Structural and Multidisciplinary Optimization manuscript No.
(will be inserted by the editor)

Stopping rule of multi-start local search for structural optimization

Makoto Ohsaki · Makoto Yamakawa

Received: date / Revised version: date

Abstract Stopping rule for multi-start local search is investigated for application to structural optimization problems with moderately large number of local optima. The size of attractor of an unknown local optimal solution is estimated using the information of already obtained solutions. A stopping rule is defined using the likelihood of obtaining the specific set of local optimal solutions. This way, characteristics of the specific optimization problem is successfully incorporated. The proposed rule is first verified using the mathematical problems in comparison with the existing rule utilizing the estimated ratio of the total size of attractors to the size of feasible domain. The rule is next applied to an optimization problem of a plane frame under constraints on inter-story drift angle and stress against static loads. Characteristics of the distribution of attractors of optimal solutions are also investigated.

Keywords Multi-start local search · stopping rule · stress constraint · building frame

1 Introduction

Among various problems of structural and multidisciplinary optimization (SMO), it is natural to expect that there exist many local optimal solutions, if constraints on various quantities such as stress, dynamic response, inelastic response, etc., are considered. Although it is not always necessary to obtain the global optimal solution for an SMO problem, it is important to estimate the number of local optimal solutions and the accuracy of the best solution obtained so far in the process of

Makoto Ohsaki
Dept. of Architecture and Architectural Engineering, Graduate School of Engineering, Kyoto University, Japan
E-mail: ohsaki@archi.kyoto-u.ac.jp

Makoto Yamakawa
Dept. of Architecture, School of Science and Technology for Future Life, Tokyo Denki University, Japan
E-mail: myamakawa@mail.dendai.ac.jp

a multi-start strategy. For this purpose, various methods of global optimization have been proposed (Le Riche and Haftka 2012; Törn and Žilinskas 1989; Zabinsky 2003). Randomized algorithms can be successfully used for obtaining approximate optimal solutions (Ohsaki 2001; Ohsaki and Katsura 2012).

For SMO problems, we do not have to expect so many local optima as observed in mathematical test problems. Therefore, for problems with continuous functions and variables, gradient based nonlinear programming (NLP) approaches can be used to obtain a solution with good accuracy. However, a multi-start strategy should also be used, because the solution obtained by NLP approach depends on the initial solution.

For combinatorial problems with integer variables, heuristic approaches, which do not require gradient information, are often used. Heuristic approaches are classified into population based approaches and those based on local search (LS), which are also applied to multiobjective problems (Ohsaki 2008). The population based approaches such as genetic algorithm (GA) and particle swarm optimization (PSO) demand substantial computational cost, because the function values of many solutions should be evaluated at each generation.

By contrast, computational cost of the methods based on LS, such as simulated annealing (SA) and tabu search (TS), is rather small for a single trial, because only one solution and its neighborhood solutions are evaluated at each iteration. Furthermore, it is important to note that the solution is always improved from the initial solution, although a multi-start strategy is necessary to obtain the global optimal solution. Muselli (1997) compared the efficiency between consecutively restarting searches and continuing searches with a single start for random search, random walk, grid search, and various covering methods (Törn and Žilinskas 1989).

The difficulty in multi-start strategy is that the number of local optimal solutions including the global optimal solution is not known *a priori*. Therefore, several stopping rules have been proposed for multi-start LS, and also for pure random search (Dorea 1983; Dorea and Goncalves 1993; Hart 1998). A simple rule may be defined based on the history of objective values; i.e., if the objective value is not improved during the prescribed number of trials, then the multi-start process is terminated, and the best solution obtained so far is conceived as the global optimal solution.

Another strategy utilizes the attractor, which is the set of solutions leading to each local optimal solution by carrying out an LS. If a sufficient amount of attractor is obtained, then the multi-start process can be terminated. Lagaris and Tsoulos (2008) used a variance of the number of expected optimal solutions to define a stopping rule. The estimated number of local optimal solutions and the expected ratio of covered attractor are also used (Lagaris and Tsoulos 2008; Boender and Rinnooy Kan 1987). Zabinsky *et al.* (2010) proposed a criterion based on trade-off between the computational cost and probability of obtaining the global optimal solution. The number of attractors is usually counted by the number of initial solutions. However, if we use a deterministic LS, then we can include the intermediated solutions between the initial and final optimal solutions as the attractors.

Several Bayesian approaches have been proposed assuming no prior information is available for the number of solutions or the sizes of attractors. However, it is important to note that the history of multi-start process, or the numbers of attractors of already found solutions, should be utilized to make an accurate decision for each specific problem. Zieliński (1981) proposed a method for minimizing the Bayesian risk of the estimated number of remaining attractors. Although he used the information of the known attractors, it is not used in the final form.

In this paper, a new stopping rule for multi-start LS is proposed for application to SMO problems with moderately large number of local optima. The size of attractor of an unknown local optimal solution is estimated from those of already obtained solutions, and a stopping rule is proposed using the likelihood of obtaining the specific set of solutions. The proposed rule is first verified using the mathematical problems in comparison with the existing rule utilizing the estimated ratio of the total size of attractors to the size of feasible domain. The rule is next applied to an optimization problem of a plane frame under constraints on inter-story drift angle and stress against static loads.

2 Local search

We consider a deterministic method for improving a single solution to find a local optimal solution. Combinatorial problem is considered for clear identification of the local optima. The vector of m integer variables is denoted by $\boldsymbol{x} = (x_1, \dots, x_m)$. The stopping rule proposed in next section can be used for any optimization method based on LS; e.g., grid search, tabu search, and greedy method for local improvement. However, we consider the basic LS, which is described as follows for a problem of minimizing the objective function $f(\boldsymbol{x})$:

Algorithm of local search (LS)

1. Sample an initial random point $\boldsymbol{x}^0$ from a uniform probability distribution in the feasible domain. Set $k = 0$.
2. Enumerate all N neighborhood solutions of $\boldsymbol{x}^k$, denoted by $\boldsymbol{y}^i = (y_1^i, \dots, y_m^i)$ $(i = 1, \dots, N)$, and compute $f(\boldsymbol{y}^i)$.
3. Select the best solution $\boldsymbol{y}^{\min}$, which has the smallest value of $f(\boldsymbol{y}^i)$ among all neighborhood solutions.
4. If $f(\boldsymbol{x}^k) > f(\boldsymbol{y}^{\min})$, let $\boldsymbol{x}^{k+1} \leftarrow \boldsymbol{y}^{\min}$, $k \leftarrow k+1$, and go to 2; otherwise, output $\boldsymbol{x}^k$ as a local optimal solution and terminate the process.

There are several definitions of neighborhood solutions. Suppose x_j can take an integer value in the set $\{1, \dots, q_j\}$, and, for simplicity, suppose the relation $1 < x_j^k < q_j$ is satisfied for the current value x_j^k of x_j. Then the following definitions may be used for the neighborhood solutions $\boldsymbol{y}^i$ $(i = 1, \dots, N)$ of $\boldsymbol{x}^k$ at the kth step:

Moore neighborhood:

$$y_j^i = \{x_j^k - 1, x_j^k, x_j^k + 1\}, \quad (j = 1, \dots, m) \tag{1}$$

Neumann neighborhood:

$$\begin{cases} y_j^i = \{x_j^k - 1, x_j^k, x_j^k + 1\} & \text{for } j = j^* \\ y_j^i = x_j^k & \text{for } j \neq j^* \end{cases} \quad (j^* = 1, \dots, m) \tag{2}$$

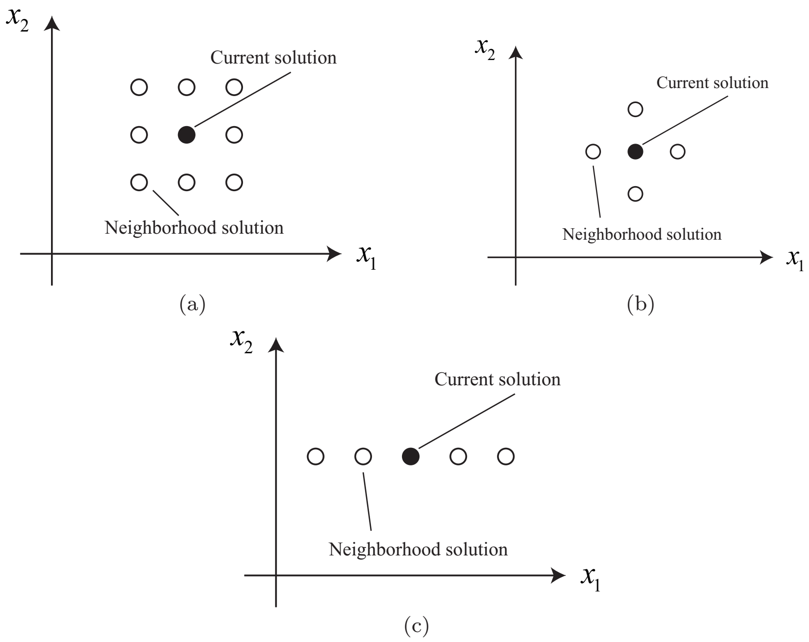

Fig. 1 Three types of neighborhood solutions; (a) Moore, (b) Neumann, (c) grid search ($j^* = 1$).

Neighborhood for grid search:

$$\left\{ \begin{array}{l} y_j^i = \{1, \ldots, q_i\} \quad j = j^* \\ y_j^i = x_j^k \quad \text{for } j \neq j^* \end{array} \right. \qquad (j^* \in \{1, \ldots, m\}) \tag{3}$$

The neighborhood solutions for the case $m = 2$ are illustrated in Fig. 1. Note that the grid search finds the best solution among the neighborhood solutions in a fixed direction. We use the Moore neighborhood in the following examples.

3 Stopping rule for multi-start LS

3.1 Existing stopping rule

Suppose we obtain w local optimal solutions $\boldsymbol{x}_1^*, \ldots, \boldsymbol{x}_w^*$ by carrying out LS t times from randomly generated initial solutions. The number of LSs that find $\boldsymbol{x}_i^*$ is denoted by n_i, i.e., $n_1 + \cdots + n_w = t$. Define X_i as *attractor* or *region of attraction* of $\boldsymbol{x}_i^*$ (Zieliński 1981; Lagaris and Tsoulos 2008), which is the set of solutions that leads to $\boldsymbol{x}_i^*$ by carrying out LS. For problems with integer variables, *attractor* and *feasible region* are to be replaced by *number of solutions in attractor* and *number of feasible solutions*, respectively.

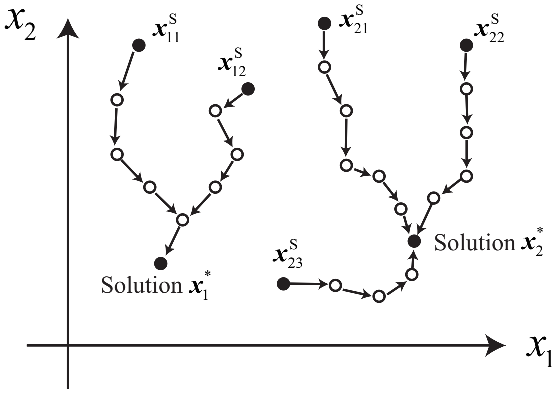

Fig. 2 Process of local search and definition of attractors ($w = 2$, $n_1 = 2$, $n_2 = 3$, $t = 5$, $T = 26$).

If we use a deterministic LS, the same solution $\boldsymbol{x}_i^*$ is found when the LS is started from any intermediate solution of the LS between the initial solution and $\boldsymbol{x}_i^*$. Therefore, all intermediate solutions along the path to $\boldsymbol{x}_i^*$ can be included in X_i. The total number of solutions visited during the t trials of LS is denoted by T, which is counted without duplication.

The definitions of parameters w, t, and T are summarized as follows for reference in the following parts:

- w: number of optimal solutions obtained by LS
- t: number of trials of LS, or number of initial solutions
- T: total number of solutions visited during t trials of LS, including initial, optimal, and intermediate solutions along the path of LS

Figure 2 illustrates the process of LSs for a problem with two variables $\boldsymbol{x} = (x_1, x_2)$. Two solutions $\boldsymbol{x}_1^*$ and $\boldsymbol{x}_2^*$ are found; i.e., $w = 2$. The solutions $\boldsymbol{x}_1^*$ and $\boldsymbol{x}_2^*$ are found from two and three initial solutions $\{\boldsymbol{x}_{11}^{\mathrm{S}}, \boldsymbol{x}_{12}^{\mathrm{S}}\}$ and $\{\boldsymbol{x}_{21}^{\mathrm{S}}, \boldsymbol{x}_{22}^{\mathrm{S}}, \boldsymbol{x}_{23}^{\mathrm{S}}\}$, respectively; therefore, $n_1 = 2$, $n_2 = 3$, and $t = n_1 + n_2 = 5$. The two and three paths leading to $\boldsymbol{x}_1^*$ and $\boldsymbol{x}_2^*$ have 10 and 16 solutions, respectively; therefore, $T = 10 + 16 = 26$.

The ratio of size s_i of X_i to the number of all feasible solutions, which is called share of X_i (Zieliński 1981), is denoted by c_i. Suppose the problem has h local optimal solutions, which are not known *a priori*. Boender and Rinnooy Kan (1987) derived the following estimate $\tilde{h}$ of the number of local optimal solutions based on Bayesian approach, where the share of attractor of solution is supposed to be uniformly distributed between 0 and 1 satisfying $c_1 + \cdots + c_h = 1$, and the intermediate solutions are not included in the attractors. Suppose we have found w local optimal solutions during t trials; i.e., $n_1 + \cdots + n_w = t$. Then $\tilde{h}$ is estimated as the mean value of posterior estimate of h as follows:

$$\tilde{h} = \frac{w(t-1)}{t-w-2} \tag{4}$$

The multi-start LS can be terminated if w is sufficiently close to $\tilde{h}$.

The ratio $\tilde{\Omega}$ of sum of size of attractors $X_1, \ldots, X_w$ to the total feasible region is also estimated as

$$\tilde{\Omega} = \frac{(t-w-1)(t+w)}{t(t-1)} \tag{5}$$

It is noted by Boender and Rinnooy Kan (1987) that the convergence of $\tilde{h}$ is very slow; therefore, they proposed another stopping rule of multi-start LS based on (5):

Rule 1 (Boender and Rinnooy Kan 1987): Terminate multi-start LS if $\tilde{\Omega} \geq e_1$ is satisfied, where e_1 is a prescribed value slightly less than 1.

3.2 New stopping rules

Note again that the ratio c_i of attractor is assumed to be uniformly distributed between 0 and 1 in the Bayesian approaches for Rule 1; i.e., property of the problem to be solved is not fully incorporated in the estimation. Here, we estimate c_i based on the record of multi-start LS. If s_i consists of the number of initial solutions leading to $\boldsymbol{x}_i^*$, then $s_i = n_i$, $s_1 + \cdots + s_w = t$, and $c_i = s_i/t$. By contrast, if s_i includes all intermediate solutions, then $s_1 + \cdots + s_w = T$ and $c_i = s_i/T$ are satisfied. In the example of Fig. 2, $s_1 = 10$ and $s_2 = 16$; therefore, $c_1 = 10/26$ and $c_2 = 16/26$. In the definitions of the following rules, the intermediate solutions are included in the attractors; i.e., the total number of solutions in attractors is T, while t is used as the total number of trials.

Suppose there is another solution x_{w+1}^* that has not been found. Let $c_i^{(k)}$ denote the ratio of size s_i of attractor X_i when there exist $w+1$ local optimal solutions with $s_{w+1} = k$; i.e.,

$$\begin{aligned} c_i^{(k)} &= \frac{s_i}{T+k}, \quad (i = 1, \ldots, w), \\ c_{w+1}^{(k)} &= \frac{k}{T+k} \end{aligned} \tag{6}$$

Note that $c_1^{(k)} + \cdots + c_{w+1}^{(k)} = 1$ is satisfied from (6) and $s_1 + \cdots + s_w = T$. Since n_i is a random variable, it is written as N_i. Then the probability of finding solutions $\boldsymbol{x}_1^*, \ldots, \boldsymbol{x}_{w+1}^*$, respectively, $N_1, \ldots, N_{w+1}$ times in t trials is written as

$$P_w^{(k)}(N_1, \ldots, N_{w+1} | c_1^{(k)}, \ldots, c_{w+1}^{(k)}, w, t) = \frac{t!}{\prod_{i=1}^{w+1} N_i!} \prod_{i=1}^{w+1} (c_i^{(k)})^{N_i} \tag{7}$$

After n_i times finding $\boldsymbol{x}_i^*$ $(i = 1, \ldots, w)$ and 0 time $\boldsymbol{x}_{w+1}^*$ in t trials; i.e., $n_1 + \cdots + n_w = t$ and $n_{w+1} = 0$, the likelihood $L_w^{(k)}$ for observing this event is computed as

$$\begin{aligned} L_w^{(k)} &= P_w^{(k)}(N_1 = n_1, \ldots, N_w = n_w, N_{w+1} = 0 | c_1^{(k)}, \ldots, c_{w+1}^{(k)}, w, t) \\ &= \frac{t!}{\prod_{i=1}^{w+1} n_i!} \prod_{i=1}^{w+1} (c_i^{(k)})^{n_i} \\ &= \frac{t!}{\prod_{i=1}^{w} n_i!} \prod_{i=1}^{w} (c_i^{(k)})^{n_i} \end{aligned} \tag{8}$$

The size s_{w+1} of attractor X_{w+1} of x^*_{w+1} is estimated based on the information of the attractors that have already been found. From our preliminary investigation, we found that the minimum value of $s_1, \ldots, s_w$ is too small as the estimate of s_{w+1}, while the maximum value is too large. Therefore, we first use a smoothing process.

The maximum and minimum numbers of attractors of the w solutions are denoted by s^{U} and s^{L}, respectively. We assume that the sizes of attractors of unknown solutions are uniformly distributed between s^{L} and s^{U}. Then, $\bar{L}_w$ is defined as the average value of $L_w^{(k)}$ for $k = s^{\mathrm{L}}, \ldots, s^{\mathrm{U}}$ as

$$\bar{L}_w = \frac{1}{s^{\mathrm{U}} - s^{\mathrm{L}} + 1} \sum_{k=s^{\mathrm{L}}}^{s^{\mathrm{U}}} L_w^{(k)} \tag{9}$$

The value of $L_w^{(k)}$ when no solution is left, i.e., $w = h$, is denoted by $L_w^{(0)}$. The following stopping rule is proposed using $\bar{L}_w$:

Rule 2: Terminate multi-start LS if $\bar{L}_w / L_w^{(0)}$ is smaller than a specified small positive value e_2.

This way, the history of trials of LS is fully incorporated to $\bar{L}_w$ and $L_w^{(0)}$ that are updated at each trial using (8) based on the updated ratios of attractors in (6).

From (6) and (8), we obtain

$$\begin{aligned} c_i^{(0)} &= c_i, \quad (i = 1, \ldots, w), \\ L_w^{(0)} &= \frac{t!}{\prod_{i=1}^{w} n_i!} \prod_{i=1}^{w} (c_i)^{n_i} \end{aligned} \tag{10}$$

Therefore, using (6) and $s_i = c_i T$, $L_w^{(k)} / L_w^{(0)}$ is reformulated as

$$\begin{aligned} \frac{L_w^{(k)}}{L_w^{(0)}} &= \prod_{i=1}^{w} \left(\frac{T}{T+k} \right)^{n_i} \\ &= \left(\frac{T}{T+k} \right)^{t} \\ &= \left(1 - \frac{k}{T+k} \right)^{t} \\ &= (1 - c_{w+1}^{(k)})^{t} \end{aligned} \tag{11}$$

which is equal to the probability of missing the $(w+1)$th solution within t trials.

Using (9) and (11), we obtain

$$\frac{\bar{L}_w}{L_w^{(0)}} = \frac{1}{s^{\mathrm{U}} - s^{\mathrm{L}} + 1} \sum_{k=s^{\mathrm{L}}}^{s^{\mathrm{U}}} (1 - c_{w+1}^{(k)})^{t} \tag{12}$$

In the example in Fig. 2 that has two local optima, $s^{\mathrm{L}} = 10$ and $s^{\mathrm{U}} = 16$, and $\bar{L}_w / L_w^{(0)}$ is computed using (11) as

$$\frac{\bar{L}_w}{L_w^{(0)}} = \frac{1}{16 - 10 + 1} \sum_{k=10}^{16} \left(1 - \frac{k}{26 + k}\right)^5 = 0.08695 \tag{13}$$

As shown in the numerical examples, $\bar{L}_w$ depends significantly on the existence of small and/or large attractor. Therefore, probability of finding another solution is overestimated, if one solution has a very large attractor. Hence, we next assume w possible choices of s_{w+1} as $s_{w+1} = s_j$ $(j = 1, \dots, w)$, and define c_i^{j*} as

$$c_i^{j*} = \frac{s_i}{T + s_j}, \quad (i = 1, \dots, w) \tag{14}$$

Using (14), L_w^{j*} is computed as

$$L_w^{j*} = \frac{t!}{\prod_{i=1}^{w} n_i!} \prod_{i=1}^{w} (c_i^{j*})^{n_i}, \quad (j = 1, \dots, w) \tag{15}$$

and $\bar{L}_w^*$ is obtained as the mean value of L_w^{j*} as

$$\bar{L}_w^* = \frac{1}{w} \sum_{j=1}^{w} L_w^{j*} \tag{16}$$

In a similar manner as (11) and (12), the likelihood ratio is obtained as

$$\begin{aligned} \frac{\bar{L}_w^*}{L_w^{(0)}} &= \frac{1}{w} \sum_{j=1}^{w} \left(\frac{T}{T + s_j}\right)^t \\ &= \frac{1}{w} \sum_{j=1}^{w} \left(1 - \frac{s_j}{T + s_j}\right)^t \\ &= \frac{1}{w} \sum_{j=1}^{w} (1 - c_{w+1}^{(s_j)})^t \end{aligned} \tag{17}$$

Using $\bar{L}_w^*$, the following rule is proposed

Rule 3: Terminate multi-start LS if $\bar{L}_w^* / L_w^{(0)}$ is smaller than a specified small positive value e_3.

In the example in Fig. 2, $\bar{L}_w^* / L_w^{(0)}$ is computed using (17) as

$$\frac{\bar{L}_w^*}{L_w^{(0)}} = \frac{1}{2} \left[\left(1 - \frac{10}{26 + 10}\right)^5 + \left(1 - \frac{16}{26 + 16}\right)^5 \right] = 0.1437 \tag{18}$$

4 Numerical examples

4.1 Mathematical problems

The proposed rule is verified using the mathematical problems in Voglis and Lagaris (2009). The TestN2 function (Lagaris and Tsoulos 2008) is also used. Note that problems with too many local optima are not considered, because we attempt to apply the rule to SMO problems with moderately large number of solutions. The test functions are listed as follows, where N^{opt} is the number of local optima, and x^{L} and x^{U} are the lower and upper bounds, respectively, of the variables:

Ackley's function ($N^{\mathrm{opt}} = 121$):

$$f(\boldsymbol{x}) = 20 + e - 20e^{-0.2\sqrt{a}} - e^{b/2}, \quad a = \frac{1}{3}\sum_{i=1}^{2} x_i^2, \quad b = \sum_{i=1}^{2}\cos(2\pi x_i),$$
$$x^{\mathrm{L}} = -5, \quad x^{\mathrm{U}} = 5$$

Guillin Hill's function ($N^{\mathrm{opt}} = 25$):

$$f(\boldsymbol{x}) = 3 + \sum_{i=1}^{2} \frac{2(x_i + 9)}{x_i + 10} \sin\left(\frac{\pi}{1 - x_i + 0.1}\right), \quad x^{\mathrm{L}} = 0, \quad x^{\mathrm{U}} = 1$$

Holder function ($N^{\mathrm{opt}} = 85$):

$$f(\boldsymbol{x}) = -(\cos x_1 \cos x_2)e^{(1-\sqrt{x_1^2+x_2^2}/\pi)}, \quad x^{\mathrm{L}} = -20, \quad x^{\mathrm{U}} = 20$$

Piccioni's function ($N^{\mathrm{opt}} = 28$):

$$f(\boldsymbol{x}) = 0.5 + \frac{\sin(x_1^2 + x_2^2)^2 - 0.5}{[1 + 0.001(x_1^2 + x_2^2)]^2} + 0.1\sin(10x_1) + 0.1\sin(10x_2),$$
$$x^{\mathrm{L}} = -5, \quad x^{\mathrm{U}} = 5$$

M0 function ($N^{\mathrm{opt}} = 64$):

$$f(\boldsymbol{x}) = \left[\sin\left(2.2\pi x_1 + \frac{\pi}{2}\right) + \sin\left(\frac{\pi}{2}x_2^2 + \frac{\pi}{2}\right)\right]\frac{(2 - x_2)(3 - x_1)}{4},$$
$$x^{\mathrm{L}} = -5, \quad x^{\mathrm{U}} = 1$$

Test2N function $n = 5$ ($N^{\mathrm{opt}} = 32$):

$$f(\boldsymbol{x}) = \frac{1}{5}\sum_{i=1}^{5}(x_i^4 - 16x_i^2 + 5x_i), \quad x^{\mathrm{L}} = -5, \quad x^{\mathrm{U}} = 5$$

The variable x_i is discretized with 100 equally spaced interval; i.e., x_i is defined in terms of the integer variable $J_i \in \{1, 2, \dots, 101\}$ as

$$x_i = x^{\mathrm{L}} + (J_i - 1)\frac{x^{\mathrm{U}} - x^{\mathrm{L}}}{100} \tag{19}$$

Let S denote the number of trials of LS to find all local optima. The results for six test functions are shown in Table 1. Since the number of local optima is known for each problem as shown in Table 1, ten sets of trials are carried out until all local optima are found. Although the numbers of solutions in Voglis and Lagaris (2009) and Lagaris and Tsoulos 2008 are based on continuous variables,

Table 1 Values when all solutions are found.

	N^{opt}	S	$\tilde{\Omega}$	$\bar{L}_w/L_w^{(0)}$	$\bar{L}_w^*/L_w^{(0)}$
Ackley's	121	969.7	0.9806	0.02391	0.02003
Guillin Hill's	25	3589.4	0.9999	0.00044	0.06901
Helder	85	443.1	0.9595	0.05069	0.03346
Piccioni's	28	453.2	0.9942	0.01245	0.07872
M0	64	2923.1	0.9987	0.00533	0.01635
Test2N	32	125.1	0.9110	0.06020	0.09360

Table 2 Distribution of ratios of attractors.

	Max.	Min.	Mean	Std. Dev.
Ackley's	0.0201	0.00275	0.00826	0.0030
Guillin Hill's	0.5468	0.00012	0.04000	0.1084
Helder	0.0147	0.00844	0.01176	0.0014
Piccioni's	0.1043	0.00431	0.03571	0.0344
M0	0.0484	0.00028	0.01562	0.0135
Test2N	0.0406	0.02439	0.03125	0.0038

Table 3 Values when stopping criteria are satisfied.

	$N_{95}^{\mathrm{opt}}(S)$	$N_{98}^{\mathrm{opt}}(S)$	$N_{995}^{\mathrm{opt}}(S)$	
Ackley's	116.4(523.9)	119.9(852.3)	121.0(1719.0)	
Guillin Hill's	10.7(50.8)	13.3(98.5)	18.7(272.6)	
Helder	84.0(378.7)	84.8(604.4)	85.0(1210.0)	
Piccioni's	21.5(99.5)	25.0(181.0)	27.2(392.8)	
M0	54.1(245.1)	58.5(418.5)	61.7(880.8)	
Test2N	31.7(144.7)	32.0(231.0)	32.0(461.0)	
	$\bar{N}_{02}^{\mathrm{opt}}(S)$	$\bar{N}_{05}^{\mathrm{opt}}(S)$	$\bar{N}_{02}^{\mathrm{opt}*}(S)$	$\bar{N}_{05}^{\mathrm{opt}*}(S)$
Ackley's	119.9(930.7)	116.4(508.0)	119.5(836.8)	117.8(568.2)
Guillin Hill's	11.6(64.6)	7.5(27.2)	24.7(4926.6)	23.9(2815.5)
Helder	84.9(558.3)	84.0(442.7)	84.7(484.8)	83.1(353.7)
Piccioni's	27.1(287.5)	23.2(124.5)	27.6(671.7)	27.8(518.2)
M0	59.3(603.9)	53.9(250.2)	63.4(1817.3)	59.5(553.5)
Test2N	32.0(186.3)	31.4(125.9)	31.9(192.6)	31.8(145.9)

we confirmed that the same numbers are obtained after discretization by (19). The value of $\tilde{\Omega}$ is close to one for Guillin Hill's, Piccioni's, and M0, and $\bar{L}_w/L_w^{(0)}$ has small values for Guillin Hill's and M0. However, $\bar{L}_w^*/L_w^{(0)}$ has the same order for the six problems.

Distribution of the ratios c_i of attractors is listed in Table 2 for the six problems. Note that the minimum values are very small for Guillin Hill's and M0, and the standard deviation is large for Guillin Hill's, Piccioni's, and M0. It is seen from the results in Tables 1 and 2 that Rules 1 and 2 are not effective for problems that have local optima with small attractors.

Stopping criteria are defined with the parameter values $e_1 = 0.95$, 0.98, and 0.995 for Rule 1, $e_2 = 0.02$ and 0.05 for Rule 2, and $e_3 = 0.02$ and 0.05 for Rule 3. The results are shown in Table 3, where N_{95}^{opt} is the average value of N^{opt} when $\tilde{\Omega} > 0.95$ (Rule 1) is satisfied, and $\bar{N}_{02}^{\mathrm{opt}}$ and $\bar{N}_{02}^{\mathrm{opt}*}$ are the average values of N^{opt} when $\bar{L}_w/L_w^{(0)} < 0.02$ (Rule 2) and $\bar{L}_w^*/L_w^{(0)} < 0.02$ (Rule 3), respectively, are

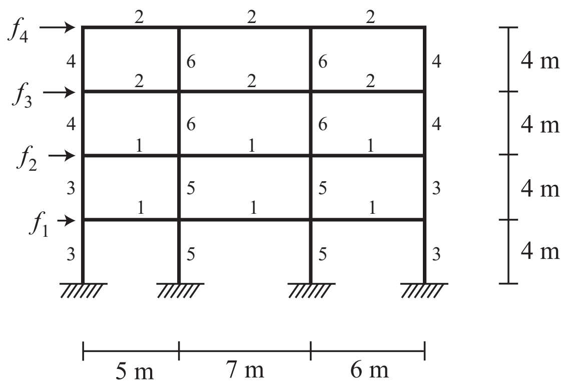

Fig. 3 A three-span four-story plane frame.

satisfied. Other parameters are defined similarly. The value in parentheses is the average number of trials to satisfy the stopping rule.

As expected, large portion of local optima is missed for Guillin Hill's and M0, if Rule 1 or 2 is used. It is possible to use a value of e_1 close to 1 for Rule 1. However, as seen in the fourth column in Table 3, the value of N_{995}^{opt} is still less than $\bar{N}_{02}^{\mathrm{opt}*}$ for Guillin Hill's and M0, while large number of unnecessary trials should be carried out for Ackley's, Helder, and Test2N. By contrast, $\bar{N}_{02}^{\mathrm{opt}*}$ is close enough to N^{opt}; i.e., Rule 3 gives better estimate of the number of local optimal solutions than Rules 1 and 2. Although it is difficult to define the value of e_3, it is important that its appropriate value does not strongly depend on property of the problem to be solved.

It is also seen from Table 3 that $\bar{N}_{05}^{\mathrm{opt}*}$ is slightly smaller than $\bar{N}_{02}^{\mathrm{opt}*}$ for all problems except Piccioni's, for which $\bar{N}_{05}^{\mathrm{opt}*} > \bar{N}_{02}^{\mathrm{opt}*}$ is satisfied. This is because N^{opt} depends on the initial random seed, and almost all solutions are found before satisfying $\bar{L}_w^*/L_w^{(0)} < 0.05$.

4.2 Optimization of plane frame

Multi-start local search is carried out for optimization of a three-span four-story plane frame as shown in Fig. 3, where the numbers beside the beams and columns are group numbers; i.e., there are six groups ($n^{\mathrm{g}} = 6$). The horizontal loads are given as $f_1, \ldots, f_4 = 180,\ 210,\ 240,\ 270$ (kN). Concentrated downward vertical load of 70 kN representing the self-weight is applied at the beam-column joint and the center of each beam.

The member section is assumed to be sandwich section with the height H; i.e., we consider wide-flange section neglecting the effect of web. The cross-sectional areas $A^{(i)}$ (m^2) of the members in the ith group are selected from the list of n^{s} different sections as

$$A_j^{(i)} = A_0^{(i)} + 0.002 J_i, \quad (i = 1, \ldots, n^{\mathrm{g}}; J_i \in \{1, \ldots, n^{\mathrm{s}}\}) \tag{20}$$

The values of $A_0^{(i)}$ ($i = 1, \ldots, n^{\mathrm{g}}$) are given as 0.019, 0.007, 0.017, 0.007, 0.021, 0.007 (m^2), and $n^{\mathrm{s}} = 10$ in the following example.

Table 4 Values when stopping criteria are satisfied.

	N^{opt}	S	$N_{95}^{\mathrm{opt}}(S)$	$N_{98}^{\mathrm{opt}}(S)$	$N_{995}^{\mathrm{opt}}(S)$
$\theta = 0.010$	12	1847.2	8.6(41.4)	11.2(70.8)	9.8(146.2)
$\theta = 0.011$	15	2158.6	7.4(36.0)	8.6(65.2)	10.8(160.4)
$\theta = 0.012$	10	696.2	6.0(29.8)	7.0(54.0)	7.4(112.6)

	$\bar{N}_{02}^{\mathrm{opt}}(S)$	$\bar{N}_{05}^{\mathrm{opt}}(S)$	$\bar{N}_{02}^{\mathrm{opt}*}(S)$	$\bar{N}_{05}^{\mathrm{opt}*}(S)$
$\theta = 0.010$	9.6(72.2)	8.6(42.2)	10.6 (1101.4)	10.2 (268.4)
$\theta = 0.011$	7.6(47.6)	5.2(18.8)	13.8(2409.0)	10.0(242.2)
$\theta = 0.012$	6.6(42.4)	5.4(19.2)	8.4(376.2)	7.0(89.8)

The second moment of area $I_j^{(i)}$ (m^4) and section modulus $Z_j^{(i)}$ (m^3), of the jth section of the members in the ith group are defined with respect to the cross-sectional areas as

$$\begin{aligned} I_j^{(i)} &= (H/2)^2 A_j^{(i)} \\ Z_j^{(i)} &= I_j^{(i)}/(H/2) = (H/2)A_j^{(i)} \end{aligned} \tag{21}$$

where $H = 0.25$ (m) in the following example.

The design variable vector consisting of the cross-sectional areas A_i ($i = 1, \dots, n^{\mathrm{g}}$) is denoted by $\boldsymbol{A}$. Let L_i denote the total length of members in the ith group. The total structural volume is defined as

$$F(\boldsymbol{x}) = \sum_{i=1}^{n^{\mathrm{g}}} A_i L_i \tag{22}$$

which is minimized under constraints on inter-story drift angle and stress. The upper-bound stress is 235 MPa, and the upper bound θ of inter-story drift angle is varies parametrically as 0.010, 0.011, and 0.012. Moore neighborhood is used; i.e., there are at most $3^6 = 729$ neighborhood solutions at each step of local search. Since the number of neighborhood solutions is very large, the Neumann neighborhood may be used for practical application.

Five sets of multi-start local search are carried out from different random seeds, where each set consists of 5000 trials of local search to find all local optimal solutions. Note that N^{opt} has the same value for five trials, respectively, for three gases of θ, and the average steps S for three cases are sufficiently smaller than the total steps 5000; therefore, we assume all solutions have been found.

Table 4 shows the mean values among five sets when Rules 1 ($e_1 = 0.95$, 0.98, 0.995), Rule 2 ($e_2 = 0.02$, 0.05), and Rule 3 ($e_3 = 0.02$, 0.05) are satisfied. Note that Rule 1 with $e_1 = 0.95$ is not appropriate, because N_{95}^{opt} in Table 4 is too small for all cases. Even for $e_1 = 0.995$, the ratio of missing optimal solutions are not small enough especially for $\theta = 0.011$. By contrast, Rule 3 with $e_2 = 0.02$ may be appropriate as the termination rule for this frame, because $\bar{N}_{02}^{\mathrm{opt}*}$ is slightly smaller than N^{opt}.

Table 5 shows the three best solutions for each case. As seen from the table, there are several solutions that has the same or almost the same objective value. Furthermore, the best solution has large number s_i of attractors, and the large number n_i of the trials to reach the solution. Table 6 shows 12 solutions obtained from a set of 5000 trials for $\theta = 0.010$. As seen from the table, the best solution is

Table 5 Three best solutions for $\theta = 0.010$, 0.011, and 0.012.

	J_1	J_2	J_3	J_4	J_5	J_6	$F(\boldsymbol{x})$	n_i	s_i
$\theta = 0.010$	7	3	4	2	5	6	2.864	1410	3558
	8	3	2	3	5	4	2.872	492	1040
	8	3	3	3	4	4	2.872	504	1109
$\theta = 0.011$	6	3	3	2	3	4	2.632	2140	4546
	4	3	3	2	7	5	2.648	699	1508
	5	3	3	3	5	3	2.656	142	290
$\theta = 0.012$	3	3	4	2	4	4	2.480	2282	4949
	3	3	3	2	5	4	2.480	193	673
	4	3	5	2	2	3	2.488	1521	3133

Table 6 All 12 solutions obtained by a set of 5000 trials for $\theta = 0.010$.

Rank	J_1	J_2	J_3	J_4	J_5	J_6	$F(\boldsymbol{x})$	n_i	s_i	S
1	7	3	4	2	5	6	2.864	1410	3558	8
2	8	3	2	3	5	4	2.872	492	2040	1
2	8	3	3	3	4	4	2.872	504	1109	3
2	8	3	2	2	5	6	2.872	7	10	1351
5	6	2	4	2	7	5	2.888	1158	2949	2
5	6	3	3	2	8	7	2.888	14	23	19
5	6	3	3	3	8	5	2.888	86	161	76
5	6	3	4	2	7	7	2.888	30	65	160
9	7	3	7	2	3	6	2.896	364	839	5
9	7	3	6	2	4	6	2.896	582	1476	10
11	8	3	5	3	3	4	2.904	268	565	4
12	8	3	7	3	2	4	2.936	85	185	59

found at the 8th trial, and it is the most frequently obtained solution with largest size of attractor. Therefore, for this frame optimization problem, the existence of local optimal solution with small attractor may be neglected.

5 Conclusions

A new stopping rule for multi-start LS has been presented for application to structural optimization problems with moderately large number of local optima. The conclusions obtained from this study are summarized as follows:

1. The proposed stopping rule incorporates the history of multi-start LS into its definition. It has been shown through examples of mathematical test problems that information such as the size of attractors of local optimal solutions of the specific optimization problem is useful for improving accuracy of the stopping rule. If a deterministic LS with integer variable is used, the intermediate solutions between the initial and optimal solution can effectively be incorporated in the attractor to the local optimal solution.
2. The ratio of likelihood of occurrence of the history of finding local optimal solutions is computed based on two types of assumption; i.e., all solutions are found, and there is one missing solution. The size of attractor of the missing local optimal solution is estimated using those of the already found solutions. It has been demonstrated in the numerical examples that the ratio of likelihood of two cases can be a good measure for stopping rule. It has also been shown

that the likelihood ratio is equivalent to the probability of missing the unknown local optimal solution within the specific number of trials.

3. The proposed rule is applicable even to the case where small attractors exist, because the small attractor has only slight effect on the definition of likelihood ratio. Effect of large attractor is also alleviated by taking average value of likelihood of obtaining the sequence of solutions corresponding to different expected size of the attractor of unknown solution. The rule can be effectively used for practical application to a frame optimization problem with discrete variables, which has moderately large number of local optimal solutions.

Acknowledgments

This work is partially supported by JSPS KAKENHI No. 25420576.

References

C. G. E. Boender and A. H. G. Rinnooy Kan, Bayesian stopping rules for multistart global optimization methods, Mathematical Programming, Vol. 37, pp. 59–80, 1987.

C. C. Y. Dorea, Expected number of steps of a random optimization method, J. Opt. Theory and Appl., Vol. 39(2), pp. 165–171, 1983.

C. C. Y. Dorea and C. R.Goncalves, Alternative sampling strategy for a random optimization algorithm, J. Opt. Theory and Appl., Vol. 78(2), pp. 401–407, 1993.

W. E. Hart, Sequential stopping rules for random optimization methods with applications to multistart local search, SIAM J. Optim., Vol. 9(1), pp. 270–290, 1998.

I. E. Lagaris and I. G. Tsoulos, Stopping rules for box-constrained stochastic global optimization, Appl. Math. and Computation, Vol. 197, pp. 622–632, 2008.

R. Le Riche and R. T. Haftka, On global optimization articles in SMO, Struct. Multidisc. Optim., Vol. 46, pp. 627–629, 2012.

M. Muselli, A theoretical approach to restart in global optimization, J. Glob. Optim., Vol. 10, pp. 1–16, 1997.

M. Ohsaki, Random search method based on exact reanalysis for topology optimization of trusses with discrete cross-sectional areas, Comput. & Struct., Vol. 79(6), pp. 673–679, 2001.

M. Ohsaki, Local search for multiobjective optimization of steel frames, Proc. 5th China-Japan-Korea Joint Symposium on Optimization of Structural and Mechanical Systems (CJK-OSM5), Jeju, Korea, 2008.

M. Ohsaki and M. Katsura, A random sampling approach to worst-case design of structures, Struct. Multidisc. Optim., Vol. 46, pp. 27–39, 2012.

M. Pant, R. Thangaraj, A. Abraham, and K. Deep, Particle swarm optimization using sobol mutation, Int. J. Simulation, Systems, Science and Technology, Vol. 10(3), pp. 89–98, 2009.

A. Törn and Žilinskas, Global Optimization, Lecture Notes in Comput. Sci., No. 350, Springer, 1989.

C. Voglis and I. E. Lagaris, Towards 'Ideal Multistart': A stochastic approach for locating the minima of a continuous function inside a bounded domain, Appl. Math. and Computation, Vol. 213, pp. 216–219, 2009.

Stochastic Adaptive Search for Global Optimization, Kluwer Academic Publishers, 2003.

Z. B. Zabinsky, D. Bulger and C. Khompatraporn, Stopping and restarting strategy for stochastic sequential search in global optimization, J. Glob. Optim., Vol. 46, pp. 273–286, 2010.

R. Zieliński, A statistical estimate of the structure of multi-extremal problems, Mathematical Programming, Vol. 21, pp. 348–356, 1981.