COMMUN. STATIST.—SIMULA., 21(1), 243-267 (1992)

SPATIAL DESIGNS WHEN THE OBSERVATIONS ARE CORRELATED

Mark F. Schilling

Department of Mathematics
California State University
Northridge, CA 91330

Key Words and Phrases: *simulated annealing*; *experimental design*

ABSTRACT

Suppose n observations are to be taken within a compact region, where the objective is to estimate the mean level of a multidimensional stationary process using the ordinary sample mean as the estimator. Simulated annealing is used to search for optimal (variance minimizing) designs for the case when the observations are correlated. The results give insight into the sensitivity of optimal designs to the strength and nature of the correlation present, extending and reinforcing previous results for the one-dimensional case. An important outcome is that designs which space out the sampling locations evenly are optimal if the correlation is low.

1. INTRODUCTION

Consider the problem of estimating a quantity whose level is unknown but assumed to be constant over a given region, by means of taking observa-

tions at several sites within the region. Frequently, the observations at various sites are dependent, with correlation being a function of distance ("isotropic"), decreasing but probably not well known otherwise. Examples may include test drilling to measure either the mean amount of recoverable oil per unit area in an oil field or the ore content in a mineral deposit, or sampling trace amounts of a chemical pollutant in a body of water. Similar situations arise in geography, ecology (e.g., transect sampling), economics, and studies of turbulence.

Let the values of the quantity of interest be represented by

$$Z_i = Z(x_{1i}, x_{2i}, \cdots, x_{pi}) = \mu + \varepsilon_i, \quad i = 1, \cdots, n,$$

where the $(x_{1i}, \cdots, x_{pi})$'s are the sampling locations and the errors ε_i are normally distributed with $E(\varepsilon_i) = 0$, $Var(\varepsilon_i) = \sigma^2$ and $Corr(\varepsilon_i, \varepsilon_j) = \gamma\rho(d_{ij})$ for $0 < \gamma < 1$, where $d_{ij} = [(x_{1i}-x_{1j})^2 + \cdots + (x_{pi}-x_{pj})^2)]^{1/2}$. Assume also that $\rho(0) = 1$ and that $\rho(d_{ij})$ decreases to zero as $d_{ij} \to \infty$. This paper concentrates on the case when each coordinate is constrained to the interval $[-1, 1]$, so the sampling region will be either a square, cube or hypercube. Multiple observations are allowed at the same position ("repeated measures" designs); when $\gamma < 1$ these measurements can give different results.

When the correlation is zero, the location of the sampling points is irrelevant, and the best linear unbiased estimator of μ is the ordinary sample mean $\bar{Z}$. If correlation is present but ρ is not known precisely, the b.l.u.e. cannot be found. However, $\bar{Z}$ still may be a quite efficient estimator, especially if the correlation is weak. It is reasonable, therefore, to study the nature of sampling designs that are optimal for estimating the process mean value μ with the sample mean, giving particular attention to the nature of these designs as the correlation is lessened.

A major motivation for analyzing the behavior of such designs as the intrapoint correlation is lowered is that, in practice, the taking of samples often begins in a relatively small region but then progressively expands to adjacent areas. Such a sampling scheme, with fixed correlation determined by physical

properties, can be translated into an alternate framework in which the region is fixed and the correlation at a given distance decreases with n simply by rescaling the region down to a standardized region as new observations are taken.

The results obtained below extend to multiple dimensions the work of Bickel and Herzberg (BH, 1979) and Bickel, Herzberg and Schilling (BHS, 1981), who found asymptotic and exact finite one-dimensional designs for estimating the mean of a stationary first order autoregressive process plus an independent error. Evidence is presented that indicates that designs in which the observation points are spaced out evenly over the design region are again favored when the correlation is low, as was found for one dimension in BH and BHS.

To determine an optimal design for a given model structure and choice of parameter values, it is necessary to minimize the design variance

$$\mathrm{Var}(\bar{Z}) = \left[\frac{1}{n} + \frac{2\gamma}{n^2}\sum_{i=2}^{n}\sum_{j=1}^{i-1}\rho(d_{ij})\right]\sigma^2 .$$

The value of γ clearly affects the variance but does not influence the optimal design itself. Finding variance minimizing solutions to the general design problem analytically in more than one dimension appears to be an intractable problem. The recently developed numerical method known as simulated annealing, however, provides an effective means of finding optimal or very nearly optimal solutions.

While the scope of the numerical studies reported here obviously cannot address all of the possible variations allowable by the model given above, an attempt is made to assess the sensitivity of the optimal design to the particular correlation structure as well as to the shape of the design region. Two specific correlation functions are studied: $\rho(d) = e^{-\lambda d}$, which is the correlation function used in BH and BHS, and $\rho(d) = e^{-\lambda d^2}$. The first of these is convex on $0 < d < \infty$ but not analytic, while the second is analytic but not convex.

2. THE ANNEALING ALGORITHM

The simulated annealing algorithm is a simple procedure useful for searching for the minimal value of an objective function V that is not convex, or at least is not known to be convex. Given an initial configuration, successive configurations are generated as a random walk over the configuration space. A configuration is accepted with probability $p = 1$ if it lowers or maintains the value of the objective function and with probability $p < 1$ if it raises the value of the objective function, where $p = e^{-\Delta V/t}$; ΔV is the increase in the value of the objective function from the old configuration to the candidate for the new configuration, and t is a parameter known as temperature because of its analogous role to ordinary kinetic temperature in actual chemical annealing. The value of t is reduced slowly, making the rule for acceptance of configurations that increase V progressively more stringent as the process evolves. Detailed information on simulated annealing can be found in Kirpatrick, Gelatt, and Vecchi (1983), Bohachevsky, Johnson and Stein (1986), and van Laarhoven and Arts (1987).

Simulated annealing has recently begun to be used in problems of statistical design (Haines (1987), Meyer and Nachtsheim (1988), Sacks and Schiller (1988)). For the present application, the objective function is $V = \mathrm{Var}(\bar{Z})$ and the configuration space is pn-dimensional, with each coordinate x_{ji} constrained to the interval $[-1, 1]$. Since this is not a problem of combinatorial optimization but rather one involving continuous variables, the specification of new configurations according to some random walk is more open than in most combinatorial problems. One possibility is to use a local measure of the topology of the objective function at each stage, to cause the random walk to favor more promising directions over less promising ones; Vanderbilt and Louie (1984) developed such a procedure which, after the beginning stages, is based on a covariance matrix of the accepted steps comprising the recent history of the random walk. That approach is not employed here due to (1) the considerable amount of computation required for each step, and (2) concern that bias-

ing the random walk by shape information would increase the risk of trapping design configurations in local minima. Instead, only one point is moved on each iteration and the step distribution is taken to be uniform in direction, except for influences due to the design region's boundary. Keeping the point in bounds requires special handling. Using a technique due to Box (1966), the new coordinates $(x_{ji}', j = 1, \cdots, p)$ of the point moved are kept within $[-1, 1]$ by the transformation

$$x_{ji}' = \sin(\arcsin(x_{ji}) + c\,\delta\,u_j)\,, \qquad j = 1, \cdots, p\,,$$

where $\mathbf{u} = (u_1, \ldots, u_p)$ is a random unit vector whose direction is uniformly distributed, the random scalar c is chosen according to a standard Cauchy distribution, and δ is a parameter which decreases proportionally with temperature. When δ becomes small, the distance that the i-th point moves is thus approximately $c\,\delta$ when the point is not near the boundary; near the boundary the steps tend to be shorter. The use of a Cauchy step size distribution was suggested by Szu and Hartley (1987), and outperforms finite variance step distributions in this application -- perhaps due to the fact that the occasional very large steps generated allow a more thorough search of the configuration space, with less trapping near suboptimal configurations.

A homogeneous version of the annealing algorithm is used, that is, many iterations are computed at each temperature value (see van Laarhoven and Aarts (1987)). When no improvement occurs in 250 consecutive iterations, both temperature and the step size parameter are decreased at a rate inversely proportional to the number of stages used in the process so far. The effect is that the probability of the acceptance of detrimental steps is adaptively determined. This approach was used successfully by Bohachevsky, Johnson and Stein (1986) to speed up the annealing algorithm in its later stages. The algorithm eventually terminates when a sufficient number of stages passes with no further lowering of the variance at four decimal places.

Initial configurations were chosen randomly according to a uniform measure. In several cases, different random configurations or even nonrandom, pat-

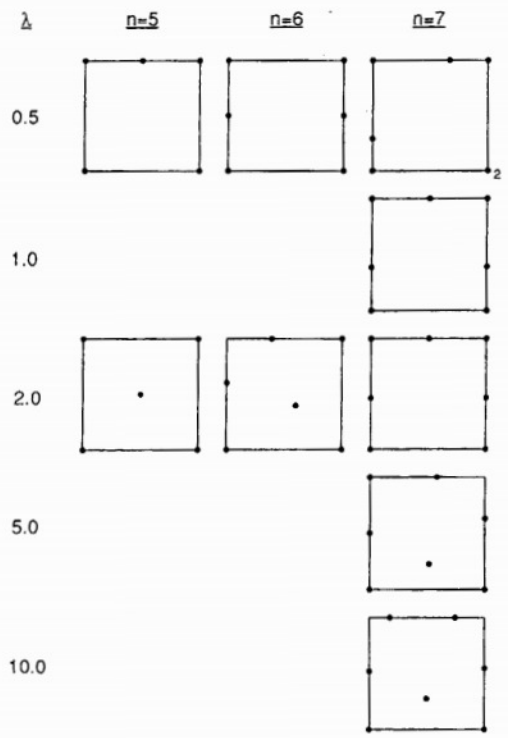

FIG. Ia. Optimal Designs for $\rho = e^{-\lambda d}$, n = 5, 6, 7.

terned configurations were tried as the initial design. In most cases the initial arrangement had no effect on the design found by the algorithm except for rotation and reflection; in no case was a design obtained which was significantly inferior to another design found on a different run with the same parameter values.

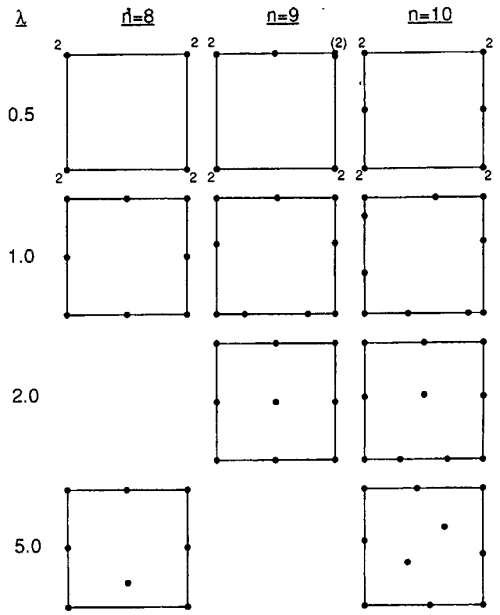

FIG. Ib. Optimal Designs for $\rho = e^{-\lambda d}$, n = 8, 9, 10.

3. RESULTS FOR TWO DIMENSIONS

The primary focus of the investigation is to determine the nature of two-dimensional sampling plans for a square sampling region. The sample sizes chosen for study were n = 5, 6, 7, 8, 9, 10, 16, 20, 25 and 36. The reasons for using these values was (1) to choose a consecutive sequence of small values to learn the effects of adding a single new point to a design, (2) to use perfect squares in order to allow comparison to square lattice designs, and (3) to select several multiples of four to take advantage of the four-way symmetry of the design region. The values selected for the correlation parameter λ range from 0.1 to 10, which covers a wide range of correlations varying from quite

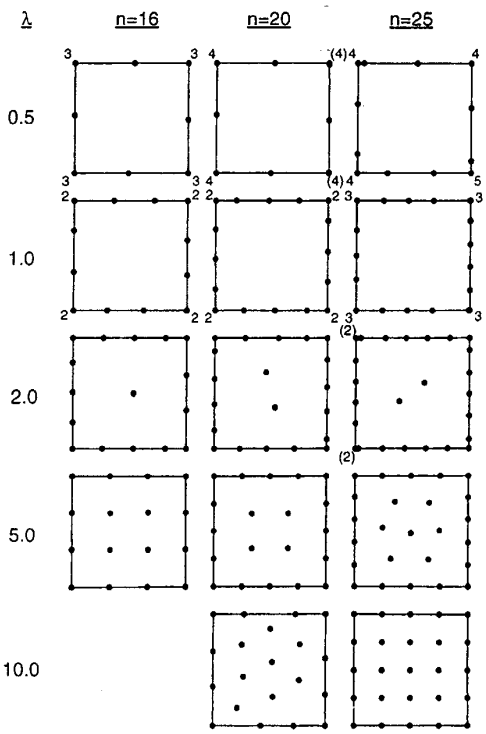

FIG. Ic. Optimal Designs for $\rho = e^{-\lambda d}$, n = 16, 20, 25.

high to almost no correlation at typical interpoint distances. The parameter γ was not varied because its value affects only the variance of the design but not the design itself.

Figures 1 and 2 display most of the designs found by the annealing algorithm. Although there is no guarantee that each design shown is the best possible, it is highly likely that the great majority of the designs given are indeed optimal and that those that are not are very close to optimal. All designs for

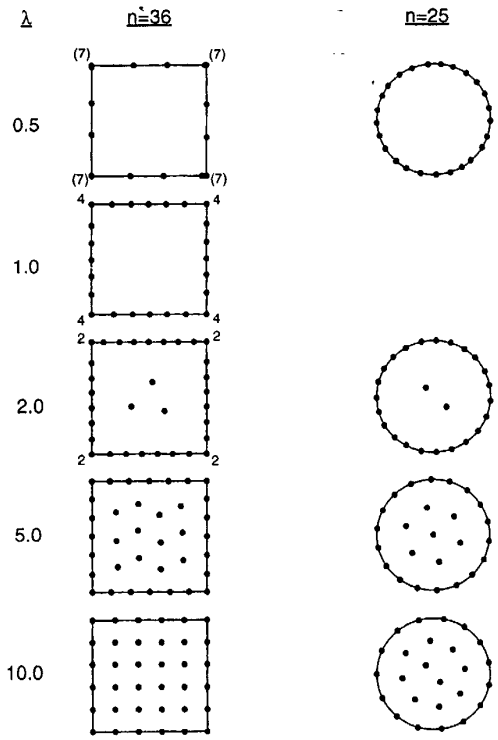

FIG. Id. Optimal Designs for $\rho = e^{-\lambda d}$, n = 36 (square), 25 (circle).

which λ is smaller than for the configurations displayed in the figures distributed the n sampling points as evenly as possible on the corners of the square. One such corner design is shown ($n = 8$, $\lambda = 0.5$). Each other design not shown is precisely the same as the design directly above it. For example, in Figure 1b, the designs for $(n, \lambda) = (9, 5)$ and $(9, 10)$ are not shown because they are the same as the design displayed for $(n, \lambda) = (9, 2)$. Unparenthesized numbers next to design points represent multiple observations which coincide

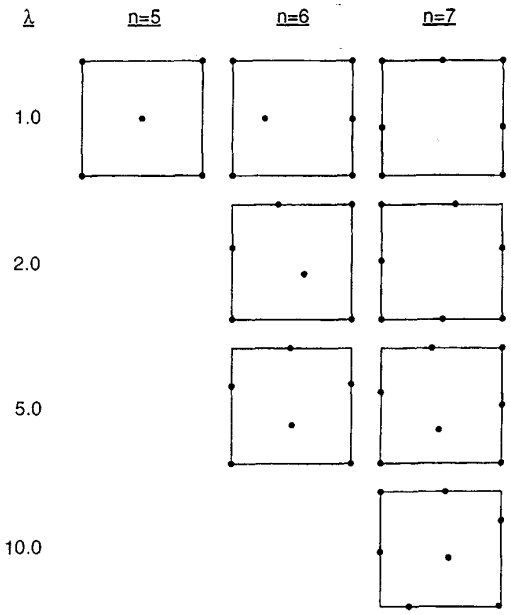

FIG. IIa. Optimal Designs for $\rho = e^{-\lambda d^2}$, n = 5, 6, 7.

to within .01. Parenthesized numbers indicate clusters which may appear to be a multiple design point but which actually include at least some points whose values vary by at least .01.

Denote the variance of the optimal design found for particular n and λ by $V_n(\lambda)$. Tables 1 and 2 below give these minimum design variances for the case $\gamma = 1$ for $\rho = e^{-\lambda d}$ and $\rho = e^{-\lambda d^2}$, respectively. The value for $\lambda = \infty$ (independence) is included for comparison.

Each simulation was run until the design variance stabilized at four decimal places and no new lower variance values were occurring. Simulations were performed using Fortran on a VAX 8550. Typical running times were

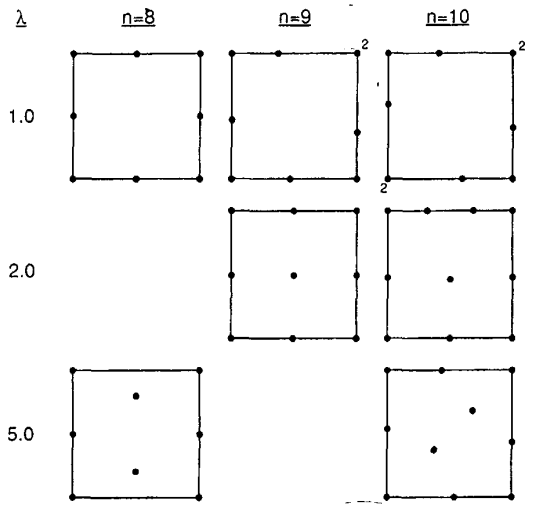

FIG. IIb. Optimal Designs for $\rho = e^{-\lambda d^2}$, n = 8, 9, 10.

about $3.5n$ seconds, involving perhaps $1200n$ iterations although variances within .001 of the minimum were normally found within only about $1.5n$ seconds and $500n$ iterations. Only one point is changed in each iteration, thus it is not surprising that the running time of the algorithm is roughly proportional to n, since the point moved must be recompared to the $n-1$ remaining design points.

Several researchers have noted that simulated annealing converges quite slowly in the late stages of the run. Running times to achieve a variance within .001 of the stable minimum usually varied only slightly according to the initial configuration used; times to actually reach the bottom were somewhat more variable.

There were some cases in which different initial designs led to distinct limiting designs with possibly unequal variances. Typically the difference between the designs found involved a slight variation in the number of points

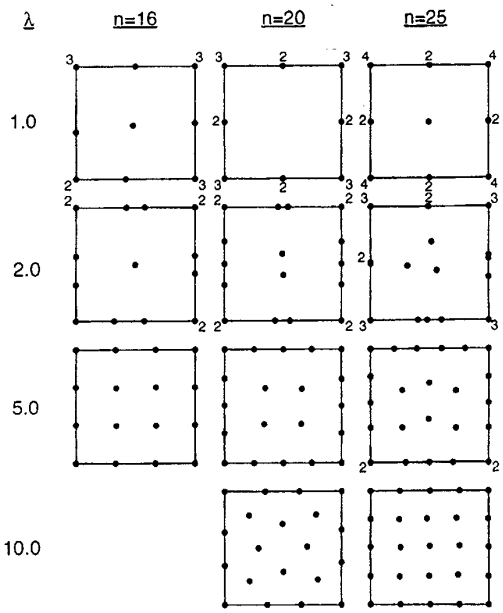

FIG. IIc. Optimal Designs for $\rho = e^{-\lambda d^2}$, n = 16, 20, 25.

to be spread around the interior of the square. In every case, however, the variances of the competing designs differed by at most .0007, and typically by much less. For example, for the case $\rho = e^{-\lambda d}$, $n = 20$, $\lambda = 10$, three different designs were found having either six, seven or eight points arranged in the interior; however, the variances of these designs were identical to four places.

The results of the design study for two dimensions can be capsulized as follows. The statements below apply to both correlation functions:

(1) For λ sufficiently small, the optimal design takes observations only at the corners of the design region, distributed numerically as evenly as possible.

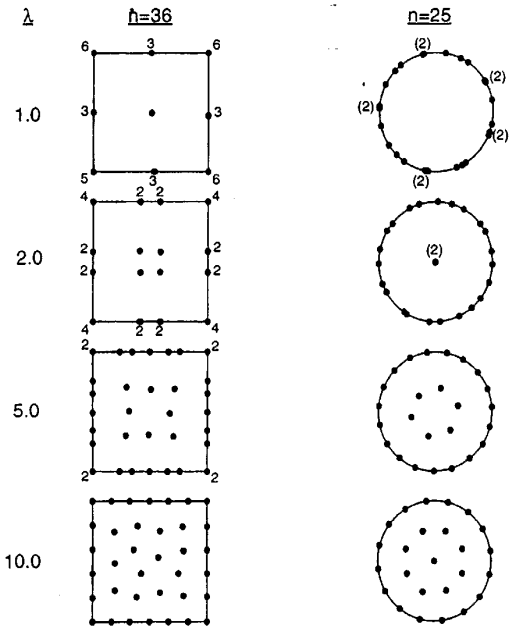

FIG. IId. Optimal Designs for $\rho = e^{-\lambda d^2}$, n = 36 (square), 25 (circle).

(2) For λ sufficiently large, the optimal design spaces the observations evenly over the design region, with several on the edges and at the corners; in particular, if n is a perfect square then the optimal design is a uniform square lattice.

(3) For intermediate values of λ, the optimal design has a higher density of points at the edges and/or corners of the design region, with few or no points in the interior.

The intuitive rationale for these results is that when λ is small, the correlation function has influence over an intermediate to long range; for the

TABLE I

Design Variances, $\gamma = 1$, $\rho = e^{-\lambda d}$

n	$\lambda = 0.1$	0.2	0.5	1.0	2.0	5.0	10.0	∞
5	.8539	.7381	.5060	.3287	.2253	.2003	.2000	.2000
6	.8510	.7335	.5033	.3163	.2048	.1676	.1667	.1667
7	.8509	.7327	.5013	.3098	.1874	.1443	.1429	.1429
8	.8478	.7272	.4947	.3032	.1728	.1268	.1250	.1250
9	.8497	.7305	.4954	.3029	.1680	.1133	.1111	.1111
10	.8489	.7294	.4934	.3015	.1643	.1035	.1000	.1000
16	.8478	.7272	.4901	.2935	.1491	.0707	.0628	.0625
20	.8478	.7272	.4895	.2921	.1463	.0613	.0506	.0500
25	.8480	.7276	.4894	.2911	.1436	.0547	.0410	.0400
36	.8478	.7272	.4888	.2899	.1409	.0473	.0298	.0278

TABLE II

Design Variances, $\gamma = 1$, $\rho = e^{-\lambda d^2}$

n	$\lambda = 0.1$	0.2	0.5	1.0	2.0	5.0	10.0	∞
5	.7096	.5442	.3494	.2492	.2060	.2000	.2000	.2000
6	.7005	.5336	.3430	.2544	.1876	.1669	.1667	.1667
7	.7037	.5349	.3361	.2477	.1740	.1438	.1429	.1429
8	.6975	.5251	.3222	.2390	.1612	.1264	.1250	.1250
9	.7012	.5310	.3306	.2408	.1549	.1131	.1111	.1111
10	.6986	.5282	.3297	.2389	.1587	.1055	.1002	.1000
16	.6975	.5251	.3222	.2359	.1564	.0845	.0647	.0625
20	.6975	.5251	.3222	.2348	.1557	.0839	.0569	.0500
25	.6980	.5259	.3233	.2339	.1545	.0840	.0513	.0400
36	.6975	.5251	.3222	.2341	.1539	.0830	.0496	.0278

corner design the great majority of the $\binom{n}{2}$ terms contributing to the covariance component of $\mathrm{Var}(\bar{Z})$ come from point pairs in which the pairs are widely separated. The penalty paid by taking multiple samples at the same location is more than offset by the high proportion of pairs that are far apart. On the other hand, when λ is large the correlation effects are felt only over the short range; it is thus important in this case to keep the *minimum* distance between points as large as possible. This leads to uniform spacing.

In particular, the optimal design samples only at the corners for $\rho = e^{-\lambda d}$ if $\lambda \leq 0.2$ and for $\rho = e^{-\lambda d^2}$ if $\lambda \leq 0.5$, regardless of n. Note that the variance values in Tables 1 and 2 for these cases are identical for $n = 8$, 16, 20 and 36. This is because each of these designs places 1/4th of the observations at each corner; since $\gamma = 1$ in the tables, repeated measurements at the same location give identical results, thus these designs merely represent multiple copies of a simple four point corner design.

Uniform square lattice configurations occur for both correlation functions when $\lambda \geq 2$ for $n = 9$, $\lambda \geq 5$ for $n = 16$, and $\lambda = 10$ for $n = 25$. For $n = 36$, the uniform square lattice is found for $\lambda = 10$ for $\rho = e^{-\lambda d}$, but a higher value of λ is required for $\rho = e^{-\lambda d^2}$. When n is not a perfect square, the designs for large λ are not lattices but still convey a strong measure of uniformity.

It is easy to visualize the progression of designs dynamically by viewing sets of figures in sequence, especially for the larger values of n. For $\rho = e^{-\lambda d}$, as λ increases, the design for a given n begins with all points in the corners, then begins to peel points off to the edges, where they space themselves regularly. Then points begin to jump to the interior, spreading out according to their number until eventually the process stabilizes with a lattice or lattice-like limiting design. Interestingly, there are no designs with multiple points at the corners that take any observations in the interior.

For $\rho = e^{-\lambda d^2}$ the situation is similar except that small sets of observation points are sometimes placed at the same or nearly the same location; the tendency to spread out evenly at the individual point scale is less dominant. This is undoubtedly due to the flat shape of the correlation function $\rho = e^{-\lambda d^2}$ in the neighborhood of $d = 0$, as opposed to the peak which is present in $\rho = e^{-\lambda d}$. Thus the primary effect of analyticity is the allowance of optimal designs that retain point clusters at locations other than the corners.

Frequently an approximately triangular lattice appears in the interior of the region; see for example $(n, \lambda) = (20, 10)$, $(25, 5)$, and $(36, 5)$ for both

Frequently an approximately triangular lattice appears in the interior of the region; see for example $(n, \lambda) = (20, 10)$, $(25, 5)$, and $(36, 5)$ for both correlation functions. It is reasonable to conjecture that this phenomenon would be even more prevalent were it not for the influence of the boundary.

A brief investigation of the nature of designs for a circular region was also conducted. Designs found for $n = 25$ are shown in Figures 1d and 2d. Solutions for the case of low correlation (high λ) appear very similar in nature to designs obtained for square sampling regions. For the case $\rho = e^{-\lambda d^2}$, $\lambda = 1$ (and to some extent $\lambda = 2$ also), the annealing program failed to find an orderly arrangement of points. Calculations show that for this value of λ, the range of significant influence of ρ covers many intrapoint distances for which ρ is concave upwards and many for which it is concave downwards. The correlation function is roughly linear in this range, hence the tradeoff between a design having many small interpoint distances and a design with some zero intrapoint distances and some moderately small intrapoint distances is about even. As a result the optimization procedure has difficulty choosing between an equally spaced design and a design which clusters points in pairs.

Wille and Vennik (1985) have employed simulated annealing to determine configurations of equal point charges which minimize the total potential energy of the system, when the point charges are constrained to lie within a circle. Although the Coulomb potential function they utilize is quite different from the exponential based correlation functions used here, the configurations they obtain are much like those exhibited in Figures 1d and 2d. Optimal arrangements thus appear to be fairly robust with respect to variations in the dependency structure. Energy minimizing configurations on a sphere are found in Wille (1986).

Sacks and Schiller have also studied the problem of design for a closely related model to the one employed here, using the same two correlation functions. Although some of their designs share the characteristics of spreading points out fairly evenly and placing many points on the boundary, these

arrangements do not reveal the same regularity as the designs found here. Apparently this is due to the fact that their design space is limited to a lattice rather than being a continuous region. Their examples also do not detect the tendency for design points to occur in pairs for the analytic correlation function $\rho = e^{-\lambda d^2}$, probably because a range of values for the parameter λ was not explored.

In the one-dimensional location problem studied in Bickel, Herzberg and Schilling (1981), the equally spaced design was shown to compete extremely favorably with the optimal design when the correlation is low. In two or more dimensions, there is more than one candidate for a design having uniform spacing. The three regular lattices are each valid choices for the spacing of the interior points. Since the design region treated here is a square, the optimal designs found for the cases when n is a square are compared to the corresponding square lattices. In addition, the case $n = 20$ is compared to a five-by-four rectangular lattice. Let the variances of these rectangular lattice designs be denoted by $L_n(\lambda)$, and define the efficiency of a lattice design relative to the corresponding optimal design by $V_n(\lambda)/L_n(\lambda)$. Tables 3 and 4 present these efficiencies for the two correlation functions $\rho = e^{-\lambda d}$ and $\rho = e^{-\lambda d^2}$ and for three values of the repeated measures parameter γ. The efficiency values are seen to be quite high in general, increasing towards 1 as $\rho \to 0$, i.e., as λ increases. Also, efficiencies of uniform designs become higher as γ becomes smaller, i.e., as the dependency between two observations at the same location decreases.

4. RESULTS FOR HIGHER DIMENSIONS

A more limited study was conducted in three and four dimensions in order to verify the extension of the behavior found for one and two-dimensional designs to higher dimensions, and to check the performance of the algorithm for such problems. The results indicate that (1) the nature of

TABLE III

Efficiencies of Uniform Rectangular Lattice Designs, $\rho = e^{-\lambda d}$

λ	$\gamma = 1.0$	0.5	0.2	n
0.1	0.980	0.982	0.987	9
0.1	0.967	0.969	0.974	16
0.1	0.964	0.966	0.971	20
0.1	0.961	0.963	0.967	25
0.1	0.957	0.959	0.962	36
0.2	0.966	0.970	0.979	9
0.2	0.942	0.947	0.956	16
0.2	0.937	0.941	0.950	20
0.2	0.932	0.935	0.943	25
0.2	0.924	0.927	0.933	36
0.5	0.954	0.962	0.975	9
0.5	0.907	0.916	0.936	16
0.5	0.895	0.904	0.923	20
0.5	0.883	0.891	0.909	25
0.5	0.867	0.873	0.889	36
1.0	0.970	0.978	0.988	9
1.0	0.904	0.920	0.946	16
1.0	0.885	0.900	0.928	20
1.0	0.866	0.880	0.909	25
1.0	0.840	0.852	0.879	36
2.0	1.000	1.000	1.000	9
2.0	0.956	0.968	0.983	16
2.0	0.938	0.953	0.973	20
2.0	0.916	0.933	0.959	25
2.0	0.881	0.899	0.930	36
5.0	1.000	1.000	1.000	9
5.0	1.000	1.000	1.000	16
5.0	0.990	0.995	0.998	20
5.0	1.000	1.000	1.000	25
5.0	0.981	0.988	0.994	36
10.0	1.000	1.000	1.000	9
10.0	1.000	1.000	1.000	16
10.0	0.998	0.999	1.000	20
10.0	1.000	1.000	1.000	25
10.0	1.000	1.000	1.000	36

TABLE IV

Efficiencies of Uniform Rectangular Lattice Designs, $\rho = e^{-\lambda d^2}$

λ	$\gamma = 1.0$	0.5	0.2	n
0.1	0.896	0.909	0.934	9
0.1	0.857	0.868	0.891	16
0.1	0.849	0.858	0.879	20
0.1	0.841	0.849	0.867	25
0.1	0.832	0.838	0.852	36
0.2	0.835	0.860	0.903	9
0.2	0.772	0.792	0.834	16
0.2	0.759	0.776	0.813	20
0.2	0.747	0.761	0.794	25
0.2	0.730	0.741	0.767	36
0.5	0.825	0.863	0.917	9
0.5	0.717	0.752	0.819	16
0.5	0.696	0.726	0.788	20
0.5	0.677	0.703	0.759	25
0.5	0.650	0.669	0.714	36
1.0	0.959	0.972	0.985	9
1.0	0.824	0.856	0.906	16
1.0	0.789	0.820	0.874	20
1.0	0.756	0.785	0.840	25
1.0	0.721	0.743	0.792	36
2.0	1.000	1.000	1.000	9
2.0	0.924	0.945	0.969	16
2.0	0.883	0.909	0.946	20
2.0	0.842	0.871	0.916	25
2.0	0.794	0.820	0.869	36
5.0	1.000	1.000	1.000	9
5.0	1.000	1.000	1.000	16
5.0	0.984	0.990	0.995	20
5.0	0.977	0.984	0.992	25
5.0	0.918	0.937	0.963	36
10.0	1.000	1.000	1.000	9
10.0	1.000	1.000	1.000	16
10.0	0.988	0.994	0.998	20
10.0	1.000	1.000	1.000	25
10.0	0.995	0.997	0.999	36

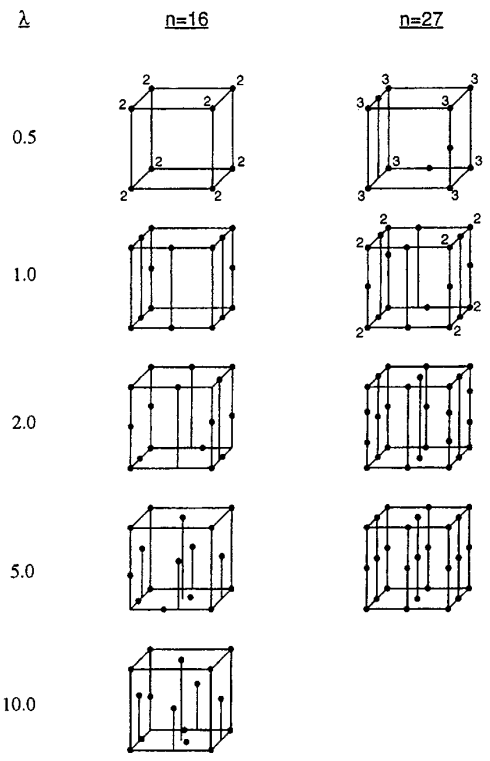

FIG. IIIa. Optimal Designs for $\rho = e^{-\lambda d}$, n = 16, 27.

efficient designs that was found for one and two dimensions carries over to higher dimensions, and (2) the algorithm runs substantially more slowly as the dimension increases. The latter fact appears to be due to the fact that in continuous design space optimization problems there is much more space to search and more directions in which to search when the dimension grows, thus considerably more iterations are needed before the design settles down.

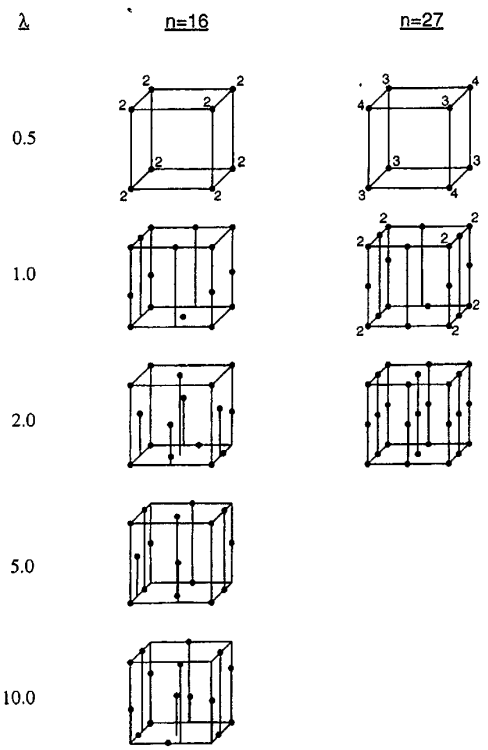

FIG. IIIb. Optimal Designs for $\rho = e^{-\lambda d^2}$, n = 16, 27.

Figures 3a, b and c display some of the designs found for three dimensions for n = 16, 27, and 40 when the sampling region is the cube $[-1,1]^3$. As before, designs for values of λ less than those shown distributed all sample points on the corners of the cube, while each other design not displayed is the same as the one directly above it. As the dependency between observations is decreased, the points begin to spread out, first onto the edges of the cube, then onto the faces, and perhaps finally entering the interior. Points occur fre-

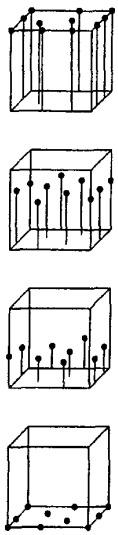

FIG. IIIc. Sectioned Display for $\rho = e^{-\lambda d^2}$, n = 40, λ = 10.

quently at or near the centers of edges or faces. In the case of $n = 27$, the optimal design for both correlation functions is a $3\times3\times3$ cubical lattice when the correlation is low enough.

A small number of annealings, not shown, were tried for $n = 64$. Once again, corner designs were obtained for high correlation and designs approached the $4\times4\times4$ cubical lattice as the correlation became low. For moderate correlation, the tendency persisted as in two dimensions for the analytic correlation function $\rho = e^{-\lambda d^2}$ to yield designs which take multiple observations at points other than at the vertices. A typical case is $\lambda = 1$, which gave a design placing four to five points on each corner, two on the center of each edge, and one in the center of each face. Again this clustering behavior was not found to occur for $\rho = e^{-\lambda d}$.

If n is not a perfect cube then the design cannot be a cubical lattice; however, the designs found still spread the points evenly over the design region in a lattice-like manner when the correlation is low. The sectioned display in Figure 3c shows this effect for the case $\rho = e^{-\lambda d^2}$, $n = 40$, $\lambda = 10$.

The basic structures of efficient designs found for one, two and three dimensions continued to appear in annealings performed in four dimensions. For $n = 16$ the limiting designs placed all observations at the corner of the hypercube for all values of λ as would be expected since there are 16 corner sites available. Experiments for $n = 81$ again showed a progression from corner designs for strong correlation to lattice designs for weak correlation. Results are not presented graphically because of the added difficulty in displaying an extra dimension and because no significant new insights are gained from these cases.

The propensity for the algorithm to converge to different designs for the same values of n, λ, and ρ is much greater in three and four dimensions than in two dimensions. This reflects not on the accuracy of the algorithm but rather on the shape of the variance function surface, which appears to have in many cases a large number of local minima taking on virtually the same variance value. For example, in three dimensions using $\rho = e^{-\lambda d^2}$ with $\lambda = 5$ for n=16, the following designs were found on separate runs:

# on corners	# on edges	# on faces	# in interior
6	5	5	0
4	8	4	0
4	7	4	1
3	11	1	1

All four designs had the same variance to four places.

5. CONCLUSIONS

The results presented suggest a reasonably simple strategy for selecting the locations at which observations should be taken to make efficient estimates

of the mean of a multidimensional process using the sample mean estimator, at least in the case when the design region is convex. If the correlation is believed to decrease rapidly with distance, a design should be chosen which is evenly spaced over the design region and its boundary, perhaps using a lattice structure if possible. If the correlation is moderate, then a somewhat higher density of points should be placed on the boundary and at any extreme points. If the correlation is high, no observations should be made in the interior of the design region.

If theoretical considerations suggest that correlation increases rapidly as the interpoint distance approaches zero, then multiple samples at the same location should not be taken except possibly at the extreme points of the design region, whereas if the correlation is believed to vary slowly near zero, then a limited number of multiple observations may be appropriate at certain locations. More study needs to be done for other region shapes and for larger sample sizes in order to determine additional guidelines on this aspect.

The evidence obtained in this paper strongly supports the theoretical and numerical results of BH and BHS that evenly spaced designs are optimal as the correlation decreases to zero. Such designs have the additional advantage of allowing validation of model assumptions--in particular, the constantness of both the mean and variance over the sampling region.

ACKNOWLEDGEMENT

I am grateful to Brad McWilliams of California State University, Northridge for his generous assistance in producing the plots. This work was supported in part by the California State University, Northridge.

BIBLIOGRAPHY

Bickel, P. J. and Herzberg, A. M. (1979). 'Robustness of design against auto-correlation in time I: asymptotic theory, optimality for location and linear regression,' *Annals of Statistics*, 7, 77-95.

Bickel, P. J., Herzberg, A. M. and Schilling, M. F. (1981). 'Robustness of design against autocorrelation in time II: optimality, theoretical, and numerical results for the first-order autoregressive process,' *Journal of the American Statistical Association*, 76, 870-877.

Bohachevsky, I. O., Johnson, M. E. and Stein, M. L. (1986). 'Generalized simulated annealing for function optimization,' *Technometrics*, 28, 209-217.

Box, M. J. (1966). 'A comparison of several current optimization methods, and the use of transformations in constrained problems,' *Computer Journal*, 9, 67-77.

Haines, L. M. (1987). 'The application of the annealing algorithm to the construction of exact optimal designs for linear-regression models,' *Technometrics*, 29, 439-447.

Kirpatrick, S., Gelatt, C. D. and Vecchi, M. P. (1983). 'Optimization by simulated annealing,' *Science*, 13, 671-680.

van Laarhoven, P. J. M. and Aarts, E. H. L. (1987). 'Simulated annealing: theory and applications,' *Mathematics and its Applications*, D. Reidel Publishing Co.

Meyer, R. K. and Nachtsheim, C. J. (1988). 'Constructing exact d-optimal experimental designs by simulated annealing,' *American Journal of Mathematics and Management Sciences*, 8, 329-359.

Sacks, J. and Schiller, Susannah (1988). 'Spatial designs.' In *Statistical Decision Theory and Related Topics IV* (S. S. Gupta and J. O. Berger, eds.), 2, 385-399.

Szu, H. and Hartley, R. (1986). 'Fast simulated annealing with Cauchy probability densities,' *Physics Letters A*, 122, 157-162.

Vanderbilt, D. and Louie, S. G. (1984). 'A monte-carlo simulated annealing approach to optimization over continuous variables,' *Journal of Computational Physics*, 56, 259-271.

Wille, L. T. (1986). 'Searching potential energy surfaces by simulated annealing,' *Nature*, 324, 46-48.

Wille, L. T. and Vennik, J. (1985). 'Electrostatic energy minimisation by simulated annealing,' *Journal of Physics A – Mathematics and General*, 18, L1113-L1117.

Received February 1991; Revised July 1991