School of Operations Research
and Industrial Engineering
College of Engineering
Cornell University
Ithaca, N.Y. 14853

Technical Report No. 714
August 1986
(Revised September 1987)

Driving Frequency Selection for Frequency Domain Simulation Experiments

by

S.H. Jacobson
A. Buss
L.W. Schruben

This research was supported by AT&T Information Systems and Air Force System Command, Rome Air Development Center, Griffiss Air Force Base, NY 13441-5700 by subcontract from Syracuse University, Syracuse, NY.

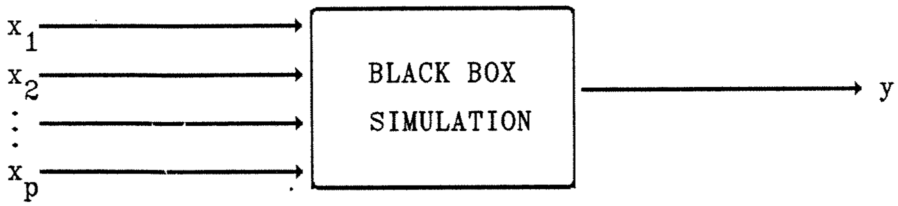
ABSTRACT

Frequency Domain Methodology has recently been applied to discrete event simulations to identify terms in a polynomial model of the simulation output. The problem of optimally selecting input frequencies is studied in this paper. This problem can be formulated as a large mixed integer linear program or as a large set of small linear programs. A fast algorithm is presented that closely approximates the optimal solution. The results obtained from the algorithm are compared to known optimal solutions.

1. INTRODUCTION:

Simulation programs can involve a large number of input factors. Analyzing such systems can be very difficult. Standard analysis techniques require a new simulation run for each setting of the input factors. This can be both expensive and time consuming. We will refer to such procedures as "run oriented" experiments. An alternative to this approach was introduced by Schruben and Cogliano [7] in which the input factors in a single run are varied according to specific sinusoidal patterns. Various output spectra can then be analyzed to obtain information about the sensitivity of the output to each of the input factors. Such experiments are called "frequency domain" simulation experiments. Frequency domain simulation experiments were initially introduced as a technique for factor screening. More recently, work has been done to apply the technique for simulation optimization using patterns in the simulation output power spectrum [6] as well as by estimating the sensitivities of the expected simulation output response from spike heights in the simulation output power spectrum.

In the frequency domain simulation experiments considered in this paper, input factors are varied according to specific sinusoidal patterns during a run of a simulation program. Input patterns other than sinusoidals are of course permissible. The only restriction on these patterns is that they form a complete orthogonal basis when viewed as a vector in R^n, where n is the number of observations collected during a run (i.e. the run size). The following diagram depicts a "black box" simulation (i.e. we input values into the simulation (black box) and observe an output value (or values)):



where $x_1, x_2, \ldots, x_p$ are the input factors and y is the output response of the simulation. We will assume that y can be modelled as a polynomial function of the input factors. This function is sometimes referred to as a meta-model of the simulation response. More specifically, if y is the simulation response and $x_1, x_2, \ldots, x_p$ are the input factors, then the response surface is given by:

$$E(y) = \beta_0 + \sum_{j=1}^{q} \beta_j \tau_j \tag{1.1}$$

where $\tau_1, \tau_2, \ldots, \tau_q$ are possible terms in the polynomial model; they are products of non-negative powers of continuous parameters x_i $(i=1,2,\ldots,p)$.

$\beta_1, \beta_2, \ldots, \beta_q$ are real valued coefficients.

An experimental technique for identifying the functional form of E(y) was presented by Schruben and Cogliano [8]. The technique requires one to specify k, the degree of the polynomial given by (1.1). One must then select frequencies at which the inputs can be changed during the running or execution of the simulation. These frequencies, called <u>driving frequencies</u>, should be chosen such that all the frequencies which indicate the presence of terms in the prospective meta-model are as spread out as possible. Such identifying frequencies will be referred to as <u>term indicator frequencies</u>.

The term indicator frequencies are obtained by standard trignometric relations between sine, cosine and their powers, as well as from aliasing (see (1.5)). For example, if a general term in the polynomial model is given by $\tau=(x_1)^{a_1}(x_2)^{a_2}\ldots(x_r)^{a_r}$, and we oscillate x_i at frequency ω_i, for $i=1,2,\ldots,r$, then the term indicator frequencies will be $S_1 \oplus S_2 \oplus \ldots \oplus S_k$, where $S_i=\{(a_i)\omega_i,(a_i-2)\omega_i,\ldots,(-a_i+2)\omega_i,(-a_i)\omega_i\}$, $i=1,2,\ldots,r$ and $S_i \oplus S_j$ is the direct sum of sets S_i and S_j (i.e. the set of all possible sums s_i+s_j where $s_i \in S_i$ and $s_j \in S_j$). To show this, consider the following term:

$$(x_1)^{a_1}(x_2)^{a_2}$$

We will oscillate factor x_1 with frequency ω_1 and factor x_2 with frequency ω_2 (i.e. $x_1(t)=x_1(0)+\alpha_1\cos(2\pi\omega_1 t)$, $x_2(t)=x_2(0)+\alpha_2\cos(2\pi\omega_2 t)$ where $t=0,1,2,\ldots$).
The following argument will still hold if we use sine to oscillate the factors. Now,

$$\cos(2\pi\omega_1 t)=(e^{2\pi\sqrt{-1}\omega_1 t}+e^{-2\pi\sqrt{-1}\omega_1 t})/2. \tag{1.2}$$

Thus we have

$$\begin{aligned}(x_1)^{a_1}= {} & (x_1(0))^{a_1}+ \\ & a_1(x_1(0))^{a_1-1}\cos(2\pi\omega_1 t)+ \\ & ((a_1)(a_1-1)/2)(x_1(0))^{a_1-2}\cos^2(2\pi\omega_1 t)+\ldots+ \\ & a_1(x_1(0))\cos^{a_1-1}(2\pi\omega_1 t)+\cos^{a_1}(2\pi\omega_1 t).\end{aligned} \tag{1.3}$$

Using (1.2), $\cos^k(2\pi\omega_1 t)$ can be computed for $k=2,3,\ldots,a_1$.

Regrouping the terms in the resulting expression gives us

$$(x_1)^{a_1} = (\cos(2\pi\omega_1 t(a_1)) + (a_1)\cos(2\pi\omega_1 t(a_1-2)) + \dots + (a_1)\cos(2\pi\omega_1 t(-a_1+2)) + \cos(2\pi\omega_1 t(-a_1))) \qquad (1.4)$$

We get a similar expression for $(x_2)^{a_2}$.

Therefore using the trignometric identity

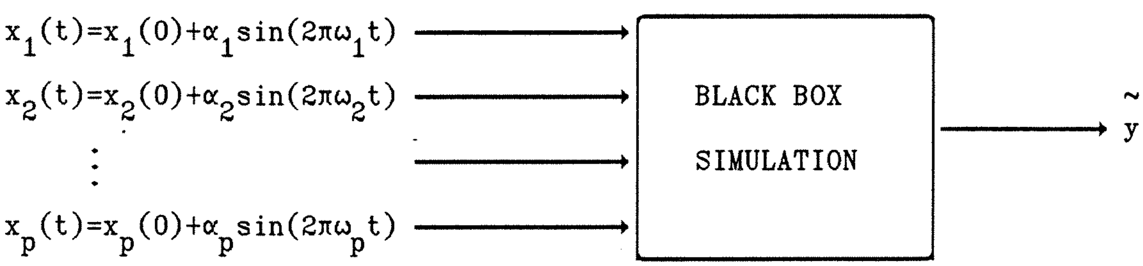
$$\cos(a+b)+\cos(a-b)=2\cos(a)\cos(b)$$

the set of term indicator frequencies is

$$S_{x_1} \oplus S_{x_2}.$$

Note that the term indicator frequencies will not change if the model given by (1.1) is generalized to contain memory.

For continuous inputs, sinusoidal oscillation patterns can be used. For discrete or qualitative inputs, randomized rectangular oscillation patterns can be used [6]. The following diagram depicts this for continuous inputs:



where $x_i(t)$ are the input factors ($i=1,2,\dots,p$) at time t ($t\geq 0$), α_i ($i=1,2,\dots,p$) defines the range over which the inputs are oscillated, called the <u>amplitude</u> <u>of</u> <u>oscillation</u>, ω_i are the frequencies at which the inputs are oscillated (i.e. the driving frequencies) and $\tilde{y}$ is the output response obtained from the procedure described above. The α_i and ω_i fully define the input power spectrum. The frequency domain technique allows us to identify the presence of the τ_j in the model, by estimating the output power spectrum and identifying significant spike heights at the term indicator frequencies of the τ_j (j=1,2,...,q). To be able to get good power spectrum spike height estimates with reasonable run lengths, an attempt should be made to space the term indicator frequencies as far apart as possible. It would be optimal to be able to evenly space these frequencies in the interval (0,.5). Therefore the criterion for selecting driving frequencies is to attempt to space as evenly as possible all term indicator frequencies, given k, the assumed upper bound on the polynomial degree that fits the simulation response. This will widen the bandwidth and permit the desirable level of power spectrum estimator variance and bias without excessively long run lengths (see Chatfield [1] p.141-145).

The main advantage of the frequency domain method is that it reduces the number of simulation runs necessary to screen factors. Keeping these run lengths short makes this technique even more attractive.

The problem discussed in this paper is the selection of driving frequencies such that the term indicator frequencies are maximally spaced (i.e. the minimum spacing between all the term indicator frequencies is maximized). Because we can sample the output response at only discrete times, certain frequencies can not be distinguished. This causes some of the frequencies to be folded or aliased onto other frequencies. All frequencies can be aliased into the interval [0,.5]. This <u>alias rule</u> is as follows:

If ω is a term frequency, then if $\omega<0$, ω become $-\omega$

if $0\leq\omega\leq.5$, ω is unchanged (1.5)

if $.5<\omega\leq1.$, ω becomes $1-\omega$

if $\omega>1$, ω becomes $\omega-1$

These rules are repetitively applied until ω is in the interval [0,.5].

Section 2 defines the problem and gives a mixed integer linear program formulation which solves it. Section 3 gives a reformulation of the problem as an exponential number (in p and k) of linear programs, each with a polynomial number (in p and k) of constraints using the concept of term indicator frequency ordering. Section 4 gives an algorithm which approximates the optimal solution. Examples are given throughout which help to clarify the concepts and ideas presented. Tables of optimal and near optimal driving frequencies are included in Appendix 3 for those interested in conducting frequency domain simulation experiments.

2. PROBLEM STATEMENT:

The problem addresses in this report is to select a set of input frequencies such that indicator frequencies for terms in the polynomial model given by (1.1) are as widely spaced as possible. We want to select a set of driving frequencies which maximizes the minimum space between the term indicator frequencies.

The problem can be stated explicitly as follows:

Let $\omega_1, \omega_2, \ldots, \omega_p$ be the driving frequencies for the polynomial model (without loss of generality, because of aliasing we can assume $0<\omega_1<\omega_2<\ldots<\omega_p<.5$).

Let $\Omega=\{\eta : \eta$ is a potentially distinct term indicator frequency for the polynomial model, including the endpoints at 0 and .5$\}$ (we denote v=cardinality(Ω) (i.e. the number of elements in Ω)).

Given a bound, k, on the degree of the meta-model, our problem is to find

$$\mathfrak{z}_k = \underset{\omega_1,\omega_2,\ldots,\omega_p}{\text{Maximize}} \ \text{Minimize}\{|\eta_i-\eta_j|: \eta_i,\eta_j\in\Omega,\ i,j=1,2,\ldots,v,\ i\neq j\} \quad (2.1)$$

We call $\mathfrak{z}_k$ the optimal spacing for a given polynomial model.

Notice that term indicator frequencies for even (odd) powers of an input factor are a proper subset of the indicator frequencies for the highest order even (odd) power term. The frequency 0 is also a term indicator frequency for any term which is an even power of an input. If frequency 0 is included in the set Ω for every such term, then $\mathfrak{z}_k=0$ for all k>1. To avoid this difficulty, we will exclude frequency 0 as such a term indicator frequency. We also consider only indicator frequencies of the highest order even power (and the highest order odd power) for each input in the prospective meta-model. The inclusion of lower order powers would necessarily make $\beta_k=0$.

A mixed integer linear program has been formulated by Cogliano [2] to solve the frequency selection problem. The constraints in the program are determined by the differences between all the term indicator frequencies for the model (1.1). Since it is not always clear whether these differences are positive or negative, absolute values of these differences are required. It is the need for these absolute values which creates the need for the 0-1 integer variables. Without the absolute values, the problem could be formulated as a linear program. The number of constraints in the mixed integer linear program is at least $O((p)^{2k})$ where p is the number of driving frequencies and k is the degree of the polynomial one is willing to assume fits the simulation response. We obtained this number by observing that, given p and q, there are

$$v = pk + \sum_{r=1}^{k-1} \sum_{j=1}^{r} 2^j \binom{r}{j} \binom{p}{j+1} + 2$$

term indicator frequencies (See Appendix 1). By taking all possible differences between these term indicator frequencies, it is straightforward to show that the number of constraints in the mixed integer linear program is of the order given above. The program is as follows:

$$\begin{aligned} &\text{Maximize} \quad Z \\ \text{Subject to} \quad & Z \le |\eta_i - \eta_j| \\ & \eta_i, \eta_j \in \Omega, \text{ for all } i \ne j, \quad i,j=1,2,\ldots,v \\ & 0 \le \omega_i \le .5 \text{ for } \omega_1 \text{ through } \omega_p \end{aligned} \tag{2.2}$$

where ω_1 to ω_p represent driving frequencies for p factors and the η_i represent term indicator frequencies for the prospective polynomial response model (1.1) of order k.

The following example demonstrates this mixed integer linear program:

Example 2.1:

Let p=2, k=2.

Therefore, we need two driving frequencies (say ω_1, ω_2)

(Without loss of generality, we will assume $0 < \omega_1 < \omega_2 \le .5$, due to the effect of aliasing)

The term indicator frequencies will be:

$$\omega_1 \qquad \omega_2 \qquad 2\omega_1 \qquad 2\omega_2 \qquad \omega_2+\omega_1 \qquad \omega_2-\omega_1.$$

The mixed integer linear program is as follows:

Maximize Z

Subject to

$$
\begin{array}{lll}
\omega_1 \geq Z & 1-2\omega_2 \geq Z & |1+\omega_1-2\omega_2| \geq Z \\
\omega_2 \geq Z & 1-3\omega_2 \geq Z & |1-2\omega_1-2\omega_2| \geq Z \\
2\omega_1 \geq Z & .5-\omega_1 \geq Z & |1-3\omega_1-\omega_2| \geq Z \\
2\omega_2 \geq Z & .5-\omega_2 \geq Z & |1-3\omega_2| \geq Z \\
\omega_1+\omega_2 \geq Z & .5+\omega_1-\omega_2 \geq Z & |1-4\omega_2| \geq Z \\
-\omega_1+\omega_2 \geq Z & |-2\omega_1+\omega_2| \geq Z & |1+\omega_1-3\omega_2| \geq Z \\
-\omega_1+2\omega_2 \geq Z & |-3\omega_1+\omega_2| \geq Z & |1+2\omega_1-2\omega_2| \geq Z \\
-2\omega_1+2\omega_2 \geq Z & |1-4\omega_1| \geq Z & |1-\omega_1-3\omega_2| \geq Z \\
1-2\omega_1 \geq Z & |1-2\omega_1-\omega_2| \geq Z & |.5-2\omega_1| \geq Z \\
1-\omega_1-\omega_2 \geq Z & |1-3\omega_1| \geq Z & |.5-2\omega_2| \geq Z \\
1-\omega_2 \geq Z & |1-\omega_1-2\omega_2| \geq Z & |.5-\omega_1-\omega_2| \geq Z
\end{array}
$$

$$0<\omega_1<\omega_2\leq .5$$

The solution to the mixed integer linear program gives us the driving frequencies $\omega_1^*=(2/14)$ and $\omega_2^*=(3/14)$, with spacing (1/14). The term indicator frequencies are:

$$
\begin{array}{ccc}
\omega_2^*-\omega_1^*=(1/14) & \omega_1^*=(2/14) & \omega_2^*=(3/14) \\
2\omega_1^*=(4/14) & \omega_2^*+\omega_1^*=(5/14) & 2\omega_2^*=(6/14)
\end{array}
$$

Solving the above mixed integer linear program can be quite difficult unless a small number of driving frequencies are desired (the above example for p=k=2 has 27 constraints containing 15 variables, of which 12 are integer (0-1). In general, for k=2 and p arbitrary, the mixed integer linear program will contain $(p^4+6p^3-p^2+9p+3)/3$ constraints containing $(p^4+6p^3-p^2+12p+6)/6$ variables, of which $(p^4+6p^3-p^2+6p)/6$ are integer (0-1)). We conjecture that this problem is NP-Complete ([3]), although this has not been demonstrated. If that is the case, an algorithm which can approximate the optimal solution in a polynomial number of steps in p and k would be desired.

3. PROBLEM ANALYSIS:

The mixed integer linear program presented in Section 2 gives us one formulation which can solve the frequency selection problem. We now present a second formulation of the problem.

If one could select driving frequencies such that all the term indicator frequencies (after aliasing) are evenly spaced in the interval (0,.5) cycles per observation or unit of time the problem would be solved. When we say that the set of term indicator frequencies are evenly spaced, we include the endpoints 0 and .5 as part of that set. However, these frequency endpoints are not considered true term indicator frequencies; they are included in the set of term indicator frequencies for spacing requirements only. If we exclude the endpoint .5, it will widen the spacing, since the constraints of the mixed integer linear program described in Section 2 will be defined using one less frequency. However, unless otherwise stated, we will assume that this endpoint is included as part of the set of frequencies which we are attempting to space out. This is a reasonable assumption to make since the power spectrum spike heights at the endpoint frequency can contain information which is of no value to us. If this frequency is not included when we attempt to space out all the indicator frequencies, it's spike height may distort neighboring frequency spike heights which are not sufficiently apart to ensure their independence. (see Jenkins and Watts [4] p.286).

Even spacing is never feasible except for very small problems. The only cases we have found where even spacing can be achieved are when the response is linear or there are only 2 factors (i.e. p=2) and the degree of the polynomial is arbitrary. In the first case, the p driving frequencies are $\omega_i = i/(2p+2)$, $i=1,2,\ldots,p$, where the spacing between the term indicator frequencies, which happen to be just the driving frequencies, is 1/(2p+2). In the second case, the two driving frequencies are ω_1=k/(2k(k+1)+2) and ω_2=(k+1)/(2k(k+1)+2), where the spacing between the term indicator frequencies is 1/(2k(k+1)+2) (see Appendix 2 for a proof of this result). Let us examine this special case of even spacing for p=2 and k=3.

Example 3.1:

Let p=2,k=3.

Let $0<\omega_1<\omega_2<.5$ represent the driving frequencies.

Therefore, the term indicator frequencies are

ω_1 ω_2 $2\omega_1$ $2\omega_2$ $\omega_2-\omega_1$ $\omega_2+\omega_1$

$3\omega_1$ $3\omega_2$ $2\omega_2-\omega_1$ $2\omega_1-\omega_2$ $2\omega_1+\omega_2$ $2\omega_2+\omega_1$.

If ω_1=k/(2k(k+1)+2)=3/26 and ω_2=(k+1)/(2k(k+1)+2)=4/26, then the term indicator frequencies are:

$\omega_2-\omega_1=1/26$	$2\omega_1-\omega_2=2/26$	$\omega_1=3/26$	$\omega_2=4/26$
$2\omega_2-\omega_1=5/26$	$2\omega_1=6/26$	$\omega_2+\omega_1=7/26$	$2\omega_2=8/26$
$3\omega_1=9/26$	$2\omega_1+\omega_2=10/26$	$2\omega_2+\omega_1=11/26$	$2\omega_2=12/26$.

The term indicator frequencies are evenly, or perfectly spaced in the interval [0,.5].

The relative positions of the term indicator frequencies in the interval (0,.5) can be quite arbitrary. There is no obvious way of predicting these positions, such that the minimum spacing can be increased or even shown to be optimal. We define an <u>ordering</u> of the term indicator frequencies as the sequence in which these frequencies fall into the interval [0,.5]. We will assume that for a particular set of driving frequencies, the ordering will not change if the driving frequencies are permuted among the different factors.

To demonstrate some orderings, suppose there are two factors and the degree of the polynomial is assumed to be k=2. Let ω_1 and ω_2 represent the driving frequencies of these two factors. The term indicator frequencies can easily be shown to be ω_1, ω_2, $2\omega_1$, $2\omega_2$, $\omega_2+\omega_1$, $\omega_2-\omega_1$ (we will assume, without loss of generality, $\omega_1<\omega_2$). If (.) represents an aliased term, then some possible orderings are:

$\omega_2-\omega_1$	ω_1	ω_2	$2\omega_1$	$\omega_2+\omega_1$	$2\omega_2$
ω_1	$2\omega_1$	$\omega_2-\omega_1$	ω_2	$\omega_2+\omega_1$	$(2\omega_2)$
$\omega_2-\omega_1$	$(2\omega_2)$	$(\omega_2+\omega_1)$	$(2\omega_1)$	ω_1	ω_2

For the case of p=2 and k arbitrary, one can observe that the term indicator frequencies position themselves in a predictable pattern, by inserting new term indicator frequencies in specific positions as k increases. No such pattern seems to exist for p>2. The number of possible orderings of v term indicator frequencies can also grow to be quite large (i.e. $O(2^v v!/p!)$ since the v term indicator frequencies can be permuted in any order (v!), each term indicator frequency can be represented by itself or its alias (2^v), and the p driving frequencies cannot be shuffled or permuted to yield a new ordering (1/p!)). These orderings have an important role in solving the frequency selection problem. To see this, let us define a <u>locally optimal solution</u>, with respect to a particular ordering, to be a set of driving frequencies such that no changes can be made in any of the driving frequencies without decreasing the minimum spacing. Therefore, each ordering has a locally optimal solution. The mixed integer linear program (2.2) selects the best such local solution, hence is solving a global optimization problem. The following summarizes this:

A set of driving frequencies is said to be locally optimal with respect to an ordering of term indicator frequencies if a small change in any or all of the driving frequencies results in a decrease in the minimum spacing of the indicator frequencies. The local optimum which has the largest minimum spacing is the optimal solution.

Let us look at the following examples of different locally optimal solutions for different orderings:

Example 3.2:

Let p=3 k=2.

Consider the following orderings:

$$\begin{array}{cccccc} \omega_1 & 2\omega_1 & \omega_3-\omega_2 & \omega_2-\omega_1 & \omega_2 & \omega_2+\omega_1 \\ \omega_3-\omega_1 & \omega_3 & \omega_3+\omega_1 & 2\omega_2 & (2\omega_3) & \omega_3+\omega_2 \end{array}$$

$$\begin{array}{cccccc} \omega_1 & 2\omega_1 & \omega_2-\omega_1 & \omega_2 & \omega_2+\omega_1 & \omega_3-\omega_2 \\ 2\omega_2 & \omega_3-\omega_1 & \omega_3 & \omega_3+\omega_1 & (2\omega_3) & \omega_3+\omega_2 \end{array}$$

The first ordering has a locally optimal solution, with driving frequencies given by: ω_1=1/28 ω_2=5/28 ω_3=8/28.

The term indicator frequencies are:

$$\begin{array}{cccc} \omega_1=(1/28) & 2\omega_1=(2/28) & \omega_3-\omega_2=(3/28) & \omega_2-\omega_1=(4/28) \\ \omega_2=(5/28) & \omega_2+\omega_1=(6/28) & \omega_3-\omega_1=(7/28) & \omega_3=(8/28) \\ \omega_3+\omega_1=(9/28) & 2\omega_2=(10/28) & (2\omega_3)=(12/28) & \omega_3+\omega_2=(13/28), \end{array}$$

which gives a spacing of (1/28).

The second ordering has a locally optimal solution, with driving frequencies given by: $\omega_1=1/32$ $\omega_2=4/32$ $\omega_3=10/32$.

The term indicator frequencies are:

$\omega_1=(1/32)$ $2\omega_1=(2/32)$ $\omega_2-\omega_1=(3/32)$ $\omega_2=(4/28)$

$\omega_2+\omega_1=(5/32)$ $\omega_3-\omega_2=(6/32)$ $2\omega_2=(8/32)$ $\omega_3-\omega_1=(9/32)$

$\omega_3=(10/32)$ $\omega_3+\omega_1=(11/32)$ $(2\omega_3)=(12/32)$ $\omega_3+\omega_2=(14/32)$,

which gives a spacing of (1/32).

Suppose we are given an ordering of the term indicator frequencies. We would like to find a set of driving frequencies which maximizes the minimum spacing between the resulting term indicator frequencies, where the relative position of the term indicator frequencies in the interval [0,.5] is fixed by the ordering. We will include 0 and .5.

We solve for these driving frequencies by solving a linear program which maximizes the minimum spacing between all the terms, subject to the constraint that the ordering of the terms does not change (this linear program is a simplification of (2.2), where the fixed ordering eliminates most of the constraints, including the 0-1 integer variables, making the number of constraints polynomial, instead of exponential, in p and k. The program is linear since only the distance between adjacent terms must be maximized, as determined by the ordering. This implies that there will be only O(v) constraints in the linear program, where v is the number of term indicator frequencies).

To illustrate this, consider the following example:

Example 3.3:

Let p=3, k=2.

Let ω_1, ω_2, and ω_3 represent the driving frequencies.

Let the order of the terms be as follows:

$$\begin{array}{cccccc} \omega_1 & 2\omega_1 & \omega_2-\omega_1 & \omega_2 & \omega_2+\omega_1 & (2\omega_3) \\ \omega_3-\omega_2 & 2\omega_2 & \omega_3-\omega_1 & \omega_3 & \omega_3+\omega_1 & (\omega_3+\omega_2) \end{array}$$

where ( ) means that the term has been aliased (we can choose an ordering arbitrarily. This particular ordering will yield a local optimum near the theoretical upper bound for the maximum of the minimum spacing of the terms).

To obtain the driving frequencies, we need to solve the following linear program:

Maximize Z

Subject to

$$\begin{array}{rlrl} \omega_1 & \geq Z & 3\omega_2-\omega_3 & \geq Z \\ -3\omega_1+\omega_2 & \geq Z & -\omega_1-2\omega_2+\omega_3 & \geq Z \\ 1-\omega_1-\omega_2-2\omega_3 & \geq Z & 1-\omega_1-\omega_2-2\omega_3 & \geq Z \\ -1\quad -\omega_2+3\omega_3 & \geq Z & -.5\quad +\omega_2+\omega_3 & \geq Z \end{array}$$

$$\omega_1\geq 0 \quad \omega_2\geq 0 \quad \omega_3\geq 0$$

where the first constraint is a result of $[\omega_1]-[0]$, $[2\omega_1]-[\omega_1]$, $[\omega_2]-[\omega_2-\omega_1]$, $[\omega_2+\omega_1]-[\omega_2]$, $[\omega_3]-[\omega_3-\omega_1]$ and $[\omega_3+\omega_1]-[\omega_3]$. The second constraint is a result of $[\omega_2-\omega_1]-[2\omega_1]$. The third constraint is a result of $[(2\omega_3)]-[\omega_2+\omega_1]$. The fourth constraint is a result of $[\omega_3-\omega_2]-[(2\omega_3)]$. The fifth constraint is a result of $[2\omega_2]-[\omega_3-\omega_2]$. The sixth constraint is a result of $[\omega_3-\omega_1]-[2\omega_2]$. The seventh constraint is a result of $[(\omega_3+\omega_2)]-[\omega_3+\omega_1]$. The eighth constraint is a result of $[.5]-[(\omega_3+\omega_2)]$.

The solution is $Z^*=1/28$ with $\omega_1^*=1/28$, $\omega_2^*=4/28$, $\omega_3^*=11/28$.

This spacing is optimal for the ordering given.

Let us consider another ordering:

Example 3.4:

$$
\begin{array}{cccccc}
\omega_2-\omega_1 & \omega_3-\omega_2 & \omega_3-\omega_1 & \omega_1 & \omega_2 & \omega_3 \\
2\omega_1 & \omega_2+\omega_1 & 2\omega_2 & \omega_3+\omega_1 & \omega_3+\omega_2 & 2\omega_3
\end{array}
$$

To obtain the driving frequencies, we need to solve the following linear program:

$$\text{Maximize } Z$$

Subject to

$$
\begin{array}{lll}
-\omega_1+\omega_2 \geq Z & \qquad & .5 \quad -2\omega_3 \geq Z \\
\omega_1-2\omega_2+\omega_3 \geq Z & & \omega_1 \geq 0 \\
2\omega_1 \quad -\omega_3 \geq Z & & \omega_2 \geq 0 \\
-\omega_2-\omega_3 \geq Z & & \omega_3 \geq 0
\end{array}
$$

The solution is $Z^*=1/30$ with $\omega_1^*=4/30$, $\omega_2^*=5/30$, $\omega_3^*=7/30$. This spacing is optimal for the ordering given but is worse than the spacing found for the previous ordering.

By looking at other orderings of the terms, one can obtain results which are locally optimal for each such ordering. To find the global optimum, one would have to consider every possible ordering, since each may yield a different locally optimal spacing. Since there are $2^v v!/p!$ possible orderings, where v is the number of term indicator frequencies, the work involved to find the global optimum would be exponential in p and k. We have observed that many of the orderings will yield zero as the solution to the linear program described above. It is not clear whether there are an exponential or polynomial number of orderings (in p and k) which yield positive optimal spacing or whether there is an easy way to identify and eliminate these zero spacing orderings.

4. THE ALGORITHM:

We have two formulations of our problem. One involves solving a mixed integer linear program, with a polynomial number of constraints. The other involves solving an exponential number of simple linear programs. Neither of these methods can be used to solve the problem very quickly or easily. It is clear that a heuristic which obtains reasonable results, in a polynomial number of operations (in p and k), would be an important step towards obtaining driving frequencies for frequency domain simulation experiments.

In this section we present an algorithm which obtains a local optimum to the problem. The algorithm is polynomial in the number of factors p and the degree of the polynomial k. A simple illustrative example follows the formal algorithm presentation.

It is clear from Cramer's Rule [5] that since the mixed integer linear program has only integer coefficients for all its variables, then the solution it finds will be rational (i.e. can be represented as fractions with integer numerators and denominators). Therefore, we can form a one-to-one relationship between the optimal driving frequencies and a set of integers. These integers, which represent the numerators of the fractions, together with a common integer denominator, will completely define the driving frequencies. We will call these integer numerators <u>driving integers</u> or driving numerators (analogously we will have <u>term indicator integers</u> or term indicator numerators). Our only constraint is that the term indicator integers for the different terms in the model do not coincide.

Our algorithm attempts to find, in sequence, p driving integers. It does this by first arbitrarily choosing a driving integer, say h_1, which will represent the numerator of the first driving frequency. To obtain the second driving integer h_2, the algorithm considers subsequent integers greater than h_1 until one is found such that the updated set of term indicator integers contains no duplicate integers. This process is repeated until p driving integers $(h_1,h_2,...,h_p)$ have been selected.

The final stage of the algorithm is to select an integer denominator. This will completely define all the driving frequencies, hence all the term indicator frequencies. The algorithm selects this integer denominator by first looking at the integer $B=2\mathrm{card}(V_p)+2$ as a possible candidate, where V_p is the set of term indicator integers resulting from the driving integers $h_1,h_2,...,h_p$. If the set $V_p(B)=\{\nu/B\in[0,.5]: \nu\in V_p$, B a positive integer$\}$ has no duplicate elements, then this is the set of driving frequencies. If $V_p(B)$ has one or more duplicate elements, then B is increased by one until no duplicate elements remain. Note: If $\min\{|.5-\nu|:\nu\in V_p(B)\}<\min\{|\nu_1-\nu_2|:\nu_1,\nu_2\in V_p(B)\}$, then it may be desireable to increase B further, especially if one wants to space the term indicator frequencies away from .5.

The following terms are defined for a given k:

$h_i\in\{1,2,...\}$ represents the i^{th} driving integer (i.e. the numerator of the i^{th} driving frequency) $(i=1,2,...,p)$

ω_i represents the i^{th} driving frequency ($0\le\omega_i\le.5$ for all $i=1,2,...,p$)

$V_i=\{(\nu\in\{1,2,...\}$: ν is a term indicator integer obtained from the driving integers $h_1,...,h_i)$ $(i=0,1,...,p)\}$, $(V_0=\phi)$

$B\in\{1,2,...\}$ is a positive integer such that $\omega_i=h_i/B$ for $i=1,2,...,p$ (i.e. the common denominator which defines the relationship between the driving integers and the driving frequencies)

The algorithm is as follows.

ALGORITHM

STAGE 0: Set k,p to their given values for the problem.

STAGE 1: Arbitrarily choose an integer h_1

(in practice, h_1 should be chosen to be at most one half the number of element in V_p. This will give the algorithm a reasonable chance to fill the integers from 1 to card(V_p) with the remaining p-1 driving integers).

V_1 is the set of term indicator frequencies determined by h_1.

For i=2 to p,

Solve: $h_i = \text{Min } h$

such that V_i has no duplicate integer elements

(once h_i is obtained, V_i is easily enumerated)

STAGE 2: Solve: $B = \text{Min } \mathcal{B}$

such that $\mathcal{B} \geq 2\text{card}(V_p)+2$ and there are no duplicate elements in the set $V_p(\mathcal{B})$, where

$$V_p(\mathcal{B}) = \begin{cases} \mathcal{V}/\mathcal{B} & \text{where } \mathcal{V}\in V_p \text{ and } 0\le \mathcal{V}/\mathcal{B} \le .5 \\ 1-\mathcal{V}/\mathcal{B} & \text{where } \mathcal{V}\in V_p \text{ and } .5<\mathcal{V}/\mathcal{B}\le 1.0 \\ \mathcal{V}/\mathcal{B}-1 & \text{where } \mathcal{V}\in V_p \text{ and } 1.0<\mathcal{V}/\mathcal{B} \end{cases}$$

and all elements of $V_p(\mathcal{B})$ must lie in $(0,.5)$.

Note: If it is desireable to space the term indicator frequencies away from .5, then one can solve the following program:

$$B = \text{Min} \quad \mathcal{B}$$

such that $\mathcal{B}\ge 2\text{cardinality}(V_p)+2$, there are no duplicate elements in the set $V_p(\mathcal{B})$, $\lfloor B/2 \rfloor \notin V_p(\mathcal{B})$ and all elements of $V_p(\mathcal{B})$ must lie in $(0,.5)$.

Set $\omega_i = h_i/B$ for $i=1,2,\ldots,p$ and stop.

5. AN EXAMPLE:

The operation of the algorithm can best be understood by working through an example. Suppose we have a quadratic response (k=2). As previously outlined, the basic strategy in stage 1 is to start with a positive integer h_1, which will be the numerator for the first driving frequency, and eliminate from consideration $2h_1$ corresponding to the quadratic term associated with the first frequency. The smallest candidate for the second numerator is h_1 plus the smallest integer not eliminated; call this number $\tilde{h}_2$. Now compute $2\tilde{h}_2$ and $\tilde{h}_2 \pm h_1$ corrseponding to quadratic and interaction terms, respectively. If these are all distinct, then $\tilde{h}_2$ is accepted as a numerator and now called simply h_2. The numbers $2h_2$ and $h_2 \pm h_1$ are eliminated from further consideration. If there is confounding with $2\tilde{h}_2$ and $\tilde{h}_2 \pm h_1$, then $\tilde{h}_2$ is replaced by h_1 plus the next largest integer not eliminated for h_1. This is continued until a numerator has been selected for each factor. If $k>2$, then there will be more term indicator integers which can be confounded, hence it will be necessary to perform more checking to ensure that all the term indicator integers are distinct.

We will consider a prospective quadratic response model in three factors (i.e. $k=2$ and $p=3$). Start with $h_1=1$ so that $V_1=\{1,2\}$ are eliminated from consideration ($h_1=1$ is the term indicator integer for x_1 and $2h_1=2$ is the term indicator integer for x_1^2). The smallest candidate for h_2 is now $1+3=4$. This numerator adds 8 (quadratic), 3 and 5 (interaction) to the list of term indicator integers, which is now $V_2=\{1,2,3,4,5,8\}$. Finally, the smallest candidate for the third numerator, h_3, is $4+6=10$, since 6 is the smallest number not eliminated. The term indicator integers added by 10 are $\{10,20,6,14,9,11\}$, so the final list is $V_3=\{1,2,3,4,5,6,8,9,10,11,14,20\}$ (see Table 1). The driving integers are $\{1,4,10\}$. Note: optimal numerators would result if the third numerator had been chosen to be 11 instead of 10, resulting in the term indicator integers $\{1,2,3,4,5,7,8,10,11,12,15,22\}$.

Stage 2 is finding the smallest integer denominator, B, corresponding to the numerators obtained in stage 1 such that the indicator frequencies h_i/B are all less than .5. Clearly a denominator which is feasible is one obtained by adding 2 to twice the largest term indicator integer. However, this is usually not best for a given set of numerators. Stage 2 attempts to exploit the aliasing effect to reduce the size of the denominator. For a second degree model, the largest term indicator integer corresponds to the quadratic term of the largest driving integer. A small denominator will result if it can be aliased to a small number. This corresponds to finding a point to "fold" the numerators about so that all term indicator integers larger than the fold point fall on vacant numbers (i.e.integers which are not term indicator integers).To illustrate stage 2 for the example above, first try to alias 20 to 7 (see figure 1). This corresponds to a fold about 13½, and 14 is also aliased to 13. The denominator corresponding to this is 20+7=27, and the bandwidth is 1/54. This is the distance from 13/27 to 1/2. The next integer we can alias 20 to is 12, resulting in a fold point of 16 and a denominator of 32 (see figure 2). The bandwidth for this denominator is 1/32. Thus, the driving frequencies are {1/32,4/32,10/32} and the term indicator frequencies are {1/32,2/32,3/32,4/32,5/32,6/32,8/32,9/32,10/32,11/32,12/32,14/32}. The best denominator corresponding to optimal numerators {1,4,11} is 28, obtained by the fold point of 14.

If the algorithm starts with a numerator other than 1 then there are other possibilities. If the algorithm starts with numerator 2, then driving integers of {2,3,10} result, with term indicator integers as in figure 3. The best fold for these term indicator integers is at 14½ and the denominator is 29, only slightly worse than the optimal denominator of 28. Some starting values require no fold. If the algorithm starts at 4, then {4,5,7} are the driving integers, 14 is the largest term indicator integer, and the denominator is 30 (see figure 4).

The algorithm may be modified to allow skipping eligible driving integers. For example, if the second eligible driving integer is chosen at each step, then (starting at 1) the resulting driving integers are {1,5,8}. Folding at 14 yields a denominator of 28, the optimal denominator (see figure 5). This variation in the algorithm can greatly improve the spacing between the resulting term indicator frequencies.

Appendix 3 gives sets of driving frequencies for various number of factors for a prospective quadratic response model. The frequencies for 2 to 7 factors are optimal, assuming a full model (all terms present). The frequencies for 8 to 21 factors are not necessarily optimal, but are the best available at present. The table gives numerators and denominators separately, and for the factors with optimal spacing, all known alternative sets of frequencies are given. For example, if there are 5 factors, the experimenter may choose driving frequencies $\{1/69, 4/69, 13/69, 19/69, 29/69\}$, $\{4/69, 5/69, 7/69, 20/69, 26/69\}$, or $\{2/69, 5/69, 11/69, 25/69, 26/69\}$. These are the only ones with optimal spacing; the choice of any other driving frequencies results in a narrower bandwidth.

Table 1 Algorithm for Frequency Selection,

First Numerator = 1

Step	Linear	Quadratic	Interaction
1	1	2	
2	4	8	3 5
3	10	20	6 9 11 14

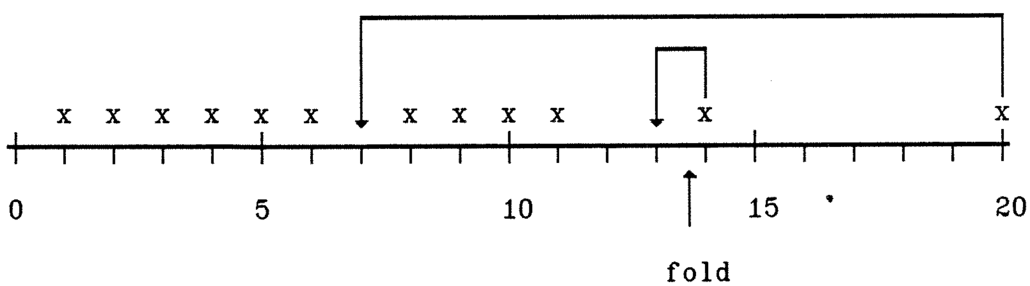


Figure 1 Denominator = 27 by "Folding" at 13½
Bandwidth = .5 - 13/27 = 1/54

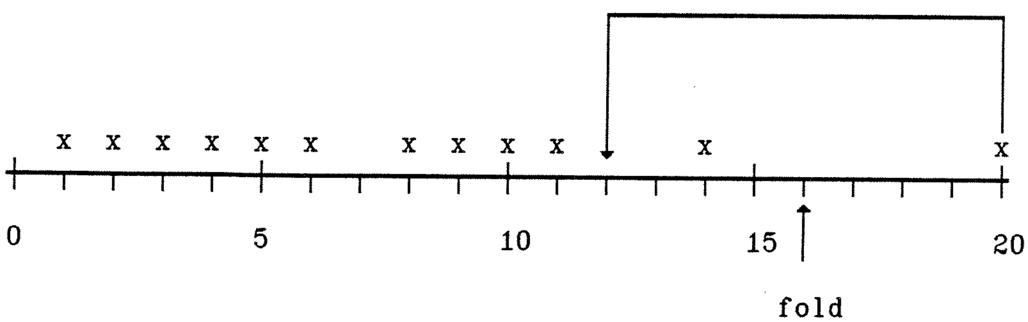
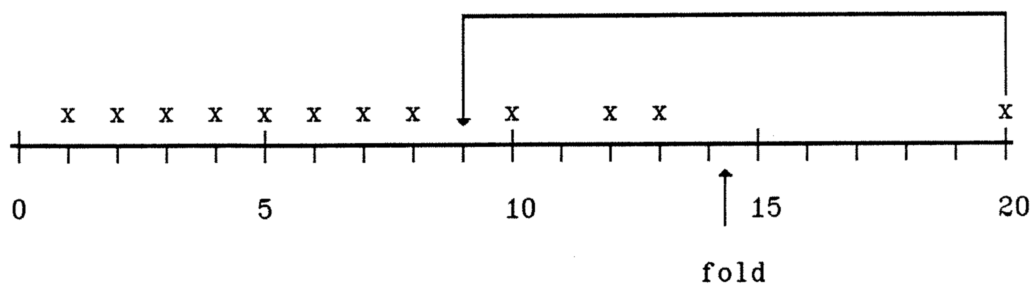


Figure 2 Denominator = 32 by Folding at 16.
Bandwidth = 1/32



<u>Figure 3</u> Starting Algorithm with Numerator = 2
Denominator = 29 by Folding at 14½
Bandwidth = 1/29

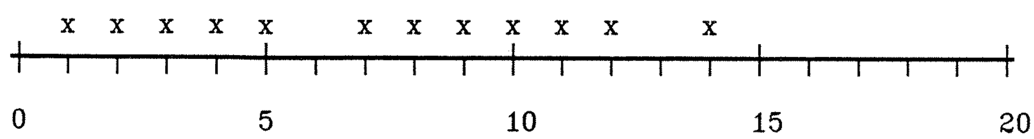
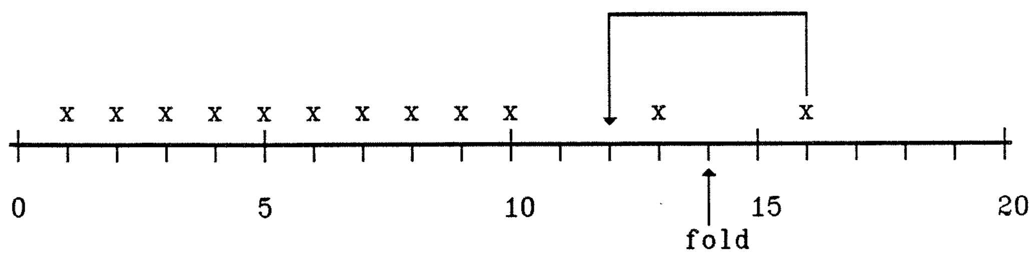


Figure 4 Starting Algorithm with Numerator = 4
No Folding is Required
Denominator = 30, Bandwidth = 1/30.



<u>Figure 5</u> Driving Numerators = {1,5,8}
Denominator = 28 by Folding at 14
Bandwidth = 1/28 (Optimal).

A summary of the 2 stages for the above example is given below:

STAGE 1:

We arbitrarily chose $h_1=1$.

Therefore, $V_1=\{1,2\}$.

i=2: If $h_2=2$, $V_2=\{1,2,2,4,1,3\}$ which has duplicate elements.

If $h_2=3$, $V_2=\{1,2,3,6,2,4\}$ which has duplicate elements.

If $h_2=4$, $V_2=\{1,2,4,8,3,5\}$ which has no duplicate elements.

Therefore, $h_2=4$ and $V_2=\{1,2,3,4,5,8\}$.

i=3: If $h_3=5$, $V_3=\{1,2,3,4,5,8,5,10,4,6,1,9\}$ which has duplicate elements.

If $h_3=6$, $V_3=\{1,2,3,4,5,8,6,12,5,7,2,10\}$ which has duplicate elements.

If $h_3=7$, $V_3=\{1,2,3,4,5,8,7,14,6,8,3,11\}$ which has duplicate elements.

If $h_3=8$, $V_3=\{1,2,3,4,5,8,8,16,7,9,4,12\}$ which has duplicate elements.

If $h_3=9$, $V_3=\{1,2,3,4,5,8,9,18,8,10,5,13\}$ which has duplicate elements.

If $h_3=10$, $V_3=\{1,2,3,4,5,8,10,20,9,11,6,15\}$ which has no duplicate elements.

Therefore, $h_3=10$ and $V_3=\{1,2,3,4,5,6,8,9,10,11,14,20\}$.

STAGE 2: For $B=\max\{21,26\}=26$,

$V_3/26= \{1,2,3,4,5,6,8,9,10,11,14,20\}/26$ which has duplicate elements since $20/26$ is aliased to $6/26$.

$V_3(27) = \{1,2,3,4,5,6,8,9,10,11,14,20\}/27$ which has no duplicate elements if it is not desireable to space the term indicator frequencies away from .5 (see the note in stage 2 of the algorithm description in Section 3). If it is desireable to space away from .5, then $\lfloor 27/2 \rfloor = 13$ which is the alias of 14.

$V_3/28 = \{1,2,3,4,5,6,8,9,10,11,14,20\}/28$ which has duplicate elements since 20/28 is aliased to 8/28.

$V_3/29 = \{1,2,3,4,5,6,8,9,10,11,14,20\}/29$ which has duplicate elements since 20/29 is aliased to 9/29.

$V_3/30 = \{1,2,3,4,5,6,8,9,10,11,14,20\}/30$ which has duplicate elements since 20/30 is aliased to 10/30.

$V_3/31 = \{1,2,3,4,5,6,8,9,10,11,14,20\}/31$ which has duplicate elements since 20/31 is aliased to 11/31.

$V_3/32 = \{1,2,3,4,5,6,8,9,10,11,14,20\}/32$ which has no duplicate.

Therefore, $(\omega_1,\omega_2,\omega_3)=(1/32,4/32,10/32)$ or $(\omega_1,\omega_2,\omega_3)=(1/27,4/27,10/27)$ depending on whether full spacing is desired at the endpoints.

In the above example, it is not clear that the global optimum has indeed been reached. The minimum spacing for these driving frequencies is 1/32. If the 12 term indicator frequencies could be evenly spaced, the minimum spacing would be 1/26.

The algorithm is an easy method to obtain driving frequencies which appear to be close to optimal. Variations in the algorithm, such as forcing certain minimum jumps in driving integers at any point in stage 1, rather than for just the first driving integer, can greatly increase the minimum spacing. Appendix 3 gives for quadratic and cubic response models ($k=2$ and $k=3$) the largest spacings we have found using this variation. We have proven, using direct enumeration and by solving the mixed integer linear program in Section 2, that the spacings, for $p=2$ up to $p=7$ are all optimal (solutions to the mixed integer linear program were obtained using MPSX on the Cornell Supercomputer). We believe that the spacings for $p=8$ and 9 may be optimal, though this has not been proven. We also believe that the spacings for $p=10$ up to 21 can still be improved. Appendix 3 contains the driving frequencies which yield the best available spacings for $p=2$ up to $p=21$, for $k=2$, and for $p=2$ up to 11, for $k=3$. Appendix 3 also gives the minimum spacing necessary to set the run length and spectral estimation window [7].

It is often the case that low indicator frequencies are required. If driving frequencies are desired with the added constraint that all term indicator frequencies lie in the interval $[0,\delta]$, where $0<\delta\leq.5$, then the integer numerators given in Appendix 3 can be used with the integer denominator $B=2h_p/\delta$ for $k=2$, and $B=3h_p/\delta$, for $k=3$, to space the term indicator frequencies in this new interval. One could also modify the last stage of the algorithm to incorporate this constraint in selecting the integer denominator B, such that an attempt is also made to evenly space the term indicator frequencies in $[0,\delta]$.

It may also be possible to modify the algorithm such that all term indicator frequencies fall into prespecified parts of the interval $[0,.5]$. This would give a possible way of dealing with gain or memory in a system, by restricting all term indicator frequencies to parts of the interval $[0,.5]$ where the gain or memory effect is approximately the same (note: some care should be taken when specifying these subintervals, since one can specify parts of the interval [0,.5] where it is impossible for all the term indicator frequencies to be forced to lie).

The algorithm can also be modified for problems where only interaction terms, or any subset of polynomial terms, are assumed to be present. For example, after performing an initial run to screen out insignificant terms, the algorithm can be modified to select driving frequencies such that only term indicator frequencies of the significant terms must be evenly spaced in the interval [0,.5]. This would tend to widen the spacing between the term indicator frequencies, hence ultimately reduce the run size n for a desired variance level of the spectral estimators.

The examples given above demonstrate some of the ways in which the algorithm can be modified. As frequency domain simulation experiments become more widely implemented, more sophisticated modifications to the frequency selection algorithm should be developed to handle their driving frequency and term indicator frequency requirements.

APPENDIX 1:

We will prove by induction on p and k that there are

$$v(p,k)= pk + \sum_{r=1}^{k-1} \sum_{j=1}^{r} 2^j \binom{r}{j}\binom{p}{j+1} +2$$

term indicator frequencies which need to be spaced. Note that the 2 added to the end of the formula for v(p,k) corresponds to the endpoint frequencies 0 and .5. Therefore, it is sufficient to prove this formula for v(p,k)-1.

Let p=k=1.

Clearly, there is only one term, which agrees with n(1,1)=1.

Let us fix p and perform the induction on k. We notice that

$$p + \sum_{j=1}^{k} 2^j \binom{k}{j}\binom{p}{j+1}$$

terms will be added if we increase k by one, where the first value represents terms with only one element (i.e. of the form $(k+1)\omega_i$, where i=1,2,...,p, and ω_i is the driving frequency for the i^{th} factor) and the second value represents terms with between 2 and k+1 elements.

Therefore, we have

$$v(p,k) + p + \sum_{j=1}^{k} 2^j \binom{k}{j}\binom{p}{j+1} = p(k+1) + \sum_{r=1}^{k} \sum_{j=1}^{r} 2^j \binom{r}{j}\binom{p}{j+1}$$

$= v(p,k+1)$ as desired.

Let us fix k and perform the induction on p. We notice that

$$k + \sum_{r=1}^{k-1} \sum_{j=1}^{r} 2^j \binom{r}{j}\binom{p}{j}$$

terms will be added if we increase p by one, where the first value represents terms with only one element (i.e. of the form $(j)\omega_{p+1}$, where $j=1,2,\ldots,k$, and ω_{p+1} is the driving frequency for the $p+1^{th}$ factor) and the second value represents terms with between 2 and k elements.

Therefore, we have

$$v(p.k) + k + \sum_{r=1}^{k-1} \sum_{j=1}^{r} 2^j \binom{r}{j}\binom{p}{j} = (p+1)k + \sum_{r=1}^{k-1} \sum_{j=1}^{r} 2^j \binom{r}{j}\binom{p+1}{j+1}$$

$= v(p+1,k)$ as desired.

Since the order in which we increase p and k does not affect the number of terms, simultaneous induction on p and k is not necessary.

APPENDIX 2:

We will prove that for $p=2$ and k arbitrary, the driving frequencies $\omega_1=k/(2k(k+1)+2)$ and $\omega_2=(k+1)/(2k(k+1)+2)$ evenly space all term indicator frequencies in the interval $[0,.5]$.

First, we will prove by induction on k, that the number of term indicator frequencies is $k(k+1)$.

For $k=1$, there are $2=1(2)$ term indicator frequencies.
Assume there are $k_1(k_1+1)$ term indicator frequencies, for some $k=k_1$. If we add a degree to our polynomial model given by (1.1) (i.e. $k=k_1+1$), then we will add $2k_1+2$ term indicator frequencies ($(k+1)\omega_1,(k+1)\omega_2$, and $\kappa\omega_1\pm(k+1-\kappa)\omega_2$ for $\kappa=1,2,\ldots,k$). Therefore, there will be $k_1(k_1+1)+2(k_1+1)=(k_1+2)(k_1+1)$ term indicator frequencies for $k=k_1+1$.
Therefore, we have proven by induction on k that there are $k(k+1)$ term indicator frequencies for $k\geq 2$ and $p=2$.

Without loss of generality, we will only consider the numerators of ω_1 and ω_2 and prove that the $k(k+1)$ term indicator integers will be the integers 1 to $k(k+1)$, with no aliasing required.

Let h_1 be the integer numerator of ω_1 and h_2 be the integer numerator of ω_2 $(h_1<h_2)$. The largest term indicator integer will be $k(h_2)=k(k+1)$. Therefore, $h_2=k+1$. The second largest term indicator integer must be $(k-1)h_2+h_1=k(k+1)-1$. This implies $h_1=k$.

We will now show that all the term indicator integers obtained from $h_1=k$ and $h_2=k+1$ exactly cover all the integers from 1 to $k(k+1)$ exactly once, with no aliasing required.

If j is even, there are $2j$ j^{th} degree term indicator integers.

They are $\{jk, jk+1, \ldots, jk+j\} \cup \{(j-2)k-1, (j-2)k+(j-1)\}$

$\cup \{(j-4)k-2, (j-2)k+(j-2)\}$

$\cup \ldots$

$\cup \{2k-((j/2)-1), 2k+((j/2)+1)\}$

$\cup \{(j/2)\}$

Then we have $j+1$ odd, hence there are $2j+2$ j^{th} degree term indicator integers.

They are $\{(j+1)k, (j+1)k+1, \ldots, (j+1)k+(j+1)\}$

$\cup \{((j-1)k-1, (j-1)k+(j)\}$

$\cup \{(j-3)k-2, (j-3)k+(j-1\}$

$\cup \ldots$

$\cup \{3k-((j/2)-1), 3k+(((j+2)/2)+1)\}$

$\cup \{k-((j)/2), k+((j+2)/2)\}$

Therefore, for $j=1$, we have the term indicator integers $\{k, k+1\}$.

For $j=2$, we have the term indicator integers $\{2k, 2k+1, 2k+2, 1\}$.

If we order these, we get $\{1, k, k+1, 2k, 2k+1, 2k+2\}$.

For $j=3$, we have the term indicator integers $\{3k, 3k+1, 3k+2, 3k+3, k+2, k-1\}$. If we order these, we get

$\{1, k-1, k, k+1, k+2, 2k, 2k+1, 2k+2, 3k, 3k+1, 3k+2, 3k+3\}$.

For $j=4$, we have the term indicator integers $\{4k, 4k+1, 4k+2, 4k+3,$

$4k+4, 2k+3, 2k-1, 2\}$. If we order these, we get

$\{1, 2, k-1, k, k+1, k+2, 2k-1, 2k, 2k+1, 2k+2, 2k+3, 3k, 3k+1, 3k+2, 3k+3, 4k,$
$4k+1, 4k+2, 4k+3, 4k+4\}$.

If we continue in this manner, we get for k even, the term indicator integers $\{1, 2, \ldots, (k/2), k-(k/2)+1, k-(k/2)+2, \ldots, k+(k/2), 2k-(k/2)+1, 2k-(k/2)+2, \ldots, 2k+(k/2), \ldots, k(k+1)\}$ which is exactly the integers from 1 to $k(k+1)$.
Similarly, for k odd, we get the set of term indicator integers to be exactly the integers from 1 to $k(k+1)$.

Therefore, if $h_1=k$ and $h_2=k+1$, the resulting term indicator integers are exactly the integers from 1 to $k(k+1)$, where each indicator integer appears exactly once. This implies that $\omega_1=h_1/(2k(k+1)+2)$ and $\omega_2=h_2/(2k(k+1)+2)$ are driving frequencies which result in the term indicator frequencies being evenly spaced in $[0,.5]$.

APPENDIX 3:

The following list of driving frequencies yield the best available spacings for quadratic response models (k=2):

# of Factors (p)	Driving Frequencies	Minimum Spacing
2	(1,4)/14 (2,3)/14	1/14
3	(1,5,8)/28 (1,4,11)/28 (3,4,13)/28 (3,5,12)/28 (8,9,13)/28 (9,11,12)/28	1/28
4	(1,4,10,17)/46 (6,8,9,13)/46 (2,3,11,18)/46 (4,5,7,20)/46 (3,5,12,16)/46 (2,9,10,15)/46 (1,6,16,19)/46	1/46
5	(1,4,13,19,29)/69 (4,5,7,20,26)/69 (2,5,11,25,26)/69	1/69
6	(1,11,28,31,35,49)/103 (3,4,13,28,40,42)/103 (4,9,10,21,37,44)/103 (6,32,40,42,43,47)/103 (8,15,18,20,29,42)/103 (10,12,15,21,28,29)/103 (11,27,35,36,48,50)/103 (16,23,24,43,45,49)/103 (20,24,30,42,45,47)/103	1/103
7	(1,4,19,31,44,53,60)/130 (4,7,9,24,30,49,59)/130 (1,9,14,21,40,46,57)/130 (1,7,18,22,27,57,60)/130 (2,3,12,29,37,50,57)/130 (7,9,17,20,32,62,63)/130 (3,7,19,28,30,43,48)/130 (1,10,16,29,34,37,41)/130 (3,8,10,27,41,42,63)/130 (2,17,22,23,30,33,59)/130 (6,9,19,20,36,41,43)/130 (6,31,40,47,51,61,64)/130 (3,21,41,49,50,54,64)/130 (10,11,36,44,49,51,63)/130 (8,33,47,48,51,53,60)/130 (11,16,17,20,46,59,61)/130 (10,11,18,23,37,53,62)/130 (19,29,30,33,54,56,61)/130 (11,14,23,24,29,31,50)/130 (21,27,34,51,56,59,60)/130 (12,17,21,27,40,47,58)/130 (31,32,38,40,43,53,57)/130 (20,33,37,42,43,58,61)/130 (23,28,38,47,49,50,63)/130	1/130
8	(10,16,29,33,38,40,41,75)/168	1/168
9	(8,30,33,39,40,44,57,59,94)/209	1/209
10	(10,12,13,27,31,59,65,94,101,110)/268 (2,15,22,33,47,56,75,83,121,126)/268	1/268
11	(19,21,22,26,32,46,55,105,117,135,166)/340	1/340
12	(27,44,56,57,59,63,77,102,124,150,155,219)/448	1/448
13	(29,37,39,43,44,55,64,77,96,166,208,211,257)/565	1/565
14	(29,42,53,57,62,63,65,79,109,140,210,242,269,310)/675	1/675
15	(20,28,32,41,42,47,58,65,145,176,179,222,247,298,342)/780	1/780
16	(29,38,46,48,49,53,71,83,110,176,217, 223,302,358,372,427)/942	1/942
17	(37,46,63,66,67,73,78,91,122,160,162, 239,309,325,377,430,495)/1052	1/1052
18	(37,53,78,86,87,89,93,107,117,155,190, 236,255,318,377,453,475,574)/1208	1/1208
19	(18,38,42,45,47,53,70,86,96,163,197, 218,299,300,372,429,469,589,620)/1398	1/1398
20	(15,29,37,40,41,46,64,100,148,167,195, 257,291,329,382,447,535,608,705,707)/1588	1/1588
21	(31,32,48,57,59,67,71,100,113,206,221, 266,315,389,407,493,570,576,716,767,774)/1834	1/1834

The following list of driving frequencies yield the best available spacings for cubic response models (k=3):

# of Factors (p)	Driving Frequencies	Minimum Spacing
2	(2,5)/26 (3,4)/26	1/26
3	(3,5,22)/70 (4,9,15)/70	1/70
	(5,8,17)/70 (15,18,23)/70	
4	(17,20,21,32)/152	1/152
5	(27,37,45,48,97)/319	1/319
6	(25,34,37,39,159,192)/600	1/600
7	(10,14,23,68,143,219,336)/1023	1/1023
8	(10,16,17,72,142,227,462,575)/1801	1/1801
9	(10,14,23,68,143,219,336,687,923)/2808	1/2808
10	(15,16,19,40,122,251,402,711,1165,1314)/4097	1/4097
11	(2,5,21,55,130,287,455,788,1160,1515,1769)/5522	1/5522

ACKNOWLEDGEMENT: The computer work for the research contained in this report was done through the Cornell National Supercomputer Facility at Cornell University, funded by the NSF. The authors would also like to thank David Jensen for his assistance in using the MPSX optimization package, as well as Doug Morrice, Robert Sargent and Enver Yucesan for their helpful comments.

BIBLIOGRAPHY:

[1] Chatfield, C., THE ANALYSIS OF TIME SERIES: AN INTRODUCTION, 3rd Edition, Chapman and Hall, New York, 1984.

[2] Cogliano, V. "Sensitivity Analysis and Model Identification in Simulation Studies: A Frequency Domain Approach", Ph.D. Thesis, School of OR&IE, Cornell University, 1982.

[3] Garey, M.R., Johnson, D.S., COMPUTERS AND INTRACTABILITY: A GUIDE TO THE THEORY OF NP-COMPLETENESS , H. Freeman, San Francisco, 1978.

[4] Jenkins, G.M., Watts, D.G., SPECTRAL ANALYSIS AND ITS APPLICATIONS, Holden-Day, San Francisco, 1968.

[5] Murty, K., LINEAR AND COMBINATORIAL PROGRAMMING , Wiley and Sons, New York, 1976.

[6] Schruben, L.W., " Simulation Optimization Using Frequency Domain Methods ", Proceedings of the 1986 Winter Simulation Conference, Washington, D.C..

[7] Schruben, L.W., Cogliano, V., " An Experimental Procedure for Simulation Response Surface Model Identification ", Communications of the ACM, Vol. 30(8), p.716-730, August 1987.