Optimization and Engineering, 2, 5–29, 2001


# Mixed Variable Optimization of the Number and Composition of Heat Intercepts in a Thermal Insulation System

MICHAEL KOKKOLARAS
*Department of Mechanical Engineering, The University of Michigan, Ann Arbor, MI, USA*
*email: michailk@engin.umich.edu*
CHARLES AUDET
*Département de Mathématiques et de Génie Industriel, École Polytechnique de Montréal, Canada*
*email: charlesa@crt.umontreal.ca*
J.E. DENNIS, JR.
*Department of Computational and Applied Mathematics, Rice University, Houston, TX, USA*
*email: dennis@caam.rice.edu*

*Received June 30, 2000; Revised December 27, 2000*

**Abstract.** In the literature, thermal insulation systems with a fixed number of heat intercepts have been optimized with respect to intercept locations and temperatures. The number of intercepts and the types of insulators that surround them were chosen by parametric studies. This was because the optimization methods used could not treat such *categorical* variables. Discrete optimization variables are categorical if the objective function or the constraints can not be evaluated unless the variables take one of a prescribed enumerable set of values. The key issue is that categorical variables can not be treated as ordinary discrete variables are treated by relaxing them to continuous variables with a side constraint that they be discrete at the solution.

A new mixed variable programming (MVP) algorithm makes it possible to optimize directly with respect to mixtures of discrete, continuous, and categorical decision variables. The result of applying MVP is shown here to give a 65% reduction in the objective function over the previously published result for a thermal insulation model from the engineering literature. This reduction is largely because MVP optimizes simultaneously with respect to the *number* of heat intercepts and the *choices from a list* of insulator types as well as intercept locations and temperatures. The main purpose of this paper is to show that the mixed variable optimization algorithm can be applied effectively to a broad class of optimization problems in engineering that could not be easily solved with earlier methods.

**Keywords:** optimization, thermal insulation, heat intercepts, categorical variables, mixed variable programming (MVP), pattern search algorithm

## 1. Introduction

Thermal insulation systems use heat intercepts to minimize the heat flow from a hot to a cold surface. In figure 1, the cooling temperature $T_i$ is a control imposed at the $i = 1, 2, \dots, n$ locations $x_i$ to "intercept" the heat. The design configuration of such a multi-intercept thermal insulation system is defined by the number of intercepts, their locations, temperatures, and the types of insulators placed between each pair of adjacent intercepts. We will refer to three insulator types: a specific material chosen from a list, a vacuum, and so-called

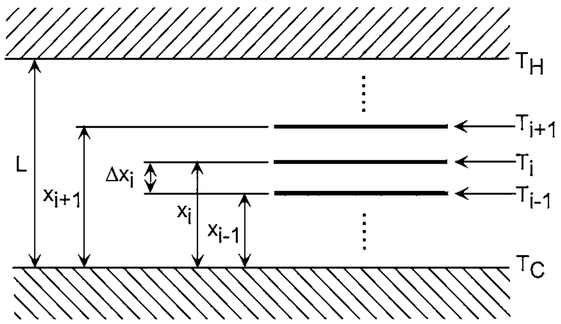


*Figure 1.* Schematic of a thermal insulation system.

multi-layer insulators. The latter consists of a number of thin films of insulating materials stacked next to each other. The number of films is called the density of a multi-layer insulator.

Hilal and Boom (1977) considered cryogenic engineering applications in which either load-bearing insulators are required in the construction of dewars, or mechanical struts are necessary in the design of solenoids for superconducting magnetic energy storage systems. In such applications, vacuum is ruled out as an insulator because the presence of material is necessary all the way between the hot and cold surfaces in order to support the mechanical loads. They formulated an objective function based on a power minimization principle to optimize the configuration of a thermal insulation with respect to the locations and the temperatures of the intercepts.

Hilal and Boom used a gradient-based optimization algorithm that could not handle categorical variables like the number of intercepts and choices of insulators between each adjacent pair of intercepts. Discrete optimization variables are categorical if the objective function or the constraints can not be evaluated unless the variables take one of a prescribed enumerable set of values. Hilal and Boom chose these categorical variables by taking the best values they found as they fixed the number of intercepts and the choice of insulators to sensible selected values and then solved for the resulting optimal temperatures and intercept locations. Thus, they used a parametric study to pick the categorical variables.

Specifically, Hilal and Boom solved the nonlinear programs obtained by fixing the number of intercepts $n$ to $1, 2,$ and $3$. Moreover, they considered only one specified uniform choice for all the layers between the intercepts in each run. After obtaining the minimum refrigeration power for a given type of insulator and a fixed number of intercepts, they performed a system cost optimization study with respect to the length $L$ of the solenoid.

Hilal and Eyssa (1980) considered a variable cross section for the mechanical supports and reported lower optimal power values than those obtained with a constant cross section.

A few years later, Chato and Khodadadi (1984) considered applications in which mechanical supports between the cold and hot ends of the thermal insulation system are optional, i.e., where vacuum is an option as an insulator type. They adopted more general models of heat transfer and used a slightly different objective function based on the entropy principle formulated by Bejan (1979). An interesting difference between their work and that of Hilal and Boom is that instead of considering explicitly different choices of insulators (i.e., different functions for the effective thermal conductivity), they assumed a general parameterized exponential relation, and ran their algorithm with different combinations of parameters in order to determine the optimal location and temperature for each intercept. We emphasize that they neither intended to represent different types of insulators by adjusting the above parameters, nor did they consider different thermal conductivities for the layers between the intercepts.

Cryogenic systems of space borne magnets have been considered more recently. The insulation efficiency of a space borne system should be high so that the available liquid helium used for cooling the intercepts evaporates with a minimal rate during the mission. Musicki et al. (1989) optimized the inlet temperatures and flow rates of the liquid helium for a specified number of intercepts and insulator thicknesses. Yamaguchi et al. (1991) studied the effect of the number of intercepts and the insulator types (by varying the density of the multi-layers insulators used as the insulating material) on the temperature distribution and the resulting heat losses. Li et al. (1989) considered the use of liquid nitrogen and neon instead of liquid helium as the cooling media for the intercepts and compared different insulator types including multi-layer insulators.

These references use models that vary in fidelity and/or geometry, but all of the associated optimization approaches share an important limitation: categorical variables, such as the number of intercepts or the type of insulators surrounding them, are not treated as optimization variables but as parameters. This necessitates the use of parametric studies to choose the categorical variables. The obvious drawback of this is that it is costly to make a large study; however it is quite important. Later, we will see that the MVP algorithm does make a choice of insulators that is seen to be obvious in retrospect, but which is not mentioned in previous studies.

Categorical optimization variables are represented here by discrete real values, but we hope by now to have convinced the reader that they differ from ordinary discrete optimization variables in a fundamental way. The key algorithmic issue is that branch and bound type algorithms are out; categorical variables can not be treated as continuous variables with a side constraint that they be discrete at the solution because the models used for the insulation systems cannot return an output value for a "relaxed" input value of, say, 1.5 intercepts or for an insulator that is an arbitrary mixture of, say, steel and aluminum.

Since continuous relaxation techniques like branch-and-bound or branch-and-cut are ubiquitous for problems with both continuous and discrete variables, we decided to not use the name, "nonlinear mixed integer programming". Instead we use the term *mixed variable programming* (MVP) for the case of a mixture of categorical, continuous, and even ordinary discrete variables. The resulting methodology enables us to optimize thermal insulation systems with respect to all the design variables and obtain a decrease in the objective function value by as much as 65%. We use the MVP optimization algorithm, a pattern search for bound constraints. Presently, we can only guarantee convergence to a stationary point when the objective function is continuously differentiable in the continuous variables, but Abramson's thesis (to appear) will extend this result using the Clarke calculus (1990) as in Audet and Dennis (2000b) for the continous case. The algorithm is outlined in

Section 4, and all the details including a convergence analysis were introduced by Audet and Dennis (2000c).

The paper is structured as follows. In the next section, we describe a thermal insulation model from the literature. In Section 3 we show how the categorical variables are included. Section 4 outlines the principal steps of the mixed variable programming algorithm, and in Section 5 we discuss the results of applying the MVP algorithm to the mixed variable design problem. Finally, we summarize the lessons learned, and we conclude that the MVP algorithm holds promise for some previously intractable engineering optimization problems.

## 2. Classical model of thermal insulation systems

The configuration of a thermal insulation system is specified by the number of heat intercepts $n$, their locations and temperatures, and the selection of insulators that surround them. Optimizing the configuration in the continuous variables alone has been studied traditionally through the following nonlinear programming problem:

$$\min_{\mathbf{x},\mathbf{T}} f(\mathbf{x}, \mathbf{T}), \tag{1}$$

where

- $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$ with $x_i$, $i = 1, 2, \ldots, n$, the location of the $i$th intercept,
- $\mathbf{T} = [T_1, T_2, \ldots, T_n]^T$ with $T_i$, $i = 1, 2, \ldots, n$, the temperature of the $i$th intercept, and
- $f : \Re^{2n} \rightarrow \Re$ is the objective function,

subject to the constraints $0 \leq x_1 \leq x_2 \leq \cdots \leq x_n \leq L$ and $T_C \leq T_1 \leq T_2 \leq \cdots \leq T_n \leq T_H$.

Hilal and Boom (1977) studied problem (1) in order to optimize the design of mechanical supports for large superconductive magnets. A schematic of the thermal insulation system for the specific application is depicted in figure 2.

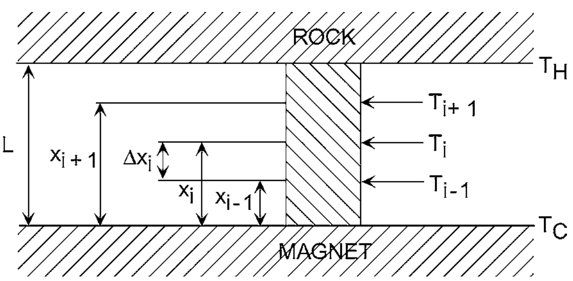


*Figure 2.* Schematic of the mechanical support for superconductive energy storage magnet systems.

The objective function represents the total refrigeration power $P$ required by the system:

$$f = P = \sum_{i=1}^{n} P_i. \tag{2}$$

This is just the sum over all the intercepts of the refrigeration power required to keep the $i$th intercept at a fixed temperature $T_i$:

$$P_i = C_i\left(\frac{T_H}{T_i} - 1\right)(q_{i+1} - q_i), \quad i = 1, 2, \ldots, n, \tag{3}$$

where $C_i$ (a function of temperature) is the thermodynamic cycle efficiency coefficient at the $i$th intercept and $q_i$ of is the heat flow from the $i$th intercept to the $(i - 1)$-th intercept. The cold ($T_C$) and hot ($T_H$) surfaces are referred to for convenience as the 0-th and $(n+1)$-th intercepts, respectively. Similarly, the position of these surfaces are represented by $x_0 = 0$ and $x_{n+1} = L$ and their temperatures are denoted by $T_0 = T_C$ and $T_{n+1} = T_H$. The heat flow $q$ is given by Fourier's law:

$$q\,dx = Ak\,dT, \tag{4}$$

where $A$ (a function of the spatial coordinates in the $z$–$y$ plane perpendicular to the $x$ coordinate) is the constant cross section area and $k$ (a function of temperature) is the effective thermal conductivity of the insulator. Therefore, the heat flow across the $i$th portion of the strut with thickness $\Delta x_i = x_i - x_{i-1}$ is given by

$$q_i = \frac{A}{\Delta x_i}\int_{T_{i-1}}^{T_i} k\,dT, \quad i = 1, 2, \ldots, n+1. \tag{5}$$

Substituting Eq. (5) into (3) yields

$$P_i = AC_i\left(\frac{T_H}{T_i} - 1\right)\left(\frac{\int_{T_i}^{T_{i+1}} k\,dT}{\Delta x_i} - \frac{\int_{T_{i-1}}^{T_i} k\,dT}{\Delta x_{i-1}}\right) \quad i = 1, 2, \ldots, n. \tag{6}$$

The constraints on the locations of the intercepts can be rewritten as $\Delta x_i \geq 0$, $i = 1, 2, \ldots, n+1$ and

$$\sum_{i=1}^{n+1} \Delta x_i = L. \tag{7}$$

Hilal and Boom incorporated the linear constraint (7) into the objective function using the Lagrangian and used a gradient-based method to compute locations and temperatures for various fixed combinations of hot and cold surface temperatures, thermodynamic cycle efficiency coefficients, number of intercepts, and insulators (considering one insulator at a time for the entire mechanical strut); they do not mention explicitly how the bound and temperature constraints are taken into account. Our explanation is that their algorithm will never pick an incompatible cooling temperature because the mathematical model of the underlying physics will yield a larger value.

## 3. Treatment of categorical variables

We will use a recent optimization algorithm for mixed variable programming (Audet and Dennis, 2000c) to include in the optimal configuration problem not only the traditional continuous variables, but also the number of intercepts $n$ and the media that surround them as categorical variables. These media can be different insulators in thermal insulation systems based on the presence of mechanical supports. If mechanical supports are optional, the media can represent vacua or multi-layer insulators; the proposed algorithm only requires a thermal conductivity distribution for each insulator candidate in order to choose the optimum. Figure 3 presents the schematic for a general thermal insulation system to be optimized by the proposed methodology. For the sake of simplicity, and without any loss of generality, it will be assumed in the rest of the paper that the media are represented solely by insulators.

The mathematical model of the thermal insulation system remains the same. However, the set of design variables is extended to include the categorical variables. In this regard, the proposed bound constrained optimization problem is formulated as

$$\begin{array}{ll} \min\limits_{n,\mathbf{I},\mathbf{\Delta x},\mathbf{T}} & \hat{f}(n, \mathbf{I}, \mathbf{\Delta x}, \mathbf{T}) \\ \text{subject to} & h(n, \mathbf{I}, \mathbf{\Delta x}, \mathbf{T}) \le 0 \end{array} \tag{8}$$

with the following optimization variables:

- $n \in N$ is the number of intercepts, with the convention that the cold and hot walls are the 0-th and $n + 1$-th intercepts respectively,
- $\mathbf{I} = [I_1, I_2, \ldots, I_{n+1}]^T$ is the vector of insulators whose $i$th component is the integer assigned to represent the insulator between the $(i - 1)$-th and $i$th intercepts,

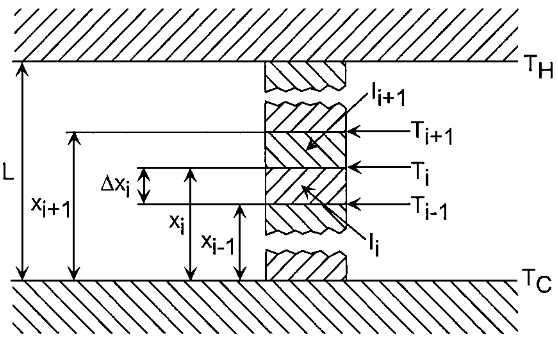


*Figure 3.* Schematic of a general thermal insulation system.

- $\mathbf{\Delta x} \in \Re^n$ is the vector whose $i$th component is the thickness of the $i$th insulator, with the convention that $\Delta x_{n+1}$ is $L - \sum_{i=1}^{n} \Delta x_i$,
- $\mathbf{T} \in \Re^n$ is the vector whose $i$th component is the temperature of the $i$th intercept with the convention that the cold and hot temperatures are $T_0$ and $T_{n+1}$.

The bound constraints $h(n, \mathbf{I}, \mathbf{\Delta x}, \mathbf{T}) \leq 0$ consist of $1 \leq n \leq n_{\max}$, $\Delta x_i \geq 0$ for $i = 1, 2, \ldots, n$ and of $T_1 \geq T_C$ and $T_n \leq T_H$. Technically, the linear constraints $\sum_{i=1}^{n} \Delta x_i \leq L$ and $T_1 \leq T_2 \leq \cdots \leq T_n$ are not handled by the theory presented in Audet and Dennis (2000c). In Abramson (to appear), linear constraints for mixed variable programming are treated using the idea presented in Lewis and Torczon (2000). It involves setting the objective function value equal to $+\infty$ when one of them is violated, and when the current iterate is sufficiently close to the boundary, the polling directions must contain directions tangent to the boundary. However, this is not an issue in the present work since in all the runs of the algorithm, the current iterate is never close to the boundary (these linear constraints are not active at the solution produced by the algorithm: the temperatures are monotonically increasing, $\Delta x_{n+1}$ is much greater than zero).

The dependence of the objective function and the bound constraints on the insulators $\mathbf{I}$ is implicit. For each insulator $I$ there corresponds an effective thermal conductivity function $k(I, T)$ that typically is given in the form of tabulated data over a certain temperature range. Before accepting a certain insulator for a certain location, the algorithm has to check if the insulator's thermal conductivity function is compatible with the temperatures of the surfaces between which it is placed.

An additional difficulty in treating the number of intercepts $n$ as an optimization variable is that this causes the number of optimization variables to possibly vary from iteration to iteration because the dimension of the vectors $\mathbf{I}$, $\mathbf{\Delta x}$, $\mathbf{T}$ depend on $n$. For any value $n$, there are $n + 1$ categorical variables and $2n$ continuous variables. The total number of variables is therefore $3n + 2$. If, for example, $n = 10$, there are 31 variables in addition to the variable $n$; 11 of which are categorical. In Section 5, we give results for the MVP algorithm applied to this more flexible formulation.

## 4. Mixed variable programming (MVP)

Pattern search algorithms were first designed as derivative-free methods for unconstrained minimization of smooth functions. Torczon (1997) proposed a formal definition of this class of algorithms and showed first order optimality results. Lewis and Torczon extended the work to bound constrained optimization (Lewis and Torczon, 1999) and more generally for problems with a finite number of linear constraints (Lewis and Torczon, 2000). In Audet and Dennis (2000c) the bound constrained algorithm is generalized to the mixed variables case, and in Audet and Dennis (2000b) the assumption on the smoothness of the objective function is lifted and appropriate first order optimality results are derived. In Audet and Dennis (2000a), pattern search algorithms are combined with filter methods for general constrained optimization.

The MVP algorithm used in this paper is the one described in Audet and Dennis (2000c). This section describes the implementation used here. Our implementation is not intended to be the most efficient one possible for this problem. Instead, we aim to show the robustness of the algorithm as well as some of the less obvious ways to incorporate the user's insight. We hope to make this more clear in context.

### *4.1. General description of the MVP algorithm*

The underlying structure of a pattern search algorithm for bound constraints is as follows. It is an iterative method that generates a sequence of feasible iterates whose objective function value is nonincreasing. At any given iteration, the objective function is evaluated at a finite number of points on a conceptual mesh in order to try to find one that yields a decrease in the objective function value. The user is free to choose most of these points, and a common aim in this choice is to avoid premature convergence to a "too local" optimizer. We assume that the initial *incumbent solution* $x_0$ yields finite values of the objective function and the constraints.

Any iteration $k$ of a pattern search method is initiated with the incumbent solution $x_k = (x_k^c, x_k^d)$, i.e., the currently best found solution partitioned into its continuous $x_k^c$ and discrete $x_k^d$ components,[1] as well as with the current mesh whose fineness, or resolution, is parameterized by a positive real number $\Delta_k$. The goal of each iteration is to obtain a new incumbent solution on the current mesh whose objective function value is strictly less (by any amount at all) than the current incumbent. On the continuous variable space, the mesh can be written

$$\left\{x_k^c + \Delta_k S z : z \in Z^{n_s}\right\} \tag{9}$$

where $S$ is a $n \times n_s$ (for some $n_s > n$) rational matrix whose nonnegative linear combination of the columns spans the whole continuous space $\Re^n$. For example, $S$ could be $[I \; -I]$ where $I$ is the identity matrix.

Exploration of the mesh is conducted in one or two phases. First, a finite search, free of any other rules imposed by the algorithm, is performed anywhere on the mesh. Any strategy can be used, as long as it searches finitely many points (possibly none). This part of the algorithm has the advantage that the user can put in place any heuristic search he/she might favor for improving the incumbent with the knowledge that, if this fails, the next phase will provide a failsafe.

If the first phase does not succeed in improving the incumbent, the second phase is called. This second phase tries to improve the incumbent solution by exploring nearby mesh points. These points are easily obtained for the continuous variables. For the discrete ones, the user must define a notion of "local optimality". For example, in the present application, the user-defined set of neighbors may contain the solutions where one intercept is added or removed (this is further developed in Section 4.2). A potentially exhaustive (but always finite) exploration in small mesh neighborhoods around $x_k$ and around the points in its set of neighbors is performed.

The first phase (called the SEARCH step) provides flexibility to the method, and in practice, it determines the global quality of the solution. The second phase (called the POLL step) follows stricter rules and guarantees theoretical convergence to a local minimizer of a quality specified by the user. The set of points visited by this phase is referred to as the *poll set*. We will discuss in context how engineering intuition can inform either phase.

There are two possible types of iterations. *Successful* iterations occur when a point having an objective function value less than the incumbent is found in either phase. The incumbent solution is then updated, and the next iterate is initiated with a (possibly) coarser (and possibly different) mesh around the newly found incumbent solution. *Unsuccessful* iterations occur otherwise, and they mean that the current iterate is a local mesh optimizer (i.e., with respect to the poll set). The next iteration is initiated at the same solution (since the incumbent did not change), but with a finer mesh on the continuous variables, and a set of neighbors "closer" (if possible) to the incumbent solution.

### *4.2. The MVP algorithm for thermal insulation systems*

This section gives details of a Matlab 5.3 implementation of the MVP algorithm to solve problem (8).

In order to increase the number of directions for which the convergence results hold, the flexibility of the definition of the spanning set is exploited. At each iteration the spanning set $S$ (that defines the current mesh) is obtained by cycling through the following. The first is constructed from the identity matrix with an additional column whose entries are all equal to $-1$. The second one is composed of $n + 1$ unit vectors where the angle between any two distinct ones is constant (this positive basis is considered in a more general class in Alberto et al. (2000)). The third and fourth ones are the negatives of the first and second.

For this application, we put practically no effort into the SEARCH step. One of the objectives of the paper is to convince the reader that the algorithm works (albeit convergence is slower) without incorporating engineering intuition in a SEARCH step. The SEARCH step here consists of at most a single function evaluation. It is invoked only after a successful iteration that did not modify the values of the discrete variables. The trial search point is then $(x_k^c + 2(x_k^c - x_{k-1}^c), x_k^d)$, which consists in looking further along a successful step.

The POLL step is done as follows. First polling is done by modifying the continuous variables at nearby points on the current mesh. Then polling is done by modifying the discrete variables (and possibly the continuous ones as well) through the user-defined set of neighbors. If the objective function value at one of these last point is not better than the incumbent value, but within the parameter $\xi$ of it then EXTENDED POLL is conducted around this promising point. This means further exploration of the mesh points by modifying the continuous variables only. A way to view this step is that the POLL on the categorical variables is an incomplete local parametric study controlled by the user. If the incomplete evaluation of a local parametric change is promising, as measured by the user's threshold $\xi$, then the algorithm refines the study until this promise is either realized or debunked on the current mesh.

In order to choose the set of neighbors, additional insight in the nature of the data is useful: the initial selection set for the insulators was taken from Hilal and Boom (1977) and included teflon, 6063-T5 aluminum, 304 stainless steel, nylon, low-carbon steel, and epoxy-fiberglass (narmco 570) both in plane and normal cloth. Thermal conductivity data for the above insulators were found in tabulated form in Barron (1966) and Handbook on

Materials (1974). The tabulated data were fitted by cubic splines using Matlab; figure 4(a) depicts the resulting curves for all of the above insulators. It can be seen that aluminum and steel should not be considered for optimum thermal insulation systems unless there are specific reasons. The thermal conductivity behavior of nylon, teflon, and epoxy-fiberglass can be observed much better in figure 4(b). It can be clearly seen that different insulators are superior in different temperature ranges. Although epoxy-fiberglass (in plane cloth) is inferior to one of nylon, teflon, and epoxy-fiberglass (normal cloth) at all temperatures, it is kept because it was used by Hilal and Boom in Hilal and Boom (1977). Figure 4(b) suggests that the insulator defining an interesting solution would consist of a sequence of (N)ylon, followed by a sequence of (T)eflon, then one of (E)poxy-fiberglass (normal cloth) and finally (T)eflon again (note that these sequences can also be empty), i.e., the variable $I$ belongs to the set

$$\mathcal{I} = \{\mathbf{I} = [I_1, I_2, \ldots, I_{n+1}]^T : \; I_i = \begin{cases} N & \text{if } i \le i_1 \\ T & \text{if } i_1 < i \le i_2 \quad \text{or} \quad i_3 < i \\ E & \text{if } i_2 < i \le i_3, \end{cases}$$
$$1 \le i \le n+1, \quad 0 \le i_1 \le i_2 \le i_3 \le n+1\}.$$

Therefore, this extra information about the nature of the problem leads to the assumption that any solution is composed of such sequences.

The set of neighboring solutions considered by the POLL step include those where

- any of the existing intercepts and the insulator to its left are removed,
- a new intercept together with an insulator to its right are added,
- the type of insulator between two intercepts is changed.

We now illustrate the flexibility of the method by presenting the rules for defining the set of neighbors by considering an example. Table 1 displays a solution with four intercepts and the 23 poll trial points that must be explored before declaring an iteration unsuccessful. The poll trial points are considered in the algorithm in the order presented in the table. Note that in the current implementation the POLL step and the iteration ends as soon as a better solution is found. The entries that differ from the current solution $x_k$ appear as boldface characters. In the remainder of the paper, we present the normalized spacing $\frac{\Delta x}{L}$ instead of $\Delta x$.

The first spanning set $S$, with a mesh size parameter $\Delta_k = 5$ generates 9 poll trial points that differ only in the continuous variables $\{x_k^c + \Delta_k s : s \in S\}$ (left part of Table 1). The other poll trial points have to be defined by the user who has an understanding of the model (right part of Table 1). The set of neighbors contains 5 more poll trial points obtained by replacing the type of an insulator by another. Of course, the variable $\mathbf{I}$ for all these poll trial points belongs to the set $\mathcal{I}^{n+1}$. It also contains the 5 poll trial points corresponding to adding an intercept (and therefore an insulator) according to the following rules assuming that the new intercept is introduced between the $i$th and the $(i-1)$-th previously existing intercepts:

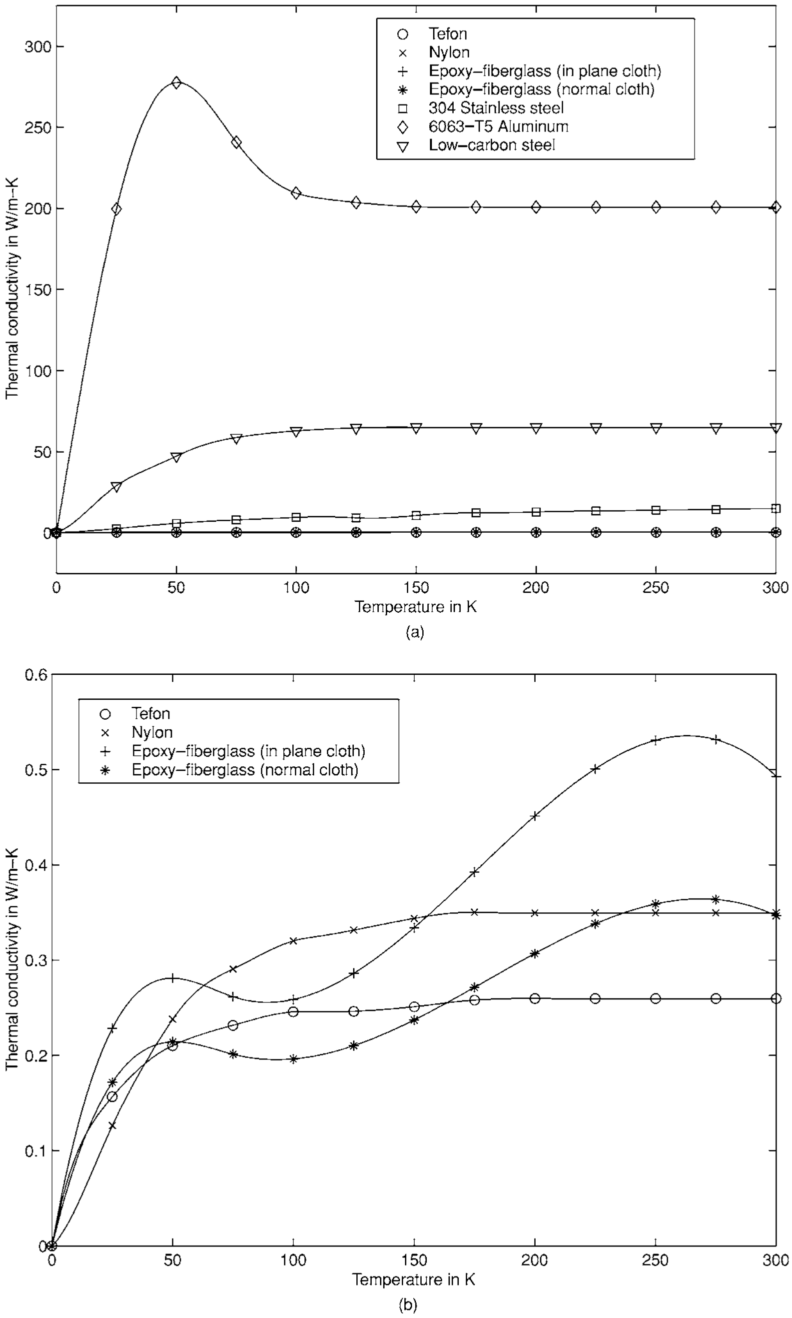


*Figure 4*. Thermal conductivity versus temperature for (a) all insulators and (b) the “better” four insulators.

*Table 1.* Example of a POLL set.

| $n$ | $\frac{\Delta x}{L}$ [%] | **T** [K] | **I** |
|---|---|---|---|
| | Current incumbent solution $x_k$ | | |
| 4 | [20 20 20 20] | [10 20 40 80] | [*TTEET*] |
| | "Continuous" mesh neighbors | | |
| 4 | [**25** 20 20 20] | [10 20 40 80] | [*TTEET*] |
| 4 | [20 **25** 20 20] | [10 20 40 80] | [*TTEET*] |
| 4 | [20 20 **25** 20] | [10 20 40 80] | [*TTEET*] |
| 4 | [20 20 20 **25**] | [10 20 40 80] | [*TTEET*] |
| 4 | [20 20 20 20] | [**15** 20 40 80] | [*TTEET*] |
| 4 | [20 20 20 20] | [10 **25** 40 80] | [*TTEET*] |
| 4 | [20 20 20 20] | [10 20 **45** 80] | [*TTEET*] |
| 4 | [20 20 20 20] | [10 20 40 **85**] | [*TTEET*] |
| 4 | [**15 15 15 15**] | [**5 15 35 75**] | [*TTEET*] |
| | Change an insulator type | | |
| 4 | [20 20 20 20] | [10 20 40 80] | [**N***TEET*] |
| 4 | [20 20 20 20] | [10 20 40 80] | [*T***E***EET*] |
| 4 | [20 20 20 20] | [10 20 40 80] | [*TT***T***ET*] |
| 4 | [20 20 20 20] | [10 20 40 80] | [*TTE***T***T*] |
| 4 | [20 20 20 20] | [10 20 40 80] | [*TTEE***E**] |
| | Add an intercept (and therefore an insulator) | | |
| **5** | [**10** 10 20 20 20] | [**5** 10 20 40 80] | [**T***TTEET*] |
| **5** | [20 **10** 10 20 20] | [10 **15** 20 40 80] | [*T***T***TEET*] |
| **5** | [20 20 **10** 10 20] | [10 20 **30** 40 80] | [*TT***E***EET*] |
| **5** | [20 20 20 **10** 10] | [10 20 40 **60** 80] | [*TTE***E***ET*] |
| **5** | [20 20 20 20 **10**] | [10 20 40 80 **190**] | [*TTEE***T***T*] |
| | Remove an intercept (and therefore an insulator) | | |
| **3** | [**25 25 25**] | [**20 40 80**] | [**TEET**] |
| **3** | [**25 25 25**] | [**10 40 80**] | [**TEET**] |
| **3** | [**25 25 25**] | [**10 20 80**] | [**TTET**] |
| **3** | [**25 25 25**] | [**10 20 40**] | [**TTET**] |

- the cooling temperature of the new intercept is $T_{\text{new}} = \frac{T_i + T_{i-1}}{2}$, rounded to the nearest integer multiple of the mesh size parameter $\Delta_k$,
- the type of the insulator associated with the new intercept is the same as the type of the $i$th insulator, i.e., $I_{\text{new}} = I_i$,
- the thickness of the insulator associated with the new intercept is $\Delta x_{\text{new}} = \frac{\Delta x_i}{2}$; to accommodate this, the thickness $\Delta x_i$ of the insulator $I_i$ is halved. Both are rounded to the nearest integer multiple of the mesh size parameter $\Delta_k$.

Finally, the last four elements of the POLL set correspond to the poll trial points where one of the four intercepts (and therefore an insulator) is removed; the thickness of the removed insulator is distributed among the remaining others, and again the lengths are rounded to the nearest integer multiple of the mesh size parameter $\Delta_k$. The purpose of rounding the values is to make sure that all trial points lie on the mesh generated by the mesh directions of $S$ (as required by the convergence theory).

The speed of the algorithm and the quality of the solution produced by it depends on the user-defined set of neighbors. For example, if one does not realize that the type of insulators of interesting solutions belong to $\mathcal{I}$, then the trial points associated with modifying an insulator could contain all the combinations of insulators that differ from the incumbent solution in any one component. In the example above, there would be ten poll points corresponding to a change of insulator type instead of five. For larger $n$, the algorithm would require significantly more function evaluations, but it would not produce a better solution.

## 5. Results and discussion

In this section, we present numerical results for our Matlab 5.3 implementation of the MVP algorithm described above to the mathematical model. First, we reproduce and compare results that were reported previously in the literature in order to validate the implementation of the algorithm. Then, we report new results in Sections 5.2–5.4.

### 5.1. *Optimization with fixed insulators and number of intercepts*

In Table 2, we compare the results presented by Hilal and Boom (H & B) when using 304 stainless steel for the entire mechanical support and $n = 1, 2$, and 3 intercepts to the results obtained by the MVP algorithm when forced to use the same insulator and respective maximum number of intercepts. We emphasize the fact that $n$ was fixed to 1, 2, and 3 in Hilal and Boom's method, while the MVP algorithm was initiated with $n = 1$ and converged to 1, 2, and 3 for the three test cases, respectively. The initial guess for the location of the intercept $\frac{\Delta x_1}{L}$ is 50%, i.e., it is initially positioned at half distance from the hot ($T_H = 300$ K) and cold

*Table 2.* Optimum temperatures, locations, and refrigeration power when using 304 stainless steel for the entire strut; $T_C = 4.2$ K.

| Algorithm | $n$ | $T_1$, [K] | $T_2$, [K], | $T_3$, [K] | $\frac{\Delta x_1}{L}$ [%] | $\frac{\Delta x_2}{L}$ [%] | $\frac{\Delta x_3}{L}$ [%] | $\frac{\Delta x_4}{L}$ [%] | $\frac{PL}{A}$, [$\frac{\text{W}}{\text{cm}}$] |
|---|---|---|---|---|---|---|---|---|---|
| H & B | 1 | 39.7 | – | – | 33.8 | 66.2 | – | – | 1927 |
| MVP | 1 | 36.2 | – | – | 32.9 | 67.1 | – | – | 1910 |
| H & B | 2 | 21.5 | 81.9 | – | 18.8 | 33.5 | 47.6 | – | 1134 |
| MVP | 2 | 18.2 | 71 | – | 18.5 | 36.3 | 45.2 | – | 1077 |
| H & B | 3 | 11.7 | 28.7 | 72.4 | 9.3 | 14.7 | 28.1 | 47.9 | 966 |
| MVP | 3 | 10.8 | 27.9 | 71.5 | 9.4 | 14.9 | 28.1 | 47.6 | 963.5 |

($T_C = 4.2$ K) surfaces. The initial guess for the temperature is 50 K. The thermodynamic cycle efficiency coefficient is a function of the temperature as follows

$$C = \begin{cases} 2.5 & \text{if } T \geq 71 \text{ K} \\ 4 & \text{if } 71 \text{ K} > T > 4.2 \text{ K} \\ 5 & \text{if } T \leq 4.2 \text{ K}. \end{cases} \tag{10}$$

All objective function values are normalized with respect to unit area and length.

Note that Hilal and Boom performed their computations in the late seventies. In this regard, and in order to have a common comparison basis, we recalculated the objective function values based on their reported variable values using our function evaluation routine, and used these numbers in Tables 2 and 3. The objective function values Hilal and Boom (1977) report (i.e., computed by their function evaluation routine) are 1781, 1265, and 948.7 for $n = 1$, 2, and 3, respectively, in Table 2, and 142, 94.5, and 71.9 for $n = 1$, 2, and 3, respectively, in Table 3.

Table 3 tabulates the results of Hilal and Boom when using epoxy-fiberglass (in plane cloth) for the entire strut and $n = 1$, 2, and 3 intercepts, and the results obtained by the MVP algorithm when forced to use the same insulator and number of intercepts. In this case, the cold surface temperature was $T_C = 1.8$ K.

Due to the presence of local optima, it is possible to converge to a different solution if the algorithm is initiated at a different starting point. For example, the implementation produces different results when the initial guess for the temperature is 150 K. In order to investigate the existence of local optima, we plot the objective function for $n = 1$ intercept over the possible temperature and location ranges. When looking at the top of figure 5, the objective function looks quite smooth. However, zooming in reveals (at the bottom of figure 5) the presence of a local optimum. It is clear that the local optimum is associated with the discontinuity at $T = 71$ K caused by the discontinuous change of the thermodynamic cycle efficiency $C$ at this point. With the initial guess of the temperature being $T_1 = 150$ K, the MVP implementation converges to the local optimum caused by the discontinuity. When we change the initial guess for $T_1$ to 50 K, the MVP implementation yields the results displayed in Table 2. Note that Hilal and Boom did not comment on

*Table 3*. Optimum temperatures, locations, and refrigeration power when using epoxy-fiberglass (in plane cloth) for the entire strut; $T_C = 1.8$ K.

| Algorithm | $n$ | $T_1$, [K] | $T_2$, [K], | $T_3$, [K] | $\frac{\Delta x_1}{L}$ [%] | $\frac{\Delta x_2}{L}$ [%] | $\frac{\Delta x_3}{L}$ [%] | $\frac{\Delta x_4}{L}$ [%] | $\frac{PL}{A}$, [$\frac{\text{W}}{\text{cm}}$] |
|---|---|---|---|---|---|---|---|---|---|
| H & B | 1 | 29.6 | – | – | 47.5 | 52.5 | – | – | 145 |
| MVP | 1 | 21.7 | – | – | 37.9 | 62.1 | – | – | 140 |
| H & B | 2 | 11.1 | 70.3 | – | 30.5 | 32.4 | 37.0 | – | 91.9 |
| MVP | 2 | 10.5 | 65.2 | – | 23.2 | 37.9 | 38.9 | – | 89.7 |
| H & B | 3 | 5.28 | 20 | 71.7 | 19.2 | 21.9 | 25.8 | 33.1 | 68.6 |
| MVP | 3 | 5.6 | 18.8 | 71.0 | 13.3 | 20.2 | 32.8 | 33.7 | 65.9 |

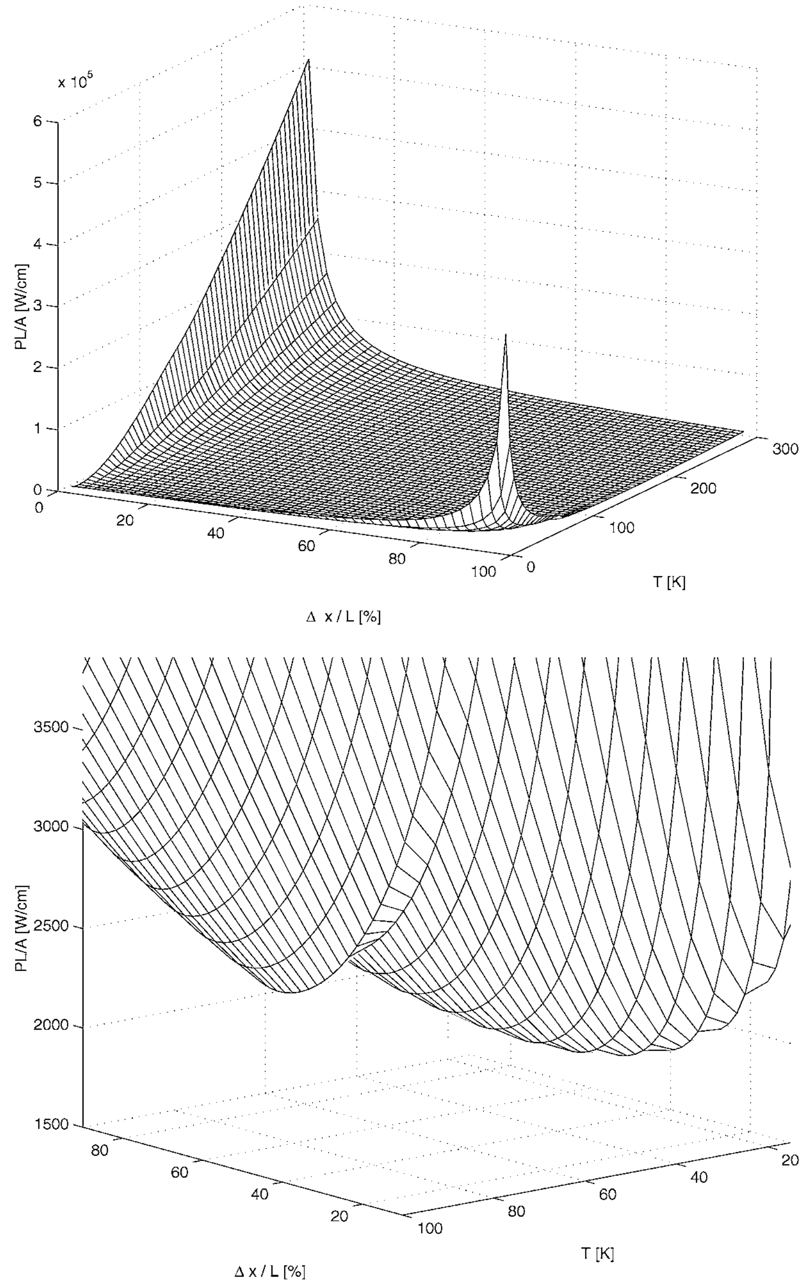


*Figure 5.* The refrigeration power objective function for $n = 1$ intercept when using 304 stainless steel for the entire strut.

either the initial guesses or a multi-start strategy and the associated convergence results of the gradient-based optimization algorithm they used. Also, they did not provide any information on the performance of their algorithm (e.g. number of function evaluations), making comparison impossible.

An alternative to switching starting points for increasing the likelihood of converging to the global optimum is to incorporate a more sophisticated SEARCH strategy and/or increase the local mesh size parameter $\Delta_k$ at successful iterations. Again we mention that we did not take this route for the reasons discussed earlier.

### *5.2. Categorical variable optimization*

Before proceeding with further numerical results, we first need to address some more computational implementation issues: As mentioned in Section 4.2, the SEARCH step is minimal; all results are obtained mostly by polling. In practice, the algorithm terminates when the following criterion is satisfied

$$\Delta_k < \delta, \tag{11}$$

where $\Delta_k$ is the mesh size parameter at the $k$-th iteration and $\delta$ is some nonnegative small tolerance. When condition (11) is met for a small $\delta$, the current solution satisfies optimality conditions on a fine mesh; a local optima is probably found or nearby. For the specific applications of this paper, when the algorithm starts to converge near a local optimizer, our experience shows that a more aggressive approach of reducing the mesh size parameter reduces the total number of function evaluations with respect to the more passive approach of simply dividing it by two. In this regard, we start with the initial mesh size parameter $\Delta_0 = 10$, and then refine the mesh according to the rule

$$\Delta_{k+1} = \frac{\Delta_k}{2^\ell} \tag{12}$$

when the $\ell$th local mesh optimizer is found. This means that when the first local mesh optimizer is found, the mesh size parameter decreases to 5, then when another is found it drops to 1.25 and finally to 0.15625. Since a final mesh size parameter of 0.15625 is sufficient for practical engineering purposes, we stop at $\ell = 4$, and accept the associated local mesh optimizer as our final solution.

The parameter $\xi$ that triggers the EXTENDED POLL step is set to 1% of the incumbent objective function value. The following initial guess and parameter values are used for all calculations in the remaining sections: $n = 1$, $\frac{\Delta x_1}{L} = 50\%$, $I_1 = N$, $I_2 = T$, $T_1 = 150$ K, $T_C = 4.2$ K, and $T_H = 300$ K.

We now apply the MVP algorithm to the optimization problem with the maximum number of intercepts $n_{\max} = 100$, which is quite large. In addition, any appropriate combination of the following three insulators can be chosen for the spaces between the heat intercepts: (N)ylon, (T)eflon, or (E)poxy-fiberglass (normal cloth). For later comparison, the available information on the nature of the problem is not exploited when defining the neighbors for the POLL step, i.e., it is not assumed that the solutions are composed insulators in the set $\mathcal{I}$.

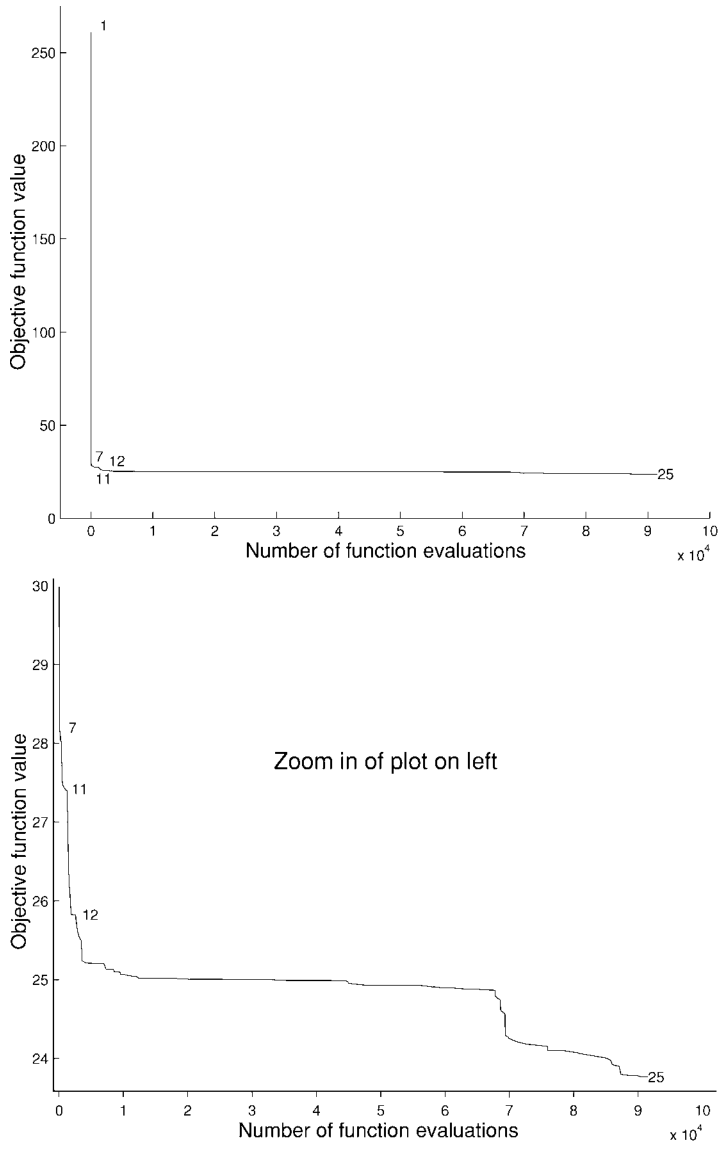


*Figure 6.* Evolution of the objective function for minimum refrigeration power when the set of neighbors does not take into account the extra information on the problem.

We depict the evolution of the objective function versus the number of function evaluations in figure 6. It can be seen on the top of figure 6 that the minimum refrigeration power is approached rapidly in the early stages of the algorithm, and then progress slows; this is typical behavior for derivative-free algorithms. The numbers next to the objective function values curve indicate the number of intercepts at local mesh minimizers, i.e., the local

*Table 4.* Configuration of the thermal insulation system for minimum refrigeration power when the set of neighbors does not take into account the extra information on the problem.

$n = 25$, $\frac{PL}{A} = 23.764682$ [$\frac{\text{W}}{\text{cm}}$], $\eta = 97002$

$\frac{\mathbf{\Delta x}}{L}$ [%] = [.3125 1.4062 1.0938 1.0938 2.0312 1.8750 1.875 3.2812 1.5625 1.5625 3.2812 2.8125 2.6562 2.6562 2.5 4.6875 4.8438 6.4062 13.9062 .7812 5. 4.0625 .625 4.5312 2.6562 22.5]

**T** [K] = [4.2188 5. 5.625 6.25 7.5 8.75 10. 12.3438 13.5938 14.8438 17.5 20. 22.5 25. 27.5 32.5 38.125 46.4062 70.9375 71.0938 86.7188 102.1875 104.8438 125.3125 139.2188]

$\mathbf{I} = [NNNNNNNNNNNNNNNNNTEEEEEEET]$

optimum number of intercepts for $\Delta_k = 10$, 5, 1.25, and 0.15625 was $n = 7$, 11, 12, and 25, respectively. The drop in the objective function can be examined better on the bottom of figure 6, where the first few hundred function evaluations are excluded.

The configuration of the thermal insulation system appears in Table 4; $\eta$ denotes the number of function evaluations. Recall that the minimum normalized power reported by Hilal and Boom (1977) was 68.6 $\frac{\text{W}}{\text{cm}}$ (Table 3); increasing the number of intercepts and, most importantly, combining different insulators decreases the objective function value by 65%. Of course, this is achieved also because insulators other then epoxy were selected by the MVP algorithm. Although Hilal and Boom considered all insulators mentioned in Section 4.2, the best value they reported was based on epoxy in plane cloth. We emphasize again this advantage of an algorithm that can treat categorical variables automatically. Hilal and Boom could not consider combinations of insulators automatically in their algorithm. They would have needed the prescience to include *the right specific cases* in their parametric studies.

Observe that the selection of insulators is optimal with respect to their thermal conductivity over the chosen temperature intervals, i.e., the solution produced by the algorithm is composed of the specified sequence of insulators, even if this was not imposed. Specifically, at the 55898-th function evaluation, the incumbent solution does not satisfy this property. It is only at the 80663-rd evaluation that the new incumbent is composed of the specific sequence of insulators. Figure 6 shows that a significant decrease in the objective function value occurs in that interval.

The speed of convergence can be improved by exploiting the available information on the nature of the problem when defining the neighbors for the POLL step, i.e., assuming that the variable **I** belongs to the set $\mathcal{I}$. These results are tabulated in Table 5. The total number of function evaluations is significantly smaller (by approximately 73%) than when the extra information on the problem is not taken into account. The gain in function evaluations compensates for a small loss (approximately 4.8%) in the objective function value. At the same time, however, there is a saving in costs due to the use of about 9% fewer intercepts. The plot of the objective function value versus the number of function evaluations appears in the top left part of figure 8.

During the run that produced the results presented in Table 4, the algorithm evaluated the objective function at some solutions outside of the set $\mathcal{I}$. Our discussion of figure 4 concluded that these solutions would not be retained by the algorithm. Therefore, both runs

*Table 5*. Optimum configuration of the thermal insulation system for minimum refrigeration power when defining the neighbors by taking into account available information on the problem.

| |
|---|
| $n = 23$, $\frac{PL}{A} = 24.863534$ [$\frac{\mathrm{W}}{\mathrm{cm}}$], $\eta = 25843$ |
| $\frac{\mathbf{\Delta x}}{L}$ [%] = [.3125 .9375 .7812 1.4062 1.25 2.3438 2.5 3.4375 1.5625 1.5625 2.9688 2.9688 5. 3.4375 3.4375 3.75 4.6875 19.375 4.6875 3.9062 5.1562 2.6562 2.8125 19.0625] |
| **T** [K] = [4.2188 4.6875 5.1562 5.9375 6.7188 8.2812 10. 12.5 13.75 15. 17.5 20.1562 25. 28.5938 32.3438 36.875 43.125 71.0938 85.7812 100.625 123.9062 137.6562 153.125] |
| $\mathbf{I} = [NNNNNNNNNNNNNNNNTEEEEEET]$ |

converge to a solution in $\mathcal{I}$, but the one in Table 5 converges more rapidly since it only considers solutions in $\mathcal{I}$.

The solutions in Tables 4 and 5 differ, a fact that points to the presence of local optima (both are local mesh optimizers for the same mesh size parameter). Different starting points and/or different definitions of the set of neighbors may lead the algorithm to different local solutions. It is also clear that obtaining an improved solution is correlated to higher computational expenses. How much more expense is necessary is not clear because the implementation of the MVP algorithm we used is intended to be as simple as possible and does not include major cost saving features, like a non trivial SEARCH strategy, a strategy for increasing the mesh size parameter $\Delta_k$ at successful iterations, or the use of a function cache for avoiding the function evaluation at the same points in different iterations. In this regard, the reader should look at the number of necessary function evaluations in a qualitative and not quantitative manner.

All runs in the following sections use the information in the set $\mathcal{I}$ for the POLL step.

### *5.3. Limiting the number of heat intercepts*

The results in the previous section suggest that after some point, additional intercepts do not improve the objective function value significantly. Moreover, if the mesh size parameter $\Delta_k$ is allowed to get smaller than 0.15625, the algorithm will converge to larger numbers of intercepts $n$. This is because the algorithm will try to emulate a multilayer insulation by choosing many intercepts: if the mesh size parameter could decrease to smaller values, it would enable the algorithm to introduce additional intercepts with marginal cooling temperature differences and at marginal distances. However, for engineering purposes, and due to cost considerations, we are satisfied with a decimal accuracy. In addition, as $n$ increases, so does the computational work since the problem size increases.

This motivates us to consider further research on strategies that terminate the optimization process sooner, but with a good solution. For this study, we accomplish this is by setting an smaller upper bound on the number of intercepts to be used, i.e., by using smaller values for $n_{\max}$. Figure 7 displays the associated objective function plot for $n_{\max} = 10$. The numbers in parentheses next to the curve on the bottom of figure 7 indicate that the number of intercepts

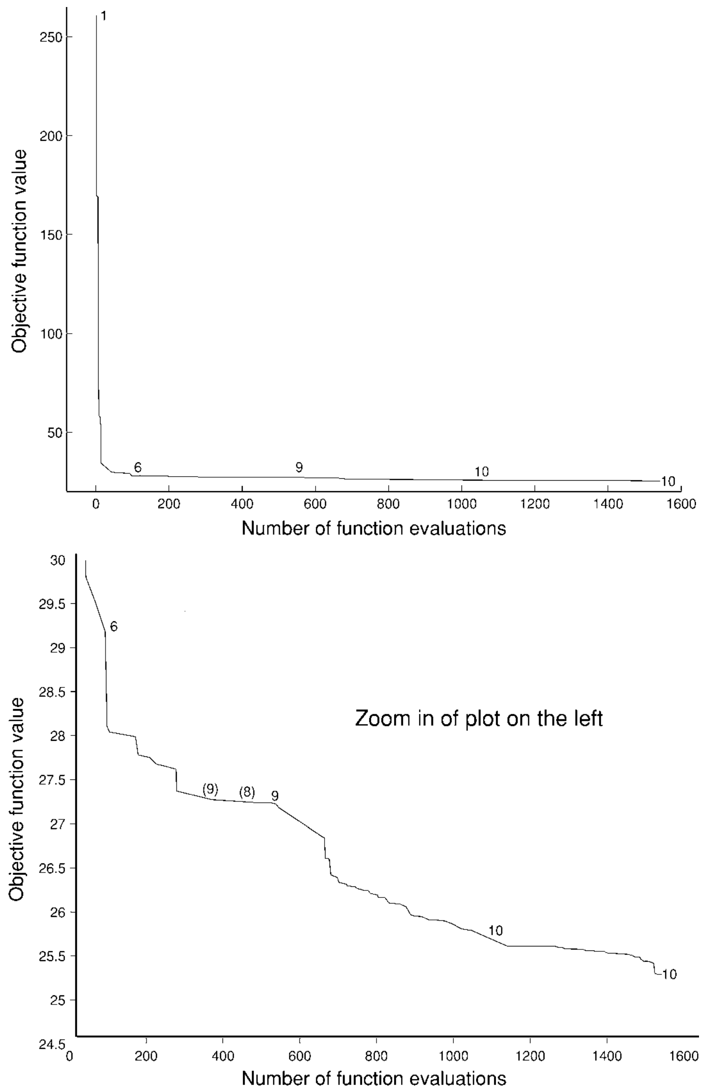


*Figure 7.* Evolution of the objective function for minimum refrigeration power when the set of neighbors takes into account the extra information on the problem; $n_{\max} = 10$.

was increased and then decreased by the algorithm, which shows that intercepts are not only added during the optimization process, but also removed. This behavior is also present in the runs of Section 5.4.

The solution appears in Table 6. It can be seen that, when hitting the upper bound for $n$, the MVP algorithm converges in 92% less function evaluations with a loss in optimality less than 2%, compared to the run with information detailed in Table 5.

*Table 6.* Optimum configuration of the thermal insulation system for minimum refrigeration power when the set of neighbors takes into account the extra information on the problem; $n_{\max} = 10$.

$n = 10$, $\frac{PL}{A} = 25.293569$ [$\frac{\mathrm{W}}{\mathrm{cm}}$], $\eta = 2020$

$\frac{\boldsymbol{\Delta x}}{L}$ [%] = [.3125 5.4688 3.9062 6.5625 5.7812 5.1562 13.2812 21.4062 8.5938 9.2188 20.3125]

**T** [K] = [4.2188 7.3438 10. 15. 20. 25. 40. 71.0938 101.25 146.25]

$\mathbf{I} = [NNNNNNNEEET]$

*Table 7.* Losses in optimality and reduction in computational work when including an extra cost term in the objective function for minimum refrigeration power.

| $\gamma$ [%] | $n$ | $f = \frac{PL}{A}$ [$\frac{\mathrm{W}}{\mathrm{cm}}$] | $\eta$ | Loss in $f$ [%] | Reduction in $\eta$ [%] |
|---|---|---|---|---|---|
| 0 | 23 | 24.863534 | 25843 | – | – |
| 0.1 | 13 | 25.728785 | 8751 | 3.5 | 66 |
| 0.5 | 8 | 27.207694 | 5883 | 9.4 | 77 |
| 1 | 7 | 27.380670 | 10103 | 10.1 | 61 |

*Table 8.* Optimum configuration of the thermal insulation system for minimum refrigeration power when the set of neighbors takes into account the extra information on the problem and including an extra cost term in the objective function.

$\gamma = 0.1\%$, $n = 13$, $\frac{PL}{A} = 25.728785$ [$\frac{\mathrm{W}}{\mathrm{cm}}$], $\eta = 8751$

$\frac{\boldsymbol{\Delta x}}{L}$ [%] = [.3125 4.0625 2.6562 4.375 5.9375 5.4688 7.5
6.5625 6.0938 8.5938 9.5312 5.625 4.5312 28.75]

**T** [K] = [4.2188 6.5625 8.2812 11.4062 16.2500 21.25 28.9062
36.5625 44.6875 57.9688 71.0938 90. 108.9062]

$\mathbf{I} = [NNNNNNNNTEEEET]$

$\gamma = 0.5\%$, $n = 8$, $\frac{PL}{A} = 27.207694$ [$\frac{\mathrm{W}}{\mathrm{cm}}$], $\eta = 5883$

$\frac{\boldsymbol{\Delta x}}{L}$ [%] = [.3125 4.0625 5.625 7.8125 9.375 7.6562 7.1875 28.4375 29.5313]

**T** [K] = [4.2188 6.5625 10.4688 16.7188 25.7812 34.6875 44.8438 104.0625]

$\mathbf{I} = [NNNNNNTET]$

$\gamma = 1\%$, $n = 7$, $\frac{PL}{A} = 27.38067$ [$\frac{\mathrm{W}}{\mathrm{cm}}$], $\eta = 10103$

$\frac{\boldsymbol{\Delta x}}{L}$ [%] = [.3125 5.7812 5.7812 6.25 9.6875 13.2812 30.1562 28.75]

**T** [K] = [4.2188 7.6562 11.8750 17.1875 26.7188 42.9688 106.0938]

$\mathbf{I} = [NNNNNNET]$

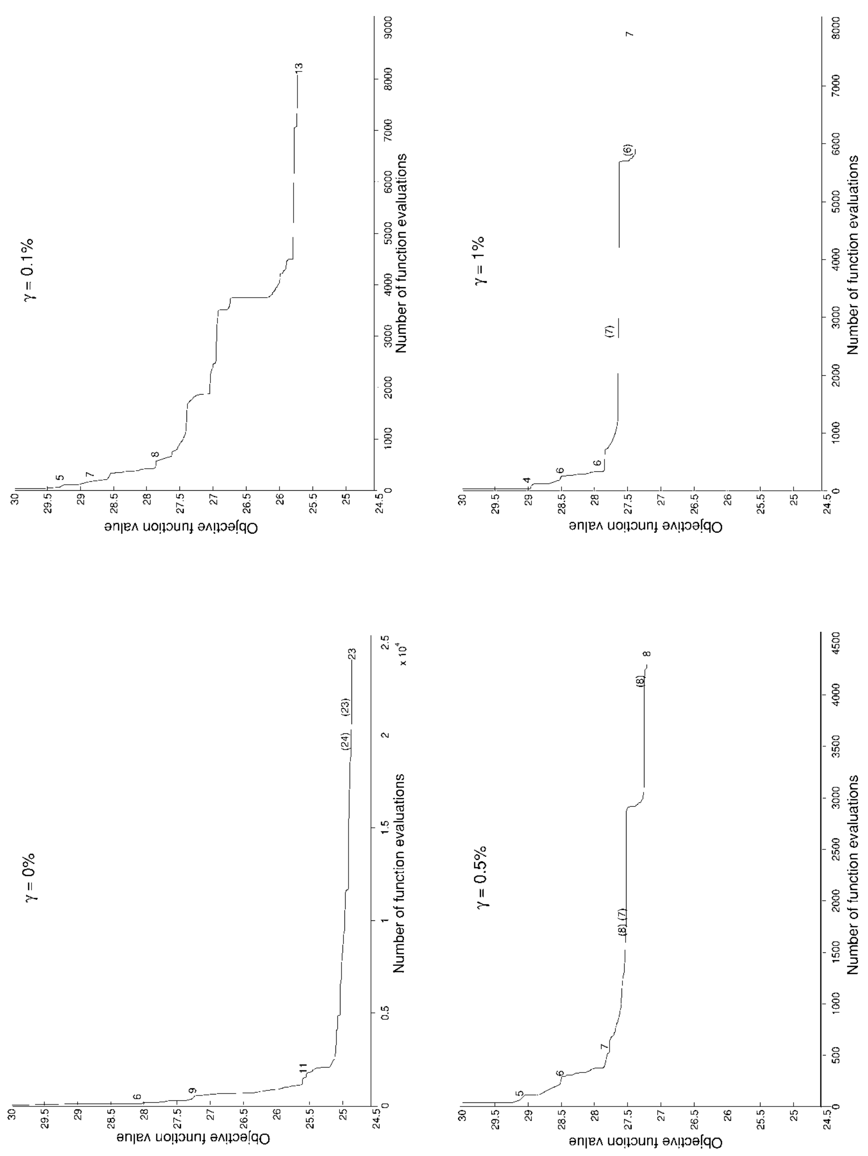


*Figure 8.* Evolution of the objective function for minimum refrigeration power when the set of neighbors takes into account the extra information on the problem and including an extra cost term of 0% (top left), 0.1% (top right), 0.5% (bottom left), and 1% (bottom right) in the objective function; $n_{\max} = 100$.

*5.4. Adding an extra cost term to the objective function*

Keeping the number of intercepts $n$ reasonably low without sacrificing optimality substantially can be modeled by including an extra cost term in the objective function such that is not added unless a certain percentage gain $\gamma$ in the objective function value can be achieved, i.e., the objective function is multiplied by $(1 + \frac{\gamma}{100}n)$. Similarly, an extra term can be added for cost or weight, especially for use in space missions, in order to keep the number of intercepts low.

Figure 8 summarizes the results obtained for different extra cost term coefficients. On the top right, bottom left and right of figure 8, a intercept is added only if the objective function is reduced by at least 0.1%, 0.5%, or 1%, respectively; the first few function evaluations are excluded to facilitate the examination of the objective function behavior. The top left one represent the case with no additional cost, described in Table 5. It is clear that a high value for the extra cost term will yield a lower number of intercepts and a higher objective function value. Note that the extra cost term coefficient cannot be larger than the extended poll triggering coefficient $\xi$.

Losses in optimality and reduction in computational work related to the extra cost term and using the polling information are tabulated in Table 7; $\gamma$ denotes the extra cost term coefficient. The optimum configurations are reported in Table 8.

## 6. Summary and conclusions

The new mixed variable programming algorithm (Audet and Dennis, 2000c) was used to optimize thermal insulation systems with respect to both continuous and categorical variables. We learned some lessons from this study about the algorithm in general and about this application in particular. We summarize some features of the algorithm together with our conclusions from this study.

- The algorithm can be applied to a broad class of optimization problems in engineering that could not be easily solved before due to the presence of categorical variables. Categorical variables are treated here as optimization variables and not as parameters. Important components of the optimal configuration, such as the number of heat intercepts and the types of insulators, are taken into consideration directly during the optimization process. The objective function value is reduced by as much as 65% compared to the previously reported result in the literature. Further improvement would be likely from treating the cross section area as an optimization variable. The implementation here could be used for optimizing the configuration of any general thermal insulation system (with or without mechanical supports, and for any kind of media) if it is provided with effective thermal conductivity data.
- Being able to treat the categorical variables exposed the need for modifications to the objective function not needed for a fixed number of intercepts. We observed that the addition of intercepts in the late stages of the algorithm reduces the objective function value only marginally. Improving the objective model caused the algorithm to add an intercept only if a certain percentage gain would be achieved in the objective. More sophisticated cost and/or weight functions can be developed for specific applications. In

retrospect, it is to be expected that the model objective will have to be refined when used with additional optimization variables.

- The algorithm requires from the user only a function evaluation routine (black box) and no derivative information. Further convergence analysis of pattern search algorithms with respect to the local smoothness of the objective function can be found in Audet and Dennis (2000b), Abramson (to appear): There are no smoothness or finite value assumptions on the objective function; the black box is allowed to return infinite or no value at all for the objective function (for example, it returns the value $+\infty$ if the discrete variables are relaxed to continuous ones). The rapid decrease of the objective function value in the early stages of the algorithm, followed by a plateau, is typical behavior for derivative-free methods. Finally, general constraints can also be handled by a "filter" version of the algorithm Audet and Dennis (2000a), and work is underway to incorporate any user supplied derivative information.
- The algorithm can be readily implemented on a parallel architecture; such an implementation would be highly scalable and either decrease computational time dramatically, or else explore design space more thoroughly for the best optimizers.
- The algorithm consists of two main components: a) the SEARCH step, which can employ any strategy based on available information of the problem (including none) and may accelerate convergence and/or lead to the global optimum, and b) the POLL step, which guarantees convergence to a point that satisfies some first order optimality conditions for any initial guess (the conditions depend on the local smoothness of the objective function at that point; see Audet and Dennis (2000b), Abramson (to appear) for more details). It is also clear that obtaining improved extrema is highly correlated with paying a higher computational price. That was demonstrated here when a higher EXTENDED POLL triggering parameter $\xi$ required more function evaluations but yielded a better solution.
- Problem specific information can lead to a less expensive POLL STEP. The algorithm determined in choosing the correct insulators that different types of insulators are optimal for different temperature ranges. In retrospect, this was obvious, and it is the sort of information the user might provide. When we used this problem specific information to define the set of categorical neighbors, the algorithm requires many fewer function evaluations.

## Acknowledgments

Work of the first author was done during a stay as a visiting research scholar at Rice University. Work of the second author was supported by NSERC (Natural Sciences and Engineering Research Council) fellowship PDF-207432-1998 during a post-doctoral stay at Rice University. All three authors were supported by DOE DE-FG03-95ER25257, AFOSR F49620-01-1-0013, The Boeing Company, Sandia LG-4253, Exxon-Mobil and CRPC CCR-9120008.

## Note

1. For notation convenience, discrete variables denote the set of categorical and integer variables. In the present application $x^c = (\mathbf{\Delta x}, \mathbf{T})$ and $x^d = (n, \mathbf{I})$.

## References


M. Abramson, “Pattern search algorithms for mixed variable general constrained optimization problems,” PhD thesis, Rice University, Department of Computational and Applied Mathematics, Houston, Texas, in progress, to appear as a CAAM Technical Report.

P. Alberto, F. Nogueira, H. Rocha, and L. N. Vicente, “Pattern search methods for molecular geometry problems,” Technical Report Preprint 00–20, Departamento de Matemática, Universidade de Coimbra, 2000.

C. Audet and J. E. Dennis, Jr., “A pattern search filter method for nonlinear programming without derivatives,” Technical Report 00–09, Department of Computational and Applied Mathematics, Rice University, Houston, Texas, March 2000a.

C. Audet and J. E. Dennis, Jr., “Analysis of generalized pattern searches,” Technical Report 00–07, Department of Computational and Applied Mathematics, Rice University, Houston, Texas, Feb. 2000b.

C. Audet and J. E. Dennis, Jr., “Pattern search algorithms for mixed variable programming,” *SIAM Journal on Optimization* vol. 11, no. 3, pp. 573–594, 2000c.

R. Barron, *Cryogenic Systems*. McGraw-Hill: New York, 1966, p. 469.

A. Bejan, “A general variational principle for thermal insulation system design,” *International Journal of Heat and Mass Transfer* vol. 22, pp. 219–228, 1979.

J. C. Chato and J. M. Khodadadi, “Optimization of cooled shields in insulations,” *ASME Transactions, Journal of Heat Transfer* vol. 106, pp. 871–875, 1984.

F. H. Clarke, “Optimization and nonsmooth analysis,” *SIAM*, Classics in Applied Mathematics, Philadelphia, PA vol. 5, 1990.

*Handbook on Materials for Superconducting Machinery, Mechanical, Thermal, Electrical, and Magnetic Properties of Structural Materials*. Metals and Ceramics Information Center, Batelle’s Columbus Laboratories, 1974.

M. A. Hilal and R. W. Boom, “Optimization of mechanical supports for large superconductive magnets,” *Advances in Cryogenic Engineering* vol. 22, pp. 224–232, 1977.

M. A. Hilal and Y. M. Eyssa, “Minimization of refrigeration power for large cryogenic systems,” *Advances in Cryogenic Engineering* vol. 25, pp. 350–357, 1980.

R. M. Lewis and V. Torczon, “Pattern search algorithms for bound constrained minimization,” *SIAM Journal on Optimization* vol. 9, pp. 1082–1009, 1999.

R. M. Lewis and V. Torczon, “Pattern search methods for linearly constrained minimization,” *SIAM Journal on Optimization* vol. 10, pp. 917–941, 2000.

Q. Li, X. Li, G. E. McIntosh, and R. W. Boom, “Minimization of total refrigeration power of liquid neon and nitrogen cooled intercepts for smes magnets,” *Advances in Cryogenic Engineering* vol. 35, pp. 833–840, 1989.

Z. Musicki, M. A. Hilal, and G. E. McIntosh, “Optimization of cryogenic and heat removal system of space borne magnets,” *Advances in Cryogenic Engineering* vol. 35, pp. 975–982, 1989.

V. Torczon, “On the convergence of pattern search algorithms,” *SIAM Journal on Optimization* vol. 7, pp. 1–25, 1997.

M. Yamaguchi, T. Ohmori, and A. Yamamoto, “Design optimization of a vapor-cooled radiation shield for LHe cryostat in space use,” *Advances in Cryogenic Engineering* vol. 37, pp. 1367–1375, 1991.