communications (2)
Project File
Other Heyser
Jones
Embry

Form OCA 60:1028

I. Summary of Progress to Date.

Several areas of research were suggested in the proposal for Grant No. MCS81-03444. Much of our effort was devoted to robust optimal designs for the estimation of a linear functional. We employed a robust linear estimator due to Speckman [65] and utilized the similarities between the form of its maximum mean square error and that of the usual best linear unbiased estimator to provide a basic framework of theorems which enable the calculation of optimal designs. These theorems were successfully applied to the following problem.

Consider the problem of the extrapolation of a function to a point outside the interval on which observations may be taken. We suppose that for every collection of points $(x_1,\ldots,x_n)$ in $[a,b]$ we may observe with error the values of an unknown function θ and that we are to provide on the basis of observed values $(y(x_1),\ldots,y(x_N))$, $y(x_j) = \theta(x_j) + \epsilon_j$, $E(\epsilon_j) = 0$, $E(\epsilon_i\epsilon_j) = \delta_{ij}$, an estimate of the value $\theta(c)$ where $c > b$. We prove that if we allow for an error in our assumed form

$$\theta(x) = \beta_0 + \beta_1 x + \ldots + \beta_{m-1}x^{m-1}$$

of the mean function of the type $\int_a^c (\theta^{(m)}(t))^2 dt \le \epsilon^2$ then the optimal design concentrates on m points in [a,b] as it would if no contamination were present, the weights at the points are calculated from the same formula as for Hoel-Levine [30] designs, and the optimal linear estimator is the usual least squares estimator. In addition, for the case $m=2$ we show that when $N\epsilon^2$ is sufficiently small the optimal design coincides with the Hoel-Levine design. Thus the Hoel-Levine design is robust. "Sufficiently small" appears to be reasonably large indicating that in many cases the Hoel-Levine design can be used without fear that reasonable deviations from the assumed model will give large mean square errors. We also have obtained preliminary numerical evidence of this behavior for $m > 2$.

The applicability of our theorems does not end here. We can with time and effort utilize them in establishing optimal designs for the estimation of derivatives at points. The course is clear though the specific details are probably, as in the extrapolation problem, quite difficult. For example, if our method of proof is to carry over with minimal changes then we would be required to understand the oscillation properties of the minimax approximant to our arbitrary continuous function from the set of functions spanned by $\{1, x^2, x^3, \ldots, x^{m-1}\}$ on an interval containing zero. This is not even a weak Chebychev system, see [32], so the properties of the error are, as far as I know, unknown. We can answer these questions at present for an endpoint but the details of finding the design are far from complete. The theorems are not model specific so other forms of mean functions and contamination can be studied. The theorems also apply to data which are stochastic process valued. In addition the proof of the extrapolation results uncovered a very mathematically interesting set of functions which are splines possessing Chebychev-like oscillation properties. Finally, when $m=2$ these functions are shown to be the trajectories of the solutions to the following non-standard control problems.

A particle of unit mass travels in the x-y plane. Its coordinates at time t are $(x(t), y(t))$ where $x'(t) \equiv -1$, and $(x(0), y(0)) = (c, 1)$. The y coordinate may be controlled. For those controls which have the property that the particle hits the line $x=a$, $c>a$, the energy expended is $\int_0^{c-a} (y''(t))^2 dt$. Denote the corridor $\{(x,y): x \in [a,b],\ |y| \le \gamma\}$ by $C(\gamma)$, where $b \in (a,c)$. Given that $E \le E_0$ is to be spent, what is the minimum $\gamma \ge 0$ for which the particle may be made to pass entirely through the corridor $C(\gamma)$ and its trajectory?

Although there are many papers on robust optimal designs, for example [13],[28],[31],[37],[39],[43],[44],[46],[47],[51],[60],[74],[76],[78], the model we employed and the criterion of interest do not match any of them. The closest is Huber's paper [31] which treats minimax extrapolation on a half-line with contamination measured in the sup norm rather than the L_2 norm. The results appear to be somewhat different although they share the property that both are supported on the same number of points. We emphasize and that as far as we know are the only ones which potentially have a direct bearing on showing that the usual Hoel-Levine designs are robust.

II. Proposed Research

For several years the principal investigator has been working toward the accumulation of a body of facts which would be useful to experimenters whose data are "time recordings", or more properly speaking, second order stochastic processes. Our intent is to continue our investigations in the optimal design of experiments focusing on this aspect. However, much of our effort will be directed toward obtaining solutions to problems involving scalar observatins. Of immediate interest is the set up described in the summary of our progress to date on MCS-8103444. In that problem the mean function θ is assumed to be a member of the Sobolev space $W_m^2[a,c]$ of functions with m-1 absolutely continuous derivatives on [a,c] and the nth in $L_2[a,c]$. Speckman's linear estimator $\ell_0'y$ minimizes the maximum mean square error

$$\sup_{\|\theta^{(m)}\|_2 \le \epsilon} E_\theta(\ell'y - \theta(c))^2.$$

It was shown that that in contrast to the case of best linear unbiased estimation one must specify beforehand the number of observations to be employed. Setting $\eta^{-1} = N\epsilon^2$ we prove that the optimal design can be characterized as follows. Define for each collection of points $a \le t_1 < t_2 < \ldots < t_m \le b$ the functions on [a,c].

$$(1)\quad \delta(x) = (s^2+\eta z^2)^{-1}[\eta z \sum_{i=1}^{m} (-1)^{i-m} \varphi_{t_i}(x) + \int_a^c h_x(s)h_c(s)ds],$$

where $h_x(s) = \frac{(x-s)_+^{m-1}}{(m-1)!} - \sum_{i=1}^{m} \varphi_{t_i}(x) \frac{(t_i-s)^{m-1}}{(m-1)!}$,